1983

Probabilistic qualitative response estimation of the spatial market for Powder River Basin coal

Henry Y. Yoshimura
The University of Montana

A PROBABILISTIC QUALITATIVE RESPONSE ESTIMATION OF THE SPATIAL MARKET FOR POWDER RIVER BASIN COAL

By

Henry Y. Yoshimura

B.A., University of Montana, 1981

Presented in partial fulfillment of the requirements for the degree of

Master of Arts

UNIVERSITY OF MONTANA

1983

Approved by:

Chairman, Board of Examiners

Dean, Graduate School

6-6-83

Date

Yoshimura, Henry Y., M.A., June 1983 Economics

A Probabilistic Qualitative Response Estimation of the Spatial Market for Powder River Basin Coal (152 pp.)

Director: Dr. Thomas M. Power TMP

The national electric utility steam coal market is characterized by competition among approximately 13 major coal supply regions. Spatial electric utility coal market analysis attempts to answer the question: what geographic region will supply the coal a given coal-fired power plant will use? The major purpose of this study is to specify and estimate an appropriate descriptive spatial market model for Powder River Basin subbituminous coal.

Since the market boundary between two competing coal supply regions is better characterized by a broad band rather than a unique sharp line, a more appropriate spatial market model is one that predicts the probability a given power plant will buy Powder River coal; the probability of a power plant using Powder River coal declines according to some cumulative probability distribution function as power plants are located farther away from the Powder River Basin. The estimation of a Powder River spatial coal market model was therefore accomplished using a probabilistic qualitative response regression model that was based upon the cumulative logistic distribution function. A sample of 438 power plants, based on 1980 data, was used for the regression estimation.

The estimated model was found to be significant in explaining coal choosing behavior. The chi-square value of the estimated model at 3 degrees of freedom was 313.252 (which is significant at the 0.0000 level). The estimated model can predict 1980 coal choosing behavior correctly 89% of the time. McFadden's R-squared was estimated to be 61.25%. Effron's R-squared was estimated to be 62.68%. Since R-squared measurements in qualitative response models are usually biased downwards, the above statistical calculations seem to indicate a good fitting model.

ACKNOWLEDGMENTS

The completion of this thesis was made possible through a combined effort and support of many individuals. I particularly thank my thesis committee (i.e. Thomas Power, Arnold Silverman, Michael Kupilik and Kay Unger) and John Duffield for their helpful guidance, comments, and criticisms.

Special gratitude and recognition goes to Charles Weichler for his selfless assistance and contributions in the area of computer graphics and programing. Many of the unique aspects of this thesis was made possible through his special talent, creative ideas, and friendly encouragement.

Finally, I dedicate this thesis to my wife Linda. Her contribution (both direct and in support) was the critical "swing ingredient" in the preparation and finalization of this document.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

A STATEMENT OF THE PROBLEM

I. Introduction

Immense deposits of low sulfur strippable coal underlying the Northern Great Plains (NGP)* states of Montana, Wyoming, North Dakota, and South Dakota have undergone expeditious development in recent history. Because of the perceived instability of foreign energy sources and the high price of substitute fuels, political and economic incentives for using large domestic coal reserves for the production of electricity have risen. For example, production of NGP coal leaped from approximately 15 million tons per year in 1976 to about 96 million tons per year in 1979. It is likely that coal deliveries will exceed 182 million tons per year in 1985 (Duffield et al., 1982). Since rapid development of NGP coal will drastically affect the entire character of the region, intense interest in forecasting the demand for NGP coal has evolved over the past decade.

The extent to which NGP coal will be developed in the future will be largely tied to growth in electricity demand since the overwhelming majority of NGP coal is sold to electric utilities (over 90% of NGP coal was sold for

* The coal fields referred to include the Powder River Basin of Montana and Wyoming and the Fort Union Basin of Montana, North Dakota and South Dakota.

electric generation in 1979). Since NGP coal is almost exclusively used for electric generation, past NGP coal demand models concentrated attention specifically on the electric utility coal market. The industrial market and the export market were examined separately (Power et al., 1976 and Duffield et al., 1982). The market for NGP electric utility coal is geographically constrained however because coal is a low value per unit weight commodity. Utility companies may buy coal from one of several coal fields located throughout the United States. Studies on NGP electric utility coal demand therefore utilized the theory of spatial markets (Hyson and Hyson, 1950) in order to define the geographical market for NGP coal. Defining the NGP spatial coal market was the essential first step in estimating NGP electric utility steam coal demand. This is because spatial market analysis provides a systematic approach for examining the determinates of electric utility coal choice. If the variables that shifted the NGP coal market boundary were known, shifts in NGP coal demand could be readily explained.

The estimation of the NGP coal geographical market involves drawing a market boundary between the NGP coal supply center (i.e. Gillette, Wyoming) and other competing coal supply centers located in different areas of the country. The spatial market boundary is usually defined as the locus of points where the total cost of utilizing an input resource purchased from competing sources are equal.

This implies that on either side of the boundary, one of the resource suppliers experience a cost advantage over their competitors. The aforementioned research found that the spatial coal market boundary is a function of air pollution regulations, relative mine mouth prices, and transportation rates.

The major shortcoming of the above spatial market analyses is that the market boundary model is a hypothetical model that has not been tested for empirical significance. Given a set of assumed hypothetical behavior and conditions the model draws hypothetical market boundaries; the market boundaries are true by definition. Since market boundaries are not built on empirical electric utility coal buying behavior, we are not sure whether or not the model will have any significance in explaining and predicting empirical world behavior. It has been found the empirical coal markets overlap considerably is some cases. New power plants in the states of Louisiana, Michigan and Texas (states that fall outside the calculated NGP market) have contracted for Powder River coal. Also, some new Nebraska and Iowa power plants (states that always fall within the predicted market boundary) use non-NGP coal (Duffield et al., 1982).

II. Proposed Research

Spatial coal market analysis attempts to answer the following question: given the power plant location, which coal among the alternative supply centers will the power plant use? It is a problem that requires us to analyze the factors that affect electric utility coal choice. First, we must identify the variables and relationships that conceivably affect coal choice. Second, we must set up a model (usually expressed in mathematical terms) that summarizes the relationship the variables have on coal choice. This model also enables us to empirically measure the strength and direction of the relationships among the variables. Third, we must confront the model with actual empirical data in order to measure and estimate the parameters of the model, and to verify the model's ability to explain and predict actual coal choice behavior.

From economic theory, we expect that for a fixed level of production, a firm will attempt to minimize costs as a strategy to maximize profits. In addition, the theory of spatial markets says that all power plants within the NGP geographical market will buy NGP coal since it is least cost. Plants outside the market will buy other coals. This apparent sharp distinction between markets however does not strictly accord with observed phenomena. The divergence between theory and fact can occur for two reasons. First, the assumptions on hypothetical behavior behind theory might not hold in the empirical world. Second, the model specification might not be appropriate for empirical estimation.

The purpose of this study is twofold. First, statistical tests of significance will be applied to the theoretical market boundary model in order to see if the theory of spatial markets has some empirical import. Second, a more appropriate model specification for the estimation of empirical spatial market phenomena will be proposed and estimated. The major focus will be upon the latter of the two research goals: to specify an <u>empirical spatial market model for Powder River coal</u>.*

Because the market boundary is better characterized by a broad band rather than a sharp line, a more appropriate spatial market model may be one that predicts the <u>probability</u> of a given plant buying Powder River coal. Such a spatial market model should tell us that the probability of a power plant choosing Powder River coal approaches 100% if the plant is closely located to Gillette Wyoming (the Powder River supply center). As plants are more closely located to competing supply centers, the probability of the plant choosing Powder River coal should decline according to a cumulative probability distribution function. This suggests that the estimation of a Powder River spatial coal market may be accomplished using a probabilistic qualitative

* This study focuses on Powder River Basin subbituminous coal because Fort Union lignite has no potential for export outside the Northern Great Plains region at this time. Lignite is lower in value per unit weight relative to other coals (Duffield et al., 1982).

response (QR) model.

QR models have never been used in the appraisement of spatial markets though there has been a recent upsurge in the use of such formulations to study discreet yes/no, either/or decision making behavior. Studies on voting behavior, choice of occupation, purchase of a consumer durable, the decision whether or not to join a union, etc. have been conducted utilizing qualitative response models (Amemiya, 1981). The use of a QR formulation in the estimation of a Powder River spatial coal market should be equally successful.

A QR model based on the logistic cumulative distribution function will be used to estimate the probabilistic spatial market. Because there are 12 other coal supply centers competing with the Powder River coal supply center, the main independent variable affecting coal choice will be the difference in total electric generating costs between using Powder River coal versus other coals as a function of plant location. This variable will be known as the cost differential. Air pollution control variables will also be included in the analysis. Observations on the location, size, construction date, and coal choice of 438 coal fired power plants coming on line between the years 1976 and 2000 have been gathered. Cost data have been supplied by Duffield et al. (1982). Computer software calculating the cost differentials have been developed by the author. The BMDP Statistical Software package (Dixon et al., 1981) available on the DEC-2060 computer at the University of Montana has the capability to run qualitative response models using maximum likelihood estimation procedures.

III. Expected Findings

It is expected that the qualitative response model will be a more appropriate, and therefore a more useful, formulation for examining spatial market phenomenon. This, in and of itself, will be a contribution to our body of knowledge. Also, statistically testing the assertions made by previous theoretical studies on spatial coal markets will enable us to establish the usefulness of the market boundary model for describing coal choosing behavior. Lastly, this research will empirically analyze the Powder River coal market thereby providing an important tool for assessing policy such as the impact of air quality regulations.

IV. Thesis Outline

A brief description of the remaining chapters in this thesis is as follows:

Chapter 2; The Theory of Spatial Markets: A Review of Past Spatial Coal Studies.

This chapter will explain the development and the structure of the theory of spatial markets. A review of past spatial coal market studies utilizing this theory will follow. The findings and methodological assumptions of previous coal market studies will be presented.

Chapter 3; Qualitative Choice in the Powder River Electric Utility Coal Market: An Empirical Probabilistic Spatial Market Model.

The qualitative response model that will be used in the empirical spatial market analysis will be developed. This chapter will describe the mathematical structure of the chosen model specification, the method through which model parameters are estimated, the interpretation of estimated regression coefficients, the statistics used to evaluate the "goodness of fit" of the model, and the selection of explanatory variables for the model.

Chapter 4; Preliminary Analysis of the Powder River Spatial Coal Market: Data Base, Cost Differential Calculations, and Simple Statistical Tests.

The data base and simplifying assumptions used in the calculation of cost differentials will be discussed. Simple statistical tests (e.g. one-way ANOVA, grouped t-tests, and Mann-Whitney tests) will be run using the power plant data base in order to test the assertions made by the theory of spatial coal markets.

Chapter 5; Estimating the Qualitative Response Spatial Market Model for Powder River Coal.

Using the model specification mentioned in Chapter 3 and the data base described in Chapter 4, the empirical spatial market model for Powder River coal will be estimated and presented. Statistical tests of the estimated parameters will be conducted. An interpretation of the estimated model will be presented. Charts showing the current 1980 Powder River spatial coal market will be shown. Concluding remarks will be made here.

V. References

Amemiya, T. "Qualitative Response Models: A Survey" Journal of Economic Literature. (December 1981, pp. 1483-1536).

Dixon, W. J. et al. BMDP Statistical Software 1981. Berkeley: University of California Press. 1981.

Duffield, J. et al. Projections of Coal Demand From the Northern Great Plains Through the Year 2010. (Final Report, OSM, May 1982).

Hyson, C. D. and Hyson, W. P. "The Economic Law of Market Areas" Quarterly Journal of Economics. (May 1950, pp. 319-327).

Power, T. et al. Montana University Coal Demand Study: Projections of Northern Great Plains Coal Mining and Energy Conversion Development. (Final Report, NSF-RANN, May 1976).

CHAPTER TWO

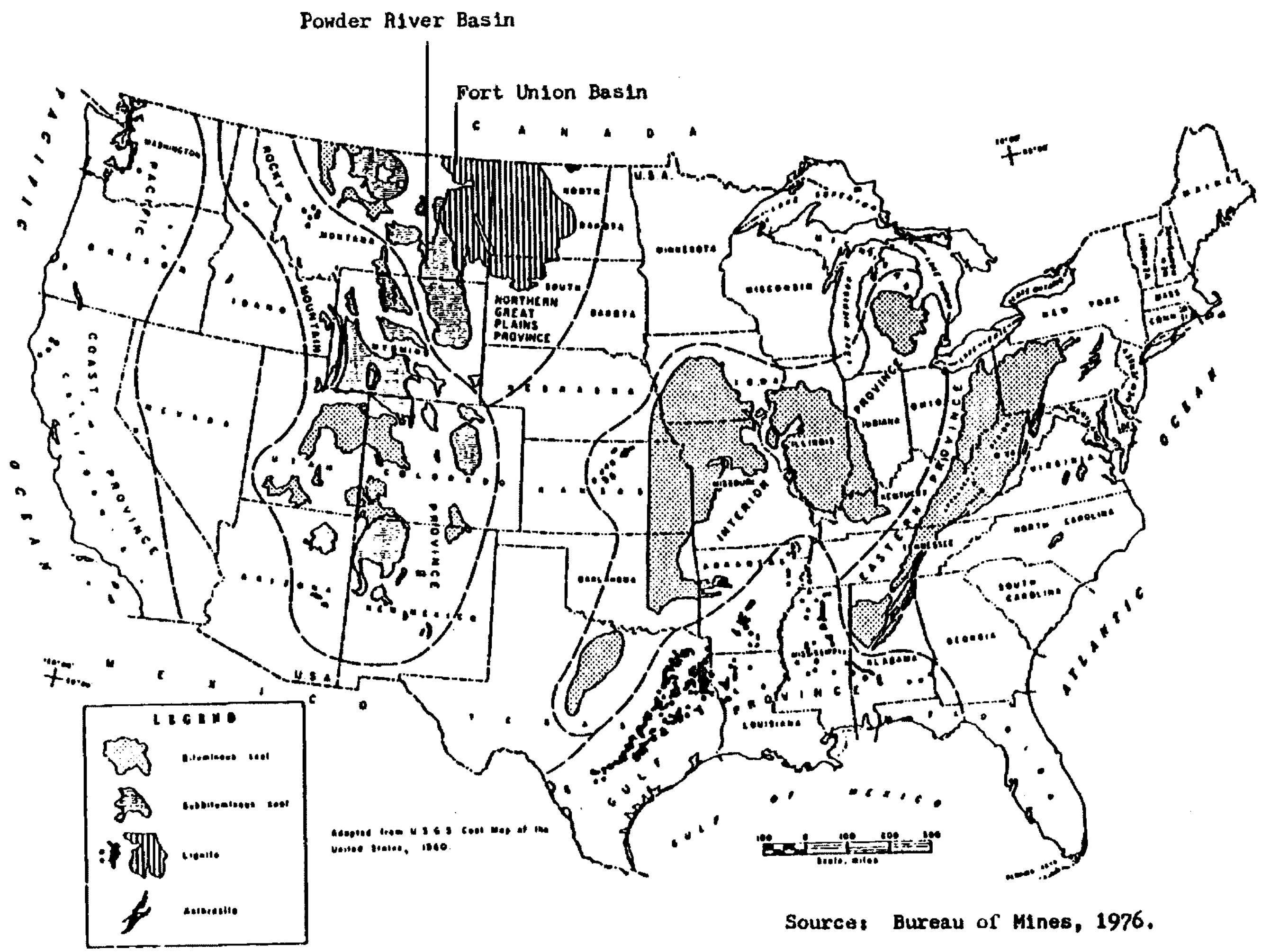
THE THEORY OF SPATIAL MARKETS: A REVIEW OF PAST SPATIAL COAL STUDIES

I. Introduction: Motivation for the Theory of Spatial Coal Markets

Since there are many developed coal fields in the United States (see Figure 2.1), different coal supply centers must compete for buyers. Assuming that coal qualities are homogeneous across regions, mine mouth coal price will obviously affect coal choice. In addition, buyers of coal are widely scattered across the continent and are usually located some distance from the sellers. Given the scattered distribution of buyers and sellers in the market, what coal will the buyer select? Distance between buyers and sellers will affect a potential buyers' choice of coal supply since coal has a relatively low market value per unit weight; transporting tons of coal across space is likely to be expensive. Intuitively, a buyer will tend to select the closest coal source assuming all other things (e.g. coal quality, and mine mouth price) being equal. The market, therefore, for a coal sold from a particular supply center is spatially constrained vis a vis competing coal supply origins.

Given this information, it seems possible to draw a somewhat unique geographical boundary between two competing

Figure 2.1



Source: Bureau of Mines, 1976.

Bituminous and subbituminous coal and lignite fields of the conterminous United States.

coal supply centers; this boundary will delineate areas where buyers prefer one coal over the other coal. Knowledge of these market boundaries will enable researchers to limit their examination of coal demand to a specific geographical region greatly simplifying the analysis. Also, if the critical factors that cause these market boundaries to shift were known, many of the determinates causing demand to shift would be known implicitly. What is needed, therefore, is a theory that will formalize the idea of market boundaries. Practitioners will then be able to define the shape and extent of market boundaries for specific commodity markets.

II. The General Theory of Spatial Markets

The above intuitive ideas on spatial market phenomena were first formalized and published under the title "The Economic Law of Market Areas" by Frank A. Fetter (1924). The market boundary between competing supply origins <u>for like goods</u> is the locus of points where the sum of price per unit and transportation cost per unit from both markets are equal. On the market boundary, a buyer is indifferent between either supply source. On either side of the boundary, one of the supply sources is strictly prefered over the other because of a cost advantage. Fetter envisioned the boundary line, that spatially separated two geographically competing markets for like goods, taking the form of a hyperbolic curve. The critical variable changing the shape of the boundary would be the price difference between both markets; Fetter assumed constant freight rates per unit distance between all points in the space being examined.

Fetter's original law was generalized by Hyson and Hyson (1950). Hysons' formulation of the "Economic Law of Market Areas" is essentially the same as Fetter's except for the assumption of constant freight rates. For example, Hyson and Hyson recognized that different modes of transport and differential topography in varying areas of the continent would cause freight rates to fluctuate between points in space. The market boundary between geographical competing markets for like goods takes the from of a hypercircle. Not only can the boundary be a hyperbolic curve; under special circumstances the boundary becomes a circle. Hyson and Hyson (1950) states the economic law of market areas as follows:

> The boundary line between the territories tributary to two geographically competing markets for like goods is a hypercircle. At each point on this curve the difference between freight costs from the two markets is just equal to the difference between the market prices, whereas on either side of this line the freight differences and the price differences are unequal. The ratio of the price difference to the ratio of the freight rates from the two markets, determine the location of the boundary line; the higher the relative price, and the lower the relative freight rate, the larger the tributary area.

As before, the locus of points where it is equally advantageous for a consumer to buy from either market is defined as the market boundary. Both price and transportation rate differences are critical in formulating and shifting the market boundary.

Campbell and Hwang (1978) showed, both theoretically and empirically, that aggregate spatial demand for a commodity is smaller than spaceless demand for the same commodity. This is especially true for commodities with a low market value per unit weight (e.g. coal) because transportation costs predominate. They show that real world conditions reflect spatial differentiation in the coal market must include a spatial element or the results will be misleading.

Abstracting from the above three articles, the theory of spatial markets makes the following assumptions:

1) _All_ buyers of a particular commodity choose to cost minimize and complete knowledge of market conditions prevail.

2) Buyers of a particular commodity have _identical_ needs and are located at varying distances away from the sellers.

3) The sellers are capable of supplying the entire geographical market and can be identified by a single point on the plane (all sellers in the market are located closely together).

4) Suppliers sell nondiscriminately to all buyers.

5) The commodity is standard or identical across different suppliers.

6) Freight charges are equal to the _distance as the crow flies_ from the market multiplied by the freight rate per unit distance between the market and the point in question.

III. Studies Estimating the Spatial Market for Electric Utility Steam Coal

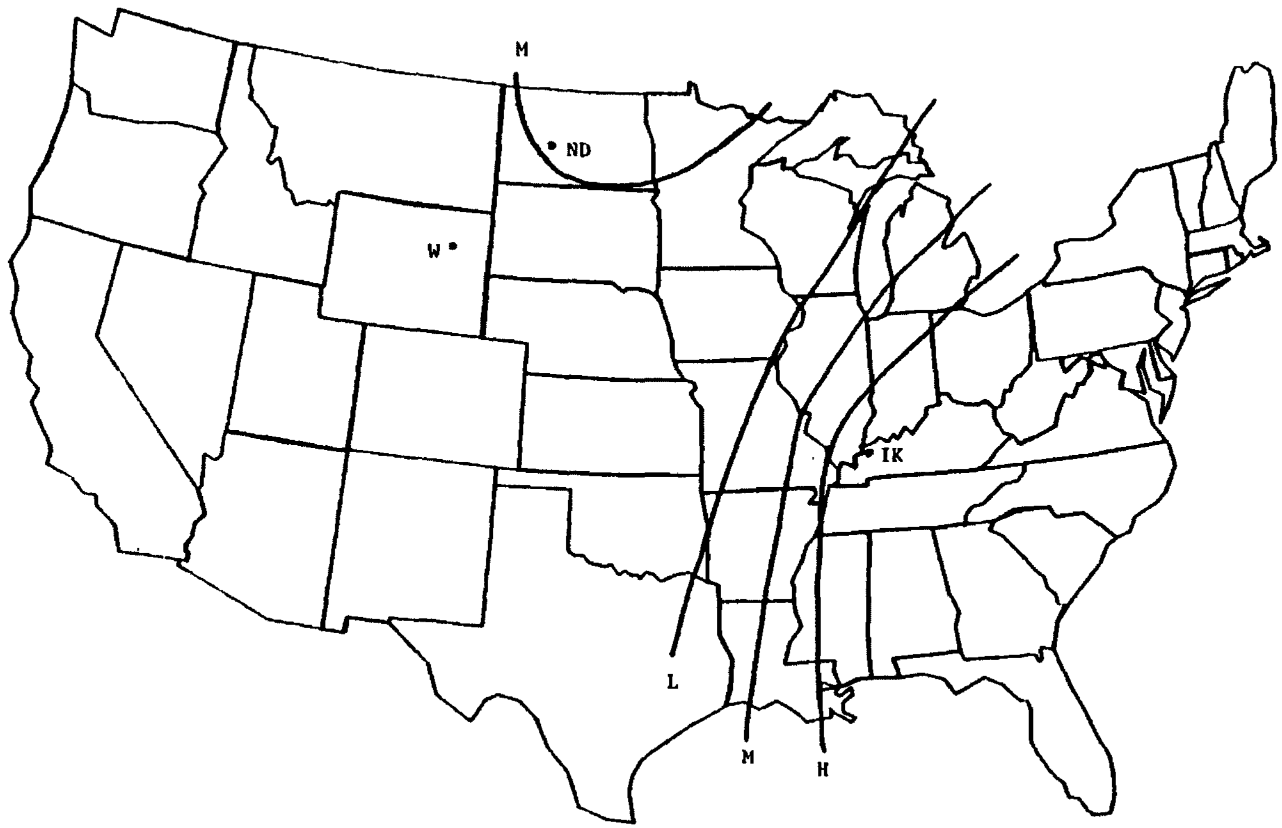
The economic law of market areas or the theory of spatial markets has been used to establish market boundaries between two competing coal suppliers. Since transportation costs and regional differences in mine mouth coal prices are important factors in coal buying behavior, the use of the theory of spatial markets to study the extent of electric utility coal markets was only natural.

Watson (1972) used the theory of spatial markets to analyze the sensitivity of wet limestone sulfur dioxide scrubbing costs on coal choice. The problem immediately facing Watson was that the theory of spatial markets assumes that the commodity being sold from different supply centers is standard or identical. Coal quality, however, varies between different supply centers in very important ways. In this case, sulfur dioxide control costs is proportional to coal sulfur content. Other important differences in coal quality between regional supply centers include coal rank (anthracite, bituminous, subbituminous, lignite, etc.) and BTU content. Boilers must be designed to burn a specific quality of coal. Instead of coal prices and transportation costs, total generating costs are used in Watson's analysis so that most of the important cost differences associated with burning a particular coal for electric generation are taken into account.

Total generating costs are a function of mine mouth prices, regional transportation rates, power plant costs, and air pollution control costs. Holding the level of pollution for both particulate and sulfur dioxide constant, Watson wanted to see how the market boundary shifts given the choice between burning low sulfur Wyoming coal (NGP coal) without scrubbing the effluent versus burning high sulfur Illinois coal and scrubbing the effluent. It was found that the spatial market was highly sensitive to the cost of scrubbing sulfur dioxide out of high sulfur coal emission. As the cost of limestone sulfur dioxide scrubbing increases, the Wyoming coal option becomes more cost attractive over a significantly larger geographical area (see Figure 2.2).

The Montana University Coal Demand Study (MUCDS) of 1976 developed a simple market definition model based upon Watson's work (Power et. al., 1976). MUCDS provided a systematic analysis defining the key factors that caused NGP coal demand to shift. Utilizing the theory of spatial markets, the critical factors that caused the market area to significantly expand or contract were identified thus helping to explain how demand for NGP coal changes. The analysis included estimating the market boundary between the NGP coal supply center (Gillette, WY) and the midwest coal supply center (Springfield, IL). The development of a NGP electric utility coal spatial market was an essential first step in a four step utility steam electric coal demand

Figure 2.2 Equal-cost Contours for Sulfur Dioxide and Fly Ash Removal Processes



IK - Illinois-Kentucky coal fields
ND - North Dakota coal fields
W - Wyoming coal fields

L - Low limestone scrubbing costs
M - Medium limestone scrubbing costs
H - High limestone scrubbing costs

Source: Watson (1972), p. 74

forecast methodology (Power, 1976):

1) Define over what geographical area (market) NGP coal can be competitively sold for electric power generation.

2) Estimate the future demand for electricity in that market area.

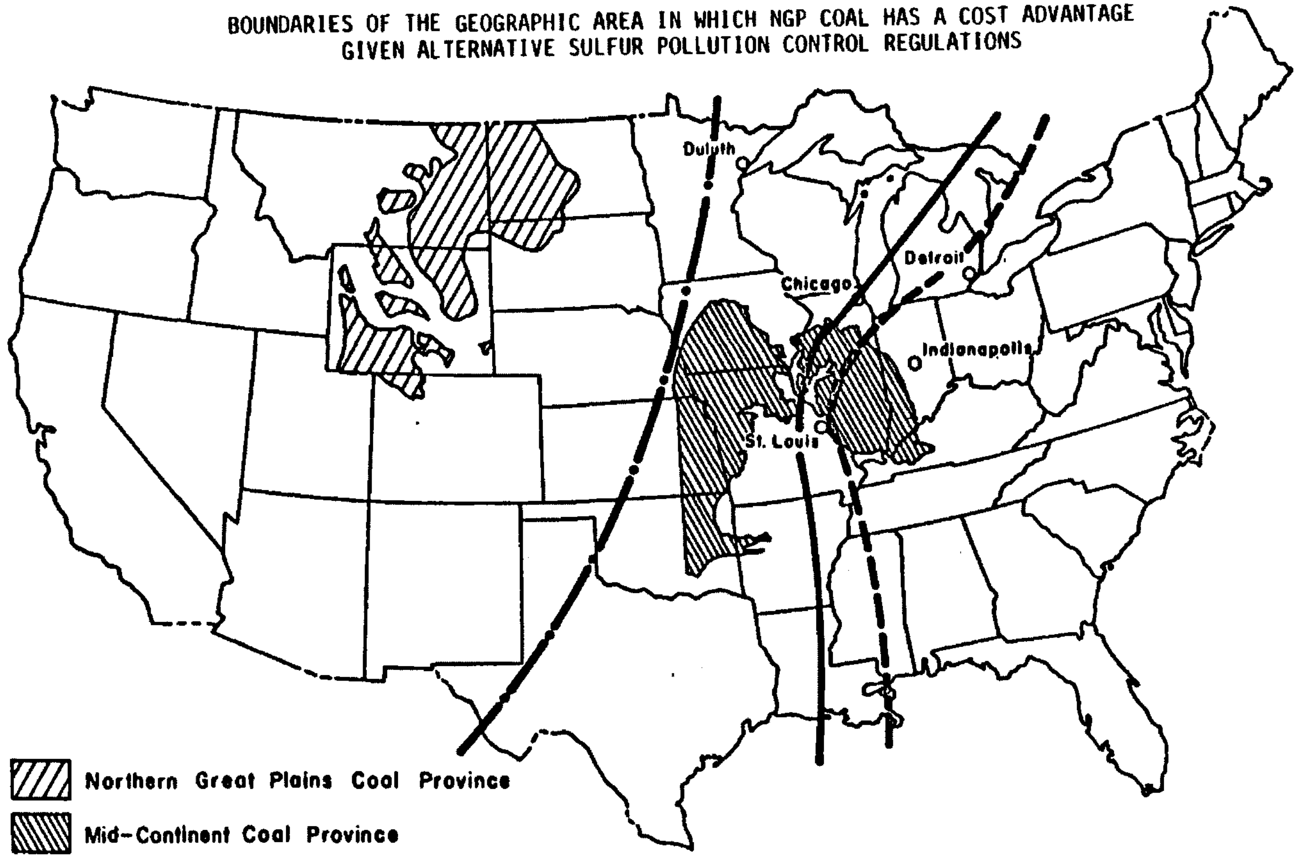
3) Determine how much of this electricity will be generated by burning coal (interfuel substitution).

4) Calculate how much of this coal will come from the NGP region (intrafuel substitution).

It was found that the demand for NGP coal was very sensitive to air pollution control regulations. If all coal fired power plants were required to have a flue gas desulfurization system (FGD), the market area for NGP electric steam coal would cover a 21 state area. If however the electric utilities were allowed to use supplementary control systems (SCS) (i.e. use tall smoke stacks and/or intermittent control systems in place of FGD), the market area shrinks to a 13 state area (see Figure 2.3). The difference in FGD pollution control costs between low sulfur coal and high sulfur coal is the source of market boundary sensitivity; it is much cheaper to scrub low sulfur coal as opposed to high sulfur coal. IF FGD is not required and SCS is permited, there is a greater cost incentive to use high sulfur Illinois coal as opposed to low sulfur NGP coal. Although the mine mouth price and transportation costs are essential in the calculation of market boundaries, MUCDS found that alternative transportation costs (all transportation was assumed to be by unit train) and relative

FIGURE 2.3

BOUNDARIES OF THE GEOGRAPHIC AREA IN WHICH NGP COAL HAS A COST ADVANTAGE
GIVEN ALTERNATIVE SULFUR POLLUTION CONTROL REGULATIONS



Northern Great Plains Coal Province

Mid-Continent Coal Province

Supplementary Control Systems Permitted on all Coal Fired Power Plants.

Flue Gas Desulfurization Required on all New Coal-Fired Plants.

Flue Gas Desulfurization Required on all Coal-Fired Plants.

pricing of alternative coal supplies did not significantly change the market definition lines; market boundaries were very insensitive to changes in prices and freight rates.

Campbell and Hwang (1978) used the theory of spatial markets to define 6 major coal market areas (the markets they defined included industrial coal as well as electric utility steam coal). Their study indicates that alternative transportation modes and differential transportation rates between regions are more important determinates in the formation of spatial coal markets than mine price differences. The paper vigorously points out however that the defined markets do not necessarily behave in an ideal manner. When one examines actual coal deliveries in the United States, one finds that geographical coal markets overlap greatly. Coal would be exported region to region across calculated market boundaries. Campbell and Hwang found that the majority of a coal supply is sold within a defined geographical market region thus partially supporting the use of the theory of spatial markets in empirical analysis. When any of the previously listed assumptions behind the theory of spatial markets are violated in the empirical world, we can expect the real world <u>not to reflect</u> a clear cut geographical market boundary solution which is suggested by theory.

By the late 1970's, significant changes in the parameters that determined the spatial market for NGP coal had occured. In 1977, congressional amendments to the Clean

Air Act took place. A new Administration backed by a substantial environmental vote took office. Mine mouth prices were increasing in real terms because of higher prices for substitute fuels, increasing extraction costs, declining labor and capital productivity, and more stringent environmental controls. Transportation costs have also substantially increased during this period. Rising costs of rail transport are making utilities question lengthening contractual agreements for cross country coal hauling by rail. Coal slurry pipelines and high voltage transmissions lines have become viable alternatives to present modes of coal transportation; these alternative modes of transport face substantial public and political opposition however (Duffield et al., 1982).

Because of these changes, another study updating the MUCDS of 1976 had commenced. The final report of this newer study was entitled "Projections of Coal Demand from the Northern Great Plains through the Year 2000" (PCDNGP, 1982). The methodological approach of the PCDNGP was exactly the same as in the MUCDS; the definition of the NGP spatial electric steam coal market was the first step in a four step NGP coal demand estimation methodology. The NGP electric steam coal market area is defined such that NGP coal is the least total generating cost coal over the lifetime of a new model coal-fired electric power plant. The PCDNGP was more sophisticated than the MUCDS in that:

1) The researchers sought to completely bound the

U.S. market for NGP coal while the MUCDS looked only at one boundary (the Illinois: NGP boundary).

2) Since electricity forecasts were disaggregated only to the state level, a method must be devised to further partition the state level forecast if the calculated market boundary bisected the state. The PCDNGP used the percentage of the state's population falling in the NGP geographical market as a proxy for the percentage of electricity, in the state level forecast, that is used in the NGP area; the spatial NGP coal market is population weighted for electricity demand forecasts. This population weighing does not affect the definition of market boundaries however.

3) The impact of real escalating costs on the market was investigated. Market areas for the years 1980, 1990, and 2000 were defined. Low, base (best guess), and high real cost escalation scenarios and their effects on the NGP coal market area were investigated.

The PCDNGP found that the future size and shape of the NGP spatial coal market was highly dependent upon how certain real costs changed over a period of time. In general, all real costs increase over time. Cost increases were not constant across all relevant cost categories however. The real cost escalation rates between power plant capital equipment, power plant operating and maintenance, transportation, and coal prices all differ. Table 2.1 contains the real cost escalation multipliers used in the PCDNGP analysis. Current costs are multiplied by the real cost escalation multiplier so that the impact of real cost increases on the size and shape of the NGP spatial market for a particular forecast year may be measured. Base real cost escalation rates were determined using a 15-year

Table 2.1

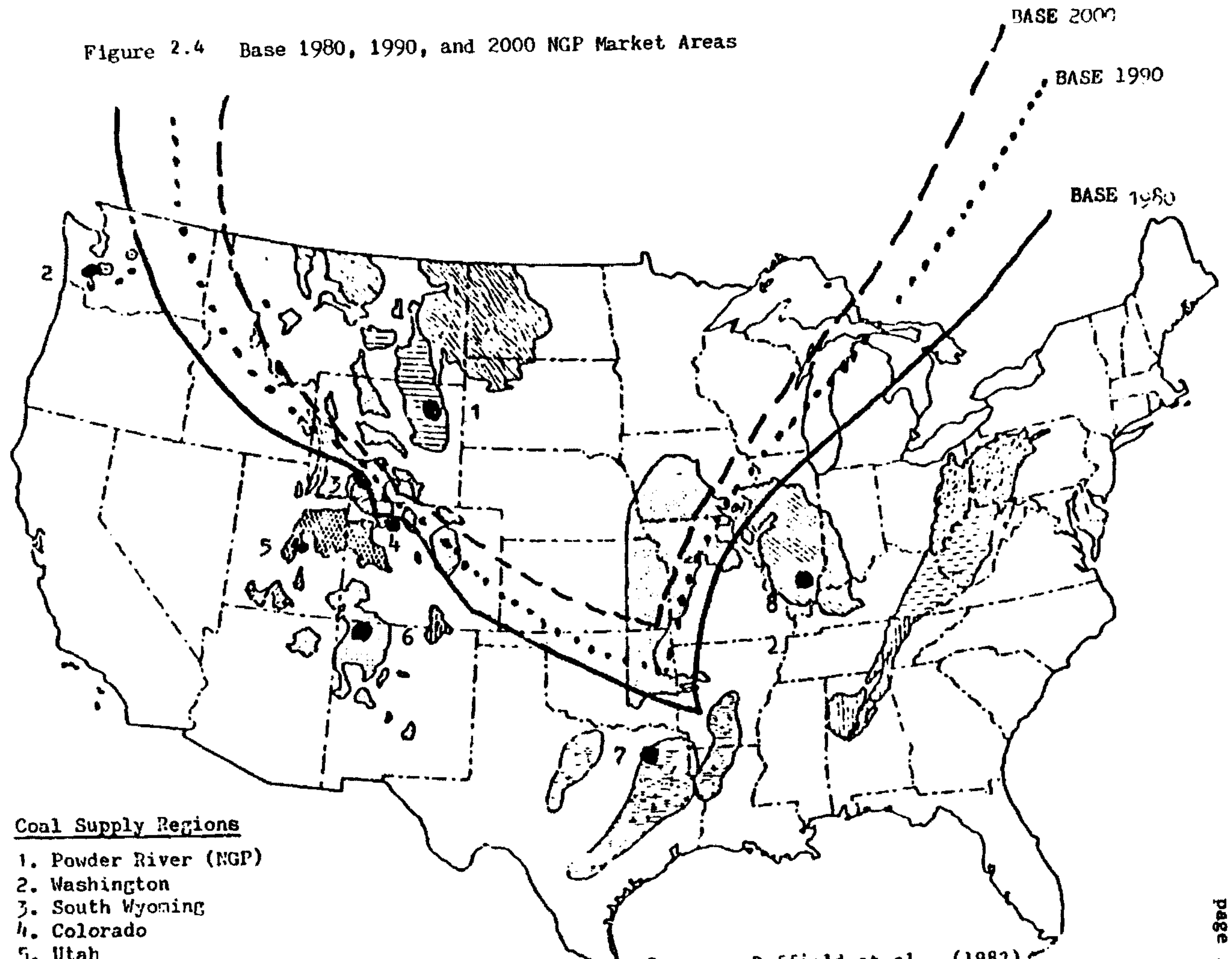
Real Cost Escalation Multipliers

Input Parameter	1980 Low	1980 Base	1980 High	1990 Low	1990 Base	1990 High	2000 Low	2000 Base	2000 High
Base Plant									
Capital	1.000	1.000	1.000	1.127	1.255	1.384	1.269	1.576	1.914
O & M	1.026	1.170	1.343	1.047	1.318	1.669	1.068	1.485	2.075
SO_2 Control									
Capital	1.000	1.000	1.000	1.010	1.051	1.094	1.020	1.105	1.196
O & M	1.026	1.170	1.343	1.047	1.318	1.669	1.068	1.485	2.075
Railroad Rates	1.462	1.619	1.826	1.927	2.284	2.782	2.540	3.221	4.238
Coal Prices									
Underground	1.013	1.17	1.505	1.023	1.399	2.123	1.033	1.672	2.995
Surface	1.000	1.11	1.236	1.041	1.353	1.795	1.083	1.649	2.606

Source: Duffield et al., (1982), p. 8-28 historical trend time series analysis. The lower 95% confidence interval prediction band of this time series analysis was defined as the low real cost escalation scenario; high real cost escalation rates are based upon the upper 95% confidence interval prediction band.

In contrast to the MUCDS, the PCDNGP found that the key swing variables determining significant shifts in the market boundaries included mining labor costs (which affects mine mouth prices) and rail transportation rates. These two variables are equally, if not more significant, in shifting the market boundary over time as air pollution control policy. It was predicted that from 1980 to 2000, the NGP electric utility steam coal market would shrink over time. In 1980, the base case spatial market included 18 states. Because of the effect of real increases in transportation costs, the market shrinks to 12 states by the year 2000 (see Figure 2.4). Given a particular forecast year, if real escalation rates in the cost of capital, operating/maintenance, transportation, and coal prices were assumed to be higher than in the base case, a larger NGP market would be witnessed. This is because escalation rates in coal prices, for a given forecast year, between capital intensive western strip mined coal and labor intensive eastern deep mined coal are different while escalation rates in all other costs are constant between regions. High escalation rates result in a larger NGP coal market because the assumed high escalation rate for the cost of labor

Figure 2.4 Base 1980, 1990, and 2000 NGP Market Areas



Coal Supply Regions

1. Powder River (NGP)
2. Washington
3. South Wyoming
4. Colorado
5. Utah
6. New Mexico
7. Texas
8. Midwest

Source: Duffield et al., (1982) p. 8-66

significantly increases the cost of underground mining; the price of deep mined coal will rise faster than the price of striped mined NGP coal. Conversely, lower real escalation rates in all costs result in a smaller NGP coal market because of lower real cost increases for labor intensive underground mines (see figure 2.5).

Finally, the study points out that the market boundary is better characterized by a broad band rather than a sharp line. As in Campbell and Hwang (1978), the PCDNGP (1982) finds that there is some market overlap; not all the generating plants on the NGP side of the boundary use NGP coal. Also, plants on the non-NGP side of the border use NGP coal. This occurs because the real world violates the simplifying assumptions spatial market theory makes. The PCDNGP (1982) market model makes the following assumptions:

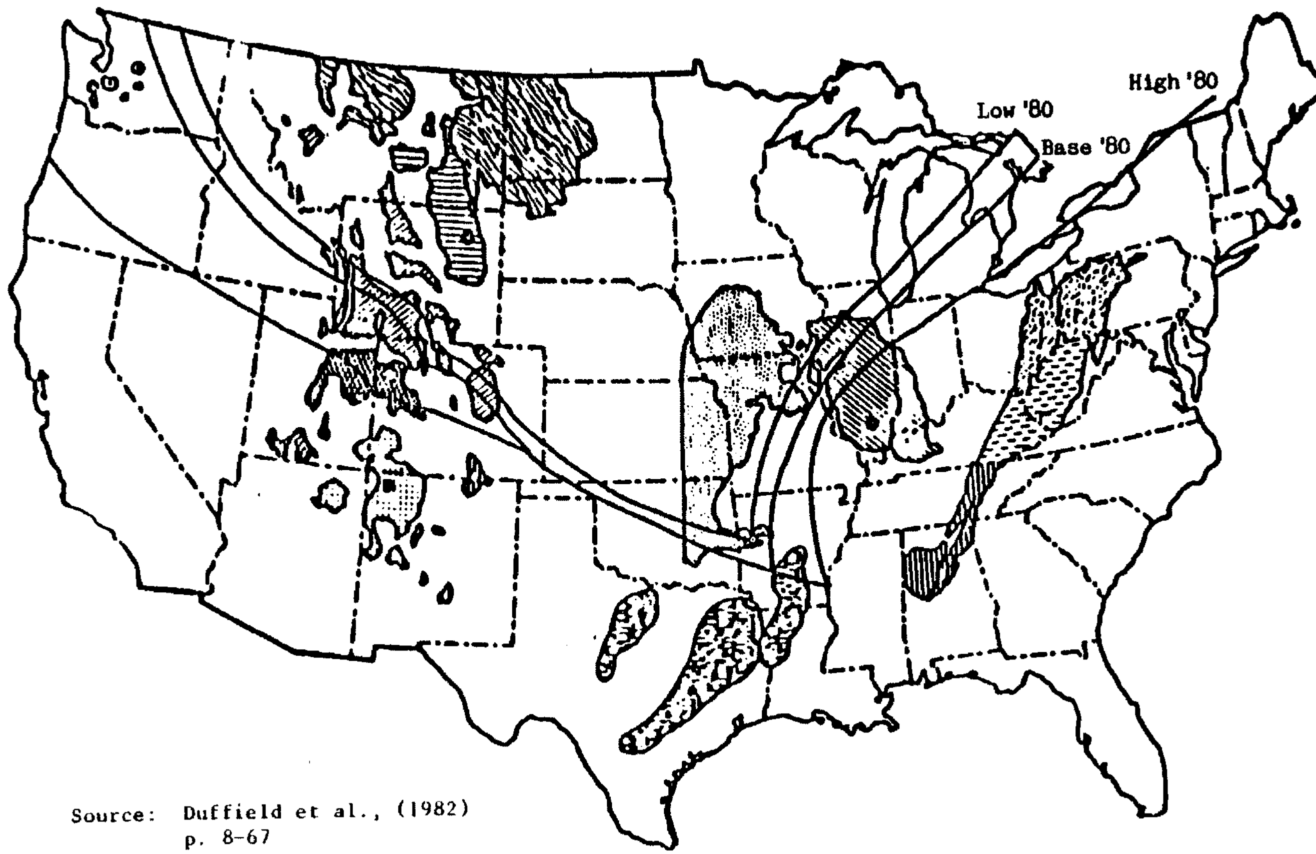
1) A utility company which decides to build a new large baseload coal-fired power plant will base its coal choice on an informed lifetime least cost analysis.

2) The power plant has a 500 megawatt net capacity and has a base load lifetime capacity factor of 65%. Plant location decisions are independent of coal choice decisions.

3) The coal regions offering significant future competition to NGP are Colorado, Illinois, New Mexico, South Wyoming, Texas, Utah, and Washington. The coal supply regions are identified by a single point known as the coal supply center. Suppliers sell nondiscriminately to buyers.

4) A coal supply region's coal is characterized by 3 factors: coal rank, BTU content, and percent sulfur. Prices are based on current long term contract prices.

Figure 2.5 1980 Base, Low, and High NGP Market Areas



Source: Duffield et al., (1982)
p. 8-67

5) All coal is transported by unit train in the model. The existing rail network is assumed complete enough to allow uniform coal distribution.

These assumptions serve to simplify the analysis by reducing the number of seemingly insignificant complicating factors. A simple analysis reduces the the amount of effort required to model spatial market phenomena and makes the research more understandable. Of course, it is hoped that the amount of explanatory power lost through simplification is minimal. It is not difficult to imagine however that one or more of the modeling assumptions will not accurately reflect real world conditions thus impairing computed results.

Although the sharp demarcation of coal buying behavior indicated by a single market boundary line does not strictly exist in reality, Duffield et al. (1982) feels that the spatial market model has explanatory and predictive power. Actual coal contract data seems to support general spatial market theory by showing that a majority of electric generating facilities using coal contract for a coal supply from within their respective market areas. Appropriate statistical tests on such a claim however have not been accomplished by the authors of the study.

Since spatial market studies are often used as an intermediate step for forecasting NGP coal demand, it might be interesting to compare spatial coal forecasts with forecasts that use a different approach. Large linear programing models that forecast national coal demand have been disaggregated to forecast regional levels of development. Two such models include the U.S. Department of Energy (1981) and ICF Inc. (1980). Table 2.2 compares spatial model forecasts to these linear programing forecasts. The linear programing model approach include the same cost categories (e.g. transportation costs) as the market boundary model approach.

It could be seen that linear programing models forecast much higher levels of coal development than the spatial studies. Proponents of spatial studies feel that large linear programing models, which were originally designed to forecast national coal demand, do not produce accurate regionally disaggregated results (Duffield et al., 1982). The existing linear models do not enable the researcher to do satisfactory sensitivity analysis because altering the population from which the data sample is drawn decreases the power of statistical tests finding significant results when the null hypothesis is false; linear models increase the possibility of Type II error (failing to accept the alternative hypothesis when it is true). The linear model therefore, lacks robustness in the face of parameter changes (Duffield et al., 1982).

IV. Summary

The ultimate purpose of spatial electric utility coal market analysis is to provide a systematic method that identifies the critical variables which influences and cause

Table 2.2 Northern Great Plains Coal Forecast Comparisons
(million tons per year)

Forecast/Market	Year 1985	1990	1995	2000	2010
(A) Montana University Coal Study:					
Electric Utility		145-202		145-374	177-493
Industrial		6-8		6-15	7-20
Export				13-26	15-50
Synfuel				30-40	90-160
Total		151-210		194-455	289-723
(B) U.S. Department of Energy: (1981)					
Residual (Utility plus industrial)	203-239	194-401	253-739		
Synthetic Fuels	12	32-43	42-141		
Total	215-251	226-444	295-880		
(C) ICF, Inc. (1980)					
Electric Utility	153-213	171-395	173-607		
Industrial	2	3	4		
Synthetic Fuel	6	16	32		
Total	161-221	190-414	209-643		

Source: Duffield et al. (1982), p. 1-43.

change in the size of the market area for a particular coal. This information can in turn be used to partially explain shifts in the demand for a specific region's coal. The economic concept of market areas or the theory of spatial markets defines market boundaries between competing supply origins of like goods as the locus of points where the sum of price per unit and transportation costs per unit from both supply centers are equal. A hypothetical buyer located on the market boundary is indifferent when it comes to choosing supply source assuming the buyer wishes to minimize cost. On either side of the market boundary, therefore, one of the supply sources is strictly preferred over the other because of a cost advantage. The mathematical specification of the market boundary is hyperbolic in nature.

Coal is a commodity characterized by low value per unit weight and is mined at different locations on the continent. Since transportation costs and regional differences in minemouth coal prices are significant, the use of the theory of spatial markets studying the extent of electric utility coal markets seemed to be a logical choice. Since, the theory of spatial markets assumes that the goods from different supply centers are identical, a serious problem arises since coals across regions are not identical. Total generating costs therefore are used in the analysis so that most of the important cost differences associated with burning a particular coal for electric generation are taken into account.

The coal market studies cited found that the critical swing variables affecting shifts in market boundaries include sulfur dioxide air pollution control costs and policy, mining labor costs (affecting mine mouth prices) and transportation rates. These studies also seem to indicate that real spatial coal markets overlap significantly while the present spatial market model defines unique hyperbolic market boundary lines. The appropriateness of this particular spatial model specification for explaining and predicting empirical world behavior is called into question. This will discussed at length in the next chapter.

V. References

Campbell, T. C. and Hwang, M. "Spatial Analysis and the Identification of Coal Markets" The Journal of Energy and Development. (Autumn 1978, Vol IV, No. 1, pp. 104-125).

Duffield, et al. Projections of Coal Demand From the Northern Great Plains Through the Year 2010. (Final Report, OSM, May 1982).

Fetter, F. "The Economic Law of Market Areas". Quarterly Journal of Economics. (May 1924, pp. 520-529).

Hyson, C. D. and Hyson, W. P. "The Economic Law of Market Areas" Quarterly Journal of Economics. (May 1950, Vol 64, pp. 319-327).

ICF, Incorporated, Forecasts and Sensitivity Analysis of Western Coal Production. (November, 1980).

Power, et al. Montana University Coal Demand Study: Projections of Northern Great Plains Coal Mining and Energy Conversion Development 1975 to 2000 A.D. (Final Report, NSF-RANN, May 1976).

U.S. Department of Energy. The 1980 Biennial Update of National and Regional Production Goals for 1985, 1990, and 1995. (March, 1981).

Watson, W. D., Jr. "Cost of Air Pollution Control in the Coal-fired Electric Power Industry" Quarterly Review of Economics and Business. (Autumn 1972, Vol 12, No. 3, pp. 63-85).

CHAPTER THREE

QUALITATIVE CHOICE IN THE NGP ELECTRIC UTILITY COAL MARKET: AN EMPIRICAL PROBABILISTIC SPATIAL MARKET MODEL.

I. Introduction: A More Systematic and Appropriate Empirical Spatial Market Model.

Spatial electric utility coal market analysis deals with the question: for a given coal fired power plant, what geographic area will supply the coal the power plant will use? The purpose of this analysis is to estimate a descriptive model which identifies the key swing variables that impact electric utility coal source choice. Up to the present, researchers analyzed coal choice behavior by directly utilizing the theory of spatial markets in a geographical coal market model known as the market boundary model. This model draws distinct market boundaries between competing coal supply centers as a function of total relative costs of burning one coal versus another coal for electric generation. Electric utilities would theoretically buy coal from the supply center within their boundary area because it was least cost.

The market boundary model analyzes electric utility coal choosing behavior using a "deductive/hypothetical" approach. This model deduces (calculates or draws) market boundary lines on the hypothesis that electric utilities will base coal choice strictly on cost minimization criteria given "simplified facts" (or assumptions) about real world conditions (see Chapter Two). The calculated market boundary is an exact mathematical result of hypothetical cost minimizing power plants operating under justifiably simplified artificial conditions. Since the model specifies exact relationships between the variables, the market boundary model solution is a deduction: a logically correct and conclusive inference. The market boundary solution, as an explanation of electric utility coal buying behavior however, can be useful only if market boundaries exist in the empirical world. It follows that the market boundary model solution will have empirical import if the hypothetical conditions and behavior assumptions are at least approximately true in the empirical world. Since the market boundary model assumes simplified world conditions, the market boundary model can be a viable explanation of empirical spatial markets insofar as the market boundary solution adequately discriminates which geographic area will supply the coal actual power plants will use.

Both Campbell and Hwang (1978) and Duffield et al. (1982) find that a certain amount of market overlap exists in the real world. That is, after hypothetical market boundaries have been drawn, real power plants were observed to buy coal from a source on the other side of the market boundary. The frequency of power plants buying coal from the supply source on the "wrong side" of the boundary was high when the plants were located close to the boundary;

the frequency of power plants buying coal from the "wrong side" of the border decreased as you moved away from the boundary toward the supply center. This phenomena caused Duffield et al. (1982) to say that market boundaries are better characterized by broad bands rather than unique sharp lines.

Because real world spatial markets overlap, the empirical significance of the calculated market boundaries are called into question. Can the calculated market boundary adequately discriminate NGP coal users from non-NGP coal users? One way to test the market boundary model for statistical significance is to gather a sample of coal fired power plants noting their geographical location and where they actually purchase their coal. One would expect those coal fired generators within the calculated market area to buy coal from within the market area. A statistical test of significance may then be applied to see if the calculated market boundary can discriminate electric utility coal choice. If the calculated market boundary does not adequately discriminate NGP coal users from non-users, one could inductively modify the model assumptions or transform the data through a process of trial and error in order to make the model fit empirical phenomena. Even if the market boundary solution were found to be significant however, at this time there is no clear procedure to formally test the market boundary solution for "goodness of fit." Duffield et al. (1982) gathered coal contract data to see if new plants within the NGP coal market area were planning to buy NGP coal. Without a formal statistical test, Duffield et al. (1982) felt that the market boundary model discriminates electric utility coal choice behavior fairly well. Formal statistical tests of significance examining the implications set forth by the market boundary model are presented in Chapter Four of this study.

A different approach to modelling spatial coal buying behavior involves an "inductive/empirical" method. In contrast to the deductive/hypothetical approach where inferences are deduced from a given set of behavior assumptions and simplified world conditions, the inductive/empirical approach describes spatial coal buying behavior by drawing inferences from <u>repeat</u> observations of <u>actual</u> coal buying behavior as a function of power plant location and other "theoretically important" variables. Both methodologies initially start with theory so that the important relationships between the variables may be identified. There is a notable difference between these two approaches however. Deductive/hypothetical models specify <u>exact</u> relationships between the dependent and independent variables; the strength and direction of the relationships between the dependent and independent variables are determined apriori. Inductive/empirical models use observable empirical data to determine, aposteriori, the strength and direction of the relationships between the dependent and independent variables; these models describe the relationships between dependent and independent factors in terms of <u>tendencies</u> rather than in exact terms because this approach recognizes the element of randomness in behavior. The inductive/empirical approach involves econometric regression techniques which use quantifiable data in order to measure the strength and direction of <u>statistical</u> correlations between the dependent and independent variables. In this case, the dependent variable of such a model is based on whether or not utility companies buy NGP coal. The explanatory variables in the model include those factors which, we feel, strongly affect electric utility coal choosing behavior; theory gives us an idea about which factors affect the dependent variable.

Since empirical market boundaries are characterized by broad bands rather than unique sharp lines, a more systematic and appropriate analysis of empirical spatial coal markets and electric utility coal choice is an inductive/empirical approach that predicts the tendency or <u>probability</u> a given plant will use NGP coal; the probability of a plant using NGP coal declines according to some cumulative distribution function as power plants are located farther away from the Northern Great Plains. Because inductive/empirical models estimate true population parameters using empirical data and inductive statistical techniques, the estimated parameters of these models can be easily and systematically tested for statistical significance and goodness of fit. In addition, for a given model specification and a given set of empirical data, a regression technique that permits us to calculate the "best fitting" model can be selected (i.e. the estimated parameters will be calculated so that the model is most consistent with observed data). The remainder of this chapter will discuss and develop an inductive/empirical approach to modelling empirical electric utility coal buying behavior.

II. A Probabilistic Qualitative Response Spatial Market Model for Powder River Coal

A class of regression models that would be appropriate for explaining empirical electric utility coal buying behavior is known as probabilistic qualitative response (QR) models. This class of models was designed to explain and predict human choice behavior where the behavioral response is observationally qualitative (discreet, categorical) rather than continuous and quantitative. The categorical response may be binomial (i.e. yes/no, success/failure, buy/not buy, etc.) or multinomial (e.g. alternative 1, 2, 3, ...). In the binomial case, individuals are assumed to face a choice between two alternatives; the choice they actually make depends upon characteristics (or attributes) of the individuals. On the basis of the choices individuals make and the attributes they possess, a qualitative response model determines the probability that an individual with particular characteristics will make one choice rather than the alternative. QR models recognize the fact that behavior can not be predicted with absolute certainty; human behavior is better expressed in terms of tendencies. As opposed to previous hypothetical spatial coal market studies, parameters of QR models are estimated using a regression technique that selects the best fitting coefficients for a given model specification and set of observed data. These estimation procedures also yield statistics which enable the practitioner to test the estimated coefficients for significance; The QR model is estimated and statistically tested simultaneously.

A QR formulation will be used to study electric utility coal choice behavior where the dependent variable is the dichotomous "buy Powder River coal/not buy Powder River" coal decision.* Let Yi = 1 when a power plant "i" is observed to buy Powder River coal and Yi = 0 when no Powder River coal is purchased. As an example, if we assume that the probability of an individual power plant making a given coal choice is a <u>linear</u> function of power plant attributes, we may write our coal choice model as:

$$Pi = Prob(Yi = 1) = F(x'B) = F(B0 + B1X1i + B2X2i + \ldots + BjXji)$$

Where:

i = 1, 2, 3, ..., N = the ith power plant unit.

---

* Instead of estimating a spatial market for both Powder River and Fort Union coal, our empirical spatial market model will focus attention on Powder River coal. Fort Union coal has no extra-regional export potential at this time.

Pi = the probability that power plant i buys Powder River coal.

F = a cumulative probability distribution function.

Xji = the jth attribute value (explanatory variable) of the ith power plant. (x' = the vector of power plant attributes Xji).

Bj = the jth model parameter (coefficients on the explanatory variables except for the constant B0).

What is critical at this point is selecting the functional form F.

A. Choosing the Mathematical Specification F.

In the literature, there are three common probability functional forms used in QR specifications (Amemiya, 1981). They include:

1) The Linear Probability Model (LP):

$$Pi = F(x'B) = x'B + Ui = B0 + B1X1i + \ldots + BjXji + Ui$$

where: Ui = an independently distributed random error term.

2) The Probit Model:

$$Pi = F(x'B) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x'B} e^{-s^2/2}\, ds$$

where: e = the Napierian logarithm (approximately 2.7183).

s = a random normal variable with zero mean and unit variance.

3) The Logit Model:

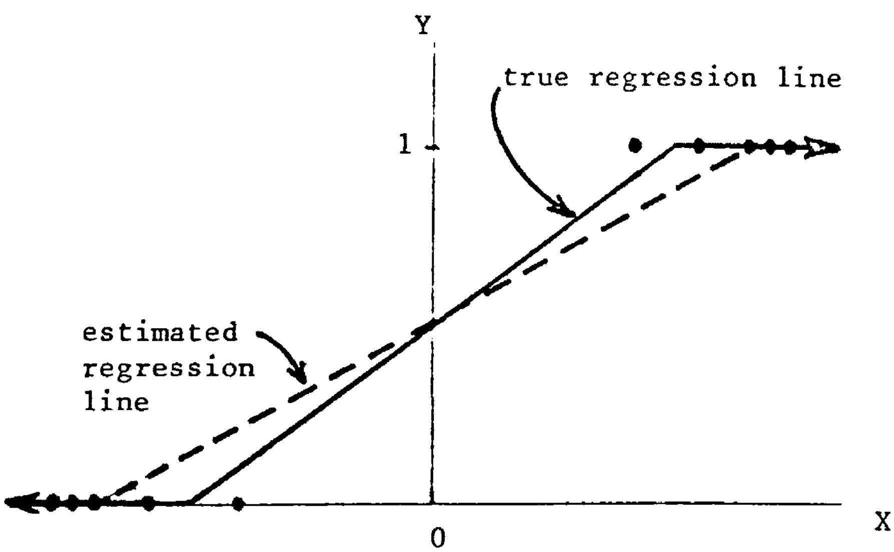
$$Pi = F(x'B) = [\ 1 + e^{-(\ x'B\ )}\ ]^{-1}$$

The linear probability model is nothing more than a simple linear regression where Yi is directly regressed on x' to get estimates of B and predictions of Pi. The beauty of this model is its computational simplicity. Ordinary least squares (OLS) may be used to estimate the coefficients of the model. There are several reasons however why the use of this model is discouraged by researchers (Pindyck and Rubinfeld, 1981). First, Pi is not constrained in the unit interval (0,1) as probabilities should. The model must be artificially constrained in this way:

$$Pi = \begin{cases} x'B \text{ when } 0 < x'B < 1 \\ 1 \text{ when } x'B > 1 \\ 0 \text{ when } x'B < 0 \end{cases}$$

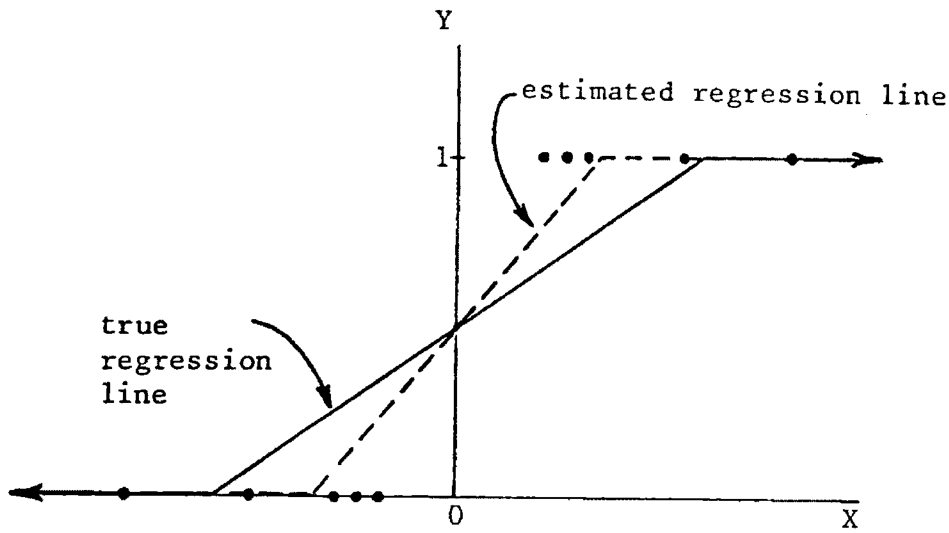
Second, the variance of the error term Ui is heteroscedastic (i.e. the variance of Ui is not constant for all Xji). The estimates of B are consistent but inefficient if OLS is used. Weighted Least Squares (WLS) may be used to yield consistent and asymptotically efficient estimates of B. However, WLS does not yield efficient estimators for small samples. Third, LP model estimates of B using any regression technique are likely to biased because constraining Pi to the unit interval causes a problem known as "data bunching." If sample observation include many extreme values of attribute Xji, the slope of the regression line will be under estimated (see figure 3.1).

Figure 3.1 The Effects of Data Bunching
(slope under estimation)



The slope may be over estimated when data are bunched differently (see figure 3.2).

Figure 3.2 The Effects of Data Bunching
(slope over estimation)



Fourth, since the random error term is not normally distributed, regular tests of significance can not be applied.

Because of these problems, other qualitative response models have been proposed. The most serious drawback of the LP model is that Pi is not automatically constrained to the unit interval. What we wish to find is a model that allows predicted probabilities to vary <u>within</u> the interval (0,1) over all attribute values x' which may range over the entire real line. Ideally, increases in the value of x' will be associated with increases or decreases in the predicted probability Pi. This suggests that a suitable model will incorporate the use of a cumulative probability distribution function. Probability values in a cumulative probability distribution function are automatically constrained within the (0,1) interval. Though there are many different cumulative probability distribution functions, only two specifications are predominately found in the literature. They include the cumulative normal probability distribution function on which the probit model is based and the cumulative logistic probability distribution function on which the logit model is based.

The cumulative logistic probability distribution function closely approximates the cumulative normal probability distribution function. In fact, both distributions are so similar that one can not distinguish then statistically unless one has an extremely large number of observations (Chambers and Cox, 1967). Figure 3.3 and Table 3.1 show the similarity between the probit and the logit formulations. The only difference is that the logistic formulation has slightly fatter tails compared to the probit (Pindyck and Rubinfeld, 1981).

Hartman (1979) mentions that in the binary choice case, the logit and probit formulations yield essentially the same results in most applications to date. Amemiya (1981) feels that the choice between the probit and logit models is unimportant because of their similarity. The main advantage of the logit model over the probit model is computational simplicity; since the logit formulation is much simpler to work with, computational costs are greatly reduced. In this

Similarities Between the Cumulative Normal and the
Cumulative Logistic Probability Distribution Functions

Figure 3.3

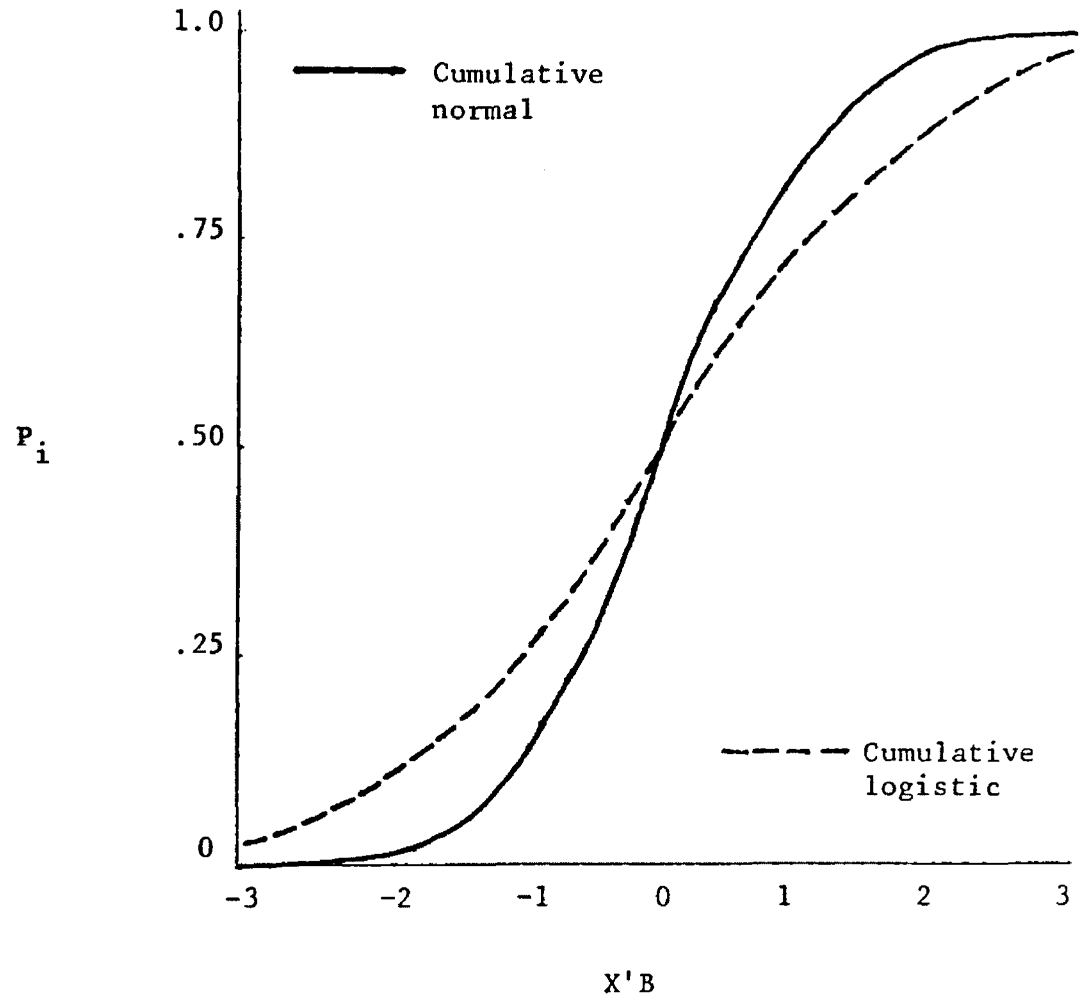


Table 3.1

X'B	Normal	Logistic
-3.0	.0013	.0474
-2.0	.0228	.1192
-1.5	.0668	.1824
-1.0	.1587	.2689
-0.5	.3085	.3775
0.0	.5000	.5000
0.5	.6915	.6225
1.0	.8413	.7311
1.5	.9332	.8176
2.0	.9772	.8808
3.0	.9987	.9526

study therefore, the logit model, based on the cumulative logistic probability distribution function, will be used to analyze electric utility coal choice behavior.

B. Model Theory and Coefficient Estimation

To see how the logit model works, imagine the following. Suppose an event E occurs when a utility company buys Powder River coal. The utility company decides to fuel their generators with Powder River coal as opposed to other coals when expected net returns of such an action are "sufficiently high." What constitutes a "sufficiently high" net return depends upon individual power companies. Now assume there exists a theoretical and unobservable index value Ii where Ii = B0 + B1X1i. Index Ii is determined by the explanatory variable X1i and is linear in the parameters B0 and B1. As Ii increases, the probability that E occurs (the buying of Powder River coal) increases as well. Since the probability that E will occur must fall between zero and one, the monotonic relationship between the index Ii and Prob(E | Ii) (i.e. the probability E occurs given index Ii) must assume the general form of a cumulative probability distribution function.

Each utility company with a particular index value Ii will make a choice between E (buy Powder River coal) and not-E (not buy Powder River coal). They make this decision by comparing their Ii with a critical cutoff value or threshold level Ii*. Stated formally, the

$$\text{utility company buys} = \begin{cases} \text{Powder River coal if } Ii > Ii^* \\ \text{non-Powder River coal if } Ii < Ii^* \end{cases}$$

The individual threshold level Ii^* is assumed to be determined by many independent factors.

By the central limit theorem, $Ii^* \sim N(\mu, \sigma^2)$. That is, Ii^* is distributed normally with mean μ and variance σ^2. This suggests that the relationship between index Ii and Prob($Ii > Ii^*$), where Prob($Ii > Ii^*$) represents the probability of E occuring given index Ii, is best described by a cumulative normal probability distribution function. As explained above, however, the cumulative logistic probability distribution function closely approximates the normal. Because of this feature, Ii^* may still be assumed to be a normal random variable when using a logit transformation (Judge, 1980). The logit model is specified as:

$$\begin{aligned} Pi &= \text{Prob}(\, E \mid Ii \,) = \text{Prob}(\, Ii > Ii^* \,) \\ &= \text{Prob}(\, B0 + B1X1i > Ii^* \,) \\ &= F(Ii) = F(\, B0 + B1X1i \,) \\ &= [\, 1 + e^{-(\, B0 + B1X1i \,)} \,]^{-1}. \end{aligned}$$

How do we estimate the logit parameters Bj? In standard regression procedures, observations on both the dependent and independent variables for each individual in the sample are required in the estimation of model parameters. In our model however, the dependent variable Pi is not observed. Instead our dependent variable is a fixed random sample of <u>independent</u> observations on whether or not utility companies bought Powder River coal. As before, let:

$$Yi = \begin{cases} 1 \text{ if Powder River coal is purchased .} \\ 0 \text{ if no Powder River coal is purchased .} \end{cases}$$

N = total number of observations .

The regressor X1i is a non-stochastic explanatory (independent) variable whose values may be continuous or discreet. Since the relationship between Yi and X1i takes the form of a logit transformation, our objective is to select coefficients Bj of the equation

$$Pi = F(\ B0 + B1X1i\ ) = [\ 1 + e^{-(\ B0\ +\ B1X1i\ )}\ ]^{-1}$$

which make it <u>most likely</u> for the above model to have given rise to the observed pattern of choices Yi (given observations on X1i). In other words, we wish to select parameters Bj which <u>maximizes the total probability</u> of observing all sample observations. In order to accomplish this, we must estimate the <u>joint probability</u> of obtaining all the observed Yi values (given X1i) <u>for each</u> possible combination of Bj and then choose the parameters Bj which maximize the joint probability of the observed sample values (Koutsoyiannis, 1979). This method of parameter estimation is known as the maximum likelihood (ML) method. The function we wish to maximize has the form:

$$L = \text{Prob}(Y1,Y2,...,YN) = \text{Prob}(Y1)...\text{Prob}(YN) .$$

We assume that each individual power plant decision independent of each other power plant decision. If we let the first n1 observations be those where Yi = 1 and the last n2 observations be where Yi = 0, then function L (also known as the likelihood function) reduces to:

$$L = P1...Pn1(1 - Pn1+1)...(1 - PN) .$$

This equation follows because Prob(Yi=1)=Pi and Prob(Yi=0)=(1-Pi).
Where $\prod$ represents the product of a number of factors, the likelihood function to be maximized may be reduced to:

$$L = \prod_{i=1}^{n1} Pi \prod_{i=n1+1}^{N} (1-Pi)$$

$$= \prod_{i=1}^{N} Pi^{Yi} \quad (1-Pi)^{(1-Yi)}$$

Since $Pi = F(B0 + B1X1i)$ where $F = [1 + e^{-(B0 + B1X1i)}]^{-1}$

and

$1-Pi = 1 - F(B0 + B1X1i)$ where $1-F = [1 + e^{+(B0 + B1X1i)}]^{-1}$, we may make the appropriate substitution:

$$L = \prod_{i=1}^{N} F(B0 + B1X1i)^{Yi} \; [1 - F(B0 + B1X1i)^{(1-Yi)}] \; .$$

By taking the natural logs of both sides, the likelihood function reduces to:

$$\ln(L) = \sum_{i=1}^{N} Yi \ln[\, F(B0+B1X1i)\,] + \sum_{i=1}^{N} (1-Yi) \ln[\, 1-F(B0+B1X1i)\,] \; .$$

To find the Bj parameters that maximize ln(L) we take the partial derivative of ln(L) with respect to the B coefficients and set them equal to zero:

$$\frac{\partial \ln L}{\partial Bj} = \sum_{i=1}^{N} Yi \frac{\partial F / \partial Bj}{F} - \sum_{i=1}^{N} (1 - Yi) \frac{\partial F / \partial Bj}{(1 - F)} \overset{set}{=} 0$$

Since F is a nonlinear function, we must use a procedure that will solve the above equation for all Bj. The most popular procedure is an iterative technique known as the method of Newton or the Newton-Raphson method. In general, Newton's method will converge to the global maximum when used to solve the above problem (Amemiya, 1981 and Judge et al., 1982). In addition, these Bj estimates have a number of desirable statistical properties; these estimators are consistent, asymptotically normally distributed, and asymptotically efficient (Pindyck and Rubinfeld, 1981).

C. Evaluating Conditional Logit Models

Since ML coefficient estimates are consistent and asymptotically normal and efficient, familiar tests of significance may be applied to test the "goodness" of the estimated coefficients. To test the hypotheses

$$H0 : Bj = 0 \quad \text{versus} \quad HA : Bj \neq 0$$

a classical t-test may be used. The test statistic is:

$$t^* = \frac{\hat{B}j}{S.E.(\hat{B}j)}$$

where: t^* = the calculated t-statistic with N-k degrees of freedom

N = the number of observations in the sample

k = the number of Bj parameters in the model

$\hat{B}j$ = the estimated coefficient being tested

$S.E.(\hat{B}j)$ = the estimated standard error of Bj

This statistic asymptotically follows a t-distribution with N-k degrees of freedom. The H0 is rejected at an appropriate level of significance if the absolute value of the test statistic is greater than the tabular t value.

Often we wish to test the worth of the estimated model as a whole. That is, we want to see if the explanatory variables Xji impact the calculated probabilities Pi. Stated formally, the test hypotheses are:

$$H0 : B1 = B2 = \ldots = B(k-1) = 0$$

versus

HA : At least one coefficient Bj other than the constant B0 contributes to the explanation of Pi.

A chi-squared test is often used to test the null hypothesis (Judge et al., 1982). The test statistic is:

$$2\ln(Lmax/Lo) = 2[\ \ln(Lmax) - \ln(Lo)\ ]\ .$$

where: Lmax = the value of the likelihood function evaluated at the maximum likelihood estimates (i.e. evaluated under the hypothesis that the estimated coefficients are significantly different than zero).

Lo = the value of the likelihood function evaluated under the hypothesis that all the coefficients except the constant term B0 are equal to zero.

This statistic follows a chi-squared distribution with k-1 degrees of freedom where k equals the number of parameters in the model. The null hypothesis is rejected if the test statistic exceeds the tabular chi-squared figure at an acceptable level of significance.

The above tests of significance only tell us if the estimated coefficients of the model significantly add to the explanatory power of the model. Neither statistic however provides an indication of <u>how much</u> the estimated model is explaining. In standard linear regression the R-squared statistic gives us a normalized number (between zero and one) that relates the proportion of the variation in the dependent variable explained by the independent variables. Given our nonlinear logit model, several pseudo R-squared measures have been proposed. McFadden (1974) proposed the following R-squared statistic for QR models.

$$\text{McFadden's R-squared} = 1 - \frac{\ln(\text{Lmax})}{\ln(\text{Lo})}$$

This statistic will be zero when ln(Lmax) is no better than the log likelihood function in which all parameters are constrained to zero except for B0. McFadden's R-squared increases to one when ln(Lmax) approaches zero. This is a convenient method. Unlike the linear regression R-squared, McFadden's R-squared <u>does</u> <u>not</u> measure the amount of variation in the dependent variable as explained by the model. McFadden's R-squared is best used to compare the <u>relative</u> worth of competing logit specifications; it can not be taken as an absolute measure of the explanatory worth of the model (Amemiya, 1981 and Judge et al., 1980).

Another measure of pseudo R-squared is more analogous to that of a standard regression problem. This statistic was proposed by Effron (1978). Let:

The Sum of Squared Residual (SSR)

$$= \sum_{i=1}^{N} (Yi - \widehat{P}i)^2$$

Total Sum of Squares (TSS)

$$= \sum_{i=1}^{N} (Yi - \overline{Y})^2$$

$$\text{where: } \overline{Y} = \frac{1}{N} \sum_{i=1}^{N} Yi$$

$$\text{Effron's R-squared} = 1 - \frac{SSR}{TSS} .$$

Though Effron's R-squared corresponds to the R-squared in the standard regression model, we must be careful interpreting its meaning. The problem surrounding Effron's measure of R-squared revolves around the calculation of the sum of squared residuals. Recall that in the regular linear regression case, the sum of squared residuals is defined as:

$$SSR = \sum_{i=1}^{N} (Yi - \widehat{Y}i)^2 .$$

where: Yi = the observed dependent variable

$\widehat{Y}i$ = the predicted dependent variable (predicted on the basis of estimated model)

In standard linear regression, both Yi and $\widehat{Y}i$ measure the same phenomena and are assumed to be continuous variables. If the model fits the observed data well, the discrepancy between Yi and $\widehat{Y}i$ should be small for all i. Effron merely replaced $\widehat{P}i$ for $\widehat{Y}i$ in his estimation of R-squared for QR models. Intuitively, the discrepancy between Yi and $\widehat{P}i$ should be small for all i if the model adequately fits observed data. Unfortunately, Yi and $\widehat{P}i$ do not measure the same phenomena, nor is Yi continuous. Yi is a discreet categorical variable that indicates whether or not a particular event occurred. $\widehat{P}i$ is a continuous variable that indicates the <u>probability</u> of a particular event occurring. Though there is a close relationship between Yi and $\hat{P}i$ (high values of $\hat{P}i$ should be associated with Yi = 1) the difference between Yi and $\hat{P}i$ does not measure the residual variation not explained by the regression. A model with a dichotomous dependent variable is not likely to produce an R-squared close to 1. Morrison (1972) shows that if the true probabilities of an event occuring were distributed evenly over an interval of the independent variable, the upper bound of R-squared would be approximately 0.3333. If the predicted probabilities are distributed at the extremes (tails) of the independent variable, the meaning of Effron's R-squared for a logit model, approaches the meaning of R-squared as in the linear regression model. It is most likely however, that both McFadden's R-squared and Effron's R-squared are biased downward: they both <u>underestimate</u> true R-squared. Like McFadden's R-squared, Effron's R-squared may be used to measure the relative worth of competing QR model specifications; Effron's R-squared should not be interpreted as an absolute measure of the amount of variation in the dependent variable explained by the regression. Amemiya (1981) and Judge et al. (1982) seem to be in agreement with this conclusion.

Another statistic used to measure the "goodness of fit" (explanatory power) of the model to observed data is the percentage of correct predictions. If $\hat{P}i$ is greater than 0.50, the probability of an event E occuring <u>is greater than</u>

the probability of it not occuring. Similarly, if $\hat{P}i$ is less than 0.50, the probability of an event E occuring <u>is less than</u> the probability of it not occuring. Therefore let:

$$\hat{Y}i = \begin{cases} 1 \text{ when } \hat{P}i > 0.50 . \\ 0 \text{ when } \hat{P}i < 0.50 . \end{cases}$$

The test statistic is:

$$\text{Proportion of correct predictions} = [N - \sum_{i=1}^{N} (Yi - \hat{Y}i)]/N .$$

The problem with this statistic is that it weighs all prediction probabilities between 0.50 and 1 similarly. The same is true for probabilities between 0 and 0.49. If the predicted probability $\hat{P}i$ is 0.01, we are almost absolutely sure that event E will not occur; our predicted $\hat{Y}i$ will equal one. However, if $\hat{P}i$ is 0.49, we predict that event E will not occur (as we did when $\hat{P}i$ = 0.01) even though there is a good chance that the event E might occur. When there are many predicted probabilities in the vicinity of 0.50, this statistic may be misleading.

D. Interpretation of Logit "Slope" Coefficients

Once the logit model has been calculated and statistically verified, we must interpret the estimated parameters of the model. There is a tendency to directly interpret the estimated Bj "slope" parameters (i.e. the parameters other than the constant term B0) as elasticities and propensities similar to that of standard linear regression models. The Bj "slope" estimates in the logit model <u>can</u> <u>not</u> be interpreted as the increase in the probability of event E occuring given a unit increase in a independent variable Xji however. To illustrate this, recall:

$$Pi = [\ 1 + e^{-(\ B0\ +\ B1X1i\ )}\ ]^{-1}.$$

Multiply both sides of this equation by $[1 + e^{-(B0\ +\ B1X1i)}]$ to get:

$$[1 + e^{-(B0\ +\ B1X1i)}]\ Pi = 1\ .$$

Divide this by Pi and then subtract 1 from the left side of the equation.

$$e^{-(B0\ +\ B1X1i)} = \frac{1}{Pi} - 1$$

or

$$e^{-(B0\ +\ B1X1i)} = \frac{1 - Pi}{Pi}$$

Invert both sides of this equation to get:

$$e^{+(B0 + B1X1i)} = \frac{Pi}{1 - Pi} .$$

Take the natural log of both sides of the equation.

$$B0 + B1X1i = \ln \frac{Pi}{1 - Pi}$$

The B1 coefficient of this bivariate logit model can be interpreted as the increase in the <u>log of the odds</u> that event E will occur given a unit change in the independent variable X1i (Judge, 1982). To solve for the effect a unit change in X1i has on the probability Pi, we must do the following:

$$\Delta \ln \frac{Pi}{1 - Pi} = \Delta (B0 + B1X1i) .$$

Since $\Delta B0 = 0$ (because B0 is a constant),

$$\Delta \ln \frac{Pi}{1 - Pi} = B1 \Delta X1i .$$

Recall that since $\ln(x/y) = \ln(x) - \ln(y)$, we may write:

$$\Delta \ln \frac{Pi}{1 - Pi} = \Delta \ln(Pi) - \Delta \ln(1-Pi).$$

Also, $\Delta \ln(x) \doteq (\Delta x)/x$*. Thus,

$$\Delta \ln(Pi) - \Delta \ln(1-Pi) \doteq \frac{\Delta Pi}{Pi} - \frac{\Delta (1-Pi)}{1 - Pi}$$

$$= \left[\frac{1}{Pi} + \frac{1}{1-Pi} \right] \Delta Pi = \frac{1}{Pi(1-Pi)} \Delta Pi \quad .$$

Therefore,

$$\Delta \ln \frac{Pi}{1 - Pi} \doteq \frac{1}{Pi(1-Pi)} \Delta Pi \quad .$$

Through substitution, we get:

$$\frac{1}{Pi(1-Pi)} \Delta Pi \doteq B1 \Delta X1i \ .$$

Multiply both sides of this equation by [Pi (1 - Pi)]. Also, since we want to find what a unit change in X1i does to the probability Pi, let $\Delta X1i = 1$. Therefore,

$$\Delta Pi \doteq B1 \left[Pi \ (1 - Pi) \right]$$

A change in Pi, as a result of a unit change in X1i, is a function of both B1 and Pi. The change in the probability of an event E occuring due to a change in a continuous explanatory variable X1i depends upon both the value of B1 and the multiplier value [Pi (1 - Pi)] which is a function

---

* This approximation is appropriate for any continuous variable X. If X is a discreet variable (i.e. a dummy variable), this equation is no longer valid.

of Pi.

Table 3.2 shows the effect of different initial Pi's on the multiplier [Pi (1 - Pi)].

Table 3.2 The Effect of P_i on the multiplier $P_i\ (1 - P_i)$

P_i	$P_i(1 - P_i)$
0.00	0.00
0.10	0.09
0.20	0.16
0.30	0.21
0.40	0.24
0.50	0.25 √
0.60	0.24
0.70	0.21
0.80	0.16
0.90	0.09
1.00	0.00

Since 0 < [Pi (1 - Pi)] < 0.25, a unit change in Xli impacts Pi by at most one quarter the value of Bi. When Pi approaches 0 or 1, a unit change in Xli impacts Pi by an extremely small fraction of the value of Bl. This makes intuitive sense. If the initial value of the choice probability Pi were very high or very low, we expect that individual i made a very definitive choice; it would take a great deal of change in the continuous explanatory variable Xli for individual i to change its mind and choose other alternatives. If however, choice probability Pi were close to 0.50 initially, we expect that the individual i made a "weak" choice between the two alternatives; i.e., the individual i is rather indifferent between the two alternatives because the incentive to choose one alternative versus another alternative was not very strong to begin with. A relatively small change in the continuous explanatory variable could sway this individual to radically change its choosing behavior.

In conclusion, a unit change in a continuous explanatory variable Xli changes the probability that a power plant will use Powder River coal by a factor of:

$$Bj[\ Pi\ (\ 1 - Pi\ )\ ]\ .$$

This interpretation is not appropriate for estimated coefficients on discreet dummy variables; coefficients on dummy variables represent "constant" parameters rather than "slope" parameters. Of course, as Bj becomes larger for a fixed Pi, the impact a change in Xli on a change in Pi becomes larger. Since Pi varies for every individual i however, the impact of a change in Xli on a change in Pi must be calculated for each individual.

E. Explanatory variables for the Spatial Powder River Coal Choice Logit Model

As indicated above, our qualitative response model will be based upon the logistic cumulative probability distribution function. This logit model has the form:

$$Pi = \text{Prob(Plant i buys Powder River coal)} = [1 + e^{-(B0+B1X1i+B2X2i+\ldots)}]^{-1}$$

We know how the model works, how the Bj parameters are estimated, how the model is statistically evaluated, and how to interpret the coefficients. We must now select appropriate explanatory variables Xji for our model.

The <u>theory</u> of spatial coal markets tells us that the relative total costs of generating electricity from Powder River coal versus another coal affects coal choice, where total generating costs are a function of mine mouth prices, regional coal transportation rates, power plant costs, and pollution control costs. Using this idea, the principle independent (explanatory) variable will be the total cost differential between competing coal supply regions. Let:

COST DIFFERENTIALi = TOTAL COSTij - TOTAL COSTi(Powder River)

where:

TOTAL COSTij = the total cost of electric generation incurred by the ith power plant if burning coal from the least cost non-Powder River coal supply center j.

TOTAL COSTi(Powder River) = the total cost of electric generation incurred by the ith power plant if burning coal from the Northern Great Plains.

If COST DIFFERENTIALi is less than zero, we would expect the probability of power plant i to buy Powder River coal to be low since the cost of using Powder River coal exceeds the cost of using non-Powder River coal. Similarly, if the COST DIFFERENTIALi is greater than zero, we would expect the probability for power plant to buy Powder River coal to be relatively high. As the COST DIFFERENTIALi approaches zero from either above or below, we would expect the probability of a utility company buying Powder River coal to approach 0.50. Since the total cost of electric generation by burning coal is a function of plant location relative to coal supply regions, the described QR formulation will represent a spatial market model. If we assume that the model is linear in the parameters and that Ci represents the cost differential, the logit model that calculates the probability power plant i buys Powder River coal has the following form:

$$Pi = [1 + e^{-(B0 + B1Ci)}]^{-1}.$$

Previous theoretical spatial coal market studies (Watson, 1972 and Power et al., 1976 and Duffield et al., 1982) have asserted that air pollution policy and the use of flue gas desulfurization (FGD) affects coal choice. Regulations on sulfur dioxide emissions obviously affect coal choice since coals from different supply centers have different sulfur content. If power plants were required to emit less sulfur dioxide than a certain ceiling level, but were permitted to use any method to meet the standard, a correlation (independent of costs) between power plants not using FGD and the use of low sulfur Powder River coal should be detected; a correlation between power plants using FGD and the use of high sulfur non-Powder River coal should be found as well. If all power plants were required to use FGD to reduce sulfur emissions, the correlation (independent of costs) between FGD and coal choice would disappear: FGD is no longer a variable. Also, air pollution policy and the use (or non-use) of FGD obviously affect total electric generating costs. Because of this, the affect of the cost differential on coal choice might vary as air pollution policy and FGD use varies; air pollution policy and the use of FGD "interacts" with the cost differential. For example, if power plants have the option to meet air pollution emission ceilings by burning low sulfur coal instead of using FGD and burning high sulfur coal, the low sulfur coal option becomes more attractive because of the high cost of FGD. In addition, FGD costs vary proportionally with the sulfur content the of coal; for any given power plant size (in megawatts) and any given level of sulfur emissions, it is less expensive to scrub low sulfur coal relative to high sulfur coal. Given these assertions, variables describing the impact of air pollution laws and FGD (and their interactions with the cost differential) on the probability of a power plant buying Powder River coal should be incorporated into the logit model.

Given the power plant sample collected for this study (discussed in Chapter 4), there are two sets of federal air pollution standards that concern us. These include the 1971 New Source Performance Standards and the 1978 Revised New Source Performance Standards. New coal fired generators larger than 73 megawatts of capacity that commenced construction after August 17, 1971 were required to emit less than 1.2 pounds of sulfur dioxide per million BTU's. The utility companies could meet this standard in a variety of ways including tall smoke stacks to FGD systems. These air quality regulations were made more stringent in 1978. New coal fired generators larger than 73 megawatts that commenced construction after September 18, 1978 must meet one of three alternative standards on sulfur dioxide emissions:

a) The plant must not emit more than 1.2 pounds of sulfur dioxide per million BTU and must acheive a 90 percent reduction of sulfur emissions.

b) If emissions are less than 0.6 pounds of sulfur dioxide per million BTU, a 70 percent reduction of sulfur emissions is required.

c) If the coal is solvent cleaned prior to burning, the standard is 1.2 pounds sulfur dioxide per million BTU ceiling emissions with 85 percent reduction of potential sulfur emissions.

The percent reduction of sulfur emissions practically requires utility companies to use a FGD system. For convenience, I will refer to the 1971 New Source Performance Standards as the NSPS, and the 1978 Revised New Source Performance Standards as the RNSPS.

The logit model will take into account the affect of FGD and air pollution control laws on the probability for a plant to use Powder River coal by utilizing the method of dummy variables. Let:

$$Ai = \begin{cases} 1 \text{ if the plant i falls under the 1978 RNSPS .} \\ 0 \text{ if the plant i falls under the 1971 NSPS.} \end{cases}$$

$$Fi^{*} = \begin{cases} 1 \text{ if FGD is used on plant i .} \\ 0 \text{ if FGD is not used on plant i .} \end{cases}$$

The logit model now has the form:

$$Pi = [\, 1 + e^{-(B0 + B1Ci + B2Fi + B3Ai)}]^{-1}.$$

Air pollution control policy faced by plant i and whether or not FGD is used by plant i affects total costs of electric generation. The dummy variables Ai and Fi may "interact" with the cost differential. The following interaction terms, therefore, must be included as variables in the logit model:

---

* Fi (the FGD dummy) is a variable only in the context of 1971 NSPS power plants. Power facilities under 1971 NSPS regulations have the option to use or not to use FGD in order to meet emission standards. A case can be made for excluding the "independent" dummy variable Fi from the above analysis. Because of sulfur emission ceiling regulations, a utility company considering the use of any particular coal must simultaneously determine whether or not they will use FGD depending upon the sulfur content of that particular coal. Since coal choice determines the value of Fi rather than Fi determining coal choice, we become hesitant to continue using Fi as an "independent" variable. Appendix D contains the results of a logit model estimation where Fi is excluded from the analysis. Since the distinction between FGD users and non-FGD users might improve the statistical "goodness of fit" of the model to observed data, we will continue to include Fi in the remainder of this analysis.

CiAi = cost differential times the dummy variable on air pollution control policy.

CiFi = cost differential times the dummy variable on FGD use.

The final logit model used in this analysis has the form:

$$Pi = [\ 1 + e^{-(B0\ +\ B1Ci\ +\ B2Fi\ +\ B3Ai\ +\ B4CiFi\ +\ B5CiAi)}\ ]^{-1}.$$

There are obviously other independent variables that affect coal choice. For example, a utility company may wish to secure their supply of coal from unpredictable interruptions (e.g. local mine strikes, regional supply depletion, changes in taxing policy such as severance taxes, the ability of coal mining operations to quickly expand production to meet regional coal demand). Due to these supply problems, power plant managers will base their coal choice on factors other than cost. Also, state level air pollution control laws have not been taken into account in this study. In essence, there are many factors that are not included in the modelling effort. Since we are interested in a simple model that utilizes only the "key swing variables" affecting coal choice, we leave these other factors out of the model. It is assumed that these other factors are relatively insignificant in explaining electric utility coal choice behavior.

III. References

Amemiya, T. "Qualitative Response Models: A Survey," Journal of Economic Literature. (December 1981, pp. 1483-1536).

Campbell, T. C. and Hwang, M. "Spatial Analysis and Identification of Coal Markets," The Journal of Energy and Development. (Autumn 1978, Vol. IV, No. 1, pp. 104-125).

Chambers, E. A. and Cox, D. R. "Discrimination Between Alternative Binary Response Models," Biometrika. (1967, Vol. 54(3-4), pp. 573-578).

Duffield, et al. Projections of Coal Demand From the Northern Great Plains Through the Year 2010. (Final Report, OSM, May 1982).

Effron, B. "Regression and ANOVA with Zero-One Data: Measures of Residual Variation," J. Amer. Statist. Assoc.. (March 1978, pp. 113-121).

Hartman, R. S. A Generalized Logit Formulation of Individual Choice. (MIT Energy Laboratory Working Paper No. MIT-EL-79-010WP, February 1979).

Judge, G. et al. The Theory and Practice of Econometrics. New York: John Wiley and Sons, 1980.

Judge, G. et al. Introduction to the Theory and Practice of Econometrics. New York: John Wiley and Sons, 1982.

Koutsoyiannis, A. Theory of Econometrics. Hong Kong: The MacMillan Press LTD, 1979.

McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behavior," in Frontiers in Econometrics. Edited by P. Zarembka. New York: Academic Press, 1974, pp.105-142.

McFadden, D. Structural Analysis of Discrete Data with Econometric Applications. Edited by C. F. Manski and D. McFadden. Cambridge, Mass.: MIT Press, 1981.

Morrison, D. G. "Upper Bounds for Correlations Between Binary Outcomes and Probabilistic Predictions," J. Amer. Statist. Assoc., March 1972, pp. 68-70.

Pindyck, R. and Rubinfeld, D. _Econometric Models and Economic Forecasts_. New York: McGraw Hill, 1981.

Power, T. et al. _Montana University Coal Demand Study: Projections of Northern Great Plains Coal Mining and Energy Conversion Development_. (Final Report, NSF-RANN, May 1976).

Rubinfeld, D. "Voting in a Local School Election: A Micro Analysis," _Review of Economics and Statistics_. (February 1977, Vol. 59, No. 1, pp. 30-42).

Watson, W. D., Jr. "Cost of Air Pollution Control in the Coal-fired Electric Power Industry," _Quarterly Review of Economics and Business_. (Autumn 1972, Vol. 12, No. 3, pp. 63-85).

CHAPTER FOUR

PRELIMINARY ANALYSIS OF THE POWDER RIVER SPATIAL COAL MARKET: DATA BASE, COST DIFFERENTIAL CALCULATIONS, AND SIMPLE STATISTICAL TESTS.

I. Introduction

Now that we have a theoretically sound and empirically oriented model, we must gather a substantial sample of observations on individual coal fired power plants for each previously listed variable (see Chapter Three). We could then estimate the model and empirically test it for statistical significance. I will first discuss how certain data was compiled for the spatial analysis.

II. The Location of Power Plants and Their Contracted Coal Sources

There are 438 additional new coal fired power plants coming on line between year 1976 and year 2000. A list of these individual plants was compiled by Green (1982). This list included information on:

1) plant and utility company name

2) plant location by state, county, city

3) the proposed on line date (the date when electric generation is to commence)

4) the nameplate megawatt capacity

5) coal sources by state and county

Using a map of the United States, the location of plants and coal sources were converted into X-Y coordinates (two dimensional Euclidean space). The origin was placed at Gillette Wyoming, the Powder River supply center. These coordinates are used to calculate distance, in miles, between power plants and coal supply centers.

Information on boiler order dates and whether or not these individual power plants utilize FGD was gathered from the following sources:

1) Green (1982)
2) Kidder, Peabody and Company (1981)
3) Komanoff (1981)
4) U.S. EPA (1983)

In addition, the United States Environmental Protection Agency (EPA) provided information on which set of air quality regulations (NSPS or RNSPS) individual coal fired generators face (EPA is empowered to enforce Federal air quality regulations).

For each individual electric generating unit in Green's (1982) sample, the above information was assembled. These data were later compiled into a computer data file. Table 4.1 summarizes the content and format of this data file. Summary statistics on the above data are contained in Table 4.2.

III. Cost Differentials: Cost Calculation Methodology

For a given power plant location, the cost differential is a number comparing the relative advantage of using Powder River coal versus the best non-Powder River coal

Table 4.1 Summary of the Plant/Mine Location Data File

Variable Name	Description	Columns (position)
1) Plant ID number	### (sequence number)	1 - 3
2) Plant location	##### state county (FIPS code)	4 - 8
3) Plant coordinates	##.####.## X Y	9 - 18
4) MW capacity	#### (nameplate capacity)	20 - 23
5) On line date	## year	24 - 25
6) Boiler order date	#### year month	26 - 29
7 - 10) Mine location*	##### state county (FIPS code)	31 - 50
14 - 17) Mine coordinates*	##.####.## X Y	52 - 91
18) FGD (Yes, No)	# (1 = Yes, 0 = No)	93
19) NSPS (Yes, No)		94
20) RNSPS (Yes, No)		95

= an integer value (0, 1, 2, 3,...., 9)

* = The first mine listed indicates the main coal source.

Table 4.2 Summary Statistics on the Power Plant Sample

1) 438 coal-fired generators, known to come on line between 1976 and 2000, are in this sample. Due to missing values on either plant location or mine location, the number of individual power plants used in the logit regression analysis equals 411.

2) These coal-fired generators are scattered across the continental United States (See Figure 4.1). A majority of the power plants are located in the Midwest. Texas has the most plants of any single state in this data sample(50 new coal fired generators).

3) The mean nameplate capacity of these electric generators is 515 megawatts (mode = 500, median = 508). The smallest plant in the sample is 20 megawatts; the largest plant is 1300 megawatts. The standard deviation is 208 megawatts.

4) THe mean on-line year for power plants in the sample is 1985 (median also equals 1985). The mode year is 1991. The standard deviation is 6.140 years.

5) The mean boiler order year** in the sample is 1975 (the median year is 1975 and the mode is 1978). These order years ranged from 1969 to 1981. The standard deviation is 3.115 years.

6) In the working sample of 411 power plants, 129 of them buy coal from the Powder River Basin as their main source of coal.*** Most of this coal comes out of Gillette, Wyoming.

7) 302 power plants in the sample use some form of FGD. 136 do not use FGD on their smoke stacks.

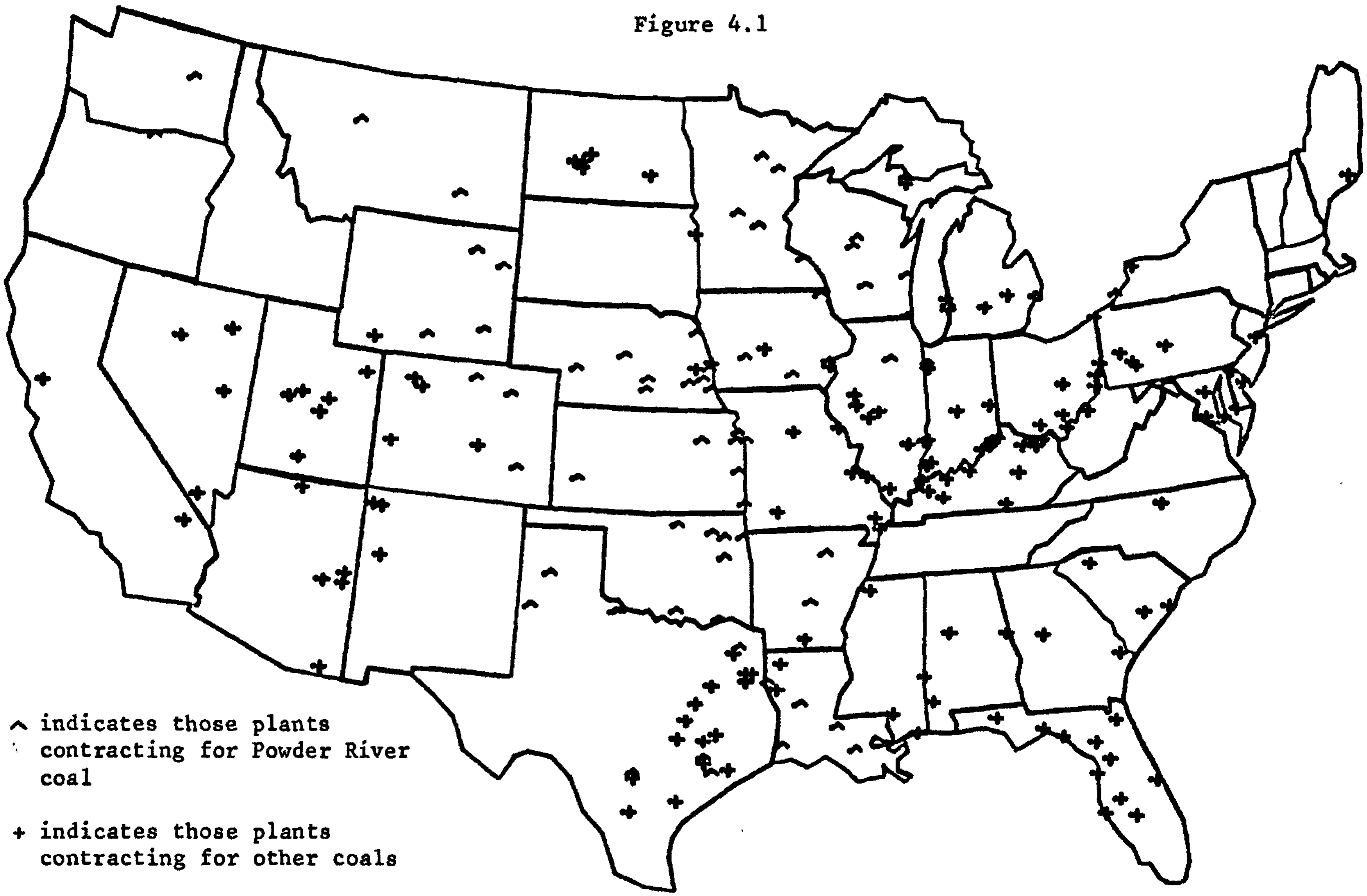
8) 206 power plants are under the 1971 NSPS air quality regulations. 232 power plants are under the 1978 RNSPS air quality regulations.

* All statistics are based upon the entire sample of 438 power plants unless otherwise noted.

** 265 power plants did not have a boiler order date entry.

*** 145 power plants buy coal from the Northern Great Plains. 13 of these plants are North Dakota lignite users. 3 of the 145 power plants use Powder River coal as a secondary coal source.

Figure 4.1



Coal-fired Electric Generators Coming On Line Between 1976 and 2000 alternative. For a given level of electricity production and given a power plant site, the cost differential is defined as:

COST DIFFERENTIALi = TOTAL COSTij - TOTAL COSTi(Powder River)

where:

TOTAL COSTij = the total cost of electric generation incurred by the ith power plant if burning coal from the least cost non-Powder River coal supply center j.

TOTAL COSTi(Powder River) = the total cost of electric generation incurred by the ith power plant if burning Powder River coal.

If coal source j is the least cost non-Powder River coal source, it is cost effective to buy from source j if the cost differential is less than zero. If the cost differential is greater than zero, it is cost effective to buy coal from Powder River. Since the cost differential is a function of comparing total costs of electric generation, we must know how these total costs are calculated. The following describes the procedure of how the total costs of electric generation resulting from the use of a particular coal were constructed. These equations are similar to those used by Duffield et. al. (1982).

1) Equation 4.1 computes the power plant costs. This calculation is specific to a coal fired generating facility i using coal from source j. Equation 4.2 computes the annualized costs, both capital and operating, for a powerplant and includes both the base plant costs and, if the plant uses FGD, sulfur dioxide control costs.

$$PCOST_{ij} = (KCOST_{ij} \times MW_i \times RATE_{ij} \times 1000) + (OP_{ij} \times MW_i \times T_i \times 1000) \quad (4.1)$$

i = 1, 2, 3,... = the ith coal fired power plant unit.

j = 1, 2, 3,... = the jth coal supply region.

PCOSTij = the annualized cost of building and operating power plant i using coal from region j. Boilers must be designed to burn a specific quality coal. Generally, power plant costs increase as coals of lower quality (e.g. lower in BTUs per weight and/or higher in sulfur content) are used.

KCOSTij = (BPCAPij + ADDCAPij +SO2CAPij) x (1 + CAPENij) = capital costs of power plant i using coal from source j in dollars per kilowatt.

BPCAPij = base plant capital costs including 1971 NSPS particulate control equipment.

ADDCAPij = additional capital costs for 1978 RNSPS power plants. These additional costs result from additional particulate control devices. 1971 NSPS plants do not face these additional costs.

SO2CAPij = sulfur dioxide pollution control capital costs. If FGD is not used by plant i, this value is zero.

CAPENij = the capacity penality (in percent) faced by the power plant if FGD is used. Since FGD reduces the electric output of the generator, the plant must increase capital expenditures to maintain its stated MWi net output.

MWi = the nameplate (net) capacity of power plant unit i in megawatts.

RATEij = real rate of annualization of KCOSTij.

OPij = (BPOMij + SO2OMij) = the operating and maintenance costs of power plant i using coal from source j in mills per kilowatt hour.

BPOMij = the base plant operating and maintenance costs. This value includes particulate control operating costs.

SO2OMij = FGD operating and maintenance costs. This value is zero if FGD is not utilized by the plant.

2) The quantity of coal required for the annual operation of a "specific" coal-fired electric generating plant is computed.

$$TONSij = (MWi \times Ti \times HRij \times 1000)/(HCj \times 2000) \quad (4.2)$$

TONSij = the quantity of coal, in tons, from source j required for the annual operation of a coal fired electric generator i of size MWi.

MWi = the nameplate (net) generating capacity of power plant unit i in megawatts.

Ti = the equivalent number of hours per year that power plant unit i operates at full capacity.

HRij = BPHRij x (1 + ENPENij) = the heat rate of a power plant i using coal from source j, in BTU/Kilowatt Hour. The heat rate describes the amount of energy (measured in BTUs) needed to produce a kilowatt hour of electricity.

BPHRij = the base plant heat rate. This value does not include the effect of sulfur emission control.

EPij = the energy penalty (in percent) faced by power plant i using coal from source j if the plant uses FGD. Energy is required to run scrubbers therefore increasing the plant's heat rate. If FGD is not being used, EPij = 0.

HCj = the heat content of coal from source j in BTU/pound.

3) The next equation calculates the distance, as the crow flies, between coal supply center j and power plant i.

$$DISTij = \sqrt{(Xj - Xi)^2 + (Yj - Yi)^2} \quad (4.3)$$

Xj = X-coordinate of the coal supply center j.

Xi = X-coordinate of the power plant i.

Yj = Y-coordinate of the coal supply center j.

Yi = Y-coordinate of the power plant i.

4) With tonnage estimates from equation (4.2) and distance estimates from equation (4.3), annual fuel costs and fixed and variable transportation costs are computed by means of equation (4.4).

$$FCOSTij = [CPj + FTCij + (VTCij \times DISTij)] \times TONSij \quad (4.4)$$

FCOSTij = the fuel costs; the sum of annual fuel costs plus the cost of transporting that fuel (both fixed and variable) to power plant i.

CPj = average coal prices from regional source j in dollars per ton.

FTCij = fixed transportation costs in dollars per ton.

VTCij = variable transportation costs in dollars per ton per DISTij.

5) Equation 4.5 computes the total cost, for plant i, of burning coal from source j.

$$Totcostij = (FCOSTij + PCOSTij)$$
or
$$Totcostij = Aij + (Bij \times Distij) \quad (4.5)$$

$$Aij = (CPj + FTCij) \times TONSij + PCOSTij$$

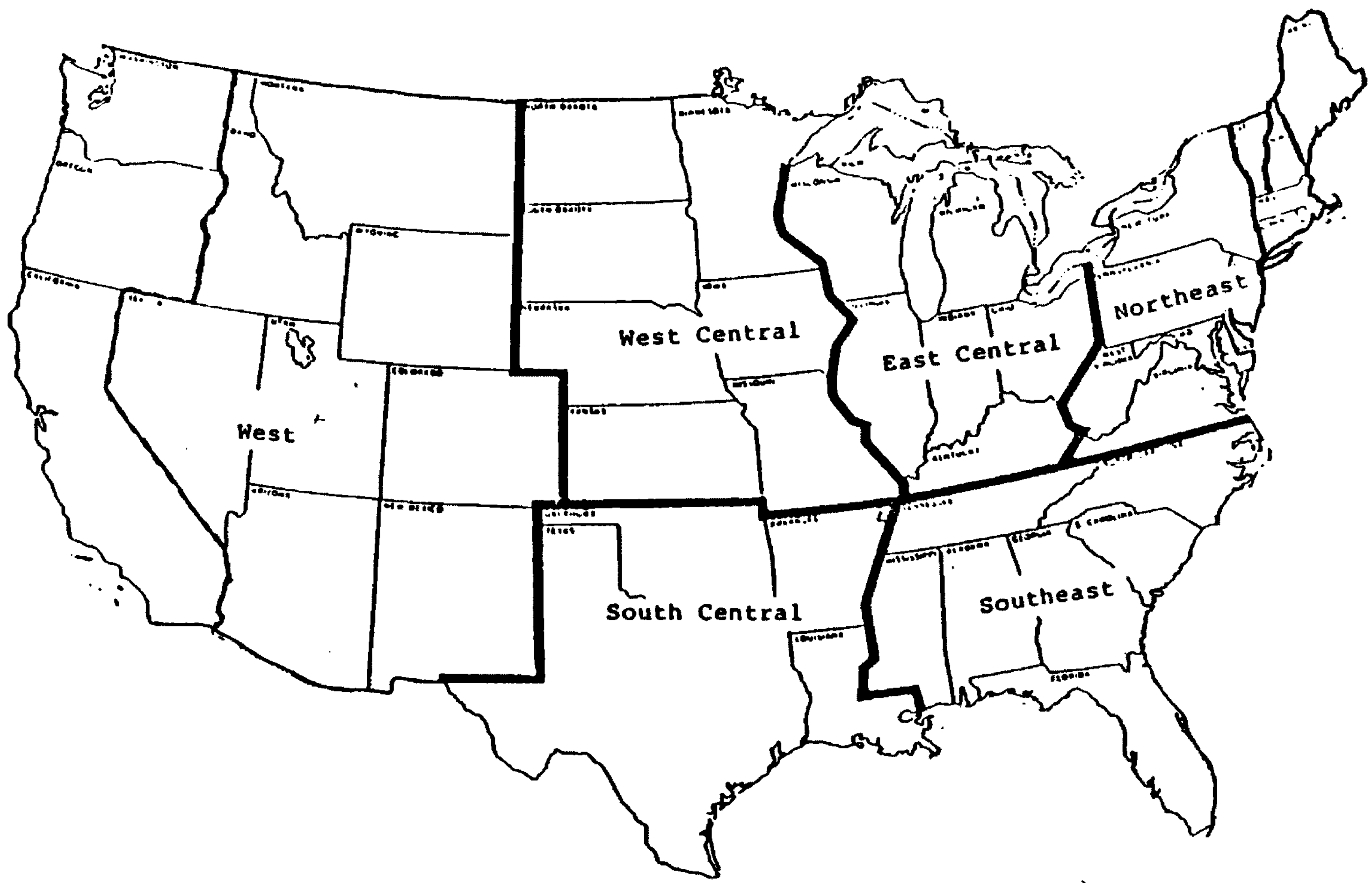
$$Bij = VTCij \times TONSij$$

For every electric power plant, thirteen separate calculations are made; each calculation describes the total cost for power plant i using coal from one of thirteen coal field supply centers in the United States including Powder River coal. By subtracting the total cost of using Powder River coal from the total cost of using the lowest cost non-Powder River coal, a cost differential is calculated for that particular power plant.

It would seem necessary to collect values for all the above mentioned variables for each individual power plant unit observation. There are two major reasons why we <u>should not</u> collect variable values for each individual power plant. The first major reason is that gathering information on each of the above mentioned variables for over 400 power plants by thirteen coal supply centers would be very costly on the analysis. Not only would the time needed to compile such a data bank be tremendous, but most of the data is not available. Information on contracted coal prices, transportation rates, and rates of return on capital are confidential within the electric utility industry. Also, data on future power plants that are presently in the planning stage does not necessarily exist. In addition, information on the coal alternatives actually considered by each utility company prior to their final coal choice is not available. For these reasons, the following generalizations are made in this study:

1) Power plant heat rates, base plant capital costs, and base plant operating/maintenance costs are functions of power plant location and coal rank. Of these variables, regional averages will be used instead of values gathered for each power plant observation. Power plant costs vary between regions in the U.S. due to regional differences in labor costs, material costs, climate, etc. Costs vary with coal rank because using low quality coal requires a more expensive boiler design. ICF Inc. (1980) calculated eight regional averages on these variables by coal rank. Figure 4.2 shows the breakdown of these regions (EPRI, 1979). Depending on the coal rank burned, a plant takes on base plant characteristics of the average plant within the region it is located. New Mexico and Arizona base plant

Figure 4.2. DATA REGIONS FOR CAPITAL COST ESTIMATES



Border between EPRI Defined Regions

Source: ICF Inc., (1980)
p. B-17

characteristics are averages of the South Central and Western regions. RNSPS plants face additional capital costs due to additional particulate control costs. Table 4.3 summarizes the base plant characteristics used in this analysis. These figures include the cost of 1971 NSPS particulate control.

2) Sulfur dioxide emission control costs (the cost of FGD) are a function of the sulfur content of the coal. In general, western states face higher pollution control costs because of Federal prevention of significant deterioration (PSD) laws. Using EPRI's regional breakdown (see Figure 4.2), average regional FGD capital costs, operating/maintenance cost, energy penalities, and capacity penalties by sulfur content of coal are presented in Table 4.4. NSPS plants and RNSPS plants utilizing FGD must include sulfur dioxide control costs in base plant costs. However, since NSPS plants are not required to use FGD in order to meet NSPS air quality standards, a problem arises in estimating the cost of sulfur emission control for this particular class of power plants. If a NSPS plant does not use FGD and contracts for Powder River coal, would the plant been required to use FGD with alternative coals? Similarly, if a NSPS plant uses FGD and contracts for non-Powder River coal, would the plant use FGD with alternative coals? The answer to this question depends upon the sulfur content of the contracted coal and the sulfur content of the alternative coal. Under the NSPS, new boilers typically do

Table 4.3 Base Power Plant Characteristics[a]

		Power Plant Region						
		North East	South East	East Central	West Central	South Central	NM AZ[b]	West
Base plant heat rates (BTU/kilowatt hour)	Bit	9592.0	9643.0	9693.0	9967.0	9920.0	9846.0	9772.0
	Sub	9863.0	9915.0	9967.0	10049.0	10200.0	10124.0	10049.0
	Lig	10669.0	10669.0	10669.0	10694.0	10644.0	10669.0	10669.0
Base plant capital costs (1980$/kilowatt)	Bit	757.0	622.0	717.0	689.0	628.0	675.0	721.0
	Sub	835.0	684.0	791.0	757.0	689.0	740.0	791.0
	Lig	831.0	831.0	831.0	830.0	831.0	831.0	831.0
Additional capital costs for RNSPS plants (1980$/kilowatt)	Bit	22.39	22.39	22.39	22.39	22.39	22.39	22.39
	Sub	23.69	23.69	23.69	23.69	23.69	23.69	23.69
	Lig	14.26	14.26	14.26	14.26	14.26	14.26	14.26
Operating and maintenance costs (1980 mills/kilowatt)	Bit	1.93	1.73	2.15	2.01	1.99	2.03	2.06
	Sub	2.00	1.77	2.03	2.03	2.01	2.05	2.08
	Lig	2.06	2.06	2.06	2.08	2.03	2.06	2.06

a. Source: ICF Inc. (1980), Duffield et al, (1982)

b. Average of South Central and Western Region data (Duffield et al, (1982)

Table 4.4 Sulfur Dioxide Emission Control Costs[a]

Power Plant Region	Item	Low Sulfur Coal (less than or equal to 0.83% sulfur)	Medium Sulfur Coal (greater than 0.83% butless than 2.5% sulfur)	High Sulfur Coal (greater than or equal to 2.5% sulfur)
Northeast Southeast East Central West Central[b] South Central	Capital cost (1980 $/kilowatt)	46.9	121.7	136.4
	Operating and Maintenance cost (1980 mills/kilowatt)	2.06	2.49	3.23
	Energy Penalty (%)	0.50	3.75	3.80
	Capacity Penalty (%)	0.50	2.60	2.65
Western[c]	Capital cost (1980 $/kilowatt)	131.8	134.8	136.4
	Operating and Maintenance costs (1980 mills/kilowatt)	2.26	2.93	3.23
	Energy Penalty (%)	4.35	4.35	4.35
	Capacity Penalty (%)	2.75	2.75	2.75

a. Source: ICF Inc. (1980), Duffield et al, (1982)

b. Does not include the states of North Dakota and South Dakota

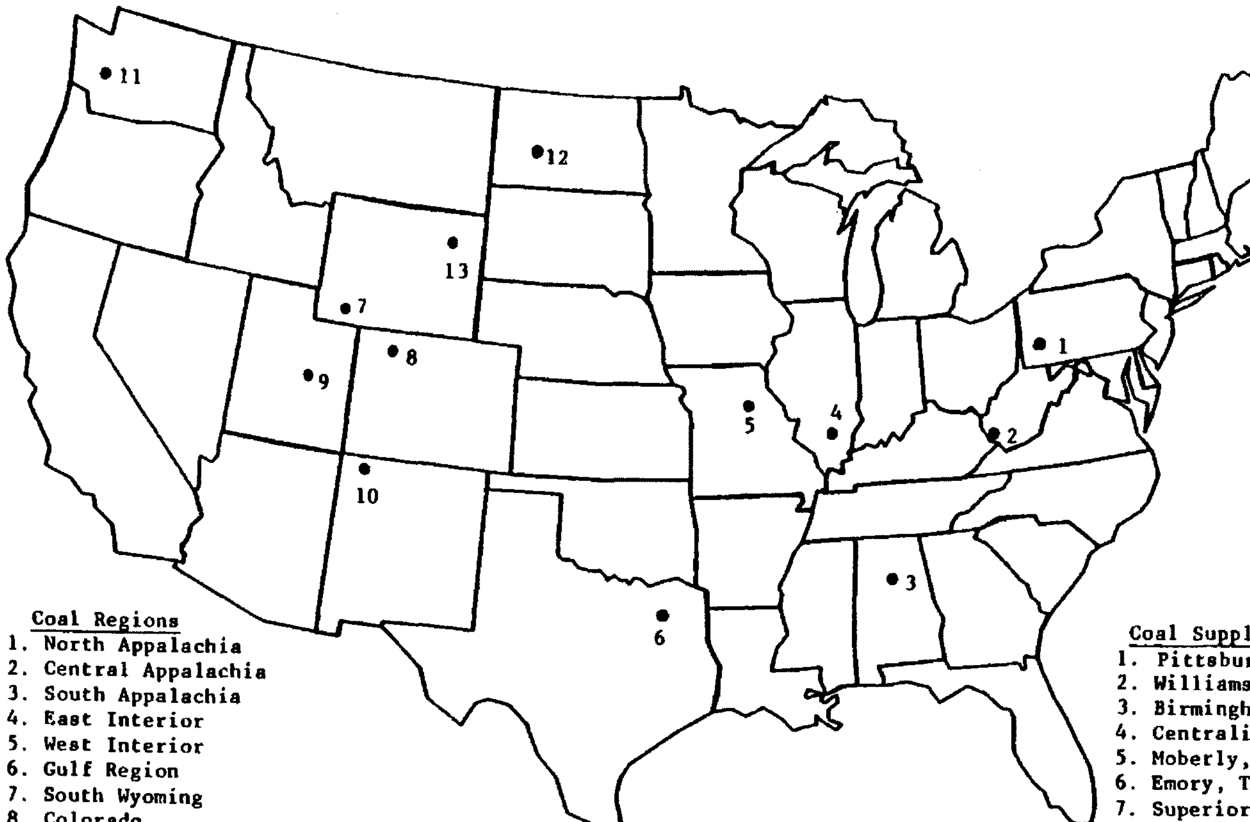
c. Western States (including North and South Dakota) face stricter air pollution regulations because of Federal PSD (Prevention of Significan Deterioration) requirements.

not have to use FGD if the coal sulfur content is less than 0.83% (Duffield et al., 1982). In this analysis, it is assumed that if NSPS plant does not use FGD but uses Powder River coal, the plant would be required to use a FGD system if the alternative coal sulfur content is greater than 0.83%. Also, if the NSPS plant does not use FGD and contracts for non-Powder River coal, it is assumed that the plant would not be required to use FGD if the alternative coal sulfur content is under 0.83%. All RNSPS plants are required to use a FGD system.

3) Coal regions offering significant future competition to Powder River are North Appalachia, Central Appalachia, South Appalachia, East Interior, West Interior, Gulf region, South Wyoming, Colorado, Utah, Southwest, Washington, and Fort Union. These coal regions can effectively be identified by a single point. Figure 4.3 identifies these major coal producing regions and their supply centers (ICF Inc., 1980). It is assumed that coal blending (power plants mixing coals purchased from different supply regions) does not occur.

4) Coal mined from the above coal regions will be characterized by three factors: average coal rank, average BTU content, and average percent sulfur. Regional mine mouth coal prices are estimated on average current long term contract steam coal prices. Table 4.5 lists coal characteristics by supply region.

Figure 4.3 Major Coal Producing Regions and Their Supply Centers



Coal Regions
1. North Appalachia
2. Central Appalachia
3. South Appalachia
4. East Interior
5. West Interior
6. Gulf Region
7. South Wyoming
8. Colorado
9. Utah
10. New Mexico
11. Washington
12. Fort Union
13. Powder River

Coal Supply Center
1. Pittsburgh, PA
2. Williamson, WV
3. Birmingham, AL
4. Centralia, IL
5. Moberly, MO
6. Emory, TX
7. Superior, WY
8. Hayden, CO
9. Huntington, UT
10. Farmington, NM
11. Centralia, WA
12. Beulah, ND
13. Gillette, WY

Source: ICF Inc. (1980), Duffield et al. (1982).

Table 4.5 Coal Quality and Price by Region[a]

	Coal Supply Regions[b]					
	North Appalachia	Central Appalachia	South Appalachia	East Interior	West Interior	Gulf Region
Mine type 1 = underground, 0 = surface	1.0	1.0	0.0	1.0	0.0	0.0
F.O.B. minemouth coal price ($/ton)	23.75	26.00	26.00	21.00	17.00	15.00
Coal Rank	Bit	Bit	Bit	Bit	Bit	Lig
Heat content (BTU/lbs)	12075.0	12200.0	12000.0	10500.0	9500.0	6300.0
Percent Sulfur	2.9	1.6	1.6	3.5	4.0	0.7
Percent ash	14.0	13.0	13.1	13.0	15.0	11.8
Distance of the supply center from Gillette,WY	1336.0	1306.0	1255.0	947.0	760.0	961.0
Supply center X coordinate (miles)	1316.0	1228.0	997.0	847.0	665.0	482.0
Supply center Y coordinate miles	-232.0	-443.0	-762.0	-423.0	-368.0	-832.0
General coal supply Region[d]	1.0	1.0	1.0	2.0	3.0	3.0

Table 4.5 Coal Quality and Price by Region (cont.)

Items	Coal Supply Regions[b] SWY	CO	UT	NM	WA	Fort Union	Powder River
Mine type 1 = underground, 0 = surface	1.0	1.0	1.0	0.0	1.0	0.0	0.0
F.O.B. minemouth coal price ($/ton)	16.50	17.50	19.75	16.00	27.50	7.25	8.80
Coal Rank	Sub-Bit[c]	Sub-Bit[c]	Bit	Sub	Sub	Lig	Sub
Heat content (BTU/lbs)	10500.0	10700.0	11500.0	10000.0	8100.0	6600.0	8660.0
Percent Sulfur	0.6	0.5	0.6	0.5	0.9	0.6	0.5
Percent ash	8.5	9.1	9.0	10.5	16.0	9.8	6.0
Distance of the supply center from Gillette, WY	250.0	278.0	453.0	552.0	875.0	276.0	0.0
Supply center X coordinate (miles)	-197.0	-102.0	-333.0	-208.0	-810.0	202.0	0.0
Supply center Y coordinate (miles)	-154.0	-258.0	-307.0	-511.0	331.0	188.0	0.0
General coal supply Region[d]	4.0	4.0	4.0	4.0	4.0	3.0	3.0

a. Source: Coal Week (1980-1981), Keystone (1980), Duffield et al, (1982)
b. Source: ICF Ind. (1980)

Table 4.5 Coal Quality and Price by Region (cont.)

c. Both Subbituminous and Bituminous coals are found in these regions

d. 1.0 = Appalachian Coal fields
   2.0 = Coal field between the Appalachian Mountains and the Mississippi River
   3.0 = Coal fields found between the Mississippi River and the Continental Divide
   4.0 = Coal fields west of the Continental Divide

5) All coal is transported by unit train. Straight line distances between supply center and power plant is sufficient to explain variable transportation costs faced by the power plant. Transportation cost coefficients (both fixed and variable) were estimated by ICF Inc. (1980) using simple linear regression. Unfortunately, the variable transportation cost coefficient is in terms of rail miles instead of air miles. Since rail lines are never straight lines between coal supply and power plant, the variable cost coefficient must be transformed from rail miles into air miles. Duffield et al. (1982) calculated rail mile/air mile ratios based on a review of actual Burlington Northern unit train shipments out of different coal supply centers to various power plant locations. These ratios reflect BN's circuitry of rail routes from coal supply centers to power generator sites. These rail mile/air mile ratios are adapted in this study so that transportation costs may be better approximated. Table 4.6 summarizes the final transportation coefficients used in this study.

The second major reason for not utilizing values of variables collected for each power plant observation involves the following. If actual power plant sizes, full load time factors, and capital fixed charge rates were collected from each power plant observation, the magnitude of the cost differential will not reflect the relative competitiveness of Powder River coal over other coals <u>across</u>

Table 4.6 Transportation Cost Parameters[a]

Region of Origin	Item Number[b]	North East	South East	East Central	West Central	South Central	SWY ID	CO	UT NV	NM AZ	WA,OR CA	MT,ND SD,WY
		Region of Destination										
All coal mined from the Appalachian Region	1	4.19	4.19	4.19	4.19	4.19	4.53*	4.53*	4.53*	4.53*	4.53*	4.19
	2	1.35	1.35	1.35	1.35	1.40	1.35	1.35	1.35	1.40	1.55	1.35
	3	0.0154	0.0154	0.0154	0.0154	0.0160	0.0166*	0.0166*	0.0166*	0.0172*	0.0191*	0.0154
All coal mined between Appalachia and the Mississippi	1	4.19	4.19	4.19	2.73	2.73	2.95*	2.95*	2.95*	2.95*	2.95*	2.73
	2	1.35	1.35	1.35	1.35	1.40	1.35	1.35	1.35	1.40	1.55	1.35
	3	0.0154	0.0154	0.0154	0.0153	0.0158	0.0165*	0.0165*	0.0165*	0.0171*	0.189*	0.0153
All coal mined between the Mississippi and the Continental Divide	1	3.29	3.29	2.73	1.04	1.04	1.12*	1.12*	1.12*	1.12*	1.12*	1.04
	2	1.35	1.35	1.35	1.35	1.40	1.30	1.30	1.30	1.40	1.55	1.30
	3	0.0153	0.0153	0.0153	0.0153	0.0158	0.0159*	0.0159*	0.0159*	0.0171*	0.0189*	0.0147
All coal mined west of the Continental Divide	1	3.55*	3.55*	2.95*	1.12*	1.12*	1.04	1.04	1.04	1.04	1.04	1.12
	2	1.35	1.35	1.35	1.35	1.40	1.30	1.30	1.30	1.40	1.55	1.30
	3	0.0165*	0.0165*	0.0165*	0.0165*	0.0171*	0.0147	0.0147	0.0147	0.0158	0.0175	0.0159*

a. Source: ICF Inc. (1980), Duffield et al, (1982)

b. 1 = Fixed transportation rate in $/ton

2 = Rail mile / Air mile ratio

3 = Variable transportation Rate in $/ton/distance

* = Includes 8% surcharge for crossing the Rocky Mountains (Doe, 1980)

all power plants within a particular region. If MWi, Ti, and RATEi were not held constant over all power plants, bias would possibly result in our qualitative response regression analysis. To illustrate this problem let us suppose that the Nebraska Public Power and Light Company wishes to build two base load RNSPS coal fired generators that are spatially adjacent to one another. The plants are exactly identical except for the megawatt capacity; plant A is 250 MW in size while plant B is 500 MW in size. Because plant B is twice the size of plant A, the total cost of electricity generation given any one coal supply is twice the amount for plant B relative to plant A (according to equations 4.1 through 4.5). Because of this, the cost differential for plant B is twice that of plant A. Suppose plant A faces a positive cost differential of $1,000,000. This means that it costs $1,000,000 less to use Powder River coal as opposed to the least cost coal alternative. Because plant B is twice the size of plant A, plant B's cost differential is $2,000,000. This cost differential does not imply that Powder River coal is twice as competitive than the next best alternative for plant B relative to plant A. The absolute magnitude of the cost differential would not necessarily measure the relative competitiveness of Powder River coal across all power plant observations. We do not want cost differentials to measure differential plant sizes, differential full time equivalents, or differential capital rates of annualization. We are interested in calculating a single variable for each power plant observation whose magnitude reflects how well Powder River coal competes with other coals as a function of differential coal quality and power plant location. To standardize the magnitude of the cost differentials therefore, the following assumptions are made in the calculation of total costs of steam coal electric generation:

1) All power plants in the data sample are new 1971 NSPS or 1978 RNSPS base load electric generators that are 500 MW in size. 500 MW is the approximate mean of the power plant size in the data sample. Also, Duffield et al. (1982) used a power plant size of 500 MW for their hypothetical model power plant basing their choice on the approximate mean of power plant sizes on order in the year 1980.

2) Power plant i will have a base load lifetime capacity factor of 65%. EPRI (1979) estimated base load capacity range from 50% to 70% with actual service depending on the availability of the specific power plant unit and the power system to which it is attached. Modern power plants are advertised to have a 70% lifetime capacity factor (Davenport, 1981). Duffield et al. (1982) therefore used a lifetime capacity factor of 65% in their analysis. Ti = (.65 x 8790hr.) = 5694 hours per year full time capacity.

3) In addition, the real rate of annualization of capital costs was calculated by Duffield et al. (1982) to

be 7.41%; the calculation of this value included the weighted cost of capital, depreciation, taxes, and tax credits.

4) Finally, all costs in this analysis are in terms of current 1980 dollars. The cost differentials will be represented in <u>one million 1980 dollars</u>.

Computer software calculating the total costs of electric generation for each power plant in the data sample has been developed by the author. Utilizing the above assumptions, thirteen separate calculations are made; each calculation describes the approximate <u>standardized</u> total cost of generating electricity resulting from using coal from one of thirteen coal fields in the U.S. Finally, a cost differential is calculated for each power plant observation. This computer program outputs a data file containing observations on:

1) The plant identification number.

2) The plant X-Y coordinates (that will later be used to create maps).

3) The plant's actual coal choice (0 = non-Powder River coal, 1 = Powder River coal).

4) The standardized cost differential.

5) Dummy variables on FGD use (0 = no FGD , 1 = yes FGD).

6) Power plant generation class based on Federal NSPS laws. (0 = 1971 NSPS plants , 1 = 1978 RNSPS plants).

IV. Statistically Testing the Theory of Spatial Coal Markets

Past spatial market studies on coal choice drew hypothetical market boundaries that delineated areas where one coal was cost effective over other coals; plants located on the market boundary, were indifferent as to what coal they preferred. Given plant location (and if utility companies are cost minimizing), we would expect generating facilities with negative cost differentials to buy non-Powder River coal; power facilities with positive cost differentials should buy coal from the Powder River Basin. To see if these expected relationships hold in the power plant data sample, several statistical tests can be employed. These tests include the one-way analysis of variance, the grouped T-test, and the non-parametric Mann-Whitney test.* The purpose of these tests is to see if the means of a single variable of two independent group samples are significantly different. If a one tailed test is employed, we can test to see if the mean of one group is significantly higher or lower than the mean of the other group. In this case, the groups being measured is the cost differential; the two groups are non-Powder River coal users and Powder River coal users.

---

* These statistical tests are discussed at length in Snedecor and Cochran, (1980).

The hypotheses being tested are:

$$H0 : M0 = M1$$

versus

$$HA : M0 \neq M1$$

where:

M0 = the population mean of the cost differential for non-Powder River coal users.

M1 = the population mean of the cost differential for Powder River coal users.

Table 4.7 gives summary statistics on the cost differential by both groups of coal users. The results of the statistical tests are as follows:

A. One-Way ANOVA Results

Table 4.8 shows the results of the one-way analysis of variance test.

Table 4.8 One-way ANOVA of the Cost Differentials Between Powder River and non-Powder River Coal Users.

Source of variation	D.F.	Sum of Squares	Mean Squares
Between groups	1	9737.0	9737.0
Within groups	409	18509.8	45.3
Total	410	28246.8	

$F(1, 409) = 215.153$

$Prob(F(1, 409) > 215.153) = 0.0000$

Table 4.7

Summary Statistics on the Standardized Cost Differential
(in one million 1980 dollars)

I. Cost differential statistics for the entire sample.

n = 411

Mean	= -4.215
Variance	= 68.895
Standard Deviation	= 8.300
Minimum	=-26.246
Maximum	= 13.248

II. Cost differential statistics for non-Powder River coal users.

n = 282

Mean	= -7.507
Variance	= 58.297
Standard Deviation	= 7.635
Minimum	=-26.246
Maximum	= 8.501

III. Cost differential statistics for Powder River coal users.

n = 129

Mean	= 2.982
Variance	= 16.628
Standard Deviation	= 4.078
Minimum	=-14.491
Maximum	= 13.248

The calculated F statistic of this test equals 215.153. The probability value of the statistic at 1 and 409 degrees of freedom is zero to at least four decimal places. Since the probability of observing a F greater than or equal to the calculated F is so low, we reject the null hypothesis H0 and accept the alternative HA. There is a significant difference between the cost differentials of both groups of coal users.

B. One Tailed Grouped T-test Results

Though there is a significant difference between the two coal user groups in terms of their cost differentials, we have not established which group generally has a larger cost differential. We expect that the mean cost differential is greater for Powder River coal users than for non-Powder River coal users. We will now conduct a one tailed t-test to establish which group has a significantly larger cost differential. The test hypotheses are stated as:

$$H0 : M0 \geq M1$$

versus

$$HA : M0 < M1$$

Table 4.9 shows the results of the grouped t-test.

Table 4.9 Grouped Cost Differential t-test by Coal User Group

Group	n	$\bar{X}$	s	Pooled Variance Estimate: t	DF	2 Tail Significance Probability
Non-Powder River	282	-7.51	7.64	-14.67	409	0.0000
Powder River	129	2.98	4.08			

Prob($|t| > 14.67$) = 0.000

Prob($t < -14.67$) = 0.000

Since the probability of observing a t less than or equal to the calculated t is extremely low, we again reject the null hypothesis H0 and accept the alternative HA. The mean cost differential of Powder River coal users is greater than the mean cost differential of non-Powder River coal users.

C. Mann-Whitney Rank Sum Test Results

Since the F-test and the t-test are heavily dependent upon normally distributed data, it is generally useful to use a robust nonparametric rank sum method to test the above mentioned hypotheses. The Mann-Whitney test, developed by Wilcoxon, will be used to test all the above mentioned hypotheses (both two tailed and one tailed tests).

Table 4.10 shows the results of the Mann-Whitney test.

Table 4.10 Mann-Whitney Test of the Cost Differential Between Coal User Groups.

Group	Mean Rank	n	Corrected for Ties: Z	Corrected for Ties: 2 Tail Significance Probability
Non-Powder River	154.21	282	-13.07	0.0000
Powder River	319.22	129		

Prob($|Z| > 13.07$) = 0.0000

Prob($Z < -13.07$) = 0.0000

The approximate normal deviate Z equals -13.0695. The probability value of the calculated Z is so low that we reject H0 in both the two tailed and one tailed case. There is a significant difference in the means of the cost differentials of both coal user groups. In fact, the cost differential of the Powder River coal user group is higher than the cost differential of the non-Powder River coal user group.

In conclusion, as we expected, there is a statistically significant difference in the cost differential between actual Powder River coal users and non-Powder River coal users. In fact, we have shown that the cost differential is generally higher for Powder River coal users as theory suggests. These statistics support the assertions made by the theory of spatial coal markets. Coal market studies based on the theory of spatial markets can significantly distinguish Powder River coal users from non-Powder River coal users. <u>To what degree</u>, however, can an analysis based upon cost differentials distinguish actual Powder River coal users from non-users? The probabilistic qualitative response spatial coal market model, that will be estimated in the following chapter, can answer this particular question.

V. References

Coal Week. "Current Contract and Spot Marker Steam Coal Prices." New York: McGraw-Hill. 1980-1981 Various Issues.

Davenport, D. Personal communication, The Montana Power Company. 1981.

D.O.E. Service Report: Preliminary 1985, 1990, and 1995 Energy Forecasts for the Annual Report to Congress 1980. Draft. SR/IA/80-16. Washington D.C. November 1980.

Duffield et al. Projections of Coal Demand From the Northern Great Plains Through the Year 2010. (Final Report, OSM, May 1982).

Electric Power Research Institute. Technical Assessment Guide. EPRI PS-1201-SR. Palo Alto, CA July 1979.

Electric Power Research Institute. CRA/EPRI Coal Market Analysis System Volume 2: Technical Background. Prepared by Charles River Associates Inc. EPRI EA-907. Palo Alto, CA July 1979.

Green, John W. Coal Power Plant Scenario 1976-2000. Natural Resource Economics Division; Economics, Statistics, and Cooperatives Service; U.S. Department of Agriculture. February 1982.

ICF Inc. The Economic and Financial Impacts of Environmental Regulations on the Electric Utility Industry. Prepared for the Edison Electric Institute. February 1980.

ICF Inc. Preliminary Evaluation of Economics of NSPS Offsets for Utility Powerplants. Prepared for: Office of Policy and Planning, EPA. Washington, D.C. June 1980.

ICF Inc. The Potential Energy and Economic Impacts of Coal Slurry Pipelines. Final Report. Prepared for: DOE. Washington, D.C. January 1980.

ICF Inc. Coal and Electric Utilities Model Documentation (Draft). Washington, D.C. May 1980.

Keystone. Keystone Coal Industry Manual. New York: McGraw-Hill. 1980.

Kidder, Peabody and Company Inc. Electric Utility Generating Equipment Status Report on Fossil Boilers. March 1981.

Komanoff, Charles. Power Plant Cost Escalation: Nuclear and Coal Capital Costs, Regulation and Economics. Komanoff Energy Associates, New York. 1981.

Snedecor, G. and Cochran, W. Statistical Methods. Seventh Edition. Ames, Iowa: The Iowa State University Press. 1980.

U.S. Environmental Protection Agency. Materials obtained from and communications with Forrest Reinhardt. Office of Policy and Planning, EPA. Washington, D.C. 1983.

CHAPTER FIVE

ESTIMATING THE QUALITATIVE RESPONSE SPATIAL MARKET MODEL FOR POWDER RIVER COAL.

I. The Estimated Logit Equation

As mentioned in Chapter 3, the logit model specification to be estimated is:*

$$Pi = [\ 1 + e^{-(B0 + B1Ci + B2Fi + B3Ai + B4CiFi + B5CiAi)}]^{-1}$$

where:

Pi = the probability of power plant i buying Powder River Coal.

Ci = the standardized cost differential for plant i represented in one million 1980 dollars.

Fi = the dummy variable on whether or not plant i uses FGD (0 = no FGD, 1 = yes FGD).

Ai = the dummy variable on which set of NSPS air quality regulations power plant i faces (0 = 1971 NSPS, 1 = 1979 RNSPS).

Bj = model parameters

e = Napierian logarithm (approximately 2.7183).

As shown in Chapter 3, the above model may be represented as:

$$\ln \frac{Pi}{1 - Pi} = B0 + B1Ci + B2Fi + B3Ai + B4CiFi + B5CiAi$$

* In reference to the footnote on page 67 in chapter 3, a logit regression equation, where the FGD dummy variable Fi is excluded from the analysis, is estimated and presented in Appendix D.

Using the maximum likelihood regression technique along with the data described in Chapter 4, the estimated equation is:

$$\ln \frac{Pi}{1 - Pi} = \underset{(.227)}{0.632^{*}} + \underset{(.057)}{0.463^{*}}Ci - \underset{(.350)}{2.664^{*}}Fi + \underset{(.294)}{1.441^{*}}Ai$$

* = significant at the 99% confidence level.

The independent interaction variables CiFi and CiAi did not enter into the model because they did not significantly explain variation in the dependent variable and because their estimated coefficients were not significantly different than zero. When CiFi was entered into the model, the improvement in the explanatory value of the model was significant only at the 38% confidence level. When CiAi was entered, the improvement in the explanatory value of the model was significant at the 10% confidence level. The calculated t-statistics for the estimated coefficients on CiFi and CiAi are -0.656 and 0.431 respectively. The coefficient on CiFi is significantly different than zero at approximately the 50% confidence level while the coefficient on CiAi is significantly different than zero somewhere below the 50% confidence level. For these reasons, coefficients on the terms CiFi and CiAi are not presented in the estimated equation; B4 and B5 can be thought of as equalling zero.

II. Test Statistics on the Estimated Equation

First, we wish to see if the independent variables of the estimated model significantly contribute to the explanation of variations in Pi. The hypotheses to be tested include:

HO : B1 = B2 = B3 = 0.0

versus

HA : At least one coefficient Bj other than the constant B0 contributes to the explanation of Pi.

A chi-squared test is employed to test these hypotheses (see chapter 3). The results of the test are as follows:

$\ln(Lo) = -255.708$

$\ln(Lmax) = -99.082$

$X^2_{(3)} = 2[\ln(Lmax) - \ln(Lo)]$

$= 313.252$

At 3 degrees of freedom, Prob($X^2_{(3)} > 313.252$) = 0.0000

The two tail significance probability of the calculated chi-squared statistic at three degrees of freedom is zero to four decimal places. We reject the null hypothesis HO and accept the alternative HA. At least one estimated coefficient (other than the constant) contributes to the explanation of Pi.

To test the significance of individual Bj coefficients, a two tailed t-test is employed (see chapter 3). The hypotheses to be tested are:

$$HO : Bj = 0$$

versus

$$HA : Bj \neq 0$$

Table 5.1 presents the results of the t-test:

Table 5.1 T-test of Individual B_j Coefficients

Term	Coefficient	Standard Error	Calculated T-statistic	2 Tail Significance Prob.
Constant B_o	0.632	0.227	2.788	0.005
Cost diff. B_1	0.463	0.057	8.085	0.000
FGD B_2	-2.664	0.350	-7.610	0.000
RNSPS B_3	1.441	0.294	4.904	0.000

All coefficients are extremely significant. The least significant coefficient, B0, would be significant even at the 99.5% confidence level. For each coefficient in our model, we reject H0 and accept the alternative HA.

Thus far, we have determined that the model and all of the coefficients are significant in explaining variation in Pi. Next, we wish to determine the fit of the model to observed data. As I have argued in Chapter 3, the interpretation of R-squared for qualitative response models must be handled with care. R-squared <u>does not necessarily</u> measure the amount of variation in the dependent variable explained by the independent variables. It should be noted however that Morrison (1972) discussed that the upper bound for R-squared, when true probabilities are distributed evenly over an interval of the independent variable, is likely to be 0.3333. Our R-squared calculations are approximately twice the hypothetical upper bound; it seems that our model is fitting observed data fairly well. We can not be sure, however, what the distribution of the true probabilities are in the power plant population. R-squared therefore, is best used to compare the relative worth of competing logit specifications. Since I do not have any competing logit models in this study, I present calculations of R-squared for the curious and for those who later may wish to compare my estimated equation with another estimation.

McFadden's R-squared = 0.6125

Effron's R-squared = 0.6268

Finally, we calculate the proportion of correct predictions to check the fit of our model (see Chapter 3). If Pi is greater than or equal to 50%, we predict that the power plant will use Powder River coal; otherwise we predict that the power plant will use another coal.

Proportion of correct predictions = 0.8905

Based on these test statistics, the model significantly explains coal choosing behavior. In fact, the model can predict correctly approximately 89% of the time. Although the proportion of correct predictions statistic is possibly biased since it weighs all prediction probabilities between 0.50 through 1.0 and 0.0 through 4.9 equally (see Chapter 3), 89% correct predictions seems to be an indication of a good fitting model.

III. Model and Coefficient Interpretation

Because of the dummy variables included in the logit specification, there are actually three models given by the estimated equation. These three models include:

1) 1971 NSPS plants <u>without</u> FGD.

2) 1971 NSPS plants <u>with</u> FGD.

3) 1978 RNSPS plants with FGD.

(FGD is mandatory on RNSPS plants)

Tables 5.2 through 5.4 summarizes these three models.

As we can see from tables 5.2 through 5.4, the effect of FGD and RNSPS on the probability of buying Powder River coal is tremendous. For a NSPS plant without FGD, the cost differential must be less than negative 1.4 million dollars before we predict that the plant will not use Powder River coal. For NSPS plants with FGD, the situation reverses; the cost differential must be greater than a <u>positive</u> 4.4 million dollars before we predict that the plant will use Powder River coal. Lastly, the cost differential on RNSPS

Table 5.2

NSPS plants without FGD

$$\ln \frac{P_i}{1 - P_i} = 0.632 + 0.463\ C_i$$

Cost Differential (in one million 1980 dollars)	-13.0	-6.5	-3.0	-1.4	-1.0	0.0	1.0	3.0	6.5	13
Probability of buying Powder River Coal	.0046	.0849	.3193	.5000	.5421	.6529	.7493	.8830	.9745	.9987

Table 5.3

NSPS plants with FGD

$$\ln \frac{P_i}{1 - P_i} = -2.0320 + 0.463\, C_i$$

Cost Differential	-13.0	-6.5	-3.0	-1.0	0.0	1.0	3.0	4.4	6.5	13
Probability of buying Powder River Coal	.0003	.0064	.0316	.0762	.1159	.1724	.3446	.5000	.7266	.9818

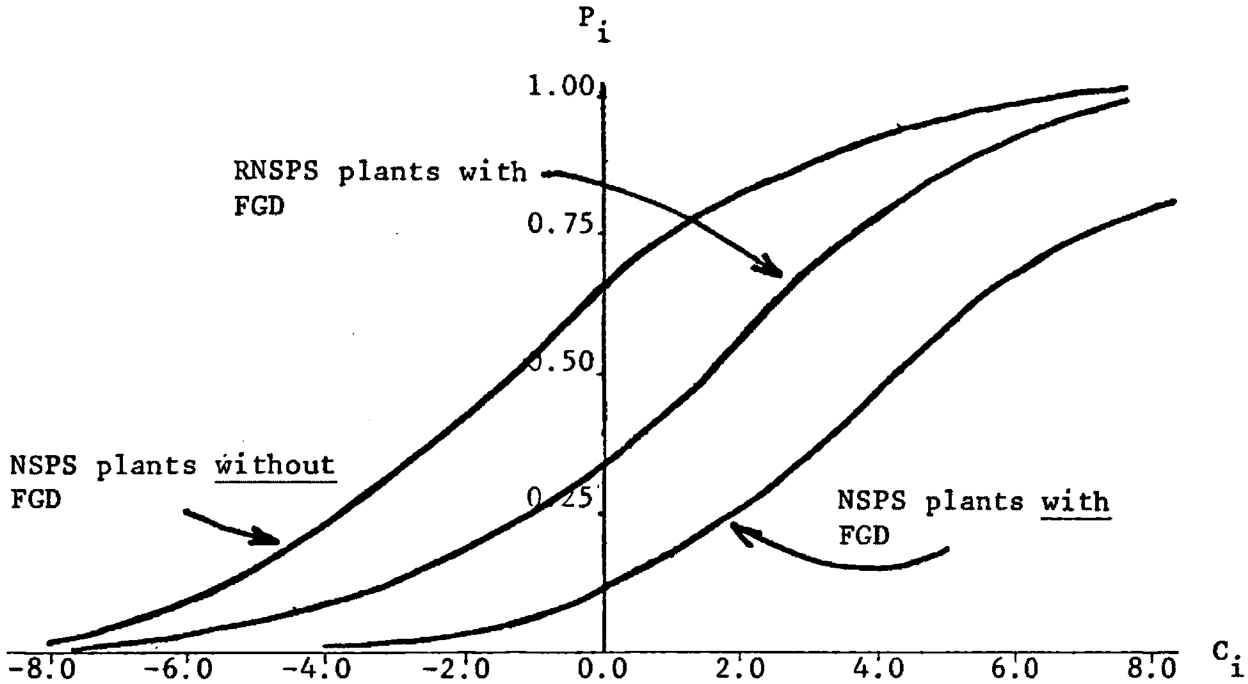
Table 5.4

RNSPS plants with FGD

$$\ln \frac{P_i}{1 - P_i} = -0.5910 + 0.463\, C_i$$

Cost differential (in one million 1980 dollars)	-13.0	-6.5	-3.0	-1.0	0.0	1.0	1.3	3.0	6.5	13
Probability of buying Powder River Coal	.0013	.0266	.1213	.2585	.3564	.4680	.5000	.6895	.9182	.9956

Figure 5.1
Logit Models for Three Power Plant Types



plants must be greater than positive 1.3 million dollars before we predict that the RNSPS plant will use Powder River coal. These results suggest that given any cost differential, NSPS plants without FGD prefer Powder River coal while NSPS plants with FGD prefer other coals. At any cost differential, RNSPS plants (all of which have FGD), will tend to prefer less Powder River coal relative to NSPS no-FGD plants. On the other hand, RNSPS plants prefer more Powder River coal than NSPS yes-FGD plants. (See figure 5.1).

These results make sense. If a NSPS plant did not use FGD, it would need to burn a low sulfur coal (such as Powder River coal) in order to meet the sulfur dioxide emission control regulation; if a NSPS plant did use FGD, the plant would most likely be scrubbing a high sulfur non-Powder River coal since scrubbing is not required on low sulfur coals according to the 1971 NSPS. For a RNSPS plant, scrubbers are required regardless of the coal used (a percentage reduction of sulfur is required). The correlation between coal choice and the FGD dummy variable no longer holds. Since the distinction Between FGD users and FGD non-users no longer exists for RNSPS plants, the "intercept" term on the RNSPS equation should be between the "intercept" terms of the two NSPS equations.

Because the interaction variables CiFi and CiAi did not enter the equation, the cost differential affects the probability of plant i buying Powder River coal similarly for each of the three models stated above.

Table 5.5 shows how a unit change (a one million dollar unit change) in Ci affects Pi. Remember that a change in Pi is a function of both Pi and Bl.

Table 5.5 The Effect of ΔCi on ΔPi
Given Bl = 0.463, and ΔCi = 1 (= one million dollars)

Pi	ΔPi
0.00	0.0000
0.10	0.0417
0.20	0.0741
0.30	0.0972
0.40	0.1111
0.50	0.1158
0.60	0.1111
0.70	0.0972
0.80	0.0741
0.90	0.0417
1.00	0.0000

If a plant were initially indifferent as to which coal it prefered (i.e. Pi = 0.50), a positive one million dollar change in the cost differential will change the prediction probability Pi by approximately 12%; the new Pi would equal 62%.

If multicollinearity is present among the independent factors of the model, one can not interpret the cost differential coefficient as suggested by table 5.5. To check the stability of the cost differential against the possibility of multicollinearity, several logit models were estimated from subsets of the power plant data. From these subsets of data, the cost differential coefficient was calculated and a 95% confidence interval, around the cost differential coefficient, was formed. If the confidence intervals of these new regressions contain the value of the estimated cost differential parameter of the original model, we can conclude that the coefficient is stable and not severely affected by collinearity.* The alternative logit models used for this test include:

1) The original model <u>minus 10</u> randomly selected observations.

2) A model where only <u>NSPS</u> plants are considered. The model specification is:

$$\ln \frac{Pi}{1 - Pi} = B0 + B1Ci + B2Fi$$

3) A model where only <u>RNSPS</u> plants are considered. The model specification is:

$$\ln \frac{Pi}{1 - Pi} = B0 + B1Ci$$

Table 5.6 summarizes the results of this test.

Table 5.6 Checking the Stability of the Cost Differential Coefficient.

Regression Model	Cost Differential Coefficiant	Standard Error	95% Confidence Level Lower	95% Confidence Level Upper
Original Model	0.463	0.057	0.3513	0.5747
Minus 10	0.459	0.057	0.3473	0.5707
NSPS	0.471	0.086	0.3051	0.6396
RNSPS	0.456	0.077	0.3051	0.6069

* Testing the stability of regression coefficients (against the adverse effects of multicollinearity) by breaking up the sample data into subsets is a common "rule of thumb" method used by statisticians.

Since the confidence intervals on the cost differential coefficient for each subset of data contains the estimated coefficient of the original model, we can be confident that the coefficient is stable and not adversely affected by multicollinear effects. We can interpret the cost differential coefficient as suggested by table 5.5.

Lastly, in our original model specification, we did not include the interaction terms CiFi and CiAi as we had originally intended. As previously mentioned, these variables were omitted because they did not significantly improve the explanatory power of the model. One reason why this might have occurred is that Federal air quality regulations were deliberately designed to prevent one coal region from benefiting at the expense of another coal region; EPA was careful to prevent the emergence of cost incentives that would give some regions advantages over other regions (Duffield et al., 1982). Federal NSPS and RNSPS regulations were designed to be uniform standards applicable throughout the U.S.; since they are uniform standards, industry would not have the incentive to develop one region as opposed to another region because of differential costs caused by environmental regulation. The variables Fi and Ai did not severely affect the <u>relative</u> cost competitiveness of one regions' coal over another. It is obvious that the variables Fi and Ai change the <u>absolute</u> total cost of power generation; the cost differential Ci, however, does not measure the <u>absolute</u> cost of power generation. The effect of the cost differential on Pi therefore would be unaffected with changes in Fi and Ai.

In conclusion, the results of the probabilistic qualitative response logit model are promising. The model fits observed data very well. the variables entering the model are significant and make sense. There are no statistical estimation problems associated with this model. Table 5.7 condenses and summarizes the logit model results and statistics.

IV. The Spatial Interpretation of the Logit Model

The purpose of this study is to estimate an <u>empirical</u> spatial market model for Powder River coal. Our estimated logit model tells us the probability of a coal fired generator buying Powder River coal given certain power plant attributes. Used as is, the logit model does not directly tell us the spatial orientation of the Powder River coal market. In order to transform our logit model so that it roughly outlines the approximate geographical Powder River coal market, we may use the estimated logit equation to calculate the probability of buying Powder River coal for electric generators whose geographical location is known. We may then plot, on a map of the Continental U.S., the location of plants whose calculated probability falls in certain ranges. If empirical geographical coal markets can be defined, spatial patterns should emerge as coal choice probabilities change. For example, we expect that plants

Table 5.7

The Estimated Logit Model on the Probability of a Power Plant Purchasing Powder River Coal.

Estimated Equation:

$$\ln \frac{P_i}{1 - P_i} = \underset{(.227)}{0.632^*} + \underset{(.057)}{0.463^*}\, C_i - \underset{(.350)}{2.664^*}\, F_i + \underset{(.294)}{1.441^*}\, A_i$$

* = significant at 99% confidence level

$X^2_{(3)} = 313.252$ Probability $(X^2_{(3)} > 313.252) = 0.0000$

McFadden's R - squared = 0.6125

Effron's R - squared = 0.6268

proportion of correct predictions = 0.8905

Equation of NSPS plants without FGD:

$$\ln \frac{P_i}{1 - P_i} = 0.632 + 0.463\, C_i$$

Equation of NSPS plants with FGD:

$$\ln \frac{P_i}{1 - P_i} = -2.0320 + 0.463\, C_i$$

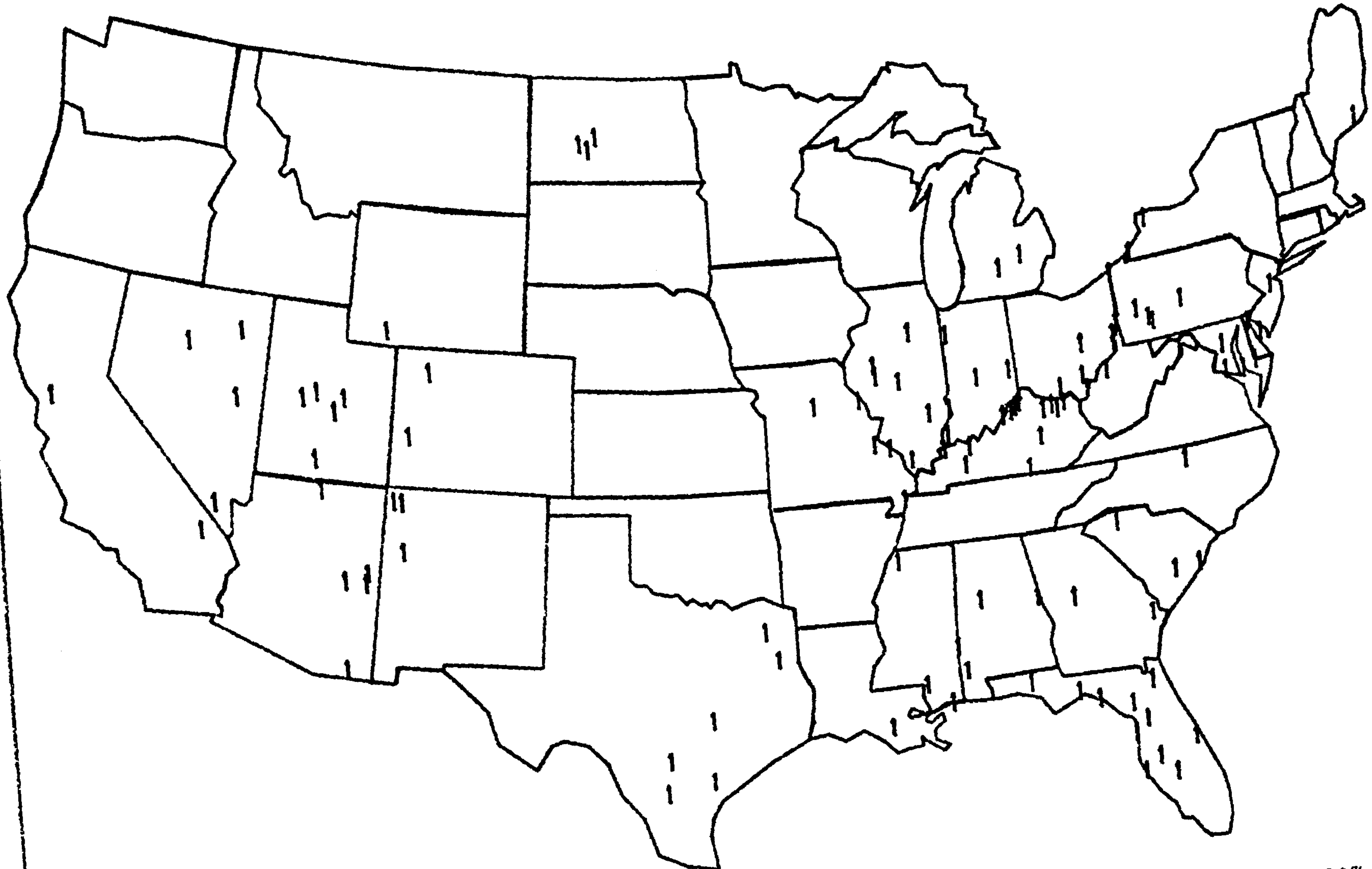
Equation of RNSPS plants:

$$\ln \frac{P_i}{1 - P_i} = -0.5910 + 0.463\, C_i$$

with a probability of buying Powder River coal between 0% and 20% to be located far away from the Powder River Basin relative to competing supply centers. As the probability of buying Powder River coal increases, the location of power plants should approach the Powder River Basin relative to competing coal supply centers; plants with a choice probability of 80% to 100% should be tightly patterned around the Powder River Basin. Figures 5.2 to 5.6 show the spatial orientation of the Powder River coal market based upon our logit model. Figure 5.2 shows the geographical location of plants (plants that are in our original data sample) who have a probability of buying Powder River coal between 0% and 20%. Figure 5.3 shows the location of plants with a choice probability of +20% to 40%. Figure 5.4 plots the position of plants in the "range of indifference" (probabilities ranging from +40% to 60%). Likewise, Figure 5.5 and 5.6 plot the location of plants in the ranges +60% to 80% and +80% to 90% respectively. As these figures indicate, the spatial orientation of the Powder River coal market behaves as expected.

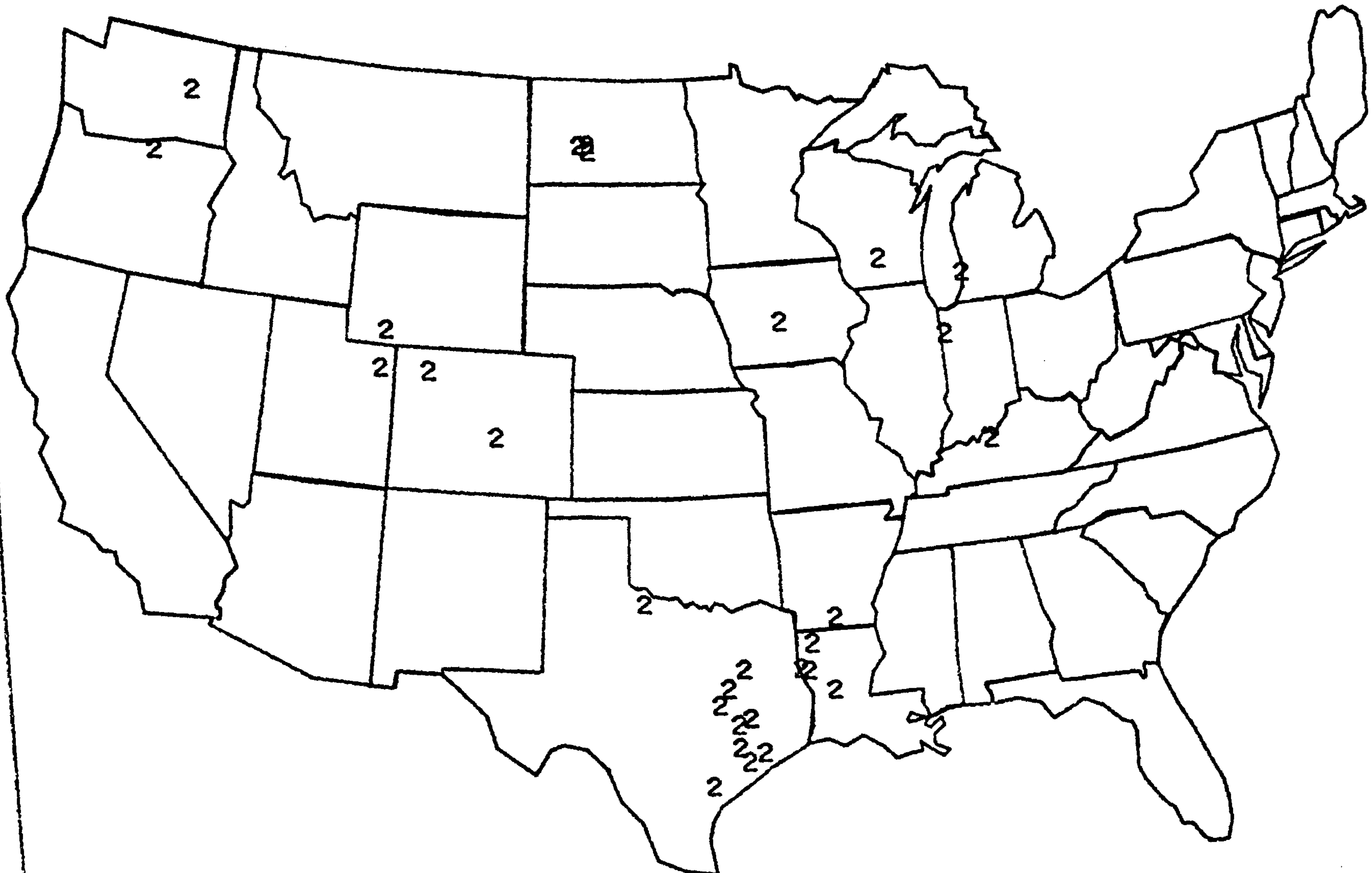
Figure 5.7 (with plastic overlays) enables the reader to see all the probability ranges together on one map (at this point, ignore the Base 1980 market boundary line). Although the spatial pattern of the Powder River coal market is visually apparent, interesting anomalies exist in the spatial pattern. The most noticeable anomalies are those where plants of different probability ranges are spatially

Figure 5.2



Coal fired generators with the probability of purchasing Powder River coal between 0% and 20%

Figure 5.3

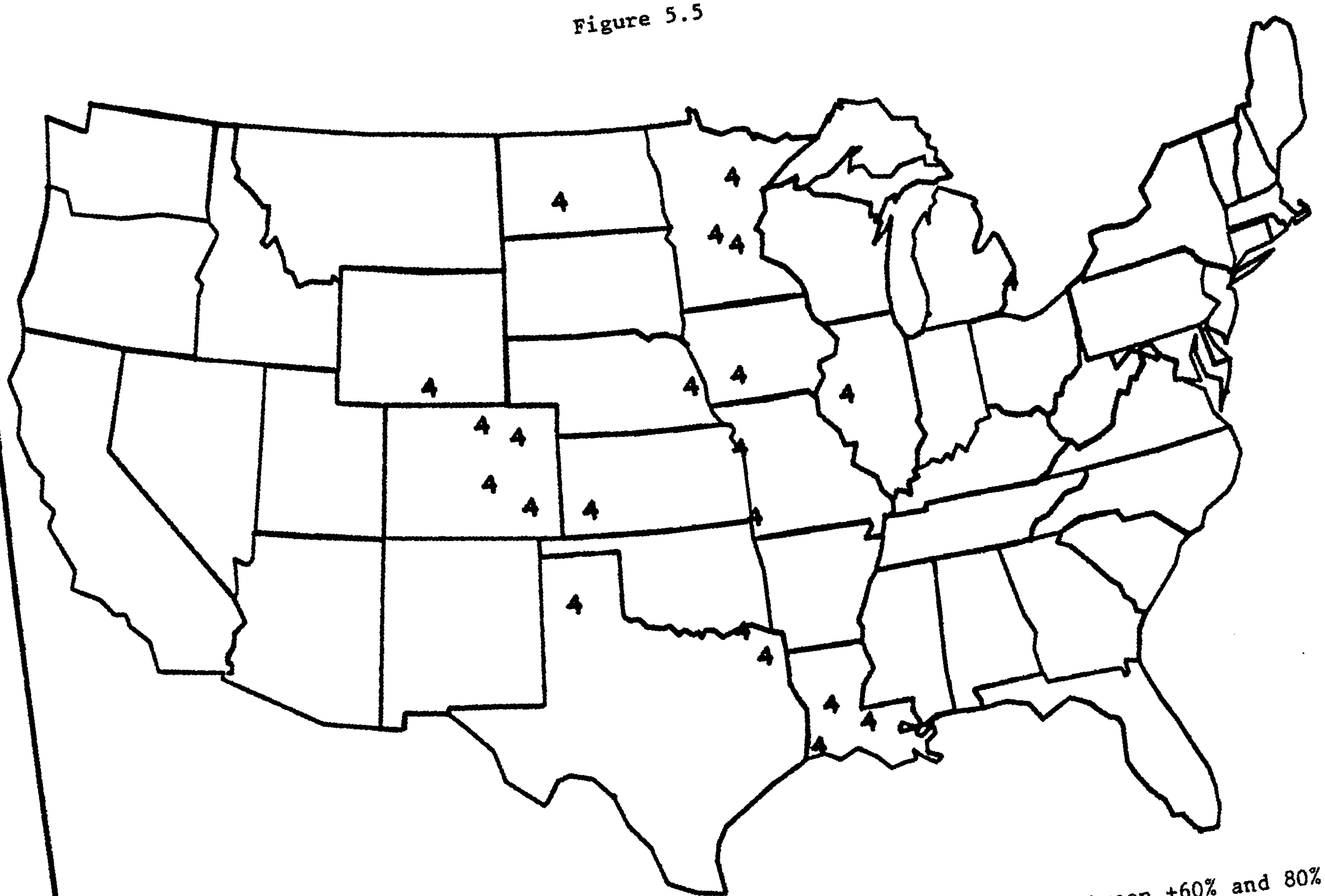


Coal fired generators with the probability of buying Powder River coal between +20% and 40%

Figure 5.4

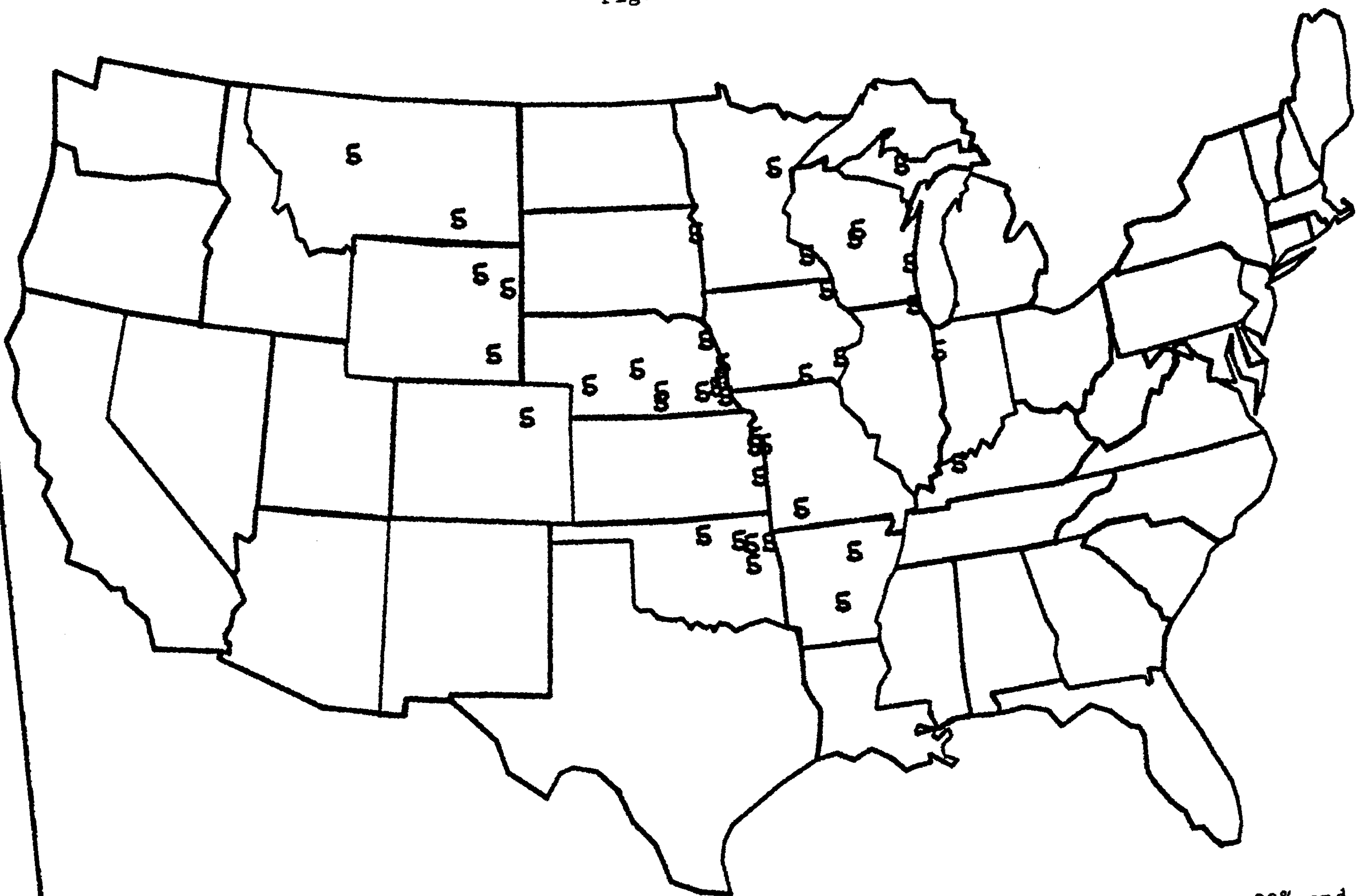
Coal fired generators with the probability of buying Powder River coal between +40% and 60%

Figure 5.5



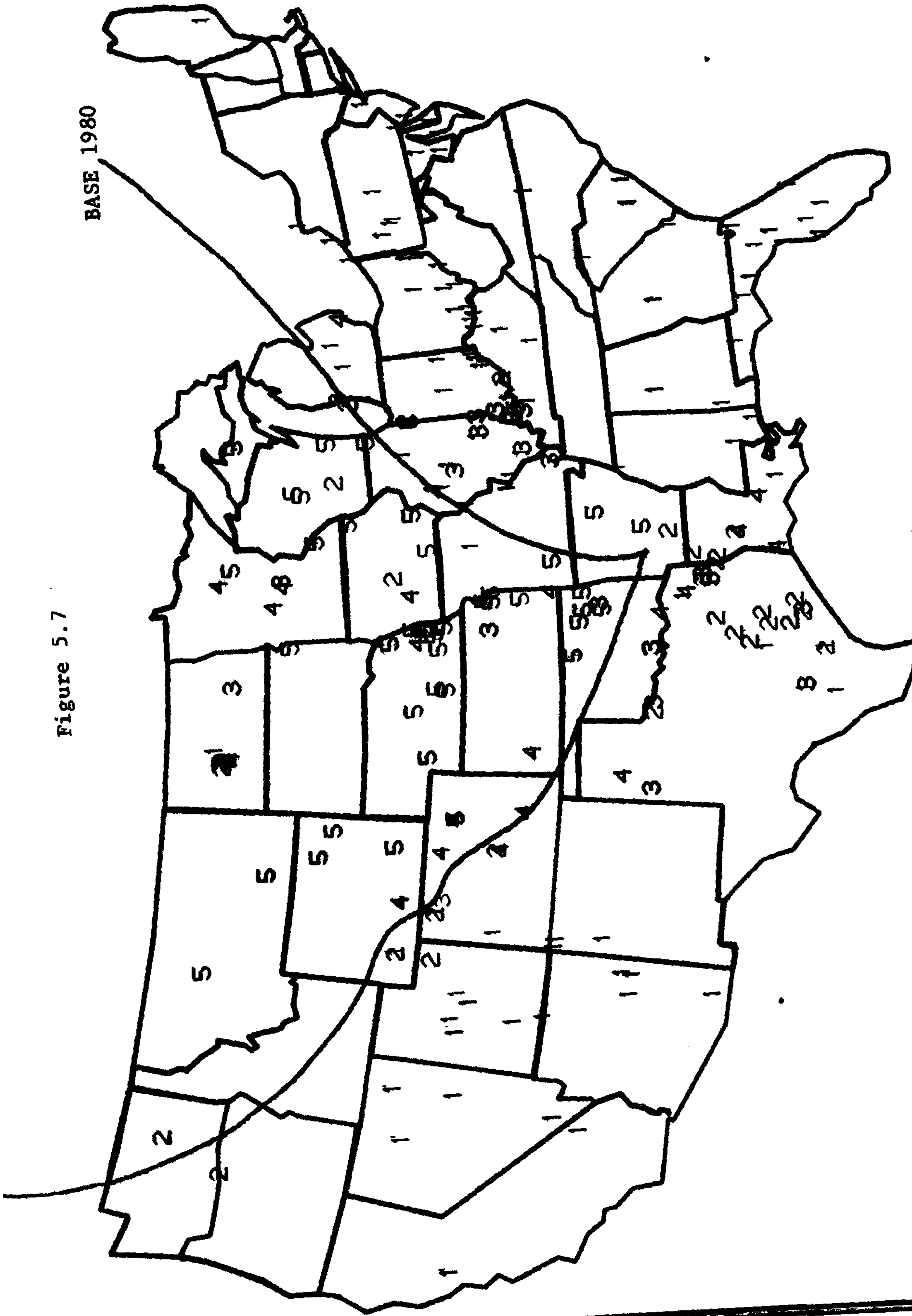
Coal fired generators with the probability of buying Powder River coal Between +60% and 80%

Figure 5.6



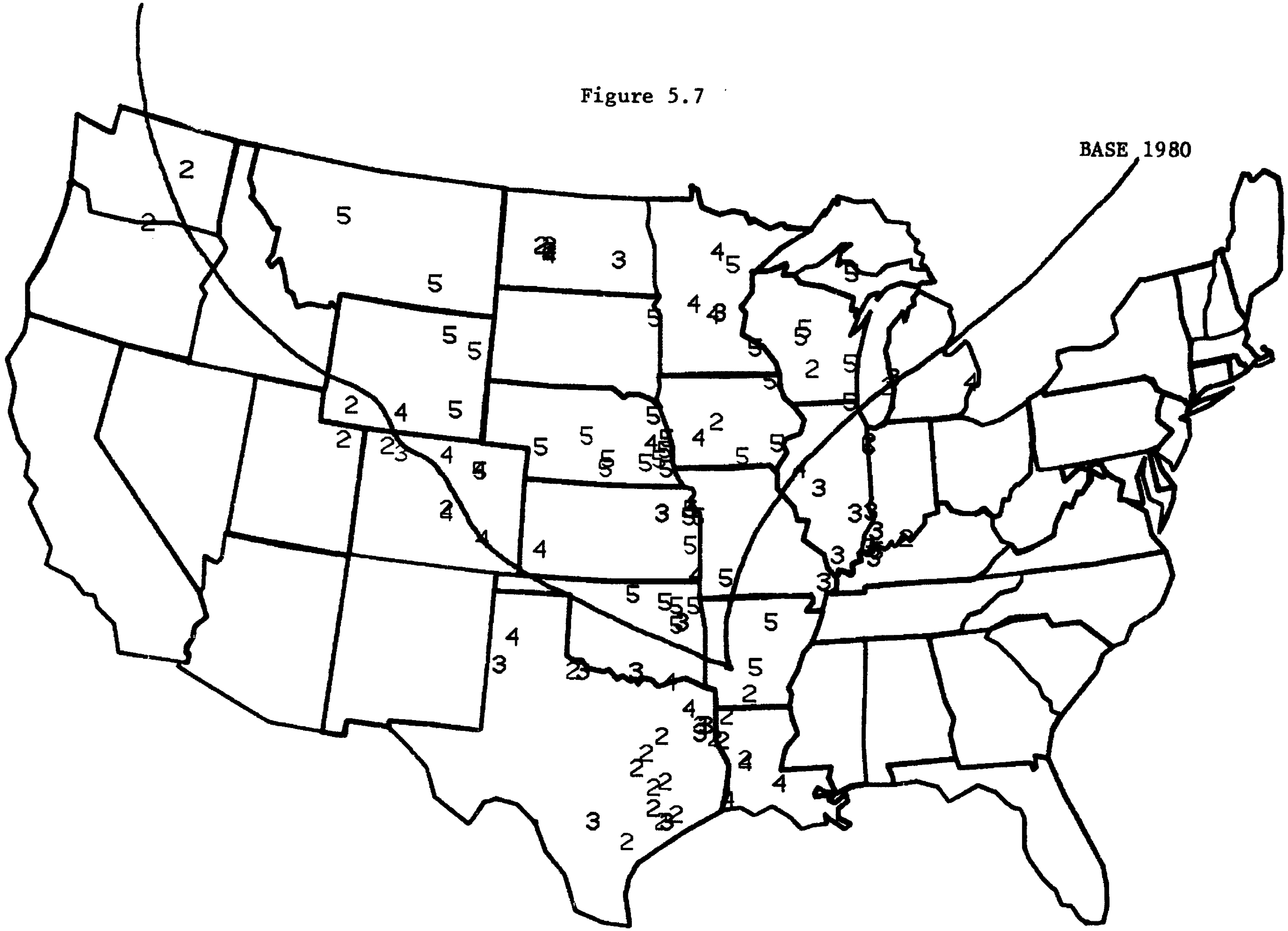
Coal fired generators with the probability of purchasing Powder River coal between +80% and 100%

Figure 5.7



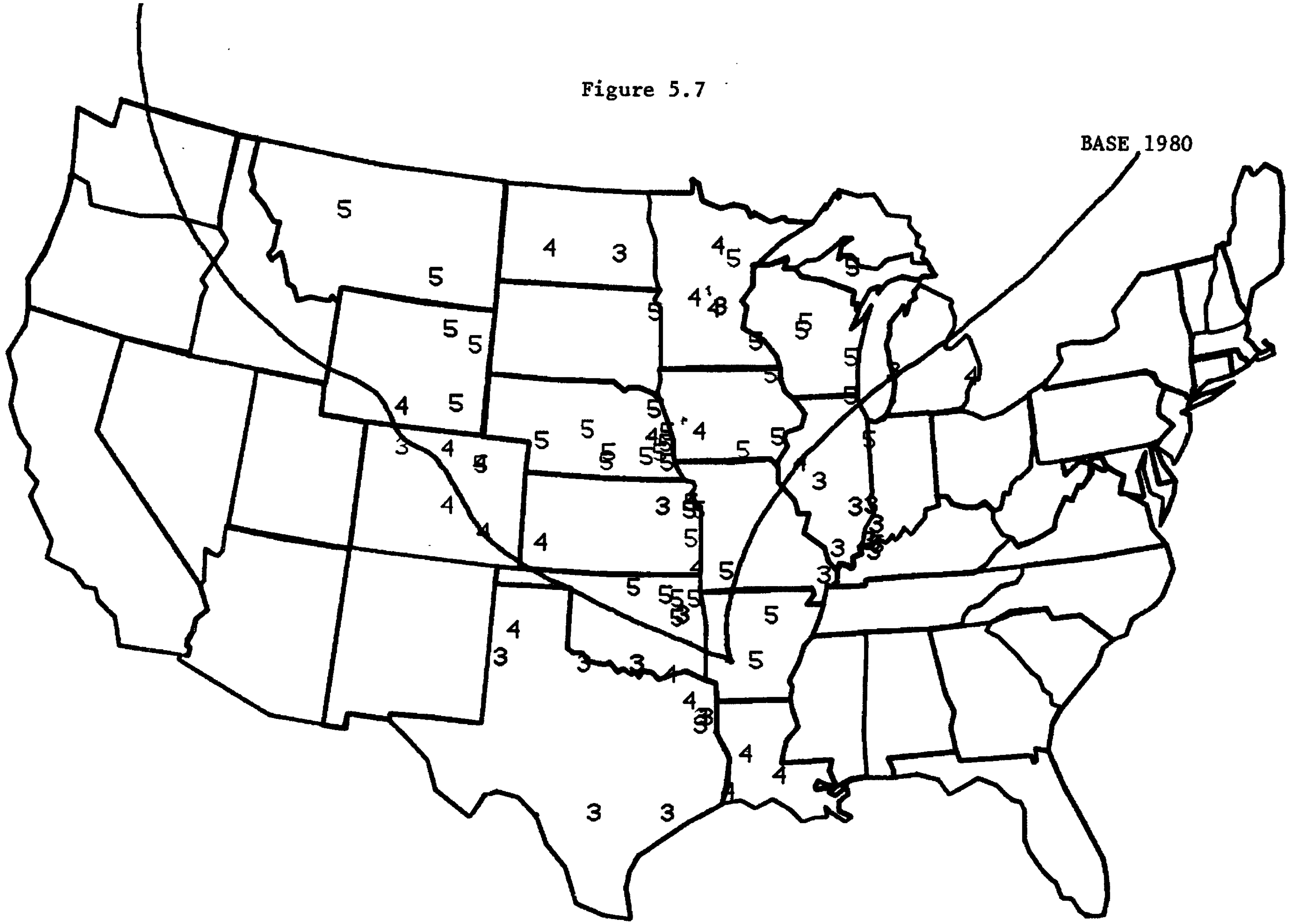
The Estimated Spatial Market for Powder River Coal
(1980 data)

Figure 5.7



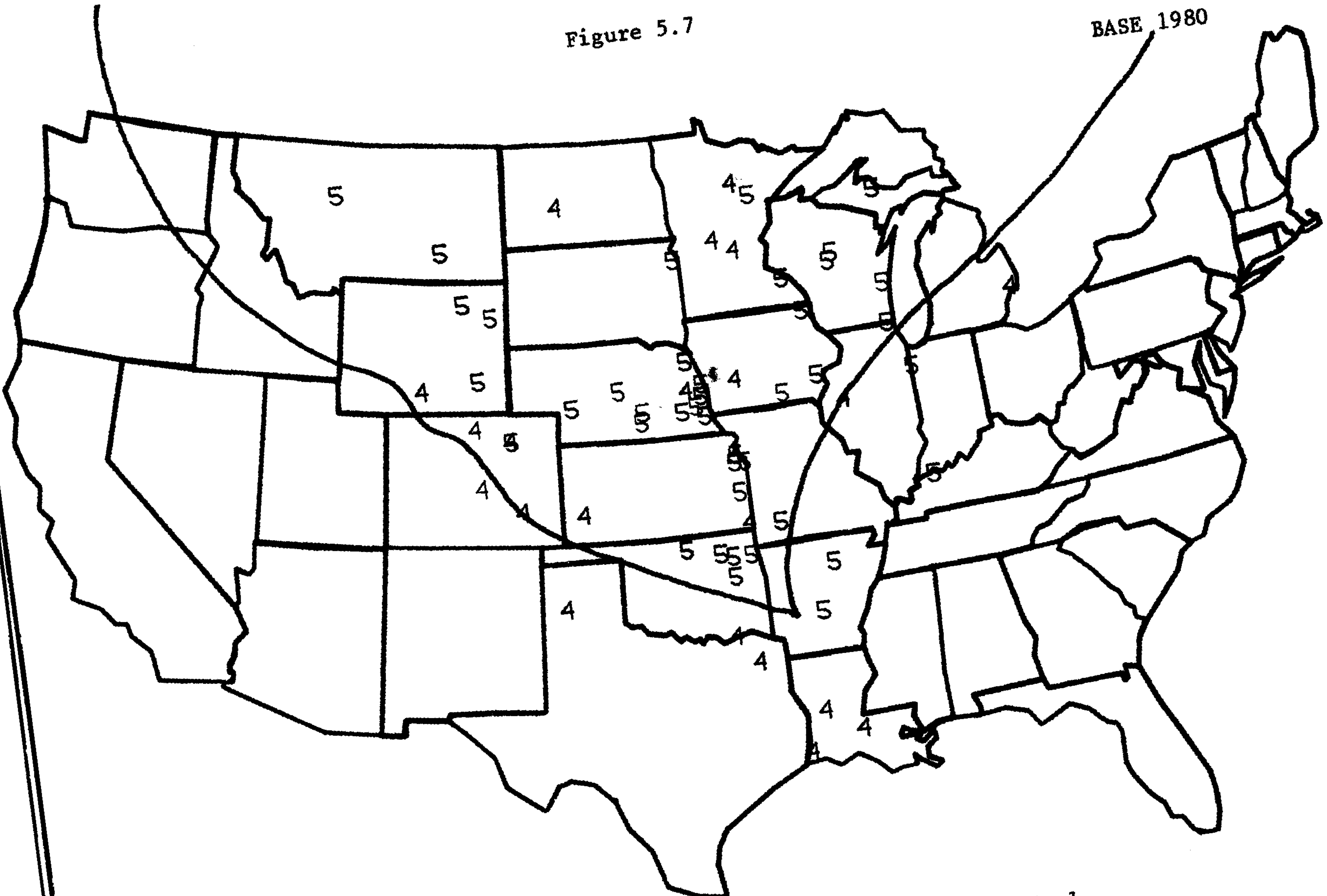
The Estimated Spatial Market for Powder River Coal
(1980 data)

Figure 5.7



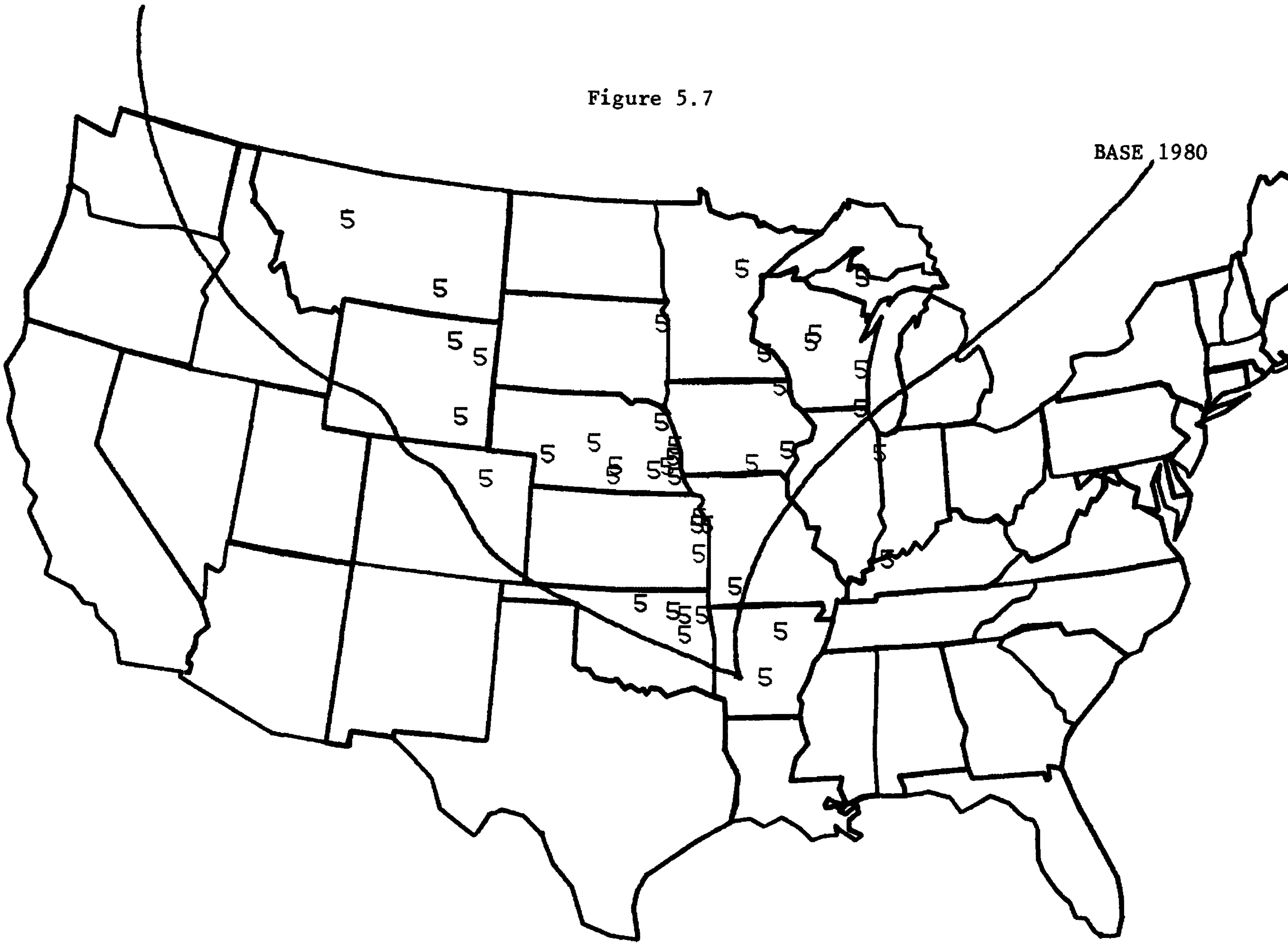
The Estimated Spatial Market for Powder River Coal
(1980 data)

Figure 5.7



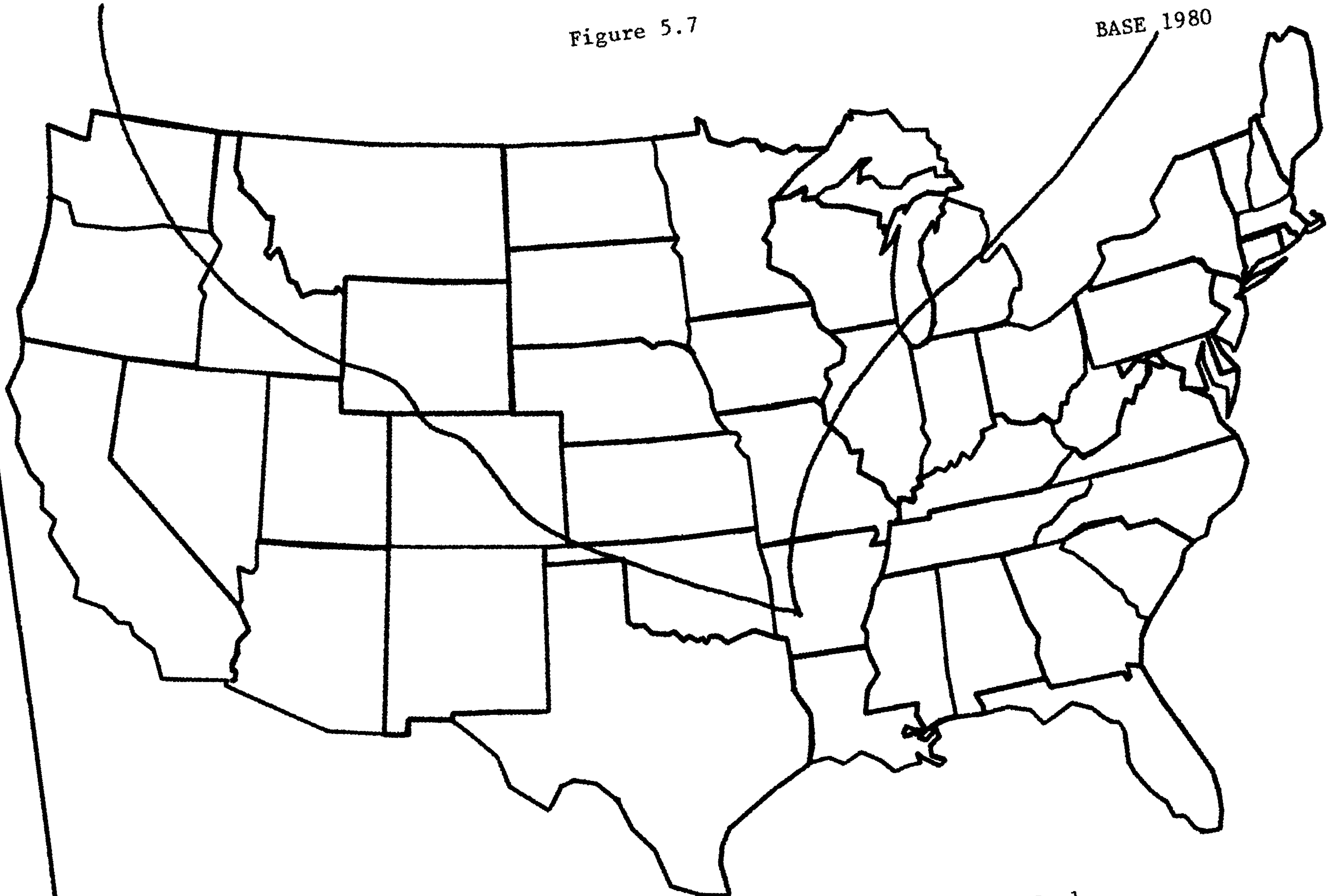
The Estimated Spatial Market for Powder River Coal
(1980 data)

Figure 5.7



The Estimated Spatial Market for Powder River Coal
(1980 data)

Figure 5.7



The Estimated Spatial Market for Powder River Coal
(1980 data)

adjacent to one another. In general, this occurs because power plants of different classes (i.e. NSPS plants versus RNSPS plants) exist in approximately the same space. For example, some plants in western Indiana and western Kentucky have a predicted probability of buying Powder River coal greater than 80%. These are older NSPS plants who find the cost of scrubbing high sulfur coal extremely expensive compared to the low sulfur coal alternative. Also, these older NSPS plants were built at a time of lower transportation rates thus giving a cost advantage to Powder River coal; older plants (especially NSPS plants who do not use FGD) therefore tend to use more Powder River coal. RNSPS facilities spatially adjacent to these particular plants tend to use local high sulfur coals. This is because the cost advantage, of using low sulfur coal in order to meet emission standards, was reduced; <u>all</u> RNSPS plants must scrub emissions. Also, these newer RNSPS plants faced higher transportation rates, at the time the initial coal choices were made, when compared to older NSPS plants; newer plants therefore tend to use local coal.

Other interesting anomalies in the spatial pattern exist at the "interface" of two coal supply centers. The interface of two coal supply centers is the area where geographical market boundaries overlap. These anomalies occur in the following areas:

1) Fort Union (North Dakota)

North Dakota lignite is mostly burned minemouth in Mercer county North Dakota while Powder River coal is cost effective in locations completely surrounding North Dakota. This is because lignite has extremely low value per unit weight making it uneconomical to be shipped any great distance.

2) Gulf Region (Louisiana, Texas)

Because of the economics of lignite transport (see number 1 above), Powder River coal can effectively compete with Texas lignite in the Gulf Region. Powder River coal has more BTU's per unit weight than Texas coal. Also, since Powder River coal does not have to be transported across the Rocky Mountains, Powder River coal can effectively compete with other western coals (e.g. New Mexico, Colorado, Utah, South Wyoming) in the Gulf Region; transportation rates are higher if the coal must be shipped across the Rocky Mountains (see chapter 4). Finally, since Powder River coal is relatively low in sulfur, Powder River coal can effectively compete with West Interior coal, East Interior coal, and Appalachian coal.

3) West Interior (Iowa, Kansas, Missouri, Nebraska, Wisconsin) and East Interior (Illinois, Indiana, Western Kentucky)

Since Powder River coal is relatively low in sulfur compared to West and East Interior coals, Powder River coal can effectively compete with coals mined in the Interior regions; the cost of scrubbing low sulfur coal is significantly less than scrubbing high sulfur coal.

The base map on figure 5.7 shows the theoretical base 1980 spatial market for NGP coal that was estimated by Duffield et al. (1982). It could be seen that the theoretical estimation of the spatial market has the same general shape as the market that is empirically estimated. The theoretical model diverges from the empirically estimated model most notably along the southern and eastern boundaries. The most likely reasons why these divergences occur is because:

1) Empirical power plants do not behave as neatly

as hypothetical power plants.

2) The theoretical spatial market model considers only new RNSPS power plants while our empirical model considers both NSPS plants and RNSPS plants.

3) The empirical spatial market analysis includes more coal supply centers than the theoretical analysis (see chapter 4).

4) The theoretical spatial market model defines market boundaries via fixed paired comparisons of total electric generating costs while the empirical model uses variable paired comparisons. For example, the theoretical market boundary between the Powder River coal supply center and the Texas lignite supply center is calculated by comparing the total costs of electric generation for hypothetical plants in the Gulf region whose coal choices are limited to Powder River coal and Texas lignite. The empirical model does not limit coal choices in this way; the empirical model compares the total cost of electric generation between the use of Powder River coal and the least cost non-Powder River coal alternative. Plants in the Gulf region might not find Texas lignite to be the least cost non-Powder River coal; plants in the Gulf region might find other non-local coals (New Mexico, Midwest, West Interior, South Appalachia) to be least cost alternatives to Powder River coal.

V. Conclusions

The empirical spatial orientation of the Powder River coal market is very evident. The potential of the estimated model has not been fully realized however. There are many avenues for further research using the estimated logit model for analyzing spatial coal markets. An interesting topic is how the market will be affected by changes in air pollution policy and changes in costs (e.g. transportation rates, coal prices, etc.). By transforming the input cost data consistent with the issue being analyzed, one could re-estimate the coal choice probabilities for all coal fired power plants. One could then plot the locations of plants, falling in certain probability ranges, to see how the data changes affects the spatial Powder River coal market.

A major shortcoming of the above analysis, and of the forecasts that might be made from the estimated model, is that the model is stochastic. The parameters defining coal markets are rapidly changing over time. Coal fired power plants are constantly slipping their on line dates and breaking coal and transportation contracts. It will not be long before the data used in the above analysis will be obsolete. Updating data and re-estimating new model parameters will be required in the near future. The author hopes that the procedure and methodology used in this study will be useful to those who will be treading this path in the coming years.

VI. References

Duffield, et al. Projections of Coal Demand From the Northern Great Plains Through the Year 2010. (Final Report, OSM, May 1982).

Morrison, D. G. "Upper Bounds for Correlations Between Binary Outcomes and Probabilistic Predictions," J. Amer. Statist. Assoc., March 1972, pp. 68-70.

APPENDIX A

A Fortran Computer Program that Calculates the Cost Differentials

```
C     PROGRAMMER:  HENRY Y. YOSHIMURA
C                  DEPARTMENT OF ECONOMICS
C                  UNIVERSITY OF MONTANA
C
C       PROGRAM:  THIS PROGRAM CALCULATES THE TOTAL COST OF ELECTRICITY
C                 GENERATION FOR EXISTING OR PLANNED COAL FIRED POWER
C                 PLANT UNITS.  GIVEN THE SIZE AND LOCATION OF THESE
C                 POWER PLANTS, THIS PROGRAM DETERMINES THE TOTAL COST
C                 OF ELECTRICITY GENERATION RESULTING FROM BURNING COAL
C                 THAT WAS MINED FROM THIRTEEN DIFFERENT COAL SUPPLY
C                 REGIONS.  BASE POWER PLANT COSTS, POLLUTION CONTROL
C                 COSTS, COAL QUALITY (E.G. COAL RANK, BTU/LBS, SULFUR
C                 CONTENT), COAL PRICE, AND TRANSPORTATION COSTS ARE
C                 ALL TAKEN INTO ACCOUNT IN THIS ANALYSIS.  THIS
C                 PROGRAM ALSO CALCULATES THE COST DIFFERENTIAL BETWEEN
C                 THE USE OF POWDER RIVER COAL AND THE LEAST COST COAL
C                 ALTERNATIVE.  A DATA FILE THAT CAN BE USED IN A
C                 QUALITATIVE RESPONSE ANALYSIS IS ALSO PRODUCED BY
C                 THIS PROGRAM.
C
C          DATE:  NOVEMBER 26, 1982
C          MODIFIED ON APRIL 26, 1983
C
      DIMENSION PLANTS( 12 , 12 ) , SULFUR( 12 , 12 )
      DIMENSION FUELS( 10 , 13 ) , TRANS( 8 , 13 )
      DIMENSION TCOSTA( 12 ) , TCOSTB( 12 )
      DIMENSION SORTA( 12 ) , SORTB( 12 )
      DIMENSION KPLREG(56)
      DATA KPLREG/3,0,0,10,6,11,0,8,1,2,2,3,3,0,0,7,4,4,5,5,4,6,1,
     1 2,1,4,5,3,5,7,5,9,1,2,10,2,3,12,4,6,11,2,0,1,3,12,3,6,9,1,2,
     2 0,11,2,4,7/
C
      WRITE( 5 , 1 )
    1 FORMAT(/,1X,'TYPE IN A FIVE LETTER CODE IDENTIFYING THIS RUN:')
      ACCEPT 5 , ALPHA
    5 FORMAT( A5 )
      WRITE( 3 , 10 ) ALPHA
   10 FORMAT(//,1X,'THIS RUN IS:',1X,A5,////)
      WRITE( 3 , 12 )
   12 FORMAT(33X,'THE TOTAL COST OF ELECTRICITY GENERATION, AS A
     1 RESULT OF',/,33X,'BURNING COAL MINED FROM THE FOLLOWING COAL
     2 SUPPLY CENTERS',/,45X,'(in 10,000,000.00 1980 dollars)',/)
      WRITE( 3 , 15 )
   15 FORMAT(2X,'ID',4X,'N.A.',5X,'C.A.',5X,'S.A.',5X,'E.C.',5X,
     1 'W.C.',5X,'S.C.',5X,'S.WY',5X,' CO ',5X,' UT ',5X,' NM ',5X,
     2 ' WA ',5X,'F.U.',5X,'NGP',4X,'COST DIFF')
C
C   INITIALIZE THE DATA MATRICES PLANTS(I,J), SULFUR(I,J), FUELS(I,J),
C   AND TRANS(I,J).  REGIONAL DATA ON POWER PLANT
C   CHARACTERISTICS ARE STORED IN PLANTS(I,J).  SULFUR DIOXIDE CONTROL
C   COSTS BY REGION AND PLANT TYPE ARE STORED IN SULFUR(I,J).  COAL
C   CHARACTERISTICS BY COAL SUPPLY REGION (INCLUDING F.O.B. PRICES ARE
C   STORED IN FUELS(I,J).  TRANSPORTATION RATES BY POWERPLANT REGION
```

```
C   ARE STORED IN TRANS(I,J).
C
        DO 20  I = 1 , 12
          DO 30  J = 1 , 12
                PLANTS( I , J ) = 0.0
                SULFUR( I , J ) = 0.0
 30       CONTINUE
 20     CONTINUE
C
        DO 33  I = 1 , 10
          DO 37  J = 1 , 13
                IF( I .GT. 8 ) GO TO 25
                TRANS( I , J ) = 0.0
 25             FUELS( I , J ) = 0.0
 37       CONTINUE
 33     CONTINUE
C
C   INITIALIZE ARRAYS TCOSTA, TCOSTB, SORTA, AND SORTB.
C   TCOSTA CONTAIN THE TOTAL COSTS OF USING NGP COAL.  TCOSTB CONTAIN
C   THE TOTAL COSTS OF USING COAL FROM OTHER NON-NGP SOURCES.  SORTA
C   AND SORTB ARE MIRROR IMAGES OF TCOSTA AND TCOSTB RESPECTIVELY.
C   SORTA AND SORTB WILL BE USED IN A ROUTINE THAT IDENTIFIES THE LEAST
C   COST ALTERNATIVE COAL SOURCE.
C
        DO 40  K = 1 , 12
          TCOSTA( K ) = 0.0
          TCOSTB( K ) = 0.0
          SORTA( K ) = 0.0
          SORTB( K ) = 0.0
 40     CONTINUE
C
C   THE REAL FIXED CHARGE RATE ON CAPITAL IS INITIALIZED HERE.  ALSO,
C   THE NUMBER OF HOURS DURING A CALANDER YEAR THAT A POWER PLANT IS
C   OPERATED AT FULL CAPACITY IS INITIALIZED.
C
        RATE = 0.07410
        TFULL = 5694.0   !ASSUMING 65% ANNUAL LOAD FACTOR
C
        WRITE(5 , 45) RATE , TFULL
 45     FORMAT( /,' THE FIXED CHARGE RATE FOR THIS STUDY IS:',F8.5,//,
     1  ' THE HOURS THAT A POWER PLANT IS OPERATED AT FULL LOAD',/,
     2  ' DURING A CALANDER YEAR IS:',F7.1,/)
C
C   OPEN INPUT AND OUTPUT DATA FILES
C
        OPEN(UNIT=20, FILE= 'COAL1.DTA')     !PLANT/MINE LOCATION FILE
        OPEN(UNIT=21, FILE='CPLA80.DTA')     !PLANT CHARACTERISTICS FILE
        OPEN(UNIT=22, FILE='CSUL80.DTA')     !SO2 CONTROL COSTS FILE
        OPEN(UNIT=23, FILE='CPRI80.DTA')     !COAL CHARACTERISTICS FILE
        OPEN(UNIT=24, FILE='CTRA80.DTA')     !TRANSPORTATION COSTS FILE
        OPEN(UNIT=01, FILE='COSDIF.DAT')     !OUTPUT FILE
C
C   READ EXTERNAL DATA FILES INTO DATA MATRICES.  THE DATA TO BE READ
C   INCLUDES POWER PLANT CHARACTERISTICS, SO2 CONTROL COSTS, COAL
```

```
C     (FUEL) CHARACTERISTICS, AND TRANSPORTATION COSTS.
C
            READ( 21 , * ) ( (PLANTS( I , J ) , J = 1 , 12) , I = 1 , 12 )
            READ( 22 , * ) ( (SULFUR( I , J ) , J = 1 , 12) , I = 1 , 12 )
            READ( 23 , * ) ( ( FUELS( I , J ) , J = 1 , 13) , I = 1 , 10 )
            READ( 24 , * ) ( ( TRANS( I , J ) , J = 1 , 13) , I = 1 ,  8 )
C
C     UNIT 20 (FILE COAL4.DTA) HAS INPUT DATA ON THE LOCATION AND SIZE OF
C     INDIVIDUAL POWER PLANT UNITS AND THE LOCATION OF THEIR MAJOR COAL
C     SUPPLY SOURCE.  DUMMY VARIABLES ON FGD USE AND PLANT GENERATION
C     CLASS (NSPS OR RNSPS PLANTS) ARE INCLUDED.  ID = POWER PLANT UNIT
C     ID NUMBER, NPLAST = POWER PLANT STATE, NPLACO = POWER PLANT COUNTY,
C     XCORP = X-COORDINATE POWER PLANT, YCORP = Y-COORDINATE POWER PLANT,
C     AMW = NAMEPLATE CAPACITY IN MEGAWATTS, MINEST = MAJOR COAL MINE
C     STATE, MINECO = MAJOR COAL MINE COUNTY, XCORM = X-COORDINATE MINE,
C     Y-COORDINATE MINE, JFGD = DUMMY ON WHETHER OR NOT THE PLANT USES
C     FLUE GAS DESULFURIZATION EQUIPMENT, JRNSPS = DUMMY ON WHETHER THE
C     POWERPLANT FACES 1971 NSPS REGULATIONS OR 1978 RNSPS REGULATIONS.
C
C
   70       READ(20,80,END=300)ID , NPLAST , NPLACO , XCORP , YCORP , AMW ,
     1      MINEST , MINECO , XCORM , YCORM , JFGD , JRNSPS
   80       FORMAT( I3 , I2 , I3 , 2F5.2 , 1X , F4.0 , 7X , I2 , I3 , 16X ,
     1      2F5.2 , 31X , I1 , 1X , I1 )
C
            AMW = 500.0     !ASSUMING A STANDARD POWER PLANT SIZE OF 500 MW.
C
C     THE FOLLOWING SECTION WILL DETERMINE WHETHER OR NOT THE POWERPLANT
C     USES NGP COAL. IF THERE ARE MISSING VALUES, READ THE NEXT CASE.
C     THE CASE IS THROWN OUT IF NO MINES WERE SPECIFIED.    ALSO,
C     IF A  POWERPLANT USES WYOMING COAL, AND IT CANNOT BE DETERMINED
C     WHETHER THE COAL CAME FROM THE NGP OR NOT, THE CASE IS THROWN OUT.
C     IF POWER PLANT SIZE AND OR COORDINATES ARE MISSING, THE OBSERVATION
C     IS DISMISSED.
C
            NOCOUN = 0 !INITIALIZE "NO COUNTY SPECIFICATION" DETECTOR
               IF(MINEST .EQ. 00) GO TO 85      !NO MINE WAS SPECIFIED
               IF(MINEST .EQ. 30) GO TO 72      !MONTANA NGP COAL
               IF(MINEST .EQ. 56) GO TO 74      !WYOMING COAL (NGP OR NOT)
               GO TO 75
C
C     IF WYOMING COAL IS USED, THE NEXT SECTION OF THIS PROGRAM WILL
C     DETERMINE IF THE COAL IS NGP WYOMING COAL.
C
   74       IF(MINECO .EQ. 005) GO TO 72    !CAMPBELL COUNTY
            IF(MINECO .EQ. 007) GO TO 72    !CARBON COUNTY
            IF(MINECO .EQ. 009) GO TO 72    !CONVERSE COUNTY
            IF(MINECO .EQ. 000) NOCOUN = 1   !NO COUNTY WAS SPECIFIED.
            IF(MINECO .NE. 000) NOCOUN = 0   !NON-NGP WYOMING COUNTY.
C
C
            IF(NOCOUN .EQ. 1) GO TO 85
C
   75       NGP = 0      !THE PLANT DOES NOT USE NGP COAL.
```

```
      GO TO 78
   72 NGP = 1      !THE PLANT DOES USE NGP COAL.
C
   78 IF(XCORP .EQ. 0.0) GO TO 85
      IF(AMW  .EQ. 0.0) GO TO 85
      GO TO 90
   85 WRITE( 3 , 86 ) ID
   86 FORMAT(1X,I3,3X,'******MISSING VALUES DETECTED, CASE DISMISSED
     1*******')
      GO TO 70
C
C     DESIGNATE THE REGION WITHIN WHICH THE POWER PLANT IS LOCATED.
C
   90 KPLANT =KPLREG(NPLAST)
C
C     THE FOLLOWING LOOP CALCULATES THE TOTAL COSTS OF USING COAL FROM
C     THIRTEEN DIFFERENT COAL CENTERS IN THE U.S.
C
      DO 150  KCOAL = 1 , 12
C
C     DETERMINE THE BASE PLANT HEAT RATE, BASE PLANT CAPITAL COST, AND
C     BASE PLANT OPERATING AND MAINTENANCE COST.  OPERATING AND MAIN-
C     TAINANCE COSTS MUST BE CONVERTED FROM MILLS TO DOLLARS.  BASE PLANT
C     CHARACTERISTICS ARE A FUNCTION OF COAL RANK AND POWER PLANT REGION.
C
      IF( FUELS(3,KCOAL) .EQ. 1.0 ) GO TO 92
      IF( FUELS(3,KCOAL) .EQ. 2.0 ) GO TO 94
      IF( FUELS(3,KCOAL) .EQ. 3.0 ) GO TO 96
      GO TO 98
C
C     BASE PLANT CHARACTERISTICS IN REGION KPLANT FOR BITUMINOUS COAL
C
   92 BPHRB = PLANTS(1,KPLANT)
      BPCAPB = PLANTS(4,KPLANT)
      ARNSPB = PLANTS(7,KPLANT)
      BPOMB = PLANTS(10,KPLANT)/1000.0
      GO TO 100
C
C     BASE PLANT CHARACTERISTICS IN REGION KPLANT FOR SUBBITUMINOUS COAL
C
   94 BPHRB = PLANTS(2,KPLANT)
      BPCAPB = PLANTS(5,KPLANT)
      ARNSPB = PLANTS(8,KPLANT)
      BPOMB = PLANTS(11,KPLANT)/1000.0
      GO TO 100
C
C     BASE PLANT CHARACTERISTICS IN REGION KPLANT FOR BITUMINOUS
C     SUBBITUMINOUS COAL BLENDS
C
   96 BPHRB = (PLANTS(1,KPLANT)+PLANTS(2,KPLANT))/2.0
      BPCAPB = (PLANTS(4,KPLANT)+PLANTS(5,KPLANT))/2.0
      ARNSPB = (PLANTS(7,KPLANT)+PLANTS(8,KPLANT))/2.0
      BPOMB = (PLANTS(10,KPLANT)+PLANTS(11,KPLANT))/2000.0
      GO TO 100
```

```
C
C   BASE PLANT CHARACTERISTICS IN REGION KPLANT FOR LIGNITE COAL
C
  98      BPHRB = PLANTS(3,KPLANT)
          BPCAPB = PLANTS(6,KPLANT)
          ARNSPB = PLANTS(9,KPLANT)
          BPOMB = PLANTS(12,KPLANT)/1000.0
C
C   BASE PLANT CHARACTERISTICS IN REGION KPLANT FOR POWER RIVER COAL
C
 100      BPHRA = PLANTS(2,KPLANT)
          BPCAPA = PLANTS(5,KPLANT)
          ARNSPA = PLANTS(8,KPLANT)
          BPOMA = PLANTS(11,KPLANT)/1000.0
C
C   DETERMINE SULFUR DIOXIDE CONTROL COSTS.  FACTORS DETERMINING SO2
C   CONTROL COSTS INCLUDE CAPITAL COSTS, OPERATING AND MAINTENANCE
C   COSTS, ENERGY PENALTIES, AND CAPACITY PENALTIES.  OPERATING AND
C   MAINTENANCE COSTS MUST BE CONVERTED FROM MILLS TO DOLLARS.  ENERGY
C   AND CAPACITY PENALTIES MUST BE CONVERTED FROM A PERCENT TO A REAL
C   NUMBER.  SO2 CONTROL COSTS ARE A FUNCTION OF COAL SULFUR CONTENT
C   AND POWER PLANT LOCATION.
C
          IF(FUELS(5,KCOAL) .GE. 2.5) GO TO 102
          IF(FUELS(5,KCOAL) .LE. 0.83) GO TO 106
          GO TO 104
C
C   SO2 CONTROL COSTS FOR REGION KPLANT WHEN THE COAL SULFUR CONTENT
C   IS GREATER THAN 2.5%
C
 102      SO2CAB = SULFUR(1,KPLANT)
          SO2OMB = SULFUR(4,KPLANT)/1000.0
          SO2EPB = SULFUR(7,KPLANT)/100.0
          SO2CPB = SULFUR(10,KPLANT)/100.0
          GO TO 110
C
C   SO2 CONTROL COSTS FOR REGION KPLANT WHEN THE COAL SULFUR CONTENT
C   IS BETWEEN 2.5% AND 0.83%
C
 104      SO2CAB = SULFUR(2,KPLANT)
          SO2OMB = SULFUR(5,KPLANT)/1000.0
          SO2EPB = SULFUR(8,KPLANT)/100.0
          SO2CPB = SULFUR(11,KPLANT)/100.0
          GO TO 110
C
C   SO2 CONTROL COSTS IN REGION KPLANT WHEN THE COAL SULFUR CONTENT IS
C   LESS THAN OR EQUAL TO 0.83%
C
 106      SO2CAB = SULFUR(3,KPLANT)
          SO2OMB = SULFUR(6,KPLANT)/1000.0
          SO2EPB = SULFUR(9,KPLANT)/100.0
          SO2CPB = SULFUR(12,KPLANT)/100.0
C
C   SO2 CONTROL COSTS IN REGION KPLANT WHEN POWDER RIVER COAL IS USED.
```

```
C
 110    SO2CAA = SULFUR(3,KPLANT)
        SO2OMA = SULFUR(6,KPLANT)/1000.0
        SO2EPA = SULFUR(9,KPLANT)/100.0
        SO2CPA = SULFUR(12,KPLANT)/100.0
C
C   DETERMINE THE TONS OF COAL NEEDED FOR THE ANNUAL OPERATION OF A
C   COAL FIRED GENERATOR.  WE WILL ALSO DETERMINE THE POWER PLANT
C   COST (PCOST) OF THE GENERATING UNIT.  COAL TONNAGE AND AND PCOST
C   ARE AFFECTED BY FEDERAL NSPS LAWS AND FGD USE.
C
        IF(JPNSPS .EQ. 1) GO TO 140
        IF(JFGD .EQ. 1) GO TO 130
C
C   HEAT RATE FACTORS, CAPITAL COSTS AND OPERATING/MAINTENANCE COSTS
C   FOR NSPS PLANTS NOT USING FGD
C
        IF((NGP .EQ. 1) .AND. (FUELS(5,KCOAL) .GT. 0.83)) GO TO 120
        GO TO 122
C
C   IF A PLANT USES POWDER RIVER COAL WITHOUT FGD, IT IS ASSUMED THAT
C   COALS OF HIGHER SULFER CONTENT MUST USE FGD.  IF THE SULFUR CONTENT
C   OF NON-POWDER RIVER COAL IS LESS THAN 0.83%, NO SCRUBBERS ARE
C   NEEDED.
C
 120    HRFB = AMW*TFULL*(BPHRB*(1.0+SO2EPB))*1000.0
        CAPB = (BPCAPB+SO2CAB)*(1.0+SO2CPB)*AMW*RATE*1000.0
        OMB  = (BPOMB+SO2OMB)*AMW*TFULL*1000.0
        GO TO 124
C
 122    HRFB = AMW*TFULL*BPHRB*1000.0
        CAPB = BPCAPB*AMW*RATE*1000.0
        OMB  = BPOMB*AMW*TFULL*1000.0
C
 124    HRFA = AMW*TFULL*BPHRA*1000.0
        CAPA = BPCAPA*AMW*RATE*1000.0
        OMA  = BPOMA*AMW*TFULL*1000.0
        GO TO 145
C
C   HEAT RATE FACTORS, CAPITAL COSTS, AND OPERATING/MAINTENANCE COSTS
C   FOR NSPS PLANTS USING FGD.
C
 130    IF((NGP .EQ. 0) .AND. (FUELS(5,KCOAL) .GT. 0.83)) GO TO 132
        GO TO 134
C
C   IF POWDER RIVER COAL IS NOT BEING USED BY THE POWER PLANT, AND THE
C   SULFUR CONTENT OF THE COAL BEING USED IS ABOVE 0.83%, IT IS ASSUMED
C   THAT POWDER RIVER COAL DOES NOT HAVE TO BE SCRUBBED.  OTHERWISE,
C   POWDER RIVER COAL MUST BE SCRUBBED.
C
 132    HRFA = AMW*TFULL*BPHRA*1000.0
        CAPA = BPCAPA*AMW*RATE*1000.0
        OMA  = BPOMA*AMW*TFULL*1000.0
        GO TO 136
```

```
C
 134     HRFA = AMW*TFULL*(BPHRA*(1.0+SO2EPA))*1000.0
         CAPA = (BPCAPA+SO2CAA)*(1.0+SO2CPA)*AMW*RATE*1000.0
         OMA  = (BPOMA+SO2OMA)*AMW*TFULL*1000.0
C
 136     HRFB = AMW*TFULL*(BPHRB*(1.0+SO2EPB))*1000.0
         CAPB = (BPCAPB+SO2CAB)*(1.0+SO2CPB)*AMW*RATE*1000.0
         OMB  = (BPOMB+SO2OMB)*AMW*TFULL*1000.0
         GO TO 145
C
C   HEAT RATE FACTORS, CAPITAL COSTS, AND OPERATING/MAINTENANCE COSTS
C   FOR RNSPS PLANTS USING FGD.
C
 140     HRFB = AMW*TFULL*(BPHRB*(1.0+SO2EPB))*1000.0
         CAPB = (BPCAPB+ARNSPB+SO2CAB)*(1.0+SO2CPB)*AMW*RATE*1000.0
         OMB  = (BPOMB+SO2OMB)*AMW*TFULL*1000.0
C
         HRFA = AMW*TFULL*(BPHRA*(1.0+SO2EPA))*1000.0
         CAPA = (BPCAPA+ARNSPA+SO2CAA)*(1.0+SO2CPA)*AMW*RATE*1000.0
         OMA  = (BPOMA+SO2OMA)*AMW*TFULL*1000.0
C
C   TOTAL COAL TONNAGE AND TOTAL POWER PLANT COSTS FOR COAL SOURCE A
C   (NGP) AND COAL SOURCE B (NONNGP) ARE CALCULATED.
C
 145     TONSA = HRFA/(FUELS(4,13) * 2000.0)
         TONSB = HRFB/(FUELS(4,KCOAL) * 2000.0)
C
         PCOSTA = CAPA + OMA
         PCOSTB = CAPB + OMB
C
C   THE RAW DIGITIZER COORDINATES CONTAINED IN DATA FILE COAL4.DTA
C   CANNOT BE USED DIRECTLY IN THE CALCULATION OF STRAIGHT LINE
C   DISTANCES BETWEEN POWERPLANT AND COAL SOURCE.  TO GET THE PROPER
C   TRANSFORMED COORDINATES IN MILES WITH THE X AXIS RUNNING EAST-WEST
C   AND THE Y AXIS RUNNING NORTH-SOUTH AND WITH THE ORIGIN CENTERED
C   ON GILLETTE WYOMING, THE FOLLOWING TRANSFORMATION IS PERFORMED.
C
         XPLANT  =  ( YCORP - 13.14 ) * 80.0
         YPLANT  =  ( XCORP -  6.74 ) * -80.0
         XMINE   =  ( YCORM - 13.14 ) * 80.0
         YMINE   =  ( XCORM -  6.74 ) * -80.0
         XLOCAT  =  ( YCORP )
         YLOCAT  =  ( XCORP * -1.0 )
C
C   THE STRAIGHT LINE DISTANCE BETWEEN THE COAL SUPPLY AND
C   THE POWER PLANT IS CALCULATED.  ACTUAL MINE COORDINATES ARE
C   USED WHENEVER POSSIBLE.
C
         XGILLE = 0.0
         YGILLE = 0.0
         IF( NGP .EQ. 1 ) GO TO 146
         GO TO 147
 146     XGILLE = XMINE
         YGILLE = YMINE
```

```
 147    DISTA = SQRT((XGILLE-XPLANT)2 + (YGILLE-YPLANT)2)
C
        XSUPPL = FUELS(8,KCOAL)
        YSUPPL = FUELS(9,KCOAL)
 149    DISTB = SQRT((XSUPPL-XPLANT)2 + (YSUPPL-YPLANT)2)
C
C   TRANSPORTATION COSTS WILL NOW BE CALCULATED.  REGION OF ORIGIN AND
C   REGION OF DESTINATION AFFECTS TRANSPORTATION COSTS AS WELL AS
C   TONNAGE AND DISTANCE.
C
        KTRANS = 1
        IF( FUELS( 10 , KCOAL ) .EQ. 2.0 ) KTRANS = 3
        IF( FUELS( 10 , KCOAL ) .EQ. 3.0 ) KTRANS = 5
        IF( FUELS( 10 , KCOAL ) .EQ. 4.0 ) KTRANS = 7
C
C   IF POWER PLANTS ARE LOCATED ON THE ROCKY MOUNTAIN FRONT, UNIT TRAINS
C   FROM THE EAST DO NOT HAVE TO CROSS THE MOUNTAINS WHILE TRAINS FROM
C   THE WEST MUST CROSS THE MOUNTAINS.
C
        IF((NGP .EQ. 1) .AND. ((KPLANT .EQ. 7) .OR. (KPLANT .EQ. 8)))
     1 GO TO 500
        GO TO 600
 500    FIXTRA = TRANS( 5 , 13 ) * TONSA
        FIXTRB = TRANS( KTRANS , 13 ) * TONSB
C
        VARTRA = TRANS( 6 , 13 ) * DISTA * TONSA
        VARTRB = TRANS((KTRANS+1) , 13 ) * DISTB * TONSB
        GO TO 700
C
C   FIXED TRANSPORTATION COSTS ARE CALCULATED.
C
 600    FIXTRA = TRANS( 5 , KPLANT) * TONSA
        FIXTRB = TRANS(KTRANS , KPLANT) * TONSB
C
C   VARIABLE TRANSPORTATION COSTS ARE CALCULATED.
C
        VARTRA = TRANS( 6 , KPLANT) * DISTA * TONSA
        VARTRB = TRANS(( KTRANS+1 ) , KPLANT) * DISTB * TONSB
C
C   PRODUCTION COSTS (COAL PRICE PER TON TIMES TONS) ARE COMPUTED.
C
 700    APROD = FUELS(2 , 13) * TONSA
        BPROD = FUELS(2 , KCOAL) * TONSB
C
C   FCOST = THE FUEL COSTS (THE SUM OF ANNUAL FUEL COSTS PLUS THE COST
C   OF TRANSPORTING THAT FUEL TO THE PLANT) IS COMPUTED.
C
        FCOSTA = APROD + FIXTRA + VARTRA
        FCOSTB = BPROD + FIXTRB + VARTRB
C
C   CALCULATE THE TOTAL COST OF USING COAL FROM SOURCE A (NGP) AND
C   SOURCE B (NON-NGP).  THE TOTAL COST WILL BE DIVIDED BY 1000000.0.
C
        TCOSTA(KCOAL) = (FCOSTA + PCOSTA)/1000000.0
```

```
      TCOSTB(KCOAL) = (FCOSTB + PCOSTB)/1000000.0
      SORTA(KCOAL) = TCOSTA(KCOAL)
      SORTB(KCOAL) = TCOSTB(KCOAL)
C
C   PICK UP THE NEXT KCOAL (COAL FIELD CENTER).
C
 150  CONTINUE
C
C   SORT TO FIND THE LEAST COST NON-NGP COAL SOURCE.  THE LEAST COST
C   NON-NGP COAL SOURCE WILL BE STORED IN SORTB( 1 ).
C
      DO 160  J = 2 , 12
        TEMP1 = SORTB( J )
        TEMP2 = SORTA( J )
          DO 170  K = J-1 , 1 , -1
              IF(SORTB(K) .LE. TEMP1) GO TO 180
              SORTB(K + 1) = SORTB(K)
              SORTA(K + 1) = SORTA(K)
 170      CONTINUE
      K = 0
 180  SORTB(K + 1) = TEMP1
      SORTA(K + 1) = TEMP2
 160  CONTINUE
C
C   CALCULATE THE COST DIFFERENTIAL BETWEEN THE LEAST COST NON-NGP
C   COAL SOURCE AND THE COST OF USING NGP COAL.
C
      CD = SORTB( 1 )  -  SORTA( 1 )
C
C   OUTPUT THE TOTAL COST INFORMATION AND THE COST DIFFERENTIAL FOR
C   EACH POWER PLANT UNIT.
C
 999  WRITE(01,190) ID, XLOCAT, YLOCAT, NGP, CD, JFGD, JRNSPS
 190  FORMAT(1X , I3 , 1X , 2F7.2 , 1X , I1 , 1X , F12.7 , 1X , I1 ,
     1  1X , I1)
      WRITE( 3 , 200 ) ID , (TCOSTB(K), K=1,12), SORTA(1), CD
 200  FORMAT(1X , I3 , 14F9.4)
C
C   GO BACK AND DO ANOTHER POWER PLANT.
C
      GO TO 70
C
 300  END
```

APPENDIX B

The BMDP Control Program That Was Used to Estimate the Logit Model Parameters

```
/PROBLEM    TITLE IS 'CURRENT 1980 LOGIT'.
/INPUT      VARIABLES ARE 7.
            FORMAT IS '(1X, I3, 1X, 2F7.2, 1X, I1, 1X, F12.7, 1X, I1,
            1X, I1)'.
            UNIT = 25.
/VARIABLE   NAMES ARE ID, XCORP, YCORP, COAL, COSTD, FGD, RNSPS.
            USE = 4 , 5 , 6 , 7.
/GROUP      CODES(4) = 1 , 0.
            NAMES(4) = PRONGP , NONNGP.
/REGRESS    DEPEND IS COAL.
            INTERVAL IS COSTD.
            CATEGORICAL IS FGD , RNSPS.
            MODEL = COSTD*FGD , COSTD*RNSPS.
            METHOD = MLR.
            ITER = 100.
/END
```

APPENDIX C

A Fortran Computer Program that Calculates Statistics which Describe the "Goodness of Fit" of the Estimated Logit Model

```
C     PROGRAMMER:    HENRY Y. YOSHIMURA
                     DEPARTMENT OF ECONOMICS
                     UNIVERSITY OF MONTANA
C
C     PROGRAM:      THIS PROGRAM CALCULATES STATISTICS DESCRIBING THE FIT
C                   OF THE LOGIT MODEL TO THE OBSERVED DATA.  THE USER
C                   MUST RUN THE PROGRAM TCOSTS.FOR (THE PROGRAM THAT
C                   CALCULATES COST DIFFERENTIALS) PRIOR TO RUNNING THIS
C                   PROGRAM.  THE USER ALSO MUST INPUT THE ESTIMATED
C                   LOGIT MODEL COEFFICIENTS IN ORDER TO USE THIS
C                   PROGRAM.
C
C     DATE:         JANUARY 20, 1983
C
         DIMENSION  COAL2 (500)
C
C     OPEN INPUT AND OUTPUT DATA FILES.  COSDIF.DAT CONTAINS THE
C     CALCULATED COST DIFFERENTIALS, FGD DUMMYS, AND AIR POLLUTION
C     POLICY DUMMYS FOR EACH POWER PLANT OBSERVATION.  PRO*.DAT
C     CONTAINS THE COORDINATES FOR POWER PLANTS WITH A ESTIMATED
C     PROBABILITY OF BUYING POWDER RIVER COAL BETWEEN 0% AND 20%,
C     +20% AND 40%, +40% AND 60%, +60% AND 80%, AND +80% AND 100%
C     RESPECTIVELY.
C
         OPEN(UNIT = 51, FILE = 'COSDIF.DAT')
         OPEN(UNIT = 20 , FILE ='PRO1.DAT')
         OPEN(UNIT = 21 , FILE ='PRO2.DAT')
         OPEN(UNIT = 22 , FILE ='PRO3.DAT')
         OPEN(UNIT = 23 , FILE ='PRO4.DAT')
         OPEN(UNIT = 24 , FILE ='PRO5.DAT')
C
         WRITE( 5 , 1 )
    1    FORMAT(1X , 'ENTER A FIVE CHARACTER IDENTIFYING CODE
     1 FOR THIS RUN',/)
         ACCEPT 2, ALPHAB
    2    FORMAT (A5)
C
         WRITE( 5 , 3 )
    3    FORMAT(/, 1X , 'ENTER THE ESTIMATED COEFFICIENTS FOR THE LOGIT
     1 MODEL',/,' IN THE FOLLOWING ORDER: B0  B1  B2  B3  B4  B5 ',
     2 /,' WHERE: ',
     3 /,' B0 = THE CONSTANT COEFFICIENT ',
     4 /,' B1 = THE COEFFICIENT ON THE COST DIFFERENTIAL C1 ',
     5 /,' B2 = THE COEFFICIENT ON THE FGD DUMMY F1 ',
     6 /,' B3 = THE COEFFICIENT ON THE AIR POLLUTION NSPS DUMMY A1 ',
     7 /,' B4 = THE COEFFICIENT ON THE INTERACTION TERM C1F1 ',
     8 /,' B5 = THE COEFFICIENT ON THE INTERACTION TERM C1A1 ',
     9 /,' ENTER THE NUMBERS ON THE SAME LINE SEPARATED BY A SPACE.',/)
         READ( 5 , * ) B0 , B1 , B2 , B3 , B4 , B5
C
         WRITE( 50 , 4 ) ALPHAB
    4    FORMAT( 1X , 'THE NAME OF THIS RUN IS:   ', A5 , ////)
         WRITE( 50 , 5 ) B0 , B1 , B2 , B3 , B4 , B5
```

```
  5      FORMAT(' THE ESTIMATED LOGIT COEFFICIENTS ARE:',/,' B0 = ',F7.4,
     1 '  B1 = ',F7.4,'  B2 = ',F7.4,'  B3 = ',F7.4 ,'  B4 = ',F7.4,
     2 '  B5 = ',F7.4,///)
C
         WRITE( 59 , 100 )
 100     FORMAT(1X,'PLANT ID        OBSERVED Y     PREDICTED Y
     1 PROBABILITY P    PRED LOG ODDS     RESIDUAL     SQUARED RESIDUAL
     2 RIGHT=1,WRONG=0')
C
C    INITIALIZE ALL COUNTER VALUES.  SWRONG = THE NUMBER OF WRONG
C    PREDICTIONS.  SRIGHT = THE NUMBER OF CORRECT PREDICTIONS.
C    SSRES = SUM OF SQUARED RESIDUALS.  SUMY = SUM OF Y1 VALUES.
C    OBSERV = NUMBER OF OBSERVATIONS.  TSS = TOTAL SUM OF SQUARES.
C    KOUNT = NUMBER OF OBSERVATIONS <INTEGER VALUE>.
C
         SWRONG = 0.0
         SRIGHT = 0.0
         SSRES  = 0.0
         SUMY   = 0.0
         OBSERV = 0.0
         TSS    = 0.0
         KOUNT  =   0
C
C    READ INPUT DATA FILE COSDIF.DAT.
C
  6      READ(01,10,END=30)IDNUM,XCORP,YCORP,COAL1,COSDIF,FGD,RNSPS
  10     FORMAT(1X, I3, 1X, 2F7.2, F2.0, 1X, F12.7, F2.0, F2.0)
C
C    CALCULATE THE PREDICTED PROBABILITY OF POWER PLANT 1 BUYING POWDER
C    RIVER COAL.
C
         ZINDEX = 2.718281828( B0+(B1*COSDIF)+(B2*FGD)+(B3*RNSPS)+
     1 (B4*COSDIF*FGD)+(B5*COSDIF*RNSPS) )
         PROB   =  ZINDEX / (1.0 + ZINDEX)
C
C    DETERMINE IF THIS PREDICTION IS CORRECT OR INCORRECT GIVEN THE
C    POWER PLANT'S ACTUAL COAL CHOICE.  UPDATE COUNTER VALUES ON THE
C    NUMBER OF WRONG AND CORRECT PREDICTIONS.
C
         PRED = 0.0
         IF( PROB .GE. 0.5 ) PRED = 1.0
         WRONG = ( COAL1 - PRED )2.0
         SWRONG = SWRONG + WRONG
         RIGHT = 0.0
         IF( WRONG .EQ. 0.0 ) RIGHT = 1.0
         SRIGHT = SRIGHT + RIGHT
C
C    CALCULATE THE ODDS OF A POWER PLANT BUYING POWDER RIVER COAL.
C
         ODDS  =  ALOG( PROB / ( 1.0 - PROB ) )
C
C    CALCULATE THE RESIDUAL COMPONENT.  UPDATE THE SUM OF SQUARED
C    RESIDUALS COUNTER.
C
```

```
      RESID = COAL1 - PROB
      SRES  = RESID2.0
      SSRES = SSRES + SRES
C
C   UPDATE THE COUNTERS FOR THE NUMBER OF OBSERVATIONS AND THE SUM OF
C   Y1 VALUES.
C
      OBSERV= OBSERV + 1.0
      KOUNT  =  KOUNT + 1
      SUMY   = SUMY + COAL1
      COAL2(KOUNT) = COAL1
C
C   OUTPUT THE CALCULATED VALUES ON POWER PLANT 1.
C
      WRITE( 50 , 20 )IDNUM,COAL1,PRED,PROB,ODDS,RESID,SRES,RIGHT
   20 FORMAT(3X,I4,13X,F2.0,13X,F2.0,10X,F6.4,10X,F9.4,6X,F8.4,8X,
     1 F8.6,15X,F2.0)
C
C   OUTPUT PLANT COORDINATES WHOSE CALCULATED PROBABILITY FALLS IN
C   PARTICULAR RANGES.
C
      XCORP = XCORP * 100.0
      YCORP = YCORP * 100.0
      IF( PROB .LE. 0.20 ) GO TO 200
      IF((PROB .GT. 0.20) .AND. (PROB .LE. 0.40)) GO TO 300
      IF((PROB .GT. 0.40) .AND. (PROB .LE. 0.60)) GO TO 400
      IF((PROB .GT. 0.60) .AND. (PROB .LE. 0.80)) GO TO 500
      GO TO 600
C
 200  WRITE( 20 , * ) XCORP , YCORP
      GO TO 6
 300  WRITE( 21 , * ) XCORP , YCORP
      GO TO 6
 400  WRITE( 22 , * ) XCORP , YCORP
      GO TO 6
 500  WRITE( 23 , * ) XCORP , YCORP
      GO TO 6
 600  WRITE( 24 , * ) XCORP , YCORP
      GO TO 6
C
C   CALCULATE THE TOTAL SUM OF SQUARES WITH THE FOLLOWING LOOP.
C
   30 YMEAN  =  SUMY / OBSERV
      DO  40  I = 1 , KOUNT
              TOTSQU = (COAL2(I) - YMEAN)2.0
              TSS    = TSS + TOTSQU
   40 CONTINUE
C
C   CALCULATE EFFRON'S R-SQUARED AND THE PROPORTION OF RIGHT AND WRONG
C   PREDICTIONS.
C
      RSQUAR  =  1.0 - (SSRES / TSS)
      PROPR   =  SRIGHT/OBSERV
      PROPW   =  SWRONG/OBSERV
```

```
C
C   OUTPUT ALL CALCULATED STATISTICS.
C
      WRITE( 50 , 50 ) OBSERV , SRIGHT , PROPR , SWRONG , PROPW ,
     1   SUMY , YMEAN , SSRES , TSS , RSQUAR
   50 FORMAT( ///,' NUMBER OF OBSERVATIONS  = ',F4.0,//,
     1 ' NUMBER OF RIGHT PREDICTIONS  = ',F4.0,/,
     2 ' PROPORTION RIGHT             = ',F12.7,//,
     3 ' NUMBER OF WRONG PREDICTIONS  = ',F4.0,/,
     4 ' PROPORTION WRONG             = ',F12.7,///,
     5 ' SUM OF OBSERVED Y        = ',F12.7,/,
     6 ' THE MEAN OF Y            = ',F12.7,/,
     7 ' SUM OF SQUARED RESIDUALS = ',F12.7,/,
     8 ' TOTAL SUM OF SQUARES     = ',F12.7,/,
     9 ' EFFRON R SQUARED         = ',F12.7)
C
      WRITE(5 , 60) RSQUAR
   60 FORMAT(1X , 'THE CALCULATED EFFRON R SQUARED =  ',F12.7)
C
      END
```

APPENDIX D

Logit Model Estimation Where Fi (the FGD dummy) is Excluded From the Analysis

---

Estimated Equation:

$$\ln \frac{P_i}{1 - P_i} = \underset{(.148)}{-0.372^*} + \underset{(.045)}{0.368^*}\, C_i - \underset{(.148)}{0.220}\, A_i + \underset{(.045)}{0.089^*}\, C_i A_i$$

* = significant at the 95% confidence level.

$X^2_{(3)} = 215.584$, Probability $(X^2_{(3)} > 215.584) = 0.0000$

Mc Fadden's R - squared = 0.4215

Effron's R - squared = 0.4628

Proportion of correct predictions = 0.8443

Equation for NSPS plants:

$$\ln \frac{P_i}{1 - P_i} = -0.372 = 0.368\, C_i$$

Equation for RNSPS plants:

$$\ln \frac{P_i}{1 - P_i} = -0.592 + 0.457\, C_i$$

By excluding the FGD dummy variable from the logit regression analysis, the estimated equation changes significantly. In relation to the estimated equation on table 5.7 (p. 119), when Fi is omitted from the analysis, we find that:

1) The sign on both the constant coefficient and the coefficient on Ai changes.

2) The coefficient on Ci is reduced by 20%.

3) The coefficient on the interaction term CiAi becomes significant at the 95% confidence level.

4) The calculated chi-squared is reduced by 30%.

5) McFadden's R-squared drops from 61% to 42%.

6) Effron's R-squared drops from 63% to 46%.

By omitting Fi from the logit equation, the estimated parameters of the model change and the estimated equation does not fit the observed data as well when compared to the analysis summarized in table 5.7.

When comparing the final estimated equations outlined on table 5.7 with the corresponding equations in this appendix, three comments can be made. First, the estimated equations for RNSPS plants both in table 5.7 and in this appendix are practically identical. This is to be expected however. Since all RNSPS plants must use FGD, the value of Fi affects only NSPS plants; Fi is a constant for all RNSPS plants and is a variable in the context of NSPS plants. Since Fi is an irrelevant variable in the context of RNSPS plants, we expect that the estimated RNSPS equation in this appendix (where Fi is omitted from the analysis) will be the same as the estimated RNSPS equation in table 5.7.

Second, the difference in the estimated equations between the analysis where the FGD dummy variable is included and the analysis where it is excluded is due exclusively to the way NSPS plants are modelled. The drop in explanatory power of the model in this appendix can be linked to the high unexplained variation in the dependent variable within the context of NSPS plants. The analysis in table 5.7 reduced this unexplained variation by introducing the FGD dummy variable; the inclusion of this dummy variable resulted in the fitting of two regression lines on NSPS plants as opposed to just one regression line. The two lines allowed more of the variation to be "explained." When the FGD dummy variable is excluded from the model specification, the result is a estimated model with lower predictive accuracy. As explained in chapters one and two of this study, the period of time through which NSPS plants came into existence was politically and economically unstable in terms of energy production. The "energy crises" of the 1970's, the instability of energy production costs and prices, the emerging federal energy independence programs subsidizing domestic energy research and production, and the anticipation of changing federal air quality regulations all occured during the early and middle 1970's. Given this unstable climate, we can expect high unexplained variation in the context of power plants built in this period.

Third, for NSPS plants, the effect of the cost differential on coal choice is less dramatic when Fi is excluded from the logit model. Compared to a model where Fi is included in the model specification, it takes a greater change in the cost differential to induce a particular change in the probability a NSPS power plant will buy Powder

River coal when the FGD dummy is excluded from the model specification. When Fi is not included in the logit model, high unexplained variability in the dependent variable seems to cause a downward shift in the coefficient on the cost differential; the model becomes more conservative in detecting changes in the choice probability given a change in the cost differential.

BIBLIOGRAPHY

Amemiya, T. "Qualitative Response Models: A Survey" Journal of Economic Literature. (December 1981, pp. 1483-1536).

Campbell, T. C. and Hwang, M. "Spatial Analysis and the Identification of Coal Markets" The Journal of Energy and Development. (Autumn 1978, Vol IV, No. 1, pp. 104-125).

Chambers, E. A. and Cox, D. R. "Discrimination Between Alternative Binary Response Models," Biometrika. (1967, Vol. 54(3-4), pp. 573-578).

Coal Week. "Current Contract and Spot Marker Steam Coal Prices." New York: McGraw-Hill. 1980-1981 Various Issues.

Davenport, D. Personal communication, The Montana Power Company. 1981.

Dixon, W. J. et al. BMDP Statistical Software 1981. Berkeley: University of California Press. 1981.

Duffield, J. et al. Projections of Coal Demand From the Northern Great Plains Through the Year 2010. (Final Report, OSM, May 1982).

Effron, B. "Regression and ANOVA with Zero-One Data: Measures of Residual Variation," J. Amer. Statist. Assoc.. (March 1978, pp. 113-121).

Electric Power Research Institute. CRA/EPRI Coal Market Analysis System Volume 2: Technical Background. Prepared by Charles River Associates Inc. EPRI EA-907. Palo Alto, CA July 1979.

Electric Power Research Institute. Technical Assessment Guide. EPRI PS-1201-SR. Palo Alto, CA July 1979.

Fetter, F. "The Economic Law of Market Areas". Quarterly Journal of Economics. (May 1924, pp. 520-529).

Green, John W. Coal Power Plant Scenario 1976-2000. Natural Resource Economics Division; Economics, Statistics, and Cooperatives Service; U.S. Department of Agriculture. February 1982.

Hartman, R. S. A Generalized Logit Formulation of Individual Choice. (MIT Energy Laboratory Working Paper No. MIT-EL-79-010WP, February 1979).

Hyson, C. D. and Hyson, W. P. "The Economic Law of Market Areas" Quarterly Journal of Economics. (May 1950, pp. 319-327).

ICF Inc. The Potential Energy and Economic Impacts of Coal Slurry Pipelines. Final Report. Prepared for: DOE. Washington, D.C. January 1980.

ICF Inc. The Economic and Financial Impacts of Environmental Regulations on the Electric Utility Industry. Prepared for the Edison Electric Institute. February 1980.

ICF Inc. Coal and Electric Utilities Model Documentation (Draft). Washington, D.C. May 1980.

ICF Inc. Preliminary Evaluation of Economics of NSPS Offsets for Utility Powerplants. Prepared for: Office of Policy and Planning, EPA. Washington, D.C. June 1980.

ICF Incorporated, Forecasts and Sensitivity Analysis of Western Coal Production. November, 1980.

Judge, G. et al. The Theory and Practice of Econometrics. New York: John Wiley and Sons, 1980.

Judge, G. et al. Introduction to the Theory and Practice of Econometrics. New York: John Wiley and Sons, 1982.

Keystone. Keystone Coal Industry Manual. New York: McGraw-Hill. 1980.

Kidder, Peabody and Company Inc. Electric Utility Generating Equipment Status Report on Fossil Boilers. March 1981.

Komanoff, Charles. Power Plant Cost Escalation: Nuclear and Coal Capital Costs, Regulation and Economics. Komanoff Energy Associates, New York. 1981.

Koutsoyiannis, A. Theory of Econometrics. Hong Kong: The MacMillan Press LTD, 1979.

McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behavior," in Frontiers in Econometrics. Edited by P. Zarembka. New York: Academic Press, 1974, pp.105-142.

McFadden, D. Structural Analysis of Discrete Data with Econometric Applications. Edited by C. F. Manski and D. McFadden. Cambridge, Mass.: MIT Press, 1981.

Morrison, D. G. "Upper Bounds for Correlations Between Binary Outcomes and Probabilistic Predictions," J. Amer. Statist. Assoc., March 1972, pp. 68-70.

Pindyck, R. and Rubinfeld, D. Econometric Models and Economic Forecasts. New York: McGraw Hill, 1981.

Power, T. et al. Montana University Coal Demand Study: Projections of Northern Great Plains Coal Mining and Energy Conversion Development. (Final Report, NSF-RANN, May 1976).

Rubinfeld, D. "Voting in a Local School Election: A Micro Analysis," Review of Economics and Statistics. (February 1977, Vol. 59, No. 1, pp. 30-42).

Snedecor, G. and Cochran, W. Statistical Methods. Seventh Edition. Ames, Iowa: The Iowa State University Press. 1980.

U.S. Department of Energy Service Report: Preliminary 1985, 1990, and 1995 Energy Forecasts for the Annual Report to Congress 1980. Draft. SR/IA/80-16. Washington D.C. November 1980.

U.S. Department of Energy. The 1980 Biennial Update of National and Regional Production Goals for 1985, 1990, and 1995. (March, 1981).

U.S. Environmental Protection Agency. Materials obtained from and communications with Forrest Reinhardt. Office of Policy and Planning, EPA. Washington, D.C. 1983.

Watson, W. D., Jr. "Cost of Air Pollution Control in the Coal-fired Electric Power Industry" Quarterly Review of Economics and Business. (Autumn 1972, Vol 12, No. 3, pp. 63-85).