GIS in Coal Transportation Modeling: Case Study of Ohio

Hanming Tu* and Jean-Michel Guldmann[§]

*Premier Research Worldwide, Ltd., 201 Burk Avenue, Ridley Park, PA 19078
[§]The Ohio State University, 190 W 17th Avenue, Columbus, Ohio 43210, USA

Abstract

In order to reduce SO_2 emissions to less than 2.5 lbs/mBtu as mandated by the Clean Air Act Amendment (CAAA) of 1990, power companies mainly using higher sulfur content coal as fuel supplies have to find their alternatives to reduce SO_2 emissions. The purpose of the study is to assess the extent of economies of scale in coal delivery. The network of coal flow consists of production sites, consumption plants, routes, and costs of coal delivery. GIS as a new tool is used to help identifying and visualizing these routes. Different transportation modes are evaluated under different price schemes. The translog function with price homogeneity restrictions is used to assess the cost structure of the coal delivery system.

I. INTRODUCTION

Since the Arab Oil Embargo started the first round of world oil price increases in 1973, the expectation for coal as an economic and secure substitute for oil has been very high. The Persian Gulf crisis in 1990-1991 once again demonstrated the important role of the Middle East in determining the world's future oil supplies and prices. The uncertain oil supply and unstable oil prices have drawn U.S. industries' and utilities' attention to international and domestic coal markets, and these industries have started switching back to coal from petroleum (EIA, 1992, 1993, 1993b, 1993c). The increase in coal use by electric utilities and other industries has brought environmental issues to the forefront, and has come along with more stringent environmental policies. Fuel market changes and environmental regulations of sulfur emissions had tremendous impacts on coal production and consumption.

Problem statement

Energy is not only an important factor for socio-economic development, but also an environmentally sensitive industry. Energy-environmental issues have important interactions. Therefore, fuel production and consumption are not only guided by social demand and economic principles, but also controlled by environmental policies. Especially, the coal industry has very important impacts on the environment, and its transportation costs are crucial to fuel consumption decision-making and the local economy.

Generally, an electric utility company is constrained to satisfy both energy demands and environmental concerns. Ohio power plants face twofold pressures: increasing energy demand and decreasing SO_2 emission allowance (Figure 1). Burning more Ohio high-sulfur-content coal may satisfy the energy demand with less coal delivery cost and more benefits for the local mining industry, but it would increase air pollution emissions, making it very difficult for companies to lower their sulfur dioxide emission levels to the range mandated by the CAAA.

What are the choices for power companies to fulfill energy demand and to satisfy environmental concerns? How will power companies react to the price changes of a specific fuel and changes in sulfur emission control?

Research background

Solutions for the problems stated in the previous section must be efficient, i.e., they must represent minimum cost networks of production, distribution and utilization for fuels in power plants. Four alternative approaches can be analyzed and compared for each company (Figure 2).

The first approach is the *Fuel Type Alternative (FTA)*, also called *Interfuel Substitution*. The powers from oil, natural gas, and nuclear powers generate very low or no sulfur content. Oil and natural gas are very practical alternative fuels for coals, but oil and gas are highly dependent on international market.

The second approach is the *Clean Coal Technology (CCT)*. Clean coal technologies can be used to convert "dirty coal" into clean coal. Sulfur is found in coal in two forms, organic and pyretic. Mechanical cleaning and crushing can reduce most pyretic sulfur. Organic sulfur can only be converted to synthetic forms (S or H_2SO_4) and separated from coal. The clean coal technologies are expensive and can only reduce sulfur content to a limited degree.

The third approach is the *Emission Quota Agreement (EQA)*. Instead of point-by-point emission control, the EPA has extended its trading emission reduction to a "bubble" concept. This approach gives more flexibility to companies to minimize their cost and to satisfy the two constraints: energy demand and SO_2 emission quota. This "bubble" policy involves source and geographical identification of emissions.

The fourth approach is the *Regional Coal Substitution (RCS)*.
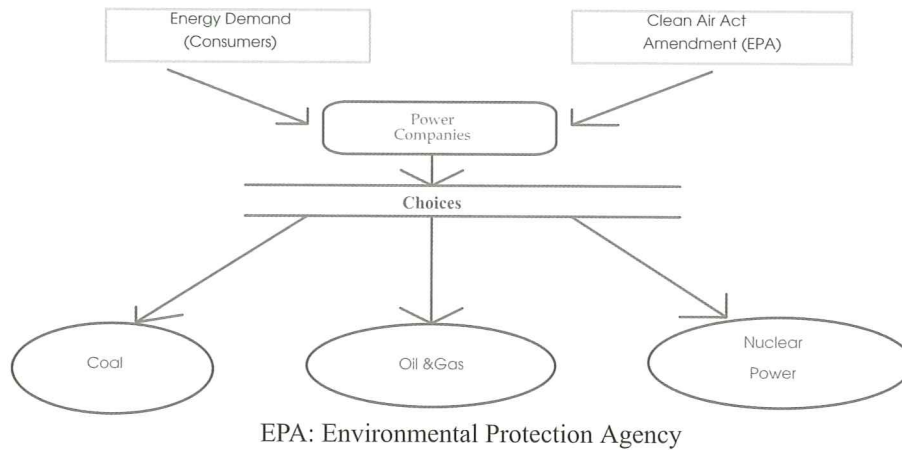
EPA: Environmental Protection Agency

Figure 1. Choices of Power Companies under Two Constraints

The sulfur content of coal differs from one region to another. When buying clean coal from other regions, power plants need to consider coal transportation costs.

In reality, one or more approaches are used in companies. Our discussion will mainly focus on the regional coal substitution.

Goal of the study

The main objective of this research is to analyze the role of transportation costs in the coal economy, under conditions of energy demand growth and the constraints of the 1990 CAAA. The basic components of the study are as follows:

1). *Coal Transportation Cost Analysis*: The delivered price of coal is the sum of the mine-mouth price, the transport rate, and the tax rate. Coal production cost (mine-mouth price) varies from mine to mine, depending on mine seam position, coal quality, and labor and operation costs, and is impacted by the depletion of coal reserves, coal policies and environmental regulations. We will analyze the share of transport costs in the delivered price when shipping coal from various regions to Ohio through various modes, and econometric models explaining the variation in total and transportation costs will

be estimated.

2). *Alternative Policies* The results of the previous analyses should suggest some alternatives available to Ohio electric utilities for achieving both goals: the SO_2 emission goals under the 1990 CAAA and energy demand. It should also clarify the choices among substituting oil or natural gas for coal, transporting clean coal from the West, and using clean coal technologies.

II. COAL TRANSPORTATION MODELLING

Transportation costs depends on factors such as distance, quantity, and transportation modes. What kind of relationship does exist among these factors and transportation costs? How strongly are these factors related to transportation costs? In order to answer these questions, many models have been built, and various analyses have been done. The major models in coal transportation cost analysis include Larwood and Benson Model (Bernknopf, 1985), Anderson Model (Transportation Service Center, 1976), Bechtel Model (Nagarvala, 1976), Zimmerman Model (Martin Zimmerman, 1977).

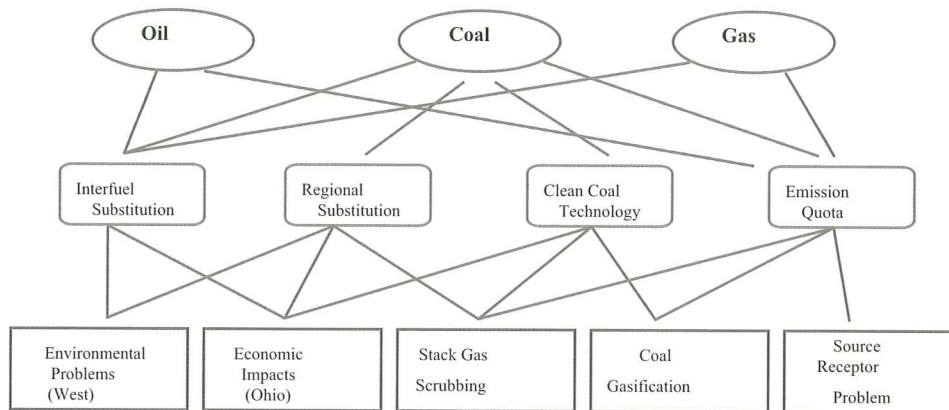Coal transportation cost is an important portion of the total

Figure 2. Economic-Environmental Interactions

cost of coal-fired electric generation. There are two basic approaches to the analysis of transportation costs: *direct modeling* and *indirect modeling*. The direct approach focuses on specific components of the coal transportation system. Models are built based on the theory that transportation cost is a function of quantity, distance, shipment delay, transport technology, and route. The indirect approach accounts for cost trade-offs between transportation cost and fuel cost. Models are estimated on the basis of economic theory. In the following section, we will build our model based on duality theory (Shephard, 1970) and on econometric approaches.

Theoretical specification

The duality of production and cost functions enables us to study production structures through the analysis of production costs. The total cost (C) of the delivered coal is a function of the quantity (Q) delivered, the input prices (P), spatial and technological factors (X), and coal characteristics (Z). The vector X should help control for the effects of variations in technology and in transportation mode, because different technologies can affect the costs of shipping operations. The Z vector should take care of the effects of heat value, ash content, and sulfur content in coal. More generally, the total cost is a function of output (O) and input prices (I). The general form of the cost function is

$$C = f(I, O) \tag{2.1}$$

The output, coal delivered, is the result of using two inputs: the coal at the mine and the transportation process. Thus, input prices include the mine-mouth price (W) and the transport rate (R). The output vector includes quantity (Q), distance (D), mode of transportation (M), ash content (A), sulfur content (S), heat value (H), and regional transshipment (B). A delay variable has been used as a proxy for transshipment in some models (Anderson, 1972; Zimmerman, 1977). We use a regional dummy variable (B) to represent regional differences in operations and transshipments, with the United States divided into nine regions. Since the cost of reducing sulfur emissions is very high under restrictive environmental standards, the sulfur content (S) in coal should be considered as a determinant of the total cost of purchased coal. Ash content is included due to its effect on the need for precipitators. Heat value is a critical determinant of the amount of energy generated by any given amount (tons) of coal. Then, the total cost (C) of the delivered coal can be specified as

$$C = f(W, R; Q, D, A, S, H, M, B) \tag{2.2}$$

Six modes (barge, truck, unit train, volume rate train, single car rate train, utility owned train) of transportation plus their combinations are used in coal delivery to Ohio power plants. Among them, barge is the most frequently used, with 55.17% of shipment, 36.57% of quantity, 28.76% of expenditures. Truck has the lowest shipment share, 0.99%, the lowest quantity share, 1.66%, and the smallest expenditure share,

1.51%. The shipment shares of trains (unit, volume rate) and their combinations are between 10% and 16%. In terms of transportation cost, the combined mode has the highest share, while truck, single car rate, and utility owned transportation individually represent less than 7% of total transportation costs. Barge is the cheapest transportation mode, 1.59 cents per ton per mile, while truck is the most expensive one, 7.9 cents per ton per mile. Single car rate trains have the longest average distance, 255.53 miles, while barges have the shortest distance, 106.11 miles.

We propose to use the translog cost function specification to estimate the relationship between total cost and those variables. This specification is termed as a *flexible functional form*, and can also be regarded as *a second-order Taylor series approximation* of the unknown function. The complete specification is a function of 38 variables:

$$
\begin{aligned}
\ln C = {} & \beta_0 + \beta_Q \ln Q + \beta_D \ln D + \beta_A \ln A + \beta_H \ln H + \beta_S \ln S \\
& + \beta_W \ln W + \beta_R \ln R + \beta_M M + \beta_B B + \beta_{QQ}(\ln Q)^2 + \beta_{DD}(\ln D)^2 \\
& + \beta_{AA}(\ln A)^2 + \beta_{HH}(\ln H)^2 + \beta_{SS}(\ln S)^2 + \beta_{WW}(\ln W)^2 \\
& + \beta_{RR}(\ln R)^2 + \beta_{QD} \ln Q \ln D + \beta_{QA} \ln Q \ln A + \beta_{QH} \ln Q \ln H \\
& + \beta_{QS} \ln Q \ln S + \beta_{QW} \ln Q \ln W + \beta_{QR} \ln Q \ln R + \beta_{DA} \ln D \ln A \\
& + \beta_{DH} \ln D \ln H + \beta_{DS} \ln D \ln S + \beta_{DW} \ln D \ln W + \beta_{DR} \ln D \ln R \\
& + \beta_{AH} \ln A \ln H + \beta_{AS} \ln A \ln S + \beta_{AW} \ln A \ln W + \beta_{AR} \ln A \ln R \\
& + \beta_{HS} \ln H \ln S + \beta_{HW} \ln H \ln W + \beta_{HR} \ln H \ln R + \beta_{SW} \ln S \ln W \\
& + \beta_{SR} \ln S \ln R + \beta_{WR} \ln W \ln R
\end{aligned}
\tag{2.3}
$$

The total cost (C) is equal to the sum of the expenditures of (1) purchasing and (2) delivering coal in each coal shipment, and will be expressed in dollars. The total consumption in ton, Q, is amount of coal shipped from a mine site to a power plant. The distance is calculated from latitudes and longitudes of mine sites and power plants by using distance equations presented in Chapter II. The distance is in miles. The mine-mouth price (W) is equal to the difference between total cost and total transportation cost, divided by total consumption (in dollar per ton). The transport rate (R) is obtained by dividing total transportation cost by the total consumption and the distance, and its unit is dollars per ton-mile. The average sulfur content (S) and ash content (A) are expressed in weight percentage. The heat value represents the quality of coal and is in mBtu per pound. The mode of transportation will be represented by dummy variable (M). M equals one if the mode is used in coal delivery; otherwise it is set to zero. Regional differences are described by a set of dummy variable (B). If coal is from region i, B_i is set to one, otherwise, to zero.

From Shephard's (1970) lemma, the factor share equations derived form the translog cost function are

$$
\begin{aligned}
S_R = \frac{\partial \ln C}{\partial \ln R} = {} & \beta_R + 2\beta_{RR} \ln R + \beta_{QR} \ln Q + \beta_{DR} \ln D + \beta_{AR} \ln A \\
& + \beta_{HR} \ln H + \beta_{SR} \ln S + \beta_{WR} \ln W
\end{aligned}
\tag{2.4}
$$

$$S_W = \frac{\partial \ln C}{\partial \ln W} = \beta_W + 2\beta_{WW} \ln W + \beta_{QW} \ln Q + \beta_{DW} \ln D + \beta_{AW} \ln A$$
$$+ \beta_{HW} \ln H + \beta_{SW} \ln S + \beta_{WR} \ln R$$

$$(2.5)$$

Generally, equation (2.3) is estimated together with equations (2.4) and (2.5). Since the shares sum to one and the errors associated with (2.4) and (2.5) are not mutually independent (Guldmann, 1990), the usual approach is to estimate only one of the two shares.

Empirical specification

We consider four alternative specifications: homogeneous translog function, homothetic translog function, Cobb-Douglas cost function, and extended homogeneous translog function (Guldmann, 1991, 1992).

Model 1: Homogeneous translog function
The homogeneous translog model includes terms for the variables Q, D, A, H, S, W, and R, and second order terms for Q, W, and R, with

$$\ln C = \beta_0 + \beta_Q \ln Q + \beta_D \ln D + \beta_A \ln A + \beta_H \ln H + \beta_S \ln S$$
$$+ \beta_W \ln W + \beta_R \ln R + \beta_{QQ}(\ln Q)^2 + \beta_{WW}(\ln W)^2 + \beta_{RR}(\ln R)^2$$
$$+ \beta_{QW} \ln Q \ln W + \beta_{QR} \ln Q \ln R + \beta_{WR} \ln W \ln R$$

$$(2.6)$$

$$S_R = \frac{\partial \ln C}{\partial \ln R} = \beta_R + 2\beta_{RR} \ln R + \beta_{QR} \ln Q + \beta_{WR} \ln W \quad (2.7)$$

$$S_W = \frac{\partial \ln C}{\partial \ln W} = \beta_W + 2\beta_{WW} \ln W + \beta_{QW} \ln Q + \beta_{WR} \ln R \quad (2.8)$$

and the price homogeneity restrictions:

$$\beta_W + \beta_R = 1, \qquad (2.9)$$
$$\beta_{WR} + 2*\beta_{WW} = 0, \qquad (2.10)$$
$$\beta_{WR} + 2*\beta_{RR} = 0, \qquad (2.11)$$
$$\beta_{QW} + \beta_{QR} = 0. \qquad (2.12)$$

Model 2: Homothetic translog function
A homothetic production function is characterized by output and input separability. The model includes all the variables of Model 1, but with the additional restriction:

$$\beta_{QW} = \beta_{QR} = 0. \qquad (2.13)$$

The functional form of the homothetic model is

$$\ln C = \beta_0 + \beta_Q \ln Q + \beta_D \ln D + \beta_A \ln A + \beta_H \ln H + \beta_S \ln S$$
$$+ \beta_W \ln W + \beta_R \ln R + \beta_{QQ} (\ln Q)^2 + \beta_{WW} (\ln W)^2$$
$$+ \beta_{RR} (\ln R)^2 + \beta_{WR} \ln W \ln R$$

$$(2.14)$$

with share functions

$$S_R = \frac{\partial \ln C}{\partial \ln R} = \beta_R + 2\beta_{RR} \ln R + \beta_{WR} \ln W \qquad (2.15)$$

$$S_W = \frac{\partial \ln C}{\partial \ln W} = \beta_W + 2\beta_{WW} \ln W + \beta_{WR} \ln R \qquad (2.16)$$

Model 3: Cobb-douglas function
A Cobb-Douglas cost function includes only variables of the first order, which implies the following additional restrictions

$$\beta_{QQ} = 0, \qquad (2.17)$$
$$\beta_{WW} = \beta_{RR} = \beta_{WR} = 0. \qquad (2.18)$$

The Cobb-Douglas cost function is

$$\ln C = \beta_0 + \beta_Q \ln Q + \beta_D \ln D + \beta_A \ln A + \beta_H \ln H$$
$$+ \beta_S \ln S + \beta_W \ln W + \beta_R \ln R \qquad (2.19)$$

with share functions

$$S_R = \frac{\partial \ln C}{\partial \ln R} = \beta_R, \qquad S_W = \frac{\partial \ln C}{\partial \ln W} = \beta_W \qquad (2.20)$$

Model 4: Extended homogeneous model
An extended homogeneous translog model has all the variables in Model 1 plus second-order terms for distance (D), heat value (H), ash content (A), and sulfur content (S). The full expression is

$$\ln C = \beta_0 + \beta_Q \ln Q + \beta_D \ln D + \beta_A \ln A + \beta_H \ln H + \beta_S \ln S$$
$$+ \beta_W \ln W + \beta_R \ln R + \beta_{QQ}(\ln Q)^2 + \beta_{WW}(\ln W)^2 + \beta_{RR}(\ln R)^2$$
$$+ \beta_{QW} \ln Q \ln W + \beta_{QR} \ln Q \ln R + \beta_{WR} \ln W \ln R$$
$$+ \beta_{DD}(\ln D)^2 + \beta_{HH}(\ln H)^2 + \beta_{SS}(\ln S)^2 + \beta_{AA}(\ln A)^2$$

$$(2.21)$$

The restrictions and share functions are the same as in Model 1.

Data sets and data processing

The utility industry is a very important component of national and local economies, and its development deeply impacts people's daily life. Many federal and state agencies are monitoring, regulating, and guiding utility companies to ensure the stability and development of the utility industry. Millions of data items are collected and published in various media (EIA, 1991; EPA, 1990; FERC, 1990). For the electric utility industry, many publications such as Weekly Coal Production, Natural Gas Monthly, and Electric Power Monthly (EPM) are available through paper, tape, and Internet accessing (EIA, 1992).

The following section describes the databases used and the information extracted from the databases.

Data sets
The Public Utilities Commission of Ohio (PUCO) compiles a data set named the Electric Fuel Component (EFC) (Ohio, 1992). Currently, the data set consists of 12 years (1982 to 1993) of monthly records for each power plant of the eight major, privately owned power companies in Ohio. There are about 80 power plants in Ohio.

There are 32 different types of records in the EFC database, which can be categorized into four groups: 1) purchase transaction data (record types 01, 02, 11); 2) plant monthly

data (record types 31-34); 3) company sale data (record types 40-46); 4) company evaluation data (record types 47-64).

The EFC data includes fuel supplies, fuel consumptions, fuel transaction records (quantity, contract term, costs, etc.), fuel types (quality, heat value, ash content, sulfur content), and the transportation modes used to complete the fuel trade.

The EFC data does not provide specific location information regarding power plants, although it provides mine location information, such as counties, states, and district regions where the fossil fuel was mined. In order to analyze transportation costs, information about power plant location must be obtained from other data sets so that distances between mine sources and power plants can be calculated.

The other two data sets to be used are **STF1C.US90** and **NADB**. The STF1C.US90 data set is a 1990 census file that contains the latitude and longitude for the centroid of each county in the United States. The latitude and longitude can be used to calculate distances from mining sites to coal consuming power plants. The NADB data set, developed by the U.S. Environmental Protection Agency, has county locations for all U.S. power plants. Thus, by merging these data sets, we will be able to obtain latitudes and longitudes of all the counties where power plants and mines are located.

Programs and data processing
Two statistical packages are used in data processing and programming: FOCUS and SAS. FOCUS is a powerful data processing and report design package running on VAX machines, but it has limited mathematical functions. FOCUS does not have built-in trigonometric functions, although it provides interface for users to add their own functions by programming in C, FORTRAN, or other computer languages. We primarily use FOCUS to extract data from the EFC data set.

SAS is powerful not only in data processing, mathematical calculation, but also in statistical modeling (SAS, 1990). We use SAS to conduct major data processing and statistical modeling. All the programs can be classified into two categories: data processing and statistical modeling.

The data sets were collected for different purposes, so the structures of these data sets and the length of corresponding variables are different. They need to be unified. Extensive amount of time and many programs are contributed to unify the formats of data and units of measures.

Distance calculation
Since information about route distances is not available in the data sets, we use the formulas for calculating the airline distances to obtain distances between mining sites and power plants. The airline distance is represented by the great circle distance, and can be computed as a function of the latitudes and longitudes of an origin and a destination. The following formula has been drawn from Love, Morris and Wesolowsky (1988):

$$D_{qr} = 24902 \bullet \frac{\pi}{180} \bullet arccos[\cos(q_1)\cos(r_1)\cos(q_2 - r_2) + \sin(q_1)\sin(r_1)]$$
(2.22)

where 24902 is the length of the equator (miles); q and r are two cities; q_1, q_2 represent city q's latitude and longitude; r_1, r_2 represent city r's latitude and longitude.

We have obtained the latitudes and longitudes of the centroids of all the counties where either power plants or mining sites are located. After identifying origin and destination relationships, we calculate distances for all the O-D pairs and output them to an intermediate file called *CTC1EFC*. More accurate distance data could be obtained from the *Standard Highway Mileage Guide* if the origin and the destination are near big cities, but in most cases, the mine sites are in a maintain area. More realistic distances could be obtained by overlaying the mine sites and power plants on river and railway GIS maps. This was not attempted due to resource and time limitations.

Estimation and results

Since all the costs and prices are adjusted to the value of 1983 (US DoL, 1984), we are only dealing with cross-section data. The sections could be company, transportation mode, and geographical region. Companies are decision-makers, and different levels of strategic planning, effective management, and efficient operation would strongly affect the cost structure of their fuel purchasing. Transportation modes represent technology and time factors involved in coal delivery. The costs of mining are mainly associated with mine deposit location, reserve, quality and mining technology. With the same technology and efficiency in mining operation, mine-mouth prices could still present big differences due to the geographical characteristics of the mine.

We have estimated the four models under four different combinations: (1) all companies, regions, and transport modes combined; (2) by company; (3) all companies with region and transport mode dummy variables; and (4) by company, with region and transport mode dummy variables. We find that the regression outcomes are essentially the same under these four approaches in terms of signs, significance, and magnitude of coefficients. The main differences can be found as follows: (a) by-company regressions display an improved overall R-square, but the significance of the coal characteristic variables (S, A, and H) varies across companies, since some companies only use 'clean' coal, and some use very dirty coal; (b) including dummy variables slightly improves the R^2.

The procedure proposed by Zellner (1962) for estimating seemingly unrelated regressions (SUR) and ordinary least square (OLS) was used in our estimation. Since the sample size is very large, it is very CPU time-consuming under SAS default setting - 0.001 convergence criterion. The program easily reaches over 1000 iterations for a single convergence.

We have chosen a moderate convergence value of 0.1.

We expect positive relationships between total cost and quantity, distance, heat value, transportation rate, and mine-mouth price, and negative relationships between total cost and sulfur content and ash content. For the purpose of comparison, we have included the results from both the OLS and SUR methods. The R^2 of the total cost function is over 95% in all models. The R^2 of the share equation is over 37% (SUR) for all models, except for Model 3. The negative R^2 of the equation in Model 3 implies that the Cobb-Douglas functional form is not appropriate for coal transportation cost analysis.

We have obtained correct signs for all the coefficients except for the coefficient of the heat value in Models 1-3. The t-test at the 5% significance level shows that we can reject the hypothesis of no relationship (H_0) for all the coefficients in the three models. In Model 4, we added the second-order terms for distance, heat value, ash content, and sulfur content. It turns out that the coefficients of ash content, heat value, and the second-order term for heat value are insignificant. The sign of the heat value's coefficient turns out to be positive, which is what we had expected. The sign of the distance variable's coefficient is negative for the first-order term, but positive for second-order one.

It is clear that Model 4 is the best of the four models. We have chosen Model 4 as our model to continue coal transportation analysis. Since the ash content coefficient is insignificant at the 5% significance level, we have dropped it from Model 4. Table 1 shows the results of re-running the modified Model 4. For comparison, the results for Model 1 are also included in Table 1. Dropping ash content has improved the model a lot in terms of the t-ratios of the individual variables and slightly in terms of the overall R^2.

In order to test the price homogeneity assumption, we have run Model 1 without price homogeneity constraints. The results are surprisingly close to those with homogeneity restrictions.

III. ELASTICITY AND POLICY ANALYSES

The models developed in previous sections provide us tools to analyze the structure of total energy consumption, individual fuel consumptions, total cost, and total transportation cost. Price-elasticity and share-elasticity are usually used to assess interfuel substitutions and input substitution (coal cost or transport cost). Policies can be formulated based on elasticity analysis.

Economies of scale analysis

Economies of scale are usually measured by the ratio of marginal cost, MC, to average cost, AC, equal to the elasticity of cost with respect to output (Guldmann, 1985), with

$$\varepsilon_Q = \frac{MC}{AC} = \frac{\partial C / \partial Q}{C / Q} = \frac{\partial \ln C}{\partial Q} \qquad (3.1)$$

There are economies of scale in coal delivery when ε_Q is less than one, while $\varepsilon_Q > 1$ implies that the power company experiences decreasing returns to scale in coal delivery. From the translog function of Model 4, we derive the elasticity of cost with respect to the quantity of coal delivered:

$$\varepsilon_Q = \beta_Q + 2\beta_{QQ} \ln Q + \beta_{QR} \ln R + \beta_{QW} \ln W \qquad (3.2)$$

Using the translog parameters of Model 4 (SUR results in Table 1), and the mean values of R and W, equation (3.2) becomes

$$\varepsilon_Q = 1.027301 - 0.00385 \ln Q \qquad (3.3)$$

Equation (3.3) shows that at average mine-mouth price and transport rate, economies of scale rise with an increasing quantity of coal delivered. At the average quantity, 19,435 tons, we find $\varepsilon_Q = 0.989248$, indicating that the coal is delivered at increasing returns to scale. It is quite straightforward to obtain the threshold value of Q for constant return to scale by solving the equation $\varepsilon_Q = 1$, which yields $Q_1 = 1,193.51$ tons. This means that any coal delivery of more than 1,193.51 tons experiences economies of scale. This represents most of the cases for the four companies.

When we discuss the elasticity of cost with respect to quantity, we assume that the other outputs remain constant. One of the other important outputs is distance. The elasticity of cost with respect to distance, in Model 4, is

$$\varepsilon_D = \frac{\partial \ln C}{\partial D} = \beta_D + 2\beta_{DD} \ln D \qquad (3.4)$$
$$= -0.26 + 0.08008 \ln D$$

When $D = 1$ mile, $\varepsilon_D = -0.26$, implying that the total cost of coal purchasing is reduced by 26% when distance increases by 1% (or 0.01 mile). This negative elasticity takes place until the distance is equal to 25.71 miles (D_0), where it is equal to zero. From D_0 upward, the total cost rises with increasing distance. The distance corresponding to unit elasticity, D_1 is 6,812,813 miles. In the sampled data, the minimum distance is 26.74 miles, which is slightly greater than D_0, and the maximum distance is 358.94 miles, thus, all the companies experience economies of scale with distance.

In order to assess the relationship between total cost and both distance and quantity, we use the concept of ray economies of scale (Baumol, 1977) to measure the behavior of the total cost when both outputs (Q, D) increase by the same percentage, with

$$\varepsilon_{QD} = \varepsilon_Q + \varepsilon_D$$
$$= \beta_Q + \beta_D + \beta_{QR} \ln R + \beta_{QW} \ln W + 2\beta_{QQ} \ln Q + 2\beta_{DD} \ln D$$
$$= 0.7673 + 0.08008 \ln D - 0.00385 \ln Q$$
$$\qquad (3.5)$$

When $D = 1$ mile and $Q = 1$ ton, the total cost of coal is slight

Table 1. Estimation results of final functional forms

Nonlinear OLS and ITSUR Summary of Residual Errors

		Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Sqr	Adj R-Sq
Model 1	OLS	TC	7.5	7,737	13.16	0.0017	0.0412	0.9991	0.9991
		SR	1.5	7,743	42.15	0.0054	0.0738	0.5201	0.5200
	ITSUR	TC	7.5	7,737	10.21	0.0013	0.0363	0.9993	0.9993
		SR	1.5	7,743	55.35	0.0071	0.0845	0.3698	0.3697
Model 4	OLS	TC	10.5	7,734	12.56	0.0016	0.0403	0.9992	0.9992
		SR	1.5	7,743	42.09	0.0054	0.0737	0.5207	0.5207
	ITSUR	TC	10.5	7,734	8.43	0.0011	0.0330	0.9994	0.9994
		SR	1.5	7,743	53.70	0.0069	0.0833	0.3885	0.3884

OLS Estimation Results, Converge=0.1

	Model 1: Price Homogeneity					Model 4: Extended Price Homogeneity			
Parameter	Estimate	Approx. Std Err	'T' Ratio	Approx. Prob>|T|	Parameter	Estimate	Approx. Std Err	'T' Ratio	Approx. Prob>|T|
B0	3.2181	0.0886	36.33	0.0001	B0	-18.1126	4.6807	-3.87	0.0001
B_Q	1.0108	0.0037	272.23	0.0001	B_Q	1.0094	0.0037	276.43	0.0001
B_R	0.5997	0.0040	148.74	0.0001	B_R	0.5977	0.0040	149.46	0.0001
B_D	0.1305	0.0015	84.66	0.0001	B_D	-0.2858	0.0213	-13.43	0.0001
B_H	-0.1219	0.0091	-13.33	0.0001	B_H	4.6186	0.9992	4.62	0.0001
B_S	-0.0254	0.0007	-35.97	0.0001	B_S	-0.0225	0.0008	-28.89	0.0001
B_QQ	-0.0024	0.0002	-15.46	0.0001	B_QQ	-0.0022	0.0002	-14.48	0.0001
B_RR	0.0305	0.0002	146.54	0.0001	B_RR	0.0306	0.0002	148.65	0.0001
B_QR	-0.0037	0.0004	-10.62	0.0001	B_QR	-0.0036	0.0003	-10.31	0.0001
					B_DD	0.0438	0.0022	19.57	0.0001
					B_HH	-0.2518	0.0532	-4.73	0.0001
					B_SS	-0.0027	0.0004	-6.30	0.0001
B_W	0.4003				B_W	0.4023			
B_WR	-0.0611				B_WR	-0.0612			
B_WW	0.0305	0.0002	146.54	0.0001	B_WW	0.0306	0.0002	148.65	0.0001
B_QW	0.0037	0.0004	10.62	0.0001	B_QW	0.0036	0.0003	10.31	0.0001

ITSUR Estimation Results, Converge=0.1

	Model 1: Price Homogeneity					Model 4: Extended Price Homogeneity			
Parameter	Estimate	Approx. Std Err	'T' Ratio	Approx. Prob>|T|	Parameter	Estimate	Approx. Std Err	'T' Ratio	Approx. Prob>|T|
B0	2.1782	0.0652	33.40	0.0001	B0	-6.9458	3.5149	-1.98	0.0482
B_Q	0.9912	0.0029	344.81	0.0001	B_Q	0.9898	0.0029	341.45	0.0001
B_R	0.4422	0.0032	138.94	0.0001	B_R	0.4581	0.0033	140.94	0.0001
B_D	0.1105	0.0011	98.69	0.0001	B_D	-0.2600	0.0160	-16.24	0.0001
B_H	-0.0556	0.0066	-8.36	0.0001	B_H	2.0651	0.7503	2.75	0.0059
B_S	-0.0189	0.0005	-36.72	0.0001	B_S	-0.0161	0.0006	-27.38	0.0001
B_QQ	-0.0020	0.0001	-17.77	0.0001	B_QQ	-0.0019	0.0001	-16.72	0.0001
B_RR	0.0193	0.0002	119.54	0.0001	B_RR	0.0202	0.0002	122.12	0.0001
B_QR	-0.0055	0.0003	-19.43	0.0001	B_QR	-0.0056	0.0003	-19.68	0.0001
					B_DD	0.0400	0.0017	23.79	0.0001
					B_HH	-0.1120	0.0400	-2.80	0.0051
					B_SS	-0.0020	0.0003	-6.14	0.0001
B_W	0.5578				B_W	0.5419			
B_WR	-0.0386				B_WR	-0.0404			
B_WW	0.0193	0.0002	119.54	0.0001	B_WW	0.0202	0.0002	122.12	0.0001
B_QW	0.0055	0.0003	19.43	0.0001	B_QW	0.0056	0.0003	19.68	0.0001

inelastic with regard to distance and quantity. By solving the equation $\varepsilon_{QD} = 1$, we obtain a log-log linear relationship of quantity with distance, with

$$\ln Q = -60.3875 + 20.7811 * \ln D \qquad (3.6)$$

Equation (3.6) indicates that the greater the distance, the larger the quantity of coal needed to maintain constant returns to scale. With less coal than this threshold, shipments display diseconomies of scale.

Elasticity of substitution

In the system of coal purchasing, the transportation rate (R) and mine-mouth price (W) are the two input prices. The sensitivities of output to input prices can be measured by the elasticity of substitution. The elasticity of substitution, ε_{WR}, and the direct elasticities of factor demand are defined (Guldmann, 1990) by

$$S_{WR} = (\beta_{WR} + S_W S_R) / (S_W S_R), \qquad (3.7)$$

$$\varepsilon_{WW} = \frac{\partial \ln C_W}{\partial W} = \frac{2\beta_{WW} + S_W (S_W - 1)}{S_W}, \qquad (3.8)$$

$$\varepsilon_{RR} = \frac{\partial \ln C_R}{\partial W} = \frac{2\beta_{RR} + S_R (S_R - 1)}{S_R}, \qquad (3.9)$$

We have calculated these elasticities for all the 7744 observations in the case of Model 4. Table 2 presents the results by using the SUR coefficients. The W and R statistics are from the original data. The shares are calculated, with a negative minimum share for the transportation cost and a bigger than one share for coal cost. There are only 36 of these unusual cases, only 0.46% of the 7744 records (Table 3).

The elasticity of substitution, S_{WR}, varies between -47.97 and 260.42, with a mean of 0.641. The elasticity of coal cost with respect to mine-mouth price, ε_{WW}, varies between -0.161 and 0.168, with a mean of -0.079, implying that increasing by 1% the mine-mouth price causes a 0.079% decrease in quantity

(tons), which must be then substituted by coals transported from other sites. The elasticity of the transportation cost with respect to the transportation rate, ε_{RR}, varies between -260.46 and 47.93, with a mean of -0.561, indicating that increasing by 1% the transportation rate results in a 0.561% volume substitution (ton*miles). The positive value of S_{WR} implies that coal cost and transportation cost are substitutes. Since the elasticity of substitution is less than one, the substitution is inelastic. Statistics for these elasticities are presented in Table 3.

The elasticity of the coal cost with respect to the transportation rate, ε_{WR}, varies between -0.168 and 0.161, with a mean of 0.079; the elasticity of the transportation cost with respect to the mine-mouth price, ε_{RW}, varies between -47.93 and 260.46, with a mean of 0.561. It is clear that when the mine-price rises in Ohio, power companies look for substitutions from other states. They are willing on average to increase by 0.561% their transportation cost when the mine-mouth price increases by 1%.

A well-behaved function must be concave in inputs, which requires the Hessian matrix, H, to be negative semidefinite, i.e., $\varepsilon_{WW} < 0$ and Det $H = (\varepsilon_{WW}\varepsilon_{RR}S_w S_R - (S_{WR})^2)/(W^2 R^2) \geq 0$. ε_{WW} is negative in 98.8% of the cases (Table 3), thus guaranteeing that the cost function is never convex, and Det H is negative in 39.27% of the cases, at which points the function is neither convex nor concave. The function is concave at all the other (60.73%) cases (Table 3).

Site specific analysis

Different mine sites have different price-elasticities and will have different responses to their own price and cross-price changes due to their unique coal characteristics, relative locations, and accessibility. If taxation were posted on burning high sulfur content coal, what would happen in regional distribution of regional coal consumption?

Table 2. Elasticity of substitution

Variable	Unit	Elasticity Analysis: Using SUR Results				
		N	Mean	Std Dev	Minimum	Maximum
W	($/ton)	7744	27.78	8.75	13.33	353.64
R	($/ton/mile)	7744	0.0348	0.0271	0	0.8979
S_R	(%)	7744	12.56%	0.0354	-13.28%	21.25%
S_W	(%)	7744	87.44%	0.0354	78.75%	113.28%
ε_{RR}		7744	-0.561	3.013	-260.463	47.926
ε_{WW}		7744	-0.079	0.034	-0.161	0.168
S_{WR}		7744	0.641	3.012	-47.966	260.423
ε_{WR}		7744	0.079	0.034	-0.168	0.161
ε_{RW}		7744	0.561	3.013	-47.926	260.463
C	(k $)	7744	714	1,034	0.14	10,455
Det		7744	-0.038	0.835	-35	0
H		7744	-33,568	2,716,310	-2.39E+08	-5.18E-06

In order to conduct site-specific analysis, all the mine sites are classified into different classes based on their spatial orders and sulfur contents. Arc/Info is used to obtain spatial orders of mine sites, and then the parameters obtained from coal transportation modelling Model 4 are used to calculate economic scales and elasticities. Figure 3 displays spatial orders, coal sites and their links. Table 4 shows the results.

The results show that coal shipments from all the sites experience economies of scale with distance and with quantity and distance. The coals with higher sulfur content experience higher elasticity of substitution. The direct elasticity of transportation rate is much higher than the direct elasticity of mine mouth price of coal.

IV. CONCLUSIONS

Statistical modeling involves theoretical and empirical model specification, extensive data processing, and model testing and econometric analyses. GIS is a very good tool to help visualizing the spatial relationship and classifying mine sites into different spatially related regions, but lacking statistical modeling tools.

The coal in Ohio is in the high-sulfur content category. The implementation of Phase I of the CAAA will force Ohio power companies to import more clean coal from other regions, especially from the East South Central one. The translog function with price homogeneity restrictions is used to assess the cost structure of the coal delivery system. Our findings can be summarized as follows:

- At the average mine-mouth price and transportation rate, economies of scale rise with the quantity of coal delivered. Economies of scale take place upward from 1,193.51 tons.
- The zero elasticity distance are 25.71 miles. Above this distance, cost increases with distance, but inelastically.

Table 3. Frequency distributions of elasticities

X	n	f	N	F
Distribution of S_R				
$-0.5 \leq X < 0$	36	0.46	36	0.46
$0 \leq X < 0.5$	7708	99.54	7744	100.00
Distribution of S_W				
$0.5 \leq X < 1$	7708	99.54	7708	99.54
$1.0 \leq X$	36	0.46	7744	100.00
Distribution of S_{WR}				
$X < -1.0$	10	0.13	10	0.13
$-1.0 \leq X \ -0.5$	6	0.08	16	0.21
$-0.5 \leq X < 0$	44	0.57	60	0.77
$0 \leq X < 0.5$	835	10.78	895	11.56
$0.5 \leq X < 1$	6,813	87.98	7,708	99.54
$1.0 \leq X$	36	0.46	7,744	100.00
Distribution of ε_{RR}				
$X < -1.0$	36	0.46	36	0.46
$-1.0 \leq X \ -0.5$	6,066	78.33	6,102	78.80
$-0.5 \leq X < 0$	1,582	20.43	7,684	99.23
$0 \leq X < 0.5$	45	0.58	7,729	99.81
$0.5 \leq X < 1$	5	0.06	7,734	99.87
$1.0 \leq X$	10	0.13	7,744	100.00
Distribution of ε_{WW}				
$-0.5 \leq X < 0$	7,648	98.76	7,648	98.76
$0 \leq X < 0.5$	96	1.24	7,744	100.00
Distribution of Det H				
$X < -1.0$	21	0.27	21	0.27
$-1.0 \leq X \ -0.5$	3	0.04	24	0.31
$-0.5 \leq X < 0$	3,017	38.96	3,041	39.27
$0 \leq X < 0.5$	4,703	60.73	7,744	100.00

Notes: $X = S_W, S_R, S_{WR}, \varepsilon_{RR}, \varepsilon_{WW}$, or *Det H*; n = observations; f = frequency
N = cumulative observations; F = cumulative frequency.

- The substitution between coal cost and transportation cost is inelastic. The mean elasticity of substitution is 0.641. Increasing by 1% the mine-mouth price causes a 0.079% of

Table 4. Economies of scale and elasticity of substitution

| Spatial Class	N	Mean	Std Dev	Min	Max	Mean	Std Dev	Min	Max
Sulfur Content						*Economic Scale to Distance*			
1	4	3.23%	21.9298	2.99%	3.48%	0.1676	0.0094	0.1589	0.1810
2	6	2.93%	50.0506	2.45%	3.79%	0.1073	0.0239	0.0805	0.1429
3	22	2.20%	82.2138	0.39%	3.66%	0.1290	0.0323	0.0698	0.1976
4	13	1.89%	129.1837	0.84%	4.04%	0.1200	0.0386	0.0495	0.1794
5	29	1.10%	63.6041	0.12%	3.29%	0.1154	0.0430	0.0283	0.2007
Economic Scale to Quantity and Distance						*Elasticity of Substitution*			
1	4	1.1640	0.0117	1.1530	1.1803	0.4162	0.2187	0.0965	0.5894
2	6	1.1011	0.0240	1.0734	1.1352	0.5838	0.1826	0.2181	0.7094
3	22	1.1200	0.0330	1.0559	1.1922	0.5729	0.1760	0.0180	0.7435
4	13	1.1127	0.0380	1.0415	1.1744	0.3045	0.2611	-0.3060	0.6905
5	29	1.1076	0.0430	1.0234	1.1927	0.0618	0.5628	-1.1252	0.7454
Direct Elasticity of Coal Price						*Direct Elasticity of Transportation Rate*			
1	4	-0.0628	0.0417	-0.1074	-0.0071	-0.3534	0.1786	-0.4820	-0.0895
2	6	-0.1351	0.0668	-0.2138	-0.0187	-0.4486	0.1221	-0.5036	-0.1995
3	22	-0.1291	0.0659	-0.2924	-0.0012	-0.4438	0.1237	-0.5050	-0.0168
4	13	-0.0478	0.0538	-0.1874	0.0151	-0.2567	0.2173	-0.5031	0.2909
5	29	-0.0612	0.1044	-0.2994	0.0334	-0.0005	0.4775	-0.5044	1.0918
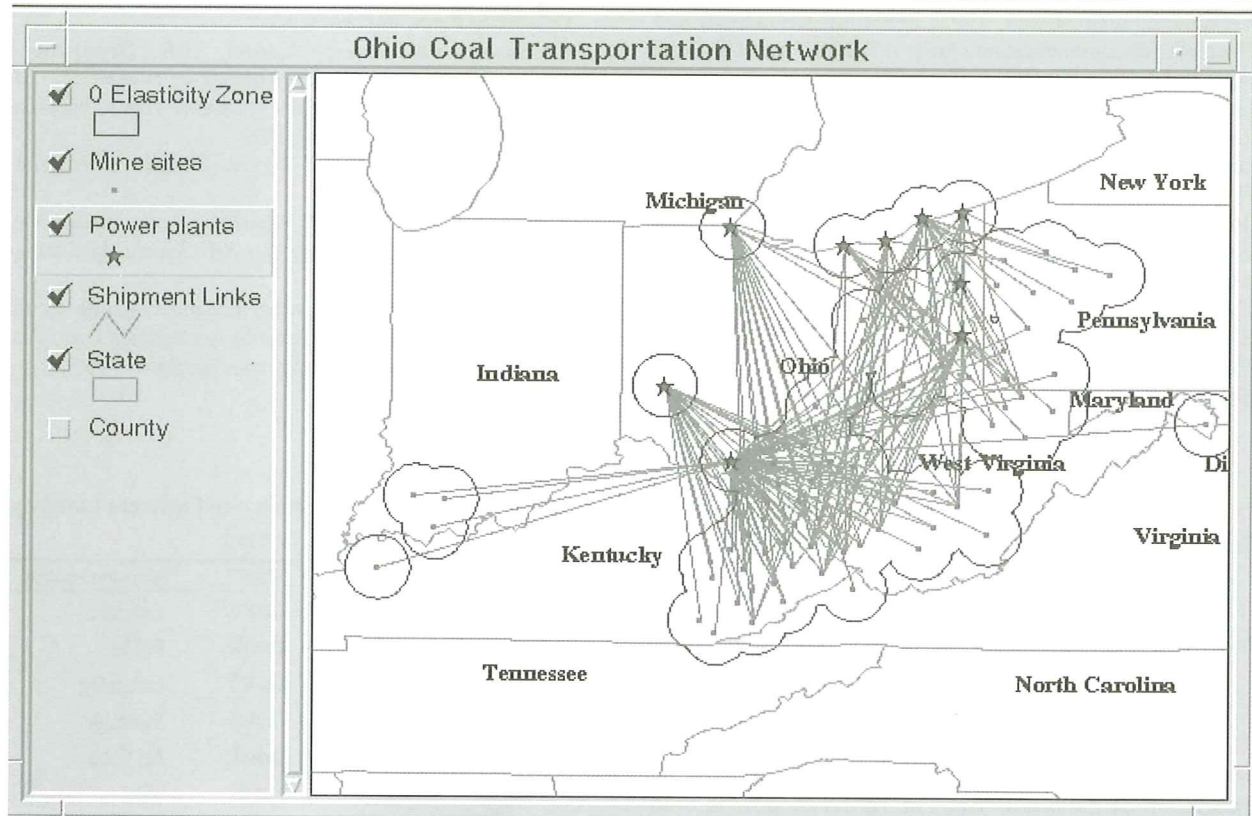
Figure 3. Coal shipment sites and their links

the quantity (tons of coal) to be substituted for the coals transported from other sites, while increasing by 1% the transportation rate results in a 0.561% volume (ton*miles) substitution.

• Regional coal substitution is possible.

ACKNOWLEDGMENTS

REFERENCES

[1] Anderson, Dennis, 1972, Models for determining Least-Cost Investments in Electricity Supply, *The Bell Journal of Econ. Mgmt. Sci.*, 3:267-299.

[2] Bernknopf, Richard Lewis, 1985, *Domestic Coal Distribution: An Interregional Programming Model for the U.S. Coal Industry.* Greenwich, Connecticut: Jai Press Inc.

[3] Baumol, W. J., 1977, On the proper cost for natural monopoly in a multiproduct industry, *American Economic Review*, 67:809-822.

[4] Energy Information Administration, 1991, *Coal Data: A Reference.*

[5] Energy Information Administration, Office of Coal, Nuclear, Electric and Alternate Fuels, U.S. Department of Energy, 1992, *Electric Power Monthly*, Washington, DC 20585.

[6] Energy Information Administration, Office of Integrated Analysis and Forecasting, U.S. Department of Energy, 1993a, *Assumptions for the Annual Energy Outlook 1993*, DOE/EIA-0527(93). Washington, D.C.: 1993.

[7] Energy Information Administration, Office of Integrated Analysis and Forecasting, U.S. Department of Energy, 1993b, *Annual Energy Outlook 1993: With Projections to 2010.* DOE/EIA-0383(93). Washington, D.C.: 1993.

[8] Energy Information Administration, Office of Energy Markets and End Use, U.S. Department of Energy, 1993c, *Short-Term Energy Outlook*, Washington, DC 20585.

[9] Environmental Protection Agency, 1990, Title IV Acid Deposition Program, *Clean Air Act Amendments of 1990*, (November 1990), pp.1-3.

[10] Federal Energy Regulatory Commission, 1990, *FERC423 Data in 1990.*

[11] Gallant, A. Ronald, 1987, *Nonlinear Statistical Models*, Wiley Series Probability and Mathematical Statistics.

[12] Guldmann, Jean-Michel, 1985, Economies of Scale and Natural Monopoly in Urban Utilities: The Case of Natural Gas Distribution, *Geographical Analysis* 17(4):302-317.

[13] Guldmann, Jean-Michel, 1990, Economies of Scale and Density in Local Telephone Networks, *Regional Science and Urban Economics,* 20:521-535.

[14] Guldmann, Jean-Michel and Donald A. Hanson, 1991, Natural Gas Market Expansion and Delivery Infrastructure Costs: The Case of New England, *Resource and Energy,* 13:57-94 .

[15]Guldmann, Jean-Michel, 1992, Modeling Residential and Business Telecommunication Flows: A Regional Point-to-Point Approach, *Geographical Analysis*, 24(2):121-141.

[16]Love, R. F., Morris, J. g., and Wesolowsky, G. O., 1998, *Facilites Location: Models & Methods*, Elsevier Science Puglishing Co. Inc., New York.

[17]Nagarvala, Phiroze J.; Ferrill, George C.; and Oliver, Leon A., 1976, *Clean Coal Energy: Source-to-Use Economics Project*. San Francisco: Bechtel.

[18]Ohio Coal Development Office, 1992, *1992 Ohio Coal Development Agenda*, Draft, January 24, 1992.

[19]SAS Institute Inc., 1990, *SAS/STAT User's Guide* Version 6, Fourth Edition.

[20]Shephard, R.W., 1970, *Theory of cost and productions*. Princeton University Press. Princeton, NJ.

[21]Transportation Service Center, U.S. Department of Transportation, 1976, *The Calculation of Comparable Modal Shipment Costs for Regional Commodity Flows*, by Davis L. Anderson. Washington, D.C.:1976.

[22] U.S. Department of Labor, Bureau of Labor Statistics. 1984. *CPI Detailed Report*.

[23] Zimmerman, Martin B., 1977, Modelling Depletion In A Mineral Industry: The Case of Coal, *The Bell Journal of Economics*, 8(1):41-65.

[24] Zellner, A., 1962, An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias, *Journal of the American Statistical Association*, 58:977-992.

Corrections

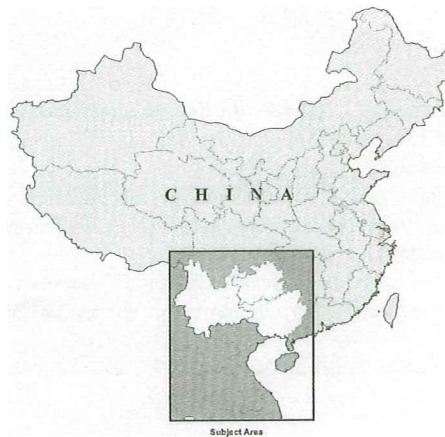The figure and table (bottom and right) were published on Volume 6, Number 2, 2000, pp. 132-133, by Luo et al.. Here are some special font error corrections. For detail description of the figure and table, see context in the paper by Luo et al. in the last issue.

(editor)

Figure 5. This map shows the spatial variation of the pronunciation of the word "rice" in Tai languages (in bold font). The spatial pattern indicates that the origin (where most change occurred) is in the Guizhou-Guangxi border. Province names are labeled with italic font and the place names are labeled with a smaller font. Places outside of China are also included to show a more complete picture of spatial pattern and direction of spread.

Table 1. The word "Rice" Spoken in Different Languages and Location

Language Group	Language	"Rice"	Location Spoken
Tai	Siamese	khaaw3	Bangkok
	Isan	khaw3	Roi Et
	S. Zhuang	khaw3	Lungming
	Lao	khaw3	Vientiane
	Tai Don	khau3	Lai Chau
	Saek	ɣawɓ	Nakhon Phanom
	Tai Dam	xaw3	Son La
	Tai Dehong	xau3	Luxi
	Tai Lue	xau3	Jinghong
	Yay	haw6	Lao Cai
	N. Zhuang	hau4	Wuming
	Bouyei	hau4	Wangmo
Kam-Sui	Sui	ʔau4	Libo
	S. Kam	qəu4	Rongjiang
	Maonan	ɦu4	Huangjiang
	Mulam	hu3	Luocheng

Note: See Figure 4 for a tree diagram of the language group. The languages are sorted in the order of decreasing similarity to the proto-Tai word of "rice"-*khau [C3]. In other words, going down the list, you are going in a general direction from a pronunciation most similar to proto-Tai (at periphery) to one most different (at origin).

[15]Guldmann, Jean-Michel, 1992, Modeling Residential and Business Telecommunication Flows: A Regional Point-to-Point Approach, *Geographical Analysis*, 24(2):121-141.

[16]Love, R. F., Morris, J. g., and Wesolowsky, G. O., 1998, *Facilites Location: Models & Methods*, Elsevier Science Puglishing Co. Inc., New York.

[17]Nagarvala, Phiroze J.; Ferrill, George C.; and Oliver, Leon A., 1976, *Clean Coal Energy: Source-to-Use Economics Project.* San Francisco: Bechtel.

[18]Ohio Coal Development Office, 1992, *1992 Ohio Coal Development Agenda*, Draft, January 24, 1992.

[19]SAS Institute Inc., 1990, *SAS/STAT User's Guide* Version 6, Fourth Edition.

[20]Shephard, R.W., 1970, *Theory of cost and productions*. Princeton University Press. Princeton, NJ.

[21]Transportation Service Center, U.S. Department of Transportation, 1976, *The Calculation of Comparable Modal Shipment Costs for Regional Commodity Flows*, by Davis L. Anderson. Washington, D.C.:1976.

[22] U.S. Department of Labor, Bureau of Labor Statistics. 1984. *CPI Detailed Report*.

[23] Zimmerman, Martin B., 1977, Modelling Depletion In A Mineral Industry: The Case of Coal, *The Bell Journal of Economics*, 8(1):41-65.

[24] Zellner, A., 1962, An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias, *Journal of the American Statistical Association*, 58:977-992.

Corrections

The figure and table (bottom and right) were published on Volume 6, Number 2, 2000, pp. 132-133, by Luo et al.. Here are some special font error corrections. For detail description of the figure and table, see context in the paper by Luo et al. in the last issue.

(editor)

Figure 5. This map shows the spatial variation of the pronunciation of the word "rice" in Tai languages (in bold font). The spatial pattern indicates that the origin (where most change occurred) is in the Guizhou-Guangxi border. Province names are labeled with italic font and the place names are labeled with a smaller font. Places outside of China are also included to show a more complete picture of spatial pattern and direction of spread.

Table 1. The word "Rice" Spoken in Different Languages and Location

Language Group	Language	"Rice"	Location Spoken
Tai	Siamese	khaaw3	Bangkok
	Isan	khaw3	Roi Et
	S. Zhuang	khaw3	Lungming
	Lao	khaw3	Vientiane
	Tai Don	khau3	Lai Chau
	Saek	ɣawɓ	Nakhon Phanom
	Tai Dam	xaw3	Son La
	Tai Dehong	xau3	Luxi
	Tai Lue	xau3	Jinghong
	Yay	haw6	Lao Cai
	N. Zhuang	hau4	Wuming
	Bouyei	hau4	Wangmo
Kam-Sui	Sui	ʔau4	Libo
	S. Kam	qəu4	Rongjiang
	Maonan	fiu4	Huangjiang
	Mulam	hu3	Luocheng

Note: See Figure 4 for a tree diagram of the language group. The languages are sorted in the order of decreasing similarity to the proto-Tai word of "rice"- *khau [C3]. In other words, going down the list, you are going in a general direction from a pronunciation most similar to proto-Tai (at periphery) to one most different (at origin).