Demand for Coal: The Problem of Aggregation

Thomas H. Stevens, Martin J. Blake, and Lawrence G. Williams

This paper demonstrates the differences that result from estimating coal demand functions using highly aggregated regional data rather than less aggregated state data. At first glance the coal demand functions based on regional data appear to explain well and seem useful in policy making. However, coal demand functions based on state data show differences in the demand structure in each state of the region. Policy decisions that may seem appropriate based on aggregated regional data may have much different impacts on individual states in the region.

Coal is a potential source of energy to satisfy increased energy demand over the next decade. Major adverse environmental and economic impacts may, however, be associated with increased coal production. Among these are: (a) land subsidence; (b) surface disturbance; (c) water and air pollution and (d) "boom town" phenomena.

In response to a general concern over the relationship between increased coal development, environmental quality, and the demand for public services, several econometric studies of the demand for coal have recently been published. The results indicate that the demand is price inelastic. Therefore, it is concluded that more stringent strip mine regulations or increased severance taxes would have a negligible effect upon coal output and employment.

These conclusions are, however, based upon econometric coal demand functions that are estimated using highly aggregated data. In this paper it is suggested that coal demand functions based upon such data may not be adequate for the formulation and evaluation of public policy. As an example, a case study of coal demand in the Four Corners Region of New Mexico, Colorado, Utah, Arizona and Nevada is developed. Conclusions for the formulation of public policies are then presented and discussed.

Thomas H. Stevens, Martin J. Blake and Lawrence G. Williams are Assistant Professor, University of Massachusetts, Assistant Professor, New Mexico State University and graduate student at New Mexico State University, respectively.

Previous Research

Several coal demand studies have recently been published [Goldstein and Smith; Libbin and Boehlje; Lin, Spore and Nephew; and Reddy]. As shown in Table 1, relatively consistent estimates of price and cross price elasticities have been obtained.

An analysis of these studies, revals, however, that coal demand contains a number of features which are difficult to model. First, the data are not extensive and are of relatively low quality. Second, specification errors may result due to the existence of price regulation, electricity pooling arrangements, and related factors. Of particular interest here is that little attention has been devoted to potential problems associated with aggregation.

In order to provide a frame of reference, a brief synopsis of each demand study cited above is necessary. Reddy used a cross-sectional data base consisting of a sample which included only those steam-electric generating plants equipped to utilize coal and one or more of the other fossil fuels. The demand for steam coal was postulated as a

TABLE 1. Estimated Price and Cross Price Elasticities of Coal Demand

Study	Region	Price Elasticity for Coal	Cross Price Elasticity Natural Gas	Oil
Reddy	U.S.A.	−.371 to −.974	.496 to 1.583	1.035 to 1.958
Lin, Spore and Nephew	Appalachia	−1.65	.144	.216
Goldstein and Smith	U.S.A.	−.480 to −.316	−.937 to −.550	.283 to .260

function of coal price, oil price, gas price and net generation of electricity by steam electric generating plants. The generating plants were aggregated according to size and fuel combinations, and a separate demand equation was estimated for each grouping. The results indicated price inelastic demand for steam coal, with the elasticity values ranging from −.371 to −.974.

Goldstein and Smith examined the price, output and employment effects of increased land reclamation requirements. Coal used for electricity generation was aggregated with that used for steel production. The results indicated that reclamation standards would have only slight effects upon output and employment.

Lin, Spore and Nephew (LSN) also focused upon the probable impacts of increased reclamation requirements on mining costs, coal prices, production and employment. The demands for steam, coking, industrial and export coal were estimated separately. The resulting demand for steam coal was both price and cross-price inelastic. The price inelasticity was attributed to difficulty of short run interfuel substitution. However, the prices of fuel inputs were weighted by their consumption in order to account for the effect of oil and gas price regulation. Moreover, the price of coal was weighted by steam electric generation in each state of the region. Consequently, the respective parameter estimates cannot be readily interpreted as elasticities. That is, the ratio representing the price variables may change, thus indicating a change in the demand for coal, but the resulting change in demand cannot be attributed solely to a change in the price of coal or the substitute fuel.

Of greater importance is that the LSN demand model contained variables for prices of coal, oil, and natural gas. Yet, Gordon [1975, 1976] found that "No more than two fossil fuels have been major potential suppliers in any one region." In the Appalachian region, the two major competing fuels differ from state to state. Aggregating state data with respect to all three fuels, when in essence one fuel may not be available in any one specific state, may result in a misspecified model.

In summary, each of these models employed a different type and degree of aggregation. Moreover, each used highly aggregated regional or national data. It is hypothesized here that such models may not be adequate for the formulation or evaluation of policy from a state or local perspective. This inadequacy is partially because the structures of both the electrical and coal industries differ among states and regions.

The Demand Models

For the purpose of testing this hypothesis, the demand for coal in the Four Corners Region was estimated using two different econometric models. The first was a single coal demand function representing the entire Four Corners Region while the second contained separate coal demand functions for each state within the region.[1]

The Regional Demand Model

The regional econometric coal demand model contained two fuel price variables: coal and natural gas.[2] The total quantity of elec-

[1]New Mexico, Colorado, Utah and Arizona and Nevada combined.

tricity generated within the region was included as a third independent variable. The model was specified as:

$$\ln Q_{sc} = B_o + B_1 \ln P_c + B_2 \ln P_g + B_3 \ln E + u_t.$$

Where Q_{sc} = quantity of coal consumed for electric power generation, in thousand tons; P_c = average price of coal as consumed by electric utilities, in cents per million BTU and deflated using the GNP implicit price deflator; P_g = average price of natural gas as consumed by the electric utilities, in cents per million BTU and deflated using the GNP implicit price deflator; E = total kilowatt hours of electricity generated in the Four Corners Region; and u_t = an error term.

This specification assumes strong separability in the production of energy (BTU's) such that: BTU = $F_1(C) + F_2(G)$ where C and G represent coal and natural gas, respectively. It also assumes that there are no substitutes for energy (BTU's) in the production of electricity. As shown by Goldstein and Smith, this implies that the production of electricity, E, requires a certain quantity of BTU = f(E). The quantity of BTU's consumed is, therefore, a function of output alone and not of labor or capital prices. We also assume that the quantity of electricity produced, E, is independent of fuel prices [Goldstein and Smith, Reddy]. To the extent that these assumptions are violated, a simultaneous equation system specification would be required. However, since our objective is to explore the problem of geographical data aggregation, a model specification consistent with that used by Goldstein and Smith, Lin, Spore and Nephew, and Reddy was utilized in this study.

Aggregated time series data were not readily available for the region. Therefore, the individual state data for the 1963-74 period were aggregated to obtain data for the entire region, giving a total of 12 observations. The ordinary least squares (OLS) results were as follows:

$$\ln Q_{sc} = \underset{\substack{(-4.383)\\ [1.6339]}}{-7.1610} - \underset{\substack{(-1.011)\\ [.3957]}}{.4000}\ln P_c$$

$$+ \underset{\substack{(1.308)\\ [.2333]}}{.3053}\ln P_g + \underset{\substack{(20.551)\\ [.0750]}}{1.5407}\ln E$$

The values in parentheses are calculated t-values and the values in brackets are standard errors of the estimates. The electricity-generating parameter estimate was significant at the .01 level while the natural gas price and the coal price variables were significant at approximately the .30 level. The R^2 value equaled .984 and the Durbin-Watson (DW) statistic was 1.68. The Durbin-Watson test was, therefore, inconclusive regarding the existence of autocorrelation of the disturbance terms at the 95 percent level.

The results indicate both price and cross-price inelasticity. These findings are generally consistent with the studies reviewed above (Table 1). From a macro (regional) perspective it, therefore, appears that policy-induced mining-cost increases may have a relatively small effect upon coal output or employment. However, this inference may not be valid from a micro (state or local) perspective for several reasons.

First, regional average price data obtained from the Bureau of Mines were used. Depending on the size of coal shipments, transportation costs and hence prices could vary substantially within the region and from plant to plant. Consequently, the impact of policy-induced increases in the average price of coal may vary substantially within the region. Second, the data on average coal price include both contract and spot prices. The average price, therefore, may not represent the market clearing price for any individual state or firm. Finally, the structure of the industry varies within the region. For example, although coal-gas competition has historically

[2]Historically, coal-gas competition has prevailed within the region. Although fuel oil was extensively used in Utah prior to 1973, Utah produces less than six percent of the total energy generated within the region.

prevailed within the region, fuel oil was extensively used in Utah prior to 1973.

Individual State Demand Functions

In order to investigate the implications of using aggregated regional data, separate demand functions were estimated for each of the states within the region (Arizona and Nevada combined, Colorado, New Mexico, and Utah). Disaggregation to a higher degree would be desirable, but was not possible due to the lack of available data. The OLS estimates for each state are given in Table 2.

As shown in Table 2, the R^2 values are high for each relationship, but two parameter estimates exhibited the wrong sign and several were statistically insignificant. Specifically, the price of coal was found to be insignificant and to have the wrong sign in both the Arizona-Nevada and Utah functions.[3] The Durbin-Watson test indicated no significant autocorrelation for the Arizona and Nevada and New Mexico models at the 95 percent level, with inconclusive results for Utah and Colorado.

Moreover, the effect of aggregation may depend not only upon the aggregation procedures adopted, but also upon the method of statistical estimation. For coal demand, the OLS technique may generate unreliable and/or inefficient parameter estimates for each state because of interactions in industry structure which are not reflected by the independent variables in the model. For example, an electricity pooling arrangement comprising two or more states would be subject to control by each state's regulatory commission.

Pooling cross section and time series data using seemingly unrelated regressions is one way to account for the effects of such unquantifiable factors in the parameter estimates. Seemingly unrelated regressions estimation is appropriate when the variables that are neglected in a set of equations are partly the same or to some degree correlated. This would result in a correlation of the disturbance terms among the equations. Although

[3]High R^2 values associated with insignificant parameter estimates may result if significant multicollinearity exists. Multicollinearity did not appear to be substantial in the Arizona-Nevada function. Multicollinearity between the price of coal and the price of gas was, however, present in the Colorado function. This may explain the high standard error associated with the coal price variable in Colorado.

TABLE 2: Estimated Coal Demand Equations, OLS Regression Technique,[a,b] States

State	Estimated Equation					R^2	DWS
Arizona & Nevada[d]	$\ln Q_{sc} =$	−21.56	+ .23 $\ln P_c$	+ 1.08 $\ln P_O$	+ 2.44 $\ln E$	.98	2.05
		(− 7.49)	(.46)[c]	(1.92)	(18.21)		
		[2.87]	[.49]	[.56]	[.13]		
Colorado	$\ln Q_{sc} =$	− 1.72	− .82 $\ln P_c$	+ 1.01 $\ln P_g$	+ .99 $\ln E$	.93	1.24
		(− .51)[c]	(− .62)[c]	(1.38)	(6.74)		
		[3.41]	[1.32]	[.73]	[.15]		
New Mexico	$\ln Q_{sc} =$	0.31	− 1.03 $\ln P_c$	+ 0.15 $\ln P_g$	+ 1.11 $\ln E$	.99	1.91
		(.12)[c]	(−1.73)	(.62)[c]	(6.61)		
		[2.66]	[.59]	[.25]	[.17]		
Utah	$\ln Q_{sc} =$	−12.49	+ .45 $\ln P_c$	+ .45 $\ln P_g$	+ 1.95 $\ln E$	.85	1.50
		(− 2.09)	(.71)[c]	(2.18)	(2.77)		
		[5.98]	[.62]	[.21]	[.70]		

[a]Figures in parentheses are calculated t-values.
[b]Figures in brackets are standard error values.
[c]Insignificant at the α = .20 level.
[d]Price of oil used as substitute fuel (PO).

the equations seem unrelated, they are in fact disturbance-related [Theil]. In the case of coal demand in the Four Corners Region, the disturbance terms may be correlated because of similar institutional relationships, regulations and other common factors that affect all states in the region. Thus, seemingly unrelated regressions estimation should give more efficient estimates of the parameters in this case.

The results obtained from the seemingly unrelated regressions technique are presented in Table 3. This technique produced more efficient parameter estimates than did OLS, as evidenced by the lower standard errors on the estimated coefficients and by only one sign that contrasted with theory. Given the results in Table 3, the demand estimates based on regionally aggregated data may not be appropriate for policy evaluation. For example, the coal price elasticity of demand is estimated to be −1.33 in New Mexico, −1.52 in Colorado and −.07 in Utah (Table 3). These do not closely correspond to the regional price elasticity estimate of −.40 based on aggregated data.

The results tend to indicate that policies formulated from a regional perspective may cause highly variable responses from specific areas within the region. That is, a policy may initially look "good" for the region but turn out "bad" for almost every state within the region.

Impacts of Using Regionally Aggregated Data

It is important to realize that these two quite different sets of results are based on the same data. Individual state data were summed to get the regional quantity of coal consumed and total kilowatt hours of electricity generated. A weighted average of individual state prices was taken to obtain the regional average price of coal and the average price of natural gas. The reason that the regional model seems to have greater explanatory ability than the separate state models is that the variability of the data used in estimating the regional model has been reduced in the process of aggregation. Some of the individual differences reflected in the state data tend to cancel when aggregated into regional figures.

Equations estimated using aggregated data may seem to perform better simply because there is less variation to explain. However, using aggregated data can be harmful if the estimated equations are used as a basis for policy decisions. The individual differences often are important in policy formulation. For example, an increase in reclamation requirements may lead to only a slight reduction in regional output. However, output may fall significantly in New Mexico and Colorado and very little in Arizona and Nevada and Utah.

TABLE 3: Estimated Coal Demand Equations Using the "Seemingly Unrelated Regression" Technique[a,b]

State	Estimated Equation											
Arizona & Nevada	$\ln Q_{sc}$	=	−20.35	−	0.01	$\ln P_c$	+	0.72	$\ln P_g$	+	2.52	$\ln E$
			[2.68]		[.45]			[.52]			[.12]	
Colorado	$\ln Q_{sc}$	=	− 0.39	−	1.52	$\ln P_c$	+	1.33	$\ln P_g$	+	.98	$\ln E$
			[2.94]		[1.11]			[.63]			[.13]	
New Mexico	$\ln Q_{sc}$	=	1.59	−	1.33	$\ln P_c$	+	.22	$\ln P_g$	+	1.04	$\ln E$
			[2.53]		[.56]			[.24]			[.16]	
Utah	$\ln Q_{sc}$	=	−10.77	−	.07	$\ln P_c$	+	.57	$\ln P_g$	+	1.89	$\ln E$
			[4.97]		[.48]			[.17]			[.60]	

[a]Figures in brackets are standard error values

[b]R^2 and "t" tests are not comparable with OLS estimates.

That is, the structure of the electrical industry in each state comprising a region may be quite different. In fact, in this case, Chow's test was used to determine whether the state estimates can be regarded as belonging to the same regression model. The results indicated that the relationships are not the same at the 95 percent confidence level.

Thus, the structure of the electrical industry in these states is basically different. A policy that may work well on one structure may work poorly on another. Alternatively, a policy that seems appropriate for an average structure may perform poorly in terms of the structure's individual components.

Conclusions

Coal demand functions based on nationally or regionally aggregated data, such as those presented in the previous literature, may not be adequate to describe specific situations within a region. Even though a policy may not appear to significantly affect coal output at the regional level, significant shifts within the region may result. Evidence has been provided which indicates that econometric coal demand functions based on regionally aggregated data may be of little value in determining the effects of coal-related policies upon the states in the Four Corners Region.

Yet as Grunfeld and Griliches point out, "in practice we do not know enough about microbehavior to be able to specify microequations perfectly. Hence, empirically estimated microrelations... (in this case individual states)... should not be assumed to be perfectly specified... Aggregation of economic variables can and, in fact, frequently does, reduce these specification errors. Hence aggregation... may provide an aggregation gain." Consequently, further research is required on appropriate micromodel specification and on the type of aggregation employed. As demonstrated here, such research may substantially improve the quality of information utilized for policy formulation and evaluation.

The specific implications of the different sets of demand functions presented may be summarized in several ways. First, environmental quality legislation which results in increased coal prices will tend to have the greatest impact upon coal output and employment in Colorado and New Mexico. Consequently, geographical variations in environmental policy may be warranted. Second, the results provide a partial justification for existing variations in coal severance tax rates between the states within the region. Third, the short-run effect of natural gas price deregulation upon the coal industry is likely to be relatively insignificant as indicated by the cross-price elasticity values reported in Tables 2 and 3.

The principal policy implication is, however, that econometric estimates based upon highly aggregated data may fail to account for economic and physical characteristics unique to each geographical area within the aggregate. Evidence has been provided which indicates that the states within the Four Corners Region must be treated as separate entities in the formulation of coal development policies. If this is not done, significant redistribution may result within the region. Geographical variation in coal development policies is, therefore, advocated as a means of avoiding or enhancing redistributional consequences.

References

Edison Electric Institute, *Statistical Year Book of the Electric Utility Industry*, New York, (1963-1974 issues).

Federal Energy Administration. *Final Report, Short-term Coal Forecast*, 1975-1980, Washington: Office of Coal, August 1975.

______, *Project Independence Report*, U.S. Government Printing Office, Washington, D.C., November 1974.

Goldstein, Morris, and Robert S. Smith, "Land Reclamation Requirements and Their Estimated Effects on the Coal Industry," *Journal of Environmental Economics and Management*, December 1976.

Gordon, Richard L., *Marketing Prospects for Western Coal,* Pennsylvania State University, prepared for the National Science Foundation, December 1976.

———, *U. S. Coal and the Electric Power Industry,* Resources of the Future, Washington, D.C., Johns Hopkins University Press, 1975.

Grunfeld, Y. and Z. Griliches, "Is Aggregation Necessarily Bad?," *Review of Economics and Statistics,* 42(1960): 1-13.

Libbin, James D. and Michael D. Boehlje, "Interregional Structure of the U.S. Coal Economy," *American Journal of Agricultural Economics,* 59(1977): 456-466.

Lin, William, Robert L. Spore, and Edmund A. Nephew, "Land Reclamation and Strip-Mined Coal Production in Appalachia," *Journal of Environmental Economics and Management,* October 1976, pp. 236-252.

Reddy, Nallapu N., "The Demand for Coal: A Cross-Sectional Analysis of Multi-Fuel Steam Electric Plants," *Industrial Organization Review,* University Publications Division of Physical Biological Sciences Ltd., Volume 3, Number 1, 1975.

Theil, Henri, *Principles of Econometrics,* John Wiley and Sons Inc., New York, 1971.

U. S. Department of Interior, Bureau of Mines, *Minerals Yearbook,* Volume on Fuels, Washington, D.C., U.S. Government Printing Office, (1953-1974 issues).

Walsh, R. G., *Some Benefits and Costs of Strip Mining Western Coal Resources,* Las Cruces, New Mexico, Department of Agricultural Business and Agricultural Economics, New Mexico State University, April 1974.