[*]Factors Influencing Inter-Modal Facility Location Decisions: Comparison of Different Empirical Estimation Procedures

Steven Peterson

School of Economic Sciences
Washington State University
103 Hulbert Hall
Pullman, WA 99164-6210
509-335-2504
skpeterson@wsu.edu

ABSTRACT

This study examines several different empirical models of inter-modal terminal characteristics using a combination of available data from the Surface Transportation Board waybill survey and Oak Ridge National Laboratories national freight transportation database in the estimation process. Empirical models included ordinary least squares, fixed-effects models, and random effects models. The data also presented issues of heteroskedasticity and autoregressive processes which are addressed in the paper. Due to the spatial context of the intermodal facilities, the data was also evaluated for spatial interactions and autocorrelation. Substantial evidence of underlying spatial relationships in the data were observed and noted. The results of this analysis identify those empirical variables and characteristics that contribute to the economic and operational sustainability of inter-modal facilities, and the relative importance of each factor. These assessments may then be used to evaluate infrastructure investment decisions, particularly by public entities. This methodology offers an unbiased framework for identifying and evaluating public and private benefits resulting from such investments, and serves as a guide for transportation policy involving inter-modal freight.

[*] The author wishes to thank the valuable assistance provided by Jim Maxwell of Burlington Northern Santa Fe Railroad, Mac Frampton at the Surface Transportation Board and Bruce Peterson at Oak Ridge National Laboratories, Center for Transportation Analysis. Further thanks are due to the helpful comments of Eric Jessup, John Yoder and Yu-can Liu. The usual caveats apply.

INTRODUCTION

The inter-modal terminal is the node linking the highway, rail, and, in many cases, port transportation networks together. At these facilities, goods are transferred between truck and rail (or ship-to-rail and ship-to-truck) for shipment to domestic markets, or through ports, to international destinations. As with other transportation infrastructure investment, public entities are interested in inter-modal transportation in facilitating the efficient movement of goods to and from municipalities, states and regions. As such, the potential exists for public agencies to provide public monies for use in infrastructure investment in inter-modal freight terminals. Plans for investment in such facilities raise three key questions: Will such an investment generate a sufficient, sustainable return? And, is public funding justified by some measure of public benefits resulting from the investment? What basic criteria need to be identified in evaluating potential locations for intermodal terminal sites?

Problem Statement

When examining freight mobility, inter-modal transportation allows the efficiencies of each transport mode to be realized across the entire move. As noted in a forthcoming publication by Casavant, et al, (2004), trucks, with low costs of collection and local distribution, but high variable costs over long distances, are combined with rail, which has high terminal and infrastructure costs, but low variable costs of movement over long distances. These inter-modal freight movements are highly dependent on the structure, location and efficiency of the transfer facilities servicing such moves. As a result, a need exists for a rigorous conceptual and mathematical model to detail the specific impacts of these various factors on the efficiency of existing, and future, inter-modal terminals.

Objective

The primary objective of this research project is to examine various candidate empirical estimation models and compare/contrast the results of these different models. This is accomplished by examining the relative impact of individual location and operational factors on the economic performance of inter-modal facilities. An important consideration is the availability of data in construction and estimation of the models.

Summary of Data and Research Methods

The method of investigation in this project was to use data from a variety of sources in order to determine the best empirical model for estimation. The data used in the study was obtained from Oak Ridge National Laboratories, Center for Transportation Analysis and from the Surface Transportation Board of the US Department of Transportation. Databases were merged together using a combination of data queries in a standard database management program and spatial queries in ArcGIS software.

The data obtained did not lend itself to estimation procedures characterized by binary or multinomial choice. As a result, logit, probit and Tobit models were excluded from the study.

The first action taken was to examine the structure of the data and determine the best modeling approach. In addition, the data was approached from both cost minimization and profit maximization frameworks. This initial specification search involved the use of multivariate adaptive regression splines in order to determine the contributory power of different variables identified in the data, and to identify potential interaction variables. During the second part of the study, the identified explanatory variables as well as characteristic indicator variables were used in ordinary least squares estimation and general linear modeling for fixed and random effects, autoregression, heteroskedasticity and spatial processes.

THEORY AND RELATED LITERATURE

Currently, there does not exist a wide body of literature addressing empirical estimation of the factors and characteristics that influence and contribute to inter-modal facility success. Most of the literature is designed to examine freight traffic flows and facility locations within a theoretical freight transportation network, or with the theoretical operating characteristics of inter-modal facilities. Other studies examine costs and prices within freight networks, but not as characteristic indicators of facility viability. Another strand of literature discusses the dynamics of location decisions within the context of manufacturing plants, or within the context of international trade.

Within the context of networks, several model types have been analyzed and implemented as in Sheu (2003), Melkote and Daskin (2001), and Barber (1975). Most of these models address issues related to theoretical location decisions and the creation and implementation of ideal transportation networks. From these base theoretical models, additional models addressing economies of scale in costs in Horner and O'Kelly (2001), flows (O'Kelly and Bryan, 1998; Fernandez, et al, 2003), and elasticities (as in Beuthe, et al, 2001) have been developed. These approaches are generally referred to as "hub network" problems. While hub networks are useful in describing inter-modal facility locations, they typically require a large number of decision variables that can be prohibitively expensive in empirical estimation. A good survey of mathematical models in freight transportation network planning can be found in Friesz (2000).

Spatial processes underlying transportation and freight networks have been noted by Nierat (1997) amongst others, though their focus is on market areas and the spatial extent of competition between terminal facilities.

One promising approach has been developed by Arnold, et al., (2004) using data for European rail and highway networks and applied to inter-modal freight terminal locations in the Iberian Peninsula. This system, known as Intermodal Terminal Location Simulation System (ITLSS), approaches the question of potential terminal location based upon flows moving through the existing transportation network and optimizing costs. This study integrated two European transportation databases to build an integrated multi-modal transportation database. However, within the context of inter-modal terminal infrastructure, the model is concerned with estimating the increase in transport supply resulting from a new terminal, or optimizing the existing inter-modal network in order to achieve transportation cost efficiencies. Another approach, developed for the analysis of grain terminals, is that of McCarl, Hilger and Uhlig (1985). This study produced a mixed-integer mathematical programming model of cost minimization that seeks to create determine the lowest cost providers of terminaling services to grain shippers.

Terminals may be viewed as collection and processing conduits for railcars; railcars are collected, stored and assembled into trains for shipment and delivery to other terminals. The terminals also route the railcars through to their final points of delivery. With inter-modal terminals, the terminals also serve as loading and unloading centers for freight trucks. As such, inter-modal terminals serve both as cost centers and profit generators, and the estimation procedure will consider both cost and profit as dependent variables. This is analogous to Casavant, et al (2004) who identify cost minimization and profit maximization (internal rate of return) as important considerations in determining economic viability where

$$Economic\ Viability = f\ (x_i\ )$$

and $f(x_i)$ is some vector of characteristics.

The empirical models being estimated in this study are designed to use the survey and transportation databases in order to construct an econometric estimation of those factors that contribute to the long-run economic viability of an inter-modal terminal. Several model specifications were considered and analyzed to aid in identifying the best explanatory model. In the context of this report, "best" refers to the model that succeeds in best identifying economic characteristics of cost minimization and profit maximization of inter-modal terminals.

Models that were considered other than ordinary least squares were the fixed effects model which has the theoretical form (Greene, 2003)

$$y_{it} = \alpha_i + \beta^{'} x_{it} + \varepsilon_{it}$$

where α_i may be a group-specific intercept term and ε_{it} is iid normal. A possible problem with this model estimation is the high probability of loss of degrees of freedom as there are 483 observations and 145 rail nodes (SPLC's). Another specification that was considered was the random effects, or error-components, model, which includes a group-specific error term. The model then has the theoretical form of

$$y_{it} = \alpha + \beta^{'} x_{it} + \upsilon_i + \varepsilon_{it}$$

where υ_i is the group-specific component of the error term , which can be rewritten as

$$\omega_{it} = \upsilon_i + \varepsilon_{it}$$

This model assumes that there are zero covariances across time periods and groups. Due to the unbalanced nature of the cross-sectional panels used in the estimations, the likelihood of heteroskedasticity was virtually assured. Therefore a random effects model that addresses heteroskedastic panels was estimated under feasible generalized least squares regression (FGLS).

Other possible effects resulting from estimation of the data may be the presence of autoregressive errors of order 1, i.e., AR (1), and/or a spatially autoregressive system. Temporal autocorrelation is a common feature of time-series data and, by extension, cross-sectional data. In regards to inter-modal terminals, it is reasonable to expect that large volumes of cars moving through a facility will be followed the next year by large volumes. In fact, the number of observations in the 1% Waybill Survey has increased substantially between 1988 and 2002 on the order of approximately 70%. As a result, errors in the estimation results are likely to persist over time. In order to address this possibility, a model was estimated using heteroskedastic panels and autoregressive errors of order 1; however, following Mizon (1995) no further correction for autocorrelation was done. Models of spatial autocorrelation follow Anselin (1988), which takes the form

$$y = \rho W_1 y + X\beta + \upsilon$$
$$\upsilon = \lambda W_2 \upsilon + \varepsilon$$
$$\varepsilon \sim N(0, \sigma^2 I_n)$$

where the W's are known weighting matrices defined either contiguously or by some measure of distance (as in this case). The other model components are the same as the FGLS model.

DATA AND ESTIMATION

In order to conduct this research data relevant to inter-modal transportation was sought. A comprehensive database of the geographical locations of inter-modal facilities in the United States and a database of rail nodes and the rail network were obtained from Oak Ridge National Laboratories. This information is also available from the National Transportation Analysis Database maintained by the US Department of Transportation's Bureau of Transportation Statistics. Data and information regarding commodity and rail shipment flows were obtained from the Surface Transportation Board, which conducts an annual survey of all railroad waybills.

Originally, the estimation process sought to use the inter-modal facilities database created by the Center for Transportation Analysis and match up railcar flows as estimated in the 1% Waybill Survey from the Surface Transportation Board. Waybill data from 1984 through 2002 was obtained from the STB for this purpose. However, no obvious connecting variable exists between the two datasets. The waybill data files are organized around freight movements between rail stations. As a result, the files contain information on the originating and terminating rail nodes such as census regions, FIPS, the freight station accounting code (FSAC), and the standard point location code, or SPLC. The SPLC was the most useable identifier since the FSAC is maintained by the individual railroads for internal accounting purposes. For facilities that share service at a rail node, use of the FSAC would have led to potential duplication of observations. The SPLC would uniquely identify a specific rail node.

This left the SPLC as the most promising candidate for linking the two databases. Unfortunately, the CTA's inter-modal facilities database did not include a SPLC identifier for each terminal. Conversations with industry personnel confirmed that more than one inter-modal terminal could have the same SPLC, in much the same way that two factories could share a ZIP code. Moreover, the SPLC's were not inter-modal specific; other types of rail equipment moved through the rail node.

In order to address these issues, the data was queried in order to separate inter-modal moves from other types of rail movements. Identifying waybill shipments that had an intermodal equipment flag for the originating and terminating SPLC accomplished this task. Within the STB data, determination of inter-modal facilities came from the "2002 Surface Transportation Board Carload Waybill Sample: 900-byte Master File Record Data Element Description." The data files contain up to 193 variables, however, the identification of observations for use in the study relied on information contained in the fields "Origin Intermodal Flag" and "Destination Intermodal Flag." Any waybill observation that satisfied the flag criteria for both the origin and destination was selected. This criterion selection noted which rail nodes were capable of handling an inter-modal move and removed rail nodes that could not handle inter-modal moves from consideration in the estimation phase of the study. The selection criteria are designated by inter-modal equipment type as follows:

"1" – Circus type ramp
"2" – Overhead crane
"3" – Side lifter
"5" – Stack train

Any observations that had other flag codes were excluded from the study.
Non-intermodal moves may be included in the observations, but they are moving through rail nodes identified as inter-modal and contributing to measures of capacity and throughput for the rail node. The results were then totaled for each observational year, such that each variable is the sum (or average in the case of miles traveled) of all activity in a specific rail node over one year. The CTA maintains an Intermodal Transportation Network that contains the inter-modal facilities database. This file identifies individual facilities that originate and terminate inter-modal shipments. The only locational identifier in this database is a latitude-longitude measure placing the facility in a geographical coordinate system (NAD 1927). The Bureau of Transportation Statistics of the US Department of Transportation, maintains the National Transportation Analysis Database, which uses the rail network information constructed by CTA. This database includes a rail network that also identifies rail nodes by SPLC with a geographical coordinate location within the United States. By performing spatial joining queries in ESRI's ArcGIS geographical software, we were able to map these rail node locations with the inter-modal terminals using a minimal distance criterion (based on conversations with personnel at railroads and inter-modal terminals, most terminals are within 1 mile of a SPLC node). This then allowed us to determine the railcar flows for the inter-modal terminals from the waybill survey accordingly.

After these steps, the file contained 934 observations for 203 rail nodes between 1994 and 2002. Observations were not included for years prior to 1994 as the inter-modal flag was not part of the STB's waybill survey. Also, information for 1997 was missing information on costs and revenues. As a result, observations prior to 1998 were dropped and a cross-sectional panel of information from 1998 through 2002 was retained with 493 observations and 144 rail nodes.

Using the STB expansion factor from the waybill survey, estimates of total carloads, tonnage and miles were created for each inter-modal location. Also, summing revenue, transit charges and miscellaneous charges created figures for total revenues. Subtracting the total variable costs from the total revenues also created an estimated profit per move. Multiplying total tons by the total miles created an additional tonmiles variable. Total miles were then averaged over all moves to approximate the average shipment length for moves serviced at each node.

Also, within the GIS system, information regarding the population of the census region where the rail node was located was obtained, as well as distances to highways and marine or inland waterway ports. Variables for the inverse distances of the nodes and highways or ports were then created. The next step involved identifying characteristic variables for the rail nodes, and by extension, the intermodal facilities contained in the waybill data. These variables were: the number of commodities serviced by the facility as determined by the different STCC's (Standard Transportation Classification Code), the number of railroads servicing the node and the number of terminating nodes each SPLC was connected with. Finally, variables for profit per carload and cost per carload were created. These last two variables were the dependent variables used in initial estimation procedures. Summary statistics for these observations can be found in Table 1.

Table 1. Summary Variable Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
avgmiles	484	1230.529	663.0951	1	3094
carloads	484	27137.79	80718.84	31	677239
tonmiles	484	667.5985	2209.839	.0067	18595.9
tons	484	446761.8	1215053	160	1.04e+07
ppc	484	187.3876	553.1618	-1240.61	4306.53
cpc	484	931.8206	529.8929	0	3570.5
invwdist	484	1.353884	2.883658	0	21.43
invhdist	484	3.655393	6.97217	.01	49.46

The following variables were then determined to be candidates for the adaptive regression splines estimation procedure:

Dependent Variables

Cost per Carload	cpc
Profit per Carload	ppc

Independent Variables

Total Carloads	carloads
Average Total Miles	avgmiles
Total Tons	tons
Total TonMiles (Millions)	tonmiles
Population	pop
Inverse Distance to Highway	invhdist
Inverse Distance to Port	invwdist

Categorical Variables

Number of Terminal Connections	termcnt
Junction Frequency Count	concnt
Number of Commodities (STCC)	comnum
Number of servicing Railroads	rrs

These variables were then input into a statistical package MARS™, or Multivariate Adaptive Regression Splines, which uses a step-wise regression procedure to identify variable contributions and interaction variables that have explanatory power.

Cost Model

The estimation results that provided the lowest standard errors identified the following variables as having the greatest explanatory importance for cost per carload: avgmiles, pop, carloads, tonmiles and termcnt. Also, rrs was identified as an important categorical variable and invhdist as an important weighting factor.

The relative variable importance as determined in the step-wise procedure is provided in Table 2.

Estimation results also indicated some higher-order interactions between the variables that were of explanatory importance.

Table 2. Variable Importance for Cost per Carload (cpc)

Variable	Cost of Omission	Importance	
AVGMILES	106000.188	100.000	||||||||||||||||||||||||||||||||||||||||||||||||||
POP	58310.043	61.886	|||||||||||||||||||||||||||||||
CARLOADS	51468.902	54.266	|||||||||||||||||||||||||||
TONMILES	39972.656	38.176	|||||||||||||||||||
TERMCNT	38263.988	35.161	||||||||||||||||||

These variables are identified below:

TonMiles x Population (Millions)	tmpop
Average Miles x Carloads	amcars
Carloads x Population	carpop
Terminal Connections x carpop	tccarpop
TonMiles x carpop	tmcarpop

Profit Model

Similar results were obtained for profit per carload, however the interaction variables were different in some instances. Variables of importance to profit per carload were identified as: avgmiles, carloads, invwdist, pop, tonmiles and tons. The results for tonmiles and tons having simultaneous importance was surprising, since this would indicated a probable collinearity problem in least-squares estimation. However the importance appears to be more relevant to the creation of the interaction terms and the tons variable was dropped in later estimation procedures. Table 3 provides a table of variable importance determined by the step-wise process for profit per car.

The count of servicing railroads and the number of terminal connections were identified as having categorical importance within the structure of the model. The variables invhdist was again identified as an important weighting variable.

Table 3. Variable Importance for Profit per Carload (ppc)

Variable	Cost of Omission	Importance	
AVGMILES	212831.938	100.000	||||||||||||||||||||||||||||||||||||||||||||||||||
CARLOADS	212494.578	99.825	||||||||||||||||||||||||||||||||||||||||||||||||||
INVWDIST	206829.469	89.935	|||||||||||||||||||||||||||||||||||||||||||||
POP	184410.797	51.959	||||||||||||||||||||||||||
TONMILES	112690.422	46.524	|||||||||||||||||||||||
TONS	97893.641	17.114	|||||||||

The higher-order variable interactions identified during the step-wise procedure were the following:

TonMiles x Population (Millions)	tmpop
Average Miles x Carloads	amcars

These are the same as the interactions noted for cost per carload model. Additional higher-order explanatory variables identified are:

Inverse Distance to Port x amcars	invwam
TonMiles x invwam	tminvwam
Population x tminvwam	ptminvwam
Average Miles x Population	ampop

Several dummy variables were then constructed based upon the results of the adaptive regression spline estimation procedure. These categorical variables reflect the characteristics of each rail node that were determined in the step-wise process to provide valuable explanatory variables. Characteristic indicator variables that were common to both the cost minimization model and the profit maximization model are:

D1	serviced by 2 railroads
D2	serviced by 3 railroads
D3	serviced by 4 railroads
D4	serviced by 5 railroads
D5	serviced by more than 5 railroads
D6	more than 100 terminal connections at the node
D7	75-99 terminal connections at the node
D8	50-74 terminal connections at the node
D9	25-50 terminal connections at the node
D10	more than 500 commodities handled at the node
D11	300-499 commodities handled at the node
D12	100-299 commodities handled at the node
D13	50-99 commodities handled at the node
D14	25-49 commodities handled at the node

The *termcnt* variable was reformulated as characteristic variables d6-d9 and applied to both models. An additional criterion identified during the adaptive regression spline procedure was a carload value that we have interpreted as a capacity or throughput threshold. For the profit model the threshold was at 3,757 carloads per year, and for the cost model it was 3,440. A dummy was created for each model that identifies carload counts greater than or equal to the threshold value.

D15	carload threshold; >= 3,757 for ppc, >= 3,440 for cpc

For the cost model, population was also identified with a threshold value of 62,527,588. A dummy for this variable was included in the profit model as well.

D16	population greater than or equal to 62,527,588

These characteristic indicator variables were then added to the previously identified independent variables and interaction terms and estimated using ordinary least squares regression. Encouragingly, all of the variables identified by the regression spline procedure are theoretically supported, except for the population variable. The interaction terms identified by the process can also be theoretically supported as indicators of higher-order processes of capacity, throughput and connectivity. Further estimation of these variable terms for movements over the entire rail network would be warranted to fully substantiate these results.

The results of the cost estimation model are presented in Table 4. Table 5 provides the results from the profit model estimation.

Table 4. Ordinary Least Square Results for Cost per Carload

Source	SS	df	MS		
Model	71329907.9	27	2641848.44	Number of obs =	484
Residual	64289941.3	456	140986.713	F(27, 456) =	18.74
Total	135619849	483	280786.437	Prob > F =	0.0000
				R-squared =	0.5260
				Adj R-squared =	0.4979
				Root MSE =	375.48

Table 5. Ordinary Least Square Results for Profit per Carload

Source	SS	df	MS		
Model	38718171.9	27	1434006.37	Number of obs =	484
Residual	109073999	456	239197.365	F(27, 456) =	6.00
Total	147792171	483	305987.931	Prob > F =	0.0000
				R-squared =	0.2620
				Adj R-squared =	0.2183
				Root MSE =	489.08

Population was dropped from both models since it was determined to be statistically insignificant. We would expect the population indicator variable to address any lingering estimation issues in this regard. Both models indicated the presence of heteroskedasticity using a Breusch-Pagan test. For the cost model the test statistic was chi2 (1) = 193.58, and for the profit model it was chi2 (1) = 204.12.

The presence of heteroskedasticity could be due to several factors; one, the nature of the data structure is quite similar to unbalanced cross-sectional panel data, and two, the possibility of model misspecification in the dependent variable, where the dependent variable should be estimated in log form. In order to analyze these probabilities, a graph of the estimated residuals v. predicted values was made for each model. Figures 1 and 2 are the resulting graphs.

The plots and histograms indicate that the heteroskedasticity problem is likely to be the result of model misspecification and that the variables should be log transformed and then run as a least-squares dummy variable fixed effects model.

Since heteroskedasticity may also be present due to the unbalanced nature of the cross-sectional panel, the model can be estimated as a fixed-effects model by adding dummy variables corresponding to the year of the observation. The dummies added to the model are:

Dy2002	for observations occurring in 2002
Dy2001	for observations occurring in 2001
Dy2000	for observations occurring in 2000
Dy1999	for observations occurring in 1999

The base year for analysis in the study is 1998.

Estimation then proceeded by implementing a fixed-effects linear panel data model for both the lognormal cost and profit functions. Due to the presence of negative values in the profit data, scaling the data for the profit model made transformation to lognormal possible.

Cost Minimization Model

A fixed-effects model was implemented with the log of cost per carload as the dependent variable and the previously noted independent variables also log transformed, except for the interaction variables. Also, dummy variables for the observation year were included to capture any information regarding changes over time or between years.

Figure 1. cpc residuals v. fitted

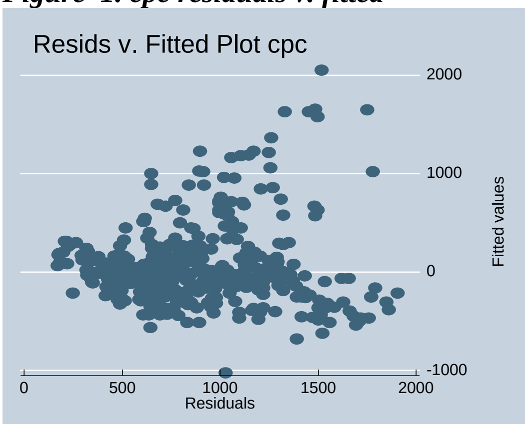
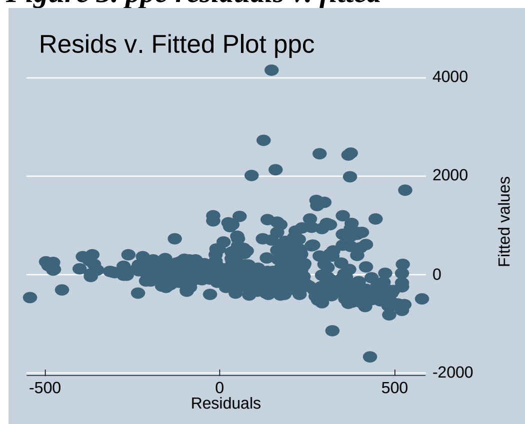


Figure 3. ppc residuals v. fitted



The structural format estimated was as follows:

$$\sum_{i=1}^{N} l\cos t = \alpha + \beta_1\left[\frac{1}{n}\sum_i lmiles\right] + \beta_2\sum_i lcars + \beta_3\sum_i ltmilese + \beta_4\sum_i carpop + \beta_5\sum_i tccarpop + \beta_6\sum_i tmcarpop + \beta_7\sum_i amcars + \beta_{i=1}^{16}(d_1...d_{16}) + \beta_y(dy1999...dy2002)$$

The results of the model estimated effects over 106 groups (the number individual rail node SPLC's in the observation file) and 483 observations. Summary regression statistics are found in Table 6.

Due to problems with multicollinearity and degrees of freedom D1 through D14 and D16 were dropped during the estimation procedure. Estimates for the other variables and constant are in Table 7.

Due to the large number of variables dropped in the fixed-effects estimation a random effects model was also run to compare against the fixed effects specification. Since the difference between the models has to do with the error terms, the empirical model specification is quite similar to the fixed effects model.

Table 6. Regression Statistics for Cost Minimization Fixed-Effects Model

```
Fixed-effects (within) regression               Number of obs      =       483
Group variable (i): splc                        Number of groups   =       145

R-sq:  within  = 0.5191                         Obs per group: min =         1
       between = 0.7641                                        avg =       3.3
       overall = 0.7272                                        max =         5

                                                F(10,328)          =     35.40
corr(u_i, Xb)  = 0.1270                         Prob > F           =    0.0000
```

Table 7. Variable Statistics for Cost Minimization Fixed-Effects Model

lcost	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
lmiles	.077291	.0545014	1.42	0.157	-.0299254	.1845074
lcars	-.4540581	.0416479	-10.90	0.000	-.5359888	-.3721274
ltmiles	.3775419	.0406632	9.28	0.000	.2975483	.4575355
amcars	2.76e-09	1.44e-09	1.91	0.057	-8.28e-11	5.60e-09
tccarpop	-3.00e-15	2.00e-15	-1.50	0.136	-6.94e-15	9.46e-16
dy1999	-.0182909	.0349473	-0.52	0.601	-.08704	.0504582
dy2000	.0745132	.0365205	2.04	0.042	.0026693	.1463571
dy2001	.049178	.0366199	1.34	0.180	-.0228614	.1212174
dy2002	.014544	.036958	0.39	0.694	-.0581606	.0872486
d15	-.0383621	.0714075	-0.54	0.591	-.1788365	.1021124
_cons	8.200405	.4958842	16.54	0.000	7.22489	9.17592

For the random effects estimation procedure the year dummy variables were dropped from the regression, so the model becomes

$$\sum_{i=1}^{N} l\cos t = \alpha + \beta_1\left[\frac{1}{n}\sum_i lmiles\right] + \beta_2\sum_i lcars + \beta_3\sum_i ltmiles + \beta_4\sum_i carpop + \beta_5\sum_i tccarpop + \beta_6\sum_i tmcarpop + \beta_7\sum_i amcars + \beta_{i=1}^{16}(d_1 ... d_{16})$$

Tables 8 and 9 detail the results of the random-effects estimation.

Table 8. Regression Statistics for Cost Random Effects Model

```
Random-effects GLS regression                   Number of obs      =       483
Group variable (i): splc                        Number of groups   =       145

R-sq:  within  = 0.5058                         Obs per group: min =         1
       between = 0.7795                                        avg =       3.3
       overall = 0.7462                                        max =         5

Random effects u_i ~ Gaussian                   Wald chi2(21)      =    812.45
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.0000
```

R-squared results were higher for the random-effects model, but the same variables were shown to be significant and close in value, although all of the dummy variables were shown to be statistically insignificant.

A Breusch-Pagan test for random effects and a Hausman specification test were run in order to compare the fixed effect and random effects models. The Breusch-Pagan results for testing the null of var(u) = 0, were rejected with a chi2(1) = 39.78. The Hausman test that the difference between the estimated coefficients in the model was not systematic (indicating the fixed effects model is a good specification) resulted in a chi2(4) = (b-B)'[(V_b-V_B)^(-1)](b-B) = 2.69. This would lead us to accept the null hypothesis of a fixed effects specification. However, a Wooldridge test for first-order serial autocorrelation was also done. The resulting test statistic was F (1, 80) = 8.197, which indicates the presence of autocorrelation in the data generating process. As a result, empirical estimation was then done as FGLS random effects with panel-specific heteroskedastic and AR (1) errors.

The FGLS model followed the same empirical specification as the random effects model above. Results obtained were more robust and provided more information regarding the dummy variable effects on the log cost dependent variable than any of the other estimated models. Results are given in Tables 10 and 11.

Table 9. Variable Statistics for the Cost Random Effects Model

lcost	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
lmiles	.1138283	.0339566	3.35	0.001	.0472747	.180382
lcars	-.4585273	.0286221	-16.02	0.000	-.5146256	-.4024291
ltmiles	.3656023	.0283328	12.90	0.000	.310071	.4211337
amcars	3.24e-09	9.54e-10	3.40	0.001	1.37e-09	5.11e-09
tccarpop	-3.33e-15	1.18e-15	-2.83	0.005	-5.64e-15	-1.03e-15
d1	-.0606717	.0552789	-1.10	0.272	-.1690163	.0476729
d2	-.0462213	.0782659	-0.59	0.555	-.1996197	.1071771
d3	-.0705685	.2038898	-0.35	0.729	-.4701852	.3290482
d4	-.0920403	.2055618	-0.45	0.654	-.494934	.3108533
d5	-.2027097	.4074171	-0.50	0.619	-1.001233	.5958131
d6	.0156458	.5762113	0.03	0.978	-1.113708	1.144999
d7	.014028	.5102735	0.03	0.978	-.9860896	1.014146
d8	.0368028	.1126274	0.33	0.744	-.1839429	.2575485
d9	.0652129	.0810616	0.80	0.421	-.093665	.2240908
d10	.2165665	.4757709	0.46	0.649	-.7159274	1.14906
d11	.1459552	.3019604	0.48	0.629	-.4458763	.7377867
d12	.0915077	.1029144	0.89	0.374	-.1102008	.2932162
d13	.0590669	.095692	0.62	0.537	-.128486	.2466198
d14	-.0447923	.0855975	-0.52	0.601	-.2125604	.1229758
d15	.0005467	.0605179	0.01	0.993	-.1180662	.1191596
d16	.1058809	.1167842	0.91	0.365	-.123012	.3347738
_cons	8.001979	.3092241	25.88	0.000	7.395911	8.608047

Table 10. Regression Statistics for Cost RE Model with Panel-Specific Errors

```
Cross-sectional time-series FGLS regression

Coefficients:  generalized least squares
Panels:        heteroskedastic
Correlation:   panel-specific AR(1)

Estimated covariances      =       106          Number of obs      =       444
Estimated autocorrelations =       106          Number of groups   =       106
Estimated coefficients     =        22          Obs per group: min =         2
                                                               avg =  4.188679
                                                               max =         5
                                                Wald chi2(19)      =  15440.61
Log likelihood             =  493.8436          Prob > chi2        =    0.0000
```

Profit Maximization Model

Empirical estimation of the profit maximization model proceeded in the same manner as the cost minimization model. Prior to estimation, the profit variable was scaled upwards in order to eliminate any negative values during the log transformation of the variable. The estimation results therefore present a scaled bias framework for estimation, but without loss of generality. Tables 13-15 present to profit maximization model estimations for the fixed effects and random effects models, without correction for heteroskedasticity. The empirical model form of the fixed effects model is much the same as the cost minimization model, although several of the identified interaction variables have changed.

The profit model for fixed effects has the form

$$\sum_{i=1}^{N} lspc = \alpha + \beta_1 \left[\frac{1}{n} \sum_i lmiles \right] + \beta_2 \sum_i lcars + \beta_3 \sum_i ltmiles + \beta_4 \sum_i tmpop + \beta_5 \sum_i tmcarpop + \beta_6 \sum_i amcars + \beta_{i=1}^{16} (d_1 ... d_{16}) + \beta_y (dy1999...dy2002)$$

Table 11. Variable Statistics for Cost RE Model with Panel-Specific Errors

lcost	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
lmiles	.210802	.0105027	20.07	0.000	.1902171	.231387
lcars	-.4730562	.0089088	-53.10	0.000	-.4905171	-.4555953
ltmiles	.3931265	.0094862	41.44	0.000	.3745338	.4117192
amcars	2.04e-09	1.98e-10	10.31	0.000	1.65e-09	2.43e-09
tccarpop	-2.16e-15	2.89e-16	-7.49	0.000	-2.73e-15	-1.60e-15
d1	-.0410579	.0115435	-3.56	0.000	-.0636827	-.018433
d2	-.043538	.0154987	-2.81	0.005	-.0739149	-.013161
d3	-.0806243	.0291989	-2.76	0.006	-.137853	-.0233956
d4	-.0595256	.2253114	-0.26	0.792	-.5011279	.3820767
d5	-.0896096	.226258	-0.40	0.692	-.5330672	.3538479
d6	-.048736	.2288964	-0.21	0.831	-.4973647	.3998928
d7	.0797256	.119428	0.67	0.504	-.1543489	.3138002
d8	.0397106	.0186917	2.12	0.034	.0030756	.0763456
d9	.0319579	.015132	2.11	0.035	.0022997	.0616161
d10	.1566769	.040954	3.83	0.000	.0764085	.2369453
d11	.112044	.0347208	3.23	0.001	.0439925	.1800954
d12	.028497	.0172807	1.65	0.099	-.0053726	.0623665
d13	.020875	.0156589	1.33	0.182	-.0098159	.051566
d14	-.0401933	.0207655	-1.94	0.053	-.080893	.0005063
d15	.0622714	.015892	3.92	0.000	.0311237	.0934191
d16	.1416368	.0329582	4.30	0.000	.07704	.2062337
_cons	7.32566	.0948447	77.24	0.000	7.139768	7.511553

while the random effects model has the empirical specification

$$\sum_{i=1}^{N} lspc = \alpha + \beta_1 \left[\frac{1}{n} \sum_i lmiles \right] + \beta_2 \sum_i lcars + \beta_3 \sum_i ltmiles + \beta_4 \sum_i tmpop + \beta_5 \sum_i tmcarpop + \beta_6 \sum_i amcars + \beta_{i=1}^{16} (d_1 ... d_{16})$$

Table 12. Regression Statistics for Profit Fixed Effects Model

Fixed-effects (within) regression		Number of obs =	484
Group variable (i): splc		Number of groups =	146
R-sq:	within = 0.3057	Obs per group: min =	1
	between = 0.0514	avg =	3.3
	overall = 0.1513	max =	5
		F(11,327) =	13.09
corr(u_i, Xb) = -0.4027		Prob > F =	0.0000

The profit maximization model indicates even more of the error variance attributable to possibly random effects in υ_i. R-squared results for both specifications are weak, although it is better for the random effects model. A Hausman test performed on the two specifications yielded a chi2 (8) = (b-B)'[(V_b-V_B)^(-1)](b-B)= -1653.30, which fails to meet the asymptotic assumptions of the Hausman statistic. The Breusch-Pagan test statistic for random effects, Var(u) = 0, was chi2(1) = 12.10. The null can then be rejected and we find that a random effects model has a better specification for the profit maximization model than the fixed effects alternative.

A Wooldridge test for serial autocorrelation was also performed on the log profit variable. The test statistic for no first-order serial autocorrelation was an F (1, 80) = 0.006. Thus the null hypothesis could not be rejected and further estimation would only consider heteroskedasticity in the errors.

Table 13. Variable Statistics for Profit Fixed Effects Model

lspc	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lmiles	-.0726327	.0377823	-1.92	0.055	-.1469598	.0016944
lcars	-.1688442	.0286636	-5.89	0.000	-.2252326	-.1124558
ltmiles	.1312282	.0281469	4.66	0.000	.0758564	.1865999
amcars	-1.05e-10	1.02e-09	-0.10	0.918	-2.12e-09	1.91e-09
ampop	-1.43e-11	1.74e-12	-8.20	0.000	-1.77e-11	-1.09e-11
tmpop	-1.66e-12	4.81e-12	-0.35	0.730	-1.11e-11	7.80e-12
dy1999	-.005882	.0240478	-0.24	0.807	-.0531898	.0414259
dy2000	-.0138638	.0251514	-0.55	0.582	-.0633428	.0356152
dy2001	-.0194293	.0251956	-0.77	0.441	-.0689952	.0301366
dy2002	.0728624	.0252433	2.89	0.004	.0232025	.1225222
d1	(dropped)					
d2	(dropped)					
d3	(dropped)					
d4	(dropped)					
d5	(dropped)					
d6	(dropped)					
d7	(dropped)					
d8	(dropped)					
d9	(dropped)					
d10	(dropped)					
d11	(dropped)					
d12	(dropped)					
d13	(dropped)					
d14	(dropped)					
d15	-.0506624	.0497352	-1.02	0.309	-.1485038	.047179
d16	(dropped)					
_cons	8.670126	.3426575	25.30	0.000	7.996035	9.344217

Table 14. Regression Statistics for Profit Random Effects Model

```
Random-effects GLS regression                  Number of obs      =       484
Group variable (i): splc                       Number of groups   =       146

R-sq:  within  = 0.3023                        Obs per group: min =         1
       between = 0.0915                                       avg =       3.3
       overall = 0.2050                                       max =         5

Random effects u_i ~ Gaussian                  Wald chi2(26)      =     85.36
corr(u_i, X)       = 0 (assumed)               Prob > chi2        =    0.0000
```

Estimation of the profit maximization model was then conducted under FGLS with heteroskedastic errors. The results are presented in Tables 16 and 17.

This last model specification produced more robust and useable information than any of the other specification types. Only two dummy variables were shown to be statistically insignificant, which may be a result of minimal observations for those classifications.

Analysis of the regression results for both the cost minimization and profit maximization models is discussed in more detail in the Results and Conclusion section of the study. Additional estimation and analysis was performed on the variables of interest to discover if there were any spatial processes affecting the estimated results.

Table 15. Variable Statistics for Profit Random Effects Model

lspc	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
lmiles	-.1008227	.0358932	-2.81	0.005	-.1711721	-.0304734
lcars	-.1838648	.0278486	-6.60	0.000	-.2384471	-.1292825
ltmiles	.1539826	.0273864	5.62	0.000	.1003062	.2076589
amcars	-2.31e-10	1.00e-09	-0.23	0.818	-2.19e-09	1.73e-09
ampop	-1.37e-11	1.73e-12	-7.91	0.000	-1.71e-11	-1.03e-11
tmpop	1.94e-13	4.64e-12	0.04	0.967	-8.91e-12	9.30e-12
d1	.0114345	.1199385	0.10	0.924	-.2236406	.2465095
d2	.0622469	.1761979	0.35	0.724	-.2830947	.4075885
d3	.2284736	.4811457	0.47	0.635	-.7145547	1.171502
d4	-.0815849	.4808399	-0.17	0.865	-1.024014	.860844
d5	.2727945	.917489	0.30	0.766	-1.525451	2.07104
d6	-.7905327	1.32239	-0.60	0.550	-3.38237	1.801305
d7	-.3212806	1.207183	-0.27	0.790	-2.687316	2.044754
d8	-.1768138	.2571747	-0.69	0.492	-.680867	.3272395
d9	-.0306182	.1758756	-0.17	0.862	-.375328	.3140915
d10	.9288914	1.06043	0.88	0.381	-1.149514	3.007297
d11	.5514398	.6756264	0.82	0.414	-.7727636	1.875643
d12	.1891518	.208879	0.91	0.365	-.2202435	.5985471
d13	.0949877	.201379	0.47	0.637	-.2997079	.4896833
d14	.1211745	.181915	0.67	0.505	-.2353723	.4777213
d15	-.0577031	.0503494	-1.15	0.252	-.1563861	.0409798
d16	.7051035	.2681969	2.63	0.009	.1794472	1.23076
_cons	8.758688	.329786	26.56	0.000	8.112319	9.405057

Table 16. Regression Statistics for Profit RE Model with Panel-Specific Errors

```
Cross-sectional time-series FGLS regression

Coefficients:  generalized least squares
Panels:        heteroskedastic
Correlation:   no autocorrelation

Estimated covariances      =       146          Number of obs      =       484
Estimated autocorrelations =         0          Number of groups   =       146
Estimated coefficients     =        23          Obs per group: min =         1
                                                               avg =  3.315068
                                                               max =         5
                                                Wald chi2(19)      =   1230.64
Log likelihood             =  271.1023          Prob > chi2        =    0.0000
```

Spatial Analysis

The spatial analysis follows the modeling framework developed by Anselin (1988) and further developed in Anselin and Bera (1998). The data was examined using a GIS compatible statistical package developed by Anselin and known as GeoDa, version 0.9.5-i that incorporates the methodology developed by Anselin and Anselin and Bera.

GeoDa has many tools for the visualization of spatial processes and a statistical component for estimating spatial lags and spatial errors in conformance with the spatial autoregressive model (SAR). Estimation is straightforward after the calculation of a weighting matrix to be used in determining ρ and λ. In this study the weighting matrix used the geographic point locations of the terminal rail nodes in order to determine distance weights between the points. A k-nearest neighbors weighting option is available, but at present, the GeoDa does not have the capability to estimate SAR models using these types of weights.

Figure 6 can be visualized as a 3 x 3 slice of the contiguous states, creating 9 distinct geographical regions. As can be seen, several regions are quite similar in the make-up of their cost structures, while other regions exhibit more variation.

Table 17. Variable Statistics for Profit RE Model with Panel-Specific Errors

lspc	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lmiles	-.1793617	.0136495	-13.14	0.000	-.2061142	-.1526092
lcars	-.208161	.0104438	-19.93	0.000	-.2286305	-.1876915
ltmiles	.2257767	.0112306	20.10	0.000	.2037653	.2477882
amcars	-1.39e-09	2.03e-10	-6.83	0.000	-1.79e-09	-9.90e-10
ampop	-9.88e-12	7.59e-13	-13.02	0.000	-1.14e-11	-8.39e-12
tmpop	6.26e-12	9.93e-13	6.31	0.000	4.32e-12	8.21e-12
d1	-.0503978	.0120043	-4.20	0.000	-.0739259	-.0268698
d2	-.039108	.0150455	-2.60	0.009	-.0685965	-.0096194
d3	.1099296	.0318799	3.45	0.001	.0474462	.1724131
d4	-.0529668	.1024389	-0.52	0.605	-.2537434	.1478098
d5	.2564055	.1041193	2.46	0.014	.0523355	.4604755
d6	-.5227481	.1183885	-4.42	0.000	-.7547853	-.2907108
d7	-.4950359	.0815225	-6.07	0.000	-.654817	-.3352548
d8	-.0601844	.016557	-3.63	0.000	-.0926354	-.0277334
d9	-.0166694	.0093938	-1.77	0.076	-.0350808	.0017421
d10	.3263136	.0552537	5.91	0.000	.2180183	.4346089
d11	.1862289	.0494561	3.77	0.000	.0892967	.2831611
d12	-.0303197	.0122948	-2.47	0.014	-.054417	-.0062223
d13	-.0682891	.0157724	-4.33	0.000	-.0992024	-.0373758
d14	-.0098904	.0135488	-0.73	0.465	-.0364455	.0166647
d15	-.1038717	.01593	-6.52	0.000	-.135094	-.0726495
d16	.4466509	.0434554	10.28	0.000	.3614798	.531822
_cons	9.360915	.1234469	75.83	0.000	9.118963	9.602866

Visual results such as these do warrant further investigation of the presence of spatial processes within the data. However, regression results were not robust statistically in determining the presence of spatial lags or errors.

The log-transformed variables from the previous regression analyses were used in the SAR regression analysis, but without the interaction or dummy variables included. Two specifications were estimated for both the cost minimization and profit maximization models using maximum likelihood: spatially lagged regression (ρ estimation), and spatially dependent errors (λ estimation). Results of the estimations are listed in Tables 20 and 21 for the cost minimization model, and Tables 22 and 23 for the profit maximization model.

The spatial lag variable, W_LCPC, and the spatial error term, LAMBDA, were determined to be statistically insignificant in determining change in log cost. However, the likelihood ratio test for spatial dependence in the spatial lag model had a value of 2.050098 and DF=1, which does indicate spatial dependence with a confidence level of 85%.

Again, the spatial lag and spatial error terms were determined to be statistically insignificant. However, in the profit model, the likelihood ratio test for spatial dependence in the error terms had a statistic = 2.315287, with DF=1, so the presence of spatial errors can be noted with a confidence level of 87%.

Figure 6. Regional Box Plots of cost in the lower 48 United States

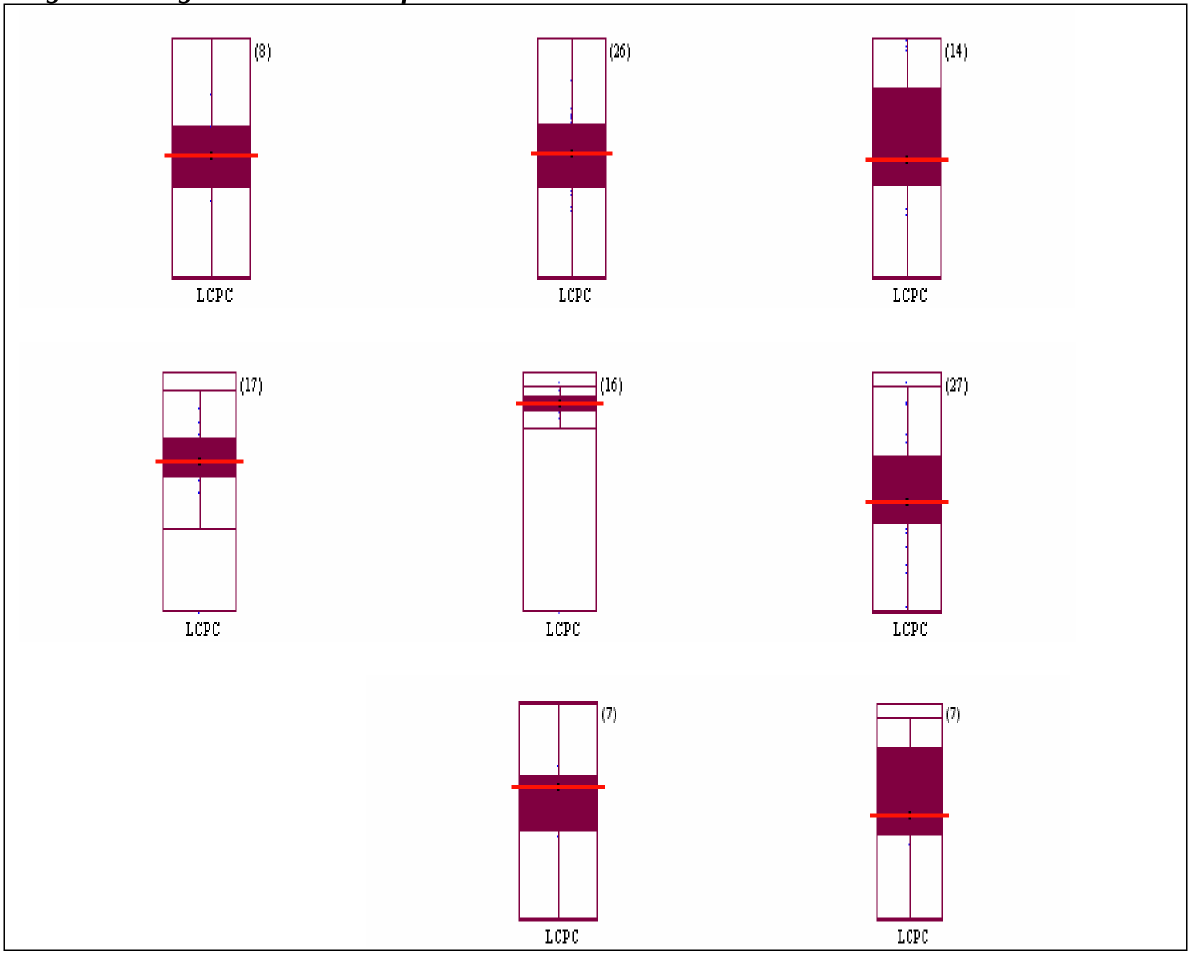


Table 20. Regression Results for Spatial Lag, Cost Model

Variable	Coefficient	Std.Error	z-value	Probability
W_LCPC	-0.1360781	0.1436525	-0.9472722	0.3435001
CONSTANT	3.87285	1.477871	2.62056	0.0087786
LMILES	0.6221565	0.1415516	4.395262	0.0000111
LCARS	-0.131833	0.09625096	-1.36968	0.1707870
LTMILES	0.1033856	0.1000521	1.033318	0.3014550
LINVW	0.04688526	0.02577761	1.818837	0.0689362
LINVH	-0.002813229	0.04119376	-0.0682926	0.9455526

Table 21. Regression Results for Spatial Errors, Cost Model

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	2.785503	1.202699	2.316044	0.0205558
LMILES	0.6159394	0.139649	4.410627	0.0000103
LCARS	-0.09654452	0.09141661	-1.056094	0.2909254
LTMILES	0.07065259	0.09757204	0.7241069	0.4690000
LAMBDA	-0.08202683	0.1664072	-0.4929284	0.6220632

Table 22. Regression Results for Spatial Lag, Profit Model

Variable	Coefficient	Std.Error	z-value	Probability
W_LSPC	0.1084052	0.1181445	0.9175645	0.3588468
CONSTANT	8.284887	0.9760228	8.488416	0.0000000
LCARS	-0.188933	0.02770839	-6.818618	0.0000000
LMILES	-0.1360784	0.03981119	-3.418094	0.0006307
LTMILES	0.1441624	0.02874566	5.0151	0.0000005
LINVW	-0.002545369	0.007362101	-0.3457394	0.7295387
LINVH	0.01515958	0.01158904	1.308096	0.1908408

Table 23. Regression Results for Spatial Errors, Profit Model

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	9.096343	0.3548825	25.63199	0.0000000
LMILES	-0.1370037	0.04045678	-3.386421	0.0007082
LCARS	-0.1949627	0.027802	-7.012544	0.0000000
LTMILES	0.149896	0.02883397	5.198591	0.0000002
LINVW	-0.00261208	0.007577508	-0.3447149	0.7303088
LINVH	0.01541025	0.01162593	1.325508	0.1850030
LAMBDA	0.2092167	0.144606	1.446805	0.1479516

RESULTS AND CONCLUSION

The results of this study are two-fold. First, the selection of a model framework several diverse data sources to estimate cost efficiencies in inter-modal facilities was made. Due to the structure of the available data the best model framework for this effort was a FGLS model with heteroskedastic and AR (1) errors for cost minimization, and a FGLS model with heteroskedastic errors for profit maximization. Evidence of spatial dependence in the data was found, although it was not found to statistically significant. The results do, however, suggest a more thorough examination of the spatial characteristics of the inter-modal freight network would be of value.

Estimates of the FGLS cost minimization model identified *lmiles, lcars, ltmiles, amcars* and *tccarpop* as significant explanatory variables. The first four are capacity and throughput measures, while *tccarpop* is a connectivity measure. Moreover, many of the characteristic indicator variables were found to be significant and are discussed below.

The average miles coefficient, *lmiles*, was positive which can be explained by the nature of railroad cost curves. While rail does have cost efficiencies over long distances, variable costs can, and do, increase as total mileage of a haul increases. Since most inter-modal rail movements are over long distances, we can expect that some components of cost will be increasing as the average length of haul increases. The tonmiles variable, *ltmiles*, also has a positive sign that may indicate another component of long-distance costs increasing over miles traveled. This may also capture the effect at an originating terminal of building large trains with more fully loaded railcars that will then move over longer distances. The positive, though small, sign on *amcars* would indicate that some increase does accrue to facility costs as the number of cars traveling long distances are built or transferred at a facility. A countering force on costs is the number of carloads, noted by the variable *lcars*; as total carloads increases, variable costs do decrease in terminal facilities. Also, the *tccarpop* variable has a negative sign. This is a more ambiguous interaction term, but it does measure the impact that increasing connectivity between terminals, coupled with high numbers of carloads in proximity to large population centers does have the effect of lowering terminal costs. In sum, these interactions lower costs of operation at the servicing inter-modal terminal facilities.

Supporting this contention, the characteristic indicator variables for the number of servicing railroads indicate that the more railroads that service a facility, more cost efficiencies are realized. Terminal connectivity did show a positive to cost in the characteristic dummies for ranges between 25 and 100 terminal connections. A facility with fewer than 25 connections is the base case, while more than 100 terminal connections did appear to have a negative impact on costs, though statistically insignificant. Also, breaking down cost efficiencies by the number of commodities serviced by a terminal found that costs increased as the number of commodities increased, while terminals handling fewer than 50 commodities realized cost efficiencies. This would indicate that intermodal facilities would benefit from specialization in handling specific commodity types in high volumes, rather than focusing on a large variety of commodity types.

Finally, it should be noted that the carload and population threshold characteristic variables (d15 and d16) were both positive. This could indicate the presence of marginal or variable costs that begin to increase after a threshold level is breeched; i.e., terminal capacity begins to approach its maximum and costly delays, traffic jams within the facility, etc. begin to wear on cost efficiencies. The population dummy may indicate that variable costs will be higher when locating a facility near a large population center.

For the profit maximization model, the same characteristic variables were of significance, as were the variables *lmiles, lcars, ltmiles*, and interaction terms *amcars, ampop* and *tmpop*. In the FGLS estimation, the average distances, the number of carloads, and the distances of from major population centers negatively affect profit. As these variables increase, profit decreases. However, tonmiles traveled and tonmiles traveled in conjunction with major population centers were positive for profit. This would indicate that large volumes of fully loaded cars moving to major population centers are profit-generating movements for the servicing railroads.

Of note, the fewer servicing railroad connections a terminal had, the lower the profit expectations. Indications from the data are increasing profitability for terminals having more than 3 servicing railroads. Also, profits were highest with terminals servicing large numbers of commodities, although, as noted above, this is accompanied by increasing costs. Finally, profits were estimated to increase if the terminal was located in proximity to a major population center. This would indicate the benefits of being near to a major attractor and generator of higher value goods that would generate large railroad profits during shipment.

REFERENCES

Anselin, Luc. *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers , Berlin, 1988.

Anselin, Luc. "An Introduction to Spatial Autocorrelation Analysis with GeoDa." Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign (2003).

Anselin, Luc, and A. Bera. "Spatial dependence in linear regression models with an introduction to spatial econometrics." A. Ullah and D.E. Giles, eds. *Handbook of Applied Economic Statistics*. New York: Marcel Dekker (1998): 237-289.

Arnold, Pierre, Dominique Peeters and Isabelle Thomas (2004). “Modelling a rail/road intermodal transportation system.” *Transportation Research Part E* 40, 255-270.

Ballis, Athanasios and John Golias (2002). “Comparative Evaluation of Existing and Innovative Rail-Road Freight Transportation Terminals.” *Transportation Research Part A* 36, 593-611.

Barber, Gerald M. (1975) “A Mathematical Programming Approach to a Network Development Problem.” *Economic Geography* 51 (2), 128-141.

Beuthe, Michel, Bart Joaquin, Jean-Francois Geerts and Christian Koul a Ndjang’ Ha (2001). “Freight Transportation Demand Elasticities: A Geographic Multimodal Transportation Network Analysis.” *Transportation Research Part E* 37, 253-266.

Casavant, Kenneth L., Eric L. Jessup and Angelica Monet. “Determining the Potential Economic Viability of Inter-Modal Truck-Rail Facilities in Washington State. (2004) Review Draft, prepared for the Washington State Transportation Commission, Washington State Department of Transportation.

Fernandez L., J. Enrique, Joaquin de Cea Ch. and Alexandra Soto O. (2003). “A Multi-Modal Supply-Demand Equilibrium Model for Predicting Intercity Freight Flows.” *Transportation Research Part B* 37, 615-640.

Friesz, Terry L. “Strategic Freight Network Planning Models.” David A. Hensher and Kenneth J. Button, eds. *Handbook of Transport Modelling*. New York: Pergamon (2000): 527-537.

Greene, William H. *Econometric Analysis, Fifth Edition*. Pearson Education, Singapore, 2003.

Griffith, Daniel A. (1982). “Dynamic Characteristics of Spatial Economic Systems.” *Economic Geography* 69 (2), 177-196.

Horner, Mark W., Morton E. O’Kelly (2001). “Embedding economies of scale concepts of hub network design.” *Journal of Transport Geography* 9, 255-265.

McCarl, Bruce A., Donald A. Hilger and J. William Uhrig. “Grain Subterminal Facility Location: A Mixed-Integer Programming Model.” Won K. Koo and Donald W. Larson, eds. *Transportation Models for Agricultural Products*. Boulder: Westview Press (1985): 13-33.

Melkote, Sanjay and Mark S Daskin (2001). “An integrated model of facility location and transportation network design.” *Transportation Research Part A* 35, 515-538.

Mizon, Grayham E. (1995). “A simple message for autocorrelation correctors: Don’t.” *Journal of Econometrics* 69, 267-288.

Nierat, Patrick (1997). “Market Area of Rail-Truck Terminals: Pertinence of the Spatial Theory.” *Transportation Research Part A* 31, 109-127.

O’Kelly, M.E., and D.L. Bryan (1998). “Hub Location with Flow Economies of Scale.” *Transportation Research Part B* 32, 605-616.

Sheu, Jiuh-Bing (2003). “Locating manufacturing and distribution centers: An integrated supply chain-based spatial interaction approach.” *Transportation Research Part E* 39, 381-397.