Forecasting Waterborne Exports With Alternative Regional Economic Models: A Statistical Analysis Based On The Charleston Port#

Joseph Pelzman*

As international trade grows in importance to the U. S. economy, local ports become more significant to their region's economic development. This fact has been recognized by a number of recent studies [8, 17, 19]. Despite this recognition, a limited effort has been made to forecast waterborne exports out of a regional port thus inhibiting any accurate estimates of future growth of port services and their associated ripple effects on the economic development of the region. Part of this problem arises from the extreme complexity of measuring interregional trading relationships and part from the absence of reliable interregional trade data.

The objective of this paper is to evaluate both the accuracy and the applicability of different regional models to the problem faced by both port administrators and development boards. Namely, what will next years commodity export traffic out of a particular port be, given that information on port activity is limited to its share of the total market, its past growth and the expected growth of total U. S. waterborne exports. The regional models evaluated must therefore cope with the constraints imposed by the data limitation as well as lend themselves to meaningful interpretation by port administrators. The regional models tested in this paper meet both requirements. In general, these models assume that waterborne exports out of a given port are either a linear, exponential or share function of total U. S. waterborne exports or some spatial distribution of commodity exports across ports.

The Charleston port was chosen as a case study primarily because it represented a growing port facility. Its exports both in terms of value and tonnage was in the early 1970's much higher than that of its major southeastern competitors, Savannah, Wilmington and Norfolk. Because of this surge in the growth of the Charleston port and the expected ripple effect on its regional economy, Charleston was judged ideal for an investigation of alternative forecasting models of a port's waterborne exports. While a

---

#Work on this paper was partially supported by a grant from the University of South Carolina Research and Productive Scholarship Fund. The author wishes to express his gratitude to an anonymous referee for helpful comments on an earlier draft.

*Economics Department, George Washington University, Washington, D.C.

similar analysis for a number of port facilities was beyond the scope of this paper, we believe the results of this study are applicable to similar port facilities constrained by identical data limitations.

The format of this paper is as follows: Section I outlines the models to be tested. Section II presents a comparison of the resulting forecasts of waterborne exports leaving the Charleston port for the years 1972, 1973 and 1974. The implications of the comparison, to policy makers, are discussed in Section III. Data sources are provided in the Appendix.

I. ALTERNATIVE REGIONAL MODELS

The regional science literature abounds with very elegant theoretical models describing the flow of commodities across states as well as regions. However, given the data constraints, many of these models had to be excluded from the analysis. The models presented below represent a group of very popular regional models which can be adapted to handle both the data constraint as well as the specific problem of forecasting waterborne exports.

The four models described below are separated into three groups. The first is based on regression growth rates. The second set of models is characterized as share models and the third group is represented by a spatial distribution model.

A. Regression Growth Models

The least complicated procedure available to a port authority by which forecasts of future waterborne exports from a regional port can be generated is based on a very naive concept of growth, namely that they are extrapolations of past trends. One such model, Regression Rate-Linear Growth (RRLG) is shown in equation (1).

$$X^{\phi}_{i,t} = a + b^{\phi}_{i}(t) \tag{1}$$

where $X^{\phi}_{i,t}$ represents the waterborne exports of commodity i originating in port ϕ.

ϕ denotes the Charleston port.

t represents the time period.

Simple extrapolation models such as the RRLG model are frequently the basis for making "casual" long-range forecasts of variables ranging from national income to population. Although they can provide quick, initial forecasts, they usually provide little forecasting accuracy.

B. Share Models

An equally simplistic model developed in the regional science literature [6, 9, 14, 20] is the Average Constant Share Model (ACS), which operates on the assumption that a ports' share of total waterborne exports at time period t + n is the same as that observed over time period (t-n, t). The projection model is written as:

$$\hat{X}^{\phi}_{i,t+n} = \alpha^{\phi}_{i,(t-n,t)} \cdot X_{i,t+n} \tag{2}$$

where

$\hat{X}^{\phi}_{i,t+n}$ = the projected waterborne exports of commodity i originating in the Charleston port (ϕ) in period t+n.

$X_{i,t+n}$ = U.S. waterborne exports of commodity i in period t+n.

$\alpha^{\phi}_{i,(t-n,t)}$ = the average constant share of waterborne exports of commodity i originating in the Charleston port (ϕ) over the period (t-n,t).

As long as independent projections of $X_{i,t+n}$ are available $\hat{X}^{\phi}_{i,t+n}$ can be estimated by this average constant share model. However, despite the fact that this model can be useful as a means of quickly formulating initial forecasts, it is constrained by the assumption that the port's "share" will be unchanged over time. The greatest weakness of this assumption is that it ignores *a priori* any net gains or losses to the regional port attributable to either a change in the product mix or a change in the relative competitive position of the particular port facility. This assumption is particularly weak given the fact that port facilities are attempting to become more competitive, either by providing containorized cargo facilities and/or providing fast service and full cargos.

A more sophisticated model designed to deal with these issues would consider differences between the growth of total U. S. waterborne exports and waterborne exports from the specific port ϕ. Such a model commonly known as the Shift-Share model (SHISHA) [6, 9, 12, 14, 20] is specified as:

$$\hat{X}^{\phi}_{i,t+n} = \left[\hat{R}_{i,(t,t+n)} + \gamma(R^{\phi}_{i,(t-n,t)} - R_{i,(t-n,t)})\right] \cdot X^{\phi}_{i,t} \tag{3}$$

where the growth rates are defined as:

$\hat{R}_{i,(t,t+n)} = X_{i,t+n}/X_{i,t}$

$R^{\phi}_{i,(t-n,t)} = X^{\phi}_{i,t}/X^{\phi}_{i,t-n}$

$R_{i,(t-n,t)} = X_{i,t}/X_{i,t-n}$

γ = scalar to adjust for differences between the historical period and the projection period. In this paper γ = 1/5, 2/5, and 3/5 for 1972, 1973 and 1974 respectively.

The basic assumption behind the shift-share model is that the differences in growth rates between the specific port and total U. S. waterborne exports in the period (t-n,t) is maintained in the projection period (t, t+n). The "share" component is $\hat{R}_{i,(t,t+n)}$ and the "shift" component is represented by $(R^{\phi}_{i,(t-n,t)} - R_{i,(t-n,t)})$. This latter component takes into account any differences in growth of waterborne exports between the specific port (ϕ) and total U. S. waterborne traffic. A positive value would indicate a more rapid growth by the specific port while a negative value would imply the reverse.

It should be emphasized that the SHISHA model is merely a measurement technique for decomposing the growth of a variable such as a port's waterborne exports. The model does not attempt to explain why the "Shift" component is positive or negative. Thus, the SHISHA model should not be viewed as a behavioral relationship. The theoretical and empirical limitations of this model are well documented in the literature [2, 6, 13, 16] and need not be described in detail in this paper. It should be sufficient to note that the major criticism of the SHISHA model include 1) that it does not take into account changes in regional structure over the period of analysis, 2) that it presupposes that structural changes in the region and the product mix of exports is independent when in fact it is interwoven, and 3) that it is not invariant with disaggregation.

C. Spatial Distribution Model

Despite their appeal to simplicity all three models presented above have one major drawback. That is, they ignore the possibility that the regional origin and the port of shipment (ϕ) of waterborne exports can be spatially interdependent. Based on the work of Beckmann [3, 4], Bos [5] and Anderson [1], however, one can make an attempt to infer the interregional flow of waterborne exports from a spatial distribution of production shares. One should recognize, however, that this spatial distribution model (SPADIS) is not free of weaknesses. The major drawback of this model, as demonstrated by Cliff and Ord [7] is the assumption that a region's production growth in sector i is independent of the growth of the same sector in other regions.

This spatial distribution model (SPADIS) is written as:[1]

$$\hat{X}^{\phi}_{i,t+n} = \left[\sum_{k} \beta^{k,\phi}_{i,(t-n,t)} \, \delta^{k}_{i,(t-n,t)}\right] - X_{i,t+n} \tag{4}$$

where:

$\delta^{k}_{i,(t-n,t)}$ = regional share parameter assumed to be represented by a spatial distribution of production, P^k_i/P_i.
Where P_i = U. S. production of good i and P^k_i = production of good i in region k, over the period (t-n,t).

$\beta^{k,\phi}_{i,(t-n,t)}$ = each regions' contribution to the waterborne exports of port ϕ. It is estimated from the following relationship:

$$X^{\phi}_{i,(t-n,t)} = \sum_k \beta^{k,\phi}_{i,(t-n,t)} \cdot P^{k}_{i(t-n,t)}/P_{i,(t-n,t)} \cdot X_{i,(t-n,t)}$$

k = customs region

While these models leave a great deal to be desired they were primarily chosen because they represent a set of very popular regional models which can very easily be adapted to deal with the problem at hand. As such they represent a set of models which can be empirically tested with the limited information available to port administrators. In the following section, the accuracy of the various models' performance in forecasting the waterborne exports out of the Charleston port is evaluated.

II. EMPIRICAL RESULTS

In an empirical investigation of the forecasting performance of the various regional models described above, one would like a fairly long time series preferably disaggregated by commodity flow and by country of destination. Current data sources, however, do not permit this. The analysis presented here, is based on a sample of total waterborne exports of the U. S. and that leaving the Charleston port disaggregated by 20 two-digit Standard Industrial Classification (SIC) groups and designated for export to a group of 19 countries.[2] Using the data for the sub-period 1967-1971, forecasts of the value of waterborne exports originating in the Charleston port were generated for the years 1972, 1973, and 1974.[3] Based on these forecasts, the various regional models were tested for the accuracy of their predictions.

For the spatial distribution model, data limitations necessitated a modification of equation (4) such that

$$\hat{X}^{\phi}_{i,t+n} = \left[\sum_k \beta^{k,\phi}_{(t-n,t)} \delta^{k}_{i(t-n,t)}\right] \cdot X_{i,t+n} \tag{4}$$

where $\beta^{k,\phi}_{(t-n,t)}$ is used as a proxy for $\beta^{k,\phi}_{i,(t-n,t)}$.[4]

Estimates of $\beta^{k,\phi}_{(t-n,t)}$ were obtained by pooling cross-section and time-series observations for the 1967-1971 period over the 20 disaggregated commodity groups.[5] The pooling procedure used is commonly termed a variance components model [15]. This generalized least squares (GLS) procedure assumes that the residuals have zero mean and common variance and that they are both serially independent and independent across cross-section units.

The accuracy of the various estimation models can be evaluated with a set of measures developed by Theil [18]. The first measure, the inequality coefficient is defined as:

$$U = \left[MSE/\sum_i A_i^2/n\right]^{1/2}$$

where $MSE = 1/n \sum_i (P_i - A_i)^2$

P_i = predicted value

A_i = actual value

MSE = mean square error

n = number of forecasts

The smaller the inequality coefficient, the more precise is the forecast.

The MSE found in the numerator of the inequality coefficient is decomposed by Theil [18, p. 32] giving us an indication of the source of estimation error. The three inequality proportions are:[6]

$U^m = (\overline{P} - \overline{A})^2/MSE$ (Central Tendency)

$U^s = (S_p - S_a)^2/MSE$ (Unequal Variation)

$U^c = 2(1\text{-}r)S_aS_p/MSE$ (Incomplete Covariation)

where

$$\overline{P} = 1/n \sum_i^n P_i$$

$$\overline{A} = 1/n \sum_i^n A_i$$

$$S_p = \left[1/n \sum_i^n (P_i - \overline{P})^2\right]^{1/2}$$

$$S_a = \left[1/n \sum_i^n (A_i - \overline{A})^2\right]^{1/2}$$

$$r = \Sigma\, (P_i - \overline{P})(A_i - \overline{A})/nS_pS_a$$

Table 1 shows the inequality coefficients and their decomposition into the bias, variance and covariance proportions for the four models tested. The inequality coefficient of the SHISHA model is by far the smallest. The data leaves no doubt that given "perfect information" the SHISHA model gives better results than the three alternative regional models in predicting the waterborne exports out of a specific port. Despite the fact that in practical application the SHISHA, ACS and SPADIS models require exogenous projections of $\hat{R}_{i,(t,t+n)}$ and $\hat{X}_{i,t+n}$ respectively, the errors associated with these independent projections would be the same for all three models. Thus, the results based on perfect information would hold in the case of "less than perfect information."

The decomposition of the inequality coefficient indicates that errors of central tendency, U^m, account for the smallest proportion of the estimation error across all models. They range from a low of 8.5 percent in the SHISHA model to a high of 26.4 percent in the SPADIS model. A more serious source of prediction error attributed to unequal variation, U^s, ranges from a low of 2.9 percent in the case of the SHISHA model to a high of 46.5 percent in the SPADIS model. The most serious source of error is attributed to incomplete covariance, U^e. Theil [18, p. 32] argues that the seriousness of this error is due to the fact that it does not lend itself to ad hoc correction. This source accounts for a high of 88.6 percent of the error in the case of the SHISHA model and for over 25 percent of the error in the three alternative models.

Although it seems that *ad hoc* adjustments to the projections of the SHISHA model will do little to improve its predicted accuracy, some improvement is possible for the three alternative models. The possibility of adjusting these predictions to reduce the errors of estimation can be evaluated by means of an alternative set of inequality coefficients [18, p. 34]. In this alternative decomposition described by Theil, the MSE is decomposed into a bias proportion U^m which measures errors of central tendency, a regression proportion U^r which measures the deviation in slopes of the regression line from 1 and a disturbance proportion U^d which measures the variance of the regression disturbances. The three inequality proportions are:[7]

$$U^m = (\overline{P} - \overline{A})^2/\text{MSE} \text{ (bias proportion)}$$

$$U^r = (Sp - rS_a)^2\text{MSE} \text{ (regression proportion)}$$

$$U^d = (1\text{-}r^2)S_a^2/\text{MSE} \text{ (disturbance proportion)}$$

The optimal linear correction described by Theil would be $\hat{\alpha} + \hat{\beta}P_i$ where each forecast is multiplied by some coefficient $\hat{\beta}$ and some constant $\hat{\alpha}$ is added. These correction coefficients are in fact the least squares

TABLE 1
Inequality Coefficients And The Bias, Variance And Covariance Proportions*

	Models			
Statistics	RRLG	ACS	SHISHA	SPADIS
U	.661	.546	.513	.616
U^m	.225	.224	.085	.264
U^s	.463	.418	.029	.465
U^c	.312	.358	.886	.271

*The descriptive measures of the forecasts used to derive U, U^m, U^s and U^c are presented in the Appendix, Table A1.

estimates of the linear regression coefficients. As Theil [18, p. 34] demonstrates applying the mean square error criterion

$$1/n\ \Sigma(\hat{\alpha} + \hat{\beta}P_i - A_i)^2$$

and minimizing with respect to $\hat{\alpha}$ and $\hat{\beta}$ yields:

$$\hat{\beta} = \Sigma(P_i - \overline{P})\ (A_i - \overline{A})/\Sigma(P_i - \overline{P})^2 = rS_a/S_p$$

and

$$\hat{\alpha} = \overline{A} - \hat{\beta}\overline{P}$$

The regression and disturbance proportions of error for the four models tested are stated in Table 2. It is apparent, that a large portion of the estimation error (from a high of 89.8 percent in the case of the SHISHA model to a low of 49.9 percent in the case of the SPADIS model) can be attributed to the disturbance proportion. Thus, the errors associated with the various regional models tested in this study are to a very high extent random in nature. The addition of the linear corrections ($\hat{\alpha}$ and $\hat{\beta}$ in Table 2) would therefore provide for a very small reduction in the errors of estimation.

III. CONCLUDING REMARKS

In the foregoing analysis an attempt was made to evaluate the performance of various regional models in predicting waterborne exports out of a regional port. These models were primarily chosen because they represent a set of very popular regional models which can be applied with the limited information available to port administrators. The determination of the most suitable model for port administrators was based entirely on forecasting performance. Although the empirical evidence is based on one port,

TABLE 2
Inequality Proportions And Correction Coefficients*

	Models			
Statistics	RRLG	ACS	SHISHA	SPADIS
U^m	.225	.224	.085	.264
U^r	.177	.202	.017	.237
U^d	.598	.574	.898	.499
$\hat{\alpha}$	7.343 E06	5.631 E06	8.316 E06	7.587 E06
$\hat{\beta}$	1.701	1.496	.912	1.701

*The descriptive measures of the forecasts used to derive U^m, U^r, U^d, $\hat{\alpha}$ and $\hat{\beta}$ are presented in the Appendix, Table A1.

we believe that the results are applicable to similar port facilities.

The basic conclusion of this study is that given information on a ports' share of total waterborne exports, its past growth and the expected growth of total U. S. waterborne exports, the SHISHA model represents, by far, the best forecasting model. While this result is, in its own right very important, what makes this study extremely useful to port administrators is that it presents a comparative evaluation of very popular regional models adapted to resolve their major forecasting problem. As such, the SHISHA model should appeal to port administrators, not only because it meets the data constraints but because it lends itself to meaningful interpretation and forecasting accuracy.

APPENDIX

The value of total U. S. waterborne exports was taken from, *U. S. Exports of Domestic Merchandise, FT 450*. The value of exports leaving the Charleston port was taken from *U. S. Exports of Domestic and Foreign Merchandise, Customs District of Exportation By Schedule B Number By Country of Destination, and Method of Transportation, EA 664*.

The seven digit trade flow data was converted to the SIC classification and aggregated to the two digit level. The cross classification used was the *U. S. Foreign Trade Statistics; Classification and Cross-Classification.*

The sample of nineteen countries chosen were determined by a frequency distribution of the seven digit export data for the Charleston district. They are: United Kingdom, Netherlands, France, West Germany, Switzerland, Spain, Greece, Malaysia, Singapore, Philippines, Hong Kong, Japan, Taiwan, Australia, Colombia, Venezuela, Ecuador, Peru and Brazil. These countries represent the group of countries to which commodity i is exported.

TABLE A1
Descriptive Measures Of The Forecasts

	Models			
Statistics	RRLG	ACS	SHISHA	SPADIS
$\overline{P}$	9.889 E06	1.239 E07	1.737 E07	9.745 E06
$\overline{A}$	2.417 E07	2.417 E07	2.417 E07	2.417 E07
S_p	1.807 E07	2.251 E07	3.462 E07	1.943 E07
S_a	3.858 E07	3.858 E07	3.858 E07	3.858 E07
r	.797	.873	.819	.857
MSE	9.079 E14	6.184 E14	5.462 E14	7.887 E14

FOOTNOTES

[1]As in the ACS model, as long as independent estimates of $X_{i,t+n}$ are available, $\hat{X}^{\phi}_{i,t+n}$ can be projected by the SPADIS model.

[2]Data sources are provided in the Appendix.

[3]Given the severe data limitation, it was decided to use a large enough base period such that a proper evaluation would be based on forecasts generated for more than one year. A comparison of forecasts to actual performances for each of these models may be obtained from the author.

[4]This substitution could possibly lead to some error. In fact, one could argue that the real test of the models forecasting performance rests on the accuracy of $\beta^{k\phi}_{(t-n,t)}$ as an estimate of $\beta^{k\phi}_{i,(t-n,t)}$.

[5]In the present study a consistent set of observations over time series and cross sections is not sufficient for an efficient estimate of either a time series or a cross-section equation, thus suggesting a pooling technique.

[6]The first term, U^m, is zero only when the mean of the estimated values equals the mean of the actual values. A positive value implies errors of central tendency. The second term, U^s, is zero if the standard deviation of the estimated and actual values are equal. When $U^s > 0$ the error is due to unequal variation. The last term is zero

when r = 1. When $U^c > 0$ the error is attributed to incomplete covariation.

[7]When $U^r > 0$ the regression line fitted between the actual and predicted values has a slope that differs from one. $U^d > 0$ implies imperfect linear correlation between the actual and predicted values.

REFERENCES

1. Anderson, James E., "External Trade and Transportation Demand,: U. S. Department of Transportation, mimeograph 1974.
2. Ashby, L. D., "The Shift and Share Analysis: A Reply," *Southern Economic Journal,* 34 (1968), 423-425.
3. Beckmann, M., *Location Theory.* New York: Random House, 1968.
4. ———, "Market Share, Distance and Potential," *Regional and Urban Economics,* 1 (1972), pp. 1-18.
5. Bos, H. C., *The Spatial Dispersion of Economic Activity,* Rotterdam: Rotterdam University Press, 1964.
6. Brown, H. James, "Shift and Share Projections of Regional Economic Growth," *Journal of Regional Science,* 9 (1969), pp. 1-18.
7. Cliff, A. D. and J. K. Ord. *Spatial Autocorrelation.* London: Pion, 1973.
8. Clifton, David S., "The Economic Impact of Georgia's Deepwater Ports," Industrial Development Division Engineering Experiment Station, Georgia Institute of Technology, 1973.
9. Esteban-Marquillas, J. M. "A Re-Interpretation of Shift-Share Analysis," *Regional and Urban Economics,* 2 (1972), pp. 249-55.
10. Greytak, David. "A Statistical Analysis of Regional Export Estimating Techniques," *Journal of Regional Science,* 9 (1969), pp. 387-395.
11. Hellman, D. A. "A Model of Regional Export Activity," *Growth and Change,* 5 (1974), pp. 12-17.
12. Hewings, G. J. D. "On The Accuracy of Alternative Models for Stepping-Down Multi-County Employment Projections to Counties," *Economic Geography,* 51 (1976), pp. 206-217.
13. Houston, D. B. "The Shift and Share Analysis of Regional Growth: A Critique," *Southern Economic Journal,* 33 (1967), 577-581.
14. James F. J. and J. Hughes, "A Test of Shift and Share Analysis as a Predictive Device." *Journal of Regional Science,* 13 (1973), pp. 223-31.
15. Maddala, G. S., "The Use of Variance Components Models in Pooling Cross-Section and Time-Series Data," *Econometrica,* 39 (1971), pp. 341-358.
16. Paraskevopoulos, C. C., "The Stability of the Regional Share Component: An Empirical Test," *Journal of Regional Science,* 11 (1971), pp. 107-112.
17. Pender, David R., and Ronald P. Wilder, "Impact of the State Ports Authority Upon the Economy of South Carolina," Division of Research, Bureau of Business and Economic Research, College of Business Administration, University of South Carolina, 1974.
18. Theil, Henri, *Applied Economic Forecasting,* Amsterdam: North Holland Publishing Co., 1966.
19. Wilder, Ronald P. and David R. Pender, "Estimates of the Economic Impact of the State Ports Authority on the Economy of South Carolina, 1977." Division of Research, Bureau of Business and Economic Research, College of Business Administration, University of South Carolina, in process 1978.
20. Zimmerman, R. "A Variant of the Shift and Share Projections Formulation," *Journal of Regional Science,* 15 (1975), pp. 29-38.
21. U. S. Department of Commerce, Bureau of the Census, *U. S. Exports of Domestic Merchandise, FT 450.* Washington: Government Printing Office, annual 1967-1974.
22. ———, *U. S. Exports of Domestic and Foreign Merchandise, Custom District of Exportation by Schedule B Number by Country of Destination and Method of Transportation, EA 664* Washington: Government Printing Office, annual 1967-1974.
23. ———, *U. S. Foreign Trade Statistics; Classification and Cross-Classification.* Washington: Government Printing Office, 1974.
24. ———, *Annual Survey of Manufacturers,* Washington: Government Printing Office, annual 1967-1974.