INTERNAL MIGRATION IN A LOW INCOME REGION: THE ATLANTIC COASTAL PLAINS

Eric Chetwynd, Jr. and Charles Richter*

In a study of migration in the Atlantic Coastal Plains Region, it was discovered that approximately thirty percent of the migrants from State Economic Areas (SEA's) in the Coastal Plains remain within this low income region.¹ We were surprised with the magnitude of this percentage and decided to examine factors which could account for intra-regional migration flows within the Coastal Plains.

The Atlantic Coastal Plains Region includes the eastern plains of North Carolina, South Carolina, and Georgia. It is a contiguous region which represents about one half of the land area of these states.

1. Theoretical Content of the Model

In exploring the literature on migration, we noticed that empirical analyses often employed gravity models or various modifications of gravity models to explain interactions among pairs of regions. Gravity models are built on the premises that distance is a hindrance to interaction between two areas and that interactions increase directly with the sizes of the areas. An early and frequently quoted study using gravity models in a demographic context was by John Q. Stewart.² Stewart borrowed from Newtonian physics to formulate "demographic forces" between two areas as a direct function of the product of the populations of the areas and as an inverse function of the distance between the areas. He also included a constant of gravity in the numerator of his expression. Since Stewart's study in 1948, a huge number of gravity models have appeared in academic journals.³ These models have been used by Stauffer,⁴ Greenwood,⁵ and Lowry⁶ to explain flows of persons between states and cities. In the Stauffer article the importance of intervening opportunities was incorporated into his model. Lowry and Greenwood modified gravity models to include more variables; both models included unemployment and an income variable. Other gravity models measured market potentials for industries.⁷ Another use for gravity models is to measure the demand for air travel between specific cities. Two such studies are relatively recent.⁸ Only a few of the many studies involving gravity models are cited in my paper.

Gravity models are used because of their empirical properties. These models have met with much success in explaining interactions between regions. The specific type of nonlinear form embodied in gravity models has often yielded satisfactory results. Although the primary justification for gravity models rests on their empirical performance, these models have a useful theoretical property in that regression coefficients obtained from these models are elasticities. This is a result of a double logarithmic transformation which is usually taken on both sides of a gravity equation to make the equation linear in its parameters.

*The authors are, respectively, a graduate student in economics at Duke University and an assistant professor of economics at the University of North Carolina. This research was partially supported by a Population Council Fellowship and by the Economic Development Administration. Neither of these agencies is responsible for the authors' conclusions.

While gravity models are useful for explaining migration flows between points in space, other types of models are used to explain outflows and inflows of migrants to areas without considering the origins or destinations of the migrants. Studies using linear regression analysis (without gravity models or logarithmic transformation) to explain flows of in-migrants or out-migrants can be found in studies by Iden and Richter,⁹ and Persky and Kain.¹⁰

Since we wish to explain point-to-point flows of migrants, we will follow the precedent of earlier studies and employ a gravity model. Of the gravity models which we have examined, the one which has a form which we believe would be most useful for our analysis was developed by Ira Lowry.¹¹ Lowry modified his gravity model by adding variables for wage rates and unemployment rates to the more standard gravity variables of distance and labor force sizes. We did not want to include some variables which were in the Greenwood¹² study because we have no a priori hypotheses regarding the directions of the signs of the regression coefficients for these variables. Greenwood included variables for years of schooling completed and a variable to measure the relative degree of urbanization of the areas.

As in the Lowry model, we assume that migration flows can be largely explained by "push" and "pull" factors in combination with gravitational forces, in the demographic sense of the term. In defining the gravity function, sizes of labor forces are used to partially explain interactions between the areas. We expect that larger labor forces are conducive to more intense interactions. As in most gravity models, we introduce a distance variable as an impediment to areal interactions.

The push and pull factors in our model are unemployment and wage rates at points of origin and destination for migrants. We anticipate that population would flow from areas with high unemployment rates to areas with low unemployment rates. Similarly, we expect people to be repelled by low wage rates and to be attracted by high wage rates. These unemployment and wage data serve as proxies for relative economic opportunities within the region.

Our model can be expressed by the following equation:

$$M_{ij} = f \left[\frac{U_i}{U_j}, \frac{W_i}{W_j}, \frac{L_i L_j}{D_{ij}} \right]$$

where:

M_{ij} = total number of migrants from point i to point j between 1955 and 1960,

U_i, U_j = average percentages of the labor force employed in i and j between 1955 and 1960,

W_i, W_j = average annual wages and salaries paid per worker in points i and j between 1955 and 1960,

L_i, L_j = average sizes of the labor forces in points i and j between 1955 and 1960,

D_{ij} = distance in miles between points i and j.

In formulating the above equation as a regression model, we gave each variable on the right side an exponent, multiplied the seven variables on the

right, and took logarithms of both sides. The exponents, which are elasticities due to the form of the model, can be estimated by regression analysis. Since we believe that migration flows from points i to j would be impeded by unemployment in j , high wages in i and the distance between i and j , we expect the regression coefficients for these variables to be negative. On the other hand, we expect migration from i to j would be encouraged by unemployment in i , high wage rates in j , and by the sizes of the labor forces in both regions, we anticipate that the regression coefficients for these variables will be positive.

2. The Regression Model

We use SEA's as our basic unit for statistical analysis. Within the Coastal Plains there are 24 SEA's and, consequently, 552 different pairs of subregions between which people can move. We examine a random sample of 150 inter-SEA moves, but omit pairs of SEA's which have zero migration flows. This exclusion was necessary because the specification of the model uses logarithms. Migration data were obtained from the 1960 Census.¹³ Data on unemployment, wages, and the labor force were estimated for the Coastal Plains by George Iden and Charles Richter. Distances were computed by noting the longitudes and latitudes of a central point in each SEA, changing the horizontal and vertical distances between these points to miles, and using the Pythagorean Theorem.

In testing our model, we put the variables in the form specified above. The following regression result was obtained:

$$\begin{aligned} \log M_{ij} = & -7.505 + 0.751 \log U_i + 0.811 \log U_j - 0.285 \log W_i + 0.300 \log W_j \\ & \quad (2.61) \quad (2.63) \quad (-1.49) \quad (1.58) \\ & +1.212 \log L_i + 1.099 \log L_j - 1.984 \log D_{ij} \\ & \quad (8.51) \quad (9.27) \quad (-17.8) \\ \bar{R}^2 = & .74 \\ & \text{(t-ratios are in parentheses)} \end{aligned}$$

With the exception of the regression coefficient for the unemployment variable at the point of destination, all the coefficients are significantly different from zero in the anticipated directions at the .10 level of significance (U_i , L_i , L_j , and D_{ij} are significant at the .01 level). About seventy-four percent of the variation in internal migration is explained by the variables used in the model.

Only the coefficient for the variable representing unemployment at the point of destination yielded an unanticipated result. It has the wrong sign and is significantly different from zero at the .01 level of significance. Although we still believe that unemployment discourages potential migrants from entering an area, the places to which people typically migrate (the cities) do have high unemployment rates. Our data show that the average unemployment rate in the urban SEA's is 4.35 percent of the labor force verse 3.72 percent in the rural areas. Unfortunately, the rates of underemployment which tend to be high in rural areas are not included in unemployment statistics. Clearly, people do not migrate to cities because these areas have high unemployment rates. Instead, they probably go to areas with increasing job opportunities, but partially because these areas are growing, they attract unskilled, jobless people who have difficulties finding jobs anywhere. Thus they raise the unemployment rate.

Expect for the above-mentioned problem, the statistical and theoretical results of the model are favorable. The following zero order correlation matrix indicated that there is no multi-collinearity problem among pairs of regressors:

	M_{ij}	U_i	U_j	W_i	W_j	L_i	L_j
U_i	.10						
U_j	-.02	-.12					
W_i	-.08	.06	-.03				
W_j	.04	-.07	.13	-.01			
L_i	.27	-.04	-.09	-.35	-.06		
L_j	.29	-.02	.05	-.04	-.30	.03	
D_{ij}	-.65	.04	.16	-.18	-.13	.16	.14

In regression models such as ours where t-tests or normal tests are made on regression coefficients, it is important that the residuals from the regression be normally distributed. If this assumption is violated, the tests are not valid. Since our model is non-linear and non-linear models often have residuals which are not normally distributed, we believe that it is important that we test the residual for normality. It is also important that we examine the residuals for serial correlation. Serial correlation could indicate that the form of a non-linear model which we specified is wrong. Fortunately, the residuals in our model appear to be well behaved.

The residuals easily passed normality tests using a chi-square test and a Kolmogorov-Smirnov test. For both tests we divided the standard normal distribution into ten class intervals of equal areas. We then counted the number of residuals (in standard deviation units) in each class. After calculations based on the differences between the expected and observed number of observations in each class, we calculated a chi-square statistic and a Kolmogorov-Smirnov statistic. When testing the hypothesis that the distribution of the residuals is normal, both statistics failed to discredit the hypothesis at commonly used levels of significance (0.10 to 0.01).

Another examination of the residuals entailed a count of the number of sign changes in the residuals after they were ordered by increasing magnitudes of actual values of the dependent variable. There were 75 sign changes out of 149 possible. Although this counting procedure does not constitute a rigorous statistical test, if the number of sign changes were considerably different from one-half of 149, we would suspect a mis-specification in the functional form of our equation.

3. Conclusions

While our model is similar in form to one of Lowry's migration models, it is interesting to note that he used a different specification to improve his statistical results. When he used a specification like the one in this paper, all of his regression coefficients had the correct signs, but neither the coefficients for the wage variables, nor the unemployment at the point of origin were significant. He modified his model to break the labor force into the civilian and military parts and found that this yielded plausible results. This modification of the labor force variable is not necessary in our model since military movements probably represent relatively few of the moves between the SEA's of the region.

In our paper we show that a model used by Ira Lowry to explain migration flows between SMSA's is capable of being extended to explain migration flows between subregions of a low income area. Both the gravity variables (labor force sizes and distance) and economic variables (wage rates and unemployment at the point of origin) can be used to explain variations in point to point migration flows in the Coastal Plains Region. Only the variable for unemployment at the point of destination of migrants behaved in an unanticipated manner, but a posterior analysis of the sign of this variable's regression coefficient indicates that perhaps this sign should not have surprised us. We believe that a model with gravity variables and economic variables is capable of explaining population movements within region as well as migration between SMSA's.

FOOTNOTES

¹Iden, George and Charles Richter, "Factors Associated with Population Mobility in the Atlantic Coastal Plains Region," Land Economics (May, 1971), pp. 189-193.

²Stewart, John Q., "Demographic Gravitation: Evidence and Applications," Sociometry (February and March, 1948).

³For a bibliography of the uses of gravity models see Walter Isard, Methods of Regional Analysis (Cambridge: The M.I.T. Press, 1960), chapter 11. Also see Gerald Carrothers, "An Historical Review of the Gravity and Potential Concepts of Human Interaction," Journal of the American Institute of Planners (May, 1956), pp. 94-102.

⁴Stauffer, Samuel A., "Intervening Opportunities and Competing Migrants," Journal of Regional Science (Spring, 1960), pp. 1-26.

⁵Greenwood, Michael J., "An Analysis of the Determinants of Geographic Labor Mobility in the United States," Review of Economics and Statistics (May, 1969).

⁶Lowry, Ira, Migration and Metropolitan Growth, (San Francisco: Chandler Publishing Company, 1966).

⁷See Chauncy D. Harris, "The Market as a Factor in the Localization of Industry in the United States," Annals of the Association of American Geographers (December, 1954), and Edgar S. Dunn, "The Market Potential Concept and the Analysis of Location," Papers and Proceedings of the Regional Science Association (1956).

⁸Howrey, E. Philip, "On the Choice of Forecasting Models for Air Travel," Journal of Regional Science (August, 1969), and Wesley H. Long, "The Economics of Air Travel Gravity Models," Journal of Regional Science (December, 1970).

⁹Iden and Richter, op. cit.

¹⁰Persky, Joseph J. and John F. Kain, "Migration, Unemployment, and Race in the Deep South," Southern Economic Journal (January, 1970).

¹¹Lowry, op. cit.

¹²Greenwood, op. cit. Greenwood also included a variable for mean yearly temperature. This would be inappropriate in our study because temperature variations are relatively minor within subregions of the Coastal Plains.

¹³U.S. Bureau of the Census, U.S. Census of Population: 1960, Subject Reports, Migration Between State Economic Areas, Final Report PC (2) - 2E. (Washington: U.S. Government Printing Office, 1967).