MIGRATION ANALYSIS: THE ROLE OF GEOGRAPHIC SCALE

Patricia Gober-Meyers*

Abstract

The purpose of this paper is to empirically evaluate the effect of geographic scale on migration analysis by conducting similar migration studies at two levels of spatial aggregation. The analyses are concerned with the interrelationship between interregional population movement in the U.S. from 1965 to 1970 and various socio-economic, demographic, and environmental factors usually associated with migration. Recognizing the interdependent nature of these relationships, the model is specified as a system of simultaneous equations, and the parameters are estimated using a two stage least squares solution. The results indicate that scale does, in fact, influence the outcome of the analyses such that the effects of migration on regional attractiveness and the conditions under which migration occurs vary from one geographic scale to another.

I. Introduction

The purpose of this paper is to investigate the role of geographic scale in influencing the relationship between migration and regional growth. It represents one of a number of questions subsumed within a larger research project the purpose of which was to examine the overall relationship between interregional migration and the regional development process.

Similar analyses of the socio-economic, demographic, and environmental conditions associated with interregional population movement in the U.S. between 1965 and 1970 are conducted at two different geographic scales: states and metropolitan-nonmetropolitan sections of states. The basic model is specified within a system of simultaneous equations because of the conceptual and methodological advantages of viewing migration and regional development as interdependent processes. It is expected that the conditions under which migration takes place and the effects of migration upon regional growth will vary from one geographic scale to another.

---

*Assistant Professor, Department of Geography, Arizona State University, United States of America.

II. The Literature

The migration literature is characterized by a general disregard for the issue of scale and its role in influencing spatial patterns and processes. Although it is recognized that analytical results are scale specific, few attempts have been made to evaluate empirically the effect of scale in altering the results of migration studies.

The issue of geographic scale involves the choice of a meaningful areal unit at which relevant information is collected and organized. The scale of geographic investigation can vary from micro-scale analysis using relatively small size areal units such as census tracts or counties to macro-scale analysis where data are collected for large size spatial units such as states or census divisions.

The issue of scale is fundamental to a definition of the migration process itself. The term, internal migration, encompasses a wide assortment of residential movements ranging from intercounty to interstate and intercensal region moves. A migration can be defined at varying geographic scales depending upon the type of boundary crossed by the migrant in moving from the old residence to the new one.

Harvey [5] discussed the scale problem at length and concluded that "inferences as to process derived from pattern analysis are not independent of the scale of the analysis." The spatial pattern derived from large-scale analysis may imply the operation of a process different from the one implied from an examination of the same phenomenon at a smaller scale. In terms of migration research, this means that factors associated with migration at one scale may be insignificant when examining migratory behavior at different scales.

Schwind [9] alluded to this problem in his examination of the relationship between migration and regional development in the United States. Inconclusive evidence regarding the pattern of regional income convergence and the role of migration in this process may, according to Schwind, be due to a confusion regarding the effects of varying areal units, the scale problem. He suggested that smaller areal units such as counties may have exhibited a trend toward income divergence while larger regions such as states and major census regions experienced a trend toward convergence.

Willis [18] evaluated the effects of various socio-economic and spatial variables on migration in Tyneside, England between 1961 and 1966 and found that occupational structure was the principal determinant of migration flows, but the nature of this relationship depended on whether the movement was within or between regions. The proportion of a region's employment involved in specific occupations was positively associated with gross in- and out-migration but negatively related to the magnitude of intra-regional migration. In real terms, this means that areas with a large relative proportion of persons in highly skilled positions experienced high rates of in- and out-migration but low rates of internal mobility. Conversely, high rates of internal mobility and low rates of in- and out-migration were recorded in areas comprised of more semi- and unskilled workers. A plausible explanation for this pattern involves the tendency for highly skilled workers to seek employment in distant locations

while those with less training operate in more spatially restricted labor markets and therefore make more short distance, intra-regional moves. Willis's results are particularly relevant to a discussion of the role of scale on migration research because they clearly show the effect of changing levels of spatial aggregation on the results of migration analyses. The relative importance of independent variables and the nature of their relationship to migration is dependent on the size of area units which, in turn, influences whether one is dealing with short or long distance moves.

The purpose of this study is to empirically evaluate the effect of scale on interregional migration by conducting identical studies at two different levels of spatial aggregation: U.S. states and the metropolitan-nonmetropolitan subsections of states. The analysis is expected to show that the effect of various socio-economic and environmental variables on the migration process may vary depending on the size of areal unit studied.

III. Data

The major source of migration data for the present study is the 15 percent public use sample from the 1970 Census of Population. This data set, comprised of individual records, provides demographic, socio-economic, and residential information about a one-in-one hundred sample of U.S. households and the persons residing therein.

Census records contain information regarding each sample individual's state of residence in 1965 and 1970. Only persons 20 years of age and older in 1970 were included in the analysis so as to exclude children whose migratory behavior reflects the desires of others and whose impact on the regional economy is small. It is possible to determine whether individuals are interstate migrants by comparing their states of residence in 1965 and 1970. The separation of migrants from non-migrants at the metropolitan-nonmetropolitan scale requires a more indirect approach. In addition to indicating the state of residence, the census records whether persons' 1965 and 1970 residences were metropolitan or nonmetropolitan in character. At this scale, those defined as migrants include all interstate migrants as well as those who moved from metropolitan to nonmetropolitan or from nonmetropolitan to metropolitan areas of the same state. The nature of the data set necessitated that areal units at this scale consist of all metropolitan areas within a given state rather than single SMSA's. Similarly, all nonmetropolitan counties are grouped together and comprise a single unit of investigation.

Unfortunately, data limitations made it impossible to include the universe of metropolitan and nonmetropolitan areas into the analysis because 1970 was the first time that the Census provided state of residence data for migration research. In order to adhere to their disclosure rule, the Census was required to delete the metropolitan-nonmetropolitan residence classification for states with too few metropolitan areas. As a result, the following twelve states were excluded from investigation: Arizona, Delaware, Hawaii, Maine, Montana, Nevada, New Hampshire, Idaho, North Dakota, Rhode Island, South Dakota, and Utah. Maryland and Connecticut were also deleted since the Census

failed to distinguish between urban metropolitan and nonmetropolitan residence and this made it impossible to study the vast majority of persons living in those states.

Data relating to the indicators of regional attractiveness are obtained from various government publications. They include: the Survey of Current Business (per capita income) [13], County and City Data Book (per capita income and population density) [11], World Weather Records, North America (January temperatures) [17], U. S. Census of Population (population density and education) [12], Vital Statistics for the U.S. (fertility) [16], Manpower Report of the President (employment growth and unemployment) [15], and Employment and Earnings, States and Areas (employment growth) [14].

IV. Study Model

The traditional view of the relationship between migration and economic change treats migration as the dependent variable and economic factors as independent variables in a single-equation regression model. More recently, the interdependent nature of the relationship between migration and economic growth has been recognized by Okun [7], Muth [6], Olvey [8], and Greenwood [2, 3, 4] who expressed the migration-economic change relationship as a system of simultaneous equations and estimated model parameters with a two or three stage least squares approach. Because of the methodological and conceptual advantages of treating migration in such a framework, the study model is represented as a system of simultaneous equations wherein net migration, an age sex race and education selectivity factor, and per capita income growth are endogenous variables. Other independent variables represent economic, demographic, and environmental conditions frequently found in models of migration. The study model is expressed in equations 1, 2, and 3.

$$Y_{1i} = a_1 + a_2Y_{2i} + a_3Y_{3i} + a_4X_{1i} \quad (1)$$

$$Y_{2i} = b_1 + b_2Y_{1i} + b_3X_{1i} + b_4X_{2i} + b_5X_{4i} + b_6X_{5i} + b_7X_{6i} \quad (2)$$

$$Y_{3i} = c_1 + c_2Y_{1i} + c_3X_{1i} + c_4X_{2i} + c_5X_{3i} + c_6X_{5i} + c_7X_{6i}$$

Where,

Y_{1i} = Growth in per capita income from 1965 to 1970 in region i

Y_{2i} = The rate of net migration from 1965 to 1970 in region i

Y_{3i} = The net benefits that accrue to region i as a result of changes in the age, sex, race, and education composition caused by interregional migration, 1965 to 1970

X_{1i} = The growth in nonagricultural employment from 1965 to 1970 in region i

X_{2i} = Average annual unemployment rate in region i for the period 1965 to 1970

X_{3i} = Median years of school completed for persons over 25 years of age in region i

X_{4i} = Crude birth rate in 1950 in region i

X_{5i} = Population density, the number of persons per square mile in 1965 in region i

X_{6i} = Climatic factor mean January temperature in the capital or largest city of region i

With the exception of the migration benefit factor, variables within the system are commonly found in migration models. The migration benefit factor was designed as a refinement to Okun's [7] age-sex selectivity factor, and it measures the change in population composition resulting from the interregional migration process. Measures of desirability were attributed to persons in the sample on the basis of the average income earned in 1967 by their age, sex, race, and educational subgroup of the national population. An aggregate regional measure was obtained by comparing the desirability of the population before migration occurred in 1965 and after it was completed in 1970. For a more complete discussion of the migration benefit variable, see Appendix A.

Reflecting their mutual interdependence, net migration, migration benefits and per capita income growth are expressed as both dependent and independent variables within the simultaneous system of equations. The two migration variables, surrogates for changes in population size and composition resulting from interregional population movements, are influenced by and, in turn, affect a region's rate of per capita income growth. The model, therefore, recognizes the two-way interaction between migration and economic growth.

Exogenous variables have been employed in past migration research and represent conditions of regional attractiveness that affect per capita income growth and/or population movement. Growth in nonagricultural employment is expected to stimulate per capita income growth and in-migration especially among the most desirable segments of the population since they are most aware of the changing distribution of job opportunities and best equipped to take advantage of them. In addition, the highest rates of net migration are expected in regions exhibiting the lowest unemployment rates, and migrant selectivity should operate so as to favor regions of low unemployment and favorable educational conditions.

The environmental factors of climate and population density are expected to affect both net migration and migration benefits. Warm climates, as evidenced by high mean January temperatures, and the availability of open space indicated by low population density are expected to be related to positive rates of net migration. Moreover, high migration benefits are expected in regions with favorable environmental circumstances because high income people can afford to make locational decisions based on noneconomic factors of which climate and population density are prime examples.

A commonly used demographic variable in migration models involves an indicator of fertility during an earlier time period. It is thought that out-migration may be induced by population pressure resulting from high fertility 15 to 20 years earlier. Accordingly, the crude birth rate in 1950 is employed as an independent variable in the net migration function, and a negative relationship is expected.

The Y variables represent factors endogenous to the system while the X's represent factors that are exogenously determined. Equations 1, 2, and 3 represent a complete set of equations, meaning that there are as many equations as there are endogenous variables. The parameters were estimated using two stage least squares regression analysis. In the first stage, each endogenous variable acts as a dependent variable in a single equation regression model with all of the exogenous factors as independent variables. In the second stage, the estimated endogenous variables replace observed ones in the original set of equations, and the coefficients are estimated using ordinary least squares. Although biased, the results of the two stage least squares solution are consistent and more efficient than if ordinary least squares had been applied. For a more thorough discussion of the methodology, see Theil (1953) and Basmann (1957).

V. Results

The results of the analysis are presented in Table 1. Only variables whose coefficients proved significantly different from zero at the .95 level of confidence were included in the table. F values, presented in parentheses, correspond to regression coefficients.

TABLE 1

REGRESSION RESULTS FOR STATES AND METROPOLITAN NON-METROPOLITAN AREAS

States*

$$Y_1 = .4214 + .0000047Y_3 \quad (F=6.2)$$

$$Y_2 = -.0575 + .2996X_1 + -.000012X_5 \quad (F=18.7) \quad (F=11.4)$$

$$Y_3 = -1454.37 + 7454.22X_1 \quad (F=9.4)$$

Metropolitan-Nonmetropolitan Areas*

$$Y_1 = .2236 + .2238X_1 \quad (F=9.2)$$

$$Y_2 = 3.66 + 41.71X_1 + -.0689X_4 + 1.79X_2 + -.0011X_5 \quad (F=7.5) \quad (F=8.0) \quad (F=4.6) \quad (F=11.2)$$

$$Y_3 = -3210.39 + 9854.86X_1 + 475.65X_2 \quad (F=6.5) \quad (F=8.3)$$

Y_1 = per capita income growth
Y_2 = net migration rate
Y_3 = migration benefits
X_1 = growth in nonagricultural employment
X_2 = unemployment rate
X_3 = median years of school completed
X_4 = fertility in 1950

X_5 = persons per square mile, population density
X_6 = mean January temperature

*All included variables had coefficients significantly different from zero at the .95 level of confidence.

The results indicate that altering the size of areal units did, indeed, cause different outcomes in the regression analyses. One of the more notable differences occurred in the per capita income growth function. In this case, only the employment growth coefficient was statistically significant at the metropolitan-nonmetropolitan scale while only the parameter of the migration benefit variable proved to be significant at the state scale.

The net migration function at both scales was similar in that coefficients for population density and employment growth were statistically significant and in the expected direction. There was a trend at both scales for population to migrate from low to high employment growth regions and from high to low density locations. The concentrated urban-industrial areas of the Northeast and Midwest failed to attract population while less dense regions of the U.S. experienced sizable in-migration during the study period. This result should not, however, be interpreted as a migratory trend away from metropolitan areas in general for there were many metropolitan areas, particularly in the South and West, with lower than average densities, and they experienced significant in-migration from 1965-1970.

Additionally, fertility and unemployment proved to be significant only at the metropolitan-nonmetropolitan scale. As expected, the sign of the coefficient for fertility was negative indicating that regions of high fertility during a previous period were characterized by out-migration during the study period. The failure of fertility to be significant at the state scale is not surprising when one considers the relatively small differentiation in 1950 birth rates among states. The most meaningful variation in fertility occurred between metropolitan and nonmetropolitan areas and, therefore, a significant relationship appeared only at this scale.

A major variation from one scale to the next involved the unemployment factor. At the metropolitan-nonmetropolitan scale, the unemployment parameter was statistically significant and in a positive direction but was insignificant at the state scale. A close examination of unemployment rates for the different data sets reveals that, in a sizable number of cases, the overall state unemployment rate disguised substantially different experiences for the metropolitan and nonmetropolitan subsections. The role of unemployment rates as a force in influencing interregional population movements did not, therefore, come to light until the smaller areal units were considered.

The positive sign of the unemployment coefficient in the net migration and migration benefit functions was unexpected since high unemployment is generally regarded as a deterrent to in-migration in general and, in particular, to high income people who are thought to possess detailed information about potential destinations. In this analysis, high unemployment rates were associated with a favorable change in the composition of the region's population. This

could be explained by a situation where unemployment rates exert a strong influence over the migratory behavior of low income persons but are relatively unimportant in affecting the behavior of those at the upper end of the income ladder. High income people are less influenced by local unemployment rates either because they have arranged for employment before moving or are moving for other than employment related reasons. On the other hand, those with lower than average incomes are more likely to be employed in the kinds of semi- and unskilled occupations that are highly dependent on local economic conditions, and it is less likely that they have secured a job before actually migrating. Naturally, this hypothesis requires further validation, but such an explanation is consistent with the findings of this analysis.

VI. Summary

The results show that outcomes of analyses conducted at varying geographic scales are different. Although several variables such as employment growth and population density were significant at both scales, important variations came to light. In the per capita income growth function, for example, different variables were related to per capita income growth, the dependent variable. Significant explanatory factors included employment growth at the metropolitan-nonmetropolitan scale and, at the state scale, changes in population composition measured by the migration benefit variable.

Fertility and unemployment were significantly related to net migration, and unemployment was associated with migration benefits only at the metropolitan-nonmetropolitan scale. A plausible explanation involves the high degree of intra-state heterogeneity ignored at the state scale but taken into account at the metropolitan-nonmetropolitan scale. Data aggregated at the state level exhibited relatively small variation while the metropolitan-nonmetropolitan data reflected the high degree of differentiation in both fertility and unemployment.

The results of this analysis should underscore the need for careful consideration of the scale factor in migration research. Studies conducted at different geographic scales may yield different results, and it is only by examining these results in the context of the areal unit being studied that we can obtain a more comprehensive view of the role of scale in migration analysis.

Appendix A

The migration benefit selectivity variable in the migration model includes the change in age, sex, race, and educational composition of the population. These factors were chosen because of their ability to differentiate those personal characteristics that are highly related to the income earning ability of an individual. On the basis of age, sex, race, and education, the study sample was divided into 168 subgroups. The subgroup consisting of white males 55 to 64 years of age with more than five years of higher education had the highest per capita income in 1967 and was assigned a desirability score of 1.0. The values of all other groups represented some proportion of that.

Average desirability scores were calculated for the in-, out-, and

non-migrants of each region and were combined in equation 4 to represent the overall regional measure of benefits or losses from migration.

$$Y_{3i} = \frac{A_{1i}(N_{1i}) + A_{2i}(N_{2i})}{N_{1i} + N_{2i}} - \frac{A_{3i}(N_{3i}) + A_{2i}(N_{2i})}{N_{3i} + N_{2i}} \qquad (4)$$

Where,

Y_{3i} = Net benefits or losses to region i as a result of interregional migration

A_{1i} = The average desirability score for the in-migrants to region i

N_{1i} = The number of in-migrants to region i

A_{2i} = The average desirability score for the non-migrants of region i

N_{2i} = The number of non-migrants in region i

A_{3i} = The average desirability score for the out-migrants from region i

N_{3i} = The number of out-migrants from region i

The first part of equation 4 represents the region's population composition in 1970 after migration has occurred. The sizes and average desirability score for in-migrants and non-migrants are used to calculate an aggregate average desirability score for the in-migrants and non-migrants. The desirability of the region's population composition in 1965, prior to migration, is determined by the second part of equation 4 which includes the sizes and average desirability scores for non-migrants and out-migrants. By subtracting the indicator of population composition before migration from the post migration measure, it is possible to arrive at an indicator of the overall impact of interregional migration, 1965 to 1970, upon each region's population composition.

In order for a region to have a large positive migration benefit score, it is necessary for the region's in-migrants to be more desirable than out-migrants and for the number of in-migrants to be large relative to the size of the non-migrant population. In regions with negative scores on the migration benefit measure, out-migrants are more desirable than in-migrants, and the size of the out-migrant population is large in relation to the number of non-migrants. It should be noted that the analysis is meant only as a short term, five year investigation of the effects of migration on regions' average income. If the purpose were to predict future impacts, then it would be unreasonable to treat men 55 to 64 years of age as highly desirable since their income is likely to be severely reduced in the succeeding ten to fifteen years. The analysis is not designed to assess the future or long run effects of migration on the regional economy but rather to measure what actually happened during the study period.

REFERENCES

1. Basmann, R. L. "A Generalized Classical Model of Linear Estimation of Coefficients in a Structural Equation." Econometrica, XX (1957), 77-84.
2. Greenwood, Michael J. "Urban Economic Growth and Migration: Their Interaction." Environment and Planning, V (January, 1973), 91-112.
3. _____ . "Simultaneity Bias in Migration Models: An Empirical Examination." Demography, XII (August, 1975), 519-36.
4. _____ . "A Simultaneous-Equations Model of Urban Growth and Migration." Journal of the American Statistical Association, LXX (December, 1975), 797-809.
5. Harvey, David. Explanation in Geography. New York: St. Martin's Press, 1969.
6. Muth, Richard F. "Migration: Chicken or Egg." Southern Economic Journal, XXXVII (January, 1971), 295-306.
7. Okun, Bernard. "Interstate Population Migration and State Income Inequality. A Simultaneous Equation Approach." Economic Development and Cultural Change, XVI (January, 1968), 297-311.
8. Olvey, Lee D. "Regional Growth and Inter-Regional Migration: Their Pattern of Interaction." Review of Regional Studies, II (Winter, 1972), 139-64.
9. Schwind, Paul J. Migration and Regional Development in the United States 1955-1960. Department of Geography Research Paper N. 133, Chicago: University of Chicago, 1971.
10. Theil, H. Estimation and Simultaneous Correlation in Complete Equation Systems. The Hague: Central Plan Bureau, 1953.
11. U. S. Bureau of the Census. County and City Data Book, 1967. Washington, D.C.: Government Printing Office, 1967.
12. _____ . U. S. Census of Population, 1970--General Social and Economic Characteristics. PC(1)-C1 through 52, Washington, D.C.: Government Printing Office, 1970.
13. U. S. Bureau of Economic Analysis. Survey of Current Business, LII (1972), 20.
14. U. S. Department of Labor. Employment and Earnings, States and Areas, 1931-71. Bulletin 1370-9, Washington, D.C.: Government Printing Office, 1973.
15. _____ . Manpower Report of the President. Washington, D.C.: Government Printing Office, 1972.
16. U.S. Public Health Service. Vital Statistics of the United States, 1950. Washington, D.C.: Government Printing Office, 1950.
17. U. S. Weather Bureau. World Weather Records, 1951-1960, North America. Washington, D.C.: Government Printing Office, 1965.
18. Willis, Kenneth G. "The Influence of Spatial Structure and Socio-Economic Factors on Migration Rates." Regional Studies, VI (March, 1972), 69-82.