DEMOGRAPHY© Volume 15, Number 3 August 1978

THE INTERACTION OF MIGRATION, INCOME, AND EMPLOYMENT IN SWEDEN

Åke Dahlberg
The National Labour Market Board, Stockholm, Sweden

Bertil Holmlund
The Industrial Institute for Economic and Social Research, Stockholm, Sweden

Abstract—A common approach in migration studies is to explain interregional migration by single-equation models. Such models are likely to suffer from simultaneous-equation bias when used in studies attempting to analyze migration over a long period of time. In this study, a simultaneous-equation approach is applied, which takes account of the interdependence between migration and income and employment changes. The four-equation model is estimated for 70 labor market areas in Sweden.

INTRODUCTION

A common approach in the international literature on migration is to explain interregional migration by means of a single-equation multiple-regression model. Such models have proved to suffer from a "simultaneous-equation bias," in particular in studies which attempt to analyze migration over a long period of time; it is reasonable to suppose that the migration studied has influenced the independent variables of the models. Besides the fact that more realistic explanatory models of migration may be constructed on the basis of simultaneous-equation models, a rough picture of some effects of migration may also emerge from this approach. For a survey of the literature on interregional migration, see Greenwood (1975a).

SPECIFICATION OF THE MODEL

The basic model used in this study consists of four structural equations, which will be further explained in the text. A similar approach has been used by Greenwood (1973, 1975b).

$$\mathrm{OM} = f_1(\overset{+}{\mathrm{IM}}, \overset{-}{\Delta\mathrm{INC}}, \overset{-}{\Delta\mathrm{EMPR}}, \overset{-}{\mathrm{INC\ 60}}, \overset{+}{\mathrm{UNEMP\ 60}}, \overset{-}{\mathrm{EMPR\ 60}}), \quad (1)$$

$$\mathrm{IM} = f_2(\overset{+}{\mathrm{OM}}, \overset{+}{\Delta\mathrm{INC}}, \overset{+}{\Delta\mathrm{EMPR}}, \overset{+}{\mathrm{INC\ 60}}, \overset{-}{\mathrm{UNEMP\ 60}}, \overset{+}{\mathrm{EMPR\ 60}}), \quad (2)$$

$$\Delta\mathrm{INC} = f_3(\overset{?}{\mathrm{OM}}, \overset{?}{\mathrm{IM}}, \overset{+}{\mathrm{EMPR\ 60}}, \overset{+}{\mathrm{DNORRL}}, \overset{-}{\mathrm{DBIGTOWN}}), \quad (3)$$

and

$$\Delta\mathrm{EMPR} = f_4(\overset{?}{\mathrm{OM}}, \overset{?}{\mathrm{IM}}, \overset{+}{\mathrm{EMPR\ 60}}, \overset{+}{\mathrm{AGE\ 60}}, \overset{+}{\mathrm{DNORRL}}, \overset{-}{\mathrm{DBIGTOWN}}). \quad (4)$$

Endogenous Variables

OM: rate of out-migration from the 70 so-called A-regions of Sweden, i.e., the number of people who moved from one

A-region to other A-regions in the country between 1961 and 1970, divided by the population figure of the sending region in 1961. (Sweden is divided into 70 regions called "A-regions," in which the population ranges between 20,000 and 1,070,000 people.)

IM: rate of in-migration to the A-regions of Sweden (the number of people who moved to one A-region from other A-regions between 1961 and 1970 divided by the population figure of the receiving A-region in 1961).

ΔINC: relative income change in the A-region level (total net income per earner in 1970 divided by the corresponding value in 1960).

ΔEMPR: relative change in employment ratio in the A-region (the employment ratio in 1970 divided by the corresponding value in 1960).

Exogenous Variables

INC 60: income per employee in 1960 in the A-region (total net income).

UNEMP 60: percentage rate of unemployment in 1960 in the A-region (the number of unemployed registered at the labor exchanges in 1960 divided by the labor force).

EMPR 60: employment ratio in the A-region in 1960 in terms of percentage (the number of employed according to the census of 1960 divided by the population).

AGE 60: percentage share of the population between 16 and 64 years of age.

DNORRL: dummy variable referring to A-regions in Norrland (that is, the north of Sweden, which is thinly populated).

DBIGTOWN: dummy variable referring to A-regions in the metropolitan areas of Stockholm, Gothenburg, and Malmoe.

The theoretical framework for the migration analysis is basically the human capital approach introduced by Sjaastad (1962) and modified by Todaro (1969). Individuals are assumed to behave as if they were computing present values of investment in migration to alternative destinations. Regions with the highest returns to migration will be chosen. The expected net present value of investment in migration from region i to region j (G_{ij}) may be expressed as

$$G_{ij} = \sum_{t=1}^{T} \frac{(P_{jt}Y_{jt} - P_{it}Y_{it})}{(1 + r)^t} - C_{ij},$$

where

P_{jt} = the probability of being employed in region j in period t,
P_{it} = the probability of being employed in region i in period t,
Y_{jt} = the average income in region j in period t,
Y_{it} = the average income in region i in period t,
C_{ij} = the costs of migration from i to j,
and
r = the rate of discount.

Individuals are thus likely to take job opportunities as well as income differentials into account in their migration decisions. We may expect, therefore, that the higher the average income is in a region, the fewer the people who will leave it. The opposite is expected to be true of in-migration. For this reason, the average income per earner at the outset of the investigation has been included as an independent variable in the equations for out- and in-migration. Furthermore, the expected future income in sending and receiving regions will affect calculated present values. The actual rise in income per earner during the period studied has accordingly been used as an additional explanatory variable in the migration equations.

The expected returns from migration are also dependent on the probabilities of finding and keeping jobs in different regions. The employment ratios in 1960 are used as proxies for regional vacancy rates, representing the probability of locating job opportunities. The growth of the employment ratios are included as proxies for expected future changes in job availability. The hypotheses are that a higher employment ratio in 1960 and a higher increase in this ratio should be linked with a lower rate of out-migration and a higher rate of in-migration.

Finding a job vacancy will not always mean that a job offer is obtained; potential employers will, to some extent, examine the applicants in order to determine their productivities. If the rate of unemployment is high, firms are likely to set higher "recruitment thresholds," thereby reducing the probability of an applicant's receiving a wage offer from a job vacancy. A higher regional unemployment rate will thus encourage out-migration and discourage in-migration.

The rate of out-migration from a region is also assumed positively dependent on the rate of in-migration (and vice versa). Regions with a high rate of in-migration will in time receive a large proportion of people who are "migration-prone" [cf. Stone (1971)]. It has been shown in various studies that people who have moved at least once are more likely to migrate than people who have never moved. The acceptance of wage offers are likely to be based on some ignorance of nonpecuniary characteristics of the new job and the new region. The actual experiences in the new locality might result in disappointment and search for a new job in the "old" region or in some other locality. It is therefore also likely that a high rate of out-migration will positively affect the rate of in-migration through a substantial amount of return migration.

Now consider the income change equation (3). According to traditional economic theory of factor price adjustments, labor mobility will lead to a tendency towards wage equalization between regions. The assumption is then that labor supply in the sending and receiving localities is changed, whereas labor demand is assumed unaffected. This hypothesis appears unrealistic in a modern market economy. Migration from a region leads to a decreasing demand for private as well as public goods and services. This is true even if the migrants are unemployed, since they normally receive unemployment relief and other transfer incomes, which are chiefly derived from other regions. Increasing migration from regions with negative net migration will probably also contribute to strengthening the negative expectations about the future development of these regions and lead to a decreasing propensity to make investments in both the private and the public sector. This indicates that the demand curve on the labor market in the sending region will be shifted towards the origin, so incomes fall with out-migration.

High in-migration to a region, analogously, can lead to an outward shift of the labor demand curve. It might be appropriate here to pay special attention to expanding regions with fully utilized social overhead capital, where in-migration may be expected to lead to substantial investment in housing, schools, roads, etc. It is not easy to make any a priori assumptions about the relative magnitudes of the migration-induced supply and demand shifts in the labor market. The empirical results will be the only guide to knowledge in this case.

Income growth in a region is, however, dependent on other factors besides migration. Thus, income growth will vary with the potentialities of utilizing agglomeration economies (economies of scale, localization economies, and urbanization economies) in different regions. As a proxy variable for such conditions, we have chosen the employment ratio in various regions at the outset (that is, in 1960 in this study). Various Swedish studies show that the employment ratio is well correlated with variables expressing the agglomeration economies of the regions. The coefficient of this variable is expected to be positive. Income growth may, however, vary between regions for reasons other than differences in migration and employment. More "institutional" conditions, e.g., changes in subsidies to certain regions or low-wage groups, which chiefly affect some parts of the country, may be of importance in this respect. To operationalize such institutional conditions is complicated, however. We have chosen here to take them into consideration by introducing dummy variables for Norrland and for metropolitan areas. The introduction of the Swedish regional policy during the late sixties provides a rationale for this; one aim of the policy was to decrease regional income differences.

The equation for the change of the employment ratio (4) is chiefly based on the same variables as the equation for income change. If we had chosen instead the number of employed as the dependent variable in this equation, it would have been reasonable to assume that the coefficient of the out-migration variable should be negative and the coefficient of the in-migration variable positive, in view of the discussion [in connection with equation (3)] of shifts in the supply and demand curves on the labor market in sending and receiving regions. On the other hand, it is more difficult to offer an opinion on changes in the employment ratio. The decisive factor for these changes is the shape of the demand and supply curves and the shifts in these curves that are caused by migration.

We have also assumed that changes in the employment ratio are affected by agglomeration factors; the employment ratio in 1960 has been chosen as a proxy variable. Furthermore, we have included the regional dummy variables, taking account of changes in the Swedish regional policy.

Employment change is probably also linked with more demographic variables. Above all, the share of the population in the most active age groups in different regions might be expected to explain some of the variations in the employment ratio. We know empirically that it is chiefly women between 25 and 55 years of age that have contributed to the increase in the employment ratio in Sweden since the beginning of the sixties. Consequently, we have also included here as an independent variable the share of the population in the most active ages (that is, about 15 to 64 years of age). It is assumed that the higher this share, the higher will be the growth of the employment ratio.

RESULTS

The equations have been estimated by the two-stage least-squares method (TSLS). All equations satisfy the order condition for identifiability—i.e., the number of excluded exogenous variables is at least as great as the number of included endogenous regressors. In the second stage of the TSLS procedure, a stepwise multiple regression technique was applied. When we use stepwise regression, multicollinearity is more easily detected.

The results of the estimations are set out in Table 1. As regards the estimates of the migration equations, the signs of five out of six coefficients correspond to our hypotheses. Only the coefficient of income growth exhibits signs that differ from what was assumed. For both equations, only the coefficients of the migration variables are significant according to Table 1. However, these results relate to the final step of the stepwise multiple regression analysis. In the equation for out-migration, significant coefficients were obtained at an earlier step (step 3) for income growth, income in 1960, and in-migration. Because of a certain degree of intercorrelation between unemployment in 1960 and, above all, income growth ($R = -0.76$) and between unemployment and in-migration ($R = -0.59$), the coefficients of the latter two variables become insignificant when unemployment in 1960 is included in the analysis (i.e., in step 4). The same is true of the equation for in-migration, in which out-migration as well as income in 1960 and unemployment in 1960 have significant coefficients at earlier steps. The results agree well with the results of other studies in which similar approaches were used [see, for example, Greenwood (1973, 1975a)].

As regards the structural equation for income growth, no definite statement was made about the signs of the migration coefficients. The results of the estimations show, however, that the coefficient of the

Table 1.—Parameter Estimates for the Structural Equations, Two-Stage Least Squares (Standard Errors in Parentheses)

Independent Variables	Dependent Variables			
	OM	IM	ΔINC	ΔEMPR
OM		8.43×10^{-1} [a] (0.36)	-1.92×10^{2} [a] (29.85)	-1.97×10 (11.04)
IM	1.18[a] (0.40)		7.36×10[a] (21.15)	1.78×10[a] (7.39)
ΔINC	2.90×10^{-3} (0.01)	-2.44×10^{-3} (0.01)		
ΔEMPR	-4.35×10^{-2} (0.04)	3.69×10^{-2} (0.04)		
INC 60	-5.38×10^{-6} (0.00)	4.55×10^{-6} (0.000)		
UNEMP 60	3.87×10^{-2} (0.06)	-3.26×10^{-2} (0.06)		
EMPR 60	-2.27×10^{-2} (0.02)	1.92×10^{-2} (0.02)	-8.57×10^{-1} (0.49)	-5.40×10^{-1} [a] (0.23)
DNORRL			-1.90 (2.25)	-5.05×10^{-1} (0.82)
DBIGTOWN			-2.27×10[a] (4.17)	-7.41×10^{-1} (1.55)
AGE 60				4.94×10^{-2} (0.25)
CNSTNT	4.67×10^{-1}	-3.95×10^{-1}	2.26×10^{2}	1.93×10
R^2	0.31	0.42	0.57	0.18

a - Significant at the 5 percent level. The number of observations is 70.

out-migration variable is negative (and significant) and the coefficient of the in-migration variable positive (and significant). These results indicate that the shift of the demand curve dominates the shift of the supply curve on the labor market in both the sending and the receiving locality. In view of what was said above, such results are not necessarily to be regarded as theoretically impossible. Moreover, similar results have emerged in studies carried out in the United States by, for example, Greenwood. A plausible additional explanation of the results might be the socioeconomical characteristics of the migrants; the typical pattern is that persons with high education and a high income migrate more often and over longer distances than others (National Central Bureau of Statistics, 1977, p. 7). As to other variables in the income equation, the coefficient of the employment ratio is

Table 2.—Ordinary Least Squares Estimates for the Structural Equations (Standard Errors in Parentheses)

Independent Variables	Dependent Variables: OM	IM	ΔINC	ΔEMPR
OM		8.66×10^{-1} [b] (0.07)	-5.32×10 [b] (14.35)	-8.57 [b] (3.86)
IM	7.95×10^{-1} [b] (0.07)		3.16×10 [b] (12.57)	8.76 [b] (3.43)
ΔINC	-3.49×10^{-3} [b] (0.001)	3.58×10^{-3} [b] (0.001)		
ΔEMPR	-4.66×10^{-3} (0.003)	5.17×10^{-3} (0.004)		
INC 60	-1.92×10^{-5} [b] (0.00)	3.17×10^{-5} [b] (0.00)		
UNEMP 60	2.79×10^{-2} [b] (0.01)	-4.45×10^{-2} [b] (0.01)		
EMPR 60	-6.28×10^{-3} (0.004)	1.45×10^{-3} (0.01)	3.65×10^{-1} (0.49)	-3.81×10^{-1} [b] (0.17)
DNORRL			-3.47 (2.38)	-9.80×10^{-1} (0.73)
DBIGTOWN[a]			-9.98 [b] (4.46)	
AGE 60				-1.58×10^{-2} (0.22)
CNSTNT	1.06	-8.88×10^{-1}	1.37×10^{2}	1.59×10
R^2	0.74	0.79	0.35	0.19

a - DBIGTOWN displayed too small partial correlations to be included in the final step of the regression analysis.

b - Significant at the 5 percent level.

negative (but insignificant), and the coefficients of the dummy variables for the two "extreme regions," Norrland and the big towns, are also negative, which means that income growth during the period studied has been greater in other types of regions.

In the equation for the change of the employment ratio, the coefficient of the out-migration variable is negative (but insignificant) and the coefficient of the in-migration variable positive (and significant). For other variables, too, the signs are in accord with the signs of the corresponding variables in the income equation. The negative sign of the employment ratio variable might be caused by spurious correlation, since it is the denominator of the dependent variable.

An alternative estimation of the struc-

Table 3.—Coefficients of the Reduced Forms for the Structural Equations

Exogenous Variables	Dependent Variables OM	IM	ΔINC	ΔEMPR
INC	-0.0000017	0.0000031	0.0000023	0.0000009
UNEMP 60	0.017100	-0.018238	-0.013753	-0.006611
EMPR 60	-0.003161	-0.001313	-0.857405	-0.540010
DNORRL	0.002263	-0.012048	-1.907911	-0.507486
DBIGTOWN	0.006970	-0.020840	-0.242674	-0.74564
AGE 60	-0.000503	0.001395	0.001037	0.049737
CONSTANT	0.116751	-0.288969	-0.192328	0.118456

tural equations was carried out by means of ordinary least squares (OLS). For the migration equations, these estimations yielded more significant coefficients, and R^2 was considerably higher. On the other hand, the differences are smaller for the other two equations. (See Table 2.)

The reduced-form coefficients were calculated on the basis of the coefficients in the structural equations that were estimated by means of the two-stage least-squares method. In contradistinction to the coefficients in the structural forms, the coefficients in these reduced forms account for "feedback" effects inherent in a simultaneous system. The results of the reduced forms are shown in Table 3.

Previous studies in which single-equation models were used to explain out- and in-migration often yielded signs for the coefficients of unemployment, income, and income growth that differed from what had been expected. The hypothesis that these deviations may be explained by simultaneous-equation bias has sometimes been put forward in the literature. It has been supported by investigations carried out by Greenwood (1973, 1975a), who used simultaneous-equation models. The present study also lends support to the hypothesis. Thus, in the equation for out-migration the expected signs of all the income and employment variables were obtained. The higher the income and the employment ratio in 1960, the less out-migration took place between 1960 and 1970, while the opposite is true of unemployment in 1960. In the equation for in-migration, the expected signs of the coefficients of income in 1960 and unemployment in 1960 were obtained, while the coefficient of the employment ratio in 1960 had an unexpected sign.

In the equation for income change, the same sign for the coefficients as in the equation for the change of the employment ratio was obtained for all the variables. Some interesting results may be noted here. For example, the higher the level of income in 1960 and the lower the unemployment, the greater the growth of both income and employment ratios during 1960 to 1970. The coefficients of the regional dummy variables show that the growth of income and employment has been smaller in Norrland and the big towns than in the other regions.

CONCLUSION

Simultaneous equations are a useful method for explaining migration over a long period of time. This approach produces a better specification of the model than do single-equation models. The empirical results reported in this study lend support to the simultaneous equation approach. Among other things, we have been able to demonstrate the interdependence between migration and income and employment changes during the period studied. Some of these results may seem unexpected. The study shows, for example, that migration from a region seems to lead to decreasing income and employment growth in the sending region. The opposite is true, however, of in-migration. Hence, the results lend no support to the hypothesis that the extensive migration that took place in Sweden during the sixties should have contributed to leveling regional differences in income and employment ratio. The results rather point to the opposite conclusion.

ACKNOWLEDGMENT

We have benefited from constructive comments of an anonymous referee.

REFERENCES

Greenwood, M. J. 1973. Urban Economic Growth and Migration: Their Interaction. Environment and Planning 5:91–112.

———. 1975a. A Simultaneous-Equation Model of Urban Growth and Migration. Journal of the American Statistical Association 70:797–810.

———. 1975b. Research on Internal Migration in the United States: A Survey. Journal of Economic Literature 13:397–433.

National Central Bureau of Statistics. 1977. Forecasting Information 1976. Stockholm: National Central Bureau of Statistics.

Stone, L. O. 1971. On the Correlation Between Metropolitan Area In- and Out-Migration by Occupation. Journal of the American Statistical Association 66:693–701.

Sjaastad, L. A. 1962. The Costs and Returns of Human Migration. Journal of Political Economy, Supplement 70 (5):80–93.

Todaro, M. P. 1969. A Model of Labor Migration and Urban Unemployment in Less Developed Countries. American Economic Review 59:138–148.