STATE AND REGIONAL RESEARCH CENTER
WORKING PAPER 88-03
OCTOBER 1988

REGIONAL LABOR MARKETS:
ECONOMETRIC ANALYSIS*

by
Michael Grosse
Glenn Nelson**

* The financial support of the Agricultural Experiment Station, Minnesota Department of Finance, and Northwest Area Foundation contributed to this study.

** Research manager, Maritz Marketing Research, Minneapolis, and professor, Department of Agricultural and Applied Economics, University of Minnesota, St. Paul, respectively.

Working papers are published without formal review within the State and Regional Research Center. The opinions expressed in Working Papers are not necessarily held by the affiliated faculty or the Council of the Center.

The University of Minnesota is committed to the policy that all persons shall have equal access to its programs, facilities, and employment without regard to race, religion, color, sex, national origin, handicap, age, or veteran status.
TABLE OF CONTENTS

I. INTRODUCTION ........................................... 1

II. MIGRATION .............................................. 2
    A Brief History of the Model’s Origin .................. 3
    Functional Form and Explanation of Variables .......... 9

III. WAGES AND EMPLOYMENT .................................. 12
    The Neoclassical Approach ............................... 13
    Limitations ............................................. 16
    A Dynamic Regional Labor Market Model ................ 19

IV. EMPIRICAL ANALYSIS ..................................... 22
    Ordinary Least Squares Estimates ....................... 24

V. SUGGESTIONS FOR FURTHER RESEARCH ..................... 32

APPENDIX I - MIGRATION DATA ............................... 36

APPENDIX II - FIRST STAGE ESTIMATES ....................... 38

ENDNOTES .................................................. 41

BIBLIOGRAPHY ............................................... 45
LIST OF FIGURES

Figure 1 - Minnesota's Regional Labor Markets. . . . . . . . . . . . . 1

Figure 2 - Rates of Migration by Age Group . . . . . . . . . . . . . . 11

Figure 3 - Output, Profit, and Quantity of Labor Demanded. . . . 13

Figure 4 - Individual Labor Supply . . . . . . . . . . . . . . . . . . . 14

Figure 5 - Emile Cheysson's Wage Adjustment Process. . . . . 18

Figure 6 - The Problem of Indentification. . . . . . . . . . . . . . . 32
I. INTRODUCTION

Regional economics is a relatively young field within the discipline of economics. Interest in regional issues has been spurred by an increasing concern for, and awareness of, uneven development within national and state boundaries. A central component of any regional economy is the regional labor market. A region's citizens and public officials typically give high priority to the growth of real wages and employment opportunities.

In Minnesota, the pace of economic growth during the last decade varied widely from region to region. Generally, areas dependent on agriculture or mining showed slow or negative growth. In contrast, the Twin Cities metropolitan area had a relatively healthy economy, and, hence, employment growth. To further understanding of how labor markets function in Minnesota, this paper develops a framework for explaining and forecasting migration, wages, and employment in Minnesota on a regional basis.

Initially, counties in the state with similar economic characteristics were combined to form 18 development regions. In the ensuing statistical analysis, these regions were pooled to provide the additional degrees of freedom required to estimate the structural form of the proposed model. Figure 1 shows the resulting six regions for which the regional labor market was estimated.

Figure 1: Minnesota's Regional Labor Markets
Four sections follow this introduction. The second section focuses on the migration component of labor supply. Migration was chosen due to its pivotal role in the interrelationship between population change and economic conditions and because it is less predictable than either natural increase or labor force participation rates. Indeed, the importance of migration appears to be increasing due to the decline in the natural rate of increase, as Wardwell and Brown note in their book, "New Directions in Urban-Rural Migration:"

"Historically, local differences in population growth rates were primarily dependent on variations in natural increase, supplemented by net migration. For example, rural areas and central cities of large metropolitan areas have shown population stability or slow growth in spite of massive out-migration because the excess births over deaths was great enough to offset the migration loss. Today, net migration has taken over, in most areas, as the prime determinant of local population."[1]

In particular, this study intends to answer the following question concerning migration: "What are the economic and demographic determinants of in- and out-migration in Minnesota?"

The third section of the paper develops the wage and employment demand portions of the labor market. The inquiry first provides the theoretical background of how wages and employment are determined on a regional level, and then presents the equations representing this theory. In addition, because labor supply forecasts are an essential input in estimating wages and employment, the study will attempt to roughly evaluate the effect of migration on regional employment and wages. Specifically, the research aspires to appraise the claim postulated by neoclassical economic theory that migration acts as an equilibrating force between regions.

The fourth section of the paper presents the empirical results and analysis for the hypothesized equations.

The final section offers suggestions for future research.

II. MIGRATION

Historically, the U.S. population has been exhibited a high degree of mobility. Recent statistics suggest little to alter this view of American life. A report by the U.S. Census Bureau shows that 46.5 million Americans, or one-fifth of the total population, changed residence between March, 1984, and March, 1985. About 16 million of these movers consisted of people migrating to destinations outside their county. The report also showed that the Midwest continues to
experience a net outflow of residents. From 1983 to 1985 the Midwest had a net loss of 211,000 residents.

In Minnesota, the pattern of migration in the 1980s is also one of net out-migration. Between 1980 and 1985, the population loss due to migration was 66,000. Focusing only on interstate migration rates, however, one misses important aspects of migration in Minnesota because net migration rates varied from region to region within the state's boundaries. For example, the iron range counties (Itasca, Lake, and St. Louis) lost over 21,000 residents to migration between 1980 and 1985. That figure represents 8 percent of their collective population. In contrast, the seven-county Twin Cities metropolitan region gained approximately 10,000 migrants over the same period.

In an attempt to understand these population shifts, this segment of the labor market model develops a model of in- and out-migration for the analysis of migration determinants. After a short historical perspective, the migration model is presented in its functional form. This is followed by a justification of the explanatory variables chosen for analysis.

A Brief History of the Model's Origins

The importance of migration in regional development has spurred great interest in the academic community. Sociologists, economists, demographers, and geographers have produced literally hundreds of migration models. Unfortunately, the cumulative empirical results are often inconsistent and conflicting. Particularly disquieting for economists has been the lack of importance of economic variables (such as wages, unemployment, and income) in explaining the decision to migrate.

The predecessors of the model presented in this paper come from two distinct species--gravity models and neoclassical theory.

Gravity Models

Gravity models are a modification of Newton's Theory of gravitational attraction. Newton's model takes the following form:[2]

\[ F(1,2) = G \frac{(M_1 \cdot M_2)}{D(1,2)^2} \]

where:
- F(1,2) is the gravitational force between particles one and two
- G is the universal gravitational constant
- M_1, M_2 are the masses of the two particles
- D(1,2) is the distance between the two particles
Gravity models used by social scientists replace Newton's variables with demographic ones, specifically:[3]

\[ I(1,2) = G \times \frac{[(W_1P_1)a + W_2P_2b]}{D(1,2)^c} \]

where:
- \( I(1,2) \) is the level of interaction between places 1 and 2
- \( D(1,2) \) is the distance between 1 and 2
- \( P_1, P_2 \) are the populations at 1 and 2
- \( W_1, W_2 \) are weights
- \( a, b, c \) are adjustment factors
- \( G \) is some appropriate constant

In general, migration between regions should be positively related to population (due simply to the rise in potential migrants), and negatively related to distance (because the cost of moving usually rises with distance.)[4] As Walter Isard points out, the gravity model is not confined to the migration of people. The object of study may be trade, commuting, or even personal correspondences.[5] Unfortunately, the model's relationship to economic theory is not explicit, and somewhat tenuous. In addition, it says nothing about the direction of migration.

**Neoclassical Theory**

Neoclassical theory, as explained by Hicks, contends that migration is the result of people maximizing their "net economic advantage."[6] Net economic advantage usually takes the form of higher wages or better employment possibilities. The theory suggests that the migration of workers will continue until wages and unemployment are equal across locales.[7] If region "A" has higher wages than region "B", then some workers in region "B" will migrate to region "A". This drives up wages in "B" and forces them down in "A". Eventually the wage disparity disappears and workers no longer have an incentive to migrate. To explain persistent regional wage disparities, Hicks reasons that "The labour market is not a perfect market, the equalizing forces do not act quickly and easily, but nevertheless they do act." In addition, permanent wage differences may occur as the result of cost of living differences, externalities associated with particular regions (such as climate), and differing skill levels between occupations.[8]

**The Lowry Model**

These two approaches to migration were formally combined by I.S. Lowry in what has come to be known as "The Lowry Synthesis." The Lowry model is as follows:[9]
\[ M(1,2) = K[U1/U2 \times W1/W2 \times (L1 \times L2 \times M1 \times M2)/D12] \]

where:
- \( M(1,2) \) is migration from 1 to 2
- \( K \) is some constant
- \( U1, U2 \) are the unemployment rate at 1 and 2
- \( W1, W2 \) are the wages at 1 and 2
- \( L1, L2 \) are the size of the civilian labor forces at 1 and 2
- \( M1, M2 \) are the size of the military labor forces at 1 and 2
- \( D12 \) is the distance between places 1 and 2

Obviously, the gravity components are labor force size and distance. Lowry anticipated the labor force coefficients to be positive because as the labor force grows, so do the number of potential migrants. For the distance variable, a negative coefficient was anticipated due to the belief that the monetary and psychological costs of migration rise with the distance between the origin and destination. The neoclassical elements are wages and unemployment. He expected migration to be positively correlated with wages and negatively related to unemployment as workers seek to maximize the compensation for their labor.

Lowry applied this model in a study of net migration of 90 SMSAs. The regression results are given below:[10]

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>-7.4154</td>
<td>.14811</td>
</tr>
<tr>
<td>log U1</td>
<td>.0371</td>
<td>.14569</td>
</tr>
<tr>
<td>log U2</td>
<td>-.78408**</td>
<td>.38870</td>
</tr>
<tr>
<td>log W1</td>
<td>-.46275</td>
<td>.37282</td>
</tr>
<tr>
<td>log W2</td>
<td>1.04615*</td>
<td>.05442</td>
</tr>
<tr>
<td>log L1</td>
<td>.68316**</td>
<td>.05870</td>
</tr>
<tr>
<td>log L2</td>
<td>.54892**</td>
<td>.02513</td>
</tr>
<tr>
<td>log M1</td>
<td>.27208**</td>
<td>.05090</td>
</tr>
<tr>
<td>log M2</td>
<td>.43151**</td>
<td>.05543</td>
</tr>
</tbody>
</table>

* significantly different from zero at .01 level
** significantly different from zero at .001 level

The model and empirical results had three important effects on subsequent research. First, the vast majority of subsequent migration models, including the model presented in this paper, follow this general format of including both economic and demographic variables. Second, the empirical evidence suggested "that total out-migration from place \( i \) can probably be forecast without regard for destinations of migrants; and the total flow of in-migration to place \( i \) can probably be forecast without regard for origins of the migrants."[11]
This is significant because to consider all possible migration alternatives is a formidable, if not impossible, task. This indicates that it may be possible to forecast migration between a particular substate region and the rest of the country. Finally, no relationship was found between migration and origin economic conditions. As one can see from the empirical results, both origin unemployment and wages did not have a statistically significant influence on migration. This somewhat surprising result gave rise to the idea that "push" and "pull" factors are often unrelated (push factors are factors related to out-migration, while pull factors are associated with in-migration).

The model presented in this paper differs from Lowry's formulation in two important respects: (1) whereas Lowry's formulation concerned net migration, this model considers in- and out-migration in separate equations and (2) the migration equations presented here are estimated simultaneously with employment change due to their interdependency, a fact Lowry recognized but did not model.

The Ins and Outs of Migration Flows

Interregional migration models can be partitioned into two groups: (1) those estimating gross migration flows and (2) those estimating net migration. Studies of gross migration include in- and out-migration as separate flows. Studies of net migration simply address the difference between in-migration and out-migration.

As mentioned above, the model presented here will analyze in- and out-migration separately. This has several advantages. First, since net migration is equal to in-migration minus out-migration, any variable which positively affected both in- and out-migration (such as population density) may appear to have no influence on net migration. The effects could cancel each other out and hide the true significance of the variable in the migration decision process. Similarly, the influence of variables with opposite effects on in- and out-migration (such as unemployment) might be exaggerated if only net migratory flows are considered.

Second, the study of gross flows allows us to test the hypothesis that origin economic conditions do not affect the decision to migrate. This has occupied the thoughts of many economists since the Lowry study was conducted; unfortunately, subsequent results have been ambiguous.

Michael Greenwood, in a survey of migration models, showed that income and unemployment have been a much more consistent indicator of in-migration than out-migration, thus lending support to the Lowry hypothesis.
In dissent, E. Miller suggests that to properly test whether out-migration is influenced by economic conditions, the population's propensity to migrate must be taken into account. For example, regions with a relatively large percentage of students or young people will experience more out-migration regardless of economic conditions due to this age group's high migration propensity. He proposes that areas with low income and low employment growth encourage out-migration and discourage in-migration. If such a situation persists, the region will eventually be dominated by people with low migration propensities. Hence, economic variables will not appear to influence the decision to migrate. Using the variables 'percent of population with a college degree' and 'percent of population born out of state' to control for migration propensity, Miller finds origin economic conditions to be statistically significant.

Miller's findings are supported by Lansing and Mueller whose study suggests that unemployment is most prevalent among the least mobile in society. Studies by W.F. Mazek and by Alperovich, Bergmann, and Eheman lend further support to this idea. Mazek controls for migration propensity by dividing the population into white and nonwhite subgroups and finds out-migration is affected by unemployment. Alperovich et al. found that by including possible destination alternatives in the migration decision, economic conditions were equally important as push and pull factors.

However, in stark contrast, a report by P.A. Morrison and D.A. Relles did not find any hidden effects due to a low propensity to migrate. They concluded that economic factors are a pull, but not a push. In brief, the jury is still out on the economic influence on out-migration.

Lastly, knowledge of gross migration flows is an important indicator of the changing composition of a population. If 500,000 move into a region and 500,000 move out, then net migration is zero. However, the characteristics of the population may be greatly altered. Such information would undoubtedly have important policy implications.

Simultaneous Equation Models

Most migration models constructed in the 1950s and 1960s to study the determinants of migration included employment or employment change as an explanatory variable. Unfortunately, migration and employment are simultaneously determined. In-migration, in response to employment growth, will shift out the labor supply curve. The new residents will also provide additional demand for local goods and services; this will shift out the labor demand curve and further increase employment. This employment growth will attract additional migrants, and the cycle repeats itself. The extent to which the supply and demand shifts increase employment will depend on the elasticity of demand for labor. M. Polese, in a study on the impact of international
migration, estimates this "employment multiplier" to be approximately .79. [24] That is, the demand created by 10 immigrants will produce 7.9 additional jobs.

In-migration will also shift the demand curve for labor because of productivity effects. Migrants are never a random selection of the population. In Minnesota, for instance, between 1975 and 1980, 48 percent of all migrants were between the ages of 20 and 34. [25] In general, migrants tend to be younger and possess a higher level of education and occupational status than both the sending and receiving regions. [26] If a region's immigrants are young and well educated, the region's marginal productivity of labor may rise, causing the labor demand schedule to shift to the right. The resulting increased demand for labor will spur additional in-migration. In-migration may affect the region's marginal productivity in other ways. For example, it may cause an increase in investment or allow a region to achieve certain economies of agglomeration. [27] These would also shift the labor demand schedule to the right.

Migration's effect on the marginal productivity of labor in a region is particularly important because it suggests that migration may not act as an equilibrating force in the supply and demand for labor.

For the sending region, the situation is much the opposite. Even if out-migrants are unemployed, their absence will reduce aggregate demand in the region. This causes further unemployment which provides additional impetus for further out-migration. J. Vanderkamp, studying out-migration effects in Canada's maritime counties, found that "For every five unemployed persons leaving the maritimes two previously employed persons become unemployed." [28] Moreover, as with in-migration, out-migration will probably change the region's marginal productivity of labor and shift the labor demand schedule, for reasons identical to those discussed above in relation to in-migration.

In recognition of the causal relation between employment growth and migration rates, in 1971 R. Muth developed a simultaneous equation model of migration and employment change. [29] This approach represents an important step in the development of migration models. As W. Mazek observes, any model not using a simultaneous-equation approach will contain a least squares bias in the coefficient estimates. This could account for a portion of the statistically insignificant results for the economic variables of earlier models. [30]

Based on this rationale, this study estimates migration simultaneously with employment.
Functional Form and Explanation of Variables

The following two equations present the variables hypothesized to influence in- and out-migration.

\[
\text{IM} = F(\text{US}_i, \text{UN}_i, \text{ISA}_i, \text{INA}_i, \text{ED}_i, \text{OUT}_i) \quad i = 1..6
\]

\[
\text{OM} = F(\text{US}_i, \text{UN}_i, \text{ISA}_i, \text{INA}_i, \text{ED}_i, \text{IN}_i, \text{EA}_i, \text{AGE}_i) \quad i = 1..6
\]

Definition of variables:

- IN and OUT represent in and out-migration, respectively.
- US\(_i\), UN\(_i\) are regional unemployment rates relative to state and national unemployment rates.
- ISA\(_i\), INA\(_i\) are regional per capita income relative to state and national per capita income.
- ED\(_i\) is the change in regional employment demand.
- AGE\(_i\) is the percentage of the population between the ages of 17 and 44 in region \(i\).
- EA\(_i\) is the percent of the population with 16 or more years of schooling in region \(i\).

Explanation of Variables

Unemployment, income, and employment demand: Regional unemployment rates relative to state and national unemployment rates, per capita income relative to state and national per capita income, and regional employment demand must be considered together. Most earlier models included only unemployment or income based on the neoclassical premise that migration will take place until income, or unemployment, becomes equal across regions. However, the cumulative results did not provide clear and consistent support for the theory. Regression results often included wages, income, and unemployment coefficients that were either the wrong sign or not significantly different from zero, or both.

M. Todaro was among the first to combine all three elements coherently. His model was an attempt to explain the persistent rural to urban migration in Kenya despite significantly higher urban unemployment.[31] He solves the dilemma by building a behavioral model of migration in which equilibrium is achieved only when expected incomes are equal across regions.[32] Migration is, therefore, a function not only of income differentials but also the probability of finding a job. In his model, the number of unemployed and the change in employment are used as proxies for employment possibilities. The model is as follows:[33]
\[ M(r-u)/UL = F((VU(t)-VR(t))/VR(t)) \quad F' > 0 \]

\[ VR(s) = \int_s^\infty YR(t)e^{rt} \, dt \]

\[ VU(s) = \int_s^\infty p(t)YU(t)e^{rt} \, dt - C(s) \]

where:

- \( M(r-u) \) is the net rural to urban migration
- \( UL \) is the current urban labor force
- \( VU(t) \) is the discounted present value of expected urban income at time \( t \)
- \( VR(t) \) is the discounted present value of expected rural income at time \( t \)
- \( YR(t) \) is rural income in period \( t \)
- \( YU(t) \) is urban income in time \( t \)
- \( p(t) \) is the probability of finding a job, which is equal to new employment divided by the number of unemployed
- \( r \) is the discount rate
- \( C(s) \) is the cost of moving at time \( s \)

Though the model is used to explain migration patterns in Kenya, it is easily adaptable to the study of interregional migration in the United States. For example, in a study of migration determinants in the southern United States S. Bowles found:

"the present value of expected income gain from moving out of the south is positively related the probability of moving, and provides a better explanation of migration than the more conventional income measure based on regional differences in current income." [34]

Expected income has been a more consistent indicator of in-migration than out-migration. [35] One possible explanation is that the age and educational characteristics of a region are not explicitly accounted for by most formulations of expected income.

**IN and OUT:** OUT is included in the IN equation because regions with a high degree of in-migration generally have high rates of out-migration. For example, in Minnesota, between 1975 and 1980, 157,000 people moved into metropolitan areas while 127,000 moved away. [36] Large outflows correlate with large inflows because people who have moved at least once have a higher migration propensity than those who have never relocated. Therefore, regions with high rates of immigration will have a larger number of potential out-migrants than regions with less in-migration. As E. Miller reasons, recent migrants
do not have strong community ties, such as family, friends, or business connections. Hence, the monetary and psychological costs of their moving again are less than those of the general population. Recent migrants are also more likely to have labor market information concerning alternative migration destinations, further reducing the costs of migration.[37]

IN is included in the out-migration equation due to the observation that a substantial portion of migration is return migration.[38] Return migrants may have been unable to secure employment in the new region, or may simply have become disillusioned with their quality of life there. Accurate numbers of return migration are difficult to find. Most published data do not distinguish between first time and repeat migrants. In one study of return migration in Copenhagen, Denmark, S. Goldstein discovered that approximately 25 percent of all immigrants were originally from Copenhagen.[39]

AGE: Migration is clearly age-selective, as is demonstrated in Figure 2 below, with young adults having the highest migration rates.[40]

Figure 2: Rates of Migration by Age Group

As cited earlier, between 1975 and 1980, 48 percent of all Minnesota's migrants were between the ages of 20 and 34; yet this same age group comprised only 26 percent of the population. This high mobility reflects a lack of home ownership, family obligations, and frequent job changes—common socio-demographic characteristics of this age group. In addition, the net present value of expected returns from migration will be higher for younger migrants. Thus, other things equal, regions with a relatively higher percentage of their population in this age group should have higher rates of out-migration.

**Education:** Most studies have hypothesized that out-migration is positively related to the level of education. This is based on the premise that an educated person will have greater opportunities outside his/her region than the general population, and can more easily overcome the cost of migration because he/she commands a higher wage rate. If true, this would have important policy implications, as M. Greenwood notes in the following passage:

"From the point of view of regions of net out-migration this observation has relevance because it may mean that these regions suffer disproportionately heavy losses of their best educated manpower. Similarly, regions of net in-migration may experience disproportionately heavy gains of such people."[41]

In addition, the inclusion of an education variable will provide better coefficient estimates for the economic variables because it will help control for a region's migration propensity.[42] For example, low out-migration from a region with high unemployment may be due to the low educational attainment of its residents, which severely restricts their opportunities outside the region.

Unemployment, income, employment demand, in-migration, out-migration, age, and education are hypothesized to represent the important determinants of migration. Like the Lowry model, the equations contain both economic and demographic determinants. The model improves on the Lowry formulation by estimating in- and out-migration separately and by estimating migration simultaneously with employment demand.

### III. WAGES AND EMPLOYMENT

This section presents a model of regional wage and employment determination. The impact of in- and out-migration will be of particular concern. That is, does migration widen or reduce regional wage and employment disparities? The section begins by presenting briefly the orthodox, neoclassical approach to understanding the wage/employment relationship. This is followed by a discussion of some of its weaknesses. Based on this information, an empirical model of employment demand and wage change is provided along with a short explanation of the explanatory variables.
The Neoclassical Approach

The theoretical foundation for most labor market models has been neoclassical economics. Neoclassical labor theory, as a framework for analysis, can be conveniently grouped into three categories: (1) the theory of labor demand, (2) the theory of labor supply, and (3) the interaction of aggregate labor demand and supply curves in determining regional wage and employment levels.

Demand for Labor

According to neoclassical theory, the amount of labor demanded by a single firm will increase until the value of the marginal product of labor just equals its respective wage rate. If the value of marginal product is greater than the wage rate, the firm can increase its profits by hiring additional units of labor until the value of the marginal product falls to the level of wages. Any further increases will cause profits to decline. If one assumes a classical production function, then the relationship between output, profit, and demand for labor can be represented graphically in the following manner:
Figure 3: Output, Profit, and the Quantity of Labor Demanded

In perfect competition, the MC curve is perfectly elastic. That is, the firm is a price taker in the labor market. The VMP curve shows how many units of labor are desired at a given wage level.[45] Hence, the downward sloping portion of the VMP curve represents the labor demand curve. The negative slope is due to the law of diminishing marginal product.

Supply of Labor

The amount of labor supplied by an individual will be determined by his/her budget constraint and indifference curves. The relationship is shown by Figure 4 below:[46]

Figure 4: Individual Labor Supply
The budget constraint represents the maximum amount of "other goods" that can be obtained for a given level of leisure. The slope is equal to the negative value of the wage rate; that is, the wage is the rate at which leisure can be turned into "other goods." Each of the indifference curves represents the combinations of "other goods" and leisure that yield the same level of utility. The slope depends on the tastes and preferences of the laborer. The "rational" individual will consume at the point where an indifference curve is just tangent to the budget line. This is the highest level of utility he/she can obtain. The resulting utility maximizing levels of leisure and "other goods" are X* and Y*, respectively.

The individual labor supply curve is found by varying the wage rate. Whether a rise in wages increases or decreases the amount of labor an individual supplies depends on which of two effects dominate—the income effect or the substitution effect.

**Aggregate Supply and Demand Curves**

The regional demand curve for labor is found by horizontally adding all individual demand curves. Likewise, the aggregate supply curve is the horizontal sum of all individual labor supply curves. The regional wage rate is determined by the intersection of the supply and demand curves.

If one specifies a supply and demand function, the equilibrium wage can be quantitatively determined. The labor supply function is some function of the population and the labor force participation rate. Assuming that the labor force participation rate is solely determined by the wage rate, as the neoclassical world does, the regional labor supply schedule may be presented as follows:

\[ ES = a_0 + a_1 IN + a_2 OUTF + a_3 NI + a_4 W \]  

where:
- **ES** is employment supplied
- **IN** is in-migration
- **OUT** is out-migration
- **NI** is the natural increase in the labor force
- **W** is the wage rate

The amount of labor demanded is determined by the marginal productivity of labor which is equal to the wage rate. In addition, long run shifts in the demand curve are certain because of changes in technology and changes in the capital to labor ratio (which, in turn, affect the marginal productivity of labor). Thus, the employment demand function may be written as follows:

\[ ED = b_0 + b_1 W + b_2 Z \]
where:
W is the wage rate
Z is a vector of demand shifters

Because \( ED = ES \) in this neoclassical framework, the equilibrium wage is found by equating equations (1) and (2):

\[
a_0 + a_1 \text{IN} + a_2 \text{OUT} + a_3 \text{NI} + a_4 W = b_0 + b_1 W + b_2 Z
\]

\[
W = 1/(b_1+a_4) \left[ (a_0-b_0) + a_1 \text{IN} + a_2 \text{OUT} + a_3 \text{NI} - b_2 Z \right]
\]

Limitations

Most analysts believe this narrow presentation of neoclassical theory is an incomplete portrayal of a labor market. Though the neoclassical school still dominates the economics profession, many important concessions to reality have been made. The following discussion addresses four principal weaknesses of the neoclassical approach: (1) nonwage costs of labor, (2) institutional influences, (3) imperfect information, and (4) discrimination.

Nonwage Costs of Labor

Perhaps the greatest shortcoming of orthodox theory is its failure to account for the nonwage costs of labor. Each firm must recruit, screen, and train new or promoted workers. Thus, in addition to wages, the price of labor includes the costs of attracting potential workers, discerning good applicants from undesirable ones, and teaching workers to perform their tasks correctly.\(^{49}\) The presence of these costs affects the labor market adjustment process. For example, in a recession a firm may not reduce its labor force because it would lose its investment in human capital embodied in those workers. This may occur even when their wage is greater than their marginal product. Similarly, a firm may not reduce wages for fear of increasing the turnover rate, and thereby increasing the company's screening and recruiting labor costs.

Clearly, the presence of these costs changes the classical relationship between marginal product and wages. As Peter Doeringer and Michael Piore state:

"The equality between the marginal product of labor and the wage of a job postulated by competitive economic theory is reduced to an equality between the discounted present value of expected costs and productivity streams calculated over the distribution of expected employment tenure for various groups within the enterprise."\(^{50}\)
Institutional Influences

Another deficiency of orthodox theory is the lack of attention to institutional influences on wages. The United States is no longer a nation of small competitive firms. The largest 500 corporations account for nearly 90 percent of all sales in the U.S. [51] The presence of these massive wage structures has led Doeringer and Piore to suggest that there are two labor markets—the "internal" labor market and the "external" labor market. The external labor market is the market of orthodox economic theory. Here, wages and employment are determined by the forces of supply and demand. Internal labor markets are found within organizations, such as a corporation. The two are connected by what Piore and Doeringer call "ports of entry and exit."

For example, an entry level position in marketing at IBM would be considered a port of entry. Aside from these ports, laborers in the external market do not have access to internal labor market job vacancies.

In the internal labor market, wages are influenced by tradition in addition to competitive forces. Each firm has a wage structure. For each job, the wage is set in relation to other jobs. For instance, middle management is paid less than upper management but more than an entry level position. Often wages are set to encourage specific promotion channels and less turnover, thus ensuring a quality supply of labor for certain job classifications and lower nonwage labor costs. Over time these wage structures become custom and are difficult, though not impossible, to alter. Even if the job function changes, its title and the wage it commands may not be altered. If competitive forces do require the wage for a specific job to change, other wages may have to be adjusted to maintain the internal consistency of the structure. [53] The system of internal wages is consistent with profit maximization; that is, firms have these systems of wages because they are more efficient (in terms of being less costly) than individual wage determination would be. [54]

Imperfect Information

Yet another reason wage rates are unlikely to be found at the intersection of the supply and demand curve is the lack of perfect information. As George Stigler observes, each worker "...faces the problem of how to acquire information on wage rates, stability of employment, and other determinants of job choice, and how to keep this information current."

Unfortunately, like any commodity, labor market information is scarce. Therefore, its acquisition costs time and money. Since a rational individual will consume only to the point where marginal benefit equals marginal cost, no rational individual will have perfect information. Imperfect information leads to greater uncertainty and longer time lags in the adjustment process. If demand for labor is continually changing, then the supply of labor will always be in the process of adjusting. Hence, a labor market equilibrium should be considered the exception rather than the rule.
Like individuals, a firm's search for the equilibrium wage continues only until the marginal cost equals the marginal benefit. This would lead to a variety of wage offers even if there were no differences between workers. The more rapidly the equilibrium wage rate changes, the greater will be the differences in wage offers. Emile Cheysson, a 19th century French economist, was one of the first to recognize the influence of imperfect producer information on the wage adjustment process. The following figure illustrates his formulation of the adjustment process to a change in the equilibrium wage rate:

Figure 5: Emile Cheysson's Wage Adjustment Process

If the wage is set above the equilibrium (OE) at OC, then an excess supply of labor will occur; labor supply is CB, and the quantity of labor demanded is CB'. Realizing its mistake, the firm will adjust wages downward, for instance to OD. At this wage a shortage occurs. The firm continues to adjust the wage until an equilibrium is reached. The important point is that the adjustment is not necessarily a smooth, orderly procession back to equilibrium. But rather, the wage can fluctuate, somewhat unpredictably, on both sides of the equilibrium.

Discrimination

Discrimination can also distort the classical view of wage determination. Prejudice against women, minorities, and others in the labor market often surfaces as either wage discrimination or occupational discrimination.

Wage discrimination occurs when two equally productive people are paid different wages. For many occupations, productivity evaluation is
largely a subjective matter, and thus easily affected by the prejudice of the employer. Current wage levels tend to reinforce prejudices because employers often base wage decisions on historical occupational patterns, rather than the marginal productivity of the firm's labor force.[59] If an occupation is dominated by women or minorities, the entire wage structure may be underpaid.

Occupational discrimination occurs when women and minorities are denied access to certain jobs. As Piore and Doeringer state:

"Race is an inexpensive screening criterion. When two racial populations differ significantly in terms of proportion of persons possessing certain desired characteristics, the most efficient hiring policy may be simply to reject all members of one racial population."[60]

The argument of Piore and Doeringer could be applied to women, though perhaps to a lesser degree. In addition to being denied entry level positions, women and minorities can systematically be denied promotions or allowed access to only certain promotion ladders.[61]

Unfortunately, wage and occupational discrimination are extremely difficult to incorporate in a regional model. Thus they are recognized here as potentially important labor market factors, but are not accounted for in the statistical model put forward.

**A Dynamic Regional Labor Market Model**

We introduce two important modifications to overcome these limitations in orthodox theory and more accurately portray regional labor market behavior. First, we explicitly define the shifters of employment demand. Second, we abandon the assumption that the labor market is continually in a state of equilibrium.

The suggested labor demand function is specified in functional form as follows:

$$ED = F( X, OUT, IN, W)$$

where:
- **ED** is the percentage change in employment demand
- **X** is the average percentage change in demand for output over the past three periods
- **OUT** is the out-migration rate
- **IN** is the in-migration rate
- **W** is the regional average percentage change in the wage rate

The following is a discussion of the explanatory variables.
Output Demand

The demand for output affects the demand for labor in both a competitive and noncompetitive product market. If the product market is competitive, an increase in product demand will push up the price of the product. This price increase leads to an outward shift in the demand for labor since the labor demand curve is, by definition, equal to the marginal product of labor times the price of output.

Firms will not, however, alter their demand for labor if the change in output demand is perceived to be a short-term aberration. The nonwage costs of labor are simply too high. Managers are more likely to base employment decisions on longer term output trends. For this reason, the average output over the past three years was chosen instead of current year output.

In- and Out-migration

The inclusion of in- and out-migration in the labor demand function recognizes the nonhomogeneity of labor force migrants. As was mentioned in section II, migrants tend to be younger and possess higher levels of education and occupational status than the average person in either the sending or receiving regions. If this holds, immigrants will tend to increase a region’s marginal productivity, shifting the labor demand curve outward. In addition, in-migration may also cause an increase in investment or allow a region to achieve certain economies of agglomeration. These would also shift the demand schedule to the right. Out-migration will have the opposite effect.

The influence of in- and out-migration upon wages is ambiguous under this formulation. For example, in-migration will shift both the labor supply and demand schedules outward. The net effect depends on the elasticities of supply and demand and the magnitudes of the respective shifts.

Wages

The average regional wage rate is a proxy for the marginal cost of labor. Rising wage rates should induce a decline in the amount of labor demanded.

A Dynamic Model

The next step is to formulate a dynamic model of wage adjustment by assuming that labor demand does not equal labor supply. We retain the assumption that there exists a long run equilibrium which may or may not be attainable. The rate of change in wages with respect to time is hypothesized to be a function of the difference between employment demanded and employment supplied:
\[
dW/dT = j(ED-ES)
\]

where:

\(j\) is the speed of adjustment coefficient.

The value of the adjustment coefficient, \(j\), is particularly influenced by the nonwage costs of labor, institutional arrangements, and the cost of labor market information.

Substituting the labor demand and labor supply equations into the wage adjustment equation and performing some algebraic manipulations,

\[
(dW/dT) - j(a_4-b_4)W = j[(a_0-b_0)+a_1X+(a_2-b_1)IN+(a_3-b_2)OUT+b_3NI]
\]

This is a nonhomogeneous first order differential equation of the form:

\[
dy/dt + a(y) = b; \text{ with solution } y(t) = yc + yp
\]

where:

\(yc = Ae; \quad A = y(0) - b/a; \quad yp = b/a\)

Letting \(Z = (a_0-b_0)+a_1X+(a_2-b_1)IN+(a_3-b_2)OUT-b_3IN\)

\[
W(t) = [W(0) - Z/(b_4-a_4)] \exp\left[ -j(b_4-a_4)t \right] + Z/(b_4-a_4)
\]

Since \((b_4-a_4) > 0\), the first term goes to zero as \(t\) approaches infinity. Hence, the wage converges toward its equilibrium value.\[62\]

This formulation implies that, while the labor market is not a perfect market in the strict neoclassical sense, it is a stable market. The wage rate, responding to the forces of supply and demand, tends over time toward a labor market equilibrium.

Including the in- and out-migration equations developed in section I, we now have four equations in our labor market system. The equations in functional form are:

\[
\begin{align*}
\text{IN} &= f(\text{ED, UN, US, INA, ISA, EA, IN}) \\
\text{OUT} &= f(\text{ED, UN, US, INA, ISA, EA, OUT, AGE}) \\
\text{ED} &= f(\text{X, IN, OUT, W}) \\
\text{W} &= f(\text{X, IN, OUT, NI})
\end{align*}
\]

The number of equations in the model could be expanded to include other important elements of the labor market. Two worthy candidates would be a labor supply equation and/or an unemployment equation. Unfortunately, time and data restrictions prevented their inclusion. Thus, the empirical analysis which follows is restricted to these four equations.
IV. EMPIRICAL ANALYSIS

The empirical analysis consists of two sections—the ordinary least squares estimates, and the two-stage least squares estimates. The estimated equations presented here do not contain several variables included in the original conceptual formulation due to their lack of significance in preliminary regressions. These include UN (unemployment in the region relative to the nation), US (unemployment in the region relative to the state), and EA (educational attainment). In addition, ISA (income in the region relative to the state) was dropped because it was highly correlated with INA (income in the region relative to the nation).

To check the stability of the parameter estimates, 20 percent of the observations were withheld from each equation in initial regressions. Then regressions were run using the full set of data to perform Chow tests of structural stability. The test statistic is:

\[ C = \frac{[RSS(\text{full}) - RSS(\text{res})]/q}{RSS(\text{res})/DF(\text{res})} \]

where:
- RSS(full) is the residual sum of squares for the full data set regressions
- RSS(res) is the residual sum of squares for the restricted data set regressions
- q is the number of observations withheld
- DF(res) is the degrees of freedom in the restricted regression

The Chow tests for structural change were significant at the 5 percent level in four of the 24 equations; region 1 - employment demand, region 2 - out-migration, region 2 - employment demand, and region 5 - out-migration. At a 5 percent level of significance, one would expect to reject the null hypothesis of no structural change in one of the 24 equations, even if the null hypothesis was valid everywhere. Rejection of the null hypothesis in these four equations reduces the amount of confidence one has in the stability of these particular coefficient estimates, and it casts some generalized suspicion in the cases of employment demand, out-migration, and region 2.

In addition to the Chow tests, scatter plots of the residuals versus the predicted values were made to assess the assumption of homogeneous variance. No scatter plot demonstrated an obvious pattern which would indicate heteroskedasticity. Though, with so few observations (particularly in regions 4, 5, and 6), the usefulness of these scatter plots is limited.

As a brief summary, the estimated equations and the expected coefficient signs are given below:
\[ \begin{align*}
\text{IN} & = a_0 + a_1 \text{ED} + a_2 \text{INA} + a_3 \text{OUT} \\
\text{OUT} & = b_0 + b_1 \text{ED} + b_2 \text{INA} + b_3 \text{IN} + b_4 \text{AGE} \\
\text{ED} & = c_0 + c_1 X + c_2 \text{IN} + c_3 \text{OUT} + c_4 W \\
W & = d_0 + d_1 X + d_2 \text{IN} + d_3 \text{OUT} + d_4 \text{NI}
\end{align*} \]

\[ a_1, a_2, a_3, b_3, b_4, c_1, c_2, d_1 > 0 \]
\[ b_1, b_2, c_3, c_4, d_4 < 0 \]
\[ a_0, b_0, c_0, d_0, d_2, d_3 \text{ indeterminate} \]

where:

\[ \begin{align*}
\text{IN} & \text{ is the rate of in-migration} \\
\text{OUT} & \text{ is the rate of out-migration} \\
\text{ED} & \text{ is the percentage change in employment} \\
W & \text{ is the percentage change in wages} \\
X & \text{ is the percentage change in gross regional product} \\
\text{AGE} & \text{ is the percent of the population between the ages of 17 and 44} \\
\text{NI} & \text{ is the natural rate of increase in the working age population} \\
\text{INA} & \text{ is per capita regional income divided by the national level of per capita income}
\end{align*} \]

This four equation model was estimated for six regions for the period 1978-1983. The data were collected from a variety of sources. The migration data come from the Internal Revenue Service (IRS) county to county migration flow data, which are derived from income tax returns filed with the IRS. These data are the sole source of county migration data which provide annual figures of in- and out-migration as separate flows. Essentially, the data were developed by matching current year income tax returns with the previous year's return using social security numbers. A more complete discussion of the IRS migration data can be found in Appendix I.

The income, employment, wage, and gross product figures were developed from data supplied by the U.S. Department of Commerce, Bureau of Economic Analysis. A detailed description of these data will be available in a forthcoming publication form the State and Regional Research Center, University of Minnesota.

The rate of natural increase in the working age population (18-65) was estimated using 1980 census figures. The post 1980 estimates were derived by "aging" the population in each region. For example, the 1981 working age population was equal to the 1980 working age population plus the number of 1980 17 year olds minus the number of 1980 65 year olds. The pre-1980 estimates were developed by essentially applying the same procedure in reverse. The working age population in 1979 was equal to the 1980 working age population minus the number of 1980 18 year olds, plus the number of 1980 66 year olds. The number of 18 and 19 year olds were adjusted slightly to reflect the migration of these young adults. Adjustments were based on the state's population of 18 and 19 year olds.
The percent of the population between the ages of 17 and 44 is derived from "Minnesota Health Statistics," a publication of the Minnesota State Board of Health, which estimates population by county, age, and sex using a cohort survival model similar to the type found in section II of this paper.

**Ordinary Least Squares Estimates**

The following are the results of the ordinary least squares estimation. The results are presented in four tables focused on in-migration, out-migration, employment, and wages; each table contains six equations representing the six regions. Each table of results is followed by a discussion of the findings. The figures in parentheses are the standard errors of the coefficients.

**In-migration**

<table>
<thead>
<tr>
<th>Region</th>
<th>Inter.</th>
<th>ED</th>
<th>INA</th>
<th>OUT</th>
<th>R**2</th>
<th>RBAR</th>
<th>CHOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.001</td>
<td>.037</td>
<td>-.006</td>
<td>.922**</td>
<td>.92</td>
<td>.91</td>
<td>.48</td>
</tr>
<tr>
<td></td>
<td>(.005)</td>
<td>(.038)</td>
<td>(.004)</td>
<td>(.066)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-.012</td>
<td>.025</td>
<td>.006</td>
<td>1.15**</td>
<td>.91</td>
<td>.90</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>(.007)</td>
<td>(.038)</td>
<td>(.004)</td>
<td>(.102)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.028</td>
<td>.169*</td>
<td>-.028**</td>
<td>1.06**</td>
<td>.89</td>
<td>.87</td>
<td>.61</td>
</tr>
<tr>
<td></td>
<td>(.014)</td>
<td>(.169)</td>
<td>(.006)</td>
<td>(.179)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-.001</td>
<td>-.029</td>
<td>-.003</td>
<td>.948**</td>
<td>.92</td>
<td>.88</td>
<td>1.34</td>
</tr>
<tr>
<td></td>
<td>(.008)</td>
<td>(.069)</td>
<td>(.011)</td>
<td>(.197)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.006</td>
<td>.155*</td>
<td>-.023*</td>
<td>1.07**</td>
<td>.89</td>
<td>.83</td>
<td>1.16</td>
</tr>
<tr>
<td></td>
<td>(.007)</td>
<td>(.055)</td>
<td>(.009)</td>
<td>(.148)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-.014</td>
<td>.098</td>
<td>-.016</td>
<td>1.37**</td>
<td>.93</td>
<td>.89</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>(.002)</td>
<td>(.055)</td>
<td>(.007)</td>
<td>(.251)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* significant at the .05 level
** significant at the .01 level

The estimated coefficients and corresponding t-ratios of the in-migration equation strongly suggest that the best predictor of the in-migration rate is the out-migration rate. Out-migration is positive and significant in every region; the value of the coefficients ranged from a low of .92 in region 1 to a high of 1.36 in region 6—a remarkably small range. The positive coefficient on out-migration supports two claims: (1) that a substantial portion of in-migration is return migration, and (2) information plays an important role in determining the destination of migrants. Implicit in the latter claim is the
assumption that the rate of out-migration is positively related to the flow of labor market information from the region with out-migration to other regions.

The employment demand coefficient is positive in all in-migration equations but region four, though only significant in regions 3 and 5. If the true parameter value was zero, and the sign of the estimated coefficient determined randomly, the probability of observing five or six positive signs in six trials would be:

\[ P(5 \text{ or } 6 \text{ positive signs}) = \left[ \left\{ \frac{6}{6} \right\} \cdot \left\{ \frac{6}{5} \right\} \right] \cdot \left( \frac{1}{2} \right)^6 = 0.11 \]

This small probability suggests the true parameter is not zero. Obviously this is not conclusive, but it does lend support to the idea that growth in employment induces more people to move to the region.

The biggest surprise in the in-migration equation was associated with the level of per capita income in the region relative to the national level (INA). The coefficient is consistently negative and sometimes significantly so (regions 3, 5, and 6). The most reasonable explanation is that per capita regional income is not directly comparable to the national level of per capita income due to large cost of living differences. One would anticipate a cost of living variable to be negatively correlated to in-migration and positively correlated to per capita income. Hence, omitting a cost of living variable introduces a negative bias in the INA coefficient.

### Outmigration

<table>
<thead>
<tr>
<th>Region</th>
<th>Intercept</th>
<th>ED</th>
<th>INA</th>
<th>IN</th>
<th>AGE</th>
<th>R**2</th>
<th>RRAR</th>
<th>Chow</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.002</td>
<td>-.036</td>
<td>.007</td>
<td>1.0**</td>
<td>.0005</td>
<td>.92</td>
<td>.90</td>
<td>.39</td>
</tr>
<tr>
<td></td>
<td>(.005)</td>
<td>(.04)</td>
<td>(.004)</td>
<td>(.075)</td>
<td>(.003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.022</td>
<td>.035</td>
<td>-.005</td>
<td>.752**</td>
<td>-.01</td>
<td>.91</td>
<td>.89</td>
<td>3.32*</td>
</tr>
<tr>
<td></td>
<td>(.01)</td>
<td>(.055)</td>
<td>(.003)</td>
<td>(.075)</td>
<td>(.016)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-.006</td>
<td>-.082</td>
<td>-.008</td>
<td>.524**</td>
<td>.098*</td>
<td>.89</td>
<td>.87</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>(.011)</td>
<td>(.06)</td>
<td>(.011)</td>
<td>(.104)</td>
<td>(.036)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.004</td>
<td>.022</td>
<td>.004</td>
<td>.86**</td>
<td>.008</td>
<td>.93</td>
<td>.88</td>
<td>2.33</td>
</tr>
<tr>
<td></td>
<td>(.008)</td>
<td>(.07)</td>
<td>(.014)</td>
<td>(.192)</td>
<td>(.017)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.012</td>
<td>-.011</td>
<td>.052*</td>
<td>1.0**</td>
<td>-.114</td>
<td>.95</td>
<td>.92</td>
<td>29.12*</td>
</tr>
<tr>
<td></td>
<td>(.008)</td>
<td>(.074)</td>
<td>(.015)</td>
<td>(.123)</td>
<td>(.055)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.009</td>
<td>-.042</td>
<td>-.008</td>
<td>.573**</td>
<td>.06</td>
<td>.88</td>
<td>.77</td>
<td>.44</td>
</tr>
<tr>
<td></td>
<td>(.013)</td>
<td>(.045)</td>
<td>(.02)</td>
<td>(.12)</td>
<td>(.066)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* significant at the .05 level
** significant at the .01 level
In the out-migration equations, the in-migration variable was most consistent with prior expectations. It was positive and significant in every region. Such results indicate that regions with high rates of in-migration have a larger proportion of potential out-migrants. That is, people who have migrated into the region are more likely to move again because, in general, the monetary and psychological costs of migrating again are lower.

Employment demand is not significant in any of the out-migration equations. This result tends to support Lowry's conclusion that employment demand is a pull, but not a push factor. If a lack of employment opportunities in a region affects primarily those who have the least ability to relocate, slow employment growth may not influence the rate of out-migration. If true, this indicates policymakers cannot look to accelerated out-migration as a cure for regional unemployment differences. Rather, differentials in employment demand affect the rate of in-migration which may gradually lessen regional unemployment differences. Consider, for example, St. Louis county in northern Minnesota from 1980 to 1984. Beginning with 1980, the percentage change in employment was -.019, -.025, -.09, and -.045, respectively. Despite the declining level of employment, the rate of out-migration remained almost constant. Starting with 1980, the rate of out-migration was .043, .046, .044, and .045, respectively. The in-migration rate, however, declined markedly over this period. Beginning with 1980, the in-migration rate was .042, .032, .024, and .026, respectively.

The income variable is only significant in region 5, and it is the unexpected sign (positive). This may be due to the same reason cited above in relation to in-migration. Another plausible explanation is that out-migration will decline as per capita income falls because potential migrants are unable to overcome the initial monetary costs of migration.

The age variable is generally positive and is significant as well as positive in region 3 (the seven county metro region). This may reflect the area's growing college age population.
### Employment Demand

<table>
<thead>
<tr>
<th>Region</th>
<th>Inter.</th>
<th>X</th>
<th>IN</th>
<th>OUT</th>
<th>OW</th>
<th>R**2</th>
<th>RBAR</th>
<th>CHOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.0001</td>
<td>.83*</td>
<td>.41</td>
<td>-.34</td>
<td>-.28*</td>
<td>.48</td>
<td>.34</td>
<td>3.88*</td>
</tr>
<tr>
<td></td>
<td>(.035)</td>
<td>(.264)</td>
<td>(2.0)</td>
<td>(1.87)</td>
<td>(.136)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>- .033</td>
<td>1.03**</td>
<td>-.066</td>
<td>.232</td>
<td>-.112</td>
<td>.87</td>
<td>.83</td>
<td>4.23*</td>
</tr>
<tr>
<td></td>
<td>(.035)</td>
<td>(.222)</td>
<td>(1.27)</td>
<td>(1.45)</td>
<td>(.135)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-.102</td>
<td>.853**</td>
<td>-.302</td>
<td>1.55*</td>
<td>.013</td>
<td>.90</td>
<td>.87</td>
<td>1.19</td>
</tr>
<tr>
<td></td>
<td>(.025)</td>
<td>(.198)</td>
<td>(.535)</td>
<td>(.712)</td>
<td>(.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-.023</td>
<td>.697</td>
<td>-3.23</td>
<td>3.56</td>
<td>-.495</td>
<td>.79</td>
<td>.63</td>
<td>5.21</td>
</tr>
<tr>
<td></td>
<td>(.108)</td>
<td>(.925)</td>
<td>(3.4)</td>
<td>(2.54)</td>
<td>(.33)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-.039</td>
<td>-.286</td>
<td>3.5</td>
<td>-2.32</td>
<td>.061</td>
<td>.29</td>
<td>-.27</td>
<td>.75</td>
</tr>
<tr>
<td></td>
<td>(.06)</td>
<td>(.401)</td>
<td>(2.9)</td>
<td>(2.6)</td>
<td>(.24)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-.053</td>
<td>1.57*</td>
<td>.324</td>
<td>.796</td>
<td>.138</td>
<td>.88</td>
<td>.78</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>(.122)</td>
<td>(.628)</td>
<td>(1.94)</td>
<td>(3.26)</td>
<td>(.074)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* significant at the .05 level  
** significant at the .01 level

In the employment demand equation, the output variable coefficient is positive and significant in four regions. Only in region 5 is the coefficient negative. An increase in output may cause a decline in the demand for labor if the increased output is due to a substitution of capital for labor. Such a scenario is not unrealistic for a heavy agricultural region such as region 5.

The in- and out-migration variables are not significant in any of the regions. This suggests that migration is not influencing regional marginal productivity of labor which, in turn, suggests that migrants are neither more nor less productive than the nonmigrant population.

The wage variable is not significant in any region. In addition, it has the wrong sign in regions 3, 5, and 6. A positive wage coefficient may be due to using the average regional wage. In some regions the average wage might rise when employment declines if those who lose their jobs tend to have earned below average wages.
### Wages

<table>
<thead>
<tr>
<th>Region</th>
<th>Inter.</th>
<th>X</th>
<th>IN</th>
<th>OUT</th>
<th>NT</th>
<th>R**2</th>
<th>RPR</th>
<th>CHOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.08</td>
<td>1.52*</td>
<td>3.72</td>
<td>-3.86</td>
<td>-3.37</td>
<td>.44</td>
<td>.30</td>
<td>.43</td>
</tr>
<tr>
<td></td>
<td>(.064)</td>
<td>(.578)</td>
<td>(3.64)</td>
<td>(3.47)</td>
<td>(2.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-.06</td>
<td>-.786</td>
<td>.04</td>
<td>.86</td>
<td>1.93</td>
<td>.12</td>
<td>-.12</td>
<td>1.77</td>
</tr>
<tr>
<td></td>
<td>(.077)</td>
<td>(.623)</td>
<td>(2.44)</td>
<td>(2.84)</td>
<td>(2.07)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.05</td>
<td>.290</td>
<td>-6.96</td>
<td>1.04</td>
<td>-2.65</td>
<td>.22</td>
<td>.01</td>
<td>3.15</td>
</tr>
<tr>
<td></td>
<td>(.09)</td>
<td>(.605)</td>
<td>(1.82)</td>
<td>(2.13)</td>
<td>(3.52)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.154</td>
<td>1.2</td>
<td>-3.26</td>
<td>.94</td>
<td>-1.18</td>
<td>.34</td>
<td>-.19</td>
<td>.88</td>
</tr>
<tr>
<td></td>
<td>(.128)</td>
<td>(1.1)</td>
<td>(4.85)</td>
<td>(4.76)</td>
<td>(2.45)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.044</td>
<td>.117</td>
<td>-6.04</td>
<td>4.09</td>
<td>2.22</td>
<td>.30</td>
<td>-.26</td>
<td>5.81</td>
</tr>
<tr>
<td></td>
<td>(.125)</td>
<td>(.741)</td>
<td>(4.54)</td>
<td>(4.64)</td>
<td>(3.82)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-0.085</td>
<td>-4.12</td>
<td>.485</td>
<td>.898</td>
<td>1.83</td>
<td>.53</td>
<td>.16</td>
<td>.17</td>
</tr>
<tr>
<td></td>
<td>(.735)</td>
<td>(3.63)</td>
<td>(12.4)</td>
<td>(20.0)</td>
<td>(18.18)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* significant at the .05 level
** significant at the .01 level

Of the four sets of equations, the wage equations explained the lowest proportion of observed variation and demonstrated the least consistency with prior expectations. The output variable coefficients were generally positive, but significant only in region 1. The natural increase in the working age population, in-migration, and out-migration were not significant in any of the regions. The joint significance of the independent variables was tested using the following F-test:

\[
F = \frac{ESS(N-K)}{RSS(k-1)}
\]

where:
- ESS is the explained sum of squares
- RSS is the residual sum of squares
- N is the number of observations
- K is the number of regressors including the intercept

The null hypothesis is that all the independent variables have coefficients equal to zero simultaneously. The regressions in regions 2, 3, 4, and 5 fail to be significant at even the .10 percent level.

These results are a plausible description of the labor market. As noted in section II, wage adjustment is but one of a myriad of mechanisms used by firms to adjust labor costs in response to excess supply or demand of labor. Hence, if wages do not respond in a
predictable fashion to excess labor supply or demand, current mean wages may actually be the best indicator of future wages.

Two-stage least squares is one of a group of techniques known as "limited information methods" which are used to estimate structural parameters of a simultaneous system of equations. They are known as limited information estimators because a single structural equation in the model can be estimated without specific knowledge of the other structural equations. All one needs to know are the exogenous variables in the entire system. Two-stage least squares (2SLS) has gained popularity because it is conceptually simple and computationally easy.

The second stage regressions results are presented below. This is followed by a short interpretation of the results. The first stage (reduced form) estimates are presented in Appendix II.

### In-migration

<table>
<thead>
<tr>
<th>Region</th>
<th>Inter.</th>
<th>ED</th>
<th>INA</th>
<th>OUT</th>
<th>R**2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.037</td>
<td>-.222</td>
<td>.005</td>
<td>1.39</td>
<td>.46</td>
</tr>
<tr>
<td></td>
<td>(.062)</td>
<td>(.4)</td>
<td>(.02)</td>
<td>(.81)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-.027</td>
<td>-.18</td>
<td>.006</td>
<td>1.36**</td>
<td>.89</td>
</tr>
<tr>
<td></td>
<td>(.012)</td>
<td>(.054)</td>
<td>(.005)</td>
<td>(.16)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.026</td>
<td>.163</td>
<td>-.03**</td>
<td>1.13**</td>
<td>.89</td>
</tr>
<tr>
<td></td>
<td>(.024)</td>
<td>(.119)</td>
<td>(.007)</td>
<td>(.304)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-.018</td>
<td>-.294</td>
<td>.044</td>
<td>.49</td>
<td>.66</td>
</tr>
<tr>
<td></td>
<td>(.036)</td>
<td>(.547)</td>
<td>(.096)</td>
<td>(1.1)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-.001</td>
<td>.23*</td>
<td>-.042*</td>
<td>1.45**</td>
<td>.76</td>
</tr>
<tr>
<td></td>
<td>(.013)</td>
<td>(.096)</td>
<td>(.019)</td>
<td>(.375)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.159</td>
<td>-.394</td>
<td>.012</td>
<td>3.46</td>
<td>-.16</td>
</tr>
<tr>
<td></td>
<td>(.336)</td>
<td>(1.15)</td>
<td>(.073)</td>
<td>(4.81)</td>
<td></td>
</tr>
</tbody>
</table>

* significant at the .05 level
** significant at the .01 level
### Outmigration

<table>
<thead>
<tr>
<th>Region</th>
<th>Inter.</th>
<th>ED</th>
<th>INA</th>
<th>IN</th>
<th>AGE</th>
<th>R**2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.035</td>
<td>.186</td>
<td>-.005</td>
<td>.658</td>
<td>-.001</td>
<td>.67</td>
</tr>
<tr>
<td></td>
<td>(.038)</td>
<td>(.258)</td>
<td>(.015)</td>
<td>(.53)</td>
<td>(.006)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.362</td>
<td>.112</td>
<td>-.005</td>
<td>.644**</td>
<td>-.026</td>
<td>.89</td>
</tr>
<tr>
<td></td>
<td>(.021)</td>
<td>(.124)</td>
<td>(.004)</td>
<td>(.132)</td>
<td>(.032)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-.011</td>
<td>-.13</td>
<td>-.024</td>
<td>.387*</td>
<td>.165*</td>
<td>.86</td>
</tr>
<tr>
<td></td>
<td>(.021)</td>
<td>(119)</td>
<td>(.014)</td>
<td>(.159)</td>
<td>(.057)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.156</td>
<td>2.48</td>
<td>-.455</td>
<td>5.6</td>
<td>.126</td>
<td>-20.39</td>
</tr>
<tr>
<td></td>
<td>(1.97)</td>
<td>(31.9)</td>
<td>(5.92)</td>
<td>(60.8)</td>
<td>(1.54)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-.021</td>
<td>-.387</td>
<td>-.005</td>
<td>.446</td>
<td>.152</td>
<td>.68</td>
</tr>
<tr>
<td></td>
<td>(.164)</td>
<td>(1.7)</td>
<td>(.25)</td>
<td>(1.8)</td>
<td>(1.12)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.046</td>
<td>.109</td>
<td>-.011</td>
<td>.275</td>
<td>.25</td>
<td>.59</td>
</tr>
<tr>
<td></td>
<td>(.046)</td>
<td>(2)</td>
<td>(.037)</td>
<td>(.415)</td>
<td>(.129)</td>
<td></td>
</tr>
</tbody>
</table>

* significant at the .05 level
** significant at the .01 level

### Employment Demand

<table>
<thead>
<tr>
<th>Region</th>
<th>Inter.</th>
<th>X</th>
<th>IN</th>
<th>OUT</th>
<th>OW</th>
<th>R**2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.031</td>
<td>.91</td>
<td>-10.6</td>
<td>9.6</td>
<td>-.21</td>
<td>-.69</td>
</tr>
<tr>
<td></td>
<td>(.36)</td>
<td>(2.03)</td>
<td>(14.5)</td>
<td>(10.1)</td>
<td>(.76)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.168</td>
<td>1.51</td>
<td>2.98</td>
<td>-6.06</td>
<td>-.386</td>
<td>.48</td>
</tr>
<tr>
<td></td>
<td>(.172)</td>
<td>(.593)</td>
<td>(5.55)</td>
<td>(7.23)</td>
<td>(.61)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-.29</td>
<td>.274</td>
<td>1.39</td>
<td>2.41</td>
<td>1.4</td>
<td>-.73</td>
</tr>
<tr>
<td></td>
<td>(.516)</td>
<td>(1.72)</td>
<td>(7.04)</td>
<td>(3.48)</td>
<td>(4.16)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.233</td>
<td>2.82</td>
<td>-13.7</td>
<td>8.22</td>
<td>-.82</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td>(.3)</td>
<td>(2.64)</td>
<td>(1.13)</td>
<td>(12.4)</td>
<td>(7.32)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-.06</td>
<td>-.331</td>
<td>10.0</td>
<td>-7.79</td>
<td>1.12</td>
<td>-2.49</td>
</tr>
<tr>
<td></td>
<td>(.164)</td>
<td>(.896)</td>
<td>(10.5)</td>
<td>(9.8)</td>
<td>(1.43)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.021</td>
<td>4.14</td>
<td>1.75</td>
<td>-.685</td>
<td>.87</td>
<td>-1.57</td>
</tr>
<tr>
<td></td>
<td>(3.78)</td>
<td>(34.1)</td>
<td>(48.7)</td>
<td>(97.4)</td>
<td>(7.73)</td>
<td></td>
</tr>
</tbody>
</table>

* significant at the .05 level
** significant at the .01 level
Without question, the two-stage least squares results are something of a disappointment. In many equations, the parameter estimates and their corresponding standard errors are far larger than is reasonable intuitive. The problem derives from the first stage estimates. In many of the first stage regressions, the adjusted R2 is close to zero, or negative (see Appendix II). In such a situation, the predicted values used in the second stage regressions on the right hand side may not accurately reflect the observed values. Therefore, the estimated coefficient in the second stage regression will not be estimating the parameter of interest.

The large standard errors are the result of the seemingly unreasonable parameter estimates which are a function of the imprecise first stage estimates. In two-stage least squares, the residual estimates are not derived from the second stage regressions. Instead, the residuals are calculated as follows. The predicted values of the right hand side endogenous variables (used to find the second stage parameter estimates) are replaced by the observed values. These, used in conjunction with the second stage coefficient estimates, generate the predicted values of the left hand side endogenous variables. The predicted values minus the observed values are equal to the estimated residuals.
Because the coefficient estimates are so unreasonable, the predicted values will be also leading, in turn, to the large residual sum of squares and the large standard errors.

When reasonable results were obtained, each coefficient estimate was within one standard deviation of the coefficient estimated by the ordinary least squares method. This leads us to believe that the size of the simultaneous equation bias is small, especially relative to the imprecision introduced by other factors such as measurement errors, omitted variables, and small sample size.

V. SUGGESTIONS FOR FUTURE RESEARCH

The empirical results provide only modest support for the set of hypotheses put forth in the conceptual framework presented in sections II and III. Three categories of additional work would contribute to a more accurate depiction of regional labor markets. First, the econometric methodology should be altered. Second, the quality of the data must be enhanced. Third, the conceptual framework requires modification.

Econometric Methods

The first econometric modification concerns the interpretation of observed employment when estimating the employment demand equation. Most labor market models assume that observed employment lies at the intersection of the demand and supply curves. The model proposed here assumes that disequilibrium in the labor market is not only possible, but probable. Lamentably, the discipline of economics has no empirical method for determining whether a region is experiencing a labor surplus or a labor shortage. Thus, one cannot be certain if the observed employment lies on the labor demand curve or the labor supply curve. The question is a sort of modified identification problem. To see this more clearly, consider Figure 6 below.

![Figure 6: The Problem of Identification](image-url)
EO represents employment observed, WO represents observed wages. The problem is, given WO and EO, determining whether the observation lies on the demand curve (point A), or the supply curve (point B). This study skirted the problem by assuming all employment observations lied along the demand curve. Considering the period was characterized by relatively high levels of unemployment, this was not an unrealistic assumption. However, at present and in the future the circumstances may be less clear. Clearly, not knowing which curve an observation lies on makes estimation difficult, if not impossible.

The second econometric alteration, already alluded to at the end of section III, concerns additional equations which, if included, would draw a more complete picture of the labor market. Certainly a labor supply and unemployment equation would add two elements of great importance to policymakers. If appropriate data series can be found, their inclusion in any subsequent model would be desirable.

The final econometric consideration pertains to the further pooling of regions for the purpose of creating additional degrees of freedom, and thereby promoting more precise coefficient estimates. Additional pooling could be justified if one believes that the true structural parameters are identical for two or more regions. Originally, regions with similar primary economic activities were pooled. This method could serve as the basis for further pooling. One might also combine those regions with similar ordinary least squares (OLS) parameter estimates. A cursory investigation of the OLS estimates does not, however, suggest any obvious candidates. Finally, Bayesian techniques may be useful because we might hypothesize similar, but not identical, values for coefficients.[63]

Quality of Data

The ability to test the hypotheses embedded in an econometric theory depends upon having high quality data in sufficient quantity. Unfortunately, small area migration data are neither precise nor plentiful. Appendix I describes the migration data used in this study. The data account for the movements of taxpayers and their dependents. This group includes about 90 percent of the population. If the 10 percent which do not pay taxes behave in the same manner as those who do, there is no problem. However, until migration data are complete, the estimated equations describe only the migration of taxpayers.

The length of the migration series can only be improved with time. One can only hope that the IRS continues to publish the data. It is the only source of information which provides annual figures of in- and out-migration as separate flows. Other sources report only net migration based on the "residual method." A longer time series would provide more stable and efficient estimates, that is, estimates with smaller standard errors.
The income data are also in need of repair because they are not an accurate proxy for the concept of expected return to working in a region. Section II presented the hypothesis that citizens made migration decisions based on the net expected present value of migrating, which was thought to be primarily determined by relative income levels and employment opportunities. However, a $10,000 Twin Cities income does not buy the same standard of living as a $10,000 income in southwest Minnesota. In order to test this present value theory of migration, regional income levels must be adjusted to account for cost of living differences, particularly differences in the costs of housing and transportation. The data necessary for these adjustments do not exist.

Conceptual Framework

Two major conceptual issues deserve consideration. First, a fundamental question needs to be answered. This is, how do firms respond to perceived labor imbalances? Specifically, are wage changes the primary tool for adjusting labor costs and for attracting and retaining a quality labor force? Piore and Doeringer, in their book on manpower analysis, cite at least 12 adjustment mechanisms:

1. Wage compensation
2. Nonwage compensation
3. Job vacancies
4. Job structure
5. Managerial structure
6. Subcontracting
7. Overtime
8. Hiring standards
9. Recruitment procedures
10. Screening procedures
11. Training
12. Internal allocative rules

In developing a theoretical framework to explain wage changes, knowledge of when and why each instrument is implemented is of critical importance. For example, consider the effect the choice of adjustment mechanism might have in determining average regional wages during a recession. If firms respond to a lack of output demand by laying off those workers with lower than average wages rather than by reducing wage rates, the average regional wage will rise. Hence, a region may quite possibly experience the odd combination of declining employment demand and rising average wages.

Estimation of wages, employment, and possibly migration on a sector level might produce parameter estimates more consistent with prior expectations. However, the mechanisms of adjustment probably do not differ greatly by sector. Rather, the adjustment process is more
likely to differ according to occupation. Because each sector is composed of several occupational classes, disaggregating the region by sector will probably not produce better results. If reliable data could be developed, dividing employment in each region into skilled and unskilled occupations would be useful, since unskilled occupations almost certainly have a different adjustment process than skilled occupations.
APPENDIX I
MIGRATION DATA

In- and out-migration data by region were compiled from income tax returns filed with the Internal Revenue Service. A tax return in year \( t \) was compared to the return filed in \( t-1 \). If the county of residence matched, the filer was labeled a non-migrant. If it did not match, the filer was recorded as an in-migrant with respect to their current residence, and an out-migrant in relation to the previous year's place of residence. Also included were all exemptions listed on the individual's tax form, except those for age and blindness. Hence, those people not considered in this data base were those who did not file a return and were not claimed as an exemption.

The data encompass a time span of six years, 1978 through 1983. Unfortunately, the 1978-1979 data are reported as the sum of those two years, while all other data are reported on an annual basis. This presents two problems with the use of ordinary and two-stage least squares estimation. The first obstacle is heteroscedasticity, and the second is an inaccurate intercept term. Both problems are demonstrated by the following set of equations:

\[
\begin{align*}
(1) & \quad \text{IM}(78) = B_0 + BX78 + e78 \\
(2) & \quad \text{IM}(79) = B_0 + BX79 + e79 \\
(3) & \quad Y = \text{IM}(78) + \text{IM}(79) = 2B_0 + B[X78 + X79] + (e78 + e79)
\end{align*}
\]

where:
- \( \text{IM}(78) \) and \( \text{IM}(79) \) are \( N \times 1 \) vectors of migration observations for 1978 and 1979, respectively
- \( B_0 \) is the intercept term
- \( B \) is a \( p \times 1 \) vector of parameters, where \( p \) is equal to the number of independent variables
- \( X78 \) and \( X79 \) are \( N \times p \) matrices of observed values of the independent variables for 1978 and 1979, respectively
- \( e78 \) and \( e79 \) are \( N \times 1 \) stochastic error terms for 1978 and 1979, respectively
- \( Y \) is an \( N \times 1 \) vector of migration observations reported in the Area Migration Flow Data, and is equal to the sum of migration in 1978 and 1979

Equation (3) depicts the structural form consistent with the data. Together \( e78 \) and \( e79 \) represent its stochastic error term. The error term represents the difference between the actual and predicted value of the dependent variable, and is comprised of two parts--fixed and random. The fixed element reflects the failure of the proposed equation to model exactly the relationship between the dependent and independent variables, often called the "lack of fit" component.[64]
For a well specified model, this part should be negligible. The random error segment reflects the random distribution of the response for any given level of the explanatory variables and also any measurement errors that occurred during data collection. By assumption, the E(ei) = 0, and the VAR(ei) = σ². If the VAR(ei) ≠ σ², B will still be unbiased but ordinary least squares will no longer be a minimum variance estimator, and the corresponding "t" and "F" statistics will not be as powerful. In equation three, the error term has the following properties:

\[
E(e_{78} + e_{79}) = E(e_{78}) + E(e_{79}) = 0 + 0 = 0 \\
\text{VAR}(e_{78} + e_{79}) = \text{VAR}(e_{78}) + \text{VAR}(e_{79}) = \sigma^2 + \sigma^2 = 2\sigma^2
\]

Obviously, \( \text{VAR}(e_{78} + e_{79}) = 2\sigma^2 \) violates the homoskedasticity assumption when the combined 1978 and 1979 observation is pooled with the observations for 1980 to 1983 which have a variance of \( \sigma^2 \). The dilemma can be resolved using a generalized least squares approach. By definition, \( \text{VAR}(Be) = \text{BVAR}(e)B \). Hence, the problem requires specifying B such that:

\[
\text{VAR}(B(e_{78} + e_{79})) = \text{BVAR}(e_{78} + e_{79})B = BB*2\sigma^2 - \sigma^2
\]

It is not difficult to perceive that the correct B in this problem is \( 1/\sqrt{2} \). The resulting equation is as follows:

\[
Y/\sqrt{2} = 2Bo/\sqrt{2} + B(X_{78} + X_{79})/\sqrt{2} + (e_{78} + e_{79})/\sqrt{2}
\]

Curing the problem of nonconstant variance still leaves a faulty intercept value. If the intercept is unbiased, its expected value should be Bo. Unfortunately, after the variance stabilizing transformation is complete, the intercept is equal to \( \sqrt{2}*Bo \). To evaluate the method used to correct this situation, consider in matrix notation the standard multiple regression equation:

\[
Y = XB + e
\]

where X is defined to be an n x (p + 1) matrix which appears:

\[
X = \begin{bmatrix}
1 & x_{11} & x_{12} & \cdots & x_{1p} \\
1 & x_{21} & x_{22} & \cdots & \\
1 & \cdots & \cdots & \cdots & \\
1 & \cdots & \cdots & \cdots & \\
1 & x_{n1} & x_{n2} & \cdots & x_{np}
\end{bmatrix}
\]

The left hand column of ones corresponds to the intercept. The second column of observations corresponds to the first predictor and parameter B1, the third column of observations is associated with the second predictor and parameter B2, and so on. Correcting the bias in the intercept value was accomplished by replacing the ones in column one which correspond to the 1978-79 observation with \( 1/\sqrt{2} \) (note that \( 1/\sqrt{2} \neq 2/\sqrt{2}Bo = Bo \).
APPENDIX II
FIRST STAGE ESTIMATES

The following tables contain the results of the first stage estimates from the two-stage least squares procedure. As noted in section IV, many equations explain only a small proportion of the variation in the dependent variable. This is particularly true of the wage equations. For this reason, the resulting predicted values used in the second stage regressions on the right hand side do not accurately reflect the observed values.

<table>
<thead>
<tr>
<th>Region</th>
<th>Inter.</th>
<th>INA</th>
<th>AGE</th>
<th>X</th>
<th>NI</th>
<th>R**2</th>
<th>RBAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.047**</td>
<td>-.001</td>
<td>-.003</td>
<td>.115</td>
<td>.6</td>
<td>.20</td>
<td>-.01</td>
</tr>
<tr>
<td></td>
<td>(.014)</td>
<td>(.014)</td>
<td>(.009)</td>
<td>(.129)</td>
<td>(.76)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.089**</td>
<td>.006</td>
<td>-.138**</td>
<td>.493**</td>
<td>.866</td>
<td>.68</td>
<td>.59</td>
</tr>
<tr>
<td></td>
<td>(.011)</td>
<td>(.008)</td>
<td>(.043)</td>
<td>(.11)</td>
<td>(.576)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.054**</td>
<td>.031</td>
<td>-.153</td>
<td>.113</td>
<td>2.25**</td>
<td>.86</td>
<td>.83</td>
</tr>
<tr>
<td></td>
<td>(.01)</td>
<td>(.027)</td>
<td>(.1)</td>
<td>(.98)</td>
<td>(.468)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.058**</td>
<td>-.033</td>
<td>.01</td>
<td>.497**</td>
<td>.726</td>
<td>.89</td>
<td>.80</td>
</tr>
<tr>
<td></td>
<td>(.028)</td>
<td>(.024)</td>
<td>(.141)</td>
<td>(.448)</td>
<td>(.013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.055*</td>
<td>-.11</td>
<td>.1</td>
<td>.03</td>
<td>2.86*</td>
<td>.55</td>
<td>.19</td>
</tr>
<tr>
<td></td>
<td>(.022)</td>
<td>(.06)</td>
<td>(.077)</td>
<td>(.1)</td>
<td>(1.43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.015</td>
<td>-.074*</td>
<td>.162*</td>
<td>.185*</td>
<td>1.39**</td>
<td>.97</td>
<td>.95</td>
</tr>
<tr>
<td></td>
<td>(.009)</td>
<td>(.019)</td>
<td>(.061)</td>
<td>(.059)</td>
<td>(.314)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* significant at the .05 level
** significant at the .01 level
### Out-migration

<table>
<thead>
<tr>
<th>Region</th>
<th>Inter.</th>
<th>INA</th>
<th>AGE</th>
<th>X</th>
<th>NI</th>
<th>R**2</th>
<th>RBAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.048**</td>
<td>.0004</td>
<td>-.001</td>
<td>.056</td>
<td>.915</td>
<td>.17</td>
<td>-.05</td>
</tr>
<tr>
<td></td>
<td>(.011)</td>
<td>(.014)</td>
<td>(.009)</td>
<td>(.135)</td>
<td>(.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.084**</td>
<td>.002</td>
<td>-.095*</td>
<td>.399**</td>
<td>.419</td>
<td>.51</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td>(.01)</td>
<td>(.008)</td>
<td>(.041)</td>
<td>(.11)</td>
<td>(.558)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.023*</td>
<td>-.016</td>
<td>.097</td>
<td>-.36</td>
<td>.733</td>
<td>.76</td>
<td>.70</td>
</tr>
<tr>
<td></td>
<td>(.009)</td>
<td>(.023)</td>
<td>(.086)</td>
<td>(.084)</td>
<td>(.402)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.059**</td>
<td>-.04</td>
<td>.026</td>
<td>.457*</td>
<td>1.08*</td>
<td>.89</td>
<td>.81</td>
</tr>
<tr>
<td></td>
<td>(.013)</td>
<td>(.028)</td>
<td>(.023)</td>
<td>(.14)</td>
<td>(.443)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.061*</td>
<td>-.034</td>
<td>-.022</td>
<td>.044</td>
<td>2.28</td>
<td>.38</td>
<td>-.11</td>
</tr>
<tr>
<td></td>
<td>(.026)</td>
<td>(.072)</td>
<td>(.093)</td>
<td>(.12)</td>
<td>(1.73)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.025*</td>
<td>-.053*</td>
<td>.159*</td>
<td>.15</td>
<td>.372</td>
<td>.83</td>
<td>.69</td>
</tr>
<tr>
<td></td>
<td>(.011)</td>
<td>(.024)</td>
<td>(.078)</td>
<td>(.076)</td>
<td>(.4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* significant at the .05 level  
** significant at the .01 level

### Employment Demand:

<table>
<thead>
<tr>
<th>Region</th>
<th>Inter.</th>
<th>INA</th>
<th>AGE</th>
<th>X</th>
<th>NI</th>
<th>R**2</th>
<th>RBAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.075**</td>
<td>.031</td>
<td>.011</td>
<td>-.11</td>
<td>2.8*</td>
<td>.72</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>(.018)</td>
<td>(.024)</td>
<td>(.016)</td>
<td>(.224)</td>
<td>(1.32)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-.083**</td>
<td>.028</td>
<td>.181*</td>
<td>.72**</td>
<td>-1.24</td>
<td>.92</td>
<td>.90</td>
</tr>
<tr>
<td></td>
<td>(.021)</td>
<td>(.016)</td>
<td>(.083)</td>
<td>(.217)</td>
<td>(1.12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-.1**</td>
<td>.028</td>
<td>.068</td>
<td>.608*</td>
<td>1.05</td>
<td>.93</td>
<td>.91</td>
</tr>
<tr>
<td></td>
<td>(024)</td>
<td>(.061)</td>
<td>(.224)</td>
<td>(.219)</td>
<td>(1.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-.171*</td>
<td>.26*</td>
<td>-.064</td>
<td>-.936</td>
<td>-1.2</td>
<td>.89</td>
<td>.80</td>
</tr>
<tr>
<td></td>
<td>(.054)</td>
<td>(.116)</td>
<td>(.098)</td>
<td>(.586)</td>
<td>(1.85)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-.148*</td>
<td>-.049</td>
<td>.564*</td>
<td>-.08</td>
<td>-.26</td>
<td>.82</td>
<td>.68</td>
</tr>
<tr>
<td></td>
<td>(.046)</td>
<td>(.127)</td>
<td>(.164)</td>
<td>(.211)</td>
<td>(3.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-.187*</td>
<td>-.201</td>
<td>.821</td>
<td>.896</td>
<td>-.09</td>
<td>.89</td>
<td>.81</td>
</tr>
<tr>
<td></td>
<td>(.081)</td>
<td>(.169)</td>
<td>(.556)</td>
<td>(.538)</td>
<td>(2.86)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* significant at the .05 level  
** significant at the .01 level
## Wages

<table>
<thead>
<tr>
<th>Region</th>
<th>Inter.</th>
<th>INA</th>
<th>AGE</th>
<th>XL</th>
<th>NI</th>
<th>R**2</th>
<th>R BAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.058</td>
<td>-.075</td>
<td>.038</td>
<td>1.5*</td>
<td>-1.62</td>
<td>.47</td>
<td>.33</td>
</tr>
<tr>
<td></td>
<td>(.044)</td>
<td>(.058)</td>
<td>(.038)</td>
<td>(.542)</td>
<td>(3.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.014</td>
<td>-.054</td>
<td>-.045</td>
<td>-.427</td>
<td>3.33</td>
<td>.21</td>
<td>-.01</td>
</tr>
<tr>
<td></td>
<td>(.048)</td>
<td>(.035)</td>
<td>(.186)</td>
<td>(.486)</td>
<td>(2.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.042</td>
<td>.017</td>
<td>.033</td>
<td>.187</td>
<td>-2.74</td>
<td>.22</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>(.078)</td>
<td>(.201)</td>
<td>(.744)</td>
<td>(.729)</td>
<td>(3.49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.106</td>
<td>-.149</td>
<td>.158</td>
<td>.834</td>
<td>.141</td>
<td>.34</td>
<td>-.18</td>
</tr>
<tr>
<td></td>
<td>(.098)</td>
<td>(.208)</td>
<td>(.177)</td>
<td>(1.05)</td>
<td>(3.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-.142</td>
<td>.692</td>
<td>-.537</td>
<td>.27</td>
<td>-12.1</td>
<td>.35</td>
<td>-.16</td>
</tr>
<tr>
<td></td>
<td>(.161)</td>
<td>(.442)</td>
<td>(.569)</td>
<td>(.731)</td>
<td>(10.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-.249</td>
<td>-.124</td>
<td>.74</td>
<td>-3.98</td>
<td>-2.61</td>
<td>.54</td>
<td>.18</td>
</tr>
<tr>
<td></td>
<td>(.523)</td>
<td>(1.09)</td>
<td>(3.58)</td>
<td>(3.46)</td>
<td>(18.4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* significant at the .05 level
** significant at the .01 level
ENDNOTES


5. Isard, p. 60.


7. Ibid, p. 76.

8. Ibid, p. 74 and 75.


10. Ibid, p. 20.


22. Lowry, p. 44.


27. Greenwood, p. 419.


29. Greenwood, p. 419.


33. Ibid, p. 141 and 142.


35. Greenwood, p. 400.

37. Miller, p. 397.


40. Lowry, p. 29.


42. Miller, p. 396.

43. Hicks, pp. 8 and 9.


46. Gisser, p. 494.


50. Ibid, pp. 77 and 78.


52. Piore, p. 2.

53. Ibid, p. 80.

54. Ibid, p. 84.


56. Ibid, p. 94.


60. Piore, p. 133.

61. Piore, p. 147.


BIBLIOGRAPHY


