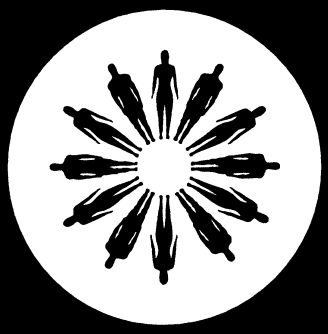


National Longitudinal Surveys

U.S. Department of Labor
Bureau of Labor Statistics

Discussion Paper

Self-Selection and Internal Migration in the United States

George J. Borjas
Stephen G. Bronars
Stephen J. Trejo

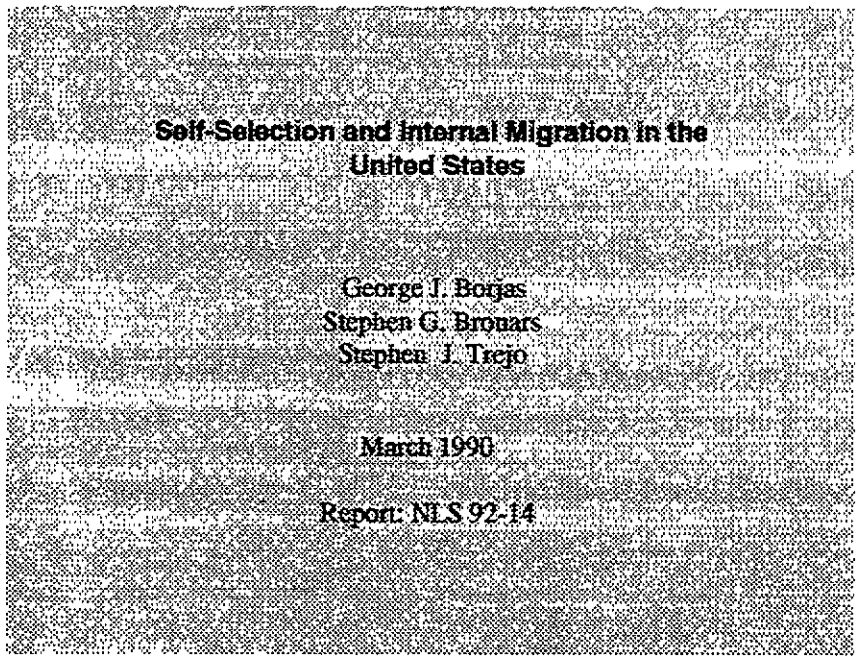
March 1990

Report: NLS 92-14

The National Longitudinal Surveys (NLS) program supports many studies designed to increase understanding of the labor market and methods of data collection. The *Discussion Papers* series incorporates some of the research coming out of the program. Many of the papers in this series have been delivered as final reports commissioned as part of an extramural grant program, under which grants are awarded through a competitive process. The series is intended to circulate the findings to interested readers within and outside the Bureau of Labor Statistics.

Persons interested in obtaining a copy of any *Discussion Paper*, other NLS reports, or general information about the NLS program and surveys should contact the Bureau of Labor Statistics, Office of Economic Research, Washington, DC 20212-0001 (202-606-7405).

NLS Discussion Papers



Self-Selection and Internal Migration in the United States

George J. Borjas
University of California at San Diego, and Unicon

Stephen G. Bronars
University of Texas, and Unicon

Stephen J. Trejo
University of California at Santa Barbara, and Unicon

March 1990

This final paper was funded by the U.S. Department of Labor, Bureau of Labor Statistics under contract number J-9-J-7-0093. Opinions stated in this paper do not necessarily represent the official position or policy of the U.S. Department of Labor.

SELF-SELECTION AND INTERNAL MIGRATION IN THE UNITED STATES

George J. Borjas, Stephen G. Bronars, and Stephen J. Trejo

Executive summary

The population of the United States is highly mobile. Since the 1960s, approximately three percent of the population moves across state lines in any given year, and 10 percent of the population moves across state lines in a five-year period. In view of the falling fertility rates of American women, this extensive mobility implies that migration has become an increasingly important source of demographic change in the various regions and a major determinant of concurrent changes in regional economic growth.

Internal migration rates are especially high among the young. In the National Longitudinal Survey of Youth, over half of the sample are currently (i.e., in 1986) residing in a state other than their state of birth, and about 40 percent are living in a different state than they were at age 14 (see Table 1). The data also indicate that approximately 8 percent of white young men and 6 percent of black and Hispanic young men move across state boundaries in any given year. Our study presents a theoretical and empirical analysis of the internal migration of young workers in the United States.

Guided by income-maximizing models, previous research has focused on "explaining" the size and direction of migration flows across areas. Generally, these studies conclude that persons migrate from low-income regions to high-income regions, and that high migration costs deter interregional mobility. We extend this framework by analyzing not only the size and direction of migration flows, but also their skill composition. In addition to regional income differentials, our model stresses the importance

TABLE 1

INTERSTATE MIGRATION RATES IN NATIONAL LONGITUDINAL SURVEYS OF YOUTH

	Whites		Blacks		Hispanics	
	Males	Females	Males	Females	Males	Females
% Living in Different State Than at Birth	52.4	50.1	40.8	43.7	57.6	57.9
% Living in Different State Than at Age 14	41.3	38.8	29.9	33.4	37.1	36.8
% Living in Different State Than in 1979	32.1	28.5	23.3	25.4	20.6	21.1
% Moving Across State Lines Per Year:						
1979-1980	8.6	8.4	6.8	7.8	6.6	5.7
1980-1981	8.2	7.8	5.8	6.6	6.8	5.7
1981-1982	9.8	7.7	5.3	8.4	8.4	7.6
1982-1983	8.7	8.1	6.6	7.3	4.9	7.2
1983-1984	5.9	4.6	3.9	4.3	3.1	2.7
1984-1985	5.6	5.2	4.6	5.0	2.9	3.1
1985-1986	8.8	7.4	5.6	5.8	4.5	5.9

of regional differences in the returns to skills. Because the impact of internal migration on economic conditions depends not only on how many persons move, but also on which persons move, we believe that these questions are as important as those that have dominated the literature.

In this framework, each region is characterized by an earnings distribution, and persons must decide which among the various earnings opportunities is best for them (net of migration costs). The potential income which could be earned by a particular worker in a particular region depends on the mean level of earnings offered in the region, as well as on the rate of return paid to the specific skills possessed by the worker. Some regions offer higher mean incomes than other regions, and, similarly, some regions offer higher prices for skills than other regions.

Income-maximizing behavior generates a well-defined equilibrium sorting in which differences in rates of return to skills among regions are the main determinants of the skill composition of internal migration flows. In particular, regions that pay higher returns to skills attract more skilled workers than regions that pay lower returns. This result implies that highly skilled workers currently residing in a region that offers a relatively low payoff for those skills, or that unskilled workers currently residing in a region that penalizes their lack of skills, are the most mismatched spatially, and therefore these workers have the greatest incentive to move. In an important sense, our theoretical approach extends earlier models by emphasizing the role of prices in allocating people among regions.

To test our theory, we use the 1979-1986 waves of the National

Longitudinal Survey of Youth. Respondents are between the ages of 14 and 22 at the time of the first interview, and the subsequent annual surveys provide a detailed history of each individual's labor force activity and geographic mobility. In our empirical work, we assume that the region of residence at age 14 is exogenously determined by parental location decisions. Because optimal location decisions for parents and children need not coincide, a spatial mismatch may arise for youths. Our empirical analysis, therefore, focuses on the study of migration patterns after age 14, as youths relocate themselves in order to correct these initial mismatches. Further, we focus on migration across state boundaries--rather than across regional or county boundaries--throughout the study.

To analyze the skill composition of the migrant flow, we use alternative measures of a worker's skills, including educational attainment, wages (both with and without adjustment for differences in demographic characteristics), and the worker's score on the Armed Forces Qualification Test. The theory predicts that highly skilled workers are most likely to leave states with low rates of return to skill, because highly skilled workers are most mismatched in these states. We estimate mover/non-mover probit models where mover status is determined by race and sex dummy variables and by skills.

The results indicate that, in general, migration rates are higher for the more skilled. However, as predicted by the theory, the data also indicate that an increase in skills has a larger impact on the migration propensity in states which offer relatively small payoffs to skill. For example, a one standard deviation increase in years of education raises the probability that a worker leaves a state with a low return to skills by

about 5.3 percent, while the same change in skills has a trivial impact (.8% percent) on the migration propensities of individuals in high return states. Therefore, it appears that high ability workers are migrating out of states with low returns to skills and low ability workers are migrating out of states with high returns to skills.

Our theoretical model also predicts that skilled workers will move to states with high returns to skills. We test this implication by viewing the change in the rate of return to skills in the state of residence between age 14 and 1986 as a choice variable. We find that this variable is positively related to skill endowments. Put differently, skills are an important determinant of the direction of internal migration flows. Moreover, these effects are economically important. A one standard deviation increase in years of education increases the rate of return to skills in the destination state by about 25 percent.

The empirical analysis, therefore, provides important insights into the internal migration process. Individuals are more likely to move the greater the mismatch between skill endowments and the returns to skills. Moreover, the direction of internal migration flows are guided by comparative advantage. Highly skilled workers tend to migrate to states that pay a larger reward to labor market skills, while unskilled workers are likely to end up in states where earnings are relatively insensitive to skills. Overall, the empirical evidence suggests that internal migration plays an important role in determining the equilibrium sorting of skills across states.

SELF-SELECTION AND INTERNAL MIGRATION IN THE UNITED STATES

George J. Borjas, Stephen G. Bronars, and Stephen J. Trejo

March 1990

This project was funded by the U.S. Department of Labor Bureau of Labor Statistics under Contract Number J-9-J-7-0093. Opinions stated in this document do not necessarily represent the official position or policy of The U.S. Departrment of Labor.

SELF–SELECTION AND INTERNAL MIGRATION IN THE UNITED STATES

George J. Borjas, Stephen G. Bronars, and Stephen J. Trejo*

I. Introduction

The population of the United States is highly mobile. Since the 1960s, approximately three percent of the population moves across state lines in any given year, and 10 percent of the population moves across state lines in a five–year period (U.S. Department of Commerce, annual). In view of the falling fertility rates of American women, this extensive mobility has become an increasingly important source of demographic change in the various regions and a major determinant of concurrent changes in regional economic growth.

The theoretical and empirical study of internal migration has a long history in labor economics (Greenwood, 1975, 1985). The substantive importance of this topic arises because internal migration helps equilibrate economic conditions across labor markets in a competitive economy. The study of the size, direction, composition, and economic impact of mobility flows across labor markets can be used to test the extent to which the neoclassical model provides a reasonable description of labor market conditions in the United States (Topel, 1986).

Guided by the income–maximizing models of Hicks (1932) and Sjaastad (1962), early empirical research focused on "explaining" the size and direction of migration flows across areas, and determining why certain groups of individuals, such as the highly educated, are more likely to migrate than others (Bowles, 1970; Greenwood, 1969; Schwartz, 1973; Vanderkamp, 1971). During the mid–1970s, economists began to focus on the role of the family in the migration decision (DaVanzo, 1976; Mincer, 1978; Polachek and Horvath, 1977; and Sandell, 1977). These studies yield a number of interesting

empirical results and theoretical insights, and suggest that Becker's (1981) approach for modeling economic behavior within the family provides a useful framework for the analysis. Despite the central role played by internal migration in the operation of a competitive economy, we believe it is fair to conclude that the study of internal migration has not been at the forefront of research in mainstream labor economics in the past decade.[1]

The recent disinterest in the study of internal migration flows is not symptomatic of a more general indifference in the study of geographic mobility. After all, the analysis of international migration flourished in recent years (Borjas, 1987, 1990; Chiswick, 1978; Freeman and Abowd, 1990). This burgeoning literature suggests that combining the Hicks–Sjaastad income–maximizing approach with the insights provided by Roy's (1951) self–selection model generates new substantive insights about the immigration process, and helps resolve many of the empirical puzzles in international migration data.

This paper presents an application of this more general approach to the analysis of internal migration flows in the United States. We argue that the Hicks–Sjaastad migration model provides a much too restrictive framework for the study of internal migration. Its key predictions are that persons migrate from low–income regions to high–income regions and that higher mobility costs deter migration. In turn, these predictions help focus the empirical work almost exclusively on "explaining" the size and direction of population flows across regions. Although the data generally support these predictions, there are many other interesting questions left unaddressed by the model and by the empirical studies.

The Hicks–Sjaastad model is restrictive because (for given migration costs) it focuses solely on the fact that mean income levels differ across regions, and these income differentials generate one–way migration flows. The theoretical approach suggested by the Roy model stresses regional differences in the returns to skills (in addition to the income differentials). We will show that these price differentials are the main determinants of the skill composition of the population flow across regions. Regions that pay higher returns to

skills attract more skilled workers than regions that pay lower returns. Our analysis thus raises fundamental questions about the skill composition of internal migrants. Because the impact of internal migration on economic conditions depends not only on how many persons move, but also on which persons move, we believe that these questions are as important as those that have dominated the literature.

Of course, the literature that estimates the returns to migration has been well aware of the fact that migrants are self–selected. The development of econometric techniques to account for selection bias (Heckman, 1979) encouraged a number of applications of this methodology to the analysis of migrant earnings. Studies by Nakosteen and Zimmer (1980), Hunt and Kau (1985), and Robinson and Tomes (1982) report the standard selectivity–corrected earnings functions in the mover and stayer samples. The evidence in these papers is inconclusive because of the sensitivity of the methodology, and also because the results are not couched within the equilibrium sorting predicted by the economic theory of selection.

This paper presents a theoretical and empirical analysis of the role self–selection plays in determining the size, direction, and skill composition of internal migration flows. Within the conceptual framework provided by the Roy model, we present an empirical analysis of internal migration flows using the National Longitudinal Survey of Youth. As shown in Table 1, internal migration is very prevalent among these young men and women. Over 50 percent of the persons in the sample are currently (i.e., in 1986) residing in a different state than at the time of birth, and about 40 percent are living in a different state than at age 14. The data also indicate that approximately 8 percent of white young men and 6 percent of black and Hispanic young men move across state boundaries in any given year.

Our empirical analysis indicates that interstate differences in the rewards to skills are a major determinant of internal migration flows. Generally, persons whose skills are most mismatched with the reward structure offered by their current state of residence are

the persons most likely to migrate out of that state. Moreover, the data indicate that these persons tend to relocate in states which offer higher rewards for their particular skills. The empirical evidence, therefore, implies that the Roy model provides a useful framework for analyzing internal migration flows in the United States.

II. Theory

Consider a country partitioned into k distinct geographic regions, indexed by i=1,...,k. To simplify the exposition, we initially assume that there are no costs of relocating across regions. After birth, therefore, individuals compare their earnings opportunities in the various regions and move to the region that maximizes their earnings. This approach provides a simple and intuitive description of the equilibrium skill sorting generated by endogenous migration flows. Further, we assume that the initial distributions of individual skills are the same for all regions. At the time of birth, therefore, individuals are randomly allocated across regions in terms of their skills. The population log earnings distribution in region i is given by:

$$\log w_i = \mu_i + v_i, \quad i=1,...,k \tag{1}$$

where μ_i is the mean income that would be observed in region i in the absence of any internal migration, and v_i is a random variable with mean zero and variance σ_i^2 that measures person–specific deviations from mean income in region i.

The assumption that the initial skill distributions are identical across regions makes the income distributions in (1) independent of initial conditions: the same earnings determination process in region i applies to all persons regardless of their region of birth. This assumption ignores the possibility that persons born in region i are distinctly different, on average, from persons born in region j, and hence the income distributions should also be subscripted for region of birth. The main results of our model are unaffected

by this complication. Hence we maintain the simplifying assumption that the initial allocation of skills does not affect the shape of the population income distributions in any of the regions.[2] However, because of regional differences in the level of resources, physical capital, and aggregate economic conditions, mean incomes μ_i and the distribution of the random variables v_i will vary across regions. Therefore, there will be regional differences in the shapes of the income distributions, even in the absence of differences in the initial skill distributions of workers.

An individual chooses to reside in region j whenever:

$$\log w_j > \max_{r \neq j} [\log w_r] \tag{2}$$

As is well known (Heckman and Honore, 1987; Roy, 1951), the characteristics of the sorting generated by the decision rule in (2) cannot be described without additional restrictions on the distribution of the random variables $v_1,...,v_k$. A simplifying assumption which allows a complete characterization of the equilibrium sorting is that individual earnings are perfectly correlated across regions, so that $\text{Corr}(v_i, v_j) = 1$ for all i,j. The population income distribution of region i can then be written as:

$$\log w_i = \mu_i + \eta_i v, \quad i=1,...,k \tag{1'}$$

This formulation of the regional differences in earnings opportunities implies that the same random variable v determines individual earnings in the various regions (up to a factor of proportionality). It is instructive to think of v as indexing the worker's ability or skills. In effect, equation (1′) assumes that the earnings determination process can be characterized by a one–factor model of ability. The factor of proportionality η_i can be thought of as the factor–loading parameter, or more generally as the "rate of return" to

skills in region i. It is useful to arbitrarily order the regions such that they they are ranked in terms of η with $\eta_1 < \ldots < \eta_k$. We also assume that v is a continuous random variable with a range defined over the real number line. This is the only restriction we make on the density of skills in the population.

Although the assumption that earnings are perfectly correlated across regions is quite strong (for it implies that the ranking of individuals by skill level is the same in all regions), it enables us to derive a number of testable implications from a multi–region selection model. Furthermore, this framework may provide a reasonably accurate representation of earnings opportunities across regions in the United States, given the relatively strong regional similarities in culture, law, and economic institutions.

It is worth stressing that we have not addressed the issue of whether the random variable v is observed or unobserved to the researcher. From an individual's point of view, this distinction is irrelevant. Individuals sort themselves across regions on the basis of all their skills, not just those that happen to be econometrically convenient. By using a one–factor model of ability, we assume that the relative prices of all skills are the same across regions, so that the composite commodity theorem allows us to focus on a single good that is being "sold" across regions. It is not difficult to analyze the migration decision within the context of a multi–factor model of ability, but we do not pursue this generalization because it detracts from the essential points that we make in this paper.

Using equations (1') and (2), it is easy to show that income–maximizing behavior implies that the equilibrium sorting of skills across the k regions is given by:

$$\text{Choose region 1: } -\infty < v < \min_{i=2,\ldots,k} \left[\frac{\mu_1-\mu_i}{\eta_i-\eta_1}\right] \tag{3a}$$

$$\text{Choose region j: } (1 < j < k) \quad \max_{i=1,\ldots,j-1} \left[\frac{\mu_i-\mu_j}{\eta_j-\eta_i}\right] < v < \min_{i=j+1,\ldots,k} \left[\frac{\mu_j-\mu_i}{\eta_i-\eta_j}\right] \tag{3b}$$

$$\text{Choose region k:} \quad \max_{i=1,\ldots,k-1} \left[\frac{\mu_i - \mu_k}{\eta_k - \eta_i}\right] < v < \infty \tag{3c}$$

In view of the assumption that v takes on values from minus to plus infinity, equations (3a) and (3c) imply that some persons will necessarily reside in the "extreme" regions 1 and k (i.e., the regions with the minimum and maximum η). It is less apparent, however, that some persons choose to locate themselves in any other region. In particular, it may well be the case that for some parameters values, equation (3b) will not be satisfied by region j, and hence no workers will choose to reside there. After all, it would seem that persons with positive v's have the most to gain by migrating to the region with the highest η, while persons with negative v's have the most to gain by moving to the region with the lowest η.

Because unpopulated regions are of no interest empirically, we restrict our attention to "interior solutions", i.e., to regions where some individuals choose to reside. Equation (3b) implies that a necessary condition for region j to be inhabited is $\frac{\mu_{j-1} - \mu_j}{\eta_j - \eta_{j-1}} < \frac{\mu_j - \mu_{j+1}}{\eta_{j+1} - \eta_j}$. This restriction can be rewritten as:

$$\mu_j > \frac{(\eta_{j+1} - \eta_j)\mu_{j-1} + (\eta_j - \eta_{j-1})\mu_{j+1}}{(\eta_{j+1} - \eta_{j-1})} \tag{4}$$

Equation (4) defines the Existence Condition that mean incomes in region j must satisfy in order to attract and retain a population. We assume that the Existence Condition is satisfied by all j (j=2,...,k–1). This assumption greatly simplifies the characterization of the equilibrium sorting. In particular, the repeated use of the Existence Condition to make pairwise comparisons of the arguments in the min(·) and max(·) terms in (3) implies that the sorting of skills across regions can be written as:

$$\text{Choose region 1: } -\infty < v < \left[\frac{\mu_1-\mu_2}{\eta_2-\eta_1}\right] \tag{5a}$$

$$\text{Choose region j: } \left[\frac{\mu_{j-1}-\mu_j}{\eta_j-\eta_{j-1}}\right] < v < \left[\frac{\mu_j-\mu_{j+1}}{\eta_{j+1}-\eta_j}\right] \quad (1 < j < k) \tag{5b}$$

$$\text{Choose region k: } \left[\frac{\mu_{k-1}-\mu_k}{\eta_k-\eta_{k-1}}\right] < v < \infty \tag{5c}$$

Figure 1 illustrates the nature of the equilibrium skill sorting when there are five regions. The least skilled workers choose the region with the lowest rate of return to skills, while the most skilled workers choose the region with the highest rate of return. Persons with intermediate levels of skills choose the intermediate regions, with the more skilled workers choosing regions with higher rates of return. In effect, income–maximizing behavior induces a positive correlation between the average skill level of the region's inhabitants and the rate of return to skills in that region. This result does not depend on any assumptions about the density of skills in the population. We summarize this theoretical implication by:

$$E(v \mid \text{choose i}) > E(v \mid \text{choose j}) \text{ if } \eta_i > \eta_j. \tag{6}$$

The assumption that earnings are perfectly correlated across regions implies that individuals who rank highly in the income distribution in one region also rank highly in the income distribution of another. Highly skilled workers, therefore, are attracted to regions with high η because these workers can then enjoy a more generous return on their superior skills. In contrast, less–skilled workers choose regions with less income inequality because this choice minimizes the economic penalty for lacking these skills. In essence, prices play an important allocative role in the internal migration decision.

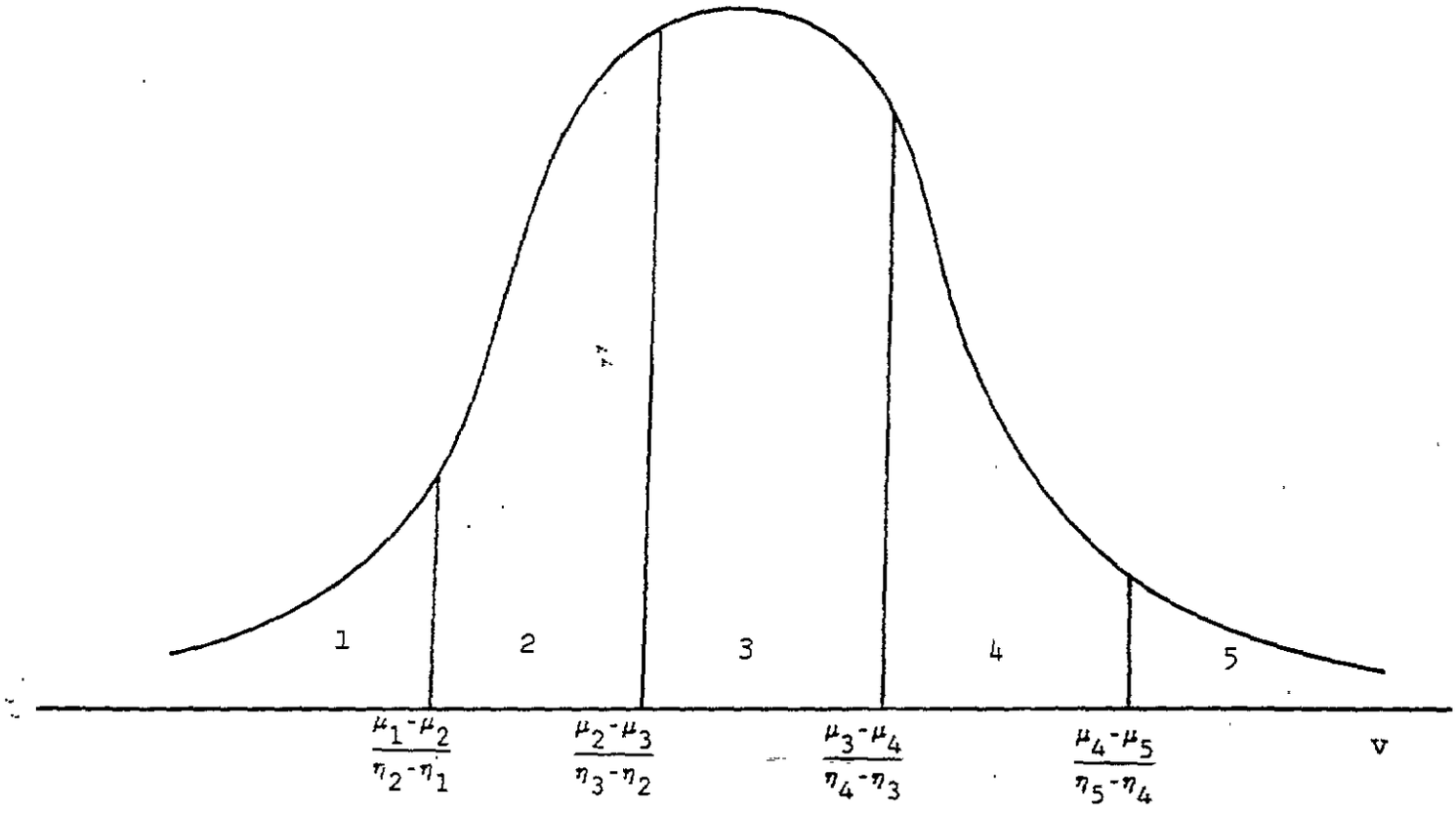


Figure 1. Residential location in a 5-region model, with zero mobility costs.

This insight helps explain the economic content of the Existence Condition. In order for region j to be inhabited, the inequality in (4) requires that mean earnings in region j exceed a weighted average of mean earnings in the "neighboring" regions, j–1 and j+1. Note that these neighboring regions need not be geographically contiguous, but are instead neighbors in an economic sense. Because neighboring regions offer relatively similar rewards for the skills of potential migrants, these are the regions which compete with region j in attracting the human capital of potential migrants.

Suppose that mean earnings in region j are below mean earnings in both neighboring regions. The Existence Condition is not satisfied and no individuals choose to locate in j. For some individuals to reside in region j, therefore, mean earnings in j must exceed mean earnings in either region j–1 or region j+1, or in both. Because these neighboring regions offer either a lower or a higher rate of return to skills than region j, they hold a natural advantage over j in attracting residents. In other words, for the same mean earnings, highly skilled individuals prefer the region with a higher rate of return to skills, while unskilled individuals prefer the region which least penalizes their lack of skills. Therefore, if mean earnings were equal in all three of these regions, or if mean earnings in j were lower than mean earnings in both of the neighboring regions, region j does not make a competitive offer to potential migrants. In contrast, a sufficiently higher mean income in region j than in either of its neighboring regions "compensates" potential migrants for region j's relative disadvantage and attracts a population.

Although the Existence Condition imposes a specific pattern of economic opportunities across regions, it is less restrictive than it seems. For instance, in a 3–state model the Existence Condition only rules out the case where the relationship between mean incomes and η is U–shaped (as well as the case where mean income is constant across regions). All 3 states can be populated if mean incomes are monotonically increasing or decreasing in η, or if mean incomes and η are related in an inverted U–shape. This restriction generalizes to a model with more than 3 regions. If the relationship between

mean incomes and η is flat or U–shaped anywhere, the region with the smallest income would not be able to compete with the neighboring regions and would cease to "exist."

This discussion suggests an important avenue for future research. Regions can "compete" in the population market only if they offer economic opportunities that compensate potential migrants for any disadvantages that accompany the offer (relative to other offers). In a more general model, therefore, the parameters that summarize the region's income distribution are themselves endogenous, and are determined simultaneously with the skill sorting that characterizes the self–selection of workers among regions. This general equilibrium model would also introduce the role played by the prices of fixed factors, such as land, in the equilibrating process. Although research in this topic is in its infancy (see, for instance, Heckman and Sedlacek, 1985; Roback, 1988), it is clear that this type of analysis should provide a much deeper characterization of labor market equilibrium.

The discussion also highlights a feature of the k–state selection model that is shared by the standard 2–state Roy model. In both models, the ranking of skill prices across regions completely determines where a region ranks in terms of the average skill level of its inhabitants. As long as the Existence Condition is satisfied, mean incomes play no role in determining the region's ranking in the skill distribution. Mean incomes, however, do affect the fraction of the population that chooses to reside in each of the regions.[3] It is obvious from equation (5), for instance, that a larger fraction of the country's population chooses to reside in region j the larger the mean income of that region.

It is worth noting that our approach not only raises a number of new substantive questions regarding the internal migration process, but also simplifies the theory and empirical analysis of internal migration flows. Earlier work has been hampered by the fact that there are k(k–1) possible migration flows in a k–region model, and the size and composition of each of these flows depend on all the parameters of the model. Given the Existence Condition, equation (5) suggests that the size and skill composition of the

population choosing region j can be completely determined from the parameters of the income distributions for that region and the neighboring regions, greatly diminishing the number of parameters that determine migration flows into any given region. In fact, our generalization of the Roy model (based on the assumption that earnings are perfectly correlated across regions) generates an ordered qualitative choice model.[4]

Our model also suggests the interesting implication that region j can be both a source of migrants and a destination for migrants. Within the context of the Hicks–Sjaastad income–maximizing model, it is difficult to provide a convincing explanation for this well–known fact. Our model shows that although mean incomes matter, they are not the sufficient statistic that completely describes the size, direction, and composition of migration flows. As long as skill prices differ across regions, the spatial missorting of individuals at the time of birth is likely to be substantial. Highly–skilled individuals, for instance, may be born in regions with low η and will migrate to regions with high η, while less able workers may be born in regions with high η and will migrate to regions with low η. Some regions, therefore, are likely to observe substantial outflows at the same time that they receive sizable inflows. Two–way population flows, therefore, are bound to occur as the mismatches caused by being born in the wrong place are corrected.

Although these insights have been derived in a model of costless mobility, the introduction of migration costs does not alter any of the key results. For concreteness, consider the migration decisions of persons born in region j. These individuals move to region i ($i \neq j$) if:

$$\log w_i - C_{ji} > \max_{r \neq i} [\log w_r - C_{jr}] \qquad (7)$$

where C_{ji} is a time–equivalent measure of the costs of migrating from region j to i, with $C_{jj} = 0$.[5] For simplicity, we assume that the costs C_{ji} are constant for all persons residing

in region j.[6] There is variation, however, in the costs of moving from j to different states (i.e., $C_{jr} \neq C_{js}$ for $r \neq s$).

Assuming initially that every other region receives at least one migrant from region j (a restriction analogous to the Existence Condition), the equilibrium sorting of individuals born in region j is given by:

$$\text{Choose region 1: } -\infty < v < \frac{\mu_1 - \mu_2 - (C_{j1} - C_{j2})}{\eta_2 - \eta_1} \tag{8a}$$

$$\text{Choose region i: } (1 < i < k) \quad \frac{\mu_{i-1} - \mu_i - (C_{j,i-1} - C_{ji})}{\eta_i - \eta_{i-1}} < v < \frac{\mu_i - \mu_{i+1} - (C_{ji} - C_{j,i+1})}{\eta_{i+1} - \eta_i} \tag{8b}$$

$$\text{Choose region K: } \frac{\mu_{k-1} - \mu_k - (C_{j,k-1} - C_{jk})}{\eta_k - \eta_{k-1}} < v < \infty \tag{8c}$$

Figure 2 illustrates the sorting of workers born in region 3 when there are five regions and it is costly to move. It is apparent that the equilibrium sorting resembles that obtained in the costless mobility model: Highly skilled workers move to regions with more earnings dispersion than region j, and less skilled workers move to regions with less earnings dispersion than j. The introduction of migration costs, however, alters the cutoff points determining who moves to which region. These thresholds now depend on mean incomes net of migration costs. This fact obviously implies that fewer persons are likely to migrate out of the region of birth.

A simple parameterization of migration costs allows us to determine exactly which regions are most likely to be affected by the fact that geographic mobility is costly. Suppose that the costs of moving from region j to region r are $C_{jr} = \bar{C}$ for $j \neq r$, and 0 otherwise. In this representation, migration costs are simply the fixed costs of moving that do not depend on the distance of the move or on any other factors which vary with the

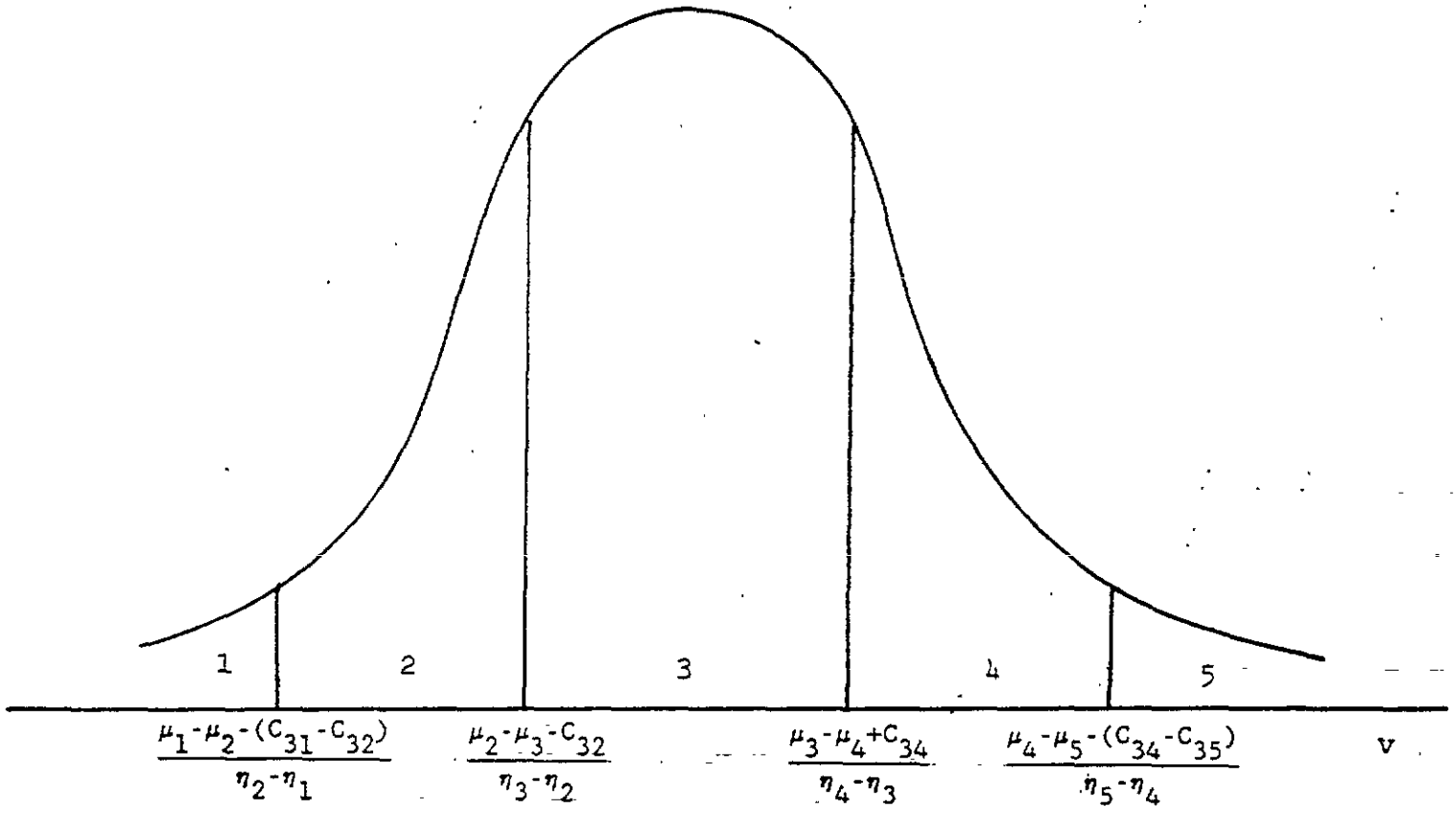


Figure 2. Residential location in a 5-region model for persons born in region 3, with positive mobility costs.

precise identities of the origin and destination. It is apparent from equation (8) or from Figure 2 that these fixed costs cancel out of all the cutoff points except those bordering the region of origin.

Consider an increase in the fixed costs of moving. Obviously, this increases the fraction of region j residents who decide not to migrate. The model also implies that for "small" changes in $\bar{C}$, the pool of workers who previously would have migrated but now decide to remain in j is drawn entirely from the sample of workers who would have moved to neighboring regions (i.e., regions with somewhat similar η's). Thus small changes in fixed migration costs do not alter the size or skill composition of the outflow from region j to non–neighboring regions. Put differently, small changes in fixed migration costs only change the migration incentives of "marginal" workers.

Of course, the larger the increase in fixed migration costs, the more likely that no one will move to the neighboring regions j–1 and j+1. Sufficiently high levels of migration costs, therefore, make it unlikely that all regions are destination choices for persons originating in j. Moreover, those regions which "drop out" of the model are the regions which most resemble j in terms of their payoffs to skills. Because migration is costly, workers who decide to leave region j are the ones who are most mismatched, and hence will move to regions that differ substantially from j in their offers of economic opportunities.

We noted earlier that our model generates two–way population flows without resorting to imperfect information or informational asymmetries among participants in the marketplace. The existence of migration costs adds further substance to this insight. As fixed migration costs increase, fewer people move to neighboring regions from any region of origin. There is a tendency, therefore, for persons to migrate to relatively "extreme" regions (i.e., regions with high or low levels of η). But these are precisely the regions where the costs of being mismatched are largest. Persons born in the extreme regions who are mismatched with their current economic opportunities have a large incentives to leave, and hence extreme regions will be the origin of sizable migration flows. For sufficiently large

levels of fixed migration costs, therefore, there will be large population inflows and outflows in the extreme regions.

Of course, the practical importance of this insight depends on the exact distribution of skills, on the level of fixed (and variable) migration costs, and on the parameters of the income distributions in the k regions. Therefore, it is difficult to quantify the importance of this tendency in the absence of additional restrictions on the model. Nevertheless, it is important to note that introducing migration costs into the analysis does not alter the fundamental insight that highly skilled workers are more likely to be residing in regions that offer higher payoffs for those skills.

III. Empirical Analysis

The theory developed in the previous section generates sharp empirical predictions for the relationship between differences in the returns to skills across regions and such diverse factors as the migration propensities of individuals, the direction of internal migration flows, and the spatial distribution of skills. To test our theory, we use the 1979–1986 National Longitudinal Survey of Youth (NLSY). In our empirical work, we assume that the region of residence at age 14 is exogenously determined by parental location decisions. Because optimal location decisions for parents and children need not coincide, a spatial mismatch may arise for youths. Our empirical analysis, therefore, focuses on the study of migration patterns after age 14, as youths themselves endogenously decide to move in order to correct these initial mismatches. Further, we focus on migration across state boundaries—rather than across regional or county boundaries — throughout the analysis.

Respondents in the NLSY are between the ages of 14 and 22 at the time of the first interview (1979), and from each of the annual surveys we obtain a detailed history of both employment behavior and geographic mobility. Because the comparison of location decisions is essential to our empirical analysis, we delete all observations for whom any

geographic information is missing in any of the surveys. In order to mitigate the impact of schooling decisions on migration flows, we also exclude individuals who left school after 1984. Finally, we restrict our analysis to native–born civilian workers who lived in the United States continuously since age 14. Movers are defined as those persons who reside in a different state in 1986 than at age 14, regardless of where they lived in any of the intervening years. Likewise, nonmovers are persons who reside in the same state in 1986 as at age 14, even if they lived in other states during the intervening years.

To analyze the skill composition of the migrant flow, we use alternative measures of a worker's skills. Our first measure of skills is simply given by the number of years of completed education (measured as of 1986). Our second measure is based on an aptitude test score. Between July and October of 1980, the Armed Services Vocational Aptitude Battery (ASVAB) was administered to about 94 percent of the NLSY respondents. The ASVAB consists of 10 tests that measure knowledge and skills in areas ranging from word knowledge and arithmetic reasoning to mechanical comprehension and electronics information. The military sums the scores of four of the tests (word knowledge, arithmetic reasoning, paragraph comprehension, and half the score in numeric operations) to create the Armed Forces Qualification Test (AFQT). The AFQT is a general measure of aptitude, and its score is standardized so that the population distribution has mean zero and a standard deviation of one.

Finally, we also use an individual's average hourly wage, defined as the ratio of annual earnings to annual hours of work, as a measure of skills. We use the panel nature of the NLSY wage data to construct individual–specific fixed effects. Therefore, our wage–based skill measure is available only in the subsample of workers who have at least two years of wage data.[7] These wage observations need not be in consecutive years.

The construction of the wage–based skill measure is as follows. Consider the earnings function:

$$\log w_{ijt} = \beta_1 \text{ STATE}_{it} + \beta_2 \text{ YEAR} + \beta_3 X_{it} + \epsilon_{ijt} \qquad (9)$$

where $\log(w_{ijt})$ gives the logarithm of individual i's hourly wage in state j in year t. The hourly wage is a function of a vector of state of residence dummies (STATE), a vector of dummy variables indicating the year of the observation (YEAR), and a vector of standardizing variables, X_{it}. This vector includes years of education completed, age, age squared, job tenure, union status, marital status, health status, metropolitan residence, and occupation and industry dummies.

The error term in (9) depends on the state/worker match. We assume that ϵ_{ijt} can be expressed as $\eta_j(v_i+u_{it})$, where the parameter η_j is indexed by state j. We interpret the residual in this earnings regression as the product of the stock of (unobservable) person–specific human capital (v_i) and the state–specific return to human capital (η_j), plus a random error term ($\eta_j u_{it}$). Therefore, differences in the returns to human capital across states generate heteroskedasticity in the earnings function.

Because we have at least two observations on the hourly wage for each individual in our sample, we can generate estimates of v_i. If all states had the same returns to skills, v_i would simply be a person–specific intercept in the regression. However, because skills are rewarded differentially across states, we allow for different intercepts for each worker/state pair in the sample. Our specification of the error term yields the restriction that worker/state intercepts are proportional across states (i.e., that ϵ_{ijt} is proportional to η_j). Note also that the vector of standardizing variables, X, includes only those characteristics that vary over the sample period. Differences in earnings due to person–specific demographic factors that are not time–varying, such as sex and race, enter the regression through v_i.

Our measure of the worker's skills is obtained by estimating (9) using data differenced from person–specific means, and then calculating the individuals' average residuals in the earnings function (appropriately weighted for state of residence).[8] This

approach yields point estimates that are identical to those that would be obtained by including thousands of worker/state intercepts in the regression. The estimated skill measure, which we call the "standardized wage", gives the typical number of standard deviations that a worker's hourly wage is above or below the mean wage for workers with similar demographic characteristics.

We also constructed a wage–based measure of skills that does not control for differences in demographic characteristics. This "unstandardized wage" measure is obtained by reestimating (9) without any variables in the control vector X, and it gives the typical number of standard deviations that a worker's wage is above or below the state average, not controlling for any observable factors.

Table 2 reports the mean of the skill variables (and of some demographic characteristics) in our data. We find that, on average, migrants have significantly higher labor market skills than do non–migrants, whether we measure skills by aptitude test results, years of education completed, or unstandardized and standardized wages.

Table 3 presents the correlation coefficients for pairwise comparisons of our four alternative skill measures. Despite the fact that our skill measures represent very different aspects of ability, they are all highly correlated across individuals. Because much of this correlation could be explained by differences across age/race/sex groups, the second panel of Table 3 presents partial correlation coefficients among the skill measures. The evidence indicates that the correlations between different measures of skills are not appreciably weaker within age/race/sex groups.

Before turning to a formal test of the theory, it is instructive to portray the pattern of migration flows across states in our data. Table 4 presents summary statistics, by state of origin, for the movers in our sample. The first column reports the number of persons who resided in each state as of age 14, and the second column gives the fraction of "natives" who left the state by 1986. The remaining columns of the table describe how the skills of migrants differ from the skills of natives who remained in the state. In order to

make interstate comparisons, we normalize the mean skill level in each state to be zero. Thus a weighted average of the skills of movers and stayers equals zero in each state.

Column 3 of Table 4 reports the AFQT scores of migrants, relative to natives. There is substantial interstate variation in the abilities of migrants. For example, Massachusetts exports workers with above average ability. The average AFQT score of migrants from Massachusetts is about .4 higher than the overall average within the state. Oklahoma, on the other hand, exports persons with lower than average AFQT scores. The remaining columns of Table 4 report similar interstate variation in the skill composition of the migrant outflow.

Table 5 complements this descriptive analysis by reporting the in–migration rate and average skill levels for in–migrants grouped by their state of destination. The first column gives the number of persons residing in each state as of 1986, while column 2 shows the fraction of these persons who moved to this state from a different state of origin. The remaining columns of Table 5 report the average skill levels of in–migrants, measured relative to the mean skill level in the migrants' state of origin.[9]

The data reveal substantial variation in the skills of in–migrants. For instance, the average person migrating to Oklahoma has an AFQT score that is .67 higher than the mean score in their native state. This does not necessarily mean that Oklahoma is "importing" workers with the highest absolute AFQT scores or years of education, but rather that Oklahoma is attracting higher than average ability persons from those states that export workers to Oklahoma.

Our model suggests that the equilibrium sorting of skills across states can be understood in terms of the parameter η_j, the state–specific returns to skills. Because dispersion in the earnings distribution is proportional to skill prices (for a given density of skills), we use the standard deviation of the wage distribution within a state to measure the rewards to skills (Juhn, Murphy, and Pierce, 1989). In view of the relatively small sample sizes in some states in the NLSY, we use the sample of adult men (working in the private

sector) in the 5/100 1980 Census data to estimate earnings dispersion within each state.

Our analysis uses two alternative measures of the state's standard deviation of earnings. The first is the unstandardized dispersion in earnings, measured by the standard deviation of the log hourly wage in the 1980 Census. The second is the standardized dispersion in earnings, measured by the standard deviation of the mean squared error in state–specific log wage equations. These wage regressions included education, age and age squared, and dummies for marital status, immigrant status, and metropolitan residence as standardizing variables.

The theoretical model predicts which workers will migrate out of their initial state of residence; explains where workers will migrate as a function of their skill endowment and the prices of skills in other states; and describes the nature of the resulting allocation of workers across states. We now conduct empirical tests of each of these implications.

In our model, migration decisions are motivated by an initial mismatch between workers and states. High ability workers are more likely to leave states where skill prices are relatively low (even though mean earnings in the state are sufficiently high to retain other workers in the state). Conversely, unskilled workers are most likely to leave states where the dispersion in earnings, and hence skill prices, are high. The theory thus predicts that an increase in skills has the greatest impact on the probability of moving in states with little income inequality, because highly skilled workers are most mismatched in these states. We estimate mover/non–mover probit models where mover status is determined by race (white, black, Hispanic) and sex dummy variables and by skills.[10] By estimating separate probit models for different groups of states of residence at age 14, we allow the magnitude of the initial mismatches to influence migration propensities.

In particular, we divide the sample into four equally sized partitions, based on the rank ordering of the unstandardized wage dispersion in each state (we obtained similar results with the standardized dispersion). Table 6 reports the effect of the alternative skill measures on the migration propensity. The first column presents the impact of each of the

skill variables on the migration probability in models estimated over the 25% of the population who, at age 14, resided in states with the lowest unstandardized wage dispersion. In addition, we report the estimated impact of a one standard deviation change in skills on the migration probability, evaluated at sample means. The remaining columns present the same statistics for individuals who, at age 14, resided in states with greater unstandardized wage dispersion.

The regressions indicate that, in general, migration rates are higher for the more skilled. The data also indicate, however, that an increase in skills has a bigger impact on the migration propensity in states with relatively little wage dispersion (columns 1 and 2). For example, a one standard deviation increase in years of education raises the probability that a worker leaves a state with low wage dispersion by about 5.3 percent, while the same change in skills has a trivial impact (.8 percent) on the migration propensities of individuals in high dispersion states. The same qualitative results hold true for AFQT scores, the unstandardized wage, and the standardized wage. Because low dispersion in earnings within a state is indicative of low returns to skills, it appears that high ability workers are much more likely to migrate out of states with relatively low returns to skills.

Our theoretical model also predicts that skilled workers will move to states with greater wage dispersion, and that unskilled workers will move to states with less dispersion. We test this implication by viewing the change in earnings dispersion between age 14 and 1986 as a choice variable. For individuals who resided in the same state in 1986 as at age 14, we define this change to be zero. For the sample of movers, we define it as the difference between wage dispersion in the state of residence in 1986 and that in the state of residence at age 14. The mean change in the unstandardized dispersion is .002, while the change in the standardized dispersion is .003 (the mean changes are .012 and .013, respectively, when averaging over the sample of movers). The data thus indicate that, on average, workers choose to migrate to states with greater wage dispersion. The theory predicts that this change in earnings dispersion should be positively related to the skills of

workers.

We begin testing this hypothesis by estimating OLS regressions of the change in dispersion on race and sex dummies and the alternative skill measures. Table 7 presents the results of these OLS regressions for the entire sample and for the subsample of movers. Column 1 reveals a strong and significant positive relationship between a worker's skills and the change in the unstandardized dispersion chosen, while column 2 reports similar results for changes in the standardized dispersion variable.

The last two columns of Table 7 present estimates of the same regressions for the subsample of movers. We again find that skill endowments are positively related to the change in the wage dispersion chosen by individuals who left their initial state of residence. Put differently, skills are an important determinant of the direction of internal migration flows. Moreover, these effects are economically important. A one standard deviation increase in years of education increases the change in dispersion chosen by individuals by about 25 percent, while a one standard deviation increase in AFQT scores yields a 10 percent increase.

An alternative way of modeling the choice of destination for internal migrants is to explain the direction but not the magnitude of the change in dispersion chosen by movers. Each migrant has two choices: move to a state with less wage dispersion, or move to a state with greater dispersion. This discrete variable is preferred to the continuous dependent variable used in Table 7 if our rank ordering of states by wage dispersion accurately represents the true rank ordering, but the magnitudes of wage dispersion and hence skill prices are estimated imprecisely.

Table 8 reports the results of a probit model of migration destination decisions estimated in the subsample of movers, where the dependent variable indicates if the worker moved to a state with greater wage dispersion than the state of origin. The estimated coefficients reveal a strong positive relationship between the alternative measures of skills and the propensity to move to a state with greater wage dispersion.

The empirical evidence, therefore, provides important insights into the internal migration process. Individuals are more likely to move the greater the mismatch between skill endowments and the returns to skills. Moreover, the direction of internal migration flows are guided by comparative advantage. Highly skilled workers are much more likely to migrate to a state with greater wage dispersion, while unskilled workers are more likely to choose a decrease in wage dispersion after they move.

The model also implies that persons locate across states according to a rank ordering of their abilities. In the absence of mobility costs, the state with the highest returns to skills will attract the most able workers. The state with the second highest returns to skills will attract workers who are less able than the least skilled worker in the first state, but who are more able than the most skilled worker in the third state. Put differently, states can be ranked equivalently according to the skill levels of their workers or by the extent of wage dispersion within the state.

An empirical test of this theoretical implication can be conducted by estimating an ordered probability model of an individual's choice of residence in 1986. Of course, the introduction of mobility costs and of a nonrandom initial distribution of persons across states (with respect to their skills) weakens the relationship between the returns to skills and the skill endowments of workers across states.

Before estimating an ordered probability model, therefore, it is important to determine if the initial distribution of persons across states is random with respect to skills. Table 9 reports the correlation coefficients between wage dispersion in a state and the level of skills for individuals living in that state as of age 14. There is a strong negative correlation between the average skills of "natives" and wage dispersion in the state. If mobility costs are sizable, it is unlikely that internal migration flows can offset the initial nonrandom distribution of skills across states, and yield a positive relationship between wage dispersion in a state and the average level of skills in that state.

We control for the nonrandom initial distribution of skills by normalizing mean

skills in each state to be zero (in effect, removing a state–specific mean from each worker's skills measure). By construction, these normalized skill measures are uncorrelated with wage dispersion in the state of origin.

Divide the population into four equally sized groups, ranked on the basis of the wage dispersion in the 1986 state of residence (location 1 has the lowest dispersion and location 4 the highest).[11] In the absence of mobility costs, individuals sort perfectly into these four groups on the basis of their labor market skills. Let v_i denote the latent (unobserved) variable representing an individual's labor market skills. Our theoretical model generates the following sorting of skills into each of the four groups of states:

$$\text{Choice of location in 1986} = \begin{cases} 1 \text{ if } v_i \leq \bar{v}_1 \\ 2 \text{ if } \bar{v}_1 < v_i \leq \bar{v}_2 \\ 3 \text{ if } \bar{v}_2 < v_i \leq \bar{v}_3 \\ 4 \text{ if } \bar{v}_3 < v_i \end{cases} \tag{10}$$

The cutoff values $\bar{v}_1$, $\bar{v}_2$, and $\bar{v}_3$ depend on the parameters of the earnings distributions.

Let Z_i denote our proxy for the skills of person i (as measured by AFQT scores, years of education, the unstandardized wage, or the standardized wage). The optimal sorting of workers across states is summarized by the ordered probit model:

$$\text{Choice of location in 1986} = \begin{cases} 1 \text{ if } v_i = \alpha_1 Z_i + \alpha_2 F_i + e_i \leq \bar{v}_1 \\ 2 \text{ if } \bar{v}_1 < v_i = \alpha_1 Z_i + \alpha_2 F_i + e_i \leq \bar{v}_2 \\ 3 \text{ if } \bar{v}_2 < v_i = \alpha_1 Z_i + \alpha_2 F_i + e_i \leq \bar{v}_3 \\ 4 \text{ if } \bar{v}_3 < v_i = \alpha_1 Z_i + \alpha_2 F_i + e_i \end{cases} \tag{11}$$

where the vector F_i includes race and sex dummy variables. The estimation procedure yields point estimates of two of the three cutoff values (one of the cutoff values is arbitrarily normalized to zero), and estimates of α_1 and α_2, the coefficients of Z_i and F_i in

(11). Our theory implies that the coefficient α_1 is positive. This implies that more skilled persons locate in states with higher returns to skills.

Table 10 reports the estimated ordered probit models. The equations presented in the first two columns indicate that the coefficient α_1 is positive and statistically significant in most of the regressions. The spatial sorting of skills in the United States, therefore, resembles that predicted by the economic theory of self–selection.

The empirical support for the theory is considerably weakened if we do not normalize skills with respect to the state of residence at age 14 (as in columns 3 and 4). However, this probably reflects the combined impact of a nonrandom initial distribution of skills across states and sizable mobility costs. Overall, therefore, the empirical evidence suggests that internal migration plays an important role in determining the equilibrium sorting of skills across states.

IV. Summary

This paper presented a systematic analysis of internal migration flows in the United States. Our research is motivated by the insight that internal migrants are not randomly selected from the population. The application of the Roy model to the study of internal migration leads to a number of new theoretical insights, raises many questions that were ignored by the previous literature, and generates a series of empirical results that expand our understanding of the internal migration process.

The main contribution of our conceptual approach is to provide a framework for simultaneously analyzing questions related to the size, direction, and skill composition of the internal migration flow. Our model implies that income–maximizing behavior generates a well–defined equilibrium sorting of skills among competing regions. In particular, regions that offer higher rewards for skills tend to attract more skilled workers. This result suggests, for instance, that highly skilled workers currently residing in a region that offers a relatively low payoff for those skills, or that unskilled workers currently

residing in a region that penalizes their lack of skills are most mismatched spatially, and are the workers most likely to move. In an important sense, our theoretical approach expands the earlier Hicks–Sjaastad framework by stressing the role of prices in allocating people among regions. Our empirical analysis, conducted on the National Longitudinal Surveys of Youth, indicates that the main implications of the model are roughly consistent with the data.

It is apparent that internal migration flows in the United States are sizable and play an important role in equilibrating economic conditions across labor markets. Although the study of internal migration flows has not been at the frontier of research in labor economics in the past decade, we believe that our theoretical and empirical analysis provides a simple framework for guiding future research, and raises a number of substantive questions that are at the core of many labor market issues.

FOOTNORES

*Univcontract from the U.S. Department of Labor.

1. We do not mean to imply that the study of internal migration flows ceased altogether. A large number of studies appeared since the mid–1970s which provide useful extensions of the empirical literature (Bartel, 1979; DaVanzo, 1978, 1983; Fields, 1976, 1979; and Schlottmann and Herzog, 1982, 1984).

2. The operational significance of this assumption is its implication that the mean income observed in region i if all persons born in region r moved to region i is the same as that observed if all persons born in region s moved to region i (where $r \neq s$).

3. Even though mean incomes do not determine the skill ranking of regions, they do affect the average skill level of a region's inhabitants.

4. For a discussion of ordered response models, see Maddala (1983), pp. 46–49.

5. If the dollar costs of moving from j to i are given by D_{ji}, then time–equivalent costs are given by the ratio D_{ji}/w_i. We assume that this ratio is "small" in deriving the condition in (7).

6. It is not difficult to model variation in mobility costs among individuals. In the simpler 2–region model with a normal distribution of skills, it can be shown that the introduction of variable mobility costs does not alter any of the implications of the analysis if earnings and mobility costs are uncorrelated, or if the variance in mobility costs is small relative to the variance in earnings.

7. In constructing wage–based skill measures, we use the following sample restrictions. Hourly wages of less than $.50 or greater than $100 are considered outliers and are excluded from the data. We also delete wage observations for individuals in any given year in which any of the following variables are missing: industry, occupation, job tenure, health status, years of education completed, school enrollment status, marital status, union status, state of residence, and whether the respondent resided in a

metropolitan area.

8. We estimate the parameters in (9), including v_i, by using a two–step generalized least squares procedure to correct for heteroskedasticity across states. We first estimated (9) using OLS, and then calculated state–specific estimates of the dispersion in earnings from the residuals. Using the estimated standard deviations of earnings by state, we reestimated (9) using GLS. It is important to note that the weights used in this procedure are state–specific, and hence as workers move from one state to another we use different weights in estimating the earnings regression.

9. We measure skills of in–migrants relative to their native states because of the rather sizable regional differences in AFQT scores, years of education, and earnings.

10. In estimating the qualitative choice models, we assume that the error distribution is normal.

11. The aggregation of the states into four groups greatly simplifies the estimation of the ordered probit model. We experimented with larger numbers of groupings, and found that convergence became more difficult as the number of groupings increased.

REFERENCE

Bartel, Ann P. "The Migration Decision: What Role Does Job Mobility Play?" American Economic Review 69 (December 1979): 775–786.

Becker, Gary S. A Treatise on the Family. Cambridge, Ma.: Harvard University Press, 1981.

Borjas, George J. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants." Journal of Labor Economics 3 (October 1985): 463–489.

Borjas, George J. "Self–Selection and the Earnings of Immigrants." American Economic Review 4 (September 1987): 531–553.

Borjas, George J. Friends or Strangers: The Impact of Immigrants on the U.S. Economy. New York: Basic Books, 1990.

Bowles, Samuel. "Migration as Investment: Empirical Tests of the Human Capital Approach to Geographical Mobility." Review of Economics and Statistics 52 (November 1979): 356–362.

Chiswick, Barry R. "The Effect of Americanization on the Earnings of Foreign–Born Men." Journal of Political Economy 86 (October 1978):897–921.

DaVanzo, Julie. Why Families Move: A Model of the Geographic Mobility of Married Couples. Rand Corporation Report R–1972–DOL, September 1976.

DaVanzo, Julie. "Does Unemployment Affect Migration? — Evidence from Micro Data." Review of Economics and Statistics 60 (November 1978): 504–514.

DaVanzo, Julie. "Repeat Migration in the United States: Who Moves Back and Who Moves On?" Review of Economics and Statistics 65 (November 1983): 552–559.

Fields, Gary S. "Labor Force Migration, Unemployment and Job Turnover," Review of Economics and Statistics 58 (November 1976): 407–415.

Fields, Gary S. "Place to Place Migration: Some New Evidence." Review of Economics and Statistics 61 (February 1979): 21–32.

Freeman, Richard B. and Abowd, John M. Immigration, Trade, and the Labor Market. Chicago: University of Chicago Press, 1990.

Greenwood, Michael J. "An Analysis of the Determinants of Geographic Labor Mobility in the United States." Review of Economics and Statistics 51 (May 1969): 189–194.

Greenwood, Michael J. "Research on Internal Migration in the United States: A Survey." Journal of Economic Literature 13 (June 1975): 397–433.

Greenwood, Michael J. "Human Migration: Theory, Models, and Empirical Studies." Journal of Regional Science 25 (November 1985): 521–544.

Heckman, James J. "Sample Selection Bias as a Specification Error." Econometrica 47 (January 1979): 153–161.

Heckman, James J. and Honore, Bo. "The Empirical Content of the Roy Model." Mimeograph, University of Chicago, 1987.

Heckman, James J. and Sedlacek, Guilherme. "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self–Selection in the Labor Market." Journal of Political Economy 93 (December 1985): 1077–1125.

Hicks, John R. The Theory of Wages. London: Macmillan, 1932.

Hunt, Janet C., and Kau, James B. "Migration and Wage Growth: A Human Capital Approach." Southern Economic Journal 51 (January 1985): 697–710.

Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce. "Wage Inequality and the Rise in Returns to Skill." Mimeograph, University of Chicago, December 1989.

Maddala, G.S. Limited–Dependent and Qualitative Variables in Econometrics. New York: Cambridge University Press, 1983.

Mincer, Jacob. "Family Migration Decisions." Journal of Political Economy 86 (October 1978): 749–775.

Nakosteen, Robert A., and Zimmer, Michael. "Migration and Income: The Question of Self–Selection." Southern Economic Journal 46 (January 1980): 840–851.

Polachek, Solomon W. and Horvath, Francis W. "A Life–Cycle Approach to Migration: Analysis of the Perspicacious Peregrinator." Research in Labor Economics 1 (197): 103–149.

Roback, Jennifer. "Wages, Rents, and Amenities: Differences Among Workers and Regions," Economic Inquiry 26 (January 1988): 23–42.

Robinson, Chris, and Tomes, Nigel. "Self–Selection and Interprovincial Migration in Canada." Canadian Journal of Economics 15 (August 1982): 474–502.

Roy, A.D. "Some Thoughts on the Distribution of Earnings." Oxford Economic Papers 3 (June 1951): 135–146.

Sandell, Steven H. "Women and the Economics of Family Migration." Review of Economics and Statistics 59 (November 1977): 406–414.

Schlottman, Alan M., and Herzog, Henry W., Jr. "Home Economic Conditions and the Decision to Migrate: New Evidence for the U.S. Labor Force." Southern Economic Journal 48 (April 1982): 950–961.

Schlottman, Alan M., and Herzog, Henry W., Jr. "Career and Geographic Mobility Interactions: Implications for the Age Selectivity of Migration." Journal of Human Resources 19 (Winter 1984): 72–86.

Schwartz, Aba. "Interpreting the Effect of Distance on Migration," Journal of Political Economy 81 (September/October 1973): 1153–1169.

Sjaastad, Larry A. "The Costs and Returns of Human Migration," Journal of Political Economy 70 (October 1962): 80–93.

Topel, Robert H. "Local Labor Markets." Journal of Political Economy 94 (June 1986): S111–S143.

Vanderkamp, John. "Migration Flows, Their Determinants and the Effects of Return Migration." Journal of Political Economy 5 (September/October 1971): 1012–1031.

TABLE 1

INTERSTATE MIGRATION RATES IN NATIONAL LONGITUDINAL SURVEYS OF YOUTH

	Whites		Blacks		Hispanics	
	Males	Females	Males	Females	Males	Females
% Living in Different State Than at Birth	52.4	50.1	40.8	43.7	57.6	57.9
% Living in Different State Than at Age 14	41.3	38.8	29.9	33.4	37.1	36.8
% Living in Different State Than in 1979	32.1	28.5	23.3	25.4	20.6	21.1
% Moving Across State Lines Per Year:						
1979-1980	8.6	8.4	6.8	7.8	6.6	5.7
1980-1981	8.2	7.8	5.8	6.6	6.8	5.7
1981-1982	9.8	7.7	5.3	8.4	8.4	7.6
1982-1983	8.7	8.1	6.6	7.3	4.9	7.2
1983-1984	5.9	4.6	3.9	4.3	3.1	2.7
1984-1985	5.6	5.2	4.6	5.0	2.9	3.1
1985-1986	8.8	7.4	5.6	5.8	4.5	5.9

TABLE 2

SUMMARY STATISTICS

	Overall Sample		Movers		Non-Movers	
Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.
Age in 1986	24.9	2.23	25.1	2.21	24.8	2.23
Percent Female	53.8		57.5		53.0	
Percent Black	26.9		22.6		27.9	
Percent Hispanic	15.0		10.4		16.1	
Percent Movers	18.3		100.0		0.0	
Length of Time in Current State (years)	5.90	2.20	3.27	2.70	6.50	1.54
Average Real Wage (1986 dollars)	$5.63	3.10	$5.79	3.24	$5.59	3.07
AFQT score	-0.40	0.89	-0.21	0.93	-0.44	0.88
Years of Education in 1986	12.03	1.98	12.36	2.18	11.95	1.92
Unstandardized Wage	-0.10	0.78	-0.06	0.79	-0.11	0.78
Standardized Wage	0.01	0.90	0.10	0.89	-0.01	0.91
Sample Size	6666		1222		5444	

TABLE 3

CORRELATION OF SKILL MEASURES
(t-statistics in parentheses)

	AFQT Score	Education	Unstandardized Wage	Standardized Wage
AFQT Score	1.00			
Education	0.57 (58.0)	1.00		
Unstandardized Wage	0.36 (27.0)	0.31 (23.3)	1.00	
Standardized Wage	0.41 (31.8)	0.34 (25.3)	0.82 (103.1)	1.00

PARTIAL CORRELATIONS OF SKILL MEASURES
CONTROLLING FOR AGE, SEX, RACE
(t-statistics in parentheses)

	AFQT Score	Education	Unstandardized Wage	Standardized Wage
AFQT Score	1.00			
Education	0.53 (58.3)	1.00		
Unstandardized Wage	0.28 (23.9)	0.28 (21.8)	1.00	
Standardized Wage	0.41 (23.9)	0.36 (19.1)	0.67 (197.8)	1.00

TABLE 4

OUT-MIGRATION RATES AND AVERAGE SKILLS OF OUT-MIGRANTS BY STATE OF ORIGIN

			Out-Migrants' Skills Relative to State Mean			
State at Age 14	N	Rate	AFQT	Education	Unstand. Wage	Stand. Wage
Alabama	246	.191	0.089	0.071	-.011	-.100
Alaska	19	.158	-.258	0.807	-.035	-.348
Arizona	70	.257	0.197	0.017	0.258	0.374
Arkansas	75	.280	-.058	-.387	0.079	0.126
California	661	.097	-.013	-.049	-.116	-.079
Colorado	122	.254	0.120	0.233	-.152	-.006
Connecticut	119	.193	0.180	0.841	0.163	0.145
Delaware	4	.750	0.053	0.000	0.503	0.307
D.C.	49	.347	0.249	0.571	0.290	0.321
Florida	240	.204	0.151	-.058	0.020	0.085
Georgia	283	.113	0.264	0.386	0.045	0.474
Hawaii	4	.750	-.135	0.000	-.231	0.016
Idaho	4	.500	-.503	-1.500	-.156	-.096
Illinois	195	.221	0.376	0.678	0.160	0.116
Indiana	118	.271	0.298	1.092	0.250	0.237
Iowa	58	.259	0.205	0.545	0.081	0.346
Kansas	35	.371	0.293	0.202	0.051	0.255
Kentucky	23	.304	-.442	-1.161	-.293	0.036
Louisiana	50	.180	0.523	0.284	0.216	0.550
Maine	4	1.000	0.000	0.000	0.000	0.000
Maryland	67	.313	0.007	-.871	0.054	-.075
Massachusetts	113	.124	0.434	1.303	0.180	0.389
Michigan	318	.255	0.161	0.465	-.007	0.098
Minnesota	148	.223	-.111	-.181	-.205	-.307
Mississippi	68	.265	0.285	0.417	-.054	0.266
Missouri	177	.254	0.074	0.021	-.111	-.040
Montana	46	.370	0.107	0.252	-.115	-.273
Nebraska	29	.448	0.109	0.430	0.092	0.091
Nevada	4	1.000	0.000	0.000	0.000	0.000
New Hampshire	6	.333	0.048	-1.500	0.117	0.017
New Jersey	258	.233	0.370	0.521	0.079	-.112
New Mexico	74	.189	0.247	0.759	-.041	0.287
New York	412	.201	0.122	-.030	0.009	-.038
North Carolina	264	.167	0.159	1.000	-.073	0.167
North Dakota	8	1.000	0.000	0.000	0.000	0.000
Ohio	411	.180	0.088	0.207	0.081	0.131
Oklahoma	103	.155	-.245	0.919	-.101	-.209
Oregon	24	.292	-.197	0.327	-.016	0.080
Pennsylvania	309	.133	0.122	0.670	0.071	0.060
Rhode Island	2	1.000	0.000	0.000	0.000	0.000
South Carolina	171	.082	0.394	1.091	0.306	0.609
South Dakota	18	.500	-.101	-.389	0.263	0.223
Tennessee	140	.136	-.001	-.062	0.034	0.391
Texas	456	.096	0.307	0.196	0.060	0.023
Utah	15	.200	-.062	0.800	-.222	0.368
Vermont	32	.063	0.307	1.375	0.256	0.111
Virginia	160	.138	0.383	0.675	0.102	0.040
Washington	84	.214	0.234	0.040	-.134	-.139
West Virginia	114	.228	-.212	-.162	-.143	-.172
Wisconsin	256	.125	0.219	0.355	0.114	0.025
Wyoming	0	--	--	--	--	--
TOTAL	6666	.183	0.144	0.282	0.027	0.062

TABLE 5

IN-MIGRATION RATES AND AVERAGE SKILLS OF IMMIGRANTS BY STATE

State in 1986	N	Rate	In-Migrants' Skills Relative to State Mean			
			AFQT	Education	Unstand. Wage	Stand. Wage
Alabama	216	.078	-.014	-.557	0.053	0.289
Alaska	30	.467	0.617	0.917	-.180	0.023
Arizona	81	.358	0.170	-.067	0.059	-.069
Arkansas	63	.143	0.104	0.375	0.216	0.103
California	709	.158	0.145	0.385	0.070	0.167
Colorado	135	.326	0.462	0.317	-.038	0.085
Connecticut	112	.143	0.429	1.177	0.467	0.539
Delaware	1	0.000	0.000	0.000	0.000	0.000
D.C.	34	.059	-.072	1.910	0.486	0.178
Florida	314	.392	0.115	-.165	0.026	0.088
Georgia	309	.188	0.327	0.991	0.316	0.180
Hawaii	7	.857	0.188	-.758	-.404	-.961
Idaho	3	.333	-.593	0.042	-.329	-.937
Illinois	190	.200	-.095	0.214	0.026	0.128
Indiana	101	.149	0.103	-.502	0.026	0.128
Iowa	49	.122	-.222	-.765	-.080	-.332
Kansas	45	.511	0.271	0.971	0.074	-.101
Kentucky	33	.515	0.112	0.783	0.045	-.057
Louisiana	66	.379	0.334	0.353	0.260	0.371
Maine	3	1.000	0.973	1.576	-.097	-.893
Maryland	70	.343	0.117	0.560	0.270	0.456
Massachusetts	108	.083	0.094	0.600	0.247	0.035
Michigan	258	.081	0.212	0.818	0.192	0.169
Minnesota	131	.122	-.001	0.466	0.078	0.236
Mississippi	65	.231	-.407	-.167	-.378	-.681
Missouri	162	.185	0.012	0.282	0.130	0.182
Montana	36	.194	0.537	0.664	-.350	0.023
Nebraska	20	.200	0.716	1.130	-.115	0.041
Nevada	14	1.000	0.251	0.599	-.240	-.026
New Hampshire	6	.333	0.227	0.890	0.125	0.478
New Jersey	230	.139	-.062	0.335	-.111	-.020
New Mexico	67	.104	0.553	0.687	0.099	-.161
New York	386	.148	0.174	0.696	-.230	-.066
North Carolina	258	.147	0.020	-.216	0.007	-.038
North Dakota	1	1.000	0.154	0.370	-.039	0.446
Ohio	366	.079	-.114	-.412	-.477	-.284
Oklahoma	108	.194	0.671	1.157	-.185	-.008
Oregon	28	.393	0.417	1.436	0.063	0.506
Pennsylvania	314	.146	0.157	0.246	-.203	-.089
Rhode Island	2	1.000	0.809	-.191	0.919	1.169
South Carolina	176	.108	0.418	0.323	0.344	0.272
South Dakota	17	.471	-.041	0.380	0.095	-.235
Tennessee	135	.104	-.048	-1.186	0.511	0.678
Texas	534	.228	0.051	0.201	0.075	0.111
Utah	17	.294	0.548	-.164	0.188	0.698
Vermont	34	.118	-.195	-1.181	-.600	-.509
Virginia	182	.242	-.215	-.055	-.024	-.040
Washington	89	.258	0.251	0.343	-.176	-.169
West Virginia	97	.093	0.145	-.073	0.620	0.672
Wisconsin	249	.100	0.302	0.316	-.300	-.008
Wyoming	5	1.000	0.024	0.479	-.008	0.523
TOTAL	6666	.183	0.144	0.282	0.027	0.062

TABLE 6

The Effect of Skills on Migration Rates
in Mover/Non-Mover Probit Models
(t-statistics in parentheses)

Skill Measure	Unstandardized Dispersion in Earnings in State at Age 14: First Quartile	Second Quartile	Third Quartile	Fourth Quartile
AFQT Score	.037 (3.32)	.059 (4.42)	.020 (1.60)	.062 (1.11)
Effect of One S.D. Change in AFQT on Migration Rate	.034	.056	.018	.012
Education	.092 (5.30)	.060 ((3.50)	.008 (0.45)	.018 (0.83)
Effect of One S.D. Change in Educ on Migration Rate	.053	.036	.004	.008
Unstandardized Wage	.095 (1.99)	.035 (0.66)	.022 (0.41)	.008 (0.13)
Effect of One S.D. Change in Un. Wage on Migration Rate	.021	.008	.005	.001
Standardized Wage	.099 (2.37)	.067 (1.51)	.038 (0.83)	.076 (1.41)
Effect of One S.D. Change in St. Wage on Migration Rate	.025	.019	.010	.016

NOTE: The Probit models also include race and sex dummy variables.

TABLE 7

THE EFFECT OF SKILLS ON THE CHOICE OF DESTINATION
(t-statistics in parentheses)

Dependent Variable: Difference in Dispersion of Earnings between State in 1986 and State at Age 14

	Entire Sample		Movers Only	
Skill Measure	Change in Unstand. Dispersion	Change in Stand. Dispersion	Change in Unstand. Dispersion	Change in Stand. Dispersion
AFQT Scores	.0012	.0017	.0042	.0058
	(4.52)	(4.88)	(3.08)	(3.45)
Education	.0005	.0006	.0014	.0018
	(4.41)	(4.27)	(2.81)	(2.68)
Unstandardized Wage	.0009	.0009	.0038	.0037
	(2.62)	(2.09)	(2.34)	(1.85)
Standardized Wage	.0007	.0007	.0029	.0027
	(2.63)	(2.09)	(2.06)	(1.55)

NOTE: The regressions also include race and sex dummy variables.

TABLE 8

THE EFFECT OF SKILLS ON THE CHOICE OF DESTINATION (PROBIT MODELS)
(t-statistics in parentheses)

Dependent Variable: Wage Dispersion Increased Between State of Residence at Age 14 and State of Residence in 1986

	Movers Only	
Skill Measure	Change in Unstandardized Dispersion	Change in Standardized Dispersion
AFQT Score	.075 (1.69)	.100 (2.23)
Education	.038 (2.21)	.042 (2.45)
Unstandardized Wage	.216 (3.98)	.113 (2.13)
Standardized Wage	.135 (2.86)	.085 (1.82)

NOTE: These Probit models also include race and sex dummy variables.

TABLE 9

CORRELATION BETWEEN SKILLS AND WAGE DISPERSION IN STATE AT AGE 14
(t-statistics in parentheses)

	Unstandardized Dispersion	Standardized Dispersion
AFQT Score	-.223 (-17.1)	-.229 (-17.9)
Education	-.088 (-6.40)	-.103 (-7.24)
Unstandardized Wage	-.033 (-2.36)	-.037 (-2.64)
Standardized Wage	-.087 (-6.03)	-.108 (-7.52)

TABLE 10

THE EQUILIBRIUM SORTING OF SKILLS ACROSS STATES
(t-statistics in parentheses)

Ordered Probit Dependent Variable: Quartile of State in 1986, where states are ordered by their dispersion in earnings

Skill Measure	Unstandardized Dispersion	Standardized Dispersion	Unstandardized Dispersion	Standardized Dispersion
AFQT Scores Differenced from State Mean	.1641 (9.58)	.0454 (6.88)		
AFQT Scores			-.1131 (-6.75)	-.0931 (-5.56)
Education Differenced from State Mean	.0294 (4.18)	.0179 (2.61)		
Education			-.0209 (-3.09)	-.0124 (-1.80)
Unstandardized Wage Differenced from State Mean	.0728 (3.70)	.0230 (2.99)		
Unstandardized Wage			.0277 (1.41)	.0392 (2.01)
Standardized Wage Differenced from State Mean	.0526 (3.08)	.0107 (1.60)		
Standardized Wage			-.0179 (-2.85)	-.0312 (-1.94)

NOTE: These coefficients are obtained from ordered probit models that also include race and sex dummy variables.