1994

Determinants of interstate migration in the United States: A search theory approach

Tubagus Feridhanusetyawan
Iowa State University

Order Number 9518376

Determinants of interstate migration in the United States: A search theory approach

Feridhanusetyawan, Tubagus, Ph.D.

Iowa State University, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Determinants of interstate migration in the United States: A search theory approach

by

Tubagus Feridhanusetyawan

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Economics
Major: Economics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1994

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to express my gratitude to my major professor, Dr. Wallace E. Huffman, for his guidance, encouragement, and financial support for my study at Iowa State. I extend my appreciation to the rest of my committee, Dr. Peter Orazem, Dr. J. Peter Mattila, Dr. Walter Enders, and Dr. Cynthia Fletcher, for their many valuable comments and suggestions at various stages. I would also to acknowledge Dr. M. Hadi Soesastro of the Centre for Strategic and International Studies, Jakarta, Indonesia, for his support of my graduate study.

I would like to remember my father, M. Wasito, who passed away in Indonesia when I was in the early part of my Ph.D. program. I would like to express my appreciation to my mother, Soemiwati, who has always believed in the importance of higher education. My greatest appreciation goes to my wonderful wife, Aurelia Dewi Hajati, who has shown me what love is all about. Finally, I would like to dedicate this dissertation to my wife and to my son, Ariel Budi Prabawa, who was born in Ames during the final stage of my study at Iowa State.

ABSTRACT

The study develops and fits two economic models of migration decisions: a commonly used point-in-time discrete choice model and a new model based on the search theory. The search model for migration views migration as a series of sequential decisions made over time under uncertainty. It captures repeated migrations and better fits the life-cycle aspects of migration decisions. The model predicts the relationship between the duration of stay, the hazard rate, and the determinants of migration. The study also develops a new approach for measuring an individual's wage performance as one determinant of migration.

The models are fitted empirically to the micro-panel data from the Panel Study of Income Dynamics (PSID). The study follows the interstate migration behavior of 915 male heads of households, who were age 19 to 45 in 1968, from 1968 to 1987. First, wage equations are fitted to create an empirical measure of each individual's wage performance. Then, based on the point-in-time model, the study fits the probability of migration using a standard probit model. Finally, from the search model of migration, the hazard rate for migration by using both a constant and a time-dependent hazard model is estimated.

The results show that better local wage performance, being married, having school-age children, being white, being a union member, and being-self employed or a farmer reduce the tendency to migrate. On the other hand, having more education, being young, and being unemployed increase the chances that an individual will migrate. With the increase in the duration of stay, the hazard rate for migration increases in the first four to six years and then decreases over time.

CHAPTER 1. INTRODUCTION

Pattern of Internal Migration in the United States

The United States has been characterized by a geographically mobile population. Historically, the U.S. population has shifted to the West. Between the Civil War and the 1950s, there was a secondary shift from the South to the industrialized cities in the North. Starting in the 1960s, however, the pattern changed so that in the 1970s the South began to have a net in-migration. There has also been a large movement from the inner city to suburban areas during the last 20 years. Workers have sought new or better jobs in different localities and have decided to live in the best place they can. Many factors ranging from regional economic growth to personal factors such as family-related reasons have contributed to this phenomenon.

The Current Population Reports (U.S. Department of Commerce, 1989) illustrates the pattern of geographical mobility in the United States from the 1960s to the late 1980s. The survey consists of data from 200,000 to 250,000 individuals in the United States. The data on mobility status were collected by cross-sectionally comparing the each person's current residence with their residence one year earlier. Table 1.1 shows the annual rates of geographic mobility, calculated as a percentage of the population who moved or resided in different places in two consecutive years from 1960 to 1990. The rate of total movers ranged from 17.0 to 20.6 percent annually for the period 1960 to 1990. Most of the movements were moves between houses in the same county; such moves accounted for between 10.3 percent to 13.7 percent of the sample. The proportion of moves between counties in the same state has been around 3.0 percent, similar to the proportion of moves between states. Moves between regions have ranged from 1.5 percent to 2.0 percent, while international migration has been less than 1.0 percent.

The dynamics of local migration has dominated total migration for the last 20 years. Most of the change in mobility rates during the last 20 years has been in the rate at which people made local moves. Because of the difficulty in tabulating moves within a local labor market using the data, the CPS report defines a local move as a move within the same county. Moves between counties in the same state and between states and the international migration are defined as long-distance migration. The rate of residential mobility fell from an average of about 20.0 percent annually in the 1960s to 16.6 percent in 1983 and then climbed to 20.2 percent in 1985 before falling again to around 17.0 percent in the late 1980s. The increase in the total annual rate of moving in the mid-1980s was primarily due to the dynamics of local moves. The rates at which people moved longer distances, namely moves between counties or states, have not changed for the last 20 years.

Table 1.1. Annual geographic mobility rates, by type of movement: 1960-87 (percent)

	1960-61	1970-71	1980-81	1990-91
Total movers	20.6	18.7	17.2	17.0
Same county	13.7	11.4	10.4	10.3
Different county				
- same state	3.1	3.1	3.4	3.2
- different state	3.2	3.4	2.8	2.9
- same region	1.5	1.4	1.3	1.5
- different region	1.7	2.0	1.5	1.4
International	0.6	0.8	0.6	0.6

Sources: U.S. Department of Commerce, Current Population Reports (1989), and U.S. Department of Commerce, Statistical Abstract of the United States (1968-93).

Regional migration patterns are important indicators of population redistribution. In-migration, out-migration, and net-migration for regions in the United States from 1965 to 1987 are presented in Table 1.2. In general, people have moved from the Northeast and the North Central regions to the South and the West for the last 25 years. According to Greenwood (1985), this pattern emerged because employment opportunities in the North and East have been lagging since the 1970s due to recessions in the mid-1970s and early 1980s. The CPS report (U.S. Department of Commerce, 1989) also indicates several factors contributing to this pattern. First, the South has more attractive tax incentives and cheaper non-unionized labor. Second, the rise of light industries such as electronics increased the demand for trucking rather than rail transportation. Third, the spread of air conditioning and the leveling of regional differences in standards of

Table 1.2. Regional in-migration, out-migration, and net-migration in the United States: 1965-90 (thousands)

Area	1965-70	1970-75	1975-80	1980-85	1985-90
Northeast					
In-migrants	1,273	1,057	1,106	2,345	2,179
Out-migrants	1,988	2,399	2,592	3,285	3,636
Net-migrants	-715	-1,342	-1,486	-940	-1,456
North Central					
In-migrants	2,024	1,731	1,993	3,766	4,295
Out-migrants	2,661	2,926	3,166	5,321	4,521
Net-migrants	-637	-1,195	-1,183	-1,555	-226
South					
In-migrants	3,142	4,082	4,204	6,798	6,839
Out-migrants	2,486	2,253	2,440	5,017	5,630
Net-migrants	656	1,829	1,764	1,781	1,209
West					
In-migrants	2,309	2,347	2,838	4,510	4,208
Out-migrants	1,613	1,639	1,945	3,795	3,603
Net-migrants	696	708	893	715	605

Sources: Greenwood (1985) and U.S. Department of Commerce, Statistical Abstract of the United States (1985-93).

living and educational opportunities, combined with the success of the civil rights movement, have influenced migration patterns.

When an economic agent migrates to follow economic opportunities, he/she will change at least marginally the economic opportunities of others in the new place. For example, a study by Greenwood and Hunt (1984) shows that the in-migration of one individual to a southern state was associated with 1.3 additional jobs, while in-migration to a western states created 0.4 additional jobs. The important point to make is that the migration does not necessarily reduce the employment opportunities of native workers.

Another important pattern in residential mobility and population redistribution in the United States has been the process of sub-urbanization. For many decades, increased farm mechanization and urbanization of economic opportunity have led to rural-to-urban migration. Since the 1970s, however, the pattern has changed. The percentage of the U.S. population residing in metropolitan areas has declined, with a corresponding change in the direction of migration. The definition of metropolitan areas had to be adjusted due to the rapid growth of areas surrounding metropolitan areas. In fact, redefinition of metropolitan statistical areas occurred in 1983. The 1989 CPS report (U.S. Department of Commerce, 1989) shows large movement from the central city to its suburbs.

According to Chalmers and Greenwood (1980), several factors have contributed to this rural-urban migration turnaround. First, the relative cost of doing business in older urban centers has been changing and the number of resource-based industries in nonmetropolitan areas has been growing. Second, rising income and wealth are increasing the demand for location-specific amenities. Third, changes in the demographic structure of the population, the labor forces and the government policies have affected migration patterns.

Literature Review

Many studies have tried to explain why people move geographically. Migration studies have maintained a strong orientation toward determinants as opposed to the consequences of migration. The development of empirical migration research has been facilitated not only by theoretical innovations but also by the availability of new micro data sets. For the past 20 years, the study of the determinants of migration has been based on point-in-time utility maximization. The empirical models used standard univariate discrete choice models such as probit and logit. Generally, the studies put the incidence of migration on the left-hand side of the equation and the variables explaining the probability of migration including individual and regional characteristics on the right-hand side. The literature review presented in this section is not exhaustive. The discussion in each chapter includes citations from related references. The first part of this review presents an overview of previous empirical studies on the migration decision. The second part presents some previous empirical applications of search theory.

Previous studies on the determinants of migration

The focus of migration studies has been in finding the determinants rather than the consequences of migration. The use of binary discrete choice models and micro-panel data have been extensive (Greenwood, 1975, 1985). Stark and Bloom (1985) reported that empirical research on labor migration has benefitted from the development of econometric techniques such as analysis for qualitative dependent variables, the model for selection bias correction, and techniques for analyzing longitudinal micro data. The importance of micro-panel data in migration studies has been recognized. The concern is that previous empirical studies of migration have not utilized micro-panel data over a long enough period to capture sequential migration decisions or repeated migration.

Selected studies on the determinants of migration for the last 20 years are presented in Table 1.3. Most of these studies used micro-panel data and a binary variable as the dependent variable. Studies 1 and 4 used aggregate data. Previous studies commonly focused on a specific determinant of migration, such as education and age, personal unemployment, or marital status, rather than on a broad range of possible determinants. Education and age have consistently been included in the migration model. When the unit of observation is the household, marital status and number of children are included in the model. Some studies have focused only on the effects of unemployment-related variables.

Table 1.3. Determinants of migration in previous studies[a]

Study	Education	Age	Race	Unemployment		Marital Status	Number of Children	Income or Wage	Climate
				Personal	Local				
1. Schwartz (1976)	X	X	-	-	-	-	-	-	-
2. Mincer (1978)	X	X	-	-	-	X	X	-	-
3. Da Vanzo (1978)	-	-	-	X	X	-	-	-	-
4. Graves (1979)	-	-	-	-	X	-	-	-	X
5. Graves and Linneman (1979)	X	X	X	X	-	X	X	X	-
6. Schlottmann and Herzog (1981)	X	X	-	X	-	-	-	-	X
7. Nakosteen and Zimmer (1980)	-	X	X	-	X	-	-	-	-
8. Goss and Schoening (1984)	X	X	-	X	-	-	-	X	-
9. Herzog and Schlottmann (1984)	X	X	-	X	X	X	-	-	-
10. Goss and Paul (1986)	X	X	X	-	-	-	-	X	-
11. Pissarides and Wadsworth (1989)	X	X	-	X	X	X	X	-	-
12. Goss and Paul (1990)	X	X	X	X	X	-	-	X	-

[a]The sign 'X' means that the study has incorporated the effect of the corresponding variable on migration decisions.

Most studies have shown the importance of education and age on the migration decision. All the studies in Table 1.3 show a positive effect of education and a negative effect of age on the migration decision. Both migration and education represent human capital investments. Migration and education might be complementary because investing in one enhances the returns to investing in the other. Young adults are more likely to move because they have the longest expected time over which to benefit from migration.

Using a binary discrete choice model, Schwartz (1976) reported the interaction between migration, education, and age. The study concluded that the shape of the earning-age function and the way education affects it could explain the empirically observed relationship of migration measures to distance moved, age, and education. The fact that there is age selectivity of migrants is also supported by Schlottmann and Herzog (1981) and an early study by Sjaastad (1962).

Models of migration behavior have commonly emphasized individual rather than family utility maximization. Mincer (1978) applied a point-in-time discrete choice model to show that family ties tend to deter migration, reduce the labor force participation and earning level of wives, and increase the employment and earning level of husbands in migrating families. Mincer found that two-earner families have stronger ties to any resident location, which deters migration. Since this study, every study on migration has incorporated family-related variables such as marital status and the number of children in the family into the model. Generally, the results have been consistent in across studies: being married and having children reduces the probability of migration.

Graves and Linneman (1979) developed a consumption-based model of the migration decision. Migration behavior was derived by classifying the goods consumed into those which are traded and nontraded. Graves and Linneman found that the changing demand for nontraded goods led to an incentive to migrate. Therefore, the probability of migrating is positively related to any variable that causes a change in the demand for nontraded goods. The empirical model is a standard probit procedure. In terms of family-related variables, Graves and Linneman showed that the existence of school-age children in the family and being married tend to deter migration. Similar results were found by Pissarides and Wadsworth (1989).

Nakosteen and Zimmer (1980) developed a standard discrete choice model of migration based on inter-location income differentials rather than on utility differentials. The objective of their research was to obtain good estimates of earning functions for migrants and nonmigrants by including selection correction in their empirical model. A standard probit procedure was used to estimate the probability of migration and to correct the earning functions. Nakosteen and Zimmer found that higher resident state employment and per capita income growth significantly reduced the probability of migration. Among mutually exclusive personal characteristics, they found that being white, young, and male increased the probability of migration.

The tendency for unemployed people to migrate has been supported by many studies. Da Vanzo (1978) showed that families whose heads were unemployed or dissatisfied with the job were indeed more likely to migrate. Therefore, unemployed persons or persons who are actively searching for a job have a higher probability to migrate. Graves and Linneman (1979), Goss and Schoening (1984), and Pissarides and Wadsworth (1989) found a similar positive effect of personal unemployment on migration.

Schlottmann and Herzog (1981) examined the demographic and socio-economic determinants of migration for employed and unemployed persons. They used binary logit models to explain the probability of migration and have found that, in general, for individuals at risk to migrate, age and education selectivity of migration were confirmed. However, for unemployed primary migrants, education selectivity was not observed. The provision of welfare services was found to have little or no impact on the migration decision of the unemployed.

Previous studies have obtained conflicting results for the direct effects of local labor market variables, especially the local unemployment rate, on the probability of migration. One possible reason for such conflict suggested by Greenwood (1975) is the simultaneity inherent in studies that use actual unemployment data. Unemployment is often measured at the end of the migration period and hence may have been affected by the migration itself. Another possible reason is that the actual and anticipated unemployment rates differ so that the actual unemployment rate is a mixed variable. By using aggregate employed and unemployed workers, Da Vanzo (1978) found an insignificant relationship between unemployment rate and migration. Herzog and Schlottmann (1984) found a positive and significant effect of the residents area unemployment rate on migration. However, a later study by Goss and Paul (1990) argued that the effects of the unemployment rate on migration were over-estimated because previous studies have ignored unemployment insurance benefits. Pissarides and Wadsworth (1989) found a positive and significant effect of personal unemployment on migration. The effect of the local unemployment rate, however, was not found to be statistically significant. They found only that at higher overall unemployment rates, migration propensities are reduced.

Another unclear result is the effect of wage rate on migration. Studies by Goss and Schoening (1984), and Goss and Paul (1986, 1990) found no significant effect of an individual's wage on his probability of migrating. Graves and Linneman (1979) included the absolute change in wages between years. Similarly, they did not find any significant effect of wages on migration. One simple reason for this result is the correlation between wages and other variables that affect migration, such as education and age. Other studies (see Table 1.3) did not include the actual wage in the equation, but incorporated the determinants of wage in the reduced form of the migration equation. The unclear effects of the wage on migration is one problem that this study is trying to solve.

The effect of resident climatic conditions on net-migration rates was studied by Graves (1979). He used the Standard Metropolitan Statistical Area (SMSA) as the geographical unit, and regressed net migration rates on local median income, unemployment, temperature, humidity, and average wind speed. The significant effect of the climatic variables showed that the omission of these variables could lead to biased estimates of other parameters. Because Graves used aggregate data, namely, the number of people moving in one SMSA, the study explained the number of people migrating rather than individual migration.

Search theory

Job search theory has been developed in the studies of unemployment and the working decision. Compared to the standard deterministic labor force participation model, job search theory treats uncertainty in the labor market explicitly. The idea is that unemployment has an investment aspect. The unemployed person makes a decision to work by responding to a stochastic wage offer. Stigler (1961, 1962) developed the first mathematical analysis of optimal search strategy of the unemployed worker. The model optimized the individual's utility subject to information constraints and other constraints.

Lippman and McCall (1979) also developed a standard job search model. Using the premise that the objective of an individual is to maximize expected income net of search costs, they classified all possible job offers into two mutually exclusive classes: acceptable and unacceptable. These classes were separated by a reservation wage. Using a similar framework, an optimal search model was developed by Devine and Kiefer

(1991). In addition to the standard model by Lippman and McCall, Devine and Kiefer derived an explicit specification of the probability of accepting an offer. Furthermore, they extended the model by incorporating the possibility for layoffs and quits. They also extended the model to cover the possibility that an individual may exit from the labor force.

Schaeffer (1985) developed a dynamic optimization model to analyze the effect of human capital on migration. He demonstrated that a dynamic model which allows for a sequence of possible moves in a individual's lifetime is superior to a static model, which views migration as one-time move. His model incorporated the accumulation of human capital, rather than a stock of human capital, on sequential migration decisions. The study predicted a positive effect of human capital accumulation and job mobility on migration. Because of the close relationship between migration and job mobility, Schaeffer suggested that the migration model should incorporate the job search. Even though his model is appealing, the empirical study based on his dynamic optimization model is cumbersome.

In general, the empirical search model can be classified into two groups: wage data based or duration data based. Because duration data are more readily available, progress has been made in developing econometric techniques for duration data. There are two common procedures: standard linear regression and the hazard function. The hazard function estimation is useful because it allows for more flexibility in specifying the duration distribution. In addition, censoring can be dealt with easily through appropriate specification of the likelihood function (Devine and Kiefer, 1991). Most studies using hazard functions incorporate explanatory variables in the model by specifying a proportional hazard. This method is simple to interpret because the effect of the explanatory variable is obtained by multiply the hazard function by a scaling factor. Another common specification is the accelerated lifetime model where the effect of an explanatory variable is to re-scale the time unit directly.

Hazard functions and statistical analysis of waiting time data are closely related. A book by Klabfleish and Prentice (1980) has been the source of much of the analysis. In general, there are two types of estimation procedures: non parametric and parametric. The differences lie in the assumptions. Kiefer (1988) summarized the nonparametric and parametric procedures needed to estimate a survivor function and a hazard function. Cox (1975) developed a partial-likelihood approach to the proportional hazard function. This approach is called a semi-parametric method because it requires a basic form for the hazard function but leaves the underlying distribution undefined. Economic theory can provide a rationale for an assumption about the duration distribution. For example, a constant reservation wage policy would lead to a constant hazard rate over time and an exponentially distributed duration. Finally, Greene (1991) developed a parametric estimation of the hazard function. The procedure follows a standard maximum likelihood estimation where the likelihood function is built upon the underlying distribution of the duration.

Even though search theory has been extensively used and developed in the area of unemployment studies, its application to migration has been relatively recent. Goss and Schoening (1984) incorporated job search time on the migration decision; however, their econometric model remained the point in time standard discrete choice model. By using a binary logit model, they found a negative effect of job search time on the probability of migration. Kiefer (1988) and Devine and Kiefer (1991) said that one potential application of search theory is in the area of migration or geographic mobility. However, published material that specifically applies search theory to migration is hard to find. One possible reason is that longitudinal data collected long enough to cover the migration history of an individual were not available until the late 1980s.

Statement of the Problems

The first problem is the use of a point-in-time decision model for migration. Migration is a long-term decision made over time. Hence, migration is a part of life-cycle decisions. Previous studies have used standard point-in-time discrete choice models. These models take a "short-term" perspective to the migration decision and do not capture the long-term, or life-cycle, perspective. A thorough treatment of life-cycle effects requires that migration be studied as an event that occurs in continuous time. A more realistic and richer specification of models of migration may be possible, and they can be fitted to micro-panel data. A study using panel data over a longer time period can capture the life-cycle characteristics of migration decisions.

Given that econometric models can only accommodate data for discrete time periods, a key issue is how to extend econometric models for multi-period or lifetime analysis of individual migration behavior. One possible extension is a multi-period discrete choice model. Assume that an individual lives in T periods and makes migration decisions in every period. The relevant issue is the probability of an individual migrating during any one of the T periods. For example, define dummy variable $M_t = 1$ if the individual migrates during period t, and $M_t = 0$ if otherwise. The probability that someone migrates at only period 3 can be expressed as $Pr(M_1 = 0, M_2 = 0, M_3 = 1, M_4 = 0, M_5 = 0, \ldots, M_T = 0)$. Under this binary choice framework, the events will have 2^T possibilities. With a normality assumption for the disturbance in the decision function, for example, this model is represented as a multi-period probit model.

Some difficulties exist in dealing with multi-period probit models. First, empirical estimation of this model is cumbersome. Several restrictions are needed to estimate the model (Avery, Hansen, and Hotz, 1983; Chamberlain, 1980, 1985b). In addition, conventional discrete choice models, such as probit or logit models, when defined for one time interval, are of different functional form when applied to another time unit. Second, in most economic models, there is no natural time unit within which agents maximize their utility, make their decisions, and take action. It is natural and analytically convenient to characterize the agent's decisions and actions in continuous time. Even if there is a decision period, there is no reason to suspect that it corresponds to the annual or quarterly data that are typically available (Heckman and Singer, 1985). Econometric estimation, however, is difficult for a continuous time model.

An alternative approach is a waiting time model. Search theory can be applied to explain an individual's time duration of residency in one place using micro-panel data. The use of search theory enables us to capture life-cycle aspects of migration decisions without the complications of the multi-period discrete choice models. In addition, search theory can explicitly incorporate uncertainty about the future in the model. Statistical analysis can then be applied to duration data.

To assess the expected gain from job or residential mobility, an individual must gather information about wage offers, quality of life, local amenities, and other nonmonetary benefits. An individual has to make a personal assessment of advantages and disadvantages with the information available. When a decision is based on incomplete information, expectations about the quality of life at a new residence may not be met, and this might lead to yet another move. The scenario fits well to what search theory provides because search theory presents sequential decision making in which uncertainty of future outcomes is treated explicitly.

The second problem is related to the stream of individual income. Previous studies have not found any clear effect of an individual's wage or income on migration. Some studies did not include the individual's actual income in the model of migration decisions. They commonly specified the reduced form migration equation, which consists of some variables determining an individual's income. There are at least two reasons for this approach. First, an individual's income has a strong correlation with his/her personal attributes affecting migration, especially education and age. Second, the use of income from one or two years of data does not capture the projected future streams. When a study included actual income or wage in the migration model, the result was not significant.

These problems might contribute to the conflicting results related to local labor market variables, such as the effect of the local unemployment rate on migration. For example, it has been predicted that an individual tends to leave a state with a high unemployment rate. When the individual is employed and his/her wage is positively correlated with the local unemployment rate, however, he/she might have higher expected wages by not moving. Therefore, incorporating the direct effect of local labor market variables on the decision to migrate might be inappropriate. In fact, a study by Da Vanzo (1978) has shown that the current local unemployment rate matters only for those who are unemployed. Other studies have not found any significant effect of local labor market variables. One reason is that an individual's utility depends on individual wage performance for given local labor market conditions, rather than directly depending on the local labor market situation itself.

What seems to be needed is a new measure of an individual's wage performance which has no or little correlation with other personal characteristics directly affecting migration. One possible measure is the residual of the wage equation, calculated as the actual wage minus the predicted wage of an individual. The predicted wage could be obtained from fitting the well-established wage equation (Topel, 1986; Tokle and Huffman, 1991). Because the wage equation is also a function of local labor market characteristics, using the residual wage might also resolve the issue of how local labor market variables affect the migration decision. The effect of local labor market variables on individual migration decisions could be traced through their effects on an individual's potential wage. First, combined with other personal characteristics, local labor market conditions will determine potential or predicted wages in the area where the person lives. Then, the realization of the individual's actual wage relative to the potential wage, which is calculated as the residual wage, may be related to the decision to move or to stay. In this case, the individual's utility and the decision to migrate is not directly a function of local labor market conditions. Therefore, the effect of local labor market conditions on migration decisions might differ across individuals depending on the effect of an individual's wage performance. This approach seems promising because it captures the individual's specific perception of how local economic conditions affect his/her utility and the migration decision.

The last problem is the effect of local amenities. The utility of being in one state is a function of local amenities, and the decision to move or to stay, therefore, should depend on local amenities. A model of migration decisions that does not include local amenities might lead to biased estimates. This study incorporates some measures of local amenities, such as the crime rate, state and national park area, and weather-related variables.

Purpose of the Study

The purpose of this study is to provide further theoretical and empirical analysis of interstate migration. Local migration (i.e., movement within a state) is excluded from the study for several reasons. First, it is hard to distinguish migration between counties within one state from local migration which does not seem very interesting. Second, it is difficult to find information on local or regional characteristics such as job growth and unemployment rates for the county level. Third, state units are the smallest common geographical area for labor laws, such as right to work, unemployment benefits, welfare benefits, etc.

This study develops economic models representing an individual's migration decisions and fits the model using micro-panel data for a sample of individuals who have been participating in a long-term survey. Two major models are pursued. First, a standard point-in-time discrete choice model of migration decisions is developed and fitted. The purpose of this study is to replicate the result from existing models using a new data set. The result of the study can be compared to that of previous migration studies that used a similar framework.

Second, the new approach, namely, incorporating search theory, will be used to derive a hazard rate for migration. Compared to previous studies on the determinants of migration using cross-sectional data, this study utilizes more information because the observations for individuals are taken from several consecutive periods of time. Because the length of time an individual lives in one state could be years, the hazard rate might change over time. Therefore, the study not only fits a constant hazard rate model but also a model with a time-dependent hazard rate. The results of this model can be compared to results from the previous model to investigate the contribution derived by using the new framework.

Objectives of the Study

Specifically, the objectives of the study are to:

1. Develop both a point-in-time discrete choice model and a dynamic search theory approach to human migration decisions. The first framework has been commonly used in the research of migration for the last 20 years. Although the second framework is common in unemployment studies, its application to migration studies is relatively new.

2. Fit a wage equation and derive annual potential wage rates for sample individuals. Then, the variable representing individual wage performance is developed. This variable is used in the equation explaining the probability of migration in the point-in-time discrete choice model and the hazard rate in the search model.
3. Perform an empirical analysis using a standard discrete choice model by fitting the equation representing the probability of migration. The result from this analysis can be compared with the result of previous migration studies, which largely used the same framework.
4. Perform an empirical study based on the search theory framework by fitting hazard functions for migration. Both constant and time-dependent hazard functions are estimated. The availability of Panel Study of Income Dynamics (PSID) data, which follows every individual from 1968 to 1988, will make it possible to estimate the model affecting individual migration decisions.

The research has important implications for reformulating the migration decision model for panel or longitudinal data sets. Compared to previous studies using a short-term view of migration, this study develops a model for a sequential migration decision over time and fits the model using 20 years of data. This study identifies variables that explain labor migration. These indicators will be useful for policy purposes dealing with labor mobility in the United States.

Organization of the Study

This study is organized as follows. The first chapter is the introduction, which consists of the literature review and the statement of the problems, purpose, and the objectives of the study. The second chapter presents the development of the economic model of migration decisions. This chapter consists of the point-in-time decision model and the application of job search theory on migration. Chapter 3 discusses the data in general. The chapter discusses the individual and state data, the pattern of migration, and the procedure used to prepare the data set for estimation. The fourth chapter presents the general econometric techniques used to fit the models based on the data. This discussion includes both the standard point-in-time discrete choice model and the search theory approach. Chapter 5 elaborates the detailed empirical specification and variable definitions corresponding to the econometric models discussed in Chapter 4. The sixth chapter presents the results of the fitted wage equation and the probability of migration based on the discrete choice model. The seventh chapter discusses the fitted hazard function for migration based on the search model using the constant and time-dependent hazard rate models. The final chapter summarizes and concludes the study.

CHAPTER 2. THE ECONOMIC MODEL

This chapter develops the economic model of the migration decision. The chapter consists of two major parts. The first part discusses the point-in-time migration decision which leads to a standard discrete choice model. This model has been commonly used in previous studies on migration. The second part develops the model for sequential migration decisions using search theory.

One Point-in-Time Decision Model

We assume that an individual lives for T periods ($t = 1, 2, 3, \ldots, T$). To fit the model of point-in-time decisions and to simplify the model, we use a discrete time model. At time $t = 1$, the individual is living at a particular location and enjoys utility represented by $Unm(1)$. At the end of period 1, the individual makes a decision about staying or migrating to a new place, called the destination. Once the individual decides to migrate, he/she resides in the new place until the end of his/her life. The migration decision is based on the present value of the expected future utility of being at the origin versus being at a (new) destination.

The expected utility for being at the origin (utility non-migrant, Unm) and the destination (utility migrant, Um) at time t are represented by $Unm(t)$ and $Um(t)$, respectively. To simplify the model, it is assumed that expected utility is constant over time so that $Unm(2) = Unm(3) = \ldots = Unm(T) = Unm$, and $Um(2) = Um(3) = \ldots = Um(T) = Um$. It is also assumed that the utility of being at the destination is net of the cost of migration. With the discount factor represented by r, the present value of the expected utility of being at the origin is:

$$Vnm = \sum_{t=2}^{T} \frac{Unm(t)}{(1 + r)^{t-1}} = \sum_{t=2}^{T} \frac{Unm}{(1 + r)^{t-1}} \qquad (2.1)$$

The present value of the expected utility of being at the destination is:

$$Vm = \sum_{t=2}^{T} \frac{Um(t)}{(1 + r)^{t-1}} = \sum_{t=2}^{T} \frac{Um}{(1 + r)^{t-1}} \qquad (2.2)$$

Given that there are D possible destinations, it is also assumed that Vm is the maximum value achievable among all possible destinations such that:

$$Vm = \text{Max } [Vm_1, Vm_2, Vm_3, \ldots, Vm_D] \qquad (2.3)$$

The individual makes a decision to migrate or to stay by comparing Vnm and Vm. We prefer this utility differential, rather than the income differential approach, because some factors affect utility directly rather than through income, (e.g., local amenities). This model is deterministic because an individual assesses future utility with certainty, and future utility is fixed. Information is complete and expectations are realized. The individual chooses to migrate if Vm > Vnm and to stay otherwise. Therefore, based on the utility assessment, in time t = 2 the individual lives in the best place possible. The decision can be summarized as:

$$\text{Decision} \begin{cases} \text{Migrate if } Vm > Vnm \\ \text{Do not migrate otherwise} \end{cases} \qquad (2.4)$$

The probability of migration can be expressed as:

$$\text{Prob(Migrate)} = \text{Prob}(Vm > Vnm) = \text{Prob}(Um > Unm)$$
$$\text{Prob(Stay)} = 1 - \text{Prob(Migrate)} \qquad (2.5)$$

The model in Equation (2.5) describes a migration decision made at one point in time at the end of period 1. This model translates the long-term expectation shown in Equation (2.1) and (2.2) into a two-period model, namely, time 1 and beyond. In other words, the model takes a short-term perspective to migration decisions. The model fits well in empirical migration studies based on data taken from a short period of time. When migration is viewed as a long-term decision made over time, an extension of the model is needed. One possible refinement is the application of job search theory.

Search Theory for Migration

Search theory frequently has been applied in studies of unemployment duration. An individual remains unemployed unless he receives wage offer that is higher than his/her reservation wage. In this migration study, we assume that an individual will move to a new location if the present value of his/her expected utility is greater in a new place than at the current residence location. Every individual is assumed to eventually migrate and to reside in a new place. Our interest, however, is in how long the individual stays in one place before moving.

In a search model, similar to a neoclassical framework, utility is maximized subject to constraints. However, the assumption of complete information is not required. The assumption of complete information is replaced by an assumption about expectations, usually rational expectations (Devine and Kiefer, 1991; Lippman and McCall, 1979). The basic feature is that in every period, every individual looks to improve his/her welfare by possibly moving. Some basic assumptions that build upon Devine and Kiefer are:

1. An individual maximizes the expected present value of utility over his/her expected finite future life span T, discounted to the present at a constant discount rate r.
2. The indirect utility flow while he/she stays at the origin is Unm (Utility nonmigrants) and, to simplify the derivation, is assumed to be constant over time.
3. During some time interval, the person may receive an opportunity (offer) to move to another place. The opportunities to move are received at a rate represented by a Poisson distribution with

parameter δ. The probability of receiving at least one offer within a short interval of length h is $\delta.h + O(h)$, where O(h) is the probability of receiving more than one offer in the interval and the limit of O(h)/h goes to zero as h goes to zero.

4. An opportunity to move is summarized by an indirect utility assessment, net of any cost of migrating, and equals Um (Utility migrants). If the offer is accepted, the individual will move from the origin to the destination and will enjoy utility represented by Um continuously over the period he/she stays in the new place.
5. Successive utility associated with offers (Um) received over the duration, or spell, of residence at the origin are independent realizations from a known distribution with finite mean (μ_{Um}) and variance σ_{Um}^2, cumulative distribution F(Um), and density f(Um).
6. Once rejected, an offer cannot be recalled. When a higher utility offer for a new destination is accepted, the individual moves to the destination and is assumed to stay there until he/she dies.

The assumption of lifetime utility maximization looks similar to the neoclassical utility-maximizing framework. The individual is essentially picking a location where he will maximize his full income, net of moving cost. Rather than focusing on utility, we are considering indirect utility where prices, full income, observed and unobserved family attributes explain utility. The continuous time and constant indirect utility flow assumptions make the derivation less complex. The search theory based migration model views migration as being caused by a push factor, i.e. a push to go to a better residence.

A finite life span of T is important for the time dependent hazard rate model. We will show later that the reservation utility increases as T decreases. In other words, the younger the individual, the smaller is his/her reservation utility, and the higher is his/her tendency to move. Similarly, the longer the individual stays in one place, the higher is his/her reservation utility, and the smaller is the tendency to move. A known utility offer distribution is important for empirical purposes. The assumption of a Poisson offer distribution is convenient and tractable. The implication is that the probability of an individual receiving an offer does not depend on the length of time spent residing at a place.

Derivation of the migration decision

Assuming that the (net) utility while at the origin is constant, offers are identically distributed, and the offer distribution and arrival rates are known and invariant, the value of utility while at the origin is defined implicitly as:

$$Vnm = \frac{Unmh}{(1 + rh)} + \frac{\delta h}{(1 + rh)} E[Max\{Vm(Um), Vnm\}] + \frac{(1 - \delta h)}{(1 + rh)} Vnm + O(h)k = Vm(Ur) \qquad (2.6)$$

where h represents a short period of time. The first term in the center portion of the equation is the discounted present value of net utility at the origin over a short time period h. The second term is the discounted expected value of utility when an offer is received with probability δh. Vm(Um) is the present value of utility when the offer is accepted and the individual migrates. The third term is the expected value of utility when no offer is received, which occurs with probability $(1 - \delta h)$. The last term captures the event when more than one offer is received, where the limit of O(h)/h goes to zero as h goes to zero and k is the value of following an optimal policy when more than one offer is received. At a maximum, Vnm in Equation (2.6) has to equal the value of the reservation utility Vm(Ur). The reservation utility policy is to accept Um and to migrate to a new place if Um > Ur, where Ur is the reservation utility.

Assuming that the agent lives over the period of time T in the future, the expected present value of Vm(Um) is:

$$Vm(Um) = Um \int_0^T Um\, e^{(-rt)}\, dt$$

$$= Um(1 - e^{-rT})/r \tag{2.7}$$

Therefore, the maximum for (2.6) while living at the origin is Vnm = Vm(Ur) = $Ur(1 - e^{-rT})/r$. Based on the derivation presented in Appendix A, Equation (2.6) can be simplified as:

$$Ur = Unm/(1 - e^{-rT}) + \delta/r\ [\int_{Ur}^{\infty} (Um - Ur)dF(Um)] \tag{2.8}$$

Define the probability that a person will migrate in a short period of time h as:

$$H.h = (\delta h + O(h)) \int_{Ur}^{\infty} f(Um)\ dUm \tag{2.9}$$

The first segment of the right-hand side of equation (2.9) represents the probability that the individual receives at least one offer to migrate. The second term represents the probability that the offer is accepted. Dividing by h and taking the limit as h -> 0, we get the hazard rate:

$$H = \delta \int_{Ur}^{\infty} f(Um)\ dUm \tag{2.10}$$

Equation (2.8) shows how the reservation utility changes as the explanatory variables change. By construction, the reservation utility is a function of the rate at which the offers are received (δ), the utility of being at the origin (Unm), discount rate (r), expected utility offers from a new destination (Unm), and expected life span (T). Equation (2.10) shows how the hazard rate is a function of Um, Ur, and δ. Given that all else is equal, a decrease in the reservation utility leads to an increase in the hazard rate. However, the reservation utility is also a function of Um and

δ, which means that in addition to their direct effect on the hazard rate, Um and δ have an indirect effect through the reservation utility. Because Unm, r, and T do not appear in Equation (2.10), their effects on the hazard rate are indirect through the change in the reservation utility.

Comparative static analysis could be undertaken to show how the changes in reservation utility (Ur) and hazard rate (H) are caused by changes in δ, Unm, r, the mean of utility offer μ_{Um}, and T. Table 2.1 shows the outcome of the comparative statics. Given the results in previous studies and the economic theory, some variables which have direct effects will be discussed.

Table 2.1. The effects of δ, Unm, r, μ_{Um}, and T on Ur and H[a]

Endogenous Variable	Exogenous Variable				
	δ	Unm	r	μ_{Um}	T
Reservation utility (Ur)	+	+	-	+	-
Hazard rate (H)	?	-	+	+	+

[a]Detailed derivations can be found in Appendix A.

An increase in the rate at which the offers are received (δ) increases the reservation utility. If δ increases, an individual will have greater confidence in having a better place to live in the future which will cause the individual to be more selective and to increase his/her reservation utility. The indirect effect on the hazard rate will in turn be negative. However, the total effect on the hazard rate turns out to be ambiguous. On one hand, with better information and more offers, the individual becomes more selective and his/her hazard rate will be lower. On the other hand, with more information and offers, the individual will have more chances to move and the hazard rate will be higher.

Human capital, such as education, that affects the chance of receiving an offer to move can be represented by the parameter δ. Higher education generally leads to more job offers and increases the hazard rate for moving. People with more education also face a larger labor market. In Becker's model of general and specific training (Becker, 1975), the conclusion is that general education increases a worker's productivity, not only in the current firm, but also in other firms/locations. Thus, the prediction is that people with more general education are more likely to move.

An increase in an individual's utility at the origin (Unm) will increase his/her reservation utility and decrease the hazard rate. The intuition behind this statement is clear. Better living conditions at the current location reduce the incentive to migrate.

Many variables may contribute to utility at the origin and can be measured and used in empirical studies. For example, an individual's wage at the origin will be important. Wage increases will increase the utility at the origin and reduce the tendency to move. However, the effect of the actual wage on migration could be mixed because the wage has a strong correlation with education and other personal characteristics that are likely to affect migration. Therefore, the individual's measurement of wage should be free from this multi-collinearity problem.

Other variables that may affect utility at the origin are personal unemployment, union membership, and local amenities. Unemployment experience may increases his hazard rate for moving. Membership in a labor union could increase job security and income and reduce the tendency to move. Local amenities, such as a low crime rate and a high percentage of state and national park in a state might affect his utility. High crime rates represents a disamenity, and the individual will most likely move away from these problems. State and national parks are positive local amenities and thus may reduce the hazard rate for moving.

The inclusion of finite life span T in the model explains the dependence of the reservation utility and the hazard rate on time. An increase on an individual's expected length of life (T) reduces his/her reservation utility and increases the hazard rate for moving. Because T decreases over time, the reservation utility increases and the hazard rate decreases over time. There are two consequences to this pattern. First, the hazard rate is higher for younger than for older adults. This assertion fits well to the fact that most migration takes place when people are young adults because they have more time to recover the cost of mobility. Therefore, age is expected to be an important variable affecting the hazard rate for moving. Second, the longer an individual stays in one place, the higher is his/her reservation utility to move. In other words, the model predicts that the longer the individual lives in one state, the smaller is his/her tendency to move.

A higher mean for expected utility at the destination (μ_{Um}) will increase both reservation utility and the hazard rate for moving. This increases the individual's confidence of having a better life in the new place, which leads to a positive effect on reservation utility and an indirect negative effect on the hazard rate. However, there is a direct positive effect of μ_{Um} on hazard rate in Equation (2.10). The direct effect is dominant so that an increase in the expected utility at a new residence increases the hazard rate for moving.

The model assumes that the expected utility corresponding to moving (μ_{Um}) is evaluated net of moving costs. All else being equal, a higher cost of moving will reduce expected utility at a new destination. Previous studies, such as Mincer (1978), suggest that family-related variables, such as the number of school-age children and marital status, are positively related to the cost of migration. Other studies have shown that being married and having school-age children reduce the chance of the family migrating. Because of the attachment to the localities, being self-employed or a farmer increases the opportunity cost and reduces the hazard rate for moving.

As the discount rate (r) increases, the future expected benefits are discounted more heavily. The comparative static result shows that an increase in r would reduce an individual's reservation utility at the origin and increase the hazard rate for moving. One reason is that the reservation utility is a function of discounted expected utility [see Equation (2.8)]. This model is different from deterministic models, such as the labor force participation model where the reservation wage is not a function of the wage offer. Because reservation utility is an increasing function of discounted expected utility in this model, reservation utility will decrease as expected utility is more heavily discounted. Because the effect of r on the hazard rate occurs only through reservation utility, the hazard rate for moving will be higher. The interpretation is when the discount rate increases, the cost of waiting is higher. Because every individual is assumed to migrate eventually, the increase in the discount rate makes the individual migrate sooner.

We do not emphasize the effect of the discount rate on an individual's migration decision in the empirical study for two reasons. First, this study explores the effects of personal and local characteristics on migration. We assumed that every individual in every state faces the same discount rate; therefore, including the discount rate in the empirical model is not necessary. Second, we assumed that the discount rate, represented by the real interest rate, has been constant over time.[1]

---

[1]Nominal and real interest rates from 1953 to 1993 are presented in Appendix E. Except for the late 1970's and early 1980's, the real interest rate was constant around 1 to 2 percent. It is possible that the movement of the real interest rate in the late 1970's and early 1980's might affect migration decisions. Future empirical research may incorporate the effect of interest rate on migration decisions.

We could rearrange the information in Equation (2.8) to obtain:

$$[Ur - Unm/(1 - e^{-rT})]r = \delta[1 - F(Ur)][E(Um|Um \geq Ur) - Ur] \qquad (2.11)$$

This equation has an appealing interpretation. The left-hand side represents the marginal cost of rejecting an offer to migrate. The right-hand side represents the marginal expected benefit from rejecting an offer to move, given that the offer would be accepted if $Um \geq Ur$. In other words, this equation equates the marginal cost and marginal expected benefit of rejecting an offer and continuing to live at the origin.

Equation (2.11) is the optimality condition for an individual's reservation utility policy. The individual stays at the origin but searches for a new place to live unless the utility offer is greater than the reservation utility. From a dynamic programming point of view, this optimality condition affirms that the optimal decision maximizes the sum of the flow of utility in the current period and the mathematical expectation of the discounted flow of utility in the future, given that the future decision is optimally made (Divine and Kiefer, 1991).

CHAPTER 3. DATA DESCRIPTION

This chapter describes the data used in this study. The observations in this study are for adult males surveyed for the Panel Study of Income Dynamics (PSID). The survey started in 1968 and followed individuals over time- hence a panel. The data contain a large amount of information on personal and family characteristics, including residence. These data are supplemented with information on the characteristics of the states in which these individuals reside, such as local labor market, cost of living, amenities, and disamenities. The first and second sections of this chapter describe the PSID and state data, respectively. The third section illustrates the pattern of migration based on the actual sample used in this study. The last section presents the procedures used to prepare the data set for empirical estimations.

The PSID Microdata

The PSID data are collected by the Institute for Social Research, The University of Michigan. The participants consist of a panel of individuals selected in 1968 who are re-interviewed every year. New individuals are added to the survey when they join the household of an original 1968 data panelist. Individuals who leave an original PSID household are followed and re-interviewed as long as they can be located. According to the PSID user's guide (Hill, 1992), the data started with approximately 5,000 family units consisting of 18,000 individuals across 40 states in 1968 and had grown to around 7,000 families and 37,500 individuals by 1988. The PSID has annual information on personal, employment, and local characteristics of every individual. It provides a wide variety of information at the family and individual levels, and some information about the location in which the households reside. The central focus is the economic and demographic data; substantial detail exists on income source and amount, employment, family composition changes, and residential location. The PSID uses a hierarchical file structure. Household, family, and individual information make up each household (or housing unit) record. Every year, newcomers and every individual who died or left the households are recorded.

The original PSID consisted of two independent samples: a cross-sectional national sample and a national sample of low-income families. The first was drawn by the Survey Research Center (SRC) and therefore is known as the SRC sample. The SRC sample was an equal probability sample of households in the 48 contiguous U.S. states. The second sample was drawn from the Survey of Economic Opportunity among lower-income families, conducted by the Bureau of the Census for the Office of Economic Opportunity. To better represent the U.S. population in general, we used the SRC sample for this study.

This study will cover the 20-year (1968-87) migration behavior of individuals and uses the SRC sample from the 1968-88 PSID data, compiled on two CD-ROMs. Because the income data are for the previous year, the study did not fully utilize the data from the last year of observation (1988). The 20-year period is long enough to capture repeat migration and to provide more qualified data on the duration of an individual's stay in one state. Every individual was followed from the year after he/she completed school to the time when he/she retired, died or was lost. We expect different migration behavior for males and females and have chosen to examine only the migration behavior of males. The data from all years 1968-88 were used to fit the search model. The point-in-time discrete choice model can be fitted by using data from two consecutive years. To increase the proportion of the migrant sample and to simplify the sampling procedure, five-year data from 1968-73 were used to fit the discrete choice model.

The sample for this study consists of 915 male individuals who were the heads of households and were between 19 and 45 years of age in 1968. Individuals who refused to participate in the survey, were in mental institutions, prisons, or religious orders, or were recorded as technical error anytime between 1968 and 87 were deleted from the sample. Individuals who joined the army are assumed to have less control of their migration behavior. Therefore, they are followed until the year before joining the army and then are treated as lost or missing. It is important to note that the starting and ending year for each observation are not the same for every individual.

State Data

The state data consist mainly of variables representing annual local labor market characteristics, cost of living, and local amenities in each U.S state from 1968 to 1987. Most of the data are taken from the Statistical Abstract of the United States (1968-1993). The state data on labor market and cost of living are to be used to predict the potential wage for each individual living in a particular state. Local amenities also enter the equation explaining the migration decision. Tables B.1 and B.2 in Appendix B illustrate the local characteristics used in this study.

Following Tokle and Huffman (1991), this study uses state unemployment and employment growth rates to represent the local labor market characteristics. The unemployment rates are taken from the unemployment rate of civilian workers. Equilibrium unemployment rates differ across states and these differences tend to persist over time. Some states, such as Michigan, West Virginia, and Mississippi, have a higher unemployment rate than the U.S. average. Kansas, Iowa, and Nebraska are known as areas with low unemployment rates. The state annual employment growth data are derived from the number of workers in non-agricultural establishment in every state from 1968 to 1987. In general, the United

States experienced increasing employment growth in the 1970s and a declining rate in the 1980s. Southern states such as Arizona, California, and Nevada have experienced a higher employment growth than have other states. In the late 1980s, however, Mississippi, Louisiana, Texas, and Oklahoma suffer a sharp drop in employment growth. Some northern states, such as Michigan, New York, Pennsylvania, and West Virginia have experienced employment growth below the U.S. average.

The local cost of living is proxied by the price of land and the percentage of urban population in a state. The data on the land prices are taken from Huffman and Evenson (1991). North-eastern states such as New Jersey, Rhode Island, Massachusetts, and Connecticut have the highest land prices. Land prices in states with fertile agricultural land such as the midwestern states are also higher than the U.S. average, but land prices in midwestern states decreased in the mid-1980s due to the farm crisis. The mountain region and the Western states, starting from South Dakota, North Dakota, and Nevada, have had lower-than-average land prices. However, California has had land prices higher than the U.S. average, as has Florida.

The percentage of urban population in every state is another indicator of the cost of living. On average, about 67 percent of the population in the U.S. lives in urban areas. The percentages differ across states, but are relatively constant over time. Naturally, Washington, DC, is the only area with 100 percent of the population living in urban areas. California and New Jersey have 90 percent of their populations living in urban areas. Other states with more than 80 percent urban population include New York, Connecticut, Rhode Island, and Massachusetts. The midwestern states and the mountain region have a relatively small percentage of urban population.

Local amenities or disamenities are represented by temperatures, crime rate, and the proportion of a state's area devoted to state and national parks. The data on state temperatures are represented by 30-year average January and July temperatures. We assumed that the temperature variability is constant over time.

Crime rates are represented by the annual number of murders and incidents of nonnegligent manslaughter per 100,000 population in each state. The crime rate in the United States increased from around 6.3 percent in 1968 to 7.8 percent in 1978 before moderately declining to 7.0 percent in 1987. Washington, DC, has an extremely high and increasing crime rate, increasing from 19 percent in the 1970s to 36 percent in the late 1980s. Similar trends occurred in Michigan and New York, where the rates increased from around 6 percent in the late 1960s to more than 11 percent in the late 1980s. Alaska, Texas, Louisiana, Florida, and Georgia have had crime rates consistently around 10 percent or higher for the period studied. The New England states and the midwestern states have had the lowest low crime rates.

The percentage of total state and national park areas in every state is also used to measure local amenities. Washington, DC, has a relatively high percentage (16 percent) of land belong to the national park system. Other states with high percentages are California, Connecticut, New Jersey, and Massachusetts. We assumed that the total areas of state and national parks were relatively constant during the 1968-87 period.

Pattern of Migration

By observing the state of residence of every individual in the sample from 1968 to 1987, we obtain a picture of observed migration behavior.[2] Most individuals in the sample did not engage in interstate migration during 1968-87. The distribution of the number of interstate moves can be

---

[2]An individual who lives in near a state border, especially in a large metropolitan area, might have a short distance (local) move that is classified as a long distance (between states) move. However, for most moves, this is not the case.

seen in Table 3.1. Among the 915 individuals in the sample, 722 (78.9 percent) stayed in the same state throughout the period. Among the 193 individuals who migrated at least once, 97 (more than 50 percent) moved only once. The frequency of movers decreased as the number of moves increased. Only five individuals moved more than 4 times during 1968-87.

By examining the sample who migrated at least once (193 individuals), we found that 91 (47.2 percent) had lived in the same state for two different subperiods, and that 72 (37.8 percent) moved back to the state where they grew up. By selecting individuals who had moved away from the state where they grew up and who had migrated at least twice in their lifetime, we found that 53.7 percent moved back to the state where they grew up. These summary results suggest that late moves are affected by prior experiences. When individuals decide to move to another state, the state where they grew up or states where they have lived before are frequent destinations.

Table 3.1. The distribution of the number of interstate moves, 1968-87

Number of move	Frequency	Percent	Frequency (Cumulative)	Percent
0	722	78.9	722	78.9
1	97	10.6	819	89.5
2	59	6.4	878	96.0
3	23	2.5	901	98.5
4	9	1.0	910	99.5
5	3	0.3	913	99.8
6	2	0.2	915	100.0

The length of time that an individual lived in one state is defined as a residence spell. A spell is complete if the time when the spell started and ended can be determined from the data. When the starting year is not observed, the spell is left-censored. Similarly, when the ending year is not observed, the spell is right-censored. The distributions of starting and ending times for residence spells are presented in Figure 3.1. It is clear that the starting and ending years for the sample varied across individuals. From Figure 3.1, we see that 844 individuals might have left-censored spells. In general, every individual has one right-censored spell because the most recent residence spell is incomplete. However, additional criteria are used to determine the status of the last spell when the ending year is not observed. When an individual died, we assumed that the last spell is incomplete. We assumed that if the individual had not died, he would have stayed in the same state. We also assumed that the last spell of individual who joined the army or was lost is complete. There is a greater chance that an individual moved to another state when he became lost or joined the army. Among 915 individuals in the sample in 1968, 56 (6.12 percent) died, 144 (15.74 percent) were lost[3], and 132 (14.4 percent) retired during 1968-87. The sample has a large proportion of

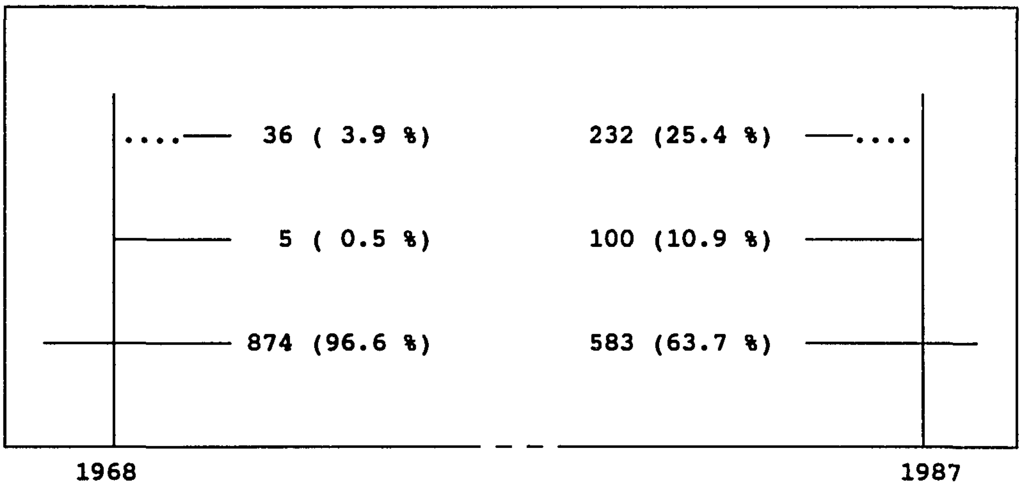

Figure 3.1. Distributions of starting and ending times of observations for the 915 males in the sample.

[3]Individuals who refused to participate in the survey at any time during 1968-87 were deleted from the sample. Individuals who were classified as 'lost' consist of those who joined the army (30.5 percent), and those who moved out from the United States and/or were really lost (69.5 percent).

left-censored spells because outcomes before 1968 were not perfectly recorded. Among the 874 individuals who started living in a state before 1968, 702 remained in the same state up to 1987. Thus, 702 spells are potentially both left- and right-censored.

We derived a total of 1,268 independent residence spells from 20 years of observations for all 915 individuals in the sample. When we consider only the completed spells which started in 1968 or after, 207 spells (16.32 percent) were both right and left completed. It is thus clear that at least 874 spells are left-censored. The PSID does not have clean information on how long the individuals had lived in a state before 1968. When we made the assumption that an individual who grew up in a state remained in the state until 1968 and defined the starting year of the spell as the year when he became age 19, the number of left-censored spells decreased. We then derived 284 completed spells (22.40 percent). However, some left-censored spells remained because some individuals moved from the state where they grew up before 1968. The starting year for the residence spell of these individuals remained unknown. In Chapter 5, we discuss the empirical procedure to estimate the starting year for these individuals. When we closed all the left-open spells, we derived a total of 496 completed spells (39.17 percent).

Procedures Applied to Prepare the Data

This section presents the procedures used to prepare the data for empirical analysis. The estimation of probit equations and hazard functions utilizes all relevant personal and local characteristics, which were taken annually. The first step in preparing the data was to select the sample of individuals from the PSID data set. Having selected the sample, information relating to the residential mobility of every individual from 1968 to 1987 was collected. This step included identifying the starting and ending years. In addition, other relevant personal characteristics corresponding to every individual were collected for every year from 1968 to 1987.

The second step was to gather annual data for state (local) characteristics covering the period 1968-1987. These variables included local labor market conditions, cost of living, temperature, and crime rate. Then, the predicted unemployment and job growth rates were calculated.

The third step was to develop the main data set used in the estimation by merging the 20-year individual and local data, based on state of residence in each year. The structure of the data remained longitudinal. In other words, we maintain the base structure from the PSID and added the characteristics of residence into it.

The fourth step was to change the structure of the data into an annual cross-sectional structure, which was needed to fit the wage equation. Then, the predicted and residual wage corresponding to every individual in every year was calculated and included in the data set. Thus the main data set consists of personal characteristics, local characteristics, potential, and residual wages, in every year for the 1968-87 period.

The final step consisted of deriving the data corresponding to the spell of residence. To do this, the structure of the data was formed back to be longitudinal. Then, the single data observations were changed from annual individual to residence spells. In other words, the data from one person for 20 years could make one, or more spells. This final procedure set the residence spell as the unit or single observation and made the data set ready for the estimation of the hazard function.

CHAPTER 4. THE ECONOMETRIC MODEL

This chapter discusses the general econometric models of migration. More precise empirical specifications of the models and definitions of the variables are presented in the next chapter. The first section of this chapter presents the discrete choice model, which consists of a standard probit model. The second section elaborates the statistical analysis of time duration data to fit the search model of migration.

The Discrete Choice Model

From Equations (2.1) and (2.2) in Chapter 2 and following Maddala (1983), we write the reduced form of the indirect utility function:

$$Um = X\ \beta_m + \epsilon_m \tag{4.1}$$

$$Unm = X\ \beta_{nm} + \epsilon_{nm} \tag{4.2}$$

The selection equation can be expressed as:

$$M^* = (Um - Unm)\lambda \tag{4.3}$$

where

M^* = variable representing the decision to move/stay,

Um = utility corresponding to migrating,

Unm = utility corresponding to staying,

X = variables explaining migration decision,

β_m, β_{nm}, λ = vectors of parameters, and

ϵ_m, ϵ_{nm} = random disturbances.

M^* in Equation (4.3) is not observable. However, a dummy variable M is observed and takes the value of 1 if an individual migrates and 0 if he stays at the same place during the observation period. We can write

$$M = 1 \text{ if } M^* > 0$$
$$= 0 \text{ otherwise.} \quad (4.4)$$

Because the migrant and nonmigrant samples are mutually exclusive, individuals can not be observed as a migrant and nonmigrant simultaneously. The probability of migration for the i^{th} individual is:

$$\begin{aligned} Pr(M_i = 1) &= Pr(M^*_i > 0) = Pr\ [(Um_i - Unm_i)\lambda > 0] \\ &= Pr\ [(\epsilon_{nmi} - \epsilon_{mi}) < X_i(\beta_m - \beta_{nm})) \\ &= F(X_iB) \end{aligned} \quad (4.5)$$

where F(.) is a distribution function of $(\epsilon_{nm} - \epsilon_m)$, and $B = (\beta_m - \beta_{nm})$. Assumed that $(\epsilon_{nm} - \epsilon_m)$ equals to ϵ, and ϵ is normally distributed with a zero mean and a constant variance, the model leads to a standard probit model. When there are n1 individuals who migrate out of a total n individuals in the sample, the likelihood function can be written:

$$\begin{aligned} L(B|M,X) &= \prod_{i=1}^{n_1} Prob\ (M_i = 1) \prod_{i=n_1+1}^{n} Prob\ (M_i = 0) \\ &= \prod_{i=1}^{n_1} F(X_iB) \prod_{i=n_1+1}^{n} (1 - FX_iB) \\ &= \prod_{i=1}^{n} F(X_iB)^{D_i} (1 - FX_iB)^{1-D_i} \end{aligned} \quad (4.6)$$

where $D_i = 1$ if an individual migrates and 0 otherwise. By maximizing the likelihood function in Equation (4.6) we could obtain the estimate of B using maximization procedure found in Maddala (1983) or Greene (1990). The log form of the likelihood function is:

$$Ln\ L = \sum_{i=1}^{n} D_i \ln F(X_i\beta) + \sum_{i=1}^{n} (1 - D_i) \ln(1 - F(X_i\beta) \quad (4.7)$$

Taking the first-order condition and setting it equal to 0 for maximization yields

$$\partial \text{Ln } L/\partial\beta = \sum_{i=1}^{n} D_i(f/F)\ X_i + \sum_{i=1}^{n} (1 - D_i)(f/(1 - F))X_i = 0 \qquad (4.8)$$

where f and F are the respective values of relevant density and cumulative distribution functions at $X_i\beta$. This equation is nonlinear in β, so that the estimate of the parameter can not be found directly. The common method for finding the maximum is the iterative procedure called the Newton-Raphson method. Judge (1988) and Greene (1990) provide the detailed derivation for obtaining the parameter estimates and their asymptotic standard errors. It can be shown that the estimators are consistent, asymptotically normally distributed, and asymptotically efficient. Having found the estimate of β, called B^*, the marginal effect of X_j on the probability of migrating is:

$$\partial Pr(M = 1)/\partial X_j = \partial F(XB)/\partial X_j = B_j^* f(XB^*) \qquad (4.9)$$

where f(.) is the normal density function.

The Search Theory Model

The search theory model has two stochastic processes embedded in it. One is in the utility at the destination, or the offer utility, and the second is in the duration. The model with constant reservation utility yields an exponential duration distribution $g(t) = H.\exp(-H.t)$ and a particular density of an accepted utility $f_a(Um) = f(Um)/[1 - F(Ur)]$ for $Um \geq Ur$. The parameter H, which is the hazard rate, is equal to the product of the offer-arrival rate and the acceptance probability, $H = \delta.[1 - F(Ur)]$, where F(.) is the cumulative distribution function of the stochastic variable Um. The densities of utilities and durations are related in the sense that they depend on common parameters, which determine

Ur and f(w), but the stochastic variable utilities and durations are independent. Therefore, the joint probability distribution of utilities and durations is given by $p(Um,t) = f_a(w).g(t)$. The distribution of utilities and durations are expected to differ across individuals due to variation in personal and local characteristics. Given these explanatory variables X and parameter β, the joint density function is $p(Um_i,t_i|X_i,\beta)$, where i represents the individual index.

However, specifying the joint density distribution of durations and utilities is difficult to implement. Both utility and its distribution are usually not observed, and even if indirect utility is observed, only those larger than the reservation utility are observed. This fact needs to be taken into account by constructing the density of accepted utility $f_a(Um)$ from f(Um). In addition, an informative specification of the joint density function of durations and utility requires some arbitrary assumptions about functional forms. One approach, or research strategy, is to focus on duration and utility separately. Focusing on duration seems better because duration data are more readily observable. The specifications of the time duration models are presented later in the section discussing the econometric models.

The econometric methods for analyzing time duration data are based on specification of a hazard function rather than a density function. It is possible to go back and forth between hazard and density functions, so the difference is merely one of convenience in specification, estimation, and interpretation. The probability distribution of duration can be specified by the distribution function $G(t) = Pr(T^* < t)$, the corresponding density function $g(t) = dG(t)/dt$, and the survivor function $S(t) = 1 - G(t) = Pr(T^* > t)$. The hazard function for migration, or the limiting probability that a spell will be completed in a short time period h, is:

$$H(t) = \lim_{h \to 0} Pr(t < T^* \le t+h|T^* > t)/h = g(t)/[1 - G(t)] \qquad (4.10)$$

H(t) can be interpreted as the rate at which spells will be completed at duration t, given that a spell lasts until time t. This hazard function is a convenient definition of duration dependence. Positive (negative) duration dependence means that the probability that a spell will end shortly increases (decreases) as the spell increases in length.

For many econometric time duration models, it is natural to analyze conditional duration distributions where the conditioning is with respect to observed variables X and unobserved parameters β. Indeed, by analogy to most regression analysis, much of the attention is focused on the effect of the regressors X on hazard rates and durations. With the explanatory variable X, we define the hazard function as:

$$\begin{aligned} H(t,X,\beta) &= \lim_{h \to 0} \Pr(t < T^* \leq t+h \mid T^* > t,X,\beta)/h \\ &= g[t \mid X(t),\beta]/[1 - G(t \mid X(t),\beta)] \end{aligned} \tag{4.11}$$

The integrated hazard conditional on X and β is:

$$\Lambda(t,\beta,X) = \int_0^t H(u \mid X,\beta)\, du \tag{4.12}$$

Equation (4.12) does not have a particular interpretation, but it is useful in practice because this integrated hazard follows a standard exponential distribution. The survivor function can be written as:

$$S(t,\beta,X) = \Pr(T^* > t \mid \beta,X) = \exp[-\int_0^t H(u \mid X,\beta)\, du] \tag{4.13}$$

and the conditional density function for duration can be written as:

$$g(t,\beta,X) = H(t,X,\beta).S(t,\beta,X) \tag{4.14}$$

It is important to specify how the explanatory variables X and parameters β actually affect the hazard function. One common and simple way to interpret those relationships is by making a proportional hazard model where the effects of X and β are multiplied. This is a re-scaled hazard function:

$$H(t,X,\beta) = Ho(t).\Phi(X,\beta) \tag{4.15}$$

where Ho(t) is a baseline hazard that corresponds to $\Phi(X,\beta) = 1$. This specification leads to the log of the conditional hazard being a linear function of t, X, and β. $\Phi(X,\beta)$ is nonnegative and the common form is $\exp(X\beta)$ with one value X in every spell t. Many studies utilize an overtime constant hazard specification normalized to Ho(t) = 1, so that:

$$H(t,X,\beta) = \exp(X\beta) \tag{4.16}$$

The partial effect of X on H(.) is:

$$\partial \text{Ln } H(t,X,\beta)/\partial X = \partial \text{Ln } H(t,X,\beta)/\partial X = \beta \tag{4.17}$$

The coefficient β can be interpreted as the constant proportional effect of X on the conditional probability of completing a spell. Because of the assumption of a constant hazard rate over time which underlies an exponentially distributed duration, the expected length of completed duration $E(T^*) = 1/H(.)$, the survivor function becomes:

$$S(t,\beta,X) = \exp[-\exp(X\beta)] \tag{4.18}$$

The economic model of migration decision based on the search theory implies that the hazard rate of moving can be a function of the expected length of remaining life. The constant hazard rate specification might not capture this aspect. The common specification of the time-dependent hazard rate is made by assuming that the distribution of the completed time spell is Weibull. Following Kiefer (1988), one form of hazard function can be specified by writing $Ho(t) = \alpha t^{\alpha-1}$, and the hazard function becomes:

$$H(t,x,\beta) = \alpha t^{\alpha-1}\exp(X\beta) \tag{4.19}$$

The effect of individual X to H(.) is similar to Equation (4.17). In addition, by taking the derivative of H(.) with respect to t, we could find:

$$\partial H(t,X,\beta)/\partial t = \alpha(\alpha - 1)t(\alpha - 2)\exp(X\beta) \tag{4.20}$$

This result means that when $\alpha > 1$, the hazard rate increases as the length of stay increases, and decreases otherwise. The expected length of completed duration is $E(T^*) = \exp(-X\beta).\Gamma[(\alpha + 1)/\alpha]$, where $\Gamma(.)$ is a gamma function.

For the purpose of estimating the time-dependent hazard rate, we specify the hazard function by following Greene (1991) as:

$$H(t,X,\beta) = \lambda/\sigma(\lambda t)^{1/\sigma-1} \tag{4.21}$$

where $\lambda = \exp(X\beta)$. Taking the derivative with respect to t, we get:

$$\partial H(t,X,\beta)/\partial t = \lambda^{1/\sigma}\ 1/\sigma\ (1/\sigma - 1)\ t^{1/\sigma-2} \tag{4.22}$$

This result means that $\partial H(t,X,\beta)/\partial t < 0$ when $\sigma > 1$, and $\partial H(t,X,\beta)/\partial t > 0$ otherwise. All else being equal, the hazard rate decreases as the length of stay increases when $\sigma > 1$, and decreases otherwise. Taking the derivative of Ln $H(t,X,\beta)$ with respect to X, we find:

$$\partial \mathrm{Ln}\ H(t,X,\beta)/\partial X = \beta/\sigma \tag{4.23}$$

The effect of X on the hazard rate will be positive if β is positive. It is important to note that when $\sigma = 1$, Equation (4.23) in the time-dependent hazard rate model is similar to Equation (4.17) in the case of a constant hazard rate. The function of σ is therefore to re-scale the effect of β. When the hazard rate is decreasing (increasing) over time, namely, when σ

is greater (smaller) than one, the marginal effect of X on the hazard rate shown in Equation (4.23) is also smaller (bigger). The survival function corresponding to the hazard rate in Equation (4.21) is:

$$S(t,X,\beta) = \exp\{-[\exp(X\beta)t]^{1/\sigma}\} \tag{4.24}$$

The issue of left-censoring turns out to be important for the time-dependent hazard function. Heckman and Singer (1985) showed that unless the hazard function is constant over time or the completed duration is exponentially distributed, left-censored spells are hard to deal with. In the exponential case, the distribution of duration after elapsed time t does not depend on t and so the distribution of the remaining duration in the left-censored spells is the same as the distribution for uncensored spells.

The problem of left-censoring is not crucial in estimating the time-dependent hazard function in unemployment studies. First, the length of a completed unemployment spell is relatively shorter than the residence spell in migration studies. Second, the data have complete information on unemployment history, or at least the length of time an individual has been unemployed at the beginning of the survey. In migration studies, however, every individual has resided in a certain place for some time before the survey took place. Because the PSID data do not have any clear information on how long an individual was at a residence before 1968, some additional procedures are needed to empirically estimate the time-dependent hazard function. Those procedures are discussed in Chapter 5.

By assuming that the transition process between residence spells follows a Semi-Markov Process, the separate spells of an individual can be taken to be independent across spells (Devine and Kiefer, 1991). In other words, spell of living in one place is an independent event drawn from all possible outcomes in a sample space.

The estimation of the model follows a maximum likelihood procedure. Greene (1991) and Kiefer (1988) provide the procedure to estimate the model. Recall that we have a hazard function $H(t,X,\beta,\sigma)$, survivor function $S(t,\beta,X,\sigma)$, and the conditional density function $g(t,\beta,X,\sigma) = H(t,X,\beta,\sigma)S(t,X,\beta,\sigma)$. When a sample of n independently completed spells t_i and explanatory variables X corresponding to them were available, the likelihood function will be $L^*(\beta,\sigma|X_i,t_i) = g\ [t_i,\beta,\sigma,X(t_i)]$. When spell t_j is censored, the only information available is that the duration was at least t_j. As a consequence, these observations contribute to the likelihood function only through the survivor function $S[t_j,\beta,\sigma,X(t_j)]$, or the probability that the duration is longer than t_j. Letting $\delta_i = 1$ if the i^{th} spell is completed and $\delta_i = 0$ otherwise, then the full log likelihood function will be:

$$\text{Ln } L(\beta,\sigma) = \sum_{i=1}^{n} \delta_i \ln g[t_i,\beta,\sigma,X(t_i)] + \sum_{i=1}^{n}(1 - \delta_i)\ln S[t_i,\beta,\sigma,X(t_i)] \qquad (4.27)$$

Based on specification of the hazard function corresponding to the distribution of the duration, we can estimate parameter β by maximizing the log-likelihood function in Equation (4.27).

To simplify the estimation procedure, Greene (1991) created a transformation, $w_i = (\ln t_i - X_i\beta)/\sigma$, such that the log-likelihood function becomes:

$$\text{Ln } L(\beta,\sigma) = \sum_{i=1}^{n} \delta_i \ln\ [g(w_i)/\sigma] + \sum_{i=1}^{n}(1 - \delta_i)\ln S(w_i) \qquad (4.28)$$

The derivatives for the first-order conditions are:

$$\partial \text{Ln } L/\partial\beta = [\delta_i \partial \text{Ln } g/\partial w_i + (1 - \delta_i)\partial \text{Ln } S/\partial w_i](-X_i/\sigma) = 0 \qquad (4.29)$$

$$\partial \text{Ln } L/\partial\sigma = [\delta_i \partial \text{Ln } g/\partial w_i + (1 - \delta_i)\partial \text{Ln } S/\partial w_i](-1/\sigma) - (\delta_i/\sigma) = 0 \qquad (4.30)$$

The iterative procedures are used to find the maximum of the likelihood function. The starting values of every iteration are the parameter estimates based on Ordinary Least Square estimates of the regression between the log of the duration and the explanatory variables. Two common iterative procedures suggested by Greene (1991) are Newton's method and the Davidson-Fletcher-Powell (DFP) method (Fletcher, 1980). For well-behaved likelihood functions, such as Tobit, logit, and probit, Newton's method is a natural choice. For the less well behaved, such as in the likelihood for the hazard functions, the DFP method is used. This study uses the convergence criterion = 0.0001. It is important to note that the maximum likelihood estimation leads to a consistent, asymptotically normally distributed, and asymptotically efficient estimator.

Heterogeneity, or individual specific effects, might influence the expected duration. Unmeasured heterogeneity can cause estimation bias in duration models (see Chamberlain, 1985a; Heckman and Singer, 1985). Following Devine and Kiefer (1991) and Greene (1991), the heterogeneity effect is represented by a random variable v that is distributed independently from both X and t, so that the survivor for exponential or Weibull model is modified into:

$$S(t|v) = v \exp\{-[\exp(X\beta)t]^{1/\sigma}\} \tag{4.31}$$

where the random variable v is a heterogeneity effect. Following Greene (1991) we assumed that v is distributed as gamma with parameters k and R, with k = R so that the expected value of v is normalized to one. Greene showed that the modified hazard function becomes:

$$H(t,X,\beta,v) = S(t)^{\theta}(\lambda/\sigma)(\lambda t)^{1/\sigma-1} \tag{4.32}$$

where $\lambda = \exp(X\beta)$, S(t) is the expected value of v over $S(t|v)$, and θ is the variance of v. When the limit of $S(t)^{\theta}$ goes to 1 as θ goes to 0 (variance of v equal to zero), the hazard function converges to the original exponential hazard function. The interpretation is that θ captures the sensitivity of the hazard function to heterogeneity. The further θ deviates from zero, the greater is the effect of heterogeneity. There is a common convergence problem in estimating the model with heterogeneity because the log-likelihood is volatile in parameter θ. The detailed procedure used to deal with heterogeneity can be found in Greene (1991).

CHAPTER 5. EMPIRICAL SPECIFICATION

This chapter presents the actual equations and definitions of variables used to fit the migration models. The specifications are based on the general econometric models discussed in previous chapters. Some traditional and some new variables are used as regressors in the migration models. Because the predicted wage is used to explain migration decisions, the first section of this chapter presents the specification of the wage equation. The second and third sections discuss the specification of the probit equation and the hazard function, respectively.

The Wage Equation and Relative Wage Performance

One important variable in the migration decision is the individual's income at his current location. A larger full income leads to a larger utility and a lower chance of migrating to another state. To get free of labor supply decisions and endogeneity of cash income, an individual's wage and time endowments are used to proxy his full income. Because everyone has the same time endowment, it becomes part of the intercept term. However, the actual wage is correlated with other personal and local characteristics that affect the migration decision. Therefore, this study develops a new measurement, the individual's relative wage performance.

To develop the relative wage performance, lets begin by writing the specification of the individual's hourly wage:

$$Ln(W_{ik}/P_k) = \beta_x X_i + \beta_z Z_k + \beta_a A_k + \varepsilon_i \tag{5.1}$$

where

W_{ik} = nominal hourly wage of individual i living in state k,

P_k = price index for inputs used to produce indirect utility in household of individual living in state k,

X_i = personal characteristics,

Z_k = state characteristics which affect labor productivity,

A_k = hedonic state amenity attributes,

ε_i = random disturbance representing luck, $E(\varepsilon_i) = 0$,

β_x = return to personal characteristics,

β_z = return to local productivity characteristics,

β_a = return to hedonic local attributes.

When the markets work well, the price of traded goods are approximately the same across states. The prices of nontraded goods, however, might differ greatly between states. We assume that the hedonic price index for goods and services purchased in state k is a linear function of the national price index, prices of state nontraded goods, and regional price. Define the state price index as:

$$\begin{aligned}
Ln(P_k) &= \alpha_1 Ln(P) + (1 - \alpha_1)Ln(PLAND_k) + \alpha_2 URBAN_k + \alpha_3 CLIMATE_k + \\
&\quad \alpha_4 REGION_k + \delta_k \\
&= \alpha_1 Ln(P) + (1 - \alpha_1)[Ln(PLAND_k)-Ln(P)] + (1-\alpha 1)Ln(P) + \\
&\quad \alpha_2 URBAN_k + \alpha_3 CLIMATE_k + \alpha_4 REGION_k + \delta_k \\
&= Ln(P) + (1 - \alpha_1)[Ln(PLAND_k)-Ln(P)] + \alpha_2 URBAN_k + \\
&\quad \alpha_3 CLIMATE_k + \alpha_4 REGION_k + \delta_k \qquad (5.2)
\end{aligned}$$

where

P = national price index for goods and services purchased by households,

$PLAND_k$ = nominal price of land in state k,

$URBAN_k$ = percentage of urban population in state k,

$CLIMATE_k$ = climatic characteristics of state k,

$REGION_k$ = regional dummy variables, and,

δ_k = random disturbance, $E(\delta_k) = 0$.

By substituting Equation (5.2) into Equation (5.1), we obtain the following wage equation:

$$\begin{aligned} Ln(W_{ik}/P) = {} & \beta_x X_i + \beta_z Z_k + \beta_a A_k + (1 - \alpha_1)[Ln(PLAND_k/P)] + \alpha_2 URBAN_k + \\ & \alpha_3 CLIMATE_k + \alpha_4 REGION_k + \varepsilon_i + \delta_k \end{aligned} \tag{5.3}$$

Equation (5.3) shows that by incorporating PLAND, URBAN, CLIMATE, and REGION on the right hand side, we can use the national price index to create the dependent variable. The parameter representing the effect of factors affecting wages are estimated as if we had used the state price index for purchased goods and services to deflate the wage.

This representation of the labor demand equation follows the convention in the labor literature of making the demand function facing an individual with given attributes perfectly elastic. The interpretation of this market wage is as the potential wage of an individual with certain personal and local characteristics. Given that state labor markets in the United States are well integrated, i.e. individuals and firms are free to move, and entry and exit occurs, the estimated coefficients on the regressors in Equation (5.3) are estimates of equilibrium compensating differentials (see Rosen, 1986). If Equation (5.3) is fitted to a panel data set of individuals living in all states then the estimated coefficients represents the United States labor market's average valuation of characteristics. In the PSID data, the nominal wage is calculated from total annual earnings from work divided by total annual hours of work. Based on Equation (5.3.) and following Tokle and Huffman (1991), the complete specification of the real market wage or real potential wage of individual i in state k and year y is:

$$\begin{aligned} Ln(W_{iky}/P_y) = {} & \beta_o + \beta_1 EDU_{iy} + \beta_2 EXP_{iy} + \beta_3 EXP_{iy}^2 + \beta_4 RACE_i + \beta_5 Ln(PLAND_{ky}/P_y) + \\ & \beta_6 URBAN_{ky} + \beta_7 JOBGR_{ky} + \beta_8 UNRATE_{ky} + \beta_9 CRIME_{ky} + \beta_{10} JAN_k + \\ & \beta_{11} JULY_k + \beta_{12} TIME_y + \beta_{13} TIME_y^2 + \beta_{14} DS_{iy} + \beta_{15} DW_{iy} + \\ & \beta_{16} DNC_{iy} + \epsilon_{iy} \qquad (5.4) \end{aligned}$$

The term ϵ_{iy} is a random disturbance and for convenience in the estimation of the hazard function, ϵ_{iy} is assumed to be distributed independently of the distribution of the duration of residence spells.[4] The description of the variables used in the wage equation can be found in Table 5.1.

The use of actual values of state employment growth and unemployment rates is based on the assumption in the economic model of migration derived above. In search theory, individuals are assumed to observe both their indirect utility and their wages perfectly. However, previous studies (Tokle and Huffman, 1991; Topel, 1986) suggest that individuals and firms respond to the expected value of local variables rather than to actual values. For this reason, another wage equation is fitted using the predicted value of employment growth ($PJOBGR_{ky}$) and the predicted value of the unemployment rate ($PURATE_{ky}$). The residual employment growth ($RSHOCK_{ky}$) and the residual unemployment rate ($RURATE_{ky}$) are included in the second equation to capture unanticipated shocks in the state labor market. The second wage equation is specified as:

$$\begin{aligned} Ln(W_{iky}/P_y) = {} & \alpha + \alpha_1 EDU_{iy} + \alpha_2 EXP_{iy} + \alpha_3 EXP_{iy}^2 + \alpha_4 RACE_i + \alpha_5 Ln(PLAND_{ky}/P_y) + \\ & \alpha_6 URBAN_{ky} + \alpha_7 PJOBGR_{ky} + \alpha_8 PURATE_{ky} + \alpha_9 RSHOCK_{ky} + \alpha_{10} RURATE_{ky} + \\ & \alpha_{11} CRIME_{ky} + \alpha_{12} JAN_k + \alpha_{13} JULY_k + \alpha_{14} TIME_y + \alpha_{15} TIME_y^2 + \\ & \alpha_{16} DS_{iy} + \alpha_{17} DW_{iy} + \alpha_{18} DNC_{iy} + \epsilon_{iy} \qquad (5.5) \end{aligned}$$

---

[4]When the distributions of residual wage and duration are independent, the conditional distribution of the duration given the residual wage is the distribution of the duration itself. Therefore, the likelihood function for estimating the hazard function can be formed based on the conditional distribution.

Table 5.1. Variable names and sample means in the wage equation

Symbol	Variable Description	Sample mean
Dependent variable:		
$Ln(W_{iy}/P_y)$	Log of real hourly wage of individual i in year y (Ln \$/hr)[a]	2.431
Personal characteristics:		
EDU_{iy}	Education of individual i in year y (in years)	12.709
EXP_{iy}	Experience of individual i in year y (Age-EDU-6) in years)	23.597
EXP_{iy}^2	Experience squared/100	6.581
$RACE_i$	Dummy variable, equal to 1 if individual i is white and 0 otherwise	0.929
State labor market:		
$JOBGR_{ky}$	Employment growth rate in state k in year y (percent)	2.257
$UNRATE_{ky}$	Unemployment rate in state k in year y (percent)	6.397
$PJOBGR_{ky}$	Predicted job growth in state k between years y and y-1 (see text)	2.213
$PURATE_{ky}$	Predicted unemployment rate in state k in year y (see text)	6.357
$RSHOCK_{ky}$	Relative employment shock in state k in year y (see text)	0.000
$RURATE_{ky}$	Residual unemployment rate in state k in year y (see text)	0.040
Cost of living and amenities:		
$Ln(PLAND_{ky}/P_y)$	Log of real price of land in state k where individual lives in year y (\$/acres)[a]	1,118.550
$URBAN_{ky}$	Proportion of urban population in state k in year y (percent)	71.771
$CRIME_{ky}$	Crime rate in state k in year y (percent)	8.090
JAN_k	Thirty-year average of January temperature in state k (degrees F.)	33.167
$JULY_k$	Thirty-year average of July temperature in state k (degrees F.)	75.648
Time and regional dummies:		
$TIME_y$	Time indicator, 1968 = 1, ..., 1987 = 20	10.123
$TIME_y^2$	Time squared/100	1.357
DS_{iy}	Dummy variable equals 1 if individual i lives in the South and 0 otherwise	0.290
DW_{iy}	Dummy variable equals 1 if individual i lives in the West and 0 otherwise	0.226
DNC_{iy}	Dummy variable equals 1 if individual i lives in the North Central region and 0 otherwise	0.272
Number of observations		15,367

[a]P_y = Gross National Product (GNP) implicit price deflator for personal consumption expenditure (1987 = 1.00).

The complete definition of variables and the sample mean values are presented in Table 5.1. Equations (5.4) and (5.5) are fitted to annual data for all individuals in the sample, treated cross-sectionally. From the 915 individuals observed over 20 years, we derived 15,367 annual observations.

Predicted job growth in state k in year y ($PJOBGR_{ky}$) is the difference between the forecasted value of the natural logarithm of the state's private sector employment in years y and y-1. The forecasts were obtained from a regression of the natural logarithm of non-agricultural employment for 1968-91 on a quadratic trend. The residuals from these regressions, e_{ky}, are indexes of time-varying local demand conditions in state k in year y. Next, the natural logarithm of national (U.S.) aggregate employment was regressed on a quadratic trend. The residual from this regression, e_y, captures the aggregate labor demand disturbance in year y. The relative local labor disturbance of state k in year y ($RSHOCK_{ky}$) is defined as $RSHOCK_{ky} = e_{ky} - e_y$. This variable expresses the current labor demand shock as a deviation from the aggregate labor demand shock.

The predicted state unemployment rate in state k in year y ($PURATE_{ky}$) measures the anticipated state equilibrium unemployment rate. This rate is obtained by regressing the state's annual unemployment rate for 1968-91 on a quadratic trend. The unemployment rates have not tended to converge over time. At any point in time, there are significant regional unemployment differences that tend to persist. For example, unemployment rates in Michigan and West Virginia are traditionally higher than rates in Iowa and Nebraska. It has been commonly accepted that unemployment rates tend to be counter-cyclical to the movement of the business cycle. The unanticipated unemployment rate is captured by the residual unemployment rate ($RURATE_{ky}$), which is calculated as the difference between the actual and predicted unemployment rates in state k in year y.

The specification of the wage equation does not include correction for selection bias. The men in the sample were screened to remove those who had unusual employment characteristics. This screening process provides a relatively homogenous group that stays in the labor force until retirement. We found that among all the observations used to fit the wage equation, the proportions of one full year of unemployment and a half-year of unemployment are only 0.1 percent and 0.96 percent, respectively. The proportion of reported wages equal to zero was only 4.4 percent of all observations. Because of these small percentages, we decided not to pursue the selection bias correction in fitting the wage equation.

Previous empirical studies provide information about the expected sign of the estimated coefficients. A higher level of education is expected to lead to higher earnings, and therefore the sign of β_1 is expected to be positive. Post schooling experience is expected to have a positive but diminishing effect on the wage. Hence, the expected signs of β_2 and β_3 are positive and negative, respectively. Previous studies have shown that being white significantly increases wages; therefore, the expected sign of coefficient β_4 is expected to be positive.

The local labor market, cost of living, and amenity factors are also expected to affect real wages. The differences in cost of living and local amenities are captured by the price of agricultural land ($PLAND_{ky}$), proportion of urban population ($URBAN_{ky}$), and regional dummy variables (DS_y, DW_y, and DNC_y). To derive the real land price from nominal land prices, we used the GNP implicit price deflator for personal consumption expenditure. In the model, land prices were entered as natural log values. When both land price and the proportion of urban population are positively correlated with wages, the signs of β_5 and β_6 are expected to be positive.

When workers or firms are immobile in the short run, local economic conditions will affect real wage rates. Therefore, state labor market variables are also included in the model. An increase in local labor demand increases market wages. The growth of non-agricultural employment ($JOBGR_{ky}$) in a state is used to capture this effect, and the expected sign of β_7 is positive. Another local labor market characteristic affecting wages is the local unemployment rate. In general, real wages are positively correlated with the local unemployment rate because localities with high unemployment rates pay a wage premium to attract workers. The expected sign of β_8 is also positive. The crime rate could affect the real wage in a similar way to the unemployment rate. The market wage consists of a wage premium to attract people to an area with a higher crime rate. Therefore, the expected sign of β_9 is positive. Following Tokle and Huffman (1991) the average temperatures during January and July are included in the model.

An individual's relative wage performance is defined as the difference between the his actual wage and his national expected market wage. The "residual wage" of individual i in year y is:

$$\Delta W_{iy} = Ln(W_{iky}/P_y) - E[Ln(W_{iky}/P_y)] \tag{5.6}$$

where $E[Ln(W_{iky}/P_y)]$ is the predicted wage obtained from Equation (5.4) and (5.5). By construction, this residual wage is not correlated with other factors affecting the market wage. The hypothesis is that a larger residual wage or better local wage performance will reduce the tendency to migrate and therefore lengthen a residence spell. In other words, when the actual wage received by an individual is less than his potential wage, he tends to move to another state. This residual wage is calculated and used in the right-hand side of the equation explaining migration. For the hazard function model, the residual wage is expressed as an annual average over a residence spell.

Because of the following characteristics, the residual wage is expected to be an important determinants of migration. First, the residual wage represents luck. The residual wage consists of a random disturbance, which explains why two persons having the same state of residence (i.e. state attributes) may have different wages. Second, the residual wage captures the unmeasured state-specific returns to factors affecting wage rates, which might differ from the general return for the United States. A positive residual wage means that the state-specific return is on average greater than the general U.S. return, so that the individual tends to stay at the current residence. When the residual wage is negative, an individual has an incentive to migrate. Therefore, the residual wage captures the unmeasured wage differentials across states which affect migrations.

Third, the residual wage incorporates tenure effects due to firm or location specific human capital associated with the current residence or the origin. Because the wage increases with the increase with the length of tenure, the probability of an individual moving decreases as his tenure accumulates in a given state. Furthermore, the location or firm- specific tenure is useful in explaining migration in the time dependent hazard rate model. The significance of the tenure effect on the residual wage can be shown by fitting a wage equation that includes a variable representing tenure. The issue is, however, more complicated because tenure specific to location or firm is endogenous to the wage. Because high-paying jobs usually last longer, more stable jobs pay more throughout the tenure within a firm, and more able workers generally stay on the job longer and receive higher wages, the pure effect of job tenure on wage is difficult to measure (Topel, 1991; Brown, 1989; Abraham and Farber, 1987). For illustration, this study shows the significance of state specific tenure on the residual wage by fitting the wage equation that includes a variable representing tenure in the model. The result can be found in Appendix C.

The Probit Equation

The standard discrete choice model is fitted to the data from 1968 to 1973. The data from the first six years were chosen in order to maintain the random selection in 1968 and to maximize the sample size. From 1968 through 1973, an average of 20 to 30 individuals out of 915 individuals (around 3 percent) in the sample migrated to another state every year. Only five or fewer individuals migrated in two consecutive years. In one year of data, the size of the migrant sample is too small. Previous migration studies used one or two years of data (e.g., Goss and Paul, 1986; Graves and Linneman, 1979). Because migration was defined as any move from the current residence, these studies treated local moves and long-distance moves equally and therefore the size of migrating sample was large. Because of the lack of data on local characteristics, these studies incorporated only personal characteristics in the models.

To increase the number of migrants, this study combined observations from the first seven years of data (1968-1973). Because the first year that an individual is observed can be as late as 1972, the year 1973 was chosen to capture every individual in the sample. To maintain the model of a point-in-time decision, the definition of migration was modified. We defined the dummy variable representing migration (M) as equal to 1 if individual moved at least once between 1969 and 1973. When an individual lived in the same state for the 1968-73 period, we set M equal to 0. As a result, the proportion of the sample with M = 1 increases to 16.4 percent.

The set of explanatory variables used to explain the migration behavior contains personal characteristics of individuals and local characteristics of the place the individuals lived before they made the final move. For those who did not move, the data are taken from the state of origin. For those who moved at least once, the data are taken from all the states the individuals lived in prior to the most recent move. When the expected utility of being in the new state is not realized, an individual will most likely make another move. In this framework, the move that counts is the last move the individual made.

The local variables that directly affect migration behavior, such as crime, local amenities, and temperature, are measured relative to the U.S. average. For example, the crime rate that is used in the migration equations, both probit equation and hazard function, is the state crime rate minus the United States average crime rate. This relative measurement is important because including the realized destination into the discrete choice model would lead to misspecification error.

Some variables are measured as an average value of several years of observations, whereas others are measured at the beginning of the period of observation. Based on Equation (4.5) in the previous chapter, the probability of migration for individual i is specified as

$$\begin{aligned} \mathrm{Prob}(M_i = 1) = F(&\alpha_0 + \alpha_1 \Delta W_i + \alpha_2 EDU_i + \alpha_3 AGE_i + \alpha_4 AGE^2{}_i + \alpha_5 CHILD_i + \alpha_6 MARR_i \\ &+ \alpha_7 DSLFARM_i + \alpha_8 RACE_i + \alpha_9 UNEMP_i + \alpha_{10} UNION_i + \alpha_{11} CRIME_i \\ &+ \alpha_{12} PARK_i + \alpha_{13} JAN_i + \alpha_{14} JULY_i) \qquad (5.7) \end{aligned}$$

The complete definition of the variable and the sample means are presented in Table 5.2. As indicated in the previous section, lower wage performance is expected to increase the tendency to migrate. Therefore, the expected sign of α_1 is negative. Table 5.2 shows that migrants have negative means for residual wages, meaning that their realized wage at the origin has been lower than their potential wage. In contrast, the means for the nonmigrant group are positive.

An individual's general education is expected to have a positive effect on migration rate. Men who have higher levels of general education participate in a larger labor market and receive more job offers. Previous studies on migration have consistently shown that education has a positive effect on migration rates.

Table 5.2. Variable names and sample means for the probit equation

Symbol	Variable Description	Sample Mean: Migrant ($M_i = 1$)	Sample Mean: Non-migrant ($M_i = 0$)
ΔW_i	Residual wage measured as an average per year in log (one of two types):		
	1. $\Delta W1_i$, derived from Equation (5.4)	-0.768	0.040
	2. $\Delta W2_i$, derived from Equation (5.5)	-0.782	0.044
EDU_i	Education measured at the starting time of observation (in years)	12.569	12.407
AGE_i	Age measured at the starting time of observation (in years)	30.050	33.818
$CHILD_i$	Number of children in school measured at the starting time of observation	0.916	1.461
$MARR_i$	Share of time being married[a]	0.455	0.905
$DSLFARM_i$	Dummy variable, equal to 1 if farmer or self-employed and 0 otherwise	0.104	0.187
$RACE_i$	Dummy variable, equal to 1 if white and 0 otherwise	0.911	0.913
$UNEMP_i$	Average annual unemployment hours (hr/year)	33.465	32.385
$UNION_i$	Dummy variable, equal to 1 if in a labor union at the beginning of observation and 0 otherwise	0.243	0.288
$CRIME_i$	Crime rate in the state of origin, relative to U.S. average	-0.052	-0.417
$PARK_i$	Percentage of state and national park areas in the state of origin, relative to U.S. average	0.027	-0.080
JAN_i	Average January temperature in the state of origin, relative to U.S. average	3.363	0.501
$JULY_i$	Average July temperature in the state of origin, relative to U.S. average	0.070	0.130
Number of observations		202	713

[a]The length of time being married and living in the state of origin divided by the length of time being in the state of origin. For example, when individual remains married over the period living in the state of origin, MARR = 1.

The signs of the coefficients for AGE and AGE^2 are expected to support the view that younger men have a higher tendency to migrate. Age reduces the net present value of the benefit from moving by reducing the benefit and increasing the cost. Young men have a longer horizon after migrating and have better chance of overcoming the costs of migration than do older men. The mean age of the sample male in Table 5.2 shows that the sample of migrants is younger than the sample of for nonmigrants.

For a male, the number of children in the school, being married, being self-employed or a farmer, and belonging to a labor union, are characteristics expected to reduce the probability of migrating. The cost of migration is higher for those who are married and have school-age children. Tenure benefits accrue with union membership tends to reduce the probability of migration. Farmers or self-employed individuals tend to have firmer ties to the origin. Therefore, being a farmer or self-employed is expected to reduce the probability of moving. The means shown in table 5.2 support these expectations. The sample of migrating males has a smaller proportion of self-employed individuals or farmers, smaller number of school-age children, and smaller share of time married than does the sample of nonmigrants. Table 5.2 also shows that the sample of migrants has a smaller proportion of men who are union members than the sample of nonmigrants. Most previous studies (see Table 1.3) have indicated that being white or unemployed increases the chances of moving. Being unemployed may reduce utility and provide some incentive to move and find a job elsewhere. Table 5.2 shows that the migrants have larger mean of average annual unemployment hour.

Locational amenities affect compensating differentials and the tendency to move. The sign of the coefficient on PARK is expected to be negative. The coefficient on CRIME is expected to be positive because a higher crime rate at one's current location reduces the cost of moving. Nonetheless, the difference between the migrant and the nonmigrant sample means for those two variables seems to be small. The average temperatures in July and January are included to capture the effect of weather-related factors affecting the migration decision.

The Hazard Function

Search theory provides us with a hazard function for migration. For a given length of time, it captures the instantaneous probability that an individual will receive an acceptable offer and move. We can specify a reduced form hazard function for individual i and residence spell j as a function of personal and local characteristics. Let the explanatory variable corresponding to that spell be X. From the previous chapter, the general form of the hazard function is given as:

$$H[t_{ij}, X(t_{ij}), \beta, \alpha, \sigma] = \lambda/\sigma(\lambda t)^{1/\sigma - 1} \quad (5.8)$$

where

λ = $\exp[X(t_{ij})\beta]$,

σ = time-dependence parameter. The hazard rate increases, stays constant, or decreases over time when σ is less than, equal to, or greater than 1,

i = 1, 2, ..., n, denotes individual i,

j = 1, 2, ..., C_i, denotes the spell j, and

C_i = maximum number of spells for individual i.

In general, the hazard function corresponding to spell t_{ij} is a function of variables representing the reduced form indirect utility function evaluated at the current residence, the cost of migration, and other variables that capture a person's opportunity of receiving an offer to move.

Search theory and previous migration studies discussed in the literature review provided some ideas for variables. Education and experience are two personal characteristics that help explain an individual's income and also directly influence the decision to migrate. People with higher education tend to receive more offers to move because of greater employment opportunity. Based on the comparative static equation discussed in Chapter 2, an increase in the rate at which offers are obtained leads to a higher reservation utility and smaller hazard rate to migrate. In addition, more educated people have a higher wage offer. Through the change in the expected utility offer, the comparative static result shows that higher education or wage offers increase the migration hazard rate. Highly educated people generally face a larger labor market than do others. Therefore, educated individuals are more likely to move between states. Education can also have nonwage effects due to gains in the efficiency of acquiring and processing information (Huffman, 1985).

Other personal characteristics affecting the cost of migration are included in the model. Being married, the number of children in school, and being self-employed or a farmer could account for some migration costs. The expected utility, net of migration cost, due to migration will be smaller for households who have children in school and are self-employed than for others. The racial characteristic is expected to matter and is included in the migration model.

Studies reviewed in previous chapters have shown the significance of personal unemployment on the migration decision. These studies have found that unemployed people have a higher tendency to migrate. This study included a variable representing personal unemployment (annual average unemployment hours) in the model. The indication of whether an individual belongs to a labor union is also included in the model. In addition, variables representing local amenities, such as the proportion of area in state and national parks, crime rate, and weather, are included.

When the length of a spell is endogenous, explanatory variables that vary over time might be endogenous. To reduce endogeneity, these explanatory variables can be measured at the beginning of each residence spell. Finally, the hazard function $H[t_{ij}|X(t_{ij}),\beta,\alpha,\sigma] = \lambda/\sigma(\lambda t)^{1/\sigma-1}$ can be formed by specifying $X(t_{ij})\beta$ in $\lambda = \exp(X(t_{ij})\beta)$ as follows:

$$X(t_{ij})\beta = \beta_o + \beta_1 \Delta W(t_{ij}) + \beta_2 EDU(t_{ij}) + \beta_3 AGE(t_{ij}) + \beta_4 AVUNHR(t_{ij}) + \beta_5 UNION(t_{ij}) + \beta_6 DSLFARM(t_{ij}) + \beta_7 DWHITE(t_{ij}) + \beta_8 MARR(t_{ij}) + \beta_9 CHILD(t_{ij}) + \beta_{10} CRIME(t_{ij}) + \beta_{11} PARK(t_{ij}) + \beta_{12} JULY(t_{ij}) + \beta_{13} JAN(t_{ij}) \tag{5.9}$$

Detailed definitions and the sample means of variables in the equation can be found in Table 5.3. The means are derived from the completed and censored spells. The censored spells in Table 5.3 consist of left-censored, right-censored, or both right- and left-censored spells. A left- and right-censored spell means that an individual lived in the same state throughout the observation period, and the exact starting year when he lived in the corresponding state is unknown. In the estimation of hazard functions with a constant hazard rate model, left-censored spells can be treated as a right-censored spell. In the time dependent hazard model, all left-open spells must be closed. When all the left-open spells are closed, the number of completed spells will increase. The detailed procedure for dealing with the time-dependent model is discussed in the next chapter.

Previous discussions provide some indication of expected signs of the parameters. Because individuals tend to move when their stream of income is not satisfactory, the residual wage will be negatively correlated with the hazard rate to migrate. Therefore, the expected sign of β_1 is negative. Education is expected to have a positive effect. Younger people have a higher tendency to migrate, so that the sign of the coefficient of AGE20 is expected to be positive. AGE20 is the time being in the one's 20s in a residence spell divided by the length of the corresponding residence spell. For example, when age at the first year of a residence spell is 25 and the residence spell is 8 years, AGE20 = (30 - 25)/8.

Table 5.3. Variable names and sample means for the hazard function

Symbol	Variable Description	Sample Means	
		Completed Spells	Censored Spells
$\Delta W(t_{ij})$	The average annual differences between actual and predicted wages in spell t_{ij} (one of two types):		
	1. $\Delta W1(t_{ij})$, derived from Eq. (5.4)	-0.183	-0.159
	2. $\Delta W2(t_{ij})$, derived from Eq. (5.5)	-0.199	-0.161
EDU(t_{ij})	Variable representing the education of individuals (one of two types):		
	1. EDUBEGIN(t_{ij}) = level of education at the beginning of a spell	14.043	12.544
	2. AVGEDU(t_{ij}) = average level of education in the spell	14.081	12.573
AGE(t_{ij})	Variable representing individual's age effect (one of two types):		
	1. AGE20(t_{ij}) = share of time being in one's 20s	0.350	0.174
	2. AGEBEGIN(t_{ij}) = age at the beginning of the spell (year)	34.473	34.341
AVUNHR(t_{ij})	Average annual unemployment hours in the spell (hr/year)	70.474	42.855
UNION(t_{ij})	Dummy variable, equals to 1 if an individual belongs to union or 0 otherwise Membership is defined in two ways:		
	1. UNIBEGIN(t_{ij}) = observed at the beginning of a spell	0.135	0.258
	2. DUNION(t_{ij}) = observed in any time in a spell.	0.184	0.390
DSLFARM(t_{ij})	Dummy variable, equal to 1 if self-employed or a farmer and 0 otherwise	0.150	0.315
RACE(t_{ij})	Dummy variable, equal to 1 if white and 0 otherwise	0.986	0.921
MARR(t_{ij})	Share of time being married in a residence spell	0.713	0.796
CHILD(t_{ij})	Number of children in school age (one of two types):		
	1. AVCHILD(t_{ij}) = annual average number in a spell	0.723	1.171
	2. CHILDBGN(t_{ij}) = number at the beginning of a spell	0.766	1.305
CRIME(t_{ij})	Average annual crime rate in the state corresponding to residence spell, relative to the U.S. average	0.195	-0.472
PARK(t_{ij})	Percentage area of state and national parks in the state corresponding to residence spell, relative to U.S. average	-0.116	-0.050
JULY(t_{ij})	Average July temperature corresponding to residence spell, relative to U.S. average	1.382	0.138
JAN(t_{ij})	Average January temperature corresponding to residence spell, relative to U.S. average	4.781	1.563
Number of spells		207	1,061

Because the effect of individual age on migration is not linear, the age at the beginning of a spell will enter the hazard function in both a linear and a quadratic form (AGEBEGIN and $AGEBEGIN^2$).

The coefficients on personal unemployment and crime rate are expected to be positive because, all else being equal, being unemployed or living in an area with a high crime rate tends to increase the probability of moving. Table 5.3 shows that the sample of completed spells has significantly larger means for average annual unemployed hours than do the censored spells. Other variables which capture the cost of migration, such as an individual being married, being self-employed or a farmer, having children in school, and being a labor union member, tended to decrease the hazard rate of migration.

In the model with a constant hazard rate, the estimation of the hazard rate is straightforward. We assumed that the completed duration is exponentially distributed and treat the left-censored spells similar to the right-censored spells. In other word, all the completed spells start in 1968 or after. We could derive 1,268 spells, with 207 spells completed and 1,061 censored. The estimation for the time-dependent hazard rate is more complicated and the procedures are presented next.

Treatments for the time-dependent hazard function

The estimation of a time-dependent hazard model requires that every spell be completed or right-censored. The model does not allow for any left-censored spells. In this study, the existence of left-censored spells is unavoidable because every individual has been in the state for some time prior to 1968. However, the PSID does not have information for how long individuals lived in the state prior to 1968. Because there has been no standard procedure for dealing with the issue of left-censoring, this study presents some empirical strategies, or treatments, to close some of the spells. Each treatment has its benefits and drawbacks. It is important to note that these treatments should be considered as exploratory procedures.

For the first treatment, we selected only completed and right-censored spells that started in 1968. In other words, we ignore the spell starting before 1968. This selection led to 207 completed and 184 right-censored spells. The benefit of this treatment is that all information for estimating the hazard rate, namely, the duration and all explanatory variables, are readily observable. The problem, however, is nonrandom selection.

For the second treatment, we tried to utilize all completed and right-censored spells including those starting before 1968. We assumed that those residing in the same state where they grew up in 1968 had remained in that same state up to 1968. This assumption means that the starting time for the spells is the year when individuals were age 19. When an individual moved before 1968 from the state where he grew up, the starting time of the first spell, and thus the duration of the spell remained unknown. In this second treatment, we did not include these unobservable spells in the estimation. The second treatment has 284 completed and 581 right-censored spells. The 284 completed spells consists of 207 spells starting in 1968 or after, and 77 spells starting before 1968. Compared to the first treatment, the number of completed spells increased significantly with the second treatment. Therefore, the problem of nonrandom selection associated with the first treatment is reduced.

There is still a problem because some of the explanatory variables, especially the actual wage, were not observed before 1968. The best information available is the value of explanatory variables in the first spell derived from the data in 1968 or earlier. The use of partially observed explanatory variables for some spells that started before 1968 could lead to measurement error problems. To reduce the measurement error problem, we fit the model twice. First, we used the spells derived from all individuals in the sample. Second, we selected only individuals who were 19 to 24 years of age in 1968. It is important to note that spells can not start until the individual is 19 years old. With the younger sample used in the analysis, the starting time of the first spell could go back as far as 1963.[5] This approach decreases the number of partially observed spells and reduces measurement error problems without making nonrandom selection error. The use of a younger population resulted in 137 completed and 141 right-censored spells.

For the third treatment, we predicted the starting time for the left-censored spells and then use all spells in estimating the time-dependent hazard function. The starting time is unknown because, prior to 1968, some individuals had migrated from the state where they grew up. In other words, the starting times for some first spells were observed and some were not. When some data are censored, we can apply Heckman's procedure to predict the unobserved starting time for those first spells (Heckman, 1979). The idea of Heckman's procedure is to predict the unobserved length of spells based on the observed group, adjusted to the sample selectivity. The selection criteria is the migration which took place before 1968.

There are 874 spells starting before 1968 in the sample. Among them, 585 are right-censored and 289 are right-closed spells. In the right-censored group, the starting time for 127 spells is unknown. Thus, 127 spells are both right- and left-censored. In the right-closed group, the starting time of 138 out of 289 spells is unknown. These 138 spells are left-censored but right-closed. Therefore, 265 potentially unknown spells out of the total 1,268 are to be estimated. Because the right-closed and the right-censored spells came from different populations, we applied Heckman's procedure twice, first for the group of right-closed spells and second for the group of right-censored spells. Note that the right-censored first spells are derived from individuals who never moved until

[5]The young males sample consists of young men in 1968. In other words, the young males sample were observed since they were 19 to 24 years of age in 1968 or until they were 38 to 43 years of age in 1987.

the end of the period of observation.

To be more specific, the procedure used to predict the unobserved starting time or the unobserved length of the first spell consists of several steps. First, we fit the equation representing the selection model for the distinction between observed and unobserved length, or duration, of the spell. Basically, we estimated a binary discrete choice model explaining why people moved or stayed prior to 1968.[6] We could use the model for migration based on the point-in-time decision developed in Chapter 2. With a normality assumption, we fit the standard probit model with a dependent variable equal to 1 if an individual had moved from the state where he grew up before 1968 and 0 otherwise. The explanatory variables are the relevant personal characteristics and local characteristics observed in the first year the person was observed and other family background characteristics such as father's education and number of siblings. The detailed variable definitions used in this selection model and the fitted equation can be found in Table D.1. in Appendix D.

Having predicted when migration occurred before 1968, we derived the selection bias correction, known as the Inverse Mill's Ratio (IMR). Then, by using the observed sample, we regressed the duration of the spell as a function of some explanatory variables with the sample selection bias correction[7]. Finally, we used the fitted equation from the observed sample to predict the duration of the spell for the unobserved sample.

---

[6]Pr (M68 = 1) = $X\beta$, where M68 = dummy variable equals to 1 if individual did not live in the same state where he grew up in 1968, and 0 otherwise, X = explanatory variables, and β = parameters.

[7]$Ln(t) = Z\alpha + \theta\lambda + \epsilon$, where t = length of residence spells, Z = explanatory variables, λ = Inverse Mill's Ratio (IMR) for selection bias correction, ϵ = random disturbance, and α, θ = parameters. The IMR is derived from the selection equation as $f(X\beta^*)/F(X\beta^*)$, where f(.) and F(.) are normal density and cumulative distribution function, and β^* is the estimate of β.

When the predicted starting time was the year before the person was 19 years old, we used the year when the person was 19. Also, when the predicted starting year was the year after the person was observed in 1968 or after, we used the year when the person was first observed as the starting year. In other words, we used the predicted spells only when necessary. Indeed, from the total of 265 potentially unknown spells, 208 are really predicted. The definition of variables used to predict the left-censored spells and the fitted equation can be found in table D.2 in Appendix D.

With Heckman's procedure, the predicted length is theoretically distributed normal. Some complications occur because, in the time-dependent hazard rate model, we have assumed that the distribution of the completed duration is Weibull. One reasonable strategy is to test whether the predicted duration of the completed spells is approximately distributed Weibull. We conducted the goodness-of-fit test on the distribution of the predicted duration. When the hypothesis that the predicted length of the spell is distributed Weibull can not be rejected, the predicted duration can be used to estimate the hazard function and the error associated with it is reduced.

Finally, similar to the second treatment, to reduce the measurement error problem, we estimated the time-dependent hazard rate model using the third treatment twice. First, we used the full sample and derived a total of 1,268 spells, with 496 spells completed. Compared to the estimation for the exponential model, the number of completed spells increased significantly from 207 to 496. Second, we selected a younger population to reduce the measurement error derived from unobserved explanatory variables prior to 1968. The selection of a younger sample, namely, the males who were 19 to 24 years old in 1968, led to 350 spells, with 191 spells completed.

CHAPTER 6. THE WAGE EQUATION AND THE PROBABILITY OF MIGRATION

The fitted wage and migration equations are presented and evaluated in this chapter. The first section presents the fitted wage equation, which will serve as an instrument in the migration equation for the actual wage at the origin. The second part presents the equation explaining the probability of migration based on the point-in-time discrete choice model.

The Potential Wage Equation

Wage equations for every individual in every year were fitted based on the 915 individuals in the sample observed over 20 years (1968-87). The data on individual characteristics were merged with the data on state characteristics. Treated cross-sectionally, there are 15,367 individual annual observations. The equations were fitted using ordinary least squares with the log of the real hourly wage as the dependent variable in the model. This wage equation is basically a wage offer equation derived from the demand side of the labor market. In general, the results are similar to those in many previous studies for labor demand of individuals. As discussed in the previous chapter, the specification does not include selection bias correction in the model. The estimates of two potential wage equations corresponding to Equations (5.4) and (5.5) are presented in Table 6.1; one with the actual and the other with the predicted state labor market characteristics.

Except for the coefficients on state labor market characteristics, the estimates of the other coefficients look similar between the two models. Because of more regressors, a slightly higher R-square result for the second equation is expected. In general, all coefficients have the expected sign and are significantly different from zero at the 5 percent level. More than the other coefficients, human capital played a major role

Table 6.1. Potential wage equation (t-values are in parentheses)

	With Actual State Characteristics (PW1)		With Predicted State Characteristics (PW2)	
INTERCEPT	1.517	(9.01)	1.350	(7.79)
EDU	0.075	(34.20)	0.075	(34.07)
EXP (AGE-EDU-6)	0.051	(20.80)	0.051	(21.83)
$EXP^2/100$	-0.098	(-19.46)	-0.098	(-19.62)
RACE	0.112	(4.80)	0.110	(4.73)
Ln(PLAND/P)	0.044	(2.81)	0.048	(3.07)
URBAN	0.279	(3.88)	0.295	(4.10)
JOBGR	0.015	(5.50)	-	
UNRATE	0.015	(3.60)	-	
PJOBGR	-		0.060	(7.46)
PURATE	-		0.043	(7.15)
RSHOCK	-		0.009	(2.92)
RURATE	-		-0.005	(-1.06)
CRIME	0.016	(6.85)	0.015	(5.94)
JAN	-0.002	(-1.59)	-0.004	(-3.39)
JULY	-0.018	(-11.88)	-0.017	(-10.42)
TIME	0.010	(1.79)	-0.012	(-1.88)
$TIME^2/100$	-0.093	(-4.07)	0.008	(-0.29)
DS	-0.061	(-2.27)	-0.089	(-3.03)
DW	-0.068	(-2.49)	-0.137	(-4.47)
DNC	-0.006	(-0.32)	-0.018	(-0.92)
n	15,367		15,367	
R-square	0.154		0.158	

in determining an individual's wages. The coefficient for human capital characteristics are different from zero at the 1 percent level. The significance of state characteristics in the wage equation shows that locality matters in determining individual wages.

We found evidence that increased schooling and work experience lead to a higher wage. A one-year increase in schooling increased the real wage by about 7.5 percent. This magnitude is consistent with results reported in previous studies (Mincer, 1974; Topel, 1986). Using a sample from four CPS data sets, Tokle and Huffman (1991) found a 5.5 percent rate of return on schooling for non metropolitan males and 7.1 percent rate of return for females.

An increase in an individual's experience had a positive effect on his real wage but at a diminishing marginal rate. The maximum effect occurred at about 26 years of experience (approximately 45 years of age) for the PW1 and PW2 equation, respectively. This pattern has been reported in many studies. Topel (1986) found the maximum at 33 years of experience. Compared to the results from Tokle and Huffman (1991), our results show a higher rate of return to schooling and the wage peaks at lower years of experience. Perhaps some of these differences are due to differences in the composition of the two samples. This study used male heads of households who were age 19-45 in 1968, while Tokle and Huffman focused on rural non farm married couples.

The results of this study also support previous finding that racial differences affect wage rates. All other measured variables held equal, the white males in this study earned about 11 percent more compared to the nonwhite males. For comparison, Topel (1986) found an 18 percent difference in wages between white and nonwhite workers. Tokle and Huffman (1991) found about a 20 percent difference using the sample of rural nonfarm households.

Wage rates also differed because of cost of living and local amenity differences. Both the real price of land (PLAND/P) and the proportion of urban population (URBAN) in a state are positively related to the real wage rate. The elasticity of the real price of land on wage rates is around 0.04 to 0.05. A 1.0 percent increase in the proportion of the state urban population contributed to about a 28 percent to 30 percent increase in real wage. These elasticities are consistent with those found in previous studies. Tokle and Huffman (1991) found the elasticity of the real price of land around 0.06 to 0.07. They also found an 18 percent to 25 percent increase due to a 1.0 percent increase in the urban population. These results support the view that wage rates in different localities should represent the difference in the prices of traded and nontraded goods. The price of traded goods between two areas should differ by the amount of transport cost. The price of nontraded goods such as housing, however, should differ by more.

State labor market characteristics had a significant effect on an individual's wage. All the coefficients on state labor market characteristics, except for the unanticipated unemployment rate, are significantly different from zero at the 5 percent level. The signs of the coefficients are as expected and provide empirical evidence to support the hypothesis developed earlier. The real wage rate seems to have incorporated compensation for state labor market condition.

State employment growth has a strong positive effect on wages. Our results show that a 1.0 percent increase in actual state employment growth led to a 1.5 percent increase in wage rates. Real wage rates seem to be more responsive to predicted state labor market characteristics than to actual characteristics. A 1.0 percent increase in predicted state job growth led to a 6.0 percent increase in the real wage rate. Employment in a state can be viewed as the result of a supply and demand equilibrium condition. When demand grows faster than supply, the equilibrium wage tends to increase. In addition, the wage premium in localities having a higher expected growth of labor demand compensates for the cost of geographical mobility. These results are stronger than those obtained by Tokle and Huffman (1991).

Actual and anticipated state unemployment rates had a positive effect on wage rates. A 1.0 percent increase in the actual unemployment rate raised wage rates by 1.5 percent. Topel (1986) found that wages increased by 2.1 percent due to a one unit increase in the probability of personal unemployment. Similar to the effect of employment growth, the effect of predicted or anticipated state unemployment rates was also found to be stronger than the actual values. A 1.0 percent point increase in the predicted state unemployment rate caused a 4.3 percent increase in the real wage rate. This increase is greater than the result obtained by Topel (1986), who found a 1.1 percent increase. For comparison, Tokle and Huffman (1991) obtained a coefficient of 1.2. The result supports the view that localities with higher unemployment rates must pay a wage premium to entice current workers to stay and to risk losing a job.

The significant effect of unanticipated labor market shocks on real wages can be seen in the second equation in Table 6.1. The coefficients for RSHOCK and RURATE are significantly different from zero at the 5 percent and 10 percent levels, respectively. This result means that the real wage rate responded to unanticipated state labor market conditions. The signs of the coefficients are as expected and are consistent with the result of Tokle and Huffman (1991). The underestimated employment growth rate and overestimated unemployment rate seem to increase the wage rate. Topel (1986) also found a positive effect of unpredicted state employment disturbance on the real wage rate.

The wage equation also shows that the crime rate is a significant determinant of the local wage. The coefficient on the crime rate is positive and significantly different from zero at the 1 percent level. A 1.0 percent increase in the crime rate led to about a 1.5 percent increase in the real wage rate. States with high crime rates must pay a premium to attract and keep workers. These results are consistent with earlier work by Roback (1982).

The significance of the coefficient for JAN and JULY shows that normal state climate affects wage rates. It is realized that these temperature variables are correlated with the regional dummy variables, so that the best measure of their effect probably occurs when the regional dummies are excluded. Nevertheless, the inclusion of the regional dummies does not seem to reduce the significance of the temperature variables. The effects of July temperatures on the real wage seem to be stronger than that the effects of January temperatures. The general indication that an increase in July temperature is associated with a decrease in real wage rates is consistent with results obtained by Tokle and Huffman (1991).

The long history that wages tend to be lower in the southern states is supported by this study. Compared to the northeastern states, wages in the southern states are about 6.1 percent to 8.9 percent lower. Residing in a western state also reduces the wage by 6.8 percent to 13.7 percent. However, wage rates in the northern-central states are not significantly different from those the northeastern states. The significance of the coefficients of the regional dummies shows the importance of unmeasured regional effects on the real wage.

The estimated coefficients on time trends in general show that the real wage has decreased since the 1970s. In the first equation using actual labor market characteristics, the real wage rate increased until reaching a peak in the early 1970s and then declined. In the second equation, however, the wage rates seem to decrease over time. The difference in the estimate between the first and second equations in Table 6.1 is expected. The use of predicted values of state labor market characteristics, which are derived by regressing the actual values on a quadratic time trend, has picked up the trend component in the wage equation.

In summary, human capital variables have strong effects on the real wages of individuals. Education and experience play major roles in determining wages. In this study, the white males earned more than the nonwhite males. The difference in cost of living also affected wages. Areas with higher costs of living offered higher wage rates. State labor market conditions, represented by unemployment rate and employment growth rate, were found to be important determinants of the real wage. All other things held equal, localities with higher unemployment rates and faster employment growth rates offered higher wage rates. In general, the real wage has declined for the last 20 years. The southern and western states experienced lower wage rates compared to the northern-central and eastern states.

The Probit Equation

The probit model is the standard discrete choice model for explaining the tendency to migrate. The probit equation is fitted to the data for the 915 individuals in the sample during the first six-year period (1968-73). It is important to note that this selection process led to the use of young individuals in fitting migration equations. The sample consists of males who were 19 to 29 years of age over 1968-73 period. During the period, 202 individuals (22.08 percent) in the sample moved to a different state. Recall that multiple-year data is used to increase the number of migrants in the sample. The estimated migration equation showing the effect of explanatory variables on the probability of migration can be seen in Table 6.2. Because the equation is nonlinear, the marginal effect of each regressor is calculated based on Equation (4.9) and presented in Table 6.3. Previous studies of migration decisions have used a similar framework. In general, the signs of the coefficients are as expected and consistent with the results of earlier migration studies.

Table 6.2. Probability of migration equation (t-values are in parentheses)

Explanatory Variables	With Actual State Characteristics (PW1)		With Predicted State Characteristics (PW2)	
INTERCEPT	1.541	(1.24)	1.520	(1.22)
ΔW1	-0.650	(-6.66)	-	-
ΔW2	-	-	-0.673	(-6.84)
EDU	0.063	(3.31)	0.063	(3.33)
AGE	-0.139	(-1.81)	-0.140	(-1.82)
AGE^2	0.156	(1.33)	0.157	(1.34)
CHILD	0.032	(0.66)	0.033	(0.68)
MARR	-1.123	(-7.02)	-1.103	(-6.86)
DSLFARM	-0.473	(-2.96)	-0.480	(-2.90)
RACE	0.403	(1.82)	0.411	(1.85)
UNEMP	0.0004	(0.70)	0.0005	(0.78)
UNION	0.051	(0.38)	0.062	(0.46)
CRIME	-0.001	(-0.31)	-0.001	(-0.17)
PARK	0.009	(0.74)	-0.001	(-0.03)
JAN	0.003	(0.56)	0.002	(0.36)
JULY	0.013	(1.02)	0.012	(0.90)
Log likelihood	-337.9		-336.4	
Number of migrants	202		202	
Number of observation	915		915	

Table 6.3. Marginal effect on the probability of migration, evaluated at the mean of the explanatory variables, based on table 6.2

Explanatory variables	Unit	Using Actual State Characteristics (PW1)	Using Predicted State Characteristics (PW2)
ΔW1	percent[a]	-0.168***	-
ΔW2	percent	-	-0.168***
EDU	year	0.016***	0.016***
AGE	year	-0.036*	-0.036*
CHILD	0, 1, 2, ...	0.008	0.008
MARR	(0-1)[a]	-0.291***	-0.283***
DSLFARM	Dummy:0,1	-0.120***	-0.120***
RACE	Dummy:0,1	0.104*	0.103*
UNEMP	hour	0.0001	0.0001
UNION	Dummy:0,1	0.014	0.016
CRIME	percent[b]	-0.0003	-0.0003
PARK	percent[b]	0.002	-0.0003
JAN	degrees F[b]	0.001	0.001
JULY	degrees F[b]	0.003	0.003

[a]The share of time spent married ranges from 0 to 1

[b]Measured relative to U.S. average

*Significant at the 10 percent level

***Significant at the 1 percent level.

For this study, let's begin by examining the wage effect. We found that an individual's actual wage relative to his predicted wage in the state is a strong determinant of the migration decision. The coefficients for $\Delta W1$ and $\Delta W2$, are negative and significantly different from zero at the 1 percent level. The use of $\Delta W1$ and $\Delta W2$, based on PW1 and PW2, does not lead to any significant difference in the estimated coefficients. A 1 percent decrease in the actual wage relative to the potential wage increased the probability of migration by 0.17 percent. We found strong empirical evidence that the tendency to migrate increases when individuals earn less than their potential locally.

A study by Goss and Paul (1986) used individuals' earnings at their current residence to explain migration. They found a similar negative effect. There is, however, an empirical problem with putting the actual earnings of individuals directly into the migration equation because actual earnings are correlated with other explanatory variables affecting migration, especially education and age. By using the "residual wage", we get free from this problem. Therefore, the use of $\Delta W1$ and $\Delta W2$ in this model is more appropriate. We conclude that the development of a new variable representing individual wage performance has led to a new and meaningful result.

An individual's education was also found to be an important determinant of migration. The coefficient on education is positive and significantly different from zero at the 1 percent level. In other words, higher education leads to a higher tendency to migrate. A one-year increase in education contributed about a 1.6 percent increase in the probability of migration. This marginal effect is stronger than the effect found by Mincer (1978). This result is consistent with result showing that more education experience lowers the cost of acquiring information which reduces migration. The view that migration and education are complementary investments (Schwartz, 1976) is also supported. In fact, a higher probability of migration for more educated individuals is common in previous migration studies (Pissarides and Wadsworth, 1989; Graves and Linneman, 1979; Goss and Paul, 1986).

The positive effect of general education on migration can also be related to Becker's model of general and specific training (Becker, 1975). He mentioned that specific training increases the productivity of a worker in his/her current job. General training increases a worker's productivity not only in the current work place but also in other firms. Therefore, workers who have greater general skills from general education can benefit from a larger labor market and have a higher tendency to migrate. The result that general education increases the probability of migrating in this study fits well with the predictions from Becker's theory.

In general, we found that as an individual gets older, he experiences a lower probability of migration. The signs of the coefficients for age and age-squared in Table 6.2 are negative and positive, respectively. Evaluated at the mean of age, the marginal effect of an added year of age on the probability of migration is around -3.6 percent. This result is consistent with the theory that individuals have less tendency to migrate when they get older. It also supports the theory that migration is an investment. As an individual becomes older, he/she faces a reduced net present value of migration because of increased costs and reduced benefits. Young people have a better chance of recovering the cost of migration because they live longer after the migration takes place. Previous studies have commonly put only a linear effect of age in migration equations. In general, they have found a negative estimate of the coefficient for age (Pissarides and Wadsworth, 1989; Schlottmann and Herzog, 1981; Graves and Linneman, 1979).

The coefficient for RACE is positive and significantly different from zero at the 10 percent level. The marginal effect shows that being white increases the chances of migration by 10 percent over being nonwhite.

Whites seem to be more likely to conduct long-distance migration, such as between-states migration. Previous studies have shown similar results (Graves and Linneman, 1979; Goss and Paul, 1986).

An adult male who was married and had school-age children tended to reduce his probability of migrating. The coefficient for the share of time married in a residence spell is negative and significantly different from zero at the 1 percent level. We found that being married reduced the migration probability by about 28 to 29 percent. This result is consistent with those of other studies, (e.g., Mincer, 1978) in showing that family ties increase the cost of migration. The number of school-age children, however, did not seem to affect male migration decision. The coefficient on the number of children in table 6.2 is not significantly different from zero at the 5 % level. Mincer (1978) found similar result showing that the significance effect of being married decreases with the inclusion of the number of children in the model. The insignificance of the coefficient for CHILD shown in table 6.2 might be caused by the use of young males sample in the analysis.

The probability of migrating across state boundaries was lower when the individual is a farmer or self-employed. The coefficient for DSLFARM is negative and significantly different from zero at the 5 percent level. The marginal effect of being a farmer or self-employed was to reduce the probability of migration by about 12 percent. This result provides significant evidence that self-employed males have stronger local ties because of land or client bases than do other males, which increases the opportunity cost of moving.

The estimated point-in-time discrete choice model does not show any significant effect of personal unemployment or being in a labor union on the migration decision. The insignificance of personal unemployment does not support the prediction of the theory. Previous studies, such as Da Vanzo (1978), have shown that unemployed people have a higher tendency to move to another labor market to find a job. When the utility of being at the origin decreases because an individual is unemployed, the tendency to migrate should be higher. The hypothesis that being in a labor union increases job security and reduces the chance of moving is also not supported. Using PSID data within the same period, Graves and Linneman (1979) found insignificant effects of unemployment hours on the migration decision.

One possible reason is the nature of the data, especially the use of young population in fitting the probit model. We will see in the next chapter that by using a hazard function representation and following individuals for 20 years, namely, 1968-87, we can get significant effects of personal unemployment and union membership on migration decision. When the hazard function is fitted to the younger sample, however, we found no significant effect of personal unemployment and union membership on migration. Therefore, these results suggests that the insignificance of personal unemployment and union membership is due to the use of younger sample in the analysis.

The estimated probit model does not show any significance effect of local amenities, such as crime rate and percentage area of state and national parks, on the tendency to migrate. Similarly, temperatures as proxy for weather-related variables did not significantly contribute to migration outcomes. One possible reason for this finding is the use of several state characteristics for those who had repeated migration. When a person migrated more than once between 1969 and 1973, the move that was counted was the last move. Therefore, the characteristics of the origin were derived from more than one state. Another reason is the use of younger sample for the analysis. In the next chapter, we will see that the fitted hazard function for migration using younger sample showed similar insignificant effects of local amenities on migration decision.

In summary, we found that an individual's actual wage relative to his potential wage is a strong migration determinant. Human capital-related characteristics, especially education and age, are also important contributors. Family-related variables, especially marital status, reduce the probability of migrating. Similarly, being self-employed or a farmer decreases the chance of migrating to another state. The fitted migration equation, however, does not provide enough support for the effects of personal unemployment, being a labor union member, the percentage of area in state and national parks, or temperatures on migration tendencies.

CHAPTER 7. THE HAZARD FUNCTION FOR MIGRATION

This chapter discusses the empirical results from the search theory-based hazard function representation of the migration decision. Because no previous empirical studies of migration have used this approach, the hazard function results are "new findings." Results are presented for both a constant hazard rate over time where completed duration is assumed to be distributed exponential and a time-dependent hazard rate where the duration is distributed Weibull.

The Constant Hazard Rate Model

The equation representing the hazard rate for migration is fitted using the maximum likelihood procedure to data for 1,268 "residence spells" derived from 915 individuals in the sample. The constant hazard rate model treats the left- and the right-censored spells equally, and we derived 1,061 censored and 207 completed spells. In Table 7.1, average values of explanatory variables over a residence spells are generally used as regressors. In Table 7.2, these same explanatory variables are measured at the beginning of each residence spell. Each table contains four equations; two use PW1 in deriving ΔW and two use PW2 in deriving ΔW. In every pair, the first equation is the estimate without considering the heterogeneity effect. The second equation incorporates heterogeneity in the model, represented by the parameter θ (see the end of Chapter 4). In general, there are strong similarities across all the estimated equations. The signs of the parameter estimates are also consistent across models. In the constant hazard rate model, σ is constrained to equal one.

Our results from the hazard function for migration have many similarities to the results using the standard discrete choice model described in the previous chapter. We found the search model approach does strengthen some of the results. Some explanatory variables which were not

Table 7.1. Hazard function of migration, constant hazard rate model, with some variables measured as average values of a residence spell (t-values are in parentheses)

Explanatory Variable	Using PW1: Without Heterogeneity	Using PW1: With Heterogeneity	Using PW2: Without Heterogeneity	Using PW2: With Heterogeneity
INTERCEPT	-6.422 (-9.26)	-6.406 (-7.66)	-6.452 (-9.31)	-6.442 (-7.71)
ΔW1	-0.244 (-2.68)	-0.273 (-2.13)	-	-
ΔW2	-	-	-0.272 (-2.98)	-0.303 (-2.36)
AVGEDU	0.095 (4.19)	0.136 (3.70)	0.095 (4.20)	0.137 (3.71)
AGE20	1.737 (9.87)	1.296 (4.58)	1.733 (9.86)	1.298 (4.60)
UNEMP	0.001 (4.58)	0.002 (2.74)	0.001 (4.59)	0.002 (2.75)
DUNION	-1.013 (-5.25)	-1.333 (-5.35)	-1.009 (-5.23)	-1.325 (-5.34)
DSLFARM	-1.304 (-6.63)	-1.631 (-6.43)	-1.307 (-6.63)	-1.632 (-6.44)
RACE	1.715 (2.98)	2.051 (3.30)	1.722 (3.00)	2.051 (3.31)
MARR	-0.691 (-4.18)	-1.017 (-3.89)	-0.668 (-4.03)	-0.99 (-3.81)
AVCHILD	-0.142 (-1.99)	-0.128 (-1.52)	-0.141 (-1.97)	-0.126 (-1.50)
CRIME	0.007 (0.30)	0.017 (0.57)	0.007 (0.34)	0.017 (0.57)
PARK	-0.016 (-0.42)	-0.009 (-0.18)	-0.017 (-0.43)	-0.009 (-0.18)
JAN	0.018 (2.61)	0.023 (2.33)	0.018 (2.58)	0.023 (2.32)
JULY	0.014 (1.10)	0.007 (0.42)	0.013 (1.08)	0.008 (0.45)
θ	-	2.608 (3.70)	-	2.590 (3.68)
Log-likelihood	-758.0	-742.1	-757.4	-741.7
Completed spells	207	207	207	207
Total spells	1,268	1,268	1,268	1,268

Table 7.2. Hazard function of migration, constant hazard rate model, with some variables measured at the beginning of a residence spell (t-values are in parentheses)

	Using PW1		Using PW2	
Explanatory Variable	Without Hetero-geneity	With Hetero-geneity	Without Hetero-geneity	With Hetero-geneity
INTERCEPT	-3.516 (-3.41)	-3.964 (-2.07)	-3.575 (-3.47)	-3.990 (-2.09)
ΔW1	-0.184 (-2.05)	-0.257 (-1.54)	-	-
ΔW2	-	-	-0.215 (-2.40)	-0.286 (-1.71)
EDUBEGIN	0.134 (5.92)	0.211 (4.89)	0.135 (5.93)	0.212 (4.90)
AGEBEGIN	-0.183 (-4.21)	-0.200 (-2.13)	-0.180 (-4.16)	-0.199 (-2.13)
AGEBEGIN^2/100	0.274 (5.04)	0.309 (2.50)	0.269 (4.97)	0.308 (2.49)
UNEMP	0.001 (4.95)	0.002 (2.15)	0.001 (4.91)	0.002 (2.16)
UNIBEGIN	-0.512 (-2.33)	-0.766 (-2.52)	-0.507 (-2.31)	-0.760 (-2.50)
DSLFARM	-1.342 (-7.02)	-1.742 (-5.98)	-1.344 (-7.03)	-1.744 (-5.99)
RACE	1.689 (2.96)	2.075 (3.16)	1.693 (2.98)	2.073 (3.17)
MARR	-1.157 (-7.40)	-1.263 (-3.83)	-1.139 (-7.30)	-1.240 (-3.78)
CHILDBGN	-0.140 (-2.34)	-0.169 (-2.14)	-0.139 (-2.32)	-0.167 (-2.12)
CRIME	0.002 (0.08)	0.052 (1.60)	0.002 (0.11)	0.052 (1.59)
PARK	-0.060 (-1.57)	-0.039 (-0.74)	-0.060 (-1.57)	-0.039 (-0.74)
JAN	0.024 (3.42)	0.020 (1.82)	0.024 (3.41)	0.020 (1.82)
JULY	0.016 (1.29)	0.022 (1.10)	0.015 (1.28)	0.022 (1.12)
θ	-	3.835 (4.48)	-	3.814 (4.47)
Log-likelihood	-794.0	-756.3	-793.5	-755.9
Completed spells	207	207	207	207
Total spells	1,268	1,268	1,268	1,268

significant in the probit equation, (e.g., personal unemployment and union membership) turn out to be significant in the hazard model. In general, the signs of the coefficients are as expected and consistent with economic theory. Based on Equation (4.17), the estimated coefficient in the constant hazard rate model can be viewed as the marginal effect of an explanatory variable on the percentage change in the hazard rate.

For the individual wage, we found that an individual's relative wage performance is a strong determinant of migration. The coefficients are consistently negative in all equations and are significantly different from zero at the 5 percent level in six out of eight equations presented in Table 7.1 and 7.2. The slight decrease in significance when heterogeneity is introduced is not surprising. The inclusion of heterogeneity, however, increases the size of the coefficient. The negative signs for the coefficients for $\Delta W1$ and $\Delta W2$ mean that when an individual's wage compares well with his potential wage, his tendency to move to another state decreases. Compared to the use of actual values, the use of predicted state labor market characteristics in constructing the potential wage (PW2) leads to slightly stronger results. We also found that the use of PW2 slightly increased the log-likelihood values.

In general, when actual wages were 1 percent higher than the potential wage, the hazard rate for migration decreased by about 0.2 percent to 0.3 percent. Compared to the results from the discrete choice model, the magnitude of the effect is slightly larger. There is significant evidence to support the hypothesis that males tend to move when their wage rate is unsatisfactory. In other words, individuals tend to move away from (remain in) a state when their wage is lower (higher) than the wages of other individuals with the same personal characteristics living in the same state. Because of location or form specific tenure in the residual wage, this result shows that individuals with a higher wage growth tend to stay longer. When utility is an increasing function of income, the result provides some evidence supporting the comparative static prediction developed in Chapter 2.

Schooling has a strong positive effect on the hazard rate of migration, given that all other things are equal. In all the models, the coefficient of schooling is positive and significantly different from zero at the 1 percent level. Schooling measured at the beginning of the spell (EDUBEGIN) has a slightly larger effect compared to the spell-average measure (AVGEDU). The difference in magnitude, however, is small because going back to school was not a significant activity of the males in the sample. We conclude that the effects of EDUBEGIN and AVGEDU on the hazard rate are similar. The inclusion of the heterogeneity effect seems to strengthen the result. A one-year increase in a male's education increases his hazard rate of moving by between 9 percent and 21 percent. Compared to results from the probit equation in the previous chapter, this magnitude is larger. Also, the result is similar to findings from previous migration studies (Mincer, 1978; Schwartz, 1976; Graves and Linneman 1979). The results support the prediction that more educated males have a larger set of job opportunities, receive higher wage, and more offers to move. The results also support the comparative static prediction discussed in Chapter 2.

We also found the evidence that a male's age plays a significant role in the migration decisions. The effect of the proportion of time spent in a residence spell for males in their 20s (AGE20) can be seen in Table 7.1. The signs of the coefficients are strongly positive and significantly different from zero at the 1 percent level. The inclusion of heterogeneity tends to reduce the size of the effect slightly. This result shows that increasing the share of the residence spell in which the individual is in his 20s increases the tendency to migrate. The marginal effect of being in his 20s is around 1.3 to 1.7. Because age changes over the duration of the spell, the second measurement of age (AGEBEGIN and AGEBEGIN^2) measures the age at the beginning of the residence spell. The effect of age measured at the beginning of the spell is presented in Table 7.2. The coefficients are generally different from zero at the 5 percent level when heterogeneity is not included in the model. The sign of the linear effect AGEBEGIN is negative, and the sign of the quadratic effect $AGEBEGIN^2$ is positive and consistent across models. These results show that when the individual is older at the beginning of the spell, his likelihood of moving is reduced, but at a diminishing rate. This result also supports the fact that most migration occurs when males are young, but out of school. The effect of age on the tendency to migrate, therefore, is similar to the result in the probit equation. Previous studies (e.g., Graves and Linneman, 1979; Pissarides and Wadsworth, 1989) have presented similar findings.

We found that an individual's personal unemployment experience, namely, annual unemployment hours, is positively related to the hazard rate for migration. It is worth mentioning that the probit model discussed in the previous chapter could not provide this evidence. Tables 7.1 and 7.2 show that the effects are small in magnitude, but significantly different from zero at the 5 percent level. The magnitude of the coefficient is larger when heterogeneity is taken into account. Comparing coefficients for the same variables, we see that using regressors measured at the beginning of a residence spell or as a spell average value has no effect on the significance of personal unemployment experience. Even though we started from different economic and econometric models, this result supports previous migration studies (Da Vanzo, 1978; Herzog and Schlottmann, 1984, 1988).

This empirical result supports the hypothesis that an increase in unemployed hours reduces a male's reservation utility for moving and increases the hazard rate. A 10-hour increase in annual unemployment hour increased the hazard rate by 1 percent to 2 percent. It is also evidence for the comparative static predictions discussed in Chapter 2 showing that when the utility at the origin worsens, the hazard rate for migration increases. A possible correlation may exist between wage performance and personal unemployment hours in a spell. However, because the wage rate is measured as an hourly wage, we concluded that the correlation between the two is trivial.

Union membership tended to reduce the hazard rate for migration. This result is different from that obtained with the probit model. The coefficient of union membership in all models is significantly different from zero at the 5 percent and 1 percent level. In terms of the magnitude of the coefficients, the effects are stronger when heterogeneity is taken into account. The use of the dummy variable DUNION led to a stronger effect compared to the use of UNIBEGIN. DUNION was set equal one if an individual belonged to a union at any time during the time spell. UNIBEGIN equals to one if the individual was in a union at the beginning of the spell. The difference in the magnitude of the effects is expected because UNIBEGIN is a subset of DUNION and therefore is more restrictive. The consistent negative signs in all the equations support the hypothesis that union-related benefits such as job security reduce the hazard rate for moving. In general, the marginal effect of union membership on the hazard rate is about -0.5 to -1.3. However, this result does not necessarily contradict Goss and Paul (1990), who found a positive relationship. One reason is for this difference is the use of different samples. Goss and Paul used a more restrictive sample, namely unemployed union members.

We found that being self-employed or a farmer decreases the hazard rate for moving. The higher opportunity cost of moving because of a farmer's attachment to the localities or land reduces the tendency to move to other states. The cost of moving is also higher for others who are self-employed because they stand to lose customers and might need to change occupations. The sign of the coefficient for DSLFARM is consistently negative and significantly different from zero at the 1 percent level across models. The use of PW1 and PW2 does not lead to any significant difference. In terms of magnitude, the inclusion of heterogeneity seems to strengthen the results. In our sample, the marginal effect of being self-employed or a farmer reduced on the hazard rate for is around -1.3 to -1.7. This effect is larger in magnitude compared to that found with the probit equation, and the signs are also consistent.

We found that being white increased the hazard rate of moving. Across all models, the coefficient of RACE is positive and significantly different from zero at the 5 percent level. The effect is stronger when heterogeneity is taken into account. This result is similar to the stylized fact of migration found in the Current Population Reports (U.S. Department of Commerce, 1989). Individuals who are white have a higher rate of long-distance moves within the United States and are more likely to move between states. On the other hand, individuals who are black have higher rates of local moving, namely, moves between places in the same county.

Compared to results from the probit model, the results from the hazard function model provide stronger evidence for the effect of family related variables on migration decision. One clear reason is the use of longer time period that follows individual for 20 years in the hazard model. In the probit model, the migration equation was fitted to a younger sample. We found that being married reduces the hazard rate for migration. The coefficients on the share of time spent married in a spell (MARR) are all negative, large in magnitudes, and significantly different from zero at the 1 percent level. Having large number of children in school also decreases an individual's hazard rate of moving. The coefficients for AVCHILD and CHILDBGN are negative and significantly different from zero at the 5 percent and 1 percent levels respectively. Because the number of children in school tends to change over the duration of a residence spell, the effects of AVCHILD and CHILDBGN are different. The measurement at the beginning of the spell leads to a larger coefficient. However, the coefficients in all models show strong negative signs. This result supports the hypothesis that family ties tend to reduce the tendency to move (Mincer, 1978). A greater number of children in school increases the cost of moving and thus reduces the hazard rate for migration. The marginal effect of an additional child is around -0.12 to -0.17.

The results do not provide strong evidence for the effects of local amenities, especially CRIME and PARK, on migration decisions. In general, the coefficients are not significantly different from zero at the 5 percent level. The signs, however, fit the prediction. A positive sign for the coefficient for CRIME in the model incorporating heterogeneity reflects the view that males living in a state with a higher crime rate have a higher tendency to move. The consistent negative signs of the coefficient for PARK show that a larger percentage of land allocated to state or national parks in a state reduces the hazard rate for moving. The estimated coefficients on weather-related variables, namely, JAN and JULY, show more promising results. Some of the coefficients on JAN and JULY are significantly different from zero at the 5 percent level.

We found significant heterogeneity effects in the model. The estimate of the heterogeneity parameter θ is significantly different from zero at the 1 percent level in all models. The effect of θ is stronger when variables are measured at the beginning of a residence spell (Table 7.2). The parameter θ measures the sensitivity of the hazard rate and the survivor function to heterogeneity. However, the inclusion of θ in the model did not significantly change the estimated coefficients of the explanatory variables. The magnitude of the estimated coefficients does change slightly, but not the signs. The largest effect is on the coefficient of the variables representing age, between the measurement at the beginning of the spells (AGEBEGIN) and the share of time spent in one's 20s (AGE20). This is to be expected, given the way heterogeneity enters the model.

In summary, the estimation of the hazard function based on search theory has led to stronger results compared to the results obtained with the point in time discrete choice model. Individual wage performance, measured by actual wage relative to potential wage, is a strong migration determinant. Education and age also play major roles in migration decisions. The more educated and younger males have higher hazard rates for moving. Being white and unemployed also increases the tendency to migrate to other states. Being a labor union member, however, reduces the chance. Family-related variables, such as the number of school-age children and being married, reduce the hazard rate for moving.

The Time Dependent Hazard Rate Model

In the constant hazard rate model, the hazard rate is constant over the duration of a residence spell, and the distribution of the completed duration is exponential. When the duration is longer than one year, this assumption seems too strong. A better model is one that allows for some changes in the hazard rate over the duration of residence spells. If we assume the distribution of duration is Weibull, a reasonably simple specification of the time-dependent hazard rate model can be fitted.

The main advantage of the time-dependent hazard rate model is the possibility of measuring the effect of the duration of residence spells on the hazard rate. The use of a Weibull distribution adds only one parameter to our model but restricts it to being the linear form between the duration and the hazard rate. In other words, the estimate of this parameter (σ) tells whether the hazard rate is linearly decreasing, constant, or increasing over time. One parameter seems to be inadequate in representing the dependence of the hazard rate on the duration of stay. However, there is another way to show the effect of the duration of stay on the hazard rate. By calculating the predicted hazard rate evaluated at the average values of the explanatory variables and the duration of stay, we can plot the predicted hazard rate for a given range of duration. This graphical description will show the movement of the predicted hazard rate as a function of the length of duration.

Based on the discussion in Chapter 5, we applied three research strategies, or treatments, in dealing with the time-dependent hazard rate. With the exception of the first treatment, the results are encouraging. The signs of the estimated coefficients are similar to those from the constant hazard rate model. Some estimated coefficients are stronger in magnitude, but others are weaker. The following discussion will present the result based on the treatments elaborated in Chapter 5. Because the effect of the explanatory variables on migration is generally consistent between the use of a constant or time-dependent hazard rate, the discussion below emphasizes the additional contribution of the time-dependent hazard model. To reduce the complication arising when explanatory variables change over time, we use the explanatory variables measured at the beginning of the residence spell.

Treatment 1

The first treatment is to select all completed and right-censored spells started in 1968 or after. The hazard function is fitted to the total of 391 spells, with 207 completed. The omission of all left-censored spells clearly leads to potential nonrandom selection problem. For example, every individual who took up residence before 1968 and remained in that state continuously until 1987 is excluded from the duration sample. Nevertheless, this first empirical strategy is conducted to examine the result when the hazard function is fitted based on the completely observed length of spells and perfectly measured explanatory variables.

The fitted equation is presented in Table 7.3. The first and third columns present the estimates using $\Delta W1$ and $\Delta W2$ without permitting

Table 7.3. Hazard function of migration, time dependent hazard model, based on all spells started in 1968 and beyond, treatment 1, (t-values are in parentheses)

Explanatory Variable	Using PW1		Using PW2	
	Without Hetero-geneity	With Hetero-geneity	Without Hetero-geneity	With Hetero-geneity
INTERCEPT	-2.886 (-2.16)	-0.278 (-0.39)	-2.933 (-2.20)	-0.276 (-0.21)
ΔW1	-0.369 (-3.09)	-0.036 (-0.57)	-	-
ΔW2	-	-	-0.386 (-3.23)	-0.048 (-0.56)
EDUBEGIN	0.028 (0.96)	-0.003 (-0.23)	0.028 (0.96)	-0.002 (-0.09)
AGEBEGIN	-0.012 (-0.20)	0.016 (0.50)	-0.011 (-0.18)	0.011 (0.20)
$AGEBEGIN^2/100$	0.009 (0.12)	-0.022 (-0.54)	0.006 (0.08)	-0.014 (-0.20)
UNEMP	0.001 (1.72)	0.000 (0.02)	0.001 (1.70)	0.000 (0.16)
UNIBEGIN	-0.155 (-0.67)	0.006 (0.06)	-0.154 (-0.67)	0.008 (0.05)
DSLFARM	-1.434 (-5.30)	-0.091 (-0.70)	-1.434 (-5.32)	-0.153 (-0.82)
RACE	1.091 (1.95)	0.021 (0.08)	1.109 (1.99)	0.017 (0.03)
MARR	-0.869 (-3.92)	-0.021 (-0.19)	-0.858 (-3.88)	0.030 (0.22)
CHILDBGN	0.013 (0.17)	-0.033 (-0.65)	0.014 (0.19)	-0.027 (-0.40)
CRIME	0.021 (0.74)	-0.008 (-0.47)	0.021 (0.72)	-0.007 (-0.27)
PARK	-0.022 (-0.44)	-0.007 (-0.25)	-0.021 (-0.43)	-0.006 (-0.15)
JAN	0.010 (1.07)	0.001 (0.25)	0.010 (1.09)	0.002 (0.19)
JULY	0.004 (0.24)	0.004 (0.57)	0.004 (0.27)	0.004 (0.32)
σ	1.170 (9.73)	0.039 (0.90)	1.168 (9.76)	0.057 (0.98)
θ	-	43.244 (0.88)	-	38.864 (0.96)
Log-likelihood	-510.5	-342.9	-510.0	-399.2
Completed Spells	207	207	207	207
Total Spells	391	391	391	391

heterogeneity. The second and the fourth columns present the model with heterogeneity incorporated. By comparing the results in the first and third columns to the results for the constant hazard model, some similarities become apparent. The signs of the coefficients are consistent, which means that the results do not violate the prediction based on the theory and that they match the results of the previous models. The coefficients for ΔW1 and ΔW2, DSLFARM, MARR, and RACE are significantly different from zero at the 5 percent level. The results show that wage performance, being self-employed or a farmer, being unemployed, racial differences, and marital status continue to be important determinants of migration outcomes.

The inclusion of heterogeneity in the model, however, leads to several problems. First, the estimation process had some difficulties in reaching convergence. Second, the estimated coefficients are not significantly different from zero, including the coefficient for heterogeneity. One possible reason is the nonrandom sample selection problem carried by this first treatment. The coefficients for heterogeneity are large in magnitude, but they are not significantly different from zero, even at the 10 percent level (see Table 7.3). Based on this finding, we conclude that the model without heterogeneity has more meaningful results.

We found the estimated σ to be larger than one, meaning that the hazard rate decreases with an increase in the duration of stay (see column 1 and 3 of Table 7.3). However, the hypothesis that Ho: $\sigma = 1$ in both equations can not be rejected at the 5 percent significant level. The t statistics are similar and equal to 1.4, which means that the hypothesis that the hazard rate is constant over time can not be rejected for this sample of duration data. This result suggests that treatment 1 does not lead to a significant improvement compared to the use of the constant hazard rate model previously discussed.

Treatment 2

The second treatments utilizes all the duration spells with a known starting year, including those started before 1968. The spells with an unknown starting year are excluded. The hazard function is fitted to 865 spells, with 284 completed. Compared to the first treatment, this selection increased the number of spells, especially completed spells. Therefore, the second treatment reduces potential sample selection problems. The estimated coefficients are presented in Table 7.4. The first two columns present the model fitted on all samples, and the last two columns show the model based on a younger male sample, namely, the males who were 19 to 24 years of age in 1968. The younger male sample provided 278 spells, with 137 spells completed.

The inclusion of heterogeneity generally improves the quality of the estimated coefficients and does not lead to any convergence problem in the computation. With an exception in the third column of Table 7.4, the estimated coefficients on heterogeneity (θ) are all positive and different from zero at the 5 percent level. Therefore, we do not present the model without heterogeneity. Because the coefficient of heterogeneity in the third column is not significant, we will use the model in the fourth column to discuss the result from the younger male sample. Similar to the results in previous estimated hazard functions, the use of ΔW1 and ΔW2 did not lead to any major differences.

There are some interesting comparisons between the results derived from the full sample and those obtained from the younger male sample (young in 1968). In general, the use of the full sample with more spells results in larger t-values. This result is expected. The signs of the coefficients are consistent between the two samples, especially when the estimated coefficient is significantly different from zero. We conclude that the use of a more restricted sample, namely, the younger population, does not alter the effect of migration determinants.

Table 7.4. Hazard function of migration, time-dependent hazard model, based on all observable residence spells, treatment 2, (t-values are in parentheses)

Explanatory Variable	Full Sample (Age 19-45 in 1968)		Young Sample (Age 19-24 in 1968)	
	Using PW1	Using PW2	Using PW1	Using PW2
INTERCEPT	-6.608 (-7.87)	-6.614 (-7.85)	-7.367 (-3.02)	-7.394 (-2.97)
ΔW1	-0.606 (-5.06)	-	-0.435 (-2.57)	-
ΔW2	-	-0.603 (-5.05)	-	-0.420 (-2.53)
EDUBEGIN	0.049 (1.74)	0.048 (1.72)	-0.001 (-0.02)	-0.005 (-0.11)
AGEBEGIN	0.271 (5.82)	0.271 (5.80)	0.501 (2.43)	0.505 (2.40)
$AGEBEGIN^2/100$	-0.323 (-4.81)	-0.323 (-4.80)	-0.918 (-2.29)	-0.926 (-2.27)
UNEMP	0.001 (1.11)	0.001 (1.12)	0.002 (1.55)	0.002 (1.59)
UNIBEGIN	-0.175 (-1.02)	-0.170 (-0.99)	0.143 (0.68)	0.170 (0.81)
DSLFARM	-1.229 (-7.22)	-1.229 (-7.21)	-1.469 (-6.45)	-1.46 (-6.44)
RACE	0.620 (1.53)	0.633 (1.55)	0.476 (0.90)	0.532 (1.01)
MARR	-1.297 (-5.06)	-1.299 (-5.07)	-0.252 (-0.68)	-0.251 (-0.68)
CHILDBGN	-0.196 (-4.29)	-0.194 (-4.27)	0.333 (1.57)	0.331 (1.56)
CRIME	0.055 (1.99)	0.056 (2.01)	0.006 (0.16)	0.008 (0.19)
PARK	-0.084 (-1.44)	-0.085 (-1.46)	-0.068 (-0.88)	-0.069 (-0.90)
JAN	-0.002 (-0.19)	-0.002 (-0.23)	-0.004 (-0.3)	-0.005 (-0.37)
JULY	0.001 (0.05)	0.001 (0.05)	0.018 (0.69)	0.019 (0.71)
σ	0.630 (7.91)	0.632 (7.97)	0.187 (2.66)	0.184 (2.62)
θ	3.876 (3.82)	3.843 (3.83)	16.364 (2.31)	16.750 (2.29)
Log-likelihood	-795.0	-795.0	-316.1	-316.3
Completed spells	284	284	137	137
Total spells	865	865	278	278

In both samples, the coefficient of age and age-squared show that the hazard rate for moving increases as people get older, but at a diminishing rate. After peaking, the hazard rate decreases as age increases. It is expected that the effects of age are different between the two samples. When the full sample is used, the peak tendency to move occurs at around 42 years of age. In the younger male sample, the peak migration rate occurs at 27 years of age. It is worth mentioning that the effect of age found here is different from the result derived from the constant hazard rate model. The constant hazard rate model led to the a negative effect of age on the hazard rate. In this model, the effect of age is increasing and then decreasing after reaching its peak. This difference does not necessarily violate the view that, in general, the hazard rate decreases as people get older.

Even though the inverted U-shape pattern fits well the stylized facts on how age affects migration decisions, the peak in the full sample does not occur when individuals are in their mid-20s. This is expected because the mean of individual's age for the full sample in 1968 was 33 years. Therefore, most migrations during 1968-87 occurred when most individuals were in their 30s or 40s. The result from the young sample fits to the fact that most migration occurred when individuals are in their 20s. The result showed that the peak of AGEBEGIN equals to 27 years. This is because the migration behavior of the young sample was observed since they were 19 to 24 years of age in 1968 until they were 38 to 43 years of age in 1987. It is important to note that the AGEBEGIN is defined as the age at the beginning of a spell, and it marks the time when the individual makes the move to end the previous, not the current, residence spell.

Another interesting difference between the full sample and the younger male sample in this analysis can be found in the effect of the number of school-age children. The number of school-age children was found to be a significant determinant of migration is the full sample. In the young male sample, however, no significant effect was found. It is not surprising because, in general, individuals in the younger sample migrated when they were in their 20s, were unmarried, or had no children.

There is also a difference in the effect of the crime rate. The coefficient for the crime rate is much more significant in the full sample compared to that in the younger male sample. It seems that the younger males were less crime conscious than the older males.

Compared to the result in the constant hazard rate model, the signs of the coefficients are consistent. The coefficients for wage performance at the origin, education, age, being self-employed or a farmer, being married, and the number of school-age children are significantly different from zero at the 5 percent level. We conclude that this result does not violate the prediction based on the economic theory or the results of previous studies. It is important to note that the effect of the crime rate in the time-dependent hazard rate model seems to be more significant. The positive sign means that a higher crime rate in the current state of residence, relative to the U.S. average, increases the hazard rate of moving.

The parameter σ marks the difference between the constant and the time-dependent hazard rate. The estimates of σ the first and second column of Table 7.4 are similar and smaller than one. The hypothesis that $\sigma = 1$ is rejected at the 5 percent level, with t-statistics of -4.6. The estimates for the younger sample shown in the fourth column show a similar result. The hypothesis that $\sigma = 1$ is also rejected for the younger male sample at the 5 percent level, with t-statistics of -11.7. These results show that there is an increasing effect of length of duration of stay on the hazard rate. Holding other explanatory variables constant, the longer an individual lives in one place, the higher is the hazard rate for moving. However, this result should be interpreted carefully. From Equation

(4.22), we know that holding X constant, σ represents the marginal effect of t on H. Because of the nonlinear relationship between the duration (t), explanatory variables (X) and the hazard rate (H), the interpretation of σ is more complicated. The interdependence between X and t has made σ not fully capture how the hazard rate changes as the duration increases.

Greene (1991) provides a more realistic way to illustrate the movement of the hazard rate as a function of the duration of a spell. The objective is accomplished by plotting the predicted hazard rate on the duration of stay. First, spells or observations are grouped based on the duration of stay. Then, based on the explanatory variables corresponding to each group, the means of the hazard rate in each group are calculated and plotted. Compared to the estimates of σ just discussed, this method illustrates the relationship between the hazard rate and the duration more realistically. The graphical representations are more realistic for at least two reasons. First, this methods allows for the change in other explanatory variables, such as age and education, corresponding to a change in the duration of stay. Second, the graphical representations are simple to understand.

The plots derived from the equations in the second and fourth column of Table 7.4 are presented in Figures 7.1 and 7.2., respectively. In general, the pattern of the hazard rate is an inverted u-shape curve. This pattern has a meaningful interpretation. With an increase in the duration of stay, the hazard rate for moving increases, reaches a peak, and then decreases. From Figure 7.1, we can see that the maximum hazard rate occurs for individuals staying in one place for four to six years. For the younger male sample shown in Figure 7.2, we can see that the peak occurs earlier, two to four years of residency in one place. The values of the scale in Figures 7.1 and 7.2. show the maximum value of the predicted hazard rate. The scale values show the maximum hazard rate for the

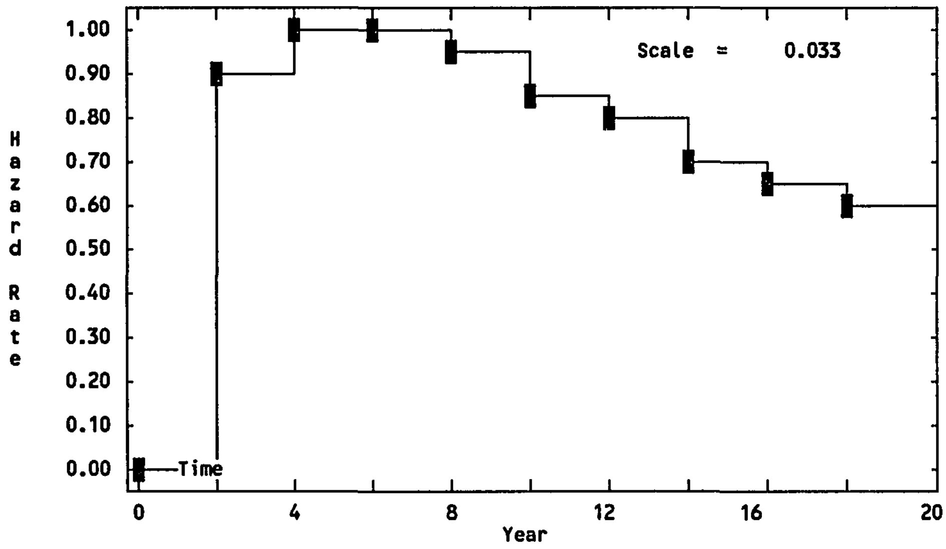

Figure 7.1. Predicted hazard rate and the duration of stay in the time-dependent hazard model, using treatment 2, for the full sample.

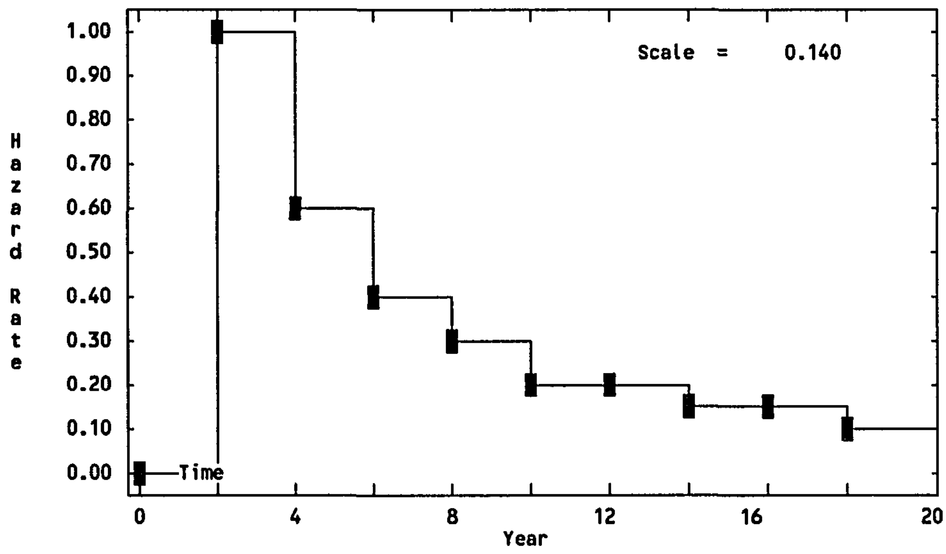

Figure 7.2. Predicted hazard rate and the duration of stay in the time dependent hazard model, using treatment 2, for the young male sample.

younger sample (0.14) is higher than that of the full sample (0.03). This results fits the prediction that younger people have a higher tendency to move.

In general, we conclude that the second treatment performs better than the first one. The coefficients are better in quality and the signs are consistent with the results from the constant hazard rate model. Among other explanatory variables, the residual wage remains a strong determinant of migration. A decrease in the actual wage relative to the potential wage increases the hazard rate for moving. Higher education consistently leads to a strong positive effect on the hazard rate. Being white also increases the hazard rate. Some factors that reduce the hazard rate for moving are age, being self-employed or a farmer, being married, and the number of school-age children. In the second treatment, a higher crime rate leads to a higher hazard rate for moving.

Treatment 3

For the third treatment, we predict the unobserved starting year of residency and then use all spells in the computation to fit the migration model. This procedure utilizes all 1,268 spells, with 496 spells completed. The number of completed spells increases from 207 in the constant hazard rate model, to 284 in the second treatment, to 496 in this treatment. In general, the estimated coefficients are statistically stronger after closing the open spells. Similar to most previous results, the use of $\Delta W1$ or $\Delta W2$ does not lead to any significant difference.

The first step of this third treatment is the estimation of an unobserved starting date for left-censored spells. Heckman's procedure is used to predict the unobserved length of some spells (Heckman, 1979). The details of the procedure can be found in Appendix D. We conducted a test that the predicted completed spells are distributed Weibull. The hypothesis that the distribution of the predicted completed spells is

Weibull can not be rejected at the 5 percent level. The Chi-squared statistics for the goodness-of-fit test is 19.7 with 18 degrees of freedom. Therefore, we conclude that the use of predicted spells for the third treatment does not violate the assumption that the distribution of the completed spells in time-dependent hazard rate is Weibull.

The estimated model for the third treatment can be found in Table 7.5. The first three columns use the full sample, and the last two utilize the younger male sample. The third column presents the model which uses the share of a male's time in residence while he is 20 to 29 years of age (AGE20) to represent age, rather than AGEBEGIN. Compared to results in previous hazard function for migration, the signs of the coefficients are consistent. The asymptotic t-statistics are larger in magnitude, showing a generally higher quality of parameter estimates. This improvement can be seen clearly in the effects of the crime rate and personal unemployment. Compared to the result from the second treatment, the coefficients for crime rate shown in the first two columns of Table 7.5 are positive and significantly different from zero at the 5 percent level. Similarly, strong effects occur for personal unemployment.

The effect of age on the hazard rate for this treatment is similar to the result from treatment 2. The pattern of the effect of age measured at the beginning of the spell is an inverted U-shape curve. The peak effect in treatment 3 occurs at 42 years of age for the full sample, and at 27 years of age in the younger male sample. Using AGE20 in the model shows that the larger the share of time an individual spent his 20s, the higher is the hazard rate for moving. The coefficient of AGE20 is positive and significantly different from 0 at the 1 percent level. We realized that using AGE20 might not be appropriate in this model. By construction, AGE20 is negatively correlated with the length of the spell, and when the hazard

Table 7.5. Hazard function of migration, time-dependent hazard model, based on all residence spells, treatment 3 (t-values are in parentheses)

Explanatory Variable	Full Sample (Age 19-45 in 1968)			Young Sample (Age 19-24 in 1968)	
	Using PW1	Using PW2	Using PW2	Using PW1	Using PW2
INTERCEPT	-6.627 (-11.44)	-6.634 (-11.47)	-3.552 (-8.56)	-3.938 (-2.12)	-3.935 (-2.11)
ΔW1	-0.543 (-6.66)	-	-	-0.509 (-3.86)	-
ΔW2	-	-0.549 (-6.76)	-0.516 (-5.99)	-	-0.499 (-3.76)
EDUBEGIN	0.044 (2.41)	0.044 (2.38)	0.090 (4.82)	0.078 (2.09)	0.077 (2.10)
AGEBEGIN	0.274 (8.41)	0.274 (8.41)	-	0.163 (1.08)	0.164 (1.07)
$AGEBEGIN^2/100$	-0.330 (-7.14)	-0.330 (-7.14)	-	-0.302 (-1.05)	-0.304 (-1.05)
AGE20	-	-	0.008[a] (5.74)	-	-
UNEMP	0.001 (1.69)	0.001 (1.71)	0.001 (2.81)	0.002 (3.93)	0.002 (3.98)
UNIBEGIN	-0.221 (-1.68)	-0.219 (-1.67)	-0.440 (-2.98)	-0.270 (-1.29)	-0.257 (-1.23)
DSLFARM	-1.193 (-9.35)	-1.191 (-9.33)	-1.333 (-9.21)	-1.658 (-8.51)	-1.664 (-8.53)
RACE	0.507 (2.06)	0.512 (2.07)	0.669 (2.43)	0.131 (0.28)	0.142 (0.30)
MARR	-1.303 (-7.33)	-1.294 (-7.33)	-1.477 (-8.15)	-0.331 (-1.31)	-0.328 (-1.28)
CHILDBGN	-0.151 (-4.59)	-0.150 (-4.57)	-0.096 (-2.66)	0.143 (1.01)	0.155 (1.09)
CRIME	0.045 (2.44)	0.046 (2.47)	0.046 (2.46)	0.050 (1.51)	0.052 (1.55)
PARK	-0.029 (-0.90)	-0.028 (-0.86)	-0.006 (-0.19)	0.001 (0.01)	0.000 (0.00)
JAN	-0.004 (-0.71)	-0.005 (-0.81)	0.007 (1.26)	-0.005 (-0.50)	-0.005 (-0.55)
JULY	0.006 (0.55)	0.006 (0.55)	0.009 (0.85)	0.007 (0.37)	0.006 (0.35)
σ	0.666 (10.90)	0.667 (10.91)	0.845 (10.89)	0.317 (4.41)	0.315 (4.39)
θ	2.664 (4.80)	2.644 (4.78)	1.587 (3.68)	7.079 (3.30)	7.166 (3.29)
Log-likelihood	-1288.0	-1287.2	-1337.0	-417.4	-417.6
Completed spells	496	496	496	191	191
Total spells	1268	1268	1268	350	350

[a]The unit of AGE20 is in percent

rate changes over time, the inclusion of this variable in the model could lead to misspecification. Nonetheless, to illustrate why the hazard rate is higher when males are in their mid-20s, we fit the effect of AGE20 (see Table 7.5, column 3).

Interesting comparisons can be found between the use of full male sample and the young male sample. For example, the effects of family-related variables are different. As shown in Table 7.5, having a large number of school-age children significantly reduces the hazard rate for migration. The table also shows that for the young male sample, the number of school-age children has no effect on the hazard rate for migration. The significance of the marital status in reducing the hazard rate is also reduced for the young male sample.

Personal unemployment turns out to be a significant determinant of migration decisions for young males. Compared to the use of large sample, the coefficient of personal unemployment in the younger male sample is larger. This result reflects the fact that unemployed young males seemed to have a higher tendency to make long-distance moves compared to the unemployed older males. This result is also consistent with the view that younger males have lower migration cost and a higher tendency to migrate.

The crime rate affects migration differently between the younger and the older males. The effect of crime rates on migration in the younger male sample is not significant, while the effect for the full sample is strongly negative. This result means that the young males were less concerned about the crime rate in making decisions to move. This is consistent with the difference in behavior toward risk between younger and older people. The young seem to be greater risk takers compared to the old.

The parameter σ shows the contribution of the time-dependent hazard rate model compared to the constant hazard model. The estimates of σ are less than one and consistent with the results from treatment 2. In all the equations, the hypothesis that $\sigma = 1$ is rejected at the 5 percent level. For the full sample, the t-statistic derived from the equation in the second column is -5.4. For the younger male sample, the t-statistics derived from the equation in the fifth column is -9.5. Other explanatory variables held equal, the hazard rate increases with an increase in duration. As discussed previously, the nonlinear relationship between the hazard rate, duration of stay, and explanatory variables complicates the interpretation of σ as the marginal effect. As we did for treatment 2, we present the more realistic graphical representation of the predicted hazard rate and the duration of stay.

The plot of the relationship between the predicted hazard rate and the duration of stay corresponding to the use of the full male sample and the young male sample (second and fifth column of Table 7.5) are presented in Figure 7.3 and 7.4, respectively. The graph shows that the effect of the duration of stay on the hazard rate is quadratic. The hazard rate first increases with a larger duration, peaks, and then decreases. This inverted U-shape form is consistent with the results from the second treatment. When the full male sample is used, the maximum hazard rate occurs when the duration of stay is four to six years. As expected, the maximum hazard rate in the young sample occurs after a shorter duration of stay. The maximum hazard takes place when the duration of stay is two years. In general, the results show that the hazard rate is small after a recent move, it increases until it reaches its peak when males have resided for six years in one place, and then it decreases. The value of the scale shows that the young sample has higher maximum predicted hazard rate (0.14) than the full sample (0.04).

The signs of the coefficients are consistent with the results in the previous treatments. Nonetheless, we found that the third treatment performs better than the previous two. The third treatment shows a stronger result for the effect of crime rates on the migration decision.

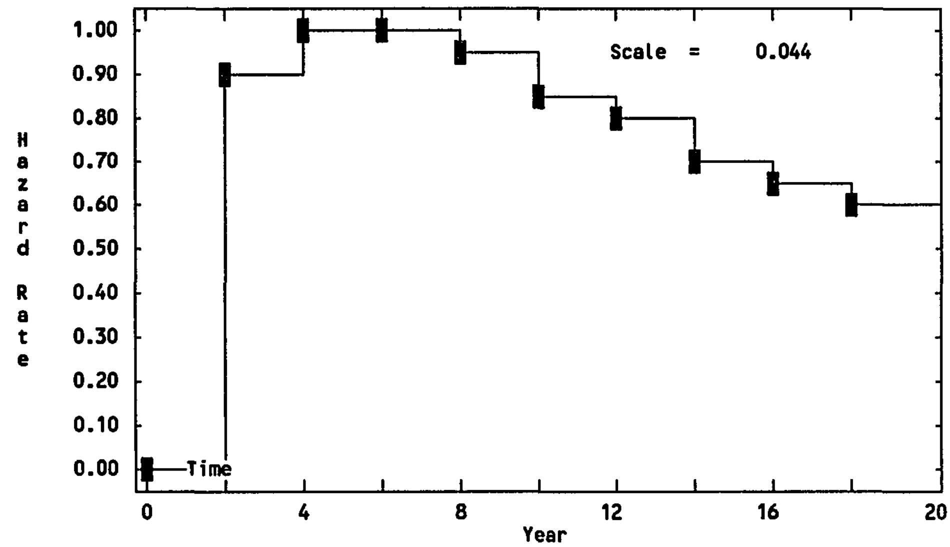

Figure 7.3. Predicted hazard rate and the duration of stay in the time-dependent hazard model, using treatment 3 for the full sample.

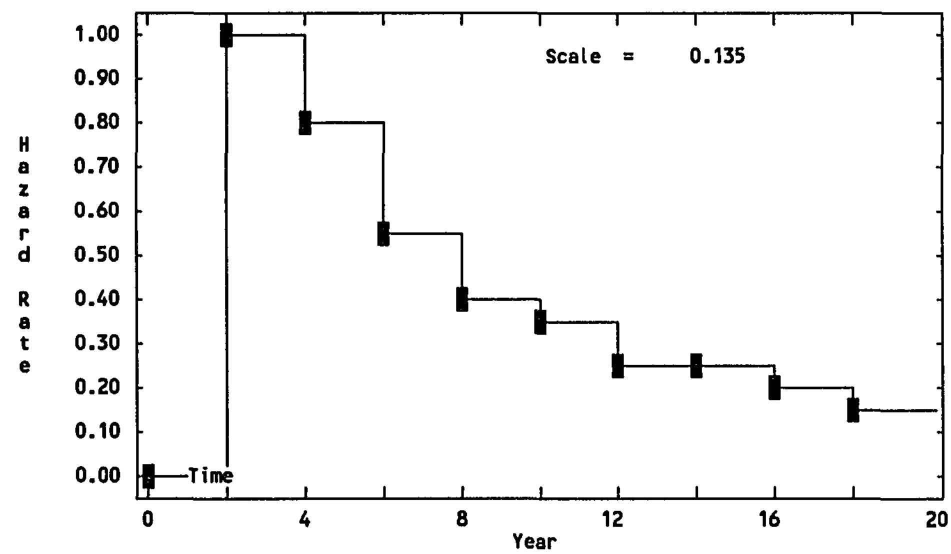

Figure 7.4. Predicted hazard rate and the duration of stay in the time-dependent hazard model, using treatment 3 for the young male sample.

Consistently, the residual wage is a strong determinant of migration. A decrease in the actual wage relative to the potential wage increases the hazard rate for moving. More education and being white also have strong positive effects on the hazard rate. Factors that reduce the hazard rate for moving are age, being self-employed or a farmer, being married, and the number of school-age children.

While temperature is a significant determinant for migration in the constant hazard rate model, none of the estimates from the time-dependent hazard model produce significant estimates. One possible reason is the use of a constant average 30-year temperature in the model. Clearly, the constant January and July temperatures fit better in the constant hazard rate model. The study did not show any significant effect of the percentage of areas belonging to state and national parks (PARK) on the migration decision.

The marginal effects of some explanatory variables on the hazard rate for migration are calculated and presented in Table 7.6. The first column presents the results from the constant hazard rate model, directly taken from the coefficients in the fourth column of Table 7.2. The second and third columns present results from the time-dependent hazard model, taken from the second and fourth of Table 7.5. In general, the signs of the marginal effects are consistent across models. A comparison of the magnitudes can be made. The marginal effect of a 1.00 percent increase in wage performance on the hazard rate for migration ranges from -0.29 percent to -1.58 percent. Compared to the result in the constant hazard model, the effect of wage performance ($\Delta W2$) is stronger in the time-dependent hazard model. The younger male sample produced the highest marginal effects showing that the young males were more responsive to wage performance in making migration decisions.

Table 7.6. Marginal effect of some explanatory variables on the hazard rate for migration

Explanatory Variable	Unit	Constant Hazard rate (From Table 7.2, Column 4)	Time-dependent Hazard Rate[a]	
			Full sample (From Table 7.5, Column 2)	Young sample (From Table 7.5, Column 5)
ΔW2	percent	-0.286*	-0.823***	-1.584***
EDUBEGIN	year	0.212***	0.066**	0.244*
AGE20	(0 - 1)[b]	1.298***c	0.947***d	-
UNEMP	hour	0.002**	0.001*	0.006***
UNIBEGIN	Dummy:0,1	-0.760**	-0.328*	-0.816
DSLFARM	Dummy:0,1	-1.744***	-1.786***	-5.283***
RACE	Dummy:0,1	2.073***	0.768**	0.451
MARR	(0 - 1)[b]	-1.240***	-1.940***	-1.043
CHILDBGN	0, 1, 2, ...	-0.167**	-0.225***	0.492
CRIME	percent[e]	0.052	0.069**	0.165

[a]The marginal effect is β/σ, see Equation (4.23)

[b]The share of time spent being married, and also being in the 20s, ranges from 0 to 1

[c]Derived from Table 7.1, column 4

[d]Derived from Table 7.5, column 3

[e]Measured relative to U.S. average

*Significant at the 10 percent level

**Significant at the 5 percent level

***Significant at the 1 percent level.

Education, personal unemployment, being white, and a higher crime rate clearly increase the hazard rate for migration. A one-year increase in education increased the hazard rate of migration by 6.6 percent to 24.4 percent. An annual 10-hour increase in unemployment augments the hazard rate by 1 percent to 6 percent. A 1 percent increase in the crime rate relative to the U.S. average increased the hazard rate by 5 percent to 17 percent. The marginal effect of being white ranged from 0.45 to 2.07. Similar to the effect of wage performance, a larger marginal effect of personal unemployment and male education are found in the young male sample. The decision to move for the young males is more responsive to unemployment experience because, all other things held equal, the cost of moving is smaller. The stronger effect of formal schooling for the young males sample is also expected. Education plays major roles because the lack of employment experience for the young sample.

The effect of the crime rate on migration was found to be smaller for the younger males. This result is consistent with the difference in risk behavior between the younger and the older males. Being young also increases the hazard rate for migration. The marginal effect of age, calculated based on the effect of AGE20, was positive. We could see on Table 7.6 that the marginal effect of being in one's 20s on the hazard rate for migration is around 0.95 to 1.30.

Being a labor union member, being self-employed or a farmer, being married, and having school-age children reduce the hazard rate for migration. The marginal effect of being a labor union member ranged from 0.33 to -0.82. The marginal effect of being self-employed or a farmer on migration was large, ranging from -1.74 to -5.28. It is worth mentioning that the effect of being self-employed or a farmer is stronger when we use the younger male sample. The marginal effect of being married ranged from -1.04 to -1.94. Having one more school-age child reduced the chance to migrate by 22.5 percent to 49.2 percent. It is clear that the marginal effects of family-related variables such as marital status and having school-age children are weaker for the younger male sample. There is a positive marginal effect of CHILDBGN in the last column of Table 7.6; however, this marginal effect is derived from a coefficient which is not significantly different from zero in the model.

There were some concerns related to the estimation of a time-dependent hazard rate. First, the use of predicted length of spell in the estimation might lead to misspecification and violation of the assumption that the distribution of the completed duration is Weibull. Second, the use of explanatory variables measured in 1968 or after in the spell started before 1968, especially when the full sample is used, could lead to measurement error problem. These problems are reduced when we use the young male sample. The young sample consists of those who were 19-24 years in 1968. Their migration behavior and their personal and local attributes in their 20s and 30s, from 1968 to 1987, were perfectly observed. The data from the young sample in 1968 are the best to fit the time dependent hazard function of migration. The consistency of the results, both in terms of the signs and the significance of the coefficients, has shown that the estimation procedure appears robust.

In summary, we conclude that using the time-dependent hazard rate improves the quality of the estimates. The main contribution of the time-dependent hazard rate is to provide perspective on the relationship between the hazard rate and the duration of stay. We found an inverted U-shape relationship. An additional contribution can be found in measuring the significance the crime rate to the migration decision.

We found consistency in the signs of the coefficients. A lower wage performance, higher education, younger age, and being white increase the hazard rate for moving. Being self-employed or a farmer, being in a labor union, being married, and having school-age children tend to reduce the hazard rate. Interesting comparisons between the use of the full sample and the younger male sample are found. The tendency to move for younger males is more sensitive to wages, education, and job-related variables, such as unemployment experience, union membership, and occupational choice, but is less sensitive to family-related variables and the crime rate. This results show the difference in risk behavior and cost of migration between the younger and older males.

CHAPTER 8. SUMMARY AND CONCLUSION

The Search Model for Migration

This study has developed a new model for determining an individual's migration decision. By applying the search theory on migration, the new model adds the multi-period dimension to previous one point-in-time decision model. Search theory allows for modelling the migration decision as a sequential decision over time. In addition, uncertainty about the future is treated explicitly. Therefore, the new model captures repeated migration and better fits the nature of migration as a life-cycle decision.

The search model views migration decisions as sequential decisions made over time under uncertainty. An individual migrates to a new place when the expected utility to move is higher than his/her reservation utility for moving. Over time, the individual receives an offer to move represented by a utility assessment corresponding to a move, net from the cost of moving. The individual remains in the same place as long as the offer is less than or equal to the reservation utility. Therefore, the concept of interest in this study is the duration of someone living in one place before he moves and the hazard rate for moving. The search model develops the relationship between the hazard rate, the duration of stay, and some explanatory variables that determine migration.

The model predicts that better living conditions at the origin reduce the hazard rate for moving and extend the duration of a stay. It predicts a positive relationship between the hazard rate and the expected utility corresponding to an offer to move. When an individual receives more offers to move, the hazard rate is expected to be higher. Another important result is the relationship between the hazard rate for moving and the expected life span. The hazard rate is higher when the expected life span after moving is longer. In other words, this relation explains the higher tendency for young adults to move.

The model is fitted empirically to a panel data set from the Panel Study of Income Dynamics (PSID) and local state characteristics from the Statistical Abstracts of the United States. The study follows the migration behavior of male heads of households, who were age 19 to 45 in 1968, for 20 years from 1968 to 1987. The econometric analysis consists of estimating the probit model based on the point-in-time discrete choice model and the hazard functions based on the search model. The empirical analysis of this study introduced the wage performance of an individual as one determinant of migration.

In general, we have estimated four models. First, we estimated the potential wage equation to derive the individual's relative wage performance. The relative wage performance was used in all equations to help explaining migration. Second, by following previous migration studies, we fitted a standard probit model corresponding to the point-in-time discrete choice framework. Third, we fitted the hazard function of migration by assuming that the hazard rate is constant over time. Fourth, we modified the assumption and estimated a time-dependent hazard rate model.

The Wage Equation and The Probability of Migration

The wage equation was fitted to develop individual wage performance, or the residual wage, by measuring how well an individual's actual wage performs relative to his potential wage. The residual wage also captures the location- or firm-specific effects on wage. The dependency of residual wages on the duration of stay is useful in fitting the time-dependent hazard model. The potential wage is fitted to 15,367 annual individual observations. The estimated wage equation provided further evidence that human capital factors, especially education and potential experience are strong determinants of an individual's wage. One year of additional education increased an individual's real wage by about 7.5 percent. The wage peak occurred at about 45 years of age. We also found that white males earned 11 percent more than the nonwhite males.

State characteristics were found to play a major role in determining individual wage rates. Wage rates differ because of differences in living cost and labor market conditions. We found that areas with higher land prices offer higher wages. The elasticity is around 0.04 to 0.05. A 1 percent increase in the percentage of urban population increased the wage rate by around 28 percent. A 1 percent increase in actual and predicted state unemployment rates leads to a 1.5 percent and 4.3 percent increase in the wage rate, respectively. A 1 percent increase in actual and predicted state job growth rates contributed to a 1.5 percent and 6.0 percent increase in the wage, respectively. The results also show that the potential wage responds to the unanticipated state unemployment and employment shocks, and state average January and July temperatures. In general, the real wage has declined since the early 1970s. The wage rates in the southern and western states are lower relatively than in the northern-central and the eastern states.

The probability-of-migration equations were fitted to 915 males observed during the 1969-73. By assuming normality, the probit model was appropriate. We found that the residual wage was a strong determinant of migration. A 1 percent increase in the actual wage relative to the potential reduces the probability of migration by 0.17 percent. This result shows that people who earn less than other with the same personal characteristics living in the same state are at higher risk of moving.

For variables other than the individual relative wage, we found similar results compared to previous studies. A 1 year increase in general education increased the probability of migration by 1.6 percent. We also found that younger males have a higher tendency to move; the marginal effect of age evaluated at the mean is -3.6 percent. The current study supported the finding of the previous results showing that white males tend to participate in interstate migration more often than nonwhite males. Being married and being self-employed or a farmer reduced the probability of migrating. The marginal effect of the share of time spent married is around -28 to -29 percent. Being self-employed or a farmer decreased the probability of moving by about 12 percent. The results from this probit model, however, do not provide enough evidence on the effects of personal unemployment, the number of school-age children, being in a labor union, and local amenities such as crime rate, percentage of area in state and national parks, and temperature. One reason is the use of the younger sample in the early period of the PSID data (1968-73) to fit the model. This finding is consistent with the result obtained from the hazard model when it was fitted to the younger male sample.

The Hazard Function for Migration

The hazard function for migration was derived from the search theory for migration. It showed how the probability of migration can be modelled as a multi-period sequential decision with uncertainty. The empirical hazard function of migration was fitted to 1,268 spells, derived from all 915 individual in the sample observed from 1968-87. There were two general models fitted in this study; a model of constant hazard rate over time and a time-dependent hazard rate. In general, we found that search theory performs better than the point-in-time standard discrete choice model in explaining the migration. The inclusion of heterogeneity in the model not only reduced the biases of the estimates but also improved their statistical significance.

The constant hazard rate model was fitted to 1,268 spells. The completed spells is derived from all the right-closed spells starting in or after 1968. Similar to the result from the discrete choice model, we found that residual wage was a strong determinant for migration. The marginal effect was -0.29. Education and age also played major roles in the decision to migrate. One year of additional general education increased the hazard rate by about 21 percent. The results also showed that being young, namely being in the mid-20s, increased the hazard rate by 1.30. The effect of racial characteristics on migration was similar to that found with the discrete choice model, namely, that being white increased the hazard rate for moving. We also found that being married and being self-employed or a farmer reduced the hazard rate for moving. The marginal effects were -1.24 and -1.74, respectively.

Compared to the probit model, the hazard function provided stronger results for the effects of personal unemployment, being in a labor union, and having school-age. We found that 10 hours of unemployment per year increases the hazard rate for moving by 2 percent. Being at the labor union and having one children at the school age reduced the hazard rate by 76 and 17 percent respectively. The constant hazard model showed the significance effect of the weather temperature, but not the other local amenities such as crime rates and percentage of areas belong to state and national park.

Time dependent hazard rate models were fitted by using several procedures or treatments. The first treatment used only the spells started in or after 1968. The second treatments utilized all spells with perfectly observed length. With the last treatment, the unobserved spell lengths were predicted, and all spells were used to fit the model. The completed spells increased from 207 in the first treatment, to 284 in the second, to 496 in the third treatment. Except for the first treatment, in general, we found stronger results with using time-dependent model. The signs of the coefficients were consistent to those found with previous models, and the quality of the estimates is higher. In the second and third treatments, we fitted the model twice- one for the full sample and once for the younger sample.

The advantage of the time-dependent hazard model is the movement of the hazard rate over time. We found that the relationship between the hazard rate and the duration of stay was an inverted U-shape curve. With the increase in the duration of stay, the hazard rate to move increased, reached the maximum, and then decreased. Results from the full sample showed that the maximum hazard occurred when individuals in one place for 6 years. A higher tendency for the young males to move caused the maximum hazard rate to occur at the shorter duration of stay, namely two to four years.

The determinants of migration are different for the full sample compared with the younger male sample. With the exceptions of family-related variables and crime rates, in general, the effects of the explanatory variables on the hazard rate for migration were stronger for the younger male sample. For example, young males were more sensitive to the wage effect. A 1 percent decrease in residual wage increased the hazard rate for the young males by 1.6 percent, compared with only 0.8 percent for the full sample. It is important to note that the effects of residual wage were stronger when we use the time-dependent hazard model. The decision to move for young males seemed to be more sensitive to unemployment experience than for older males. In general, the marginal effect of a 10-hour increase in unemployment per year on the hazard rate was 1 percent for the full sample, compared with 6 percent for the young male sample. The effects of occupational choice were also different between the two samples. The marginal effect of being self-employed or a farmer on the hazard rate for the younger males was around -5.28, compared with -1.79 for the full sample.

Education plays an important role in the migration decision. More educated people have a higher tendency to move. When the accumulation of employment experience is small, the marginal effect of education on the hazard rate for migration is higher. This effect was shown to be stronger when the young male sample is used. The marginal effect of one year of education on the decision to migrate was 0.24 for the young male sample and 0.07 for the full sample. The marginal effect of being a union member was -0.82 for the young males and -0.33 for the full sample. However, being a union member was not a significant determinant for migration for the young male sample.

The effects of family-related variables on migration were weaker in the younger male sample than in the full sample. It is expected that being married and having school-age children do not affect the migration decision for the young male sample. The marginal effect of the proportion of time spent married in the spell on the hazard rate was -1.94 and -1.04 for the full sample and the younger male sample, respectively. Having one school-age child reduced the hazard rate by 22.5 percent. When the younger male sample was used, the effect of the number of school-age children was not significant. The younger males also seemed to be less sensitive to the crime rate in the state. For the full sample, a 1 percent increase in the crime rate relative to the U.S. average led to a 7 percent increase in the hazard rate. For the younger male sample, the effect was not statistically significant.

There has been no standard procedure to fit the time-dependent hazard rate when the length of some spells and the value of some explanatory variables were not observed. We had to develop several research strategies to fit the models. We realize that there may be some problems associated with those empirical procedures. However, the consistency of the results shows that the parameter estimates are quite robust.

REFERENCES

Abraham, Katharine G. and Henry S. Farber. "Job Duration, Seniority, and Earnings." American Economic Review 77 (1987): 278-97.

Avery, Robert B., Lars Peter Hansen, and V. Joseph Hotz. "Multiperiod Probit Models and Orthogonality Condition Estimation." International Economic Review 24 (1983): 21-35.

Becker, Gary S. Human Capital. 2nd ed. New York: Columbia University Press for the NBER, 1975.

Brown, James N. "Why Do Wages Increase with Tenure? On-the-Job Training and Life-Cycle Wage Growth Observed within Firms." American Economic Review 79 (1989): 971-91.

Chalmers, James A. and Michael J. Greenwood. "The Economics of Rural to Urban Migration Turnaround." Social Science Quarterly 61 (1980): 524-44.

Chamberlain, Gary. "Analysis of Covariance with Qualitative Data." Review of Economic Studies 17 (1980): 225-38.

Chamberlain, Gary. "Heterogeneity, Omitted Variable, Duration Dependence." In Longitudinal Analysis of Labor Market Data, edited by James J. Heckman and Burton Singer, 3-38. Cambridge: Cambridge University Press, 1985a.

Chamberlain, Gary. "Panel Data." In Handbook of Econometrics, edited by Z. Griliches and M.D.Intriligator. Vol 2, 1247-1318. Amsterdam: North-Holland Pub. Co., 1985b.

Council of Economic Advisers, The. "The Annual Report of the Council of Economic Advisers." In Economic Report of The President. Washington, D.C.: U.S. Government Printing Office, 1994.

Cox, David R. "Partial Likelihood." Biometrica 62 (1975): 269-76.

Da Vanzo, Julie. "Does Unemployment Affect Migration?-Evidence from Micro Data." The Review of Economics and Statistics 60 (1978): 504-14.

Devine, Theresa J. and Nicholas M. Kiefer. Empirical Labor Economics: The Search Approach. New York: Oxford University Press, 1991.

Fletcher, Roger. Practical Methods of Optimization. New York: Wiley, 1980.

Goss, Ernst P. and Chris Paul. "Age and Work Experience in the Decision to Migrate," Journal of Human Resources 21 (1986): 397-405.

Goss, Ernie and Chris Paul. "Union Membership, Economic Rents, and Migration Behavior." Journal of Labor Research 11 (1990): 347-55.

Goss, Ernst P. and Niles C. Schoening. "Search Time, Unemployment, and the Migration Decision." Journal of Human Resources 19 (1984): 570-79.

Graves, Philip E. "A Life-cycle Empirical Analysis of Migration and Climate, by Age." Journal of Urban Economics 6 (1979): 135-47

Graves, Philip E. and Peter D. Linneman. "Household Migration: Theoretical and Empirical Results." Journal of Urban Economics 6 (1979): 383-404.

Greene, William H. Econometric Analysis. New York: Macmillan Publishing Company, 1990.

Greene, William H. Limdep: Reverence Manual Version 6.0. Bellport, New York: Econometric Software, Inc., 1991.

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

Greenwood, Michael J."Human Migration: Theory, Models and Empirical Studies."Journal of Regional Science 25 (1985): 521-44.

Greenwood, Michael J. and G. L. Hunt. "Migration and Inter-regional Employment Redistribution in the United States." American Economic Review 74 (1984): 957-69.

Heckman, James J. "Sample Selection Bias as a Specification Error." In Female Labor Supply, edited by J. Smith, 204-48. Princeton: Princeton University Press, 1980.

Heckman, James J. and Burton Singer. "Social Science Duration Analysis." In Longitudinal Analysis of Labor Market Data, edited by James J. Heckman and Burton Singer, 39-110. Cambridge: Cambridge University Press, 1985.

Herzog, Henry W., Jr. and Alan M. Schlottmann. "Labor Force Mobility in the United States: Migration, Unemployment, and Remigration." International Regional Science Review 9 (1984): 43-58.

Herzog, Henry W. Jr. and Alan M. Schlottmann. "Local Business Cycle Severity and Induced Labor Mobility: The Case of Construction Workforce." Construction Management and Economics 6 (1988): 247-58.

Hill, Martha S. The Panel Study of Income Dynamics: A User's Guide. Newbury Park, CA: Sage Publication, 1992.

Huffman, Wallace E. and Robert E. Evenson. Science for Agriculture: A Long-term Perspective. Ames, Iowa: Iowa State University Press, 1991.

Huffman, Wallace E. "Human Capital, Adaptive Ability, and the Distributional Implications of Agricultural Policy." American Journal of Agricultural Economics 67 (1985): 429-34.

Judge, George G., R. Carter Hill, William E. Griffiths, Helmut Lutkepohl, and Tsoung-Chao Lee. Introduction to the Theory and Practice of Econometrics. New York: Wiley, 1988.

Kalbfleish, John D. and Robert L. Prentice. The Statistical Analysis of Failure Time Data. New York: Wiley, 1980.

Kiefer, Nicholas M. "Economic Duration Data and Hazard Function." Journal of Economic Literature 26 (1988): 649-79.

Lippman, S.A. and J.J. McCall, ed. Studies in The Economics of Search, Amsterdam: North-Holland Pub. Co., 1979.

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

Mincer, Jacob. "Family Migration Decision". Journal of Political Economy 86 (1978): 749-773.

Mincer, Jocob. Schooling, Experience and Earnings. New York: Columbia University Press for the NBER, 1974.

Nakosteen, Robert A. and Michael Zimmer. "Migration and Income: The Question of Self Selection." Southern Economic Journal 46 (1980): 840-51.

Pissarides, Christopher A. and Jonathan Wadsworth. "Unemployment and The Inter-Regional Mobility of Labor." The Economic Journal 99 (1989): 739-55.

Roback, Jennifer. "Wages, Rents, and the Quality of Life." Journal of Political Economy 90 (1982): 1257-78.

Rosen, Sherwin. "The Theory of Equalizing Differences." In Handbook of Labor Economics, edited by Orley Ashenfelter and Richard Layard, 641-92. Amsterdam: North-Holland, 1986.

Schaeffer, Peter. "Human Capital Accumulation and Job Mobility." Journal of Regional Science 25 (1985): 103-14.

Scholttmann, Alan M. and Henry W. Herzog, Jr. "Employment Status and the Decision to Migrate." Review of Economics and Statistics 63 (1981): 590-98.

Schwartz, Aba. "Migration, Age, and Education." Journal of Political Economy 84 (1976): 701-719.

Sjaastad, Larry A. "The Cost and Return of Human Migration." Journal of Political Economy 70 (1962): S80-93.

Stark, Oded and David E. Bloom. "The New Economics of Labor Migration." American Economic Review 75 (1985): 173-78.

Stigler, G. J. "The Economics of Information." Journal of Political Economy 69 (1961): 213-25.

Stigler, G. J. "Information in the Labor Market." Journal of Political Economy 70 (1962): S94-105.

Tokle, Joane G. and Wallace E. Huffman. "Local Economic Conditions and Wage Labor Decisions of Farm and Rural Nonfarm Couples." American Journal of Agricultural Economics 73 (August 1991): 652-70.

Topel, Robert H. "Local Labor Markets." Journal of Political Economy 94 (1986): 111-43.

Topel, Robert H. "Specific Capital, Mobility and Wages: Wages Rise with Job Seniority." Journal of Political Economy 99 (February 1991):145-76.

U.S. Department of Commerce, Bureau of the Census. Current Population Reports, series P-20, no 430, April 1989.

U.S. Department of Commerce, Bureau of the Census. Statistical Abstract of the United States. Washington, DC: U.S. Government Printing Office. 1968-1993.

U.S. Department of Commerce, Bureau of Economic Analysis. Survey of Current Business. Washington, DC: U.S. Government Printing Office. 1965-1993.

APPENDIX A. THE DERIVATION OF THE SEARCH MODEL

The Derivation of the Reservation Utility Equation

From Equation (2.6) in Chapter 2, we have:

$$Vnm = \frac{Unm\ h}{(1+rh)} + \frac{\delta h}{(1+rh)} E[Max\{Vm(Um),Vnm\}] + \frac{(1-\delta h)}{(1+rh)} Vnm + O(h)\ k = Vm(Ur) \quad (2.6)$$

We could rearrange to:

$$Vnm = \frac{Unm}{(r+\delta)} + \frac{\delta}{(r+\delta)} E[Max\{Vm(Um),Vnm\}] + \frac{O(h)}{(1+rh)(r+\delta)h} = Vm(Ur) \quad (2.6a)$$

The limit of the last part of the right hand side goes to zero as h->0. From Equation (2.7) we have $Vm(Um) = Um(1-e^{-rT})/r$, and the maximum of Vnm while living at the origin is $Vm(Ur) = Ur(1-e^{-rT})/r$. Therefore, we can write Equation (2.6a) as:

$$\frac{UrC}{r} = \frac{Unm}{(r+\delta)} + \frac{\delta C}{r(r+\delta)} E[Max\{Um,Ur\}], \quad \text{where: } C = (1-e^{-rT}) \quad (2.6b)$$

$$UrC = Unm + \frac{\delta C}{r} E[Max\{Um,Ur\}] - \frac{Ur}{r}$$

$$UrC = Unm + \frac{\delta C}{r} [\int_0^{Ur} UrdF(Um) + \int_{Ur}^{\infty} UmdF(Um) - \int_0^{Ur} UrdF(Um) - \int_{Ur}^{\infty} UrdF(Um)]$$

With $C = (1-e^{-rT})$, this leads to Equation (2.8):

$$Ur = \frac{Unm}{(1-e^{-rT})} + \frac{\delta}{r} [\int_{Ur}^{\infty} (Um-Ur)dF(Um)] \quad (2.8)$$

The Effect of δ, Unm, r, μ_{Um}, and T on Ur

Let: $[*] = [\int_{Ur}^{\infty} (Um-Ur)dF(Um)\ ] > 0$, from Equation (2.8), and

Taking derivative of Equation (2.8) with respect to δ and Ur,

$$dUr = \frac{[*]}{r} d\delta - \frac{\delta}{r} [1-F(Ur)]\ dUr$$

$$\frac{dUr}{d\delta} = \frac{[*]\ /\ r}{[1+\delta/r[1-F(Ur)]} > 0$$

Taking derivative of Equation (2.8) with respect to Unm and Ur,

$$dUr = \frac{dUnm}{(1-e^{-rT})} - \frac{\delta}{r} [1-F(Ur)]\ dUr$$

$$\frac{dUr}{dUnm} = \frac{1}{(1-e^{-rT})[1+\delta/r[1-F(Ur)]} > 0$$

Taking derivative of Equation (2.8) with respect to r and Ur,

$$dUr = -\frac{UnmTe^{-rT}}{(1-e^{-rT})^2} dr - \frac{\delta[*]}{r^2} dr - \frac{\delta}{r} [1-F(Ur)]\ dUr$$

$$\frac{dUr}{dr} = -\frac{[UnmTe^{-rT}/(1-e^{-rT})^2 + \delta[*]]}{[1+\delta/r[1-F(Ur)]} < 0$$

Taking derivative of Equation (2.8) with respect to μ_{Um} and Ur,

$$dUr = \frac{\delta}{r} [1-F(Ur)]\ d\mu_{Um} - \frac{\delta}{r} [1-F(Ur)]\ dUr$$

$$\frac{dUr}{d\mu_{Um}} = \frac{\delta\ [1-F(Ur)]}{[r + \delta[1-F(Ur)]} > 0 \text{ and } < 1$$

Taking derivative of Equation (2.8) with respect to T and Ur,

$$dUr = -\frac{bre^{-rT}}{(1-e^{-rT})^2}\, dT - \frac{\delta}{r}\, [1-F(Ur)]\, dUr$$

$$\frac{dUr}{dT} = -\frac{(bre^{-rT})/(1-e^{-rT})^2}{[1+\delta/r[1-F(Ur)]} < 0$$

The Effect of δ, Unm, r, μ_{Um}, and T on H

For Equation (2.10) the relation between Ur and hazard rate (H) is $H = \delta\ \phi(Ur)$, where

$$\phi(Ur) = \int_{Ur}^{\infty} f(Um)\, dUm = 1-F(Ur) > 0, \text{ and } \partial\phi/\partial Ur = -\ f(Ur).$$

We take partial derivative of Equation (2.10) with respect to δ, Unm, r, μ_{Um}, and T respectively, and then utilize the information from above to sign the equation.

$$\frac{\partial H}{\partial \delta} = \phi(Ur) - \delta\ f(Ur)\ \frac{\partial Ur}{\partial \delta} \quad \text{(ambiguous)}$$

$$\frac{\partial H}{\partial Unm} = -\ \delta\ f(Ur)\ \frac{\partial Ur}{\partial Unm} < 0$$

$$\frac{\partial H}{\partial r} = -\ \delta\ f(Ur)\ \frac{\partial Ur}{\partial r} > 0$$

$$\frac{\partial H}{\partial \mu_{Um}} = - \delta \; f(Ur) \frac{\partial Ur}{\partial \mu_{Um}} < 0$$

$$\frac{\partial H}{\partial T} = - \delta \; f(Ur) \frac{\partial Ur}{\partial T} > 0$$

APPENDIX B. SOME STATE LOCAL CHARACTERISTICS

Table B.1. Employment growth, unemployment rate, and land prices, for some selected years

State	Employment Growth (percent)			Unemployment Rate (percent)			Real Land Price ($/acre)		
	1968	1978	1987	1968	1978	1987	1968	1978	1987
Alabama	1.6	5.2	3.0	4.5	6.3	7.8	358	679	555
Arizona	5.7	10.1	3.5	3.6	6.1	6.2	186	238	273
Arkansas	2.2	5.2	2.8	4.2	6.3	8.1	525	807	525
California	4.1	6.8	3.7	4.5	7.1	5.8	1,376	1,417	1,843
Colorado	5.0	8.3	0.4	3.0	5.5	7.7	246	430	253
Connecticut	2.8	5.1	2.5	3.7	5.2	3.3	1,660	2,287	2,498
Delaware	4.0	3.7	5.8	3.1	7.6	3.2	1,081	1,968	1,565
D. C.	2.0	2.1	2.3	2.2	8.5	6.3	1,097	1,924	1,523
Florida	6.1	8.1	5.3	2.8	6.6	5.3	1,253	2,208	2,128
Georgia	3.6	6.7	4.0	3.3	5.7	5.5	404	907	542
Idaho	2.1	7.5	1.5	4.3	5.7	8.0	516	923	507
Illinois	1.8	3.2	2.8	3.0	6.1	7.4	1,323	2,735	1,079
Indiana	2.2	4.3	3.7	3.2	5.7	6.4	1,124	2,211	934
Iowa	2.1	3.6	3.2	2.4	4.0	5.5	1,105	2,361	773
Kansas	3.5	4.7	2.0	2.7	3.1	4.9	412	620	299
Kentucky	2.1	4.3	4.2	3.9	5.2	8.8	510	950	571
Louisiana	2.2	7.0	-2.3	4.8	7.0	12.0	822	1,371	788
Maine	3.3	5.6	2.5	4.5	6.1	3.2	228	487	590
Maryland	3.6	5.6	3.8	3.2	5.6	4.2	1,112	1,881	1,482
Massachusetts	3.3	5.6	2.5	4.5	6.1	3.2	940	1,332	1,549
Michigan	3.9	5.6	1.6	4.3	6.9	8.2	636	1,140	596
Minnesota	3.4	5.6	3.4	3.2	3.8	5.4	445	1,072	485
Mississippi	3.0	6.1	1.9	4.5	7.1	10.2	455	783	477
Missouri	1.4	4.8	2.5	3.4	5.0	6.3	466	872	435
Montana	2.6	5.5	-0.4	4.7	6.0	7.4	142	264	150
Nebraska	3.5	2.7	2.1	2.4	2.9	4.9	449	746	382
Nevada	7.0	12.8	6.6	5.0	4.4	6.3	135	207	181
New Hampshire	2.1	7.2	4.8	1.8	3.8	2.5	329	871	1,138
New Jersey	1.4	4.0	1.8	4.6	7.2	4.0	2,082	3,141	2,812
New Mexico	1.1	6.8	0.6	5.1	5.8	8.9	193	295	186
New York	2.1	2.8	1.9	3.5	7.7	4.9	414	672	515
North Carolina	4.6	4.5	4.2	3.2	4.3	4.5	741	1,177	881
North Dakota	2.6	5.7	1.2	4.1	4.6	5.2	225	466	286
Ohio	3.9	3.9	2.5	2.9	5.4	7.0	874	1,789	866
Oklahoma	2.8	6.4	-1.3	3.5	3.9	7.4	423	670	374
Oregon	3.6	7.4	3.8	4.4	6.0	6.2	826	1,431	878
Pennsylvania	2.0	3.3	2.6	3.2	6.9	5.7	651	1,528	1,510
Rhode Island	1.5	4.1	2.2	3.6	6.6	3.8	1,351	2,614	2,613
South Carolina	2.4	5.0	4.0	4.3	5.7	5.6	473	822	556
South Dakota	3.0	4.3	2.0	3.0	3.1	4.2	234	382	203
Tennessee	4.3	5.3	4.2	3.6	5.8	6.6	475	893	622
Texas	4.3	7.2	-0.7	2.7	4.8	8.4	547	769	786
Utah	2.4	7.1	0.9	5.2	3.8	6.4	209	443	381
Vermont	2.9	7.1	5.0	3.6	5.7	3.6	326	656	803
Virginia	3.9	5.2	3.7	2.7	5.4	4.2	547	1,017	825
Washington	4.8	8.3	4.5	4.3	6.8	7.6	896	1,565	905
West Virginia	0.8	3.4	0.2	5.5	6.3	10.8	225	583	427
Wisconsin	2.3	4.8	3.2	3.4	5.1	6.1	479	1,043	561
Wyoming	3.0	8.9	-6.9	3.9	3.3	8.6	79	151	87
Alaska	5.2	0.6	-5.1	9.1	11.2	10.8	44	184	309
Hawaii	6.1	4.9	4.5	2.9	7.7	3.8	366	763	992
Average U.S.	3.1	5.6	2.3	3.8	5.7	6.2	628	1,113	833

Table B.2. State percentage of urban population, crime rates, areas of state and national park, and temperature, for some selected years

State	Urban Population (percent) 1968	1978	1987	Crime Rate (percent) 1968	1978	1987	Park (Percent)	Average Temp. (degree F.) Jan.	July
Alabama	58	60	60	11.8	13.3	9.3	0.2	50.8	82.2
Arizona	79	83	86	6.3	9.4	7.5	2.9	52.3	92.3
Arkansas	49	51	53	8.1	9.1	7.6	0.1	39.9	82.1
California	90	91	92	6.0	11.7	10.6	4.6	56.0	69.0
Colorado	78	80	82	5.4	7.3	5.8	0.7	29.5	73.3
Connecticut	78	79	79	2.5	4.2	4.9	6.2	25.2	73.4
Delaware	71	71	72	7.7	6.7	5.1	0.7	31.2	76.0
D. C.	100	100	100	19.1	19.9	36.2	16.1	35.2	78.9
Florida	79	84	85	11.9	11.0	11.4	1.1	53.2	81.3
Georgia	59	62	63	13.9	14.4	11.8	0.1	41.9	78.6
Idaho	53	54	56	2.3	5.4	3.1	0.4	29.9	74.6
Illinois	83	83	84	8.1	9.9	8.3	0.9	21.4	73.0
Indiana	64	64	65	4.7	6.2	5.6	0.2	26.0	75.1
Iowa	56	58	60	1.7	2.6	2.1	0.1	18.6	76.3
Kansas	65	67	68	3.7	5.7	4.4	0.1	29.6	81.4
Kentucky	51	51	52	8.9	9.0	7.5	0.2	32.5	77.6
Louisiana	66	68	68	9.5	15.8	11.1	0.1	52.4	82.1
Maine	51	48	45	3.0	2.7	2.5	0.5	21.5	68.1
Maryland	76	80	81	9.3	8.2	9.6	1.6	32.7	76.8
Massachusetts	84	84	84	3.5	3.7	3.0	5.5	29.6	73.5
Michigan	74	71	71	7.3	10.6	12.2	0.7	23.4	71.9
Minnesota	66	67	69	2.2	2.0	2.6	0.3	11.2	73.1
Mississippi	43	47	47	9.9	12.6	10.2	0.2	45.7	81.9
Missouri	69	69	69	8.8	10.4	8.3	0.4	25.9	78.5
Montana	53	53	53	3.3	4.8	4.1	1.2	18.7	69.3
Nebraska	60	63	65	2.3	3.0	3.5	0.3	20.2	77.7
Nevada	79	84	87	5.5	15.5	8.4	1.4	32.2	69.5
New Hampshire	57	53	51	1.4	1.4	3.0	1.2	19.9	69.5
New Jersey	89	89	89	5.1	5.4	4.6	5.7	31.8	74.4
New Mexico	69	72	73	6.2	10.2	10.1	0.1	34.8	78.8
New York	86	85	84	6.5	10.3	11.3	0.8	21.1	71.4
North Carolina	44	47	50	9.7	10.8	8.1	1.4	40.5	78.5
North Dakota	42	48	52	1.1	1.2	1.5	0.0	6.7	70.4
Ohio	75	74	74	5.3	6.9	5.8	0.7	28.9	75.4
Oklahoma	67	67	68	6.4	8.5	7.5	0.2	35.9	82.1
Oregon	66	68	70	3.2	5.0	5.6	0.1	38.9	67.7
Pennsylvania	72	70	69	4.0	6.2	5.4	1.1	31.2	76.5
Rhode Island	87	87	86	2.4	4.0	3.5	1.6	28.2	72.5
South Carolina	46	53	54	13.6	11.5	9.3	0.3	44.7	81.0
South Dakota	44	46	49	3.8	1.9	1.8	0.2	12.4	74.0
Tennessee	58	60	61	8.7	9.4	9.1	1.3	39.6	82.1
Texas	79	80	80	10.6	14.2	11.7	0.5	44.0	86.3
Utah	79	84	86	2.9	3.7	3.3	2.4	28.6	77.5
Vermont	33	33	33	2.6	3.3	2.7	2.7	16.6	69.6
Virginia	62	65	68	8.3	8.8	7.4	1.3	39.9	78.4
Washington	72	73	76	3.6	4.6	5.6	2.3	39.1	64.8
West Virginia	39	37	36	5.5	6.8	4.8	1.0	32.9	74.5
Wisconsin	65	65	65	2.2	2.5	3.5	0.2	18.7	70.5
Wyoming	60	62	64	6.3	7.1	2.0	3.9	26.1	68.9
Alaska	46	61	67	10.5	12.9	10.1	0.9	21.8	55.7
Hawaii	82	86	89	2.8	6.7	4.8	0.6	72.6	80.1
Average U.S.	66	67	68	6.3	7.8	7.0	1.5	32.2	75.5

APPENDIX C. LOCATION-SPECIFIC TENURE IN THE RESIDUAL WAGE

To show the importance of location-specific tenure in the residual wage, we fitted the wage equation with the duration of stay in a state up to year y (T_{iy}, in years) as an additional explanatory variable. Because T_{iy} is endogenous, we also fitted the equation with the predicted tenure, PT_{iy}, as an instrumental variable. We used T_{iy} derived from the procedures described in Appendix D. Table C.1. below presents the results. The first column presents the wage equation with the actual tenure, the second column shows the estimated equation to derive the instrumental variable PT_{iy}, and the third column presents the wage equation with the predicted tenure.

The result shows that the estimated coefficient for actual tenure (T_{iy}) is significantly different from zero at the 1 percent level. The positive sign means that wage increases with tenure. The second equation demonstrates the endogeneity of tenure by showing the tenure as a function of other regressors in the wage equation. The third equation shows a higher t-statistics on the coefficient for the predicted tenure (PT_{iy}), even though the magnitude of the coefficient is similar to that of the first equation. The procedure presented in this section is not meant to perfectly measure the effect of location-specific tenure on wages. We used this procedure to illustrate the importance of tenure in the residual wage.

Table C.1. The estimate of wage equation by including the effect of location specific tenure (t-statistics are in parentheses)

	Wage Equation with Actual Tenure	Equation to Predict Tenure	Wage Equation with Predicted Tenure
Dependent variable:	$Ln(W_{iky}/P_y)$	T_{iy}	$Ln(W_{iky}/P_y)$
Regressor:			
INTERCEPT	1.431 (8.23)	45.744 (25.38)	1.397 (8.05)
EDU	0.072 (31.38)	0.009 (0.40)	0.075 (34.18)
EXP (AGE-EDU-6)	0.047 (17.89)		0.050 (20.43)
$EXP^2/100$	-0.098 (-19.61)		-0.100 (-19.98)
RACE	0.114 (4.91)	0.560 (2.23)	0.113 (4.87)
Ln(PLAND/P)	0.046 (2.94)	3.766 (26.44)	0.027 (1.72)
URBAN	0.317 (4.40)		0.332 (4.60)
PJOBGR	0.065 (8.03)		0.065 (8.06)
PURATE	0.042 (6.86)		0.042 (6.86)
RSHOCK	0.009 (3.00)		0.009 (2.99)
RURATE	-0.005 (-1.01)		-0.005 (-1.00)
CRIME	0.015 (6.12)	0.106 (4.75)	0.015 (5.90)
JAN	-0.004 (-3.15)	-0.075 (-6.51)	-0.003 (-2.67)
JULY	-0.018 (-10.92)	-0.026 (-1.68)	-0.018 (-11.00)
TIME	-0.010 (-1.53)		-0.008 (-1.26)
$TIME^2/100$	-0.013 (-0.48)		-0.013 (-0.46)
DS	-0.086 (-2.94)	-1.980 (-7.54)	-0.081 (-2.78)
DW	-0.138 (-4.51)	0.121 (0.42)	-0.144 (-4.71)
DNC	-0.014 (-0.70)	-2.363 (-10.84)	-0.007 (-0.38)
T (actual tenure)	0.005 (5.71)		
PT (predicted tenure)			0.005 (5.97)
AGEBEGIN[a]		-2.702 (-46.01)	
$AGEBEGIN^2/100$		2.917 (32.03)	
CHILDBGN[a]		-3.267 (-37.92)	
TPARK[a]		-0.435 (-8.12)	
DADED[b]		-0.197 (-9.36)	
MOMKID[b]		-0.151 (-0.67)	
MARR[a]		3.062 (63.90)	
n	15,367	15,367	15,367
R-square	0.160	0.505	0.160

[a]For definitions, see Table 5.3

[b]For definitions, see Appendix D.

APPENDIX D. PREDICTING UNOBSERVED RESIDENCE SPELLS

The PSID data have no clear information on how long individual has lived in the state before 1968. Therefore the information on the starting time of every first spell started before 1968 is not perfectly observed. For those who lived in the same state where they grew up in 1968, we assumed that the beginning of their first spell is the year when they were 19 years old. For those who lived in a state different from the state where they grew up, the length of their first spells are not observed. Therefore, the selection process which separates between the observed and the unobserved length of first spells is the migration which took place before 1968.

When some observations are censored by a selection process, we could use Heckman's procedure to predict the unobserved by utilizing the observed ones (Heckman, 1979). The selection equation is the equation explaining the decision to migrate or to stay prior to 1968. To simplify the model, we fitted a standard probit model to estimate this selection equation. The explanatory variables were personal and local characteristics that are expected to affect migration before 1968.

Among 874 left-open spells which started before 1968, 585 are right-censored and 289 are right-closed spells. Because these two groups came from different populations of spells, we should estimate probit equations separately. In the right-censored group, the unobserved length of spells are 127. In the right-closed group, the unobserved length of spells are 138. Therefore there were 265 potentially unobserved spells. When the prediction of starting time was before the year when the person was 19, we used the year when he was 19 as the starting point. Similarly, when the prediction of the starting time was after the year when the individual was observed at first time in 1968 or later, we used the year when the individual was observed as the start. The last step reduced the number of unobserved spells from 265 to 208, out of the total 1268. The estimates of probit equations which explain migration decision before 1968 can be found in Table Appendix D.1. The table consists of two equations, one for the right-closed group and the other for the right-censored group. These equations are the selection equation which can be used to derive the Inverse Mill's Ratio (IMR), defined as $f(X\beta^*)/F(X\beta^*)$. The β^* is the vector of estimated coefficients in probit equation, X is the corresponding explanatory variables, and f(.) and F(.) are the corresponding normal density and distribution function.

To predict the unobserved length of spells, we utilized the information from the observed spells. Based on the observed spells, we regressed the log of spell on some explanatory variables, including the selection bias correction, IMR. Having fitted the equation for the observed spells, we used the estimated coefficients to predict the unobserved length of spells. The prediction equations corresponding to right-censored and right-closed group are presented in Table Appendix D.2.

Table D.1. Probit selection equation for predicting unobserved length of spell, t-statistics are in parentheses

Variable Name	Description	Right-closed Spells	Right-censored Spells
Dependent variable:			
M68	Dummy variable, equal to 1 if the person moved from the state he grew up before 1968 (the duration of the first spell is not observed), and equal to 0 otherwise		
Explanatory variables:			
CONSTANT	Intercept	-5.384 (-2.24)	-7.060 (-2.93)
DADEDU	Father's education (years)	0.048 (1.70)	-0.022 (-1.10)
MOMKID	Number of mother's children	-1.307 (-0.34)	0.275 (1.62)
EDU68	Education in 1968 (years)	0.011 (0.39)	0.059 (2.64)
AGE68	Age in 1968 (years)	0.168 (1.46)	0.153 (1.45)
$AGE68^2$	$AGE68^2/100$	-0.216 (-1.22)	-0.170 (-1.11)
RACE	Dummy variable, equal 1 if white and 0 otherwise	-0.587 (-2.03)	-0.694 (-3.28)
DSLFARM	Dummy variable, equal 1 if self employed or a farmer and 0 otherwise	-0.224 (-0.99)	-0.193 (-1.39)
Ln(PLAND/P)	State price of land (thousand $/acre)	0.156 (0.66)	0.056 (0.29)
URBAN	State percentage of urban population (percent)	2.208 (1.81)	1.221 (1.32)
PURATE	State predicted unemployment rate (percent)	-0.176 (-2.61)	0.172 (2.34)
PJOBGR	State predicted job growth	0.220 (2.17)	0.311 (2.58)
DNC	Regional dummy, equal 1 if living in the North Central, and 0 otherwise	-0.203 (-0.64)	0.361 (1.47)
DW	Regional dummy, equal 1 if living in the West and 0 otherwise	0.028 (0.06)	0.068 (0.16)
DS	Regional dummy, equal 1 if living in the South, and 0 otherwise	0.107 (0.23)	0.206 (0.49)

Table D.1. Probit selection equation for predicting unobserved length of spell, t-statistics are in parentheses (continued)

Variable Name	Description	Right-closed Spells	Right-censored Spells
CRIME	State crime rate relative to US average	0.005 (0.11)	-0.035 (-0.78)
PARK	Percentage of area belong to state and national park relative to US average	0.023 (0.37)	-0.042 (-0.63)
JAN	January temperature relative to US average	-0.004 (-0.24)	0.102 (0.76)
JULY	July temperature relative to US average	-0.043 (-1.67)	-0.013 (-0.66)
GRUPFARM	Dummy variable, equal 1 if individual grew up in farm and 0 otherwise	0.191 (1.03)	-0.215 (-1.38)
Log-likelihood		-200.0	-306.1
Number of observation		289	585
Number of M68=1 (Move)		138	127
Number of M68=0 (Do not move)		151	458

Table D.2. Prediction equation of unobserved length of residence spells, based on the observed group, t-statistics are in parentheses

Variable Name	Description	Right-closed Spells	Right-censored Spells
<u>Dependent variable:</u>			
LT	Log of the duration of residence spells, from the observed group		
<u>Explanatory variables:</u>			
CONSTANT	Intercept	-27.952 (-2.93)	-8.325 (-3.44)
ΔW2	Average annual residual wage in a spell	0.088 (3.66)	0.045 (4.62)
EDU68	Education in 1968 (years)	-0.011 (-1.22)	-0.000 (-0.18)
AGE68	Age in 1968 (years)	3.255 (2.60)	1.251 (4.02)
$AGE68^2$	$AGE68^2/100$	-13.897 (-2.29)	-5.578 (-3.77)
$AGE68^3$	$AGE68^3/1000$	2.650 (2.10)	1.108 (3.65)
$AGE68^4$	$AGE68^4/10000$	-0.188 (-1.92)	-0.082 (-3.55)
UNEMP	Average annual unemployment hour in a spell (hr/year)	0.0002(0.74)	-0.0000(-0.01)
DSLFARM	(see Table D.1)	0.213 (3.52)	0.021 (1.87)
RACE	(see Table D.1)	0.170 (1.37)	0.076 (2.47)
CHILD68	Number of school-age children	-0.009 (-0.61)	-0.001 (-0.34)
Ln(PLAND/P)	(see Table D.1)	0.165 (2.54)	0.061 (4.02)
URBAN	(see Table D.1)	-0.529 (-1.16)	-0.020 (-0.24)
PURATE	(see Table D.1)	0.208 (6.27)	0.063 (8.72)
PJOBGR	(see Table D.1)	-0.061 (-1.20)	0.030 (2.18)
DNC	(see Table D.1)	0.060 (0.71)	-0.076 (-3.79)
DW	(see Table D.1)	0.042 (0.38)	0.004 (0.12)
DS	(see Table D.1)	0.056 (0.49)	0.025 (0.89)
CRIME	(see Table D.1)	-0.014 (-1.05)	-0.082 (-2.40)
PARK	(see Table D.1)	-0.057 (-2.51)	-0.010 (-1.75)

Table D.2. Prediction equation of unobserved length of residence spells, based on the observed group, t-statistics are in parentheses (continued)

Variable Name	Description	Right-closed Spells	Right-censored Spells
JAN	(see Table D.1)	-0.046 (-1.17)	-0.007 (-6.29)
JULY	(see Table D.1)	0.036 (3.60)	0.013 (7.62)
DADEDU	(see Table D.1)	-0.014 (-1.28)	-0.002 (-0.90)
MOMKID	(see Table D.1)	0.005 (0.04)	-0.006 (-0.27)
λ	Selection bias correction (IMR)	-0.456 (-1.45)	-0.059 (-0.78)
R-square		0.904	0.845
Number of observation		151	458

APPENDIX E. NOMINAL AND REAL INTEREST RATES, 1953-93

Table E.1. Nominal and real interest rates, 1953-93

Year	Nominal Rates, 3 month T-Bills in percent (i)	GNP Implicit Price Deflator[a] (p)	Inflation in percent[b] (π)	Real Rates in percent[c] (r)
1953	1.93	26.39	0.87	1.05
1954	0.95	26.48	0.33	0.63
1955	1.75	26.68	0.76	0.99
1956	2.66	27.48	2.97	-0.30
1957	3.27	28.26	2.78	0.47
1958	1.84	28.78	1.81	0.03
1959	3.41	29.26	1.68	1.70
1960	2.93	29.72	1.56	1.35
1961	2.38	29.87	0.48	1.89
1962	2.78	30.21	1.15	1.61
1963	3.16	30.55	1.13	2.00
1964	3.55	30.90	1.12	2.40
1965	3.95	31.36	1.47	2.44
1966	4.88	32.30	2.97	1.85
1967	4.32	33.19	2.72	1.56
1968	5.34	34.48	3.82	1.47
1969	6.68	36.03	4.39	2.19
1970	6.46	37.75	4.67	1.71
1971	4.35	38.90	2.99	1.31
1972	4.07	40.08	2.98	1.06
1973	7.04	43.26	7.65	-0.56
1974	7.89	48.66	11.75	-3.46
1975	5.84	51.62	5.90	-0.06
1976	4.99	54.21	4.91	0.07
1977	5.27	56.85	4.75	0.49
1978	7.22	61.37	7.65	-0.40
1979	10.04	67.21	9.09	0.88
1980	11.51	74.16	9.85	1.51
1981	14.03	79.88	7.43	6.15
1982	10.69	83.80	4.79	5.63
1983	8.63	87.60	4.43	4.02
1984	9.58	90.70	3.48	5.90
1985	7.48	94.60	4.21	3.14
1986	5.98	97.00	2.51	3.39
1987	5.82	101.60	4.63	1.13
1988	6.69	106.10	4.33	2.26
1989	8.12	111.00	4.51	3.45
1990	7.51	117.50	5.69	1.72
1991	5.42	121.30	3.18	2.17
1992	3.45	125.30	3.24	0.20
1993	3.02	128.10	2.21	0.79

Sources: U.S. Department of Commerce, Survey of Current Business (1965-93), and The Council of Economic Advisers (1994)

[a]Personal Consumption Expenditure, 4th quarter

[b]$\pi_t = [(\mathrm{Ln}(p_t) - \mathrm{Ln}(p_{t-1})] \times 100$

[c]Calculated based on: $(1 + i) = (1 + r) \times (1 + \pi)$.