

A Poisson hurdle model of migration frequency

Alok K. Bohara and Randall G. Krieg*

University of New Mexico, University of Northern Iowa

Abstract. The United States is characterized by relatively high migration rates. This study examines whether high migration rates are due to many persons relocating or to fewer persons migrating more frequently. To examine migration frequency we use a Poisson hurdle model to discern factors disposing persons to make an initial move and to move more frequently than others. We compare a migration frequency equation to a binomial equation to examine the differences and similarities in both approaches to estimating migration behavior. Our findings show that traditional studies that model the migration decision as the potential for one move overlook the motivation for subsequent migration which accounts for a large share of the labor mobility in the United States. Other results indicate that while gender, education, and wages are important in the initial decision to migrate, these factors no longer play a role once individuals overcome the hurdle of an initial migration.

1. Introduction

High rates of migration play an important role in facilitating economic growth in capitalist countries by reallocating labor to its most productive use. According to the U.S. Bureau of the Census the annual intercounty migration rate in the late 1980s ranged between 17 percent and 18 percent (U.S. Department of Commerce 1990). Does this high migration rate represent the single movement of many persons or the repeated movement of a small number of persons? What determines whether some individuals choose to migrate more frequently than others? Is it crucial for an industrialized country to experience a relocation of many workers, as opposed to frequent movement of fewer workers?

Because the U.S. Census data are cross-sectional, we use a longitudinal data set, the *Panel Study of Income Dynamics* (PSID), to answer these questions. We examine the migration behavior of individuals from 1977 to 1987. Unlike traditional studies of individual migration, our study explains not only why persons make an initial move but also why some have a proclivity to move more frequently than others. Migrants in this study are those who change county of residence during a given year.

* The authors are associate professors at the University of New Mexico (Department of Economics, Albuquerque, NM 87131; Internet: bohara@unm.edu) and the University of Northern Iowa (Department of Economics, Cedar Falls, IA 50614-0401; Internet: krieg@uni.edu). Authors are listed alphabetically.

Because we are the first to estimate migration frequency econometrically, the relevant literature is limited. Prior to the development of comprehensive panel data sets a number of studies touted the importance of multiple migration and the frequency of migration by individuals. These studies are limited to descriptive analysis due to data limitations (Goldstein 1964; Taeuber 1961; Taeuber *et al.* 1961; Rowntree 1957). Using population registers of Copenhagen Goldstein finds that a small segment of the population moved frequently, while most of the population exhibited a high degree of residential stability. Taeuber (1961) uses duration-of-residence data from the *Current Population Survey* of May 1958 to examine multiple migration of individuals. Taeuber finds that 25 percent of adults in the United States live their entire lives in a single place and that nearly half of the entire population had lived more than 20 years in the current place of residence. These results are consistent with those of Goldstein, suggesting that high migration rates for the nation do not result from many persons migrating but from fewer persons migrating more frequently. Similar results are found by Taeuber *et al.* (1961) who use residence histories collected in the May 1958 supplement to the *Current Population Survey*. Their results also show that most adults experience a marked residential stability despite high rates of total residential mobility.

Myers *et al.* (1967) test the Cornell mobility model, i.e., that additional time spent in one area deepens ties to the location and consequently increases the cost of moving. Their findings support this theory, and they conclude that individual histories of movement are important for empirical analysis of migration.

The more recent development of comprehensive longitudinal data sets provides fertile ground for the study of migration frequency. While not modeling migration frequency per se, a number of studies of repeat migration have been done concerning persons who are moving for the second time and whether such moves are return moves to an area where the family lived before or onward moves to a place other than a family's prior residence (Morrison and DaVanzo 1986; Grant and Vanderkamp 1985; DaVanzo 1983, 1978; Schlottmann and Herzog 1982; Kau and Sirmans 1977). Such studies demonstrate the importance of accounting for multiple moves when explaining the causes and effects of migration. Reasons for multiple moves may include career change, unrealized expectations, changing regional conditions, imperfect information flows, and various life-cycle effects such as marriage and divorce.

2. Theoretical framework and data

Table 1 gives a frequency distribution of persons according to how many moves they made from 1977 to 1987. The predominant feature of the distribution is the strong concentration of observations at zero in an otherwise smooth distribution. Over four times as many persons chose not to migrate as compared to those who decided to migrate only once. After this hurdle is crossed and we consider those who moved once compared to those making multiple moves, however, the rate of decay remains relatively smooth. Overwhelming numbers of persons chose not to move at all. Those that do move are more likely to undertake future moves.

Table 1. Frequency distribution of the number of moves

	Number of Moves							
	0	1	2	3	4	5	6	7
Individuals	1591	372	178	87	32	12	4	3
% frequency	69.8	16.3	7.8	3.8	1.4	.5	.2	.1

These observations lead us to believe that the statistical model determining the binary outcome of the number of moves as zero or one differs from that determining the frequency of moves. In standard count data models such as the Poisson these two processes are forced to be identical. Therefore, instead of the standard Poisson model we estimate a model that allows the binary choice process to differ from the frequency estimation. We believe that the migration decision distinguishes persons in an important way. Persons who uproot themselves and possibly other family members by migrating are different from those unwilling to undertake such a venture. Because of this observed difference, our model is separated into two parts.

First we estimate whether individuals decide to migrate. Conditional on the migration decision, we estimate the frequency of moves. It is natural to allow the statistical model governing the binary outcome (move or not to move) to differ from that which determines the frequency of migration. This is known as a *hurdle model* (Mullahy 1991). Any model that does not allow such conditional structure, e.g., a standard Poisson, will result in inconsistent estimates (Grogger and Carson 1991). We estimate both the decision to make an initial move and to make additional moves using the same explanatory variables in each equation. This allows us to compare the estimated coefficients of these variables in both equations directly. If the models yield significantly different results, we must conclude that simply modeling the migration decision as the potential for one move (as in traditional studies) omits the motivation of why many persons make frequent additional moves.

2.1. The data

The sample for our study is taken from the PSID from 1977 to 1987. We select only those between the ages of 18 and 65 who are nonstudents, eliminating those likely to move solely for schooling or retirement. Our sample includes 2,279 individuals. The PSID data set is longitudinal and contains microdata for a representative sample of individuals and their family units residing in the United States. This data set is ideal for this study because it contains information for each year studied concerning personal characteristics of individuals as well as their county of residence.

3. Econometric model and methodology

The methodology adopted is based on a hurdle Poisson specification (Mullahy 1986; Bohara and Krieg 1995). First the decision to migrate is governed by the binary out-

come of whether an individual migrates or not. Once the migration decision has been made, then the number of times an individual migrates within a given time frame is allowed to follow a truncated-at-zero Poisson model. When an individual makes a migration decision implying a positive realization, it is known as *crossing a hurdle*.

The joint likelihood function for the hurdle Poisson model is:

$$L = \prod_{t=1}^n [\Pr(y_{1t} = 0)]^{y_{1t}} [1 - \Pr(y_{1t} = 0)]^{1-y_{1t}} [\Pr(y_{2t}|y_{2t} > 0)]^{y_{1t}} \quad (1)$$

where:

- y_{1t} = 1 if migrant and 0 for nonmigrant; and
- y_{2t} = The frequency of migration.

The first two terms represent the binary outcome representing migration decision, and the last term captures the conditional Poisson process.

Equation (2) gives the joint likelihood function:

$$L = \prod_{t=1}^n [F(X_{1t}\beta_1)]^{y_{1t}} [1 - F(X_{1t}\beta_1)]^{1-y_{1t}} \left[\lambda_t^{y_{2t}} / ((\exp(\lambda_t) - 1) y_{2t}!) \right]^{y_{1t}} \quad (2)$$

where the decision whether or not an individual migrates is represented by a probit function $F(\cdot)$. The set of independent variables X_{1t} affects the migration decision. The frequency of migration is influenced by X_{2t} through a link function, $\lambda_{2t} = \exp(X_{2t}\beta_2)$ which ensures that the expected frequencies are positive. The hurdle Poisson model is also suitable for taking into account the overdispersion or underdispersion of data (Mullahy 1986). Detailed definitions of variables are presented in Table 2.

The variables determining the migration decision include the following:

- X_{1t} ;
- CONSTANT;
- EDU;
- WAGE;
- DSPWORK;
- AGE;
- DMALE;
- DUNION;
- DWHITE;
- DOWNHOUSE;
- DBORN;
- NCHILD;
- DPROF;
- DMARR;
- DDISAB; and
- EMPLOYER.

Characteristics contributing to residential stability include a working spouse, age, union membership, home ownership, living in county of birth, number of children,

Table 2. Variable names and definitions

Variables	Definitions	Mean/std. dev.
DMIG	= 1 if migrated, 0 otherwise	.302 (.459)
FMIG	= Frequency of migration	.536 (1.008)
EDU/10	= Years of education	1.253 (.278)
WAGE/10	= Hourly wage in dollars	.607 (.408)
DSPWORK	= 1 if spouse works, 0 otherwise	.480 (.500)
AGE/10	= Age	3.427 (1.043)
DMALE	= 1 if male, 0 female	.855 (.352)
DUNION	= 1 if union member, 0 otherwise	.210 (.408)
DWHITE	= 1 if white, 0 otherwise	.714 (.452)
DOWNHOUSE	= 1 if owns a house, 0 otherwise	.563 (.496)
DBORN	= 1 if born in county of residence, 0 otherwise	.669 (.471)
NCHILD	= Number of children under 17	1.340 (1.364)
DPROF	= 1 if professional, 0 otherwise	.935 (.246)
DMARR	= 1 if married, 0 otherwise	.723 (.448)
DDISAB	= 1 if disabled, 0 otherwise	.068 (.253)
EMPLOYER/100	= Length of job tenure with current employer	.683 (.842)
WGROWTH	= Growth rate of wage	1.504 (1.772)
FEMP	= Number of employer changes	5.101 (1.776)
FMARR	= Number of marital dissolutions	.182 (.426)
FOCC	= Number of occupation changes	4.381 (2.214)

marriage, physical disability, and number of months employed by current employer. A working spouse increases the cost of a move due to an additional job search and possible lower spouse earnings resulting from a move. As individuals age, their expected future gains from migration decline, decreasing the incentive to move. Union membership promotes job stability and thus decreases the need to migrate to find alternative employment. Homeowners face an additional cost when moving due to the need to sell a house. Owners of homes also may have higher moving expenses due the accumulation of durable goods. As interpreted by studies of return migration (DaVanzo 1983), living in the county of birth (DBORN) implies that a future move will not be a return move. These persons, if they decide to move, must give up location-specific human capital and relocate to a place where they have less labor market information. A spouse and children add not only to the pecuniary cost of a

move, but also to psychic costs such as acclimation to new friends, environment, and schools. Number of months working for an employer proxies job-specific human capital acquired on the job that may be sacrificed when making a move. Being a highly educated male tends to increase migration proclivity; thus, we expect positive signs for DMALE and EDU. The same set of independent variables is used to estimate the initial migration equation, X_{1t} , and the frequency equation, X_{2t} . This allows a direct comparison of the two equations

4. Empirical results

The Broyden, Fletcher, Goldfarb, and Shannon algorithm in Gauss is used to optimize the log-likelihood function. The results are presented in Table 3.

Of primary importance in our discussion of the results is whether the causes of an initial migration are the same as those of more frequent migration, i.e., do the explanatory variables have the same influence in both equations? The equality of coefficients between the binary decision equation and the conditional Poisson equation is tested using the likelihood ratio test. The number of equality restrictions is 15, which is the number of independent variables including the constant term. The calculated chi-squared value is 33.58 which is greater than the critical value of 25.00 at the 5 percent level and with 15 degrees of freedom. This indicates that the null hypothesis that explanatory variables have the same influence in both equations is rejected. That is, explaining why persons make an initial move and explaining why some persons move more frequently than others should be done using different statistical models. Thus, modeling the migration decision as the potential for one move omits the motivation of why many persons move, i.e., why they make subsequent moves. To capture this additional information, we must account for migration frequency. Thus, we will confine our discussion to the unrestricted model.

Age significantly influences both the initial migration decision and migration frequency at the 1 percent level. The significance of age is consistent with human capital theory, i.e., that persons will be less willing to invest if the length of the return will be shorter. The number of children influences migration frequency to a higher degree than it influences an initial move.

Union membership deters migration, reflecting increased job attachment. This is consistent with Freeman and Medoff's argument that because unions handle worker grievances, there will be lower quit rates among members (1982). Living in the county of birth also decreases the tendency to migrate. This is not surprising; those living in their county of birth are not potential return migrants, i.e., if they relocate, they typically face a loss of more location-specific human capital than other potential migrants. In addition, those living in an area for a long period of time may develop more personal ties. This suggests that the cost of reorienting children to new schools, friends, and other relevant psychic costs that accompany moves increase, on the margin, with additional moves. Being white increases mobility, although race has a lesser impact once the hurdle of an initial move is crossed.

Most striking among the results is that while education, wage, and gender influence initial migration, they have no effect on migration frequency. Once individuals

Table 3. Maximum likelihood estimates of the hurdle Poisson model (standard error)

Variables	Unrestricted model		Restricted model
	Binary outcome (decision to migrate)	Conditional Poisson (frequency of migration)	Binary outcome (coefficients restricted to Poisson coefficients)
Constant	.269 (.252)	1.500*** (.355)	.692* (.409)
EDU	.220* (.134)	-.254 (.184)	.094 (.351)
WAGE	.138* (.078)	.028 (.115)	.103 (.073)
DSPWORK	-.004 (.076)	-.083 (.125)	-.035 (.153)
AGE	-.250*** (.038)	-.227*** (.067)	-.282*** (.040)
DMALE	.230** (.110)	.070 (.144)	.151 (.117)
DUNION	-.236*** (.081)	-.325* (.177)	-.292*** (.089)
DWHITE	.593*** (.077)	.279* (.153)	.574*** (.101)
DOWNHOUSE	-.462*** (.072)	-.181 (.122)	-.418*** (.071)
DBORN	-.659*** (.064)	-.322*** (.088)	-.614*** (.059)
NCHILD	-.056** (.027)	-.127*** (.048)	-.082*** (.024)
DPROF	-.091 (.117)	-.120 (.152)	-.110 (.160)
DMARR	-.071 (.104)	.078 (.154)	-.104 (.092)
DDISAB	.038 (.122)	-.005 (.172)	.026 (.067)
EMPLOYER	-.067 (.046)	.024 (.078)	-.043 (.059)
Average Log Likelihood \times n	-.86754 \times 2279		-.882908 \times 2279

*** Statistically significant at the 1 percent level

** Statistically significant at the 5 percent level

* Statistically significant at the 10 percent level

cross the threshold of an initial move, education, wages, home ownership, and gender no longer matter. Thus, crossing this hurdle distinguishes individuals in an important way. For example, once women migrate, their migration behavior is no different from that of men. Education, which has been shown to be important in the decision to migrate in past studies of the determinants of migration, is shown here to be an important determinant in the initial decision to relocate only. Once persons decide to move, education no longer plays a role in the frequency of moves. Thus, by distinguishing initial migration and migration frequency, we are able to discern the role different variables play in migration decisions more clearly.

5. Conclusion

Several insights have been gained from the study of migration frequency. Although the migration rate in the United States is high, this largely represents the relatively

high mobility of a small portion of the population. Even in an industrialized society such as the United States it is not necessary for a large number of persons to be mobile when there are fewer persons who are willing to move more frequently.

Given the prominence of frequent migration, it is important that we model migration behavior with this in mind. Our results indicate that modeling only the decision to make a single move leaves out important information that affects multiple moves by the same individuals.

Factors that explain a more traditional model of the migration decision do not necessarily influence migration frequency.¹ The most striking examples are education and gender. In traditional estimations of one time moves men and those with high levels of education are significantly more predisposed to migrate. This study shows that once one overcomes the hurdle of migrating once, gender and education no longer play a role in future migration decisions. Our results indicate that frequent movers are less rooted, tending to live in a county other than their birth, to have fewer children, and to not belong to unions. Age and race are also factors. Whites are more likely to move frequently, as are younger persons.

It is critical that research be done to further develop models of migration frequency. Factors examined here as well as other considerations influencing migration frequency deserve further study, given the large facilitating role played by frequent moves in the efficient redistribution of labor resources.

References

- Bohara, Alok K., and Randall G. Krieg, "A Zero-Inflated Poisson Model of Migration Frequency," *International Regional Science Review* (forthcoming).
- DaVanzo, Julie, "Repeat Migration in the U.S. Who Moves Back and Who Moves On?" *The Rand Paper Series* (1978), p. 59-61.
- DaVanzo, Julie, "Repeat Migration in the United States: Who Moves Back and Who Moves On?" *Review of Economics and Statistics*, 65 (1983), pp. 552-559.
- Freeman, Richard B., and James L. Medoff. *What do Unions Do?* (New York: Basic Books, 1982).
- Goldstein, Sidney, "The Extent of Repeated Migration: An Analysis Based on the Danish Population Register," *Journal of the American Statistical Association*, 59 (December 1964), pp. 1121-1132.
- Kau, James B., and C.F. Sirmans "The Influence of Information Cost and Uncertainty on Migration: A Comparison of Migrant Types," *Journal of Regional Science*, 17, no. 1 (1977), pp. 89-96.
- Grant, E. Kenneth, and John Vanderkamp, "Migrant Information and the Remigration Decision: Further Evidence," *Southern Economic Journal*, 51, no. 4 (1985), pp. 1202-1215.
- Grogger, J.T., and R.T. Carson, "Models for Truncated Counts," *Journal of Applied Econometrics*, 6 (1991), pp. 225-238.
- Morrison, P., and Julie DaVanzo, "The Prism of Migration: Dissimilarities Between Return and Onward Movers," *Social Science Quarterly*, 67, no. 3 (1986), pp. 504-516.
- Mullahy, John, "Specification and Testing of Some Modified Count Data Models," *Journal of Econometrics*, 33 (1986), pp. 341-365.

¹ As pointed out by an anonymous referee, moves (particularly multiple moves) may be influenced by whether a change of employment status is made by an employer. This is important in the case of the military. Data limitations preclude inclusion of such variables here.

- Myers, George C., Robert McGinnis, and George Masnick, "The Duration of Residence Approach to a Dynamic Stochastic Model of Internal Migration: A Test of the Axiom of Cumulative Inertia," *Eugenics Quarterly*, 14 (1967), pp. 121-126.
- Rowntree, John A., "Internal Migration, A Study of the Frequency of Movement of Migrants," *Studies on Medical and Population Subjects*, No. 11 (London: Her Majesty's Stationery Office, 1957).
- Schlottmann, Alan M., and Henry W. Herzog, "Home Economic Conditions and the Decision to Migrate: New Evidence for the U.S. Labor Force," *Southern Economic Journal*, 48, no. 4 (1982), pp. 950-961.
- Taeuber, Karl E., "Duration-Residence Analysis of Internal Migration in the United States," *Milbank Memorial Fund Quarterly*, 39 (1961), pp. 1-16.
- Taeuber, Karl E., William Haenszel, and Monroe G. Sirken, "Residence Histories and Exposure Residences for the United States Population," *Journal of the American Statistical Association*, 56 (1961), pp. 824-834.
- U.S. Department of Commerce, Bureau of the Census, "Perspectives on Migration Analysis," *Current Population Report*, P-23, no. 166 (August 1990).

with a model posited by Durlauf. The worst neighborhoods, however, face strong negative feedback effects. The negative feedback effects in several indicators (real mean family income, percent of persons with incomes below the poverty line, and percent of families with female heads) for the worst neighborhoods are consistent with regression toward neighborhood-specific mean values.

The effects of control variables are also briefly discussed, including variables related to Wilson's conjectures regarding neighborhood change and variables indicating the local government fiscal environment.

Findings suggest that large shocks or interventions are needed for a disadvantaged neighborhood to improve sufficiently for its negative feedback effects to diminish to the levels faced by most neighborhoods. The long-term effects of small shocks or interventions are likely to be reduced 40 to 80 percent by negative feedback effects associated with changes in quality.

1. Introduction

Neighborhood quality is an important concern for public policy makers and researchers because of the potential relationship between the quality of neighborhoods in which children grow up and their propensity to succeed. An implication of this potential relationship may be:

[S]hocks or policy interventions that positively affect individuals will have positive multiplier effects within neighborhoods through peer influences and across generations through family influences (Case and Katz 1991).

The potential benefit of exploiting these multiplier effects is one important reason to determine if they exist. It is also important in public policy assessments and allocations to know if such multiplier effects do not exist.

These multiplier effects are positive feedback loop effects in neighborhood quality. *Feedback loops* are dynamic processes that affect themselves. Positive feedback loops are self-reinforcing processes. Negative feedback loops are self-restraining processes. The adjectives positive and negative refer to the sign of the relationship between changes, not the normative connotation or the direction of the changes.¹ Feedback loops, whether positive or negative, are important in neighborhood evolution.

¹ Changes in the money supply or government fiscal policy frequently described as having multiplier effects are examples of positive feedback loops. One example of negative feedback loop effects is the automatic stabilizer effects of government activities that increase consumption during recessions (such as unemployment insurance, welfare, and reduced taxes) and decrease consumption during booms (such as reduced unemployment and welfare payments and increased taxes). Tietenberg (1992) gives several examples of positive and negative feedback loops. Positive feedback loops include investment that generates profits that can be used for additional investment, global warming that releases greenhouse gases that increase global warming, and shortages that lead to hoarding which leads to increased shortages. Negative feedback loops include population growth (that leads to increased pollution that leads to higher death rates and retards population growth) and self-regulating environmental processes or ecosystems.