The Effect of Expected Income on Individual Migration Decisions

John Kennan and James R. Walker[1]

University of Wisconsin-Madison and NBER

July 2001

[1]Department of Economics, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706; jkennan@ssc.wisc.edu and walker@ssc.wisc.edu. The National Science Foundation (Kennan) and the NICHD (Walker) provided research support. We are grateful to Ken Wolpin, Michael Keane, Derek Neal, Peter Arcidiacono, Phil Haile, Jim Ziliak and seminar and conference participants at Yale, Iowa, Duke, Rochester and Wisconsin for helpful comments.

1 Introduction

Migration is a complex process that has been studied in several branches of the social sciences. Even within the economics literature, the number of conjectured determinants and the range of consequences considered is impressively diverse (see Greenwood [1997] and Lucas [1997] for recent surveys). In this paper we focus on expected income as the main economic influence on migration.

We build an economic model of individual migration decisions, and estimate it on panel data (the NLSY79). Estimating a structural dynamic model of the destination-specific streams has apparently not been done before, perhaps because the required computations have not been feasible.[2] Our basic empirical question is the extent to which people move for the purpose of improving their income prospects. Work by Keane and Wolpin (1997) and by Neal (1999) indicates that individuals make surprisingly sophisticated calculations regarding schooling and occupational choices. Given the magnitude of geographical wage differentials, and given the findings of Topel (1986) and Blanchard and Katz (1992) regarding the responsiveness of migration flows to local labor market conditions, we would expect to find that income differentials play an important role in individual migration decisions.[3]

We model individual decisions to migrate as a job search problem in which welfare benefits or other alternative sources of income act as a floor, insuring workers against bad job search outcomes.[4] A worker can draw a wage only by visiting a location, thereby incurring a moving cost. Locations are distinguished by known differences in mean wages, amenity values and alternative income sources. A worker starts the life-cycle in some home location and must determine the optimal sequence of moves before settling down. There is a two-dimensional ranking of locations, *ex ante*: some places have high mean wages, and others have attractive fallback options (both adjusted for amenity values). In addition we allow for a bias in favor of the home location.

The decision problem is too complicated to be solved analytically, so we use a discrete approximation that can be solved numerically, following Rust (1994). The parameters of the model include a discount

[2] Holt (1996) estimated a dynamic discrete choice model of migration, but his framework modeled the move/stay decision and not the state-specific flows.

[3] Blanchard and Katz (1992, p.2), using average hourly earnings of production workers in manufacturing, by state, from the BLS establishment survey, describe a pattern of "strong but quite gradual convergence of state relative wages over the last 40 years." For example, using a univariate AR(4) model with annual data, they find that the half-life of a unit shock to the relative wage is more than 10 years. Similar findings were reported by Barro and Sala-i-Martin (1991) and by Topel (1986).

[4] This differs from the standard job search model in which unemployment benefits are treated as a subsidy received while search continues. In our model, welfare provides a safety net in case the search fails.

factor, a risk aversion coefficient and a home premium summarizing individual preferences; moving costs, including a fixed cost and a cost that is proportional to distance; means and variances of wages in each location; a relative variance parameter governing the extent to which individual wage offers are correlated across locations and a persistence parameter governing the relative importance of permanent and transitory components of wages. We also allow for differences in location size, measured by the population in origin and destination locations.

2 An Optimal Search Model of Migration

We model migration as an optimal search process. The basic assumption is that wages are local prices of individual skill bundles. The individual knows the wage in the current location, but not in other locations, and in order to determine the wage at each location, it is necessary to move there, at some cost. In each location there is also a fallback option, such as welfare or family support, and the value of this is known.

The model aims to describe the migration decisions of young workers in a stationary environment. The wage offer in each location may be interpreted as the best offer available in that location. Wages are permanent, so the only chance of getting a better offer is to move to a new location. It may be that wage differentials across locations equalize amenity differences, but a stationary equilibrium with heterogeneous worker preferences and skills still requires migration to redistribute workers from where they happen to be born to their equilibrium location. Alternatively, it may be that wage differentials are slow to adjust to location-specific shocks, because gradual adjustment is less costly for workers and employers. In that case, our model can be viewed as an approximation in which workers take current wage levels as a rough estimate of the wages they will face for the foreseeable future. In any case, the model is intended to describe the partial equilibrium response of labor supply to wage differences across locations; from the worker's point of view the source of these differences is immaterial, provided that the differences are permanent. A complete equilibrium analysis would of course be much more difficult, but our model can be viewed as a building-block toward such an analysis.

Suppose there are J locations, and individual i's wage W_{ij} in location j is is a random variable with distribution function F_{ij}. The fallback option is b_j, and thus income in location j is $y_{ij} = \max [W_{ij}, b_j]$. Migration decisions are made so as to maximize the expected discounted value of lifetime utility, subject to budget constraints. Consider a person with "home" location h, who is in location R this period and in location j next period. The flow of utility in the current period for such a person is specified as

$$u_h(C;\ell,j) = \alpha \frac{C^{1-\gamma}-1}{1-\gamma} + a_\ell + \kappa\chi_{\{j=h\}} - \delta_0 \chi_{\{j\neq\ell\}} - \delta_1 M(\ell,j)$$

The notation is as follows. C is consumption in the current period and $\gamma \geq 0$ is a constant relative risk aversion coefficient. The value of amenities is a_ℓ, and there is a premium κ that allows each individual to have a preference for their native location (χ_A is used as an indicator meaning that the statement A is true). The disutility associated with moving is an affine function of the distance $M(\ell,j)$ from origin to destination.

In general, the level of assets is an important state variable for this problem, but we will focus our analysis on a special case in which assets do not affect migration decisions. Suppose the marginal utility of income is constant ($\gamma = 0$ in the specification above), and individuals can borrow and lend without restriction at a given interest rate. Then expected utility maximization reduces to maximization of expected lifetime income, net of moving costs, with the understanding that the value of amenities is included in income, and that both amenity values and moving costs are measured in consumption units.[5] This is a natural benchmark model, although of course it imposes strong assumptions.[6]

There is little hope of solving this problem analytically. In particular, the Gittins index solution of the multiarmed bandit problem cannot be applied because there is a cost of moving.[7] But by using a discrete approximation of the wage distribution in each location, we can compute the value function and the optimal decision rule by standard dynamic programming methods, following Rust (1994).

Let F_j be the wage distribution function in location j. We approximate this by a discrete distribution over n points, as follows. Let $a_s^j = F_j^{-1}\left(\frac{s}{n} - \frac{1}{2n}\right)$, where s = 1,2,...,n. Then F_j is approximated by a uniform distribution over the set $\{a_s^j\}_{s=1}^n$. For example, if n = 10, the approximation puts probability 1/10 on the 5th, 15th, ... 95th percentiles of the distribution F.

[5] Note that this neatly sidesteps the question of whether moving costs should be specified as "psychic" costs that directly reduce utility, or as monetary costs that reduce disposable income. With constant marginal utility of income, there is no meaningful difference between these two specifications.

[6] Even if the marginal utility of consumption is not constant, one can still compute the increase in current-period consumption needed to just offset the utility cost of moving, and use this to translate the utility cost into an income equivalent. Then the optimal migration problem can be viewed as maximization of net lifetime income, and this will be a good approximation if the compensating variation in consumption is roughly constant. But this argument rests on the assumption that the individual can borrow against future income (including income generated by a move) in order to sustain current consumption.

[7] See Banks and Sundaram (1994) for an analysis of the Gittins index in the presence of moving costs.

2.1 The Value Function

Consider a person currently in location ℓ, with a J-vector ω summarizing what is known about wages in all locations. Here ω_j is either 0 or an integer between 1 and n, with the interpretation that if $\omega_j = s > 0$, then the wage in location j is known to be a^j_s, and if $\omega_j = 0$ then the wage in location j is still unknown, so that if the person moves to j, the wage will be a^j_s with probability 1/n, for $1 \le s \le n$. The value function for a native of location h can be written in recursive form as

$$
V_h(\ell,\omega) = \begin{cases} \dfrac{1}{n}\sum\limits_{s=1}^{n} V_h(\ell,\omega_1,\ldots,\omega_{\ell-1}, s, \omega_{\ell+1},\ldots,\omega_J) & \text{if } \omega_\ell = 0 \\[2ex] \max\limits_{j} \left[u_h\Big(y_\ell(\omega_\ell) - \Delta(\ell,j);\ \ell,j \Big) + \beta\, V_h(j,\omega) \right] & \text{if } \omega_\ell > 0 \end{cases}
$$

where $y_j(\omega) = \max[b_j, a^j_{\omega_j}]$, and $\Delta(\ell,j)$ is the monetary cost of moving from ℓ to j.

We compute V_h by value function iteration. It is convenient to use $V_h(\ell,\omega) \equiv 0$ as the initial estimate, so that if T is the number of iterations, the result gives the optimal policy for a (rolling) T-period horizon.

3 Empirical Implementation

An important limitation of the discrete dynamic programming method is that the number of states is typically large, even if the search problem is relatively simple. If there are J locations and the discrete approximation of the wage distribution has n points of support, the number of states is $J(n+1)^J$. For example a model with J=5 and n=10 has 805,255 states. Although value functions for such a model can be computed in a few hours, estimation of the structural parameters requires that the value function be computed many times. Estimation becomes infeasible unless the number of structural parameters is small.

Ideally, locations would be defined as local labor markets. The smallest geographical unit identified in the NLSY is the county, but we obviously cannot let J be the number of counties, since there are over 3,100 counties in the U.S. Indeed, even restricting J to the number of states still far exceeds current computational capabilities. To aggregate locations beyond the state level (e.g. Census Regions) is uninterpretable; for example, we lose the ability to identify the effects of state benefit systems. Consequently, we define locations as states, but restrict the information available to each individual, as explained below.

3.1 Outline of the Estimation Method

We first expand the model to allow for unobserved heterogeneity in individual payoffs. Let $\zeta = (\zeta_1, \zeta_2, \ldots, \zeta_J)$ be a vector of idiosyncratic utility adjustments that are known to the worker before the migration decision is made in each period (but not observed by the econometrician). We assume that each component ζ_j is drawn independently according to a distribution function π; also, these draws are independent across individuals and over time. The individual's value function is then given by

$$V_h(\ell,\omega,\zeta;\theta) = \begin{cases} \sum_{s=1}^{n} p_s V_h(\ell,\omega_1,\ldots,\omega_{\ell-1},s,\omega_{\ell+1},\ldots,\omega_J,\zeta;\theta) & \text{if } \omega_\ell = 0 \\ \max_j \left[u_h\left(y_\ell(\omega_\ell) - \Delta(\ell,j); \ell, j \right) + \zeta_j + \beta \bar{V}_h(j,\omega;\theta) \right] & \text{if } \omega_\ell > 0 \end{cases}$$

where θ is the vector of unknown parameters and the expected value function $\bar{V}$ is defined by

$$\bar{V}(j,\omega;\theta) \equiv \int V(j,\omega,\zeta;\theta)\, d\pi(\zeta)$$

If π is the Type 1 Extreme Value distribution[8] then, using arguments due to McFadden (1973) and Rust (1987) we can show that the function $\bar{V}$ satisfies

$$\bar{V}_h(\ell,\omega;\theta) = \begin{cases} \sum_{s=1}^{n} p_s \bar{V}_h(\ell,\omega_1,\ldots,\omega_{\ell-1},s,\omega_{\ell+1},\ldots,\omega_J;\theta) & \text{if } \omega_\ell = 0 \\ \log\left(\sum_{j=1}^{J} \exp\left[v_{hj}(\ell,\omega;\theta) \right] \right) & \text{if } \omega_\ell > 0 \end{cases}$$

where

$$v_{hj}(\ell,\omega;\theta) = u_h\left(y_\ell(\omega_\ell) - \Delta(\ell,j);\ \ell, j \right) + \beta \bar{V}_h(j,\omega;\theta)$$

This gives the probability, $\Pr(d(j)=1 \mid h, \ell, \omega)$, that a native of h in location ℓ with information ω will move to location j:

---

[8] A random variable X is exponentially distributed if exp(-X) is uniformly distributed on [0,1]. Repeating this yields an extreme value distribution: Y has the extreme value distribution if exp(-Y) is exponentially distributed.

$$Pr(d(j) = 1 | h,\mathsf{R},\omega;\theta) = \frac{\exp[v_{hj}(\mathsf{R},\omega;\theta)]}{\sum_{\tau=1}^{J} \exp[v_{h\tau}(\mathsf{R},\omega;\theta)]}$$

3.2 A Limited History Approximation

When the number of locations is moderately large, the model becomes computationally infeasible.[9] This is a common problem with discrete dynamic programming models, and various devices have been proposed to deal with it. In our context it seems natural to use an approximation that takes advantage of the timing of migration decisions. So far, we have assumed that information on the value of human capital in alternative locations is permanent, and so if a location has been visited previously, the wage in that location is known, no matter how much time has passed. This means that the number of possible states increases exponentially with the number of locations visited. In practice, however, the number of people seen in many distinct locations is very small. Thus by restricting the information set to include only wages seen in recent locations, it is possible to drastically shrink the state space while retaining most of the information actually seen in the data. Specifically, we suppose that the number of wage observations cannot exceed M, with M < J, so that it is not possible to be fully informed about wages at all locations. Then if the wage distribution in each of J locations has n points of support, the number of states is $(Jn)^M$, since this is the number of possible M-period histories describing the locations visited most recently, and the wages found there. For example, if J is 50 and n is 5 and M is 2, the number of states is 62,500, which is manageable.

This approximation reduces the number of states in the most obvious way: we simply delete most of them. Someone who has "too much" wage information in the big state space is reassigned to a less-informed state. Individuals makes the same calculations as before when deciding what to do next, and the econometrician uses the same procedure to recover the parameters governing the individual's decisions. There is just a shorter list of states, so two people who have different histories may be in different states in the big model, but they are considered to be in the same state in the reduced model. In particular, two people who have the same recent history are in the same state, even if their previous histories were

[9]And it will remain so, even if computers improve: for example, if a location is a State, and the wage distribution has 5 support points, then the number of dynamic programming states is 40,414,063,873,238,203,032,156,980,022,826,814,668,800.

different (and two people who have different wage information now may have the same wage information following a move).

In order to compute the likelihood function using this approximation, it is convenient to redefine notation. Let $R = (R^0, R^1, \ldots R^{M-1})$ be an M-vector containing the sequence of recent locations (beginning with the current location), and let ω be the corresponding sequence containing recent wage information. Then the probability that an individual in state (R,ω) will move to location j can again be written in the form

$$Pr(d(j) = 1 | h, R, \omega; \theta) = \frac{\exp[v_{hj}(R,\omega;\theta)]}{\sum_{\tau=1}^{J} \exp[v_{h\tau}(R,\omega;\theta)]}$$

where v_j is now defined as

$$v_{hj}(R,\omega;\theta) = u_h\left(y_{R^0}(\omega_{R^0}) - \Delta(R^0, j) \right) + \beta \bar{V}_h\left((j, R^0, R^1, \ldots, R^{M-2}), (\omega^j, \omega^0, \omega^1, \ldots, \omega^{M-2}); \theta \right)$$

with

$$\bar{V}_h(R,\omega;\theta) = \begin{cases} \dfrac{1}{n} \sum_{s=1}^{n} \bar{V}_h(R, (s, \omega^1, \omega^2, \ldots, \omega^{M-1}); \theta) & \text{if } \omega^0 = 0 \\ \log\left(\sum_{j=1}^{J} \exp[v_{hj}(R,\omega;\theta)] \right) & \text{if } \omega^0 > 0 \end{cases}$$

3.3 Population Effects

It has long been recognized that location size matters in migration models.[10] California and Wyoming cannot reasonably be regarded as just two alternative places, to be treated symmetrically as origin and destination locations. To take one example, a person who moves to be close to a friend or relative is more likely to have friends or relatives in California than in Wyoming. A convenient way to model this in our framework is to allow for more than one draw from the distribution of preference shocks in each location. Specifically, we assume that the number of draws per location is an affine function of the number of people already in that location, and that migration decisions are controlled by the maximal draw for each location. This leads to the following modification of the logit function describing migration probabilities:

[10] See T. Paul Schultz (1982).

$$Pr(d(j) = 1 | h, R, \omega; \theta) = \frac{(1 + \delta_3 n_j) \exp[v_{hj}(R, \omega; \theta)]}{\sum_{\tau = 1}^{J} (1 + \delta_3 n_\tau) \exp[v_{h\tau}(R, \omega; \theta)]}$$

Here n_j denotes the population in location j, and the (nonnegative) parameter δ_3 can be interpreted as the number of additional draws per person.

3.4 Computation

Since the parameters are embedded in the value function, computation of the gradient and hessian of the loglikelihood function is not a simple matter (although in principle these derivatives can be computed in a straightforward way using the same iterative procedure that computes the value function itself). We maximize the likelihood using an "amoeba" algorithm that implements the downhill simplex method of Nelder and Mead. This method does not use derivatives, and it seems appropriate for problems such as this in which there is no reason to expect that the loglikelihood function is concave. In practice the method works well for the models we have estimated so far; in particular, it is robust to large changes in the starting values of the parameters. On the other hand, the method is slow, and so we also use gradient methods to speed up the computations, particularly when doing sensitivity analysis.[11]

4 Migration and Welfare

We analyze the migration decisions of low income women at risk to receive AFDC. This is a natural application of our model, because location-specific benefits in the model are most directly related to welfare benefits (AFDC and Food Stamps) within each state.

The recent literature on welfare-induced migration is summarized by Meyer (1999). While the consensus view from earlier work reviewed by Moffitt (1992) was that differences in welfare benefits across states had a significant effect on migration decisions, subsequent studies by Levine and Zimmerman (1995) and by Walker (1994) found little or no effect. Meyer argued that by paying careful attention to the determinants of welfare participation, the ambiguity in these results can be resolved in favor of a significant (but small) effect of welfare on migration. Gelbach (2000) also found a significant effect, arguing that previous studies had failed to properly account for dynamic selection effects. None of

[11] An example of our (fortran) computer program can be found at www.ssc.wisc.edu/~jkennan/research/mbr8.f90.

these studies contains a complete dynamic choice model, however, and we believe that our model can provide a more systematic analysis.

4.1 Earnings and Welfare Benefits

We restrict the estimation sample to women from the non-military subsample of the NLSY79 with twelve or fewer years of education. The observational window begins in the year the woman is first single with a dependent child. To be included in the sample, information on residence must be observed for at least two periods. We follow these respondents either until the end of their single parenthood, the end of the sample period (1992) or the first wave in which they are not interviewed. There were 1,728 people satisfying these restrictions, and we have data on 10,101 location decisions (i.e. person-years).

For each respondent, the wage in each State is measured as the sum of annual wage and salary income divided by total weeks worked, for all years residing in that State. We use PUMS data from the 1990 Census to estimate wage distributions for each State. Benefits correspond to the combined AFDC and Food Stamp benefits for a family of 3 in 1989. Appendix Table 1 reports parameter values used in estimation.

Table 1 shows that the differences in benefits across states are large: for example the highest annual benefit (excluding Alaska and Hawaii) is $7,568 in California, and the lowest is $3,426, in Alabama (in 1983 dollars).[12] In the third column of the table, these differences are adjusted for differences in living costs across states, using the ACCRA cost of living index (**http://www.coli.org/).** Even after this adjustment, the differences remain large. The last column of the table shows the wage percentile in the 1990 PUMS data corresponding to the benefit level, by State. The typical situation is that less than 50% of single women with children earn more than the benefit level.

[12]Using a slightly different definition of benefits, and after adjusting for cost of living differences, Gelbach reports a benefit level of $9,912 in Connecticut, and $4,654 in Mississippi (in 1997 dollars).

4.2 Partial Likelihood Estimates

We condition on the estimated earnings distributions for each state and maximize the partial likelihood to obtain estimates of the utility function parameters. We fix β at 0.9.

The results in Table 2 show that differences in expected income are a significant determinant of migration decisions for this population, but this effect cannot be accurately measured without controlling for other influences on migration. There are 10,101 person-years in the data, and there are 380 interstate moves. This is an annual migration rate of 3.76%, and the first column in Table 2 matches this rate by setting the probability of moving to each of J-1 locations to a constant value, namely $\frac{1}{J-1}\frac{389}{10,171}$, with J = 51.[13] The second column shows the effect of income with no other variables included, but the next three columns show that population size, distance, and home location all have highly significant effects on migration.[14] The last two columns show the effect of income, controlling for these other effects; the last column uses wage and benefit numbers adjusted for cost of living differences across States, while the previous column uses unadjusted data.

One way to interpret the estimated coefficients is to consider what the migration rate would be in the absence of some of the effects measured in Table 2. Suppose for example that there were no differences in population or income across States, and that there were no distance between States. Then the migration rate would be much higher: the coefficient estimates in the last column of Table 2 imply an annual migration probability of 16.8% for someone currently in the home location, and a probability of 23.2% for someone who is currently in a different location.

The estimated effect of expected income is quantitatively important. For example, a permanent increase in income of \$4,000 per year in the current location reduces the annual migration probability from 3.82% to 2.00%, while an income reduction of \$4,000 increases the migration probability to 7.19%. Thus large differences in expected income lead to large migration flows (since these probability differences are cumulated over time).

How Big are the Moving Costs?

---

[13]In other words the estimate of δ_0 solves the equation $\frac{e^{-\delta_0}}{1+(J-1)e^{-\delta_0}} = \frac{1}{J-1}\frac{389}{10,171}$; the solution is δ_0 = log(489100) - log(389).

[14]The χ^2 statistics in the table are for likelihood ratio tests of the form $2\log(L^U/L^R) \sim \chi^2(r)$, where r is the number of restrictions embodied in L^R relative to L^U.

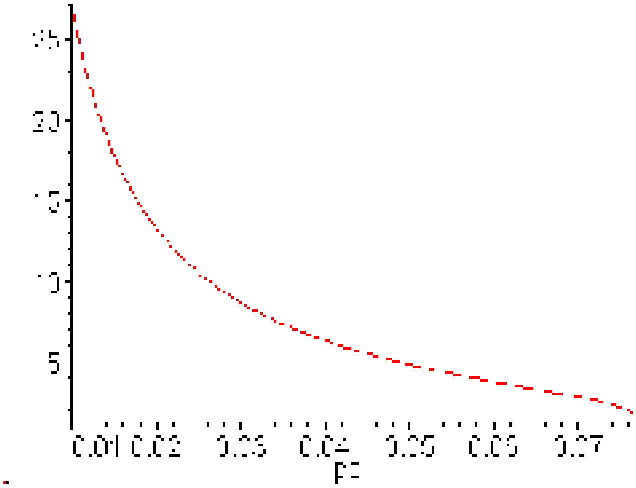
Since the utility function is linear in income, we can translate the estimated moving cost into a dollar equivalent. The result of this calculation is δ_0/α = \$306,483, with the interpretation that the (lump-sum) compensation needed to just offset the cost of a move is enormous: other things equal, an offer of \$300,000 would not be enough to persuade someone to move.

It may seem that this large moving cost is an artifact of the specification of the model. For example, in the absence of any moving cost, allowing preference shocks to be drawn randomly over J locations implies a migration probability of (J-1)/J, so that with J = 51, nearly everybody moves every period. The first column of Table 2 shows how large the fixed cost of moving has to be in relation to the preference shocks, in order to reduce the migration rate from 50/51 to the observed rate of 3.82%, when all other influences on migration are suppressed. The second column shows that the migration cost must be increased slightly when the influence of income is taken into account – income differences promote some moves and retard others, and on balance the effect is evidently in favor of an increase in migration. More importantly, the estimated moving cost is very large in relation to the income coefficient. When the effects of population and distance and the home premium are taken into account, the estimated moving cost is smaller in relation to income, but it is still very large.

To understand why the estimated moving cost is so big, it is helpful to consider an example in which income differentials and moving costs are the only influences on migration decisions. Suppose there are just two income levels, and let y be the present value of the difference between the two incomes. Then it is straightforward to show that the probability of staying in the current location is

$$\begin{aligned} \lambda_L &= \frac{e^{\delta_0}}{e^{\delta_0} + J_L - 1 + J_H e^{\alpha y}} \\ \lambda_H &= \frac{e^{\delta_0}}{e^{\delta_0} + J_H - 1 + J_L e^{-\alpha y}} \end{aligned} \qquad \textbf{(11)}$$

where λ_L is the probability of staying in a low-income location and J_L is the number of such locations, and similarly for λ_H and J_H. Given data on λ_L and λ_H, these equations can be solved for δ_0 and αy, and the ratio $C = \delta_0/\alpha y$ represents the moving cost as a multiple of the income differential. Note that if $C \le 1$ then $\lambda_L \le \frac{1}{1 + J_H}$, implying a migration rate that is much higher than the rate seen in the data. On the other hand if $C > 1$, then the migration rate can be made arbitrarily small, by increasing both δ_0 and α (in fixed proportion). In other words, as long as the moving cost exceeds the income differential (even by a tiny



amount) the model can fit a low migration rate by scaling up both the moving cost and the marginal utility of income (which is the same thing as scaling down the preference shocks). Thus if income differentials and the moving cost are both large relative to the preference shocks, the overall migration rate may be low even though the moving cost is not much bigger than the income differential. But an implication of this scenario is that the migration rate from low-income locations should be much larger than the migration rate from high-income locations, and this is evidently at odds with the data.

The relationship between observed migration rates and estimated moving costs in this simplified model is illustrated in Figure 1. It is assumed that the average migration rate matches the data, meaning that $\frac{\lambda_L + \lambda_H}{2} = 1 - \frac{389}{10,171}$. The equations for λ_L and λ_H are solved for δ_0 and αy, and the implied value of C is plotted against $\psi / \lambda_H - \lambda_L$. The result is that if λ_L and λ_H, are close, then the moving cost must be a large multiple of the income differential. So this example suggests that the large estimated moving cost implied by the estimates in Table 2 arises because the empirical relationship between migration rates and income levels is relatively weak.

There are of course important influences on migration decisions that are not included in our model, and a reasonable interpretation of the results is that, on average, the omitted variables strongly favor staying in the current location. If this is so, a more complete model would yield a more plausible estimate of the moving cost. Nevertheless, our estimate of the effect of income differentials is valid provided that the variation in the omitted variables across States is not correlated with the income differentials.

Conclusion

We have developed a tractable econometric model of optimal migration in response to income differentials across locations.

Empirical results show a significant effect of income differentials on migration, for unskilled single women with dependent children who are eligible for AFDC.

References

Barro, Robert J. and Xavier Sala-i-Martin, "Convergence across States and Regions," *Brookings Papers on Economic Activity*, 1991, **1**, 107-158.

Blanchard, Olivier Jean and Lawrence F. Katz, "Regional Evolutions," Brookings Papers on Economic Activity, 1992, **1**, 1-37.

Greenwood, Michael J. "Internal Migration in Developed Countries," in *Handbook in Population and Family Economics Vol. 1B*, edited by Mark R. Rosenzweig and Oded Stark. New York: North Holland. 1997.

Holt, Frederick (1996) "Family Migration Decisions: A Dynamic Analysis," unpublished paper, University of Virginia.

Keane, Michael P. and Kenneth I. Wolpin, "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), June 1997, 473-522.

Levine, Phillip B., and David J. Zimmerman "An Empirical Analysis of the Welfare Magnet Debate Using the NLSY," *Journal of Population Economics* 2000.

Lucas, Robert E. B. "Internal Migration in Developing Countries," in *Handbook in Population and Family Economics Vol. 1B*, edited by Mark R. Rosenzweig and Oded Stark. New York: North Holland. 1997.

Meyer, Bruce D. "Do the Poor Move to Receive Higher Welfare Benefits?" Northwestern University, October 1999.

Moffitt, Robert (1992), "Incentive Effects of the United States Welfare System: A Review," *Journal of Economic Literature* 30 (March): 1-61.

Neal, Derek, "The Complexity of Job Mobility of Young Men," *Journal of Labor Economics*, April, 1999, 237-261.

Rust, John (1987) "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55 (5), September 1987, 999-1033.

Rust, John (1994), "Structural Estimation of Markov Decision Processes,"in *Handbook of Econometrics*, Volume IV. Edited by Robert F. Engle and Daniel L. McFadden. New York: Elsevier.

Rust, John (1996), "Numerical Dynamic Programming," in *Handbook of Computational Economics,* North Holland, H. Amman, D. Kenrick, and J. Rust eds.

Schultz, T. Paul (1982) "Lifetime Migration within Educational Strata in Venezula: Estimates of a Logistic Model," *Economic Development and Cultural Change* 30: 559-593.

Topel, Robert H., "Local Labor Markets," *Journal of Political Economy*, 1986, 94(3), part 2, S111-S143.

Walker, James R. (1994) "Migration Among Low-Income Households: Helping the Witch Doctors Reach Consensus," Discussion Paper, #94-1032, Institute For Research on Poverty.

Table 0: Interstate Migration Flows, NLSY79								
	Less than High School		High School		Some College		College	
No. of people	1768		3534		1517		1435	
Horizon (years)	5	13	5	13	5	13	5	13
No. of movers	334	423	598	771	327	376	441	469
Repeat moves	239	434	313	653	167	264	196	261
Repeat moves as % of all moves	41.7	50.6	34.4	45.9	33.8	41.3	44.4	35.7
Movers (%)	18.9	23.9	16.9	21.8	21.6	24.8	30.8	32.7
Moves Per Mover	1.7	2.0	1.5	1.8	1.5	1.7	1.4	1.6
Return Migration (% of all moves)								
Return - Home	22.9	24.0	20.6	24.1	16.0	17.5	12.4	13.4
Return - Else	5.4	12.4	2.7	7.2	2.8	5.9	2.5	3.3
Movers who return home (%)	39.2	48.7	31.4	44.5	24.2	29.8	17.9	20.9
Return-Home: % of Repeat	54.8	47.5	60.1	52.5	47.3	42.4	40.3	37.5

Table 1: Wages and Benefits, by State

Single Women with Children, 1989

	Observations, NLSY (Person-years)	Benefits	Adjusted Benefits	Wage Percentile
Alabama	385	3,426	3,604	55.6%
Alaska	67	9,765	7,232	73.9%
Arizona	81	5,061	4,894	56.7%
Arkansas	207	4,258	4,517	63.4%
California	1,066	**7,568**	**6,877**	70.6%
Colorado	129	5,667	5,667	57.5%
Connecticut 910	259	7,297	5,948	67.0%
Delaware	10	5,332	4,972	59.6%
DC	102	5,739	4,625	55.4%
Florida	394	5,023	4,954	53.2%
Georgia	648	4,897	5,036	59.7%
Hawaii	5	8,381	6,325	66.7%
Idaho	3	5,139	5,291	64.3%
Illinois	311	5,448	5,316	61.0%
Indiana	116	5,032	5,228	50.7%
Iowa	31	5,748	5,879	62.6%
Kansas	75	6,126	6,353	68.4%
Kentucky	56	4,394	4,665	63.1%
Louisiana	134	4,123	4,136	72.2%
Maine	129	6,048	6,307	68.3%
Maryland	193	5,806	5,729	52.1%
Massachusetts	425	6,735	5,874	64.7%
Michigan	394	6,774	6,485	69.6%
Minnesota	107	6,687	6,643	61.5%
Mississippi	165	**3,445**	**3,601**	62.6%
Missouri	336	5,013	5,365	59.2%
Montana	72	5,516	5,524	66.0%
Nebraska	27	5,545	5,921	46.5%
Nevada	5	5,313	4,910	54.8%
New Hampshire	0	6,445	5,313	46.3%
New Jersey	284	6,029	5,178	53.3%
New Mexico	81	4,839	4,807	60.1%
New York	546	6,890	6,342	67.7%
North Carolina	405	4,858	4,918	50.9%
North Dakota	0	5,700	5,840	80.0%
Ohio	621	5,294	5,343	63.6%
Oklahoma	83	5,284	5,331	58.9%
Oregon	37	6,271	6,092	66.9%
Pennsylvania	323	5,806	5,705	60.1%
Rhode Island	12	6,629	5,988	73.4%
South Carolina	414	4,277	4,405	48.6%
South Dakota	15	5,565	5,803	62.5%
Tennessee	174	3,958	4,196	52.5%
Texas	728	4,065	4,113	57.8%
Utah	26	5,632	5,807	57.5%
Vermont	39	7,345	6,768	71.4%
Virginia	229	5,477	5,553	56.6%
Washington 532	123	6,552	6,554	63.8%
West Virginia	74	4,694	4,788	67.9%
Wisconsin	348	6,581	6,860	66.4%
Wyoming	1	5,516	5,491	58.6%
All States	10,101	5,670	5,479	

Table 2: Interstate Migration of Unskilled Single Women with Dependent Children

Disutility of Moving	7.1650	7.2086	7.0036	6.2652	5.0441	5.0938	5.1090
	0.0528	*0.0569*	*0.0529*	*0.0823*	*0.0813*	*0.0847*	*0.0805*
Population			0.1067	0.1110	0.1316	0.1234	0.1243
			0.0088	*0.0086*	*0.0129*	*0.0129*	*0.0128*
Distance (1000 miles)				0.8044	0.7133	0.7108	0.7090
				0.0824	*0.0802*	*0.0802*	*0.0805*
Home Premium				----------	0.3822	0.3874	0.3873
					0.0143	*0.0144*	*0.0144*
Income(/$10K)		0.0725		----------	----------	0.1315	
		0.0273				*0.0326*	
"Real" Income (ACCRA)				----------	----------	----------	0.1633
							0.0331
Loglikelihood	-3044.494	-3040.368	-2975.605	-2929.924	-2684.209	-2676.910	-2674.822
χ^2 (1)		8.252	129.778	91.362	491.430	14.598	18.774
p-value		0.0041				0.0001	0.00001

Notes:
Estimated asymptotic standard errors are given in italics below the coefficient estimates.
The length of the horizon is 40 years, with discount factor $\beta = .9$
The wage distributions have 3 points of support

Appendix Table 1: Population Size and
Empirical Approximation of Income Distribution By State(in 1983 Dollars)

State	Population	Proportion Censored	Benefits	Income Value Cell 2	Income Value Cell 3	Cut point between Cells 2 and 3
Alabama	3893888	.5560	3426	6048	11290	8452
Alaska	401851	.7391	9765	12903	17339	13548
Arizona	2718215	.5669	5061	8065	15323	10484
Arkansas	2286435	.6340	4258	7132	12215	9677
California	23667902	.7059	7568	10081	17742	13558
Colorado	2889964	.5746	5497	8065	14516	10452
Connecticut	3107576	.6695	7294	11694	19355	10081
DC	638333	.5536	5739	10484	16935	13710
Florida	9746324	.5324	5023	8065	14113	10081
Georgia	5463105	.5969	4897	7258	13710	10484
Hawaii	964691	.6667	8381	10484	15323	12860
Ido	943935	.6429	5139	6956	11290	8065
Illinois	11426518	.6103	5448	8468	15323	11290
Indiana	5490224	.5068	5032	8065	13710	10214
Iowa	2913808	.6264	5748	8008	14516	10081
Kansas	2363679	.6842	6126	7963	13145	9677
Kentucky	3660777	.6310	4394	7032	14253	10323
Louisiana	4205900	.7224	4123	6202	11129	8165
Maine	1124660	.6829	6048	8402	13710	12097
Maryland	4216975	.5208	5806	9677	17742	12097
Massachusetts	5737037	.6474	6735	10161	17742	13306
Michigan	9262078	.6955	6774	9677	18548	12093
Minnesota	4075970	.6150	6687	8871	16129	12097
Mississippi	2520638	.6261	3445	4919	9677	6452
Missouri	4916686	.5919	5013	7279	12903	9677
Montana	786690	.6596	5516	8380	13520	9677
Nebraska	1569825	.4648	5545	7661	12903	12016
Nevada	800493	.5476	5313	8065	16129	11290
New Hampshire	920610	.4630	6445	12903	21774	16129

Population Size and
Empirical Approximation of Income Distributions By State (continued)

State	Population	Proportion Censored	Value 1 (Benefits)	Value 2	Value 3	Cut point between Cells 2 and 3
New Jersey	7364823	.5333	6029	9677	17432	13710
New Mexico	1302894	.6011	4839	7258	12097	8871
New York	17558072	.6771	6890	9677	17742	13242
North Carolina	5881766	.5087	4858	7558	12903	9677
North Dakota	652717	.8000	5700	6876	8871	8085
Ohio	10797630	.6355	5294	8871	16129	12594
Oklahoma	3025290	.5887	5284	6894	12097	9677
Oregon	2619399	.6685	6271	8076	15323	10484
Pennsylvania	11863895	.6009	5806	8856	14516	11290
Rhode Island	947154	.7342	6629	8548	16129	12097
South Carolina	3121820	.4860	4277	7258	12431	9677
South Dakota	690768	.6250	5565	6713	11290	8131
Tennessee	4580367	.5251	3958	7089	12903	9677
Texas	14229191	.5777	4065	6452	12903	9258
Utah	1461037	.5747	5632	8065	14516	10484
Vermont	511456	.7143	7345	10349	17889	14919
Virginia	5346818	.5660	5477	8065	14516	10484
Washington	4132156	.6376	6552	8871	17889	12903
West Virginia	1921005	.6790	4964	6984	13731	9806
Wisconsin	4705767	.6644	6581	9677	16197	11631
Wyoming	469557	.5862	5516	7447	15121	10094