Modeling Individual Migration Decisions

John Kennan and James R. Walker*
University of Wisconsin-Madison and NBER

March 2012

Abstract

We summarize recent research that formulates life cycle models of migration which are estimated using longitudinal data. These models consider multiple destinations and multiple periods. The framework offers a unified view applicable to internal and international migration flows, however data limitations severely hinder studies of international migration. As is common in modeling life cycle decision–making, strong assumptions are imposed. Yet, most critical assumptions are empirically testable. The primary advantage is that these models offer an interpretable economic framework for evaluating policy alternatives and other counterfactual thought–experiments that offer insight on behavioral determinants and tools for improved policy making.

1 Introduction

We review empirical analyses of migration decisions, using life-cycle models to interpret migration histories. The starting points are Schultz (1961), who considered migration as a form of investment in human capital, and DaVanzo (1983), who documented the richness of individual migration histories, pointing out that although most individuals never move, those who do are likely to move again, often returning to a home location.[1] This means that migration decisions should be viewed as a sequence of location choices, where the individual knows that there will be opportunities to modify or reverse moves that do not work out well.

Sjaastad's (1962) treatment of migration as an investment emphasizes the dynamic aspect of migration – expected costs and payoffs to migration change over time. Viewed within a life cycle perspective, individuals (or families) decide whether and when to move. Allowing households to make multiple migration decisions substantially increases

*Department of Economics, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706; jkennan@ssc.wisc.edu, walker@ssc.wisc.edu. We thank an anonymous referee and the editors, Amelie F. Constant and Klaus F. Zimmermann for helpful comments.

the model's complexity. Decisions made in previous periods (e.g., savings, education, marriage, fertility) determine choices available in the current period, and expectations of future events also influence current decisions. Within this perspective migration and fertility choices are connected through the "primitives" of the decision making process: current opportunities (determined in part by decisions made in the past), expectations (anticipated future events and outcomes), and preferences (values assigned to different outcomes). The economic perspective thus provides a unified framework connecting several important demographic behaviors.

Any of a large number of factors may influence an individual to move from their current location. The special nature of the "home" location suggests that home will be a favored destination of natives who have moved away. A common explanation for a sequence of moves builds on the idea of differences between expected and realized outcomes and the role of learning. An initial move is motivated by anticipated outcomes. The move is made and the realized outcome is experienced. If the realized outcome is greatly different than expected, a second, corrective move may be required (which may be to home, given its the special nature) or learning may take place and the updated assessment of the costs and payoffs (at all locations) may induce onward migration.[2] This perspective also provides an explanation for the commonly observed pattern that places of high in-migration also experience high out-migration.

Prospective migrants may be uncertain about payoffs and costs, and indeed, prospective migrants may differ in their attitudes toward risk, liquidity to finance a move and insurance against poor labor market outcomes following a move. To the extent that moving entails a large monetary cost, financing a move becomes important. Here families can play a central role in providing information and in helping to finance a move. Families can also provide an important form of insurance. Social insurance programs provide a safety net for some families, while family members provide an important backstop. These types of intra-family exchanges have been extensively modeled and investigated in developing countries. In particular the role of information, credit constraints and intergenerational transfers have received the most attention.

As in all forward-looking models of behavior, specification of how expectations are formed plays a critical role in the model specification. Most commonly individuals are assumed to have rational expectations (and thus use all available information). When agents are not well informed a model of learning is necessary to determine subjective probabilities of events and payoffs. The optimal policy may involve experimentation early in the process to gain information relevant for subsequent payoffs. Such models can help rationalize chain and return migration.

Selection is an important element in every empirical study of migration.[3] As is well-understood, the fundamental problem is that many patterns may stem from true causal forces or may arise from unobserved individual characteristics. For example, if we observe large positive economic gains accruing to migrants, is this evidence of the gain anticipated by the migrant or does the gain reflect characteristics of the migrant that might have also led to above average payoffs even if the individual had not moved?

A theoretical framework is needed to control for selection, which is much more difficult than in many other applications, where panel data can generate observations on the same person in different states (albeit at different points in time). Here, most individuals are seen only in a single location, and even the movers move only a few times; no one is observed in all locations. Descriptive studies are useful for highlighting important flows and temporal patterns, but the ubiquity of selection means that an analytic framework is needed to unravel behavioral mechanisms.

In this paper we describe recent research modeling of individual migration decisions. These models consider multiple destinations and multiple periods: individuals may move more than once, may move to several distinct locations or may return to a previously visited location. In addition, the models recognize the special status of the home location. We proceed by first outlining the main elements of the models, and then present empirical applications.

2 Components of a model of migration

The basic idea of a (dynamic) migration model is that there are a number of alternative (mutually exclusive) payoff flows, associated with different locations, with options to switch between these alternatives, subject to a moving cost. At any given time, the individual is in a particular location, and must choose whether to stay there or go somewhere else, exchanging the current payoff flow for one of the alternatives. The future payoffs in the current location are generally uncertain, and the alternative payoffs even more so.

The individual acts so as to maximize the expected present value of the realized payoffs, net of moving costs.

The model is most conveniently formulated in the language of dynamic programming, since it involves solving variants of the same choice problem over and over again. Suppose there are J locations, and the payoff flow in each location is a random variable with a known distribution. Let x be the state vector (which includes current location and age, as well as all currently available information that helps to predict future payoffs, as discussed below). The utility flow for someone who chooses location j is specified as $u(x,j)+\zeta_j$, where ζ_j is a random variable that is assumed to be independent and identically distributed across locations and across periods, and independent of the state vector, representing influences on migration decisions that are not included in the model. Let $p(x'|x,j)$ be the transition probability from state x to state x', if location j is chosen. The decision problem can be written in recursive form as

$$V\left(x,\zeta\right)=\max_{j}\left(v(x,j)+\zeta_j\right)$$

where

$$v(x,j)=u(x,j)+\beta\sum_{x'}p(x'|x,j)\bar{v}(x')$$

and

$$\bar{v}(x)=E_{\zeta}V\left(x,\zeta\right)$$

and where β is the discount factor, and E_{ζ} denotes the expectation with respect to the distribution of the J-vector ζ with components ζ_j. The interpretation is as follows. The payoff shocks (for all locations) are realized before the location decision is made, and the distinction between $v\left(x,j\right)$ and $\bar{v}\left(x\right)$ is that one represents the continuation value for each alternative choice, while the other represents the expectation of the optimized continuation value, taken before the payoff shocks have been realized.

The above formulation can be used to describe both finite-horizon and infinite-horizon problems. In the finite-horizon case, it is convenient to remove age from the state vector, and write the model as

$$\begin{aligned}
V_s(x,\zeta) &= \max_j \left(v_s(x,j)+\zeta_j\right) \\
v_s(x,j) &= u_s(x,j)+\beta\sum_{x'} p_s(x'|x,j)\bar{v}_{s-1}(x') \\
\bar{v}_s(x) &= E_\zeta V_s(x,\zeta)
\end{aligned}$$

where s is the number of periods remaining, with $\bar{v}_0(x) = V_0(x,\zeta) = 0$. Thus the decision for $s = 1$ is a simple static choice problem, and once this has been solved, the decision problem for $s = 2$ can be specified explicitly, and the solution of this can be used to specify the decision problem for $s = 3$, and so on. In other words the general solution can be obtained by backward induction.

The model is implemented by specifying the function $u(x,j)$ as well as the transition probabilities, up to a vector of unknown parameters θ. Ideally, the specification should parsimoniously capture the most important features of the choice problem, giving plausible interpretations of the main features of the data, while also facilitating the prediction of behavior in situations not actually seen in the data (such as alternative policy environments).

A major limitation of models of this kind is that the solution cannot be computed unless special assumptions are made. For one thing, computation of the function $\bar{v}$ involves an integral over all possible realizations of the payoff shocks, and if there are many locations this is infeasible unless the payoff shocks are drawn from a distribution for which a closed-form solution for this integral is known (which in practice means the generalized extreme value distribution).

We assume that ζ_j is drawn from the Type I extreme value distribution. In this case, following McFadden (1974) and Rust (1987), we have

$$\exp\left(\bar{v}_s(x)\right) = \exp\left(\bar{\gamma}\right)\sum_{k=1}^{J}\exp\left(v_s(x,k)\right)$$

where $\bar{\gamma}$ is the Euler constant. Let $\rho_s(x,j)$ be the probability of choosing location j, when the state is x, with s periods remaining. Then

$$\begin{aligned}\rho_s(x,j) &= \exp\left(\bar{\gamma}+v_s(x,j)-\bar{v}_s(x)\right)\\ &= \frac{\exp\left(v_s(x,j)\right)}{\sum_{k=1}^{J}\exp\left(v_s(x,k)\right)}.\end{aligned}$$

A more general computational issue is that solution of the dynamic programming problem requires computation of the continuation values from all decision nodes that might possibly be reached, and the number of such nodes explodes as the number of possible states increases[4]. This severely limits the number of locations that can be considered. For instance, the natural specification of a location is a local labor market, but it is impossible to compute such a model for a large economy; in applications to the U.S. economy, locations must be aggregated to the level of States, or even Census regions. Even then, there is an outrageous number of decision nodes, but since the vast majority of these are almost never reached, the solution can be well approximated by just ignoring most of them. Thus we assume that the individual retains information about wage draws in at most two locations (even if more locations have actually been visited).

Another important general consideration is that the initial conditions of the decision problem are typically the result of some previous decisions, which means that even if the stochastic components of payoffs are randomly assigned *ex ante*, the distribution of these components in the data is contaminated by selection bias. This is of course especially true for models of migration by older people. But it may be reasonable to assume away the initial conditions problem in models that begin at the point of entry to the labor force.

3 An Empirical Model

In Kennan and Walker (2011) we develop a dynamic model of migration decisions, and estimate it using data on young white male high school graduates in the U.S. We follow respondents of the NLSY79 from age 20 until their mid-30s. We define locations as States. We include a utility premium for workers residing in their "home" location, defined as the State of residence at age 14. The details of the model are outlined below, followed by a summary of the empirical results. We also present results for white male college graduates.

3.1 Payoff Flows

Let $\ell = \left(\ell^0, \ell^1\right)$ be a vector recording the current and previous locations (with the convention that $\ell^1 = 0$ if there is no previous location), and let $\omega = \left(\omega^0, \omega^1\right)$ be a vector recording wage information at these locations. The state vector x consists of ℓ, ω and age. The flow payoff for someone whose "home" location is h is specified as

$$\tilde{u}_h\left(x, j\right) = u_h\left(x, j\right) + \zeta_j$$

where

$$u_h\left(x, j\right) = \alpha_0 w\left(\ell^0, \omega^0\right) + \sum_{k=1}^{K} \alpha_k Y_k\left(\ell^0\right) + \alpha^H \chi\left(\ell^0 = h\right) - \Delta_\tau\left(x, j\right)$$

Here the first term refers to wage income in the current location. This is augmented by the nonpecuniary variables $Y_k\left(\ell^0\right)$, representing amenity values. The parameter α^H represents a premium that allows each individual to have a preference for their native location ($\chi\left(A\right)$ denotes an indicator meaning that A is true). The cost of moving from ℓ^0 to j for a person of type τ is represented by $\Delta_\tau\left(x, j\right)$. The unexplained part of the utility flow, ζ_j, may be viewed as either a preference shock or a shock to the cost of moving, with no way to distinguish between the two.

3.2 Wages

Workers know their wage in the current location, are assumed to have rational expectations, and to know the distribution of offered wages at all other locations. The wage of individual i in location j at age a in year t is specified as

$$w_{ij}(a) = \mu_j + \upsilon_{ij} + G(X_i, a, t) + \eta_i + \varepsilon_{ij}(a)$$

where μ_j is the mean wage in location j, υ is a permanent location match effect, $G(X, a, t)$ represents a (linear) time effect and the effects of observed individual characteristics, η is an individual effect that is fixed across locations, and ε is a transient effect. We assume that η, υ and ε are independent random variables that are identically distributed across individuals and locations. We also assume that the realizations of η and υ are seen by the individual. The age component and the fixed effect are common to all locations and consequently do not influence migration decisions. Thus since the current realization of the transient wage component is known only after the current location has been chosen, migration decisions are driven exclusively by the State means (μ_j) and the match specific component (υ_{ij}) between worker i and location j.

For computational reasons, we model the worker-location component as a discrete distribution with three points of support (low, middle, and high). Even this simple model gives workers two motivations to migrate: to leave a bad local labor market (a low μ_j) or a bad location match, (a low υ_{ij}). The incentives to migrate are strong. For example, the 90-10 differential across State means is about \$4,700 a year (in 2010 dollars) and the value of replacing a bad location match with a good one is about \$17,000 a year.

3.3 Moving Costs

Let $D\left(\ell^0, j\right)$ be the distance from the current location to location j, and let $\mathbb{A}(\ell^0)$ be the set of locations adjacent to ℓ^0 (where States are adjacent if they share a border). The moving cost is specified as

$$\Delta_\tau\left(x,j\right)=\left(\gamma_{0\tau}+\gamma_1 D\left(\ell^0,j\right)-\gamma_2\chi\left(j\in\mathbb{A}\left(\ell^0\right)\right)-\gamma_3\chi\left(j=\ell^1\right)+\gamma_4 a-\gamma_5 n_j\right)\chi\left(j\neq\ell^0\right)$$

We allow for unobserved heterogeneity in the cost of moving: there are several types, indexed by τ, with differing values of the intercept γ_0. In particular, there may be a "stayer" type, meaning that there may be people who regard the cost of moving as prohibitive. The moving cost is an affine function of distance, but moves to a previous location may be less costly, and moves to an adjacent location may also be less costly (because it is possible to change States while remaining in the same general area). In addition, the cost of moving is allowed to depend on age, a. Finally, we allow for the possibility that it is cheaper to move to a large location, as measured by population size n_j. It has long been recognized that location size matters in migration models (see e.g. Schultz, 1982). For example, a person who moves to be close to a relative is more likely to have relatives in California than in Wyoming. One way to model this in our framework is to allow for more than one draw from the distribution of payoff shocks in each location. Alternatively, location size may affect moving costs – for example, relatives might help reduce the cost of the move. In practice, both versions give similar results.

3.4 Transition Probabilities

The state transition probabilities $p_s(x' \mid x, j)$ for this model are straightforward. First, if no migration occurs this period, then the state remains the same except for the age component. If there is a move to a previous location, the current and previous locations are interchanged. And if there is a move to a new location, the current location becomes the previous location, and the new location match component of wages is drawn at random. In all cases, age is incremented by one period (equivalently, s is decremented by one period).

3.5 Results

We show the basic estimation results from Kennan and Walker (2011), along with estimates of the same model using data for college graduates. The estimates in Table 1 show that expected income is an important determinant of migration decisions, for both education groups. Even though the overall migration rate is much higher for college graduates, the parameter estimates are quite similar for the two samples, aside from a substantially lower estimated migration cost for college graduates.

The importance of the home location is clearly shown in Table 1, especially for the high school sample. This attachment to home reduces out migration and induces return migration. It helps to explain why most people never move, despite large spatial wage differences; it also implies that the losses from forced migration (such as the migration due to hurricane Katrina) are very large.[5]

3.6 How Big are the Moving Costs?

There are big differences in wages across States, and the estimated dispersion of the worker-location component of wages is also quite large. Yet migration rates are low: the interstate migration rate for white men in the NLSY is 2.9% for high school graduates, and while the rate for college graduates is much higher (8.6%), this still seems low in relation to the estimated wage gains. A natural reaction is to infer that moving costs must be very high. And indeed if it is assumed that the moving cost is the same for everyone, the estimated model indicates that the cost is on the order of \$300,000 (for high-school graduates). Yet people do move, and those who move tend to move again, and it is hard to believe that they are paying costs of this magnitude every time they move.

The answer to this riddle is that people are heterogeneous. For some people, at some times, the moving cost is very high; for some people, at some times, the cost is quite low. The model allows us to quantify the extent of this heterogeneity. In particular, we can estimate the average moving costs for those who actually move. The estimates for the high school sample are given in Table 2. There is considerable variation in these costs, but for a typical move the cost is negative. The interpretation of this is that the typical move is not motivated by the prospect of a higher future utility flow in the destination location, but rather by unobserved factors yielding a higher current payoff in the destination location, compared with the current location. That is, the most important part of the estimated moving cost is the difference in the payoff shocks. In the case of moves to the home location, on the other hand, the estimated cost is positive; most of these moves are return moves, but where the home location is not the previous location the cost is large, reflecting a large gain in expected future payoffs due to the move.

3.7 Why do College Graduates Move so Much?

It is well known that the migration rate for skilled workers is much higher than the rate for unskilled workers; in particular the migration rate for college graduates is much higher than the rate for high school graduates (see, for example, Topel (1986), Greenwood (1997), Bound and Holzer (2000), Wozniak (2010), Molloy et al. (2011)). Malamud and Wozniak (2009), using draft risk as an instrument for education, find that an increase in education causes an increase in migration rates (the alternative being that people who go to college have lower moving costs, so that they would have higher migration rates even if they did not go to college).[6] The model described in Table 1 can be used to simulate the extent to which the differences in migration rates for college graduates can be explained by differences in expected incomes, as opposed to differences in moving costs. This distinction affects the interpretation of measured rates of return on investments in college education. For example, if college graduates move more because the college labor market has higher geographical wage differentials, then a substantial part of the measured return to college is spurious, because it is achieved only by paying large moving costs.

Table 3 shows the observed annual migration rates for the high school and college graduate samples along with the migration rates predicted by the estimated model, where these rates are computed by using the model to simulate the migration decisions of 100 replicas of each person in the data. The extent to which the large observed difference in migration rates can be attributed to differences in geographical wage dispersion can be measured by simulating the migration decisions that would be made by one group if they faced the same wage dispersion as the other group. Thus, according to the model, the migration rate of high school graduates would increase considerably if they faced the higher wage dispersion seen by college graduates, but the migration rate in this simulation is still only about 4% per year, compared with 8.6% for the college sample. The reverse experiment gives a similar result: the migration rate for college graduates facing the high school wage process would still be more than twice the observed rate for high school graduates. Thus although geographical wage dispersion can explain a nontrivial part of the difference in the explained migration rates, the model attributes the bulk of this difference to other factors, such as differences in moving costs.

3.8 Spatial Labor Supply Elasticities

The estimated model can be used to analyze labor supply responses to geographical wage differentials. We are interested in both the magnitude and the timing of these responses. For example, Blanchard and Katz (1992) found that the half life to a unit shock to the relative wage is more than a decade. Studies by Barro and Sala-i Martin (1991) and Topel (1986) report similar findings. Given that college graduates move more often than high school graduates, it is also interesting to ask whether the greater mobility of college graduates is associated with a more elastic response to geographic wage differences.

Since the model assumes that the wage components relevant to migration decisions are permanent, it cannot be used to predict responses to wage innovations in an environment in which wages are generated by a stochastic process. Instead, it is used to answer comparative dynamics questions: the estimated parameters are used to predict responses in a different environment.

The first step is to take a set of young white males who are distributed over States as in the 1990 Census data, and allow the population distribution to evolve, by iterating the estimated transition probability matrix (given the observed wages). The transition matrix is then recomputed to reflect wage increases and decreases representing a 10% change in the mean wage of an average 30-year-old, for selected States, and the population changes in this scenario are compared with the baseline simulation. Supply elasticities are measured relative to the supply of labor in the baseline calculation. For example, the elasticity of the response to a wage increase in California after 5 years is computed as $\frac{\Delta L}{\Delta w}\frac{w}{L}$, where L is the number of people in California after 5 years in the baseline calculation, and ΔL is the difference between this and the number of people in California after 5 years in the counterfactual calculation.

Figure 1 shows the results for three large States that are near the middle of the one-period utility flow distribution. The high school results show substantial responses to spatial wage differences, occurring gradually over a period of about 10 years. The wage responses for college graduates are larger (with a supply elasticity around unity), and the length of the adjustment period is longer. This is consistent with the hypothesis that college graduates face substantially lower moving costs than high school graduates. The reason for the long adjustment period is that wage differences are just one of many influences on migration decisions. Tilting the wage difference in favor of a particular location therefore has a relatively small effect on migration probabilities, but since this effect is permanent, while the payoff shocks are transient, the cumulative effect of wage differences is substantial in the longer run.

4 Welfare Migration

At various times during the last thirty years, the existence of "welfare magnets" has surfaced in public policy debates and particularly during welfare reform in the mid-1990s in the United States.[7] Prior to the reform, Aid to Families with Dependent Children (AFDC) was the primary source of income support. States followed federal guidelines but were otherwise free to determine benefits schedules. A state offering relatively high income support could function as a "magnet" for low-income population; retaining those already living in the State and attracting poor from relatively low–benefits States. Indeed, during the welfare reform debates there was concern that competition among States would become a "race to the bottom" in setting benefit levels. In Kennan and Walker (2010) we develop a dynamic model to investigate the migration decision making of welfare-eligible mothers.[8]

Our motivation for estimating the model was to investigate migration flows that might be induced by alternative welfare benefit policy regimes not seen in the data. Specifically, we investigated the implied migration flows if all states set benefits equal to Mississippi's (the lowest benefit) or California's (the highest benefit). Our estimates suggest that income has a significant but quantitatively weak influence on migration.

We find that migration adjusts slowly, but with slightly greater responsiveness to earnings than to benefits. This is to be expected as everyone is affected by wages, but high-wage women are not much affected by welfare benefits. In particular, women with favorable individual fixed effects are unlikely to be on benefits.

In the second set of counterfactual experiments we investigate migration responses to uniform benefit levels for all states. Differences in AFDC benefits are seen as the driving force behind welfare-induced migration. Investigation of a national benefit level is also interesting because the result is *a priori* ambiguous - implementing a national welfare benefit may serve to increase or decrease migration rates.[9] One striking feature of our results is the insensitivity of migration to substantial changes in either benefits or wages.[10] Another rather surprising result is that uniform benefits increase migration. A natural intuition is that State variation in benefit levels should increase migration relative to a uniform benefit regime. The intuition is correct if the State benefit levels are independent of other influences on migration. However, benefits are in fact negatively correlated with these other influences, and thus serve to dampen migration flows.

Women are eligible for welfare benefits only if they are single, with dependent children, and if their earnings are low. Thus a complete specification of the value function would require a model of marriage and divorce, including a theory of how the marital surplus is divided, and of how likely it is that the surplus disappears, so that the marriage breaks up. This is a tall order. Moreover, to incorporate the temporal change in benefits requires a significant extension to our model – we must model beliefs about future benefits. For a discussion and an application of such forward-looking behavior that does not consider migration see Keane and Wolpin (2002a,b). In addition, a woman who is out of the labor force (either because she is collecting welfare or because she is married and doing non-market work) forgoes the human capital accumulation associated with labor market experience. Thus a fully specified model should encompass the relationship between current work and future wages, as in Shaw (1989); Imai and Keane (2004). In particular, the opportunity cost of being on welfare may be considerably higher than the current wage. Thus a more complete model would require a much larger state space than that used here, with marital status, number of children, and accumulated market work experience treated as state variables. Our model can be viewed as a simplification based on two approximations. First, when a welfare-eligible woman marries, she receives no surplus, either because the surplus is negligible, or because her share is negligible. Second, the experience associated with non-market work yields the same increment of human capital as the same amount of market work experience.

5 Household Migration Decisions

Mincer (1978) recognized that when married individuals have distinct preferences and different opportunities across locations, the most preferred location for a couple may differ from the locations preferred by each individual. Mincer assumed that couples maximize the sum of their incomes. The couple has to decide where to live and whether to stay together. This gives rise to notions of "tied-movers" and "tied-stayers" along with predictions on who should remain married and who should divorce. One interesting prediction is that migration is more likely after a divorce (as the newly unwed individuals move from their "tied" locations). Mincer also noted that these forces become stronger as women's labor force participation and earnings increase.

Gemici (2011) extends Mincer's static formulation with a model of household decision making that considers multiple locations over multiple periods. Gemici seeks to quantify the inhibiting effect of location ties on labor mobility, wage growth, and marital stability. The couple acts so as to maximize the expected present value of the sum of their consumption levels, allowing the possibility that this might imply that divorce is optimal. Consumption is a linear combination of a private good, a public good that is produced by the marriage, and leisure, with weights that depend on the duration of the marriage and the presence of children.

Location is defined as one of the nine Census regions in the United States. Each period each member of a household residing in location ℓ receives a wage offer from an alternative location with some probability. The household must then decide on where to live and whether to work, allowing for the possibility that the spouses might live apart to take advantage of attractive wage offers in distinct locations. The model is estimated by the method of simulated moments, using data from the Panel Study of Income Dynamics.

6 Immigration

The analysis of individual immigration decisions is in principle no different from the analysis of internal migration, but in practice empirical work in this area is severely limited by the scarcity of panel data sets spanning national borders.Thom (2010) and Lessem (2011) model migration between Mexico and the United States using one of the few available data sets, from the Mexican Migration Project. Thom (2010) uses a two-location model with borrowing constraints and a concave utility of consumption.[11] We focus on Lessem's model, which emphasizes income differences and family links as the main variables of interest, while allowing for repeat and return migration, and also allowing for other idiosyncratic influences on migration choices. Most of the migrants in the data set are illegal, and so the cost of migrating depends on border enforcement activity. Variations in border patrol man-hours along different sections of the border are used to estimate the relationship between cost and migration decisions. The richness of the model can be illustrated by considering the effect of increased enforcement on the length of spells in the U.S. For someone who is already in the U.S. illegally, tougher border enforcement means that it will be more difficult to return to the U.S. after returning to Mexico. Thus one somewhat paradoxical effect of increased enforcement is that it tends to keep illegal immigrants in the country for longer periods of time. Although Lessem's empirical results indicate that the magnitude of this effect is not large, the estimates illustrate the ability of the model to quantify behavioral responses that are just not accessible without a well-specified model of how migration decisions are made. Such a model of course entails making various restrictive assumptions, but the assumptions can be examined and modified, enabling the analysis to move beyond mere qualitative descriptions.

Lessem's model is estimated by maximum likelihood.[12] As in Gemici's model, when spouses have the option of living apart in order to take advantage of income opportunities in different locations, the number of relevant contingencies in the decision problem becomes unmanageable. Lessem assumes that one spouse (generally the husband) is the "primary" mover, and restricts the choice set so that the other spouse cannot choose to live in one of the U.S. locations unless the primary mover is also there; in addition, she assumes that the decisions of the two spouses are made in sequence, thereby restricting the number of choices that must be considered at each node of the decision problem. In the data it is rare to find the wife working in the U.S. while the husband stays in Mexico. Thus by excluding this option Lessem makes the required computations feasible without losing much in terms of realism. As a result, she can measure the extent to which migration is influenced by family ties, and in particular the relevance of family ties for return migration decisions.

In terms of substantive empirical results, the most important feature of Lessem's model is that it gives a coherent analysis of the extent to which increases in Mexican wages relative to U.S. wages leads to a reduction in the flow of immigrants to the U.S. At this stage, the model does not deal with general equilibrium effects of migration, but a well-specified model of the supply side of the labor market is an essential first step toward a full-blown general equilibrium analysis. An important feature of Lessem's model is that it allows for uneven development across regions within Mexico, so that emigration to the U.S. and internal migration within Mexico can be analyzed within the same decision problem.

Lessem's estimates imply that an increase in wages in Mexico reduces migration to the United States and increases return migration from the United States. A ten percent increase in Mexican wages reduces the amount of time that individuals spend in the United States by approximately nine percent. Increased border enforcement decreases not only immigration from Mexico, but also return migration from the United States. Simulations indicate that a fifty percent increase in enforcement would reduce the amount of time that migrants spend in the United States by up to nine percent, depending on the allocation of additional enforcement at the border.

7 Conclusion

We have presented an analytical framework capable of modeling individual migration decisions over the life cycle. Through our choice of examples we show that the framework can include family and household determinants as well as standard economic factors. The causal linkages among core demographic processes (such as migration, child bearing, household formation, marriage and divorce) can potentially be recovered assuming appropriate longitudinal (event history) data are available.

The framework requires parsimonious empirical specifications which frequently mandates strong functional form and or distributional assumptions. Yet, these assumptions can and should be empirically investigated. The most demanding challenge imposed by the framework is the specification of a parsimonious yet flexible specification that recovers the primary features of the data. The payoff from this analytical approach is the ability to trace out lifecycle and distributional consequences of policies or (social or economic) environments that are conjectured and not estimable using existing data. The increasing availability of longitudinal household microdata suggest a wealth of future research opportunities to investigate. Even though the work to date is quite limited, it shows that this approach is empirically fruitful and adds to our understanding of individual life cycle decision making.

Notes

[1]See Dierx (1988) for what we believe is the earliest model of return migration, while Thom (2010) offers a modern perspective as does the chapter on CIRCULAR MIGRATION.

[2]See Pessino (1991) for a model of learning applied to data on migration in Peru.

[3]John suggested we cite Bauer or Bayer look in IZA DP.

[4]This is Bellman's “Curse of Dimensionality”.

[5]Interestingly, recent work by Gregory (2011) suggests that rebuilding subsidies offered post–Katrina by the United States federal government incurred relatively small deadweight losses due to the low income elasticity of return migrants to New Orleans. See also the chapter on NATURAL DISASTERS and MIGRATION.

[6]Notowidigdo (2010) interprets the difference in migration rates between skilled and unskilled workers in terms of differential responses to local demand shocks. When there is an adverse local shock, house prices decline. Low-wage workers spend a large fraction of their income on housing, so the decline in the price of housing substantially reduces the incentive to migrate, while this effect is less important for high-wage workers. At the same time, public assistance programs respond to local shocks, and these programs benefit low-wage workers (although the relevance of this in explaining the differential migration rates for high-school and college graduates is doubtful, especially for men).

[7]Please see the chapter on WELFARE MIGRATION.

[8]The effect of welfare benefits and particularly AFDC on migration decisions of poor women with children has a long history. The common perception, dating back to the English Poor Laws of the 19^{th} Century, is that high welfare benefits attract poor (welfare–eligible) individuals from other locations, and retain poor among the local population. However, benefits may help finance

a move for liquidity constrained households (i.e., access to welfare benefits may lessen poverty traps). The empirical evidence on the influence of AFDC benefits is mixed. See Brueckner (2000) for a summary of the literature. In principle, the same issue faces countries within the European Union, there however, language differences and cultural differences may serve as barriers to migration.

[9]Consider setting a national welfare benefit level high relative to the distribution of state level benefits. Migration rates will decline as individual who left states with low benefits no longer need to move to secure a higher income floor. Conversely, if national benefits are set relatively low, migration will increase as people now have a stronger incentive to move to better (higher wage) labor markets. As discussed below, whether migration flows increase or decrease with a uniform benefit level set at the mean of the State distribution of benefits depends on whether State benefits are independent of other influences on migration.

[10]Using different methods, Giulietti et al. (ming) find that flows within the European Union are insenitive to unemployment benefit generosity.

[11]The main state variable in Thom's model is the level of assets. Initially, young people in Mexico don't have enough money to get across the border. They save enough to pay the cost, and then they migrate. In the U.S. they build up assets to the point where they would rather live in Mexico, because they can consume at a reasonably high level, and staying in the U.S. to build up a higher asset level is not worth it, because they have diminishing marginal utility, and a preference for living in Mexico. So they go back. Then after returning to Mexico, it may happen that the asset level falls to a point where it is optimal to return to the U.S.

[12]Thom (2010) uses the method of simulated moments.

Table 1: **Interstate Migration, White Male High School and College Graduates**

	High School		College	
	$\hat{\theta}$	$\hat{\sigma}_\theta$	$\hat{\theta}$	$\hat{\sigma}_\theta$
Utility and Cost				
Disutility of Moving (γ_0)	4.794	0.565	3.583	0.686
Distance (γ_1) (1000 miles)	0.267	0.181	0.483	0.131
Adjacent Location (γ_2)	0.807	0.214	0.852	0.130
Home Premium $\left(\alpha^H\right)$	0.331	0.041	0.168	0.019
Previous Location (γ_3)	2.757	0.357	2.374	0.178
Age (γ_4)	0.055	0.020	0.084	0.024
Population (γ_5) (millions)	0.654	0.179	0.679	0.116
Stayer Probability	0.510	0.078	0.221	0.058
Cooling (α_1) (1000 degree-days)	0.055	0.019	0.001	0.011
Income (α_0)	0.314	0.100	0.172	0.031
Wages				
Wage intercept	-5.133	0.245	-6.019	0.496
Time trend	-0.034	0.008	0.065	0.008
Age effect (linear)	7.841	0.356	7.585	0.649
Age effect (quadratic)	-2.362	0.129	-2.545	0.216
Ability (AFQT)	0.011	0.065	-0.045	0.158
Interaction(Age,AFQT)	0.144	0.040	0.382	0.111
Transient s.d. 1	0.217	0.007	0.212	0.007
Transient s.d. 2	0.375	0.015	0.395	0.017
Transient s.d. 3	0.546	0.017	0.828	0.026
Transient s.d. 4	1.306	0.028	3.031	0.037
Fixed Effect 1	0.113	0.036	0.214	0.024
Fixed Effect 2	0.296	0.035	0.660	0.024
Fixed Effect 3	0.933	0.016	1.020	0.024
Location match (τ_v)	0.384	0.017	0.627	0.016
Loglikelihood	-4214.160		-4902.453	
	4274 observations		3114 observations	
	432 men,124 moves		440 men, 267 moves	
Source: Estimates for high school graduates are from Table 2 of Kennan and Walker (2011). Estimates of college graduates are calculated from Kennan (2011).				

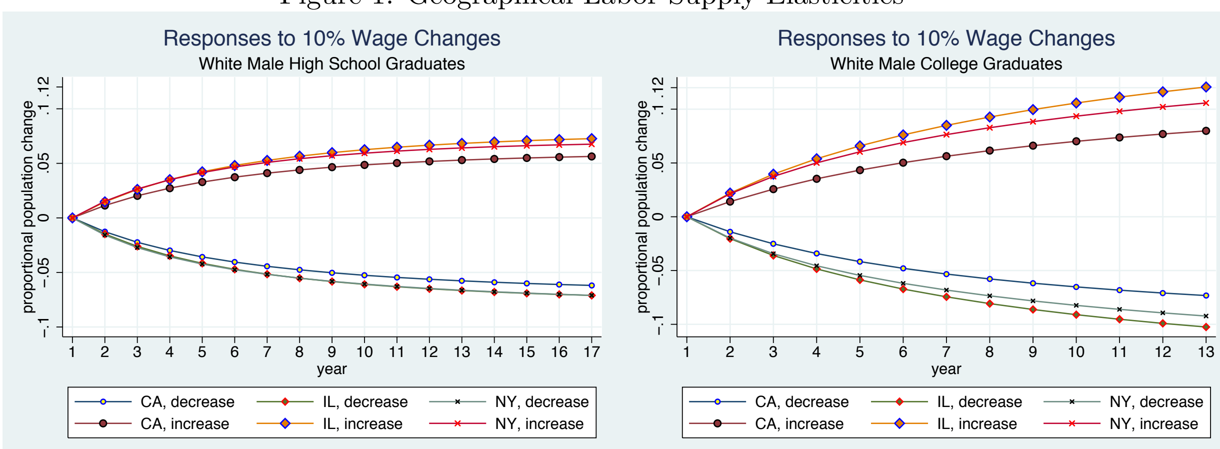
Table 2: **Average Moving Costs**

		Move Origin and Destination			
		From Home	To Home	Other	Total
Previous Location	None	-\$147,619	\$138,095	-\$39,677	-\$139,118
	Home	—	\$18,686	-\$124,360	-\$9,924
	Other	-\$150,110	\$113,447	-\$67,443	-\$87,413
	Total	-\$147,930	\$25,871	-\$97,656	-\$80,768
Source: Table 4 of Kennan and Walker (2011).					

Table 3: **Interstate Migration, White Male High School and College Graduates**

	High School	College Wages	College	HS Wages
	$\hat{\theta}$		$\hat{\theta}$	
Utility and Cost				
Disutility of Moving (γ_0)	4.794		3.570	
Distance (γ_1) (1000 miles)	0.267		0.482	
Adjacent Location (γ_2)	0.807		0.852	
Home Premium $\left(\alpha^H\right)$	0.331		0.167	
Previous Location (γ_3)	2.757		2.382	
Age (γ_4)	0.055		0.085	
Population (γ_5) (millions)	0.654		0.678	
Stayer Probability	0.510		0.227	
Cooling (α_1) (1000 degree-days)	0.055		0.001	
Income (α_0)	0.314		0.172	
Wages				
Location match (τ_v)	0.384		0.634	
NLSY Data				
Observations	4,274		3,114	
Migration Rate	2.90%		8.57%	
Simulated Observations	427,429	427,421	311,571	311,428
Migration Rate	3.16%	3.96%	8.60%	8.23%
Source: Estimates of high school graduates from Table 2 of Kennan and Walker (2011). Estimates of college graduates from Kennan (2011).				

Figure 1: Geographical Labor Supply Elasticities



References

Barro, R. J. and Sala-i Martin, X. (1991). Convergence across states and regions. *Brookings Papers on Economic Activity*, 1:107–158.

Blanchard, O. J. and Katz, L. F. (1992). Regional evolutions. *Brookings Papers on Economic Activity*, 1:1–37.

Bound, J. and Holzer, H. J. (2000). Demand shifts, population adjustments, and labor market outcomes during the 1980s. *Journal of Labor Economics*, 18:20–54.

Brueckner, J. K. (2000). Welfare reform and the race to the bottom: Theory and evidence. *Southern Economic Journal*, 66:505–525.

DaVanzo, J. (1983). Repeat migration in the united states: Who moves back and who moves on? *Review of Economics and Statistics*, 65:552–559.

Dierx, A. H. (1988). A life-cycle model of repeat migration. *Regional Science and Urban Economics*, 18(3):383–397.

Gemici, A. (2011). Family migration and labor market outcomes. Technical report, New York University, Department of Economics.

Giulietti, C., Guzi, M., Kahanec, M., , and Zimmermann, K. F. (forthcoming). Unemployment benefits and immigration: Evidence from the eu. *International Journal of Manpower*.

Greenwood, M. J. (1997). Internal migration in developed countries. In Rosenzweig, M. R. and Stark, O., editors, *Handbook of Population and Family Economics*, volume 1B. North Holland.

Gregory, J. (2011). The impact of rebuilding grants and wage subsidies on the resettlement choices of hurrican katrina victims. Technical report, University of Michigan, Department of Economics.

Imai, S. and Keane, M. P. (2004). Intertemporal labor supply and human capital accumulation. *International Economics Review*, 45(2):601–641.

Keane, M. P. and Wolpin, K. I. (2002a). Estimating welfare effects consistent with forward–looking behavior. part i: Lessons from a simulation exercise. *Journal of Human Resources*, 37(3):570–599.

Keane, M. P. and Wolpin, K. I. (2002b). Estimating welfare effects consistent with forward–looking behavior. part ii: Lessons from a simulation exercise. *Journal of Human Resources*, 37(3):600–622.

Kennan, J. (2011). Higher education subsidies and human capital mobility. Technical report, University of Wisconsin–Madison.

Kennan, J. and Walker, J. R. (2010). Wages, welfare benefits and migration. *Journal of Econometrics*, 156(1):229–238.

Kennan, J. and Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, 79(1):251–251.

Lessem, R. (2011). Mexico–u.s. immigration: Effects of wages and border enforcement. Technical report, Carnegie Mellon University, Tepper School of Business.

Malamud, O. and Wozniak, A. (2009). The impact of college education on geographic mobility: Identifying education using multiple components of vietnam draft risk. Unpublished, University of Notre Dame.

McFadden, D. D. (1974). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*. Academic Press, New York.

Mincer, J. (1978). Family migration decisions. *Journal of Political Economy*, 86:749–773.

Molloy, R., Smith, C. L., and Wozniak, A. (2011). International migration in the united states. *Journal of Economic Perspectives*, 25(3):173–196.

Notowidigdo, M. J. (2010). The incidence of local labor demand shocks. Unpublished, MIT.

Pessino, C. (1991). Sequential migration theory and evidence from peru. *Journal of Development Economics*, 36:55–87.

Rust, J. P. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica*, 55(5):999–1033.

Schultz, T. P. (1982). Lifetime migration within educational strata in venezuela: Estimates of a logistic model. *Economic Development and Cultural Change*, 30:559–593.

Schultz, T. W. (1961). Investment in human capital. *The American Economic Review*, 51(1):1–17.

Shaw, K. (1989). Life cycle labor supply with human capital accumulation. *International Economic Review*, 30:431–456.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy*, 70:80–89.

Thom, K. (2010). Repeated circular migration. Technical report, Department of Economics, New York University.

Topel, R. H. (1986). Local labor markets. *Journal of Political Economy*, 94(3 part 2):S111–S143.

Wozniak, A. (2010). Are college graduates more responsive to distant labor market opportunities? *Journal of Human Resources*, 45(4):944–970.