# Intergenerational altruism in the migration decision calculus: Evidence from the African American Great Migration

John Gardner*

December 2017

**Abstract:** It is widely believed that many migrations are undertaken at least in part for the benefit of future generations. To provide evidence on the effect of intergenerational altruism on migration, I estimate a dynamic residential location choice model of the African American Great Migration in which individuals take the welfare of future generations into account when deciding to remain in the Southern United States or migrate to the North. I measure the influence of altruism on the migration decision as the implied difference between the migration probabilities of altruistic individuals and myopic ones who only consider current-generation utility when making their location decisions. My preferred estimates suggest that intergenerational altruism explains between 24 and 42% of the Northward migration that took place during the period that I study, depending on the generation.

**Keywords:** Altruism, intergenerational altruism, migration, immigration, Great Migration, dynamic discrete choice.

**JEL codes:** J61, D64, R23.

*Department of Economics, University of Mississippi, 371 Holman Hall, University MS, 38677. Email: jrgardne@olemiss.edu, Phone: +1-662-915-7291; ORCID: 0000-0003-2946-4282.

# 1 Introduction

The idea that part of what motivates people to bear the financial and psychological costs of moving away from home is that, by doing so, they can better the lives of their children, grandchildren, etc. is part of our collective folk wisdom. Perhaps because of the ubiquity of this belief, there has been little work to quantify the influence of intergenerational altruism on the propensity to migrate. In this paper, I estimate an intergenerational model of the migration decision during the African American Great Migration, a period spanning 1915-1970 during which millions of blacks left their birthplaces in the Southern US in favor of cities in the North (see Tolnay, 2003 for a comprehensive review). My preferred estimates suggest that intergenerational altruism explains between 20 and 40% of the migration observed during this period, depending on the generation.

Many studies have examined intergenerational dimensions of migration. In an early contribution, Chiswick (1977) finds that second-generation immigrants fare better in the labor market than their parents and children. Borjas (1993) develops and tests a model of dynastic selection into migration, finding evidence of negative selection among second-generation immigrations whose parents hail from countries with high income inequality. Deutsch, Epstein and Lecker (2006) argue that the intergenerational earnings patterns first identified by Chiswick can be explained by reciprocal parent-child altruism and show that, consistent with their theoretical model, the differences in earnings across successive generations of immigrants to Israel cannot be explained by differences in observable characteristics. Caponi (2011) develops an intergenerational structural model of the migration decision, finding that altruism helps explain why second-generation immigrants accumulate more human capital than their first- and third-generation counterparts.

There is also a small literature on whether altruism motivates migration. Tcha (1995, 1996) develops and tests models with reciprocal altruism, finding that they improve the ability to explain rural-urban migrant flows in Korea and the US. Berman and Rzakhanov (2000) test their hypothesis that immigrants are self-selected on intergenerational altruism using data on the fertility of Eastern European immigrants after a policy change that dramatically increased access to Israel. The relationship between altruism and migration has also been investigated in the context of remittances (see, e.g., Lucas and Stark, 1985; Shen, Docquier and Rapoport, 2009). Papers in this strand of the literature present models in which some agents are willing to migrate in order to remit portions of their earnings back to family members in the source country.

These papers suggest that migration is an important source of intergenerational dynamics in economic outcomes, and one that may be motivated in part by altruism towards future generations. However, they tell us little about how much intergenerational altruism affects migrant flows. To provide insight into this question, I use intergenerational panel data on three generations of black American families to estimate a simplified Barro-Becker (1989, c.f. Becker and Tomes, 1986;

Becker and Barro, 1988) model in which parents view the residential location decision in part as an investment in future generations. The estimated parameters of the model allow me to compute the counterfactual probabilities of Northward migration among Southern-born blacks that would arise if the migration decision were made only on the basis of current-generation utility—that is, if parents neglected future generations' welfare when making residential location choices. In turn, comparing observed with counterfactually myopic migration probabilities allows me to estimate the effect of intergenerational altruism on the propensity to migrate.

Estimates of my simplest specification imply that intergenerational altruism increased the probability of migrating North from 10% to 13% for members of the first Southern-born generation of my sample, and from 12% to 13% for members of the Second. These estimates imply that most Southerners would not have migrated regardless of their altruism towards future generations. Furthermore, descriptive regressions show that observable characteristics have limited ability to explain the migrations that did occur. Accordingly, I also estimate models that allow location preferences to depend on unobserved state variables. These models imply that most Northward migrations were undertaken by members of families with strong unobserved preferences for the North among whom altruism is much more influential. I find that intergenerational altruism increased the probability of migrating by an average of about 16% for members of this subpopulation (from 22% to 38% among for the first generation and from 50% to 66% for the second), implying that altruism explains between 24 and 42% of the observed migrations, depending on the generation. Though in theory intergenerational altruism can encourage some parents to move while discouraging others, I show that in practice the influence of altruism on migration behavior occurs primarily through the first channel.

## 2 Theoretical motivation

The underlying idea behind this paper's approach is straightforward. First, I estimate a dynamic model of residential location choice in which Southern-born individuals decide to remain in the South or migrate to the North, taking the effect of their decision on the welfare and location decisions of future generations into account. Second, I use the estimated parameters of the model to ask how each generation's location decisions would have differed if they myopically ignored future generations' welfare when making their own residential location decisions.

A simple model helps motivate the idea behind this approach, clarify key concepts, and establish notation. The focus here is on the underlying idea; I detail my specific empirical implementation below. Let $g \in \{1, 2, \dots\}$ index generations of a dynasty or family and $l \in \{s, n\}$ denote the choice of South or North. Suppose that generation $g$ of the dynasty is born in location $b_g \in \{s, n\}$ and endowed with observed characteristics $x_g$ and idiosyncratic (that is, uncorrelated across generations) and additively separable preferences $\varepsilon_{lg}$ over locations. Beginning with generation $g$, the decision problem facing the dynasty is to choose a sequence $(l_g, l_{g+1}, \dots)$ of residential locations to solve

$$\max_{l_g, l_{g+1}, \dots} \left\{ E_g \sum_{h=g}^{\infty} \lambda^{h-g} [u_{lh}(x_h, b_h) + \varepsilon_{lh}] \right\}, \tag{1}$$

where $u_{lg}$ is the deterministic part of the flow utility accruing to generation $g$ in location $l$, $\lambda \in (0,1)$ is the intergenerational discount factor or rate of intergenerational altruism, the expectation is taken with respect to information known to generation $g$, and I assume for simplicity that each generation has one child (an assumption which I relax in estimation).

The somewhat unwieldy problem of simultaneously choosing the optimal locations $(l_g, l_{g+1}, \dots)$ for each generation becomes easier to analyze when it is transformed into a sequence of generation-specific location choices, made under the assumption that future generations choose their locations optimally. To this end, let $(l_g^*, l_{g+1}^*, \dots)$ solve (1), and denote the maximized solution to this problem by the value function

$$V(x_g, b_g) = E_g \left\{ \sum_{h=g}^{\infty} \sum_{l \in \{s,n\}} \lambda^{h-g} 1(l_h^* = l) [u_{lh}(x_h, b_h) + \varepsilon_{lh}] \right\}, \tag{2}$$

where $1(\cdot)$ is the indicator function. Finally, assume that the state variables $x_g$ evolve according to a first-order Markov process and let the conditional valuation function

$$v_{lg}(x_g, b_g) = u_{lg}(x_g, b_g) + \lambda E[V(x_{g+1}, b_{g+1} = l_g) | x_g, b_g] \tag{3}$$

represent the total utility accruing to generation $g$ from choosing location $l$, net of the idiosyncratic preference component $\varepsilon_{lg}$, under the assumption that after $g$ chooses $l$ all future generations choose their locations optimally. Then the decision facing generation $g$ can be expressed simply as

$$\max_{l \in \{s,n\}} [v_{lg}(x_g, b_g) + \varepsilon_{lg}] \tag{4}$$

and the value function can be obtained recursively through

$$E[V(x_g, b_g)] = E \left\{ \max_{l \in \{s,n\}} [v_{lg}(x_g, b_g) + \varepsilon_{lg}] \right\}. \tag{5}$$

Within this framework, a natural measure of the influence of intergenerational altruism on migration is the average difference in the probability of migrating North between altruistic and myopic agents, or

$$\Delta(x_g, b_g = s) = E\{1[v_{ng}(x_g, b_g = s) + \varepsilon_{ng} > v_{sg}(x_g, b_g = s) + \varepsilon_{sg}] \\ - 1[u_{ng}(x_g, b_g = s) + \varepsilon_{ng} > u_{sg}(x_g, b_g = s) + \varepsilon_{sg}]\}, \quad (6)$$

where the expectation is taken over the $\varepsilon_{lg}$. The effect of intergenerational altruism on migration, thus measured, operates through the dependence of the conditional valuation functions on birth location. If $u_{lg}(x_g, b_g = l) \neq u_{lg}(x_g, b_g = l')$, perhaps because moving is costly (or equivalently, because location preferences depend on birthplace), parents may migrate to their own detriment to spare future generations the cost of doing so. Alternatively, parents may avoid migrating, again to their own detriment, to spare future generations who may not share their location preferences a burdensome choice between remaining in a location for which they are poorly suited or migrating back.

# 3 Data

Because some of the modeling choices that I make are motivated by data availability, it is useful to begin with a description of the available data and their advantages and disadvantages. My implementation draws on data from the Three-Generation National Survey of Black American Families (Jackson and Tucker, 1997), a nationally-representative survey of black American families conducted between 1978 and 1981. These data consist of cross-sectional responses to the same survey questions from members of three different generations of a total of 510 families. The primary advantage of this dataset is its intergenerational structure, which allows me to follow these families through almost the entire Great Migration period (as I discuss below, the intergenerational panel structure is also useful for identifying models with unobserved heterogeneity; see Kasahara and Shimotsu, 2009 and Hu and Shum, 2012). The data also contain important demographic variables, including state of birth, state of residence when surveyed, age, gender, educational attainment, and fertility.[1] However, their retrospective cross-sectional nature means that the data are not informative about the timing of variables such as migration, education, or fertility. In addition, because the survey focuses on sociological questions, data on earnings and employment are limited and often missing.

[1] I code respondents' current and birth locations as Southern according to the Census Bureau's definition of the Southern region. According to this definition, Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia and West Virginia are Southern states. I code all other locations within the US as belonging to the North. I then classify Southern-born respondents as migrants if they lived in the North at the time of the survey. Because the data do not contain information on the timing of migration, it is unavoidably possible under this scheme that the actual migration decision was made by respondents' parents.

At the time the surveys were administered, some younger third-generation respondents may have had incomplete histories of education, fertility, and migration. The modest sample size presents a tradeoff between standardizing on ages to ensure complete demographic histories and retaining enough data to permit meaningful inference. Any right-censoring likely understates the probability of migrating among third-generation respondents, and hence the relationship between migration and their observed characteristics. Keeping the likely direction of any resulting bias in mind, I err on the side of inclusion, omitting only third-generation respondents who were younger than 18 when surveyed.

After performing these sample selections, I am left with 1,099 observations from members of 443 families (or 916 observations from the 370 families for whom all members meet the sample requirements). Table 1 provides sample characteristics for these observations. The median respondent from the first generation was born in 1910 and 70 years old when the survey was administered. The median second-generation respondent was born in 1933 and 47 years old at the time of the survey. The median third-generation respondent was born in 1954 and was 24.5 when surveyed. Members of the first and second generations are substantially more likely to be female, almost certainly a consequence of higher mortality among men. I return to this point, which suggests that the sample distribution of gender does not accurately reflect the population distribution earlier in life, in Section 4.4.

The summary statistics in Table 1 also provide a sense of the likelihood of Northward migration. Averaging across generations, 86% of individuals choose to remain in the South. Members of the second generation are about 7% more likely to migrate than the sample average, although the similarity between the migration behavior of first- and third-generation respondents may partially owe to the younger ages at which the latter were interviewed.[2] Unsurprisingly, the table also shows that successive generations obtained more education, with median years of completed schooling increasing from 7 to 11 to 12 across the three generations. Similarly, the median number of children declines from 6 to 4 to 1, though right-censoring among the third generation probably impacts the observed rates of change in education and fertility.

To provide preliminary evidence on how key covariates are related to location preferences, I present in Table 2 a series of linear models of the probability that a Southern-born individual migrates North. For simplicity, these models use binary measures of education (an indicator for having above the generation-specific median years of schooling) and fertility (an indicator for having greater than the median number of children); using full sets of education and fertility indicators produces similar results. In the first specification, which pools observations from all three genera-

[2] These modest migration rates are not at odds with the notion that Northward migration was widespread. Although a relatively small fraction of any given generation migrated North, overall about 35% of families eventually left the South; this figure is roughly comparable to outmigration rates documented elsewhere (see, for example, Tolnay, 2003).

tions, education is the only statistically significant predictor of migration. The covariates included in this regression only explain about 1% of the variation in migration. The second specification interacts the regressors with time, revealing some intergenerational heterogeneity in how they relate to migration. These generation-specific regressors still only explain about 3% of the variation in migration. For the third specification, which includes all interactions between time, gender, education, and fertility, the $R^2$ only increases to .04. The low predictive power of human capital and fertility—arguably two of the most important choices made over the life cycle—shows that the migration decision is highly idiosyncratic and made in large part on the basis of factors not observed in the data.

# 4 Specification, identification, and estimation

To provide robust evidence on the effect of intergenerational altruism on Northward migration, I estimate a series of models of the decision to remain in the South or migrate North during the Great Migration period. While the models that I estimate differ in their implementation details, the common structure that they follow is motivated by empirical aspects of the Great Migration and the conditions necessary to identify the components of the models.

## 4.1 The intergenerational environment

After omitting observations with missing values for the key covariates, there are only 11 North-South migrations, making it virtually impossible to model the location decision problem for those born in the North. Instead, I model the decision of Southern-born agents to remain in the South or migrate North under the assumption that the North is absorbing—after a Northward migration, all future generations remain in the North, effectively terminating the dynamic decision problem. For this reason, I also omit respondents who were born in the North.[3]

As Tolnay (2003) notes, the Great Migration was effectively over by the mid 1970s, rendering the dynamic residential location decision problem faced by Southerners during this period nonstationary. To model this nonstationarity, I assume that forces such as the Civil Rights Movement and the interregional diffusion of labor, capital, and technology brought the Northern and Southern labor markets into a steady-state equilibrium, leaving agents indifferent between living in either location. In this case, after the 1970s (which corresponds roughly to the fourth generation in my dataset), Southern- and Northern-born agents would have faced the same expected lifetime utility, eliminating the dynamic aspect of the location choice problem.

[3]This is not a strong assumption. The lack of return migration necessarily means that residents of the North either have preferences for living there, or face costs of Southward migration, that approach infinity. Treating the North as absorbing simply obviates the need to estimate these extreme parameters.

To interpret intergenerational altruism in a dynastic model, I assume that children's endowments are not fully realized until they have matured to adulthood, at which point parents receive utility from the welfare of their children. Accordingly, I set the intergenerational discount factor $\lambda$ to $.95^{25} \approx .28$. This value is in line with those used in other studies. Caponi (2011), studying Mexican immigration between 1994 and 2008, assumes a value of $.9615^{30} \approx .31$. Glover and Heathcote (2011) use wealth data during the Great Recession period to calibrate a parameter of .31. In an empirical model of educational investment in children, Heckman and Raut (2016) estimate a parameter of .44. Though, to avoid overstating the influence of altruism on migration, I use a value on the low end of those in other studies, I show in Section 5 that my estimates are (at least locally) insensitive to the assumed rate of intergenerational altruism.[4]

In the theoretical framework outlined above, I make the simplifying assumption that each generation has exactly one child. This is obviously unrealistic, and fertility may alter the incentives to migrate both by changing the cost of migrating and by increasing the number of potential future beneficiaries of each generation's location decision. At the same time, while the descriptive statistics in Table 1 show a clear decline in mean fertility over time, it is unlikely that members of earlier generations derived more utility from the welfare of their children, grandchildren, etc. To capture the changing tradeoff between the quantity and quality of children simply, I define a fertility indicator $f_g$ for whether the number of children $k_g$ born to generation $g$ is above the generation-specific median $\tilde{k}_g$. I then calculate the ratio $\mu_g = E(k_g|k > \tilde{k}_g)/E(k_g|k \leq k_g)$ of children between those above and below the median and inflate the intergenerational discount factor for relatively fertile members of each generation by $\mu_g$ (that is, I discount the future by an effective rate of $[(1-f_g)+f_g\mu_g]\lambda$).

## 4.2 Flow utility functions

As Magnac and Thesmar (2002) note, while dynamic discrete choice models are generically underidentified, given a discount factor and a distribution governing the idiosyncratic preference terms, the differences in the flow utility functions are identified under relatively weak conditions. Accordingly, I assume that the $\varepsilon_{lg}$ are iid draws from a Type I Extreme Value distribution and normalize the deterministic component of the utility of the South to zero, so that the total utility of living in

[4]Using a calibrated discount factor is standard practice (see Magnac and Thesmar, 2002; Arcidiacono and Miller, 2017). Because the steady-state assumption is equivalent to a finite-time horizon, the discount factor can technically be identified under the assumption that the flow utility functions are stable across time, though identification in this case is achieved only tentatively through the nonlinear manner in which the discount factor enters the likelihood function. In contrast, models such as that in Heckman and Raut (2016) recover altruistic preferences directly when parents make investments that can only benefit them through their children, making them much better-suited to estimate the altruism parameter.

the South is $u_{sg}(x_g, b_g) + \varepsilon_{sg} = \varepsilon_{sg}$ for all $g$.[5] I then parameterize the flow utility of living in the North using variations on the primary specification

$$u_{ng}(x_g, b_g) + \varepsilon_{ng} = x_g'\beta - c \cdot 1(b_g = s) + \varepsilon_{ng}, \tag{7}$$

where $x_g$ is a vector of observable characteristics, $\beta$ is a vector of parameters, and $c$ is the cost of migrating from the South to the North.[6]

To capture the relationship between migration and other important life-cycle decisions such as human capital accumulation and fertility, I include in $x_g$ indicators $e_g$ and $f_g$ for having above the generation-specific mean years of education and number of children (as I discuss below, I discretize these measures in order to estimate with greater precision how they evolve from parent to child).[7] To allow for the possibility that the likelihood of Northward migration differs by gender (for example because men face a higher regional wage difference or have greater bargaining power over location choices), I also include an indicator $m_g$ for being male. In addition, my primary specification includes a linear intergenerational time trend that allows the attractiveness of the North to change over time. The coefficients on these variables should be viewed as "reduced-form structural" parameters that reflect geographic differences in real wages as well as regional amenities and other non-pecuniary preferences.

In any model of migration, disentangling moving costs from pure location preferences requires some sort of normalization. This is particularly true when migration terminates the dynamic decision problem, since in this case moving costs cannot be identified by comparing the location choices of otherwise-similar individuals born in different locations. In my model, the cost of moving is identified by variation in the behavior of otherwise similar members of different generations under the assumption that the parameters of the current utility functions are stable over time (that is, $\beta$ and $c$ are not indexed by $g$).[8] Although this assumption is restrictive, the finite time hori-

[5] The use of the Type I Extreme Value distribution is common because it generates analytical expressions for the conditional valuation functions; in the static case, it is equivalent to using a logit model.

[6] In principle it is possible to allow the cost of migration to depend on the covariates. In practice, the lack of power of the covariates to explain observed migrations means that these covariate-specific costs are not well identified.

[7] Although they are likely choice variables, I treat education and fertility as exogenously endowed in order to focus on the migration decision. Note however that jointly choosing education, fertility and location is equivalent to choosing education and fertility with the understanding that the optimal location will depend on these choices.

[8] To see this formally, consider a simplified model without covariates in which $u_{ng}(b_g) = \beta_0 - c1(b_g = s)$ and $u_{sg} = 0$. Since the dynamic decision problem ends after the third generation, the North-South difference in conditional valuation functions for Southern-born members of the third generation is $v_{n3}(s) - v_{s3}(s) = \beta_0 - c$. Since the North is an absorbing state, it follows from properties of the Type I Extreme Value distribution (see, e.g., Rust, 1987) that for the second generation this difference is $v_{n2}(s) - v_{s2}(s) = u_{n2}(s) + \lambda E[V_3(n) - V_2(s)] = \beta_0 - c + \lambda\{\beta_0 + \gamma - \log[1 + \exp(\beta_0 - c)] - \gamma\}$, where $\gamma \approx .5772$ is Euler's constant. Since $\beta_0$ appears in the $v_{l2}$ independently of the $\beta_0 - c$ term (and since the probability that generation $g$ migrates is the same as the probability that $v_{ng} > v_{sg}$) the latter can be identified from maximum likelihood on the third generation and the former from maximum likelihood on the second. Note that, by this logic, $\lambda$ is technically identified from maximum likelihood on $v_{1n}(s) - v_{1s}(s) =$

zon of the model and the inclusion of an intergenerational time trend (and in some specifications, an interaction between time and the migration cost) allows location preferences to vary across generations.

### 4.3 Conditional valuation functions

Under the assumption that the North is an absorbing state, all Northern-born residents remain in the North with probability one, receiving expected flow utility of

$$E[u_{ng}(x_g, b_g = n) + \varepsilon_{ng} | x_g, b_g = n] = u_{ng}(x_g, b_g = n) + \gamma, \tag{8}$$

where $\gamma \approx .5772$ is Euler's constant. One consequence of this is that it suffices to consider only the location decisions of Southern-born individuals when estimating the parameters of the model. Another consequence is that the conditional valuation function associated with migrating North for any Southern-born generation is simply

$$\begin{aligned} v_{ng}(x_g, b_g = s) &= u_{ng}(x_g, b_g = s) \\ &\quad + E\left\{ \sum_{h=g}^{2} [(1-f_h) + f_h \mu_h] \lambda^{h-g} [u_{n,h+1}(x_{h+1}, b_{h+1} = n) + \gamma] \Big| x_g, b_g = s \right\}. \end{aligned} \tag{9}$$

Furthermore, under the assumption that the North and South reached a steady-state equilibrium by the fourth generation, the decision problem facing members of the third generation is static, so that

$$v_{l3}(x_3, b_3) = u_{l3}(x_3, b_3) \tag{10}$$

for all $x_3$, $b_3$, and $l \in \{s, n\}$. Combining these observations, for Southern-born members of generations $g < 3$ we also have that

$$\begin{aligned} v_{sg}(x_g, b_g = s) &= u_{sg}(x_g, b_g = s) + [(1-f_g) + f_g \mu_g] \lambda E\left[ \max_{l \in \{s,n\}} v_{l,g+1}(x_{g+1}, b_{g+1} = s) | x_g, b_g = s \right] \\ &= u_{sg}(x_g, b_g = s) \\ &\quad + [(1-f_g) + f_g \mu_g] \lambda E\left[ \log\left( \sum_{l \in \{s,n\}} \exp[v_{l,g+1}(x_{g+1}, b_{g+1} = s)] \right) + \gamma \Big| x_g, b_g = s \right], \end{aligned}$$

where the second equality follows from the assumption that the $\varepsilon_{lg}$ are draws from a Type I Extreme Value distribution (Rust, 1987).

$\beta_0 - c + (\lambda + \lambda^2)(\beta_0 + \gamma) - \lambda\{\log[\exp(v_{n2}(s)) + \exp(v_{s2}(s))]\}$, though only through the nonlinear way that it enters into this function.

## 4.4 State transitions

Representing the conditional valuation functions, which involve expectations of future value terms, requires knowledge of the transitions $f_{g+1|g}(x_{g+1}|x_g)$ between the observed state variables. Estimating these transitions poses two challenges. First, the samples for each generation are somewhat small, making it difficult to estimate transitions between a large number of state variables with any precision. To maximize the sizes of the samples available to estimate the transitions, I use discretized measures of education and fertility and pool Northern- and Southern-born individuals. Second, the overrepresentation of women in the samples of older generations means that the observed transitions do not reflect the actual transitions between gender, fertility, and education states. To address this problem, I first specify the probability that generation $g+1$ belongs to fertility-education state $\tilde{x}_{g+1} \in \{1,\ldots,4\}$ conditional on an indicator $m_{g+1}$ for whether generation $g+1$ is male and indicators $f_g$ and $e_g$ for generation $g$'s relative fertility and education using the generation-specific multinomial logit structure

$$P(\tilde{x}_{g+1} = k|m_{g+1}, f_g, k_g) = \frac{\exp(\rho_{k0g} + \rho_{k1g} m_{g+1} + \rho_{k2g} f_g + \rho_{k3g} e_g)}{\sum_{j=1}^{4} \exp(\rho_{j0g} + \rho_{j1g} m_{g+1} + \rho_{j2g} f_g + \rho_{j3g} e_g)}, \tag{11}$$

for $g \in 1,2$. I then assume that male and female children are equally likely and compute the transitions between each of the eight gender, fertility, and education states $x_{g+1} \in \{1,\ldots,8\}$ as

$$f_{g+1|g}(x_{g+1}|x_g) = f_{g+1|g}(\tilde{x}_{g+1}, m_{g+1}|x_g) = .5P(\tilde{x}_{g+1}|m_{g+1}, f_g, k_g). \tag{12}$$

## 4.5 Estimation

Let $\theta = (\beta, \rho)$ denote the parameters of the model (that is, the parameters of the flow utility and state transition functions, respectively). Given the assumptions that the North is an absorbing state and that the $\varepsilon_{lg}$ are iid, the likelihood of observing a family's sequence of location decisions and observable characteristics, conditional on the observable characteristics of the first Southern-born generation, is

$$\begin{aligned} \ell\left(l_1, \{x_g, l_g\}_{g=2}^{3} | x_1, b_1 = s; \theta\right) = P(l_1|x_1, b_1 = s; \theta) f_{2|1}(x_2|x_1; \rho) \\ \times P(l_2|x_2, b_2 = s; \theta) f_{3|2}(x_3|x_2; \rho) P(l_3|x_3, b_3 = s; \theta), \end{aligned}$$

which upon taking logs becomes

$$\log\left[\ell\left(l_1,\{x_g,l_g\}_{g=2}^{3}|x_1,b_1=s;\theta\right)\right]=\sum_{g=1}^{3}\log[P(l_g|x_g,b_g=s;\theta)]$$
$$+\sum_{g=1}^{2}\log[f_{g+1|g}(x_{g+1}|x_g;\rho)]. \quad (13)$$

The additive separability of the $f_{g+1|g}$ terms in (13) implies that $\rho$ can be estimated in a first step via the multinomial logit model defined in (11). The estimated transition functions, presented in Appendix Tables 10 and 11, agree with common sense. For example, male children born to first-generation parents are most likely to share their parents (relative) education and fertility, with the exception of those born to high-education, high-fertility parents, among whom there is more heterogeneity.[9] The apparently lower state persistence between generations two and three is likely a consequence of right-censored education and fertility histories for members of the third generation (as evidenced by the relatively high likelihood of transiting to a low-fertility, low-education state for this generation).

Substituting the parameterization (7) of the flow utility functions into the expressions for the $v_{lg}$ derived in Section 4.3 and using the assumption that the $\varepsilon_{lg}$ are Type I Extreme Value gives

$$P(l_g|x_g,b_g=s;\theta)=\frac{1(l_g=s)+1(l_g=n)\exp[v_{ng}(x_g,b_g=s;\theta)-v_{sg}(x_g,b_g=s;\theta)]}{1+\exp[v_{ng}(x_g,b_g=s;\theta)-v_{sg}(x_g,b_g=s;\theta)]}. \quad (14)$$

Thus, the remaining parameters $\beta$ of the model can be obtained in a second stage by maximizing the sample analog of

$$E\left(\sum_{g=1}^{3}\log[P(l_g|x_g,b_g=s;\beta,\hat{\rho})]\right) \quad (15)$$

with respect to $\beta$.[10]

# 5 Estimates

Table 3 summarizes estimates of the primary specification of the model, in which the flow utility function is given by (7) with $x_g=(1,m_g,f_g,e_g,g)$ containing a constant, indicators for being male and having above-median fertility and education, and a linear intergenerational time trend. The

[9] By design, the estimated transitions to state space elements for generation $g+1$ do not depend on the gender of the parent observed in generation $g$.

[10] Because maximizing the expected sum is equivalent to maximizing the expected summand, this is the dynamic analog of a standard logistic regression model, treating the future value terms as observed.

estimated flow utility function parameters are presented in the rightmost panel of the table. The estimates imply that the cost of migrating is the single most important determinant of the migration decision, exceeding the magnitudes of the remaining parameters of the model by at least a factor of four. This large estimated moving cost is consistent with the descriptive evidence presented in Table 2 that observed characteristics have low power to explain observed location choices. It is also consistent with evidence from empirical models of contemporary migration behavior (see, for example, Kennan and Walker, 2011; Bishop, 2012), which find that high moving costs must be invoked to explain why migrations are relatively uncommon.

Though their explanatory power is limited, the coefficients on the observables included in the flow utility functions agree both with intuition and the descriptive regressions in Table 2. The estimated coefficient on male is small and statistically insignificant. Fertility may affect migration directly through location preferences or indirectly through the number of future generations affected by the location decision. The estimated coefficient on fertility suggests that this direct effect is modestly negative.[11] The positive coefficient on education is consistent with the idea that the return to skill, particularly for blacks, was higher in the North (it is also consistent with prior evidence that more educated Southerners were more likely to move North; see, for example, Tolnay, 2003). The positive coefficient on $g$ (time) reflects increasing rates of migration across the generations, after accounting for the changing distributions of observable characteristics. Though most of observed state variables are individually insignificant, a Wald test of their joint significance easily rejects the null (as does a test of the joint significance of all of the flow utility function parameters).

The left panel of the Table 3 summarizes the predicted migration probabilities of altruistic individuals who take future generations' welfare into account when making their location decisions and myopic individuals who only take their own utility into account, as well as the differences $\Delta(x_g, b_g = s)$, defined in (6), between these two.[12] The altruistic probabilities, presented under the label "dynamic," approximate the observed probabilities reasonably well, given the parisomonious model.[13] Both these and the myopic probabilities, labelled "static," vary as expected across observed strata.

I interpret the differences between the migration probabilities of altruistic and myopic individuals as the contribution of intergenerational altruism to the decision to migrate. As the table

[11]This may suggest that larger families prefer the South, that fertility and preferences for the South are positively correlated with an unobserved factor, or since fertility and location decisions may be made jointly, that living in the North discourages fertility.

[12]I compute the dynamic probabilities as $\exp[v_{gn}(x_g, b_g = s; \hat{\theta}) - v_{gs}(x_g, b_g = s; \hat{\theta})] \cdot \{1 + \exp[v_{gn}(x_g, b_g = s; \hat{\theta}) - v_{gs}(x_g, b_g = s; \hat{\theta})]\}^{-1}$ and the static probabilities as $\exp[u_{gn}(x_g; b_g = s; \hat{\theta}) - u_{gs}(x_g, b_g = s; \hat{\theta})] \cdot \{1 + \exp[u_{gn}(x_g, b_g = s; \hat{\theta}) - u_{gs}(x_g, b_g = s; \hat{\theta})]\}^{-1}$.

[13]The predicted rates for the third generation are considerably higher than the observed rates. However, since some of the migration histories for this generation are likely right censored, this can be viewed as a benefit of using stable utility parameters and allowing for a time trend.

shows, the absolute effects of altruism on the probability of migration are modest; on average, altruism increases the probability of migration by about 2.7% for Southern-born members of the first generation and about 1% for Southern-born members of the second (the effect is zero for the third generation by the time-horizon assumption). In relative terms, altruism explains about 23% (=.03/.13) of the migration behavior of the first generation and about 8% of that behavior for the second (though it should be noted that neither of these average differences are statistically significant). The effects of altruism vary as expected with education. More educated individuals, whose descendants are likely to be relatively well educated and therefore have stronger preferences for the North, are more likely to migrate in order to improve the welfare of future generations. Although the estimated direct effect of fertility on migration is negative, the effect of intergenerational altruism on migration varies positively with fertility, suggesting that the indirect effect described above dominates.

## 5.1 Unobserved heterogeneity

Both the descriptive regressions from Table 2 and the estimated utility parameters from Table 3 show that, though they are relevant, the observed state variables have limited ability to explain the migration decision. This naturally raises the question of how that decision depends on unobserved factors. Unobserved heterogeneity may mask important relationships between intergenerational altruism and migration behavior among subpopulations with unobserved preferences that make them more likely to migrate. Furthermore, omitted state variables that are correlated with those included in the model may lead to inconsistent utility parameter estimates that misstate the relationship between altruism and migration.

Empirical dynamic discrete choice models typically model unobserved heterogeneity using a finite-mixture approach in which each panel unit is assumed to belong to one of a finite set of unobserved classes that represent permanent unobserved state variables (or a discrete approximation to such variables; see Kasahara and Shimotsu, 2009, Arcidiacono and Miller, 2010, and Hu and Shum, 2012 for discussions of this approach in the dynamic discrete choice context). One drawback to this approach is that the process through which it recovers unobserved-class-specific parameters from variation in observed variables can be opaque (see, for example, the identification procedures discussed in Kasahara and Shimotsu, 2009 and Hu and Shum, 2012).[14] To provide transparent evidence on the relationship between intergenerational altruism and migration in the presence of unobserved heterogeneity, I use a limited-sample approximation that obviates the need

[14]Nonparametric identification of the model that I estimate below follows from remark 3 of Kasahara and Shimotsu (2009). A model where preferences depended on further lags of the location decision, introducing state dependence, would not be identified through their construction without observations on further generations. However, because the very existence of the black population in the Southern US was a consequence of slavery, the inclusion of a generational time trend in the model helps to account for state dependence and duration effects.

for formal finite-mixture methods to recover unobserved-class-specific parameters and migration probabilities. The idea behind this approach is that, since migration is relatively rare, families who ever leave South reveal substantial information about the unobserved state variables that may have influenced their migration decisions.

To see how this approach works, consider the simplest finite-mixture structure, in which each family belongs to one of two unobserved classes.[15] Because the generation-specific migration probabilities are small and do not vary much with observed state variables, if there is meaningful unobserved heterogeneity under this assumption, it must be the case that members of one class almost never migrate (call this class "type 2," or "stayers") while members of the other migrate with modest probability (call this class "type 1," or "movers").[16] Since stayers never leave the South, families that ever migrate must therefore be movers with probability one. Thus, movers' preferences and migration probabilities are approximately identified from the location choices of families that eventually leave the South.[17] While the estimates resulting from this limited-sample approximation are very similar to those from a formal finite-mixture model, it is much easier to see how the approximation method uses variation in observables to identify location preferences in the presence of unobserved heterogeneity (in addition, Gardner, 2017 uses Monte-Carlo simulations to show in a treatment effects context that the approximation works well even when the unobserved heterogeneity is not binary).

Table 4 presents the preference parameters and migration probabilities obtained by estimating the model among the subsample of families that ever leave the South (in the interest of precision, I use the full sample to estimate the state transitions, which is equivalent to assuming that the observed state variables transition independently of the unobserved classes). These estimates are for members of the mover class; stayers remain in the South with probability one. The estimated coefficients on fertility and education increase in absolute value relative to the full-sample estimates presented above in Table 3, showing the increased influence of these variables for movers. In addition, the coefficient on the intergenerational time trend increases substantially, consistent with dynamic selection bias arising because the families that remain long in the South are those with stronger preferences for living there, understating the changing attractiveness of the North over time. Finally, and as expected, the estimated constant term (which represents pure preferences for the North) is larger, and the estimated moving cost smaller, for movers.

[15]The assumption of binary heterogeneity is particularly appropriate for small samples such as mine, but can also be viewed as an approximation of a higher-dimensional unobserved state variable.

[16]To see this, suppose that a fraction $\pi$ of the population migrate with probability $p_1$ and a fraction $(1-\pi)$ migrate with probability $p_2 < p_1$, so that the observed probability is $p = \pi p_1 + (1-\pi)p_2$. Unless $p_1 \approx p_2 \approx p$ (in which case the unobserved heterogeneity is irrelevant), $p$ small implies that $p_2 \approx 0$ and $p_1 \approx p/\pi > p$.

[17]Though some "movers" will remain in the South by chance, they will be missing at random from the limited estimation sample without affecting the consistency of the parameter estimates. Although "stayers"' preferences are not identified under this approximation, they are also uninteresting, since members of this group never migrate.

The estimated dynamic and static migration probabilities, and the differences between them, are substantially higher for movers than those obtained using the full sample. This is unsurprising since all observed migrations are undertaken by members of this subpopulation, who account for roughly 35% of the sample. However, the larger estimated effects of intergenerational altruism for movers have an important behavioral interpretation. The rarity of observed migrations means that most Southern-born individuals are not potential migrants. Since they would remain in the South regardless of their concern for future generations, intergenerational altruism has no impact on their migration behavior. Among truly potential migrants, intergenerational altruism is much more influential. As the estimates in Table 4 show, altruism increases the probability of migration by about 16% for first- and second-generation movers, on average (both average effects are significant at the 10% level or lower).[18] The corresponding population averages, from Table 3, are 3% and 1%, respectively. In relative terms, these estimates imply that 42% of first generation migrants, and 24% of second generation ones, would not have moved North absent intergenerational altruism.

The estimates in Tables 3 and 4 provide different perspectives on the relationship between altruism and migration. The estimated population-average effects of altruism in Table 3 show how intergenerational altruism affects the total flow of migrants from the South to the North, while the larger estimated effects for movers presented in Table 4 show the influence of intergenerational altruism on the migration decisions of migrants themselves. The former effect might be of interest to an analyst trying to predict how migrant flows will respond to a policy that benefits migrants' children, while the latter might be of interest to a behavioralist trying to understand how migration decisions are made.

## 5.2 Alternative specifications

In this section, I estimate a number of variations on the primary specifications presented above in order to assess the robustness of my estimates to different modeling assumptions. As I note in Section 4.1, the value of $\lambda = .95^{25} \approx .28$ for the intergenerational discount factor that I use is on the low end of those assumed, estimated, or calibrated in other studies with intergenerational models. To determine the sensitivity of the estimated effects of intergenerational altruism to the assumed intergenerational discount factor, I use the full sample to reestimate my primary specification using a substantially larger value of $\lambda = .4$ (this is near the value estimated in Heckman and Raut, 2016, the highest in my literature review). Table 5 presents the results. The majority of the coefficient estimates are comparable to the original full-sample results reported in Table 3, with the exception of the constant term and estimated moving cost, both of which are appreciably lower on the real line

[18] These mover-specific altruism effects imply population-average effects of about $.16 \times .35 = .056$, which the model without unobserved heterogeneity approximates reasonably well for first-generation Southerners, but understates for the second generation.

when the altruism parameter is larger. The estimated effects of altruism, however, are quite similar, though slightly smaller.[19] The results in Table 6, which repeats this exercise using the limited-sample approximation for unobserved heterogeneity, are much the same, with estimated effects of intergenerational altruism that are comparable to, though slightly smaller than, those reported in Table 4. Together, the estimates in Tables 5 and 6 suggest that my estimates are insensitive to the assumed rate of intergenerational altruism, at least within a sensible range of values comparable to those used in other studies.

The purpose of the limited-sample approximation developed above is to allow for unobserved heterogeneity in a way that makes it as clear as possible how preferences over unobserved state variables are identified from variation in observed variables. To show that this approach is well founded, I also estimate an analogous formal finite-mixture model in which every member of each family belongs to an unobserved type $\tau \in \{1,2\}$, the proportion of families that belong to type 1 is $\pi$, and the flow utility functions for living in the North have unobserved-type-specific fixed effects, so that

$$u_{ng}(x_g, \tau, b_g) = \beta_{0\tau} + x_g'\beta - c \cdot 1(b_g = s), \tag{16}$$

for $\tau \in \{1,2\}$, where $x_g = (m_g, f_g, e_g)$ no longer includes a constant.[20]

The estimated preference parameters and migration probabilities for this model are presented in Table 7. The estimated parameters are very similar to those for the limited-sample approximation results in Table 4 (many of the standard errors are smaller, which is likely a consequence of the approximation method's inefficient use of only a fraction of the sample). The constant term for type 2 families is highly negative, implying that members of this type almost never leave the South (this is consistent with the intuition behind the limited-sample approximation). Furthermore, the estimated proportion $\hat{\pi}$ of the population that belongs to type 1 is about .35, which corresponds closely to the fraction of families, utilized in the limited-sample approximation, that ever leave the South. Finally, the estimated migration probabilities and effects of intergenerational altruism are very similar to those obtained under the approximation method (the migration probabilities in the table are for type 1 families; for type 2 families, both the dynamic and static migration probabilities

[19] Though it seems counterintuitive that increasing the discount factor decreases the estimated effects of altruism, this is consistent with the changes in the estimated moving cost. The higher the discount factor, the lower the migration cost that parents are willing to migrate to spare future generations, which in a nonlinear model can imply a smaller impact of altruism on migration.

[20] To estimate the model, I assume as before that the $f_{g+1|g}$ are independent of $\tau$, and maximize the sample analog of

$$E\left[\log\left(\pi \prod_{g=1}^{3} P(l_g|x_g, b_g = s; \beta_{01}, \beta, \hat{\rho}) + (1-\pi) \prod_{g=1}^{3} P(l_g|x_g, b_g = s; \beta_{02}, \beta, \hat{\rho})\right)\right],$$

where $\beta_{01}$ and $\beta_{02}$ are the type-specific constants, $\beta$ are the remaining utility function parameters (excluding a constant), and $\hat{\rho}$ are the transition function parameters. Although in principle all of the parameters could be indexed by $\tau$, the limited power of observables to explain migration makes these type-specific parameters difficult to identify, particularly for the group with lower migration probabilities.

are numerically zero).

While the models estimated so far assume that the cost of migrating is stationary while allowing the relative utility of living in the North to change over time, it is possible that changes in the cost of migrating, rather than the relative attractiveness of the North, best explain increasing migration rates over time. To examine this possibility, I also estimate the model under the assumption that pure location preferences are stationary, but the cost of migrating is a linear function of time, so that

$$u_{ng}(x_g, b_g) = x_g'\beta - (c_0 + c_1 g) \cdot 1(b_g = s), \tag{17}$$

where $x_g = (1, m_g, f_g, e_g)$ no longer includes an intergenerational time trend.[21] To focus on movers, among whom the relationship between altruism and migration is more interesting, I only estimate this specification among the subsample of families who ever migrate.

As Table 8 shows, the coefficients on the observed state variables are comparable to those for the model with a constant moving cost presented in Table 4. As expected, the estimated moving cost starts higher than in the constant-cost model, but declines steeply to account for increasing rates of migration across the generations. For members of the first generation, the estimated effects of intergenerational altruism on the probability of migrating are similar to those for the constant-cost model. For members of the second generation, the effects are much smaller, since declining migration costs reduce the benefit that second generation Southerners can bestow upon their children by migrating. Although both the cost- and time-trend specifications suggest that intergenerational altruism figures heavily into the migration decision calculus (at least for members of the first generation), a Vuong test implies that the original time-trend specification provides a better fit.[22]

## 5.3 Decomposing the effects of altruism on migration

As I note in Section 2, intergenerational altruism can influence the migration decision through two distinct channels. Most obviously, parents may be willing to invest in their children by migrating to the detriment of their own flow utility in order to spare future generations the cost of migrating. On the other hand, parents with idiosyncratic preferences for the North might avoid migrating in order to benefit future generations who are unlikely to share those preferences. This second channel is potentially important in a model where the North is an absorbing state (or more generally one in

[21] It is of course possible to include time trends in pure location preferences as well as the cost of migrating. However, since the time trends are identified purely by functional form assumptions, the resulting estimates are too imprecise to be of much use.

[22] Following the procedure in Wooldrdige (2010, Ch. 13), I implement this test by regressing the differences in maximized likelihoods $\ell_i(\hat{\theta}^c) - \ell_i(\hat{\theta}^t)$ between the cost- and time-trend models for each family $i$ on a constant to test their difference from zero. The estimated coefficient of -.015 is significant with a p-value of .016.

which the cost of return migration is high), since Northern-born children who are poorly suited for the North have no choice but to remain there.

To decompose the estimated effects of altruism into components arising through these channels, I use the estimated parameters of my primary specification (both with and without unobserved heterogeneity) to estimate the static and dynamic migration probabilities that would obtain if each generation made its location decision with the understanding that, while the North would remain an absorbing state, all future Southern-born generations could migrate North at no cost. These conditions eliminate any incentive to migrate in order to benefit future generations, while preserving the incentive to remain in the South in case future generations are poorly suited for the North. Consequently, the differences between dynamic and static migration probabilities under these conditions isolate the contribution of the second channel to the total effect of intergenerational altruism on the migration decision.

The top panel of Table 9 shows the dynamic and static migration probabilities, and the differences between them, implied by the estimated parameters of the primary specification (Table 3, estimated using the full sample) when future generations can migrate at no cost. Under these conditions, altruism decreases the probability of migrating by about 2% for members of the first and second generations, on average. Comparing these effects to their counterparts under costly migration in Table 3 implies that, for the average Southerner, the altruistic incentive to migrate outweighs the altruistic incentive not to, but only marginally. Taking this into account, the estimated population-average total effects of altruism presented previously understate the degree to which the decisions of migrants themselves were motivated by altruism. Emphasizing this point, the bottom panel of the table uses the estimated parameters of the limited-sample approximation (Table 4) to perform the same analysis for the subsample of movers. For this group, the incentives to remain in the South are much smaller, implying that altruism overwhelmingly affects migration through the first channel. What drives this result is not that movers prefer the North, but that they know their descendants are likely to as well, attenuating the incentive to remain in the South to hedge against the possibility that future generations will be poorly suited for the North.

# 6 Conclusion

Writing in 1962, Sjaastad argued that migration can be viewed as an investment worth making only if the benefits exceed the costs. A quarter-century later, Becker and Barro (1989) argued that parenthood can be viewed through much the same lens. In this paper, I combine these insights in a dynastic model of the migration decision.

My empirical model suggests that intergenerational altruism figures prominently in that decision, explaining a substantial fraction of the Northward flow of Southern-born blacks during the

African American Great Migration. It also suggests that, owing to large moving costs and strong location preferences, most families would have remained in the South regardless of their concern for future generations' welfare. Among the remainder of the population, altruism increased the probability of migrating by about 16%. Furthermore, the influence of altruism on migration operates almost exclusively by encouraging parents to migrate at their own detriment in order to improve the wellbeing of future generations—few would-be migrants were discouraged from migrating by the possibility that future generations would be poorly suited to the North.

Just as the confluence of factors of which the Great Migration was a product no longer prevail, altruism may affect contemporary foreign immigrants or internal migrants differently than it did Southern-born blacks in the first half of the 20th century. However, another appeal to Sjaastad's (1962) logic suggests that the decision faced by the Great Migrants may not be that different than the one faced today by, say, a potential immigrant from Mexico to the US—a costly and difficult long-distance move offering a commensurate, though uncertain, benefit. While my estimates of the effects of intergenerational altruism on the propensity to migrate are unlikely to apply to modern economies without modification, therefore, they may still inform contemporary behavioral and policy analysis.


# Acknowledgements

I thank Robert Miller, George-Levi Gayle, an anonymous reviewer, and the editor for their helpful comments and suggestions.


# Conflict of interest

The author has no conflict of interest to declare.

# References


Arcidiacono, P. and R. A. Miller (2010) CCP estimation of dynamic discrete choice models with unobserved heterogeneity. Econometrica 79: 1823-1867

Arcidiacono, P. and R. A. Miller (2017) Identifying dynamic discrete choice models off short panels. Working paper

Barro, R. and G. Becker (1989) Fertility choice in a model of economic growth. Econometrica 57(2) 481-501

Becker, G. and R. Barro (1988) A reformulation of the economic theory of fertility. Q J Econ 103(1): 1-25

Becker, G. and N. Tomes (1986) Human capital and the rise and fall of families. J Labor Econ 4(3): S1-S39

Berman, E. and Z. Rzakhanov (2000) Fertility, migration and altruism. NBER working paper no. 7545

Bishop, K. C. (2012) A dynamic model of location choice and hedonic valuation. Working paper

Borjas, G. (1993) The intergenerational mobility of immigrants. J Labor Econ 11(1): 113-135

Caponi, V. (2011) Intergenerational transmission of abilities and self-selection of Mexican immigrants. Int Econ Rev 52(2): 523-547

Chiswick, B. (1977) Sons of immigrants: Are they at an earnings disadvantage? Am Econ Rev 67(1): 376-380

Deutsch, J., G. Epstein, and T. Lecker (2006) Multi-generation model of immigrant earnings: Theory and application. Res Labor Econ 24(05): 217-234

Gardner, J. (2017) A simple approximation to the average effect of the treatment on the treated in panel settings with selective enrolment. Appl Econ Lett: In press

Glover, A. and J. Heathcote (2011) Intergenerational redistribution in the Great Recession. NBER working paper

Heckman, J. and L. Raut (2016) Intergenerational long term effects of preschool - Structural estimates from a discrete dynamic programming model. J Econom 191(1): 169- 175

Hu, Y. and M. Shum (2012) Nonparametric identification of dynamic models with unobserved state variables. J Econom 171: 32-44

Jackson, J. S. and M. B. Tucker (1997) Three-generation national survey of black American families, 1978-1981 [computer file]. Inter-university Consortium for Political and Social Research [distributor]

Kasahara, H. and K. Shimotsu (2009) Nonparametric identification and estimation of finite mixture models of dynamic discrete choices. Econometrica 77(1): 135-175

Kennan, J. and J. R. Walker (2011) The effect of expected income on individual migration decisions. Econometrica 79(1): 211-251

Lucas, R. E. B. and O. Stark (1985) Motivations to remit: Evidence from Botswana. J Political Econ 93(5): 901-918

Magnac, T. and D. Thesmar (2002) Identifying dynamic discrete decision processes. Econometrica 70(2): 801-816

Rust, J. (1987) Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. Econometrica 55(5): 999-1033

Shen, I.-L., F. Docquier, and H. Rapoport (2009) Remittances and inequality: A dynamic migration model. J Econ Inequal 8(2): 197-220

Sjaastad, L. A. (1962) The costs and returns of human migration. J Political Econ 70(5): 80-93

Tcha, M. (1996) Altruism and migration: Evidence from Korea and the United States. Econ Dev Cult Change 44(4): 859-878

Tcha, M. J. (1995) Altruism, household size and migration. Econ Lett 49(4): 441-5

Tolnay, S. E. (2003) The African American "Great Migration" and beyond. Annu Rev of Sociol 29(1): 209-232

Wooldridge, J. (2010) Econometric Analysis of Cross-Section and Panel Data (2nd ed.). MIT, London

Table 1: Descriptive statistics

| | Generation | N | Mean | Median | SD | Min | Max |
|---|---|---|---|---|---|---|---|
| Age | 1 | 443 | 70.48 | 70 | 9.05 | 49 | 96 |
| | 2 | 400 | 46.55 | 47 | 8.18 | 28 | 75 |
| | 3 | 256 | 25.49 | 24.5 | 5.80 | 18 | 53 |
| | All | 1099 | 51.29 | 51 | 19.42 | 18 | 96 |
| Birth year | 1 | 443 | 1909.40 | 1910 | 9.04 | 1883 | 1931 |
| | 2 | 400 | 1933.18 | 1933 | 8.08 | 1906 | 1951 |
| | 3 | 256 | 1954.19 | 1955 | 5.80 | 1927 | 1963 |
| | All | 1099 | 1928.49 | 1929 | 19.33 | 1883 | 1963 |
| Male | 1 | 443 | 0.28 | 0 | 0.45 | 0 | 1 |
| | 2 | 400 | 0.29 | 0 | 0.45 | 0 | 1 |
| | 3 | 256 | 0.43 | 0 | 0.50 | 0 | 1 |
| | All | 1099 | 0.32 | 0 | 0.47 | 0 | 1 |
| South | 1 | 443 | 0.89 | 1 | 0.32 | 0 | 1 |
| | 2 | 400 | 0.82 | 1 | 0.38 | 0 | 1 |
| | 3 | 256 | 0.88 | 1 | 0.32 | 0 | 1 |
| | All | 1099 | 0.86 | 1 | 0.35 | 0 | 1 |
| Grades | 1 | 443 | 6.81 | 7 | 3.36 | 0 | 17 |
| | 2 | 400 | 10.46 | 11 | 3.18 | 0 | 17 |
| | 3 | 256 | 12.54 | 12 | 2.15 | 5 | 17 |
| | All | 1099 | 9.47 | 10 | 3.83 | 0 | 17 |
| Kids | 1 | 443 | 6.33 | 6 | 3.97 | 0 | 23 |
| | 2 | 400 | 4.77 | 4 | 3.02 | 0 | 18 |
| | 3 | 256 | 1.11 | 1 | 1.36 | 0 | 7 |
| | All | 1099 | 4.55 | 4 | 3.76 | 0 | 23 |

Notes—Samples include all Southern-born individuals aged 18 or older. The “South” is defined by the Census Bureau as Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia and West Virginia. Southern-born individuals are categorized as having migrated to the North if they did not live in a Southern state at the time of the survey.

Table 2: Migration probabilities

| | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Intercept | $0.13^{***}$ | $0.13^{***}$ | $0.12^{***}$ |
| | (0.02) | (0.02) | (0.02) |
| Male | 0.00 | 0.03 | −0.00 |
| | (0.02) | (0.04) | (0.06) |
| Fert. | −0.04 | $-0.12^{***}$ | $-0.09^{*}$ |
| | (0.02) | (0.03) | (0.04) |
| Ed. | $0.07^{**}$ | $0.10^{**}$ | $0.09^{*}$ |
| | (0.02) | (0.03) | (0.04) |
| Male*Gen. 2 | | −0.03 | 0.10 |
| | | (0.05) | (0.09) |
| Fert.*Gen. 2 | | $0.13^{**}$ | $0.14^{**}$ |
| | | (0.04) | (0.05) |
| Ed*Gen. 2 | | −0.02 | −0.00 |
| | | (0.05) | (0.05) |
| Male*Gen. 3 | | −0.04 | −0.01 |
| | | (0.05) | (0.07) |
| Fert.*Gen. 3 | | $0.11^{*}$ | $0.12^{*}$ |
| | | (0.05) | (0.06) |
| Ed*Gen. 3 | | $-0.10^{*}$ | −0.11 |
| | | (0.05) | (0.06) |
| Male*Fert. | | | 0.01 |
| | | | (0.09) |
| Male*Ed. | | | 0.16 |
| | | | (0.10) |
| Male*Fert.*Ed | | | −0.19 |
| | | | (0.15) |
| Male*Fert.*Gen. 2 | | | −0.20 |
| | | | (0.13) |
| Male*Ed.*Gen. 2 | | | −0.24 |
| | | | (0.14) |
| Male*Fert.*Ed*Gen. 2 | | | 0.31 |
| | | | (0.21) |
| Male*Fert.*Gen. 3 | | | −0.08 |
| | | | (0.14) |
| Male*Ed.*Gen. 3 | | | −0.06 |
| | | | (0.14) |
| Male*Fert.*Ed*Gen. 3 | | | 0.17 |
| | | | (0.22) |
| $R^2$ | 0.01 | 0.03 | 0.04 |
| Num. obs. | 916 | 916 | 916 |

Notes—Dependent variable is an indicator for migrating North. Coefficients estimated by OLS. Standard errors in parentheses. Education and Fertility are indicators for having above the generation-specific median number of children and, respectively, years of education. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Table 3: Primary specification

| | Parameters | | Dynamic Gen. 1 | Dynamic Gen. 2 | Dynamic Gen. 3 | Static Gen. 1 | Static Gen. 2 | Static Gen. 3 | Difference Gen. 1 | Difference Gen. 2 | Difference Gen. 3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Constant | -0.187 | All | 0.128 | 0.132 | 0.159 | 0.102 | 0.123 | 0.159 | 0.027 | 0.009 | 0.000 |
| | (0.501) | | (0.017) | (0.015) | (0.021) | (0.039) | (0.029) | (0.021) | (0.033) | (0.034) | (0.000) |
| Male | -0.019 | Fem., Low fert., Low ed. | 0.111 | 0.125 | 0.148 | 0.096 | 0.120 | 0.148 | 0.015 | 0.005 | 0.000 |
| | (0.203) | | (0.022) | (0.021) | (0.028) | (0.032) | (0.024) | (0.028) | (0.019) | (0.020) | (0.000) |
| Kids | -0.428 | Fem., Low fert., High ed. | 0.179 | 0.197 | 0.226 | 0.151 | 0.185 | 0.226 | 0.028 | 0.012 | 0.000 |
| | (0.382) | | (0.033) | (0.028) | (0.033) | (0.047) | (0.032) | (0.033) | (0.028) | (0.031) | (0.000) |
| Educ. | 0.516 | Fem., High fert., Low ed. | 0.090 | 0.087 | 0.102 | 0.065 | 0.081 | 0.102 | 0.025 | 0.006 | 0.000 |
| | (0.187) | | (0.018) | (0.019) | (0.034) | (0.040) | (0.036) | (0.034) | (0.041) | (0.037) | (0.000) |
| Cost | 2.058 | Fem., High fert., High ed. | 0.157 | 0.150 | 0.160 | 0.104 | 0.129 | 0.160 | 0.053 | 0.021 | 0.000 |
| | (0.764) | | (0.030) | (0.028) | (0.049) | (0.061) | (0.053) | (0.049) | (0.062) | (0.061) | (0.000) |
| Time | 0.248 | Male, Low fert., Low ed. | 0.109 | 0.122 | 0.146 | 0.094 | 0.118 | 0.146 | 0.015 | 0.005 | 0.000 |
| | (0.197) | | (0.026) | (0.024) | (0.029) | (0.033) | (0.025) | (0.029) | (0.019) | (0.020) | (0.000) |
| | | Male, Low fert., High ed. | 0.176 | 0.194 | 0.222 | 0.148 | 0.182 | 0.222 | 0.028 | 0.012 | 0.000 |
| Wald (X) | 9.946 | | (0.039) | (0.036) | (0.037) | (0.049) | (0.036) | (0.037) | (0.029) | (0.032) | (0.000) |
| p | 0 | Male, High fert., Low eds. | 0.088 | 0.085 | 0.100 | 0.063 | 0.080 | 0.100 | 0.025 | 0.006 | 0.000 |
| Wald (All) | 258.477 | | (0.021) | (0.020) | (0.032) | (0.040) | (0.035) | (0.032) | (0.042) | (0.037) | (0.000) |
| p | 0 | Male, High fert., High ed. | 0.154 | 0.147 | 0.157 | 0.102 | 0.127 | 0.157 | 0.052 | 0.020 | 0.000 |
| | | | (0.036) | (0.032) | (0.048) | (0.061) | (0.053) | (0.048) | (0.064) | (0.062) | (0.000) |

Notes—Estimation sample consists of all Southern-born individuals. Wald (X) denotes the Wald statistic for the joint significance of the coefficients on the observable covariates (gender, fertility and education); Wald (All) denotes the Wald statistic for the joint significance of all parameters; p denotes the p-value for the Wald test. Standard errors reported in parentheses are based on 500 nonparametric bootstrap replications. Estimates assume an intergenerational discount factor of $\lambda = .95^{25}$. Low/high fert. and ed. denote having above the generation-specific median number of children and years of schooling, respectively. All denotes an average across the observable covariate distribution.

Table 4: Unobserved heterogeneity (limited-sample approximation)

| | | | Migration probabilities | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Dynamic | | | Static | | | Difference | | |
| | Parameters | | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 |
| Constant | 0.404 | All | 0.384 | 0.660 | 0.808 | 0.221 | 0.497 | 0.808 | 0.164 | 0.163 | 0.000 |
| | (0.802) | | (0.046) | (0.045) | (0.055) | (0.099) | (0.094) | (0.055) | (0.090) | (0.118) | (0.000) |
| Male | -0.075 | Fem., Low fert., Low ed. | 0.331 | 0.644 | 0.835 | 0.234 | 0.554 | 0.835 | 0.098 | 0.091 | 0.000 |
| | (0.338) | | (0.066) | (0.071) | (0.060) | (0.088) | (0.077) | (0.060) | (0.059) | (0.070) | (0.000) |
| Kids | -0.909 | Fem., Low fert., High ed. | 0.482 | 0.772 | 0.903 | 0.360 | 0.696 | 0.903 | 0.122 | 0.076 | 0.000 |
| | (0.623) | | (0.067) | (0.056) | (0.040) | (0.103) | (0.066) | (0.040) | (0.071) | (0.059) | (0.000) |
| Educ. | 0.612 | Fem., High fert., Low ed. | 0.324 | 0.579 | 0.670 | 0.109 | 0.333 | 0.670 | 0.215 | 0.246 | 0.000 |
| | (0.285) | | (0.066) | (0.066) | (0.118) | (0.116) | (0.149) | (0.118) | (0.124) | (0.173) | (0.000) |
| Cost | 1.591 | Fem., High fert., High ed. | 0.481 | 0.727 | 0.790 | 0.185 | 0.480 | 0.790 | 0.296 | 0.247 | 0.000 |
| | (1.126) | | (0.088) | (0.071) | (0.099) | (0.152) | (0.166) | (0.099) | (0.153) | (0.181) | (0.000) |
| Time | 1.403 | Male, Low fert., Low ed. | 0.315 | 0.627 | 0.824 | 0.221 | 0.535 | 0.824 | 0.094 | 0.092 | 0.000 |
| | (0.348) | | (0.085) | (0.091) | (0.071) | (0.097) | (0.098) | (0.071) | (0.058) | (0.070) | (0.000) |
| | | Male, Low fert., High ed. | 0.463 | 0.759 | 0.896 | 0.343 | 0.680 | 0.896 | 0.120 | 0.079 | 0.000 |
| Wald (X) | 8.023 | | (0.092) | (0.073) | (0.047) | (0.116) | (0.088) | (0.047) | (0.070) | (0.062) | (0.000) |
| p | 0 | Male, High fert., Low ed. | 0.308 | 0.561 | 0.654 | 0.102 | 0.317 | 0.654 | 0.205 | 0.244 | 0.000 |
| Wald (All) | 24.542 | | (0.080) | (0.080) | (0.128) | (0.116) | (0.154) | (0.128) | (0.121) | (0.171) | (0.000) |
| p | 0 | Male, High fert., High ed. | 0.462 | 0.712 | 0.777 | 0.174 | 0.461 | 0.777 | 0.289 | 0.251 | 0.000 |
| | | | (0.102) | (0.081) | (0.110) | (0.153) | (0.173) | (0.110) | (0.152) | (0.181) | (0.000) |

Notes—Estimation sample consists of Southern-born members of families that ever migrate from the South to the North. Wald (X) denotes the Wald statistic for the joint significance of the coefficients on the observable covariates (gender, fertility and education); Wald (All) denotes the Wald statistic for the joint significance of all parameters; p denotes the p-value for the Wald test. Standard errors reported in parentheses are based on 500 nonparametric bootstrap replications. Estimates assume an intergenerational discount factor of $\lambda = .95^{25}$. Low/high fert. and ed. denote having above the generation-specific median number of children and years of schooling, respectively. All denotes an average across the observable covariate distribution.

Table 5: Primary specification ($\lambda = .4$)

| | | | Migration probabilities | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Dynamic | | | Static | | | Difference | | |
| | Parameters | | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 |
| Constant | -0.384 | All | 0.127 | 0.128 | 0.165 | 0.107 | 0.130 | 0.165 | 0.020 | -0.001 | 0.000 |
| | (0.293) | | (0.017) | (0.015) | (0.021) | (0.041) | (0.030) | (0.021) | (0.035) | (0.034) | (0.000) |
| Male | -0.027 | Fem., Low fert., Low ed. | 0.111 | 0.122 | 0.152 | 0.099 | 0.123 | 0.152 | 0.012 | -0.001 | 0.000 |
| | (0.203) | | (0.023) | (0.021) | (0.028) | (0.033) | (0.024) | (0.028) | (0.020) | (0.019) | (0.000) |
| Kids | -0.355 | Fem., Low fert., High ed. | 0.180 | 0.194 | 0.230 | 0.155 | 0.189 | 0.230 | 0.026 | 0.005 | 0.000 |
| | (0.361) | | (0.033) | (0.028) | (0.033) | (0.048) | (0.032) | (0.033) | (0.029) | (0.031) | (0.000) |
| Educ. | 0.507 | Fem., High fert., Low ed. | 0.086 | 0.082 | 0.112 | 0.072 | 0.090 | 0.112 | 0.014 | -0.008 | 0.000 |
| | (0.182) | | (0.018) | (0.018) | (0.034) | (0.044) | (0.038) | (0.034) | (0.045) | (0.037) | (0.000) |
| Cost | 1.822 | Fem., High fert., High ed. | 0.156 | 0.144 | 0.173 | 0.114 | 0.141 | 0.173 | 0.043 | 0.003 | 0.000 |
| | (0.516) | | (0.031) | (0.027) | (0.050) | (0.065) | (0.056) | (0.050) | (0.067) | (0.063) | (0.000) |
| Time | 0.245 | Male, Low fert., Low ed. | 0.109 | 0.119 | 0.149 | 0.097 | 0.120 | 0.149 | 0.012 | -0.001 | 0.000 |
| | (0.199) | | (0.026) | (0.023) | (0.029) | (0.034) | (0.026) | (0.029) | (0.020) | (0.019) | (0.000) |
| | | Male, Low fert., High ed. | 0.176 | 0.190 | 0.225 | 0.151 | 0.185 | 0.225 | 0.025 | 0.004 | 0.000 |
| Wald (X) | 9.850 | | (0.039) | (0.035) | (0.037) | (0.049) | (0.036) | (0.037) | (0.029) | (0.031) | (0.000) |
| p | 0 | Male, High fert., Low ed. | 0.084 | 0.080 | 0.109 | 0.070 | 0.088 | 0.109 | 0.014 | -0.008 | 0.000 |
| Wald (All) | 241.403 | | (0.021) | (0.019) | (0.033) | (0.043) | (0.037) | (0.033) | (0.045) | (0.037) | (0.000) |
| p | 0 | Male, High fert., High ed. | 0.153 | 0.140 | 0.169 | 0.111 | 0.137 | 0.169 | 0.042 | 0.003 | 0.000 |
| | | | (0.037) | (0.031) | (0.049) | (0.064) | (0.055) | (0.049) | (0.068) | (0.063) | (0.000) |

Notes—Estimation sample consists of all Southern-born individuals. Wald (X) denotes the Wald statistic for the joint significance of the coefficients on the observable covariates (gender, fertility and education); Wald (All) denotes the Wald statistic for the joint significance of all parameters; p denotes the p-value for the Wald test. Standard errors reported in parentheses are based on 500 nonparametric bootstrap replications. Estimates assume an intergenerational discount factor of $\lambda = .4$. Low/high fert. and ed. denote having above the generation-specific median number of children and years of schooling, respectively. All denotes an average across the observable covariate distribution.

Table 6: Unobserved heterogeneity (limited-sample approximation, $\lambda = .4$)

| | | | Migration probabilities | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Dynamic | | | Static | | | Difference | | |
| | Parameters | | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 |
| Constant | -0.105 | All | 0.385 | 0.653 | 0.814 | 0.217 | 0.500 | 0.814 | 0.168 | 0.153 | 0.000 |
| | (0.498) | | (0.046) | (0.045) | (0.052) | (0.101) | (0.094) | (0.052) | (0.092) | (0.118) | (0.000) |
| Male | -0.086 | Fem., Low fert., Low ed. | 0.330 | 0.642 | 0.841 | 0.230 | 0.557 | 0.841 | 0.100 | 0.085 | 0.000 |
| | (0.339) | | (0.067) | (0.072) | (0.058) | (0.088) | (0.077) | (0.058) | (0.060) | (0.069) | (0.000) |
| Kids | -0.898 | Fem., Low fert., High ed. | 0.482 | 0.771 | 0.907 | 0.355 | 0.698 | 0.907 | 0.127 | 0.072 | 0.000 |
| | (0.625) | | (0.067) | (0.056) | (0.038) | (0.104) | (0.065) | (0.038) | (0.072) | (0.059) | (0.000) |
| Educ. | 0.610 | Fem., High fert., Low ed. | 0.326 | 0.566 | 0.683 | 0.109 | 0.338 | 0.683 | 0.217 | 0.228 | 0.000 |
| | (0.283) | | (0.068) | (0.066) | (0.113) | (0.118) | (0.150) | (0.113) | (0.128) | (0.172) | (0.000) |
| Cost | 1.102 | Fem., High fert., High ed. | 0.489 | 0.720 | 0.798 | 0.183 | 0.485 | 0.798 | 0.306 | 0.235 | 0.000 |
| | (0.784) | | (0.089) | (0.072) | (0.095) | (0.154) | (0.165) | (0.095) | (0.157) | (0.182) | (0.000) |
| Time | 1.435 | Male, Low fert., Low ed. | 0.312 | 0.622 | 0.829 | 0.215 | 0.536 | 0.829 | 0.096 | 0.086 | 0.000 |
| | (0.356) | | (0.085) | (0.092) | (0.070) | (0.097) | (0.098) | (0.070) | (0.058) | (0.070) | (0.000) |
| | | Male, Low fert., High ed. | 0.461 | 0.755 | 0.899 | 0.336 | 0.680 | 0.899 | 0.125 | 0.075 | 0.000 |
| Wald (X) | 8.067 | | (0.092) | (0.074) | (0.046) | (0.116) | (0.088) | (0.046) | (0.071) | (0.062) | (0.000) |
| p | 0 | Male, High fert., Low ed. | 0.307 | 0.545 | 0.664 | 0.101 | 0.320 | 0.664 | 0.207 | 0.226 | 0.000 |
| Wald (All) | 26.416 | | (0.081) | (0.079) | (0.124) | (0.116) | (0.155) | (0.124) | (0.125) | (0.170) | (0.000) |
| p | 0 | Male, High fert., High ed. | 0.467 | 0.702 | 0.784 | 0.171 | 0.464 | 0.784 | 0.297 | 0.238 | 0.000 |
| | | | (0.103) | (0.082) | (0.106) | (0.153) | (0.173) | (0.106) | (0.155) | (0.182) | (0.000) |

Notes—Estimation sample consists of Southern-born members of families that ever migrate from the South to the North. Wald (X) denotes the Wald statistic for the joint significance of the coefficients on the observable covariates (gender, fertility and education); Wald (All) denotes the Wald statistic for the joint significance of all parameters; p denotes the p-value for the Wald test. Standard errors reported in parentheses are based on 500 nonparametric bootstrap replications. Estimates assume an intergenerational discount factor of $\lambda = .4$. Low/high fert. and ed. denote having above the generation-specific median number of children and years of schooling, respectively. All denotes an average across the observable covariate distribution.

Table 7: Unobserved heterogeneity (finite-mixture maximum likelihood)

| | | | Migration probabilities (Type 1) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Dynamic | | | Static | | | Difference | | |
| | Parameters | | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 |
| Constant (Type 1) | 0.344 | All | 0.331 | 0.526 | 0.642 | 0.188 | 0.375 | 0.642 | 0.142 | 0.151 | 0.000 |
| | (1.242) | | (0.050) | (0.095) | (0.116) | (0.075) | (0.083) | (0.116) | (0.066) | (0.098) | (0.000) |
| Constant (Type 2) | -16.989 | Fem., Low fert., Low ed. | 0.274 | 0.499 | 0.671 | 0.190 | 0.409 | 0.671 | 0.083 | 0.090 | 0.000 |
| | (2.719) | | (0.065) | (0.115) | (0.141) | (0.072) | (0.099) | (0.141) | (0.043) | (0.063) | (0.000) |
| Male | -0.192 | Fem., Low fert., High ed. | 0.450 | 0.683 | 0.811 | 0.332 | 0.594 | 0.811 | 0.119 | 0.090 | 0.000 |
| | (0.303) | | (0.077) | (0.106) | (0.114) | (0.096) | (0.098) | (0.114) | (0.061) | (0.058) | (0.000) |
| Kids | -0.869 | Fem., High fert., Low ed. | 0.264 | 0.430 | 0.461 | 0.090 | 0.225 | 0.461 | 0.174 | 0.205 | 0.000 |
| | (0.556) | | (0.055) | (0.103) | (0.139) | (0.075) | (0.103) | (0.139) | (0.083) | (0.136) | (0.000) |
| Educ. | 0.748 | Fem., High fert., High ed. | 0.452 | 0.634 | 0.643 | 0.172 | 0.380 | 0.643 | 0.279 | 0.254 | 0.000 |
| | (0.289) | | (0.092) | (0.116) | (0.142) | (0.123) | (0.140) | (0.142) | (0.128) | (0.162) | (0.000) |
| Cost | 1.791 | Male, Low fert., Low ed. | 0.238 | 0.451 | 0.627 | 0.163 | 0.364 | 0.627 | 0.075 | 0.088 | 0.000 |
| | (0.934) | | (0.076) | (0.125) | (0.151) | (0.071) | (0.107) | (0.151) | (0.042) | (0.061) | (0.000) |
| Time | 1.079 | Male, Low fert., High ed. | 0.404 | 0.641 | 0.780 | 0.291 | 0.547 | 0.780 | 0.113 | 0.094 | 0.000 |
| | (0.400) | | (0.093) | (0.118) | (0.123) | (0.098) | (0.108) | (0.123) | (0.058) | (0.061) | (0.000) |
| | | Male, High fert., Low ed. | 0.229 | 0.383 | 0.414 | 0.075 | 0.193 | 0.414 | 0.153 | 0.190 | 0.000 |
| $\log\left(\frac{1-\pi}{\pi}\right)$ | 0.627 | | (0.070) | (0.110) | (0.138) | (0.067) | (0.099) | (0.138) | (0.078) | (0.126) | (0.000) |
| | (0.190) | Male, High fert., High ed. | 0.405 | 0.589 | 0.598 | 0.147 | 0.336 | 0.598 | 0.258 | 0.252 | 0.000 |
| Wald (X) | 11.075 | | (0.103) | (0.124) | (0.143) | (0.113) | (0.136) | (0.143) | (0.121) | (0.160) | (0.000) |
| p | 0 | | | | | | | | | | |
| Wald (All) | 85.271 | | | | | | | | | | |
| p | 0 | | | | | | | | | | |

Notes—Estimation sample consists of all Southern-born individuals. $\pi$ denotes the proportion that belong to type 1 (estimated to be $\hat{\pi} = [1+\exp(.627)]^{-1} \approx .348$). Wald (X) denotes the Wald statistic for the joint significance of the coefficients on the observable covariates (gender, fertility and education); Wald (All) denotes the Wald statistic for the joint significance of all parameters; p denotes the p-value for the Wald test. Standard errors reported in parentheses are based on 250 nonparametric bootstrap replications. Estimates assume an intergenerational discount factor of $\lambda = .95^{25}$. Estimated dynamic and static migration probabilities, and their differences, are effectively zero for all type 2 individuals. Low/high fert. and ed. denote having above the generation-specific median number of children and years of schooling, respectively. All denotes an average across the observable covariate distribution.

Table 8: Moving-cost trend (limited-sample heterogeneity approximation)

| | | | Migration probabilities | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Dynamic | | | Static | | | Difference | | |
| | Parameters | | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 |
| Constant | 2.004 | All | 0.446 | 0.567 | 0.816 | 0.266 | 0.541 | 0.816 | 0.180 | 0.025 | 0.000 |
| | (1.343) | | (0.041) | (0.036) | (0.055) | (0.138) | (0.103) | (0.055) | (0.130) | (0.121) | (0.000) |
| Male | -0.110 | Fem., Low fert., Low ed. | 0.374 | 0.588 | 0.830 | 0.272 | 0.574 | 0.830 | 0.103 | 0.014 | 0.000 |
| | (0.321) | | (0.070) | (0.072) | (0.067) | (0.115) | (0.077) | (0.067) | (0.080) | (0.072) | (0.000) |
| Kids | -0.638 | Fem., Low fert., High ed. | 0.519 | 0.718 | 0.896 | 0.397 | 0.704 | 0.896 | 0.122 | 0.014 | 0.000 |
| | (0.691) | | (0.064) | (0.060) | (0.047) | (0.124) | (0.063) | (0.047) | (0.092) | (0.062) | (0.000) |
| Educ. | 0.568 | Fem., High fert., Low ed. | 0.420 | 0.451 | 0.721 | 0.165 | 0.416 | 0.721 | 0.256 | 0.035 | 0.000 |
| | (0.276) | | (0.070) | (0.059) | (0.107) | (0.171) | (0.171) | (0.107) | (0.190) | (0.179) | (0.000) |
| Cost | 2.991 | Fem., High fert., High ed. | 0.577 | 0.599 | 0.820 | 0.258 | 0.557 | 0.820 | 0.319 | 0.042 | 0.000 |
| | (1.844) | | (0.085) | (0.076) | (0.087) | (0.199) | (0.172) | (0.087) | (0.216) | (0.180) | (0.000) |
| Cost×Time | -1.286 | Male, Low fert., Low ed. | 0.349 | 0.561 | 0.814 | 0.250 | 0.547 | 0.814 | 0.099 | 0.014 | 0.000 |
| | (0.449) | | (0.091) | (0.088) | (0.075) | (0.126) | (0.098) | (0.075) | (0.077) | (0.073) | (0.000) |
| | | Male, Low fert., High ed. | 0.491 | 0.695 | 0.885 | 0.371 | 0.681 | 0.885 | 0.120 | 0.014 | 0.000 |
| Wald (X) | 5.689 | | (0.091) | (0.075) | (0.052) | (0.139) | (0.084) | (0.052) | (0.089) | (0.065) | (0.000) |
| p | 0 | Male, High fert., Low ed. | 0.393 | 0.424 | 0.698 | 0.150 | 0.390 | 0.698 | 0.244 | 0.034 | 0.000 |
| Wald (All) | 19.087 | | (0.082) | (0.069) | (0.119) | (0.173) | (0.180) | (0.119) | (0.182) | (0.175) | (0.000) |
| p | 0 | Male, High fert., High ed. | 0.549 | 0.573 | 0.803 | 0.237 | 0.530 | 0.803 | 0.312 | 0.043 | 0.000 |
| | | | (0.094) | (0.083) | (0.099) | (0.202) | (0.184) | (0.099) | (0.209) | (0.181) | (0.000) |

Notes—Estimation sample consists of Southern-born members of families that ever migrate from the South to the North. Wald (X) denotes the Wald statistic for the joint significance of the coefficients on the observable covariates (gender, fertility and education); Wald (All) denotes the Wald statistic for the joint significance of all parameters; p denotes the p-value for the Wald test. Standard errors reported in parentheses are based on 500 nonparametric bootstrap replications. Estimates assume an intergenerational discount factor of $\lambda = .4$. Low/high fert. and ed. denote having above the generation-specific median number of children and years of schooling, respectively. All denotes an average across the observable covariate distribution.

Table 9: Migration probabilities with costless future migration

| | Primary specification | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Dynamic | | | Static | | | Difference | | |
| | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 |
| All | 0.082 | 0.100 | 0.159 | 0.102 | 0.123 | 0.159 | -0.020 | -0.024 | 0.000 |
| | (0.023) | (0.014) | (0.021) | (0.039) | (0.029) | (0.021) | (0.018) | (0.018) | (0.000) |
| Fem., Low fert., Low ed. | 0.083 | 0.104 | 0.148 | 0.096 | 0.120 | 0.148 | -0.012 | -0.015 | 0.000 |
| | (0.024) | (0.018) | (0.028) | (0.032) | (0.024) | (0.028) | (0.010) | (0.009) | (0.000) |
| Fem., Low fert., High ed. | 0.135 | 0.165 | 0.226 | 0.151 | 0.185 | 0.226 | -0.016 | -0.020 | 0.000 |
| | (0.036) | (0.025) | (0.033) | (0.047) | (0.032) | (0.033) | (0.012) | (0.013) | (0.000) |
| Fem., High fert., Low ed. | 0.040 | 0.054 | 0.102 | 0.065 | 0.081 | 0.102 | -0.025 | -0.028 | 0.000 |
| | (0.016) | (0.016) | (0.034) | (0.040) | (0.036) | (0.034) | (0.026) | (0.022) | (0.000) |
| Fem., High fert., High ed. | 0.069 | 0.091 | 0.160 | 0.104 | 0.129 | 0.160 | -0.035 | -0.038 | 0.000 |
| | (0.028) | (0.024) | (0.049) | (0.061) | (0.053) | (0.049) | (0.035) | (0.033) | (0.000) |
| Male, Low fert., Low ed. | 0.082 | 0.103 | 0.146 | 0.094 | 0.118 | 0.146 | -0.012 | -0.015 | 0.000 |
| | (0.025) | (0.020) | (0.029) | (0.033) | (0.025) | (0.029) | (0.009) | (0.009) | (0.000) |
| Male, Low fert., High ed. | 0.132 | 0.163 | 0.222 | 0.148 | 0.182 | 0.222 | -0.016 | -0.020 | 0.000 |
| | (0.039) | (0.029) | (0.037) | (0.049) | (0.036) | (0.037) | (0.012) | (0.013) | (0.000) |
| Male, High fert., Low ed. | 0.039 | 0.053 | 0.100 | 0.063 | 0.080 | 0.100 | -0.024 | -0.027 | 0.000 |
| | (0.016) | (0.016) | (0.032) | (0.040) | (0.035) | (0.032) | (0.025) | (0.022) | (0.000) |
| Male, High fert., High ed. | 0.068 | 0.089 | 0.157 | 0.102 | 0.127 | 0.157 | -0.034 | -0.038 | 0.000 |
| | (0.029) | (0.024) | (0.048) | (0.061) | (0.053) | (0.048) | (0.034) | (0.033) | (0.000) |
| | Unobserved heterogeneity | | | | | | | | |
| | Dynamic | | | Static | | | Difference | | |
| | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 1 | Gen. 2 | Gen. 3 |
| All | 0.212 | 0.491 | 0.808 | 0.221 | 0.497 | 0.808 | -0.009 | -0.006 | 0.000 |
| | (0.086) | (0.087) | (0.055) | (0.099) | (0.094) | (0.055) | (0.016) | (0.010) | (0.000) |
| Fem., Low fert., Low ed. | 0.228 | 0.550 | 0.835 | 0.234 | 0.554 | 0.835 | -0.006 | -0.004 | 0.000 |
| | (0.081) | (0.076) | (0.060) | (0.088) | (0.077) | (0.060) | (0.010) | (0.007) | (0.000) |
| Fem., Low fert., High ed. | 0.354 | 0.694 | 0.903 | 0.360 | 0.696 | 0.903 | -0.007 | -0.002 | 0.000 |
| | (0.097) | (0.065) | (0.040) | (0.103) | (0.066) | (0.040) | (0.008) | (0.006) | (0.000) |
| Fem., High fert., Low ed. | 0.098 | 0.324 | 0.670 | 0.109 | 0.333 | 0.670 | -0.011 | -0.010 | 0.000 |
| | (0.090) | (0.137) | (0.118) | (0.116) | (0.149) | (0.118) | (0.028) | (0.016) | (0.000) |
| Fem., High fert., High ed. | 0.170 | 0.472 | 0.790 | 0.185 | 0.480 | 0.790 | -0.015 | -0.008 | 0.000 |
| | (0.132) | (0.157) | (0.099) | (0.152) | (0.166) | (0.099) | (0.023) | (0.012) | (0.000) |
| Male, Low fert., Low ed. | 0.215 | 0.532 | 0.824 | 0.221 | 0.535 | 0.824 | -0.006 | -0.004 | 0.000 |
| | (0.091) | (0.097) | (0.071) | (0.097) | (0.098) | (0.071) | (0.009) | (0.007) | (0.000) |
| Male, Low fert., High ed. | 0.337 | 0.678 | 0.896 | 0.343 | 0.680 | 0.896 | -0.006 | -0.002 | 0.000 |
| | (0.111) | (0.087) | (0.047) | (0.116) | (0.088) | (0.047) | (0.008) | (0.006) | (0.000) |
| Male, High fert., Low ed. | 0.092 | 0.308 | 0.654 | 0.102 | 0.317 | 0.654 | -0.011 | -0.009 | 0.000 |
| | (0.091) | (0.142) | (0.128) | (0.116) | (0.154) | (0.128) | (0.028) | (0.017) | (0.000) |
| Male, High fert., High ed. | 0.160 | 0.453 | 0.777 | 0.174 | 0.461 | 0.777 | -0.014 | -0.008 | 0.000 |
| | (0.132) | (0.164) | (0.110) | (0.153) | (0.173) | (0.110) | (0.023) | (0.013) | (0.000) |

Notes—Dynamic migration probabilities calculated using conditional valuation functions in which moving is costly for the current generation but costless for all future generations; static migration probabilities calculated as above. Standard errors in parentheses based on 500 nonparametric bootstrap replications. Low/high fert. and ed. denote having above the generation-specific median number of children and years of schooling, respectively. All denotes an average across the observable covariate distribution.

# A Transition functions

Table 10: Transitions between generations 1 and 2

| | | Generation 2 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | FLL | FLH | FHL | FHH | MLL | MLH | MHL | MHH |
| Generation 1 | FLL | 0.124 | 0.148 | 0.165 | 0.063 | 0.158 | 0.138 | 0.139 | 0.065 |
| | | (0.021) | (0.021) | (0.022) | (0.016) | (0.025) | (0.028) | (0.026) | (0.018) |
| | FLH | 0.078 | 0.214 | 0.077 | 0.131 | 0.099 | 0.201 | 0.065 | 0.135 |
| | | (0.016) | (0.025) | (0.015) | (0.021) | (0.026) | (0.033) | (0.016) | (0.029) |
| | FHL | 0.132 | 0.069 | 0.260 | 0.039 | 0.172 | 0.065 | 0.223 | 0.040 |
| | | (0.024) | (0.015) | (0.026) | (0.010) | (0.031) | (0.018) | (0.033) | (0.014) |
| | FHH | 0.108 | 0.130 | 0.159 | 0.104 | 0.137 | 0.121 | 0.134 | 0.107 |
| | | (0.020) | (0.025) | (0.027) | (0.025) | (0.034) | (0.027) | (0.029) | (0.032) |
| | MLL | 0.124 | 0.148 | 0.165 | 0.063 | 0.158 | 0.138 | 0.139 | 0.065 |
| | | (0.021) | (0.021) | (0.022) | (0.016) | (0.025) | (0.028) | (0.026) | (0.018) |
| | MLH | 0.078 | 0.214 | 0.077 | 0.131 | 0.099 | 0.201 | 0.065 | 0.135 |
| | | (0.016) | (0.025) | (0.015) | (0.021) | (0.026) | (0.033) | (0.016) | (0.029) |
| | MHL | 0.132 | 0.069 | 0.260 | 0.039 | 0.172 | 0.065 | 0.223 | 0.040 |
| | | (0.024) | (0.015) | (0.026) | (0.010) | (0.031) | (0.018) | (0.033) | (0.014) |
| | MHH | 0.108 | 0.130 | 0.159 | 0.104 | 0.137 | 0.121 | 0.134 | 0.107 |
| | | (0.020) | (0.025) | (0.027) | (0.025) | (0.034) | (0.027) | (0.029) | (0.032) |

Notes—Row and column labels represent (gender, fertility, education) triples (FLL, e.g., denotes a female with fewer than the generation-specific median number of children and less than the generation-specific median years of schooling). Estimation sample includes both Southern- and Northern-born individuals. Transition probabilities estimated using a multinomial logit model that allows the intergenerational transition between education and fertility states to depend on the gender of the child in addition to the parent's education and fertility, then assumes that male and female children are equally probable (see main text for details). Standard errors in parentheses are based on 500 nonparametric bootstrap replications.

Table 11: Transitions between generations 2 and 3

| | | Generation 3 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | FLL | FLH | FHL | FHH | MLL | MLH | MHL | MHH |
| Generation 2 | FLL | 0.135 | 0.143 | 0.160 | 0.063 | 0.241 | 0.117 | 0.083 | 0.059 |
| | | (0.023) | (0.025) | (0.029) | (0.019) | (0.027) | (0.023) | (0.019) | (0.018) |
| | FLH | 0.189 | 0.225 | 0.049 | 0.037 | 0.291 | 0.158 | 0.021 | 0.030 |
| | | (0.022) | (0.024) | (0.013) | (0.011) | (0.030) | (0.029) | (0.008) | (0.014) |
| | FHL | 0.136 | 0.096 | 0.219 | 0.049 | 0.254 | 0.081 | 0.117 | 0.047 |
| | | (0.021) | (0.019) | (0.027) | (0.016) | (0.026) | (0.018) | (0.025) | (0.015) |
| | FHH | 0.219 | 0.172 | 0.076 | 0.033 | 0.325 | 0.117 | 0.033 | 0.026 |
| | | (0.027) | (0.024) | (0.019) | (0.012) | (0.031) | (0.025) | (0.013) | (0.014) |
| | MLL | 0.135 | 0.143 | 0.160 | 0.063 | 0.241 | 0.117 | 0.083 | 0.059 |
| | | (0.023) | (0.025) | (0.029) | (0.019) | (0.027) | (0.023) | (0.019) | (0.018) |
| | MLH | 0.189 | 0.225 | 0.049 | 0.037 | 0.291 | 0.158 | 0.021 | 0.030 |
| | | (0.022) | (0.024) | (0.013) | (0.011) | (0.030) | (0.029) | (0.008) | (0.014) |
| | MHL | 0.136 | 0.096 | 0.219 | 0.049 | 0.254 | 0.081 | 0.117 | 0.047 |
| | | (0.021) | (0.019) | (0.027) | (0.016) | (0.026) | (0.018) | (0.025) | (0.015) |
| | MHH | 0.219 | 0.172 | 0.076 | 0.033 | 0.325 | 0.117 | 0.033 | 0.026 |
| | | (0.027) | (0.024) | (0.019) | (0.012) | (0.031) | (0.025) | (0.013) | (0.014) |

Notes—Row and column labels represent (gender, fertility, education) triples (FLL, e.g., denotes a female with fewer than the generation-specific median number of children and less than the generation-specific median years of schooling). Estimation sample includes both Southern- and Northern-born individuals. Transition probabilities estimated using a multinomial logit model that allows the intergenerational transition between education and fertility states to depend on the gender of the child in addition to the parent's education and fertility, then assumes that male and female children are equally probable (see main text for details). Standard errors in parentheses are based on 500 nonparametric bootstrap replications.