Spatial Influences in Upward Mobility*

Garrett Anstreicher
University of Wisconsin-Madison

August 20, 2022

Abstract

This paper extends a canonical model of intergenerational human capital investment to a geographic context in order to study the role of migration in determining optimal human capital accumulation and income mobility in the United States. The main result is that migration is considerably influential in shaping the high rates of economic mobility observed among children from low-wage areas, with human capital investment behavioral responses being important to consider. Equalizing school quality across locations does more to reduce interstate inequality in income mobility than equalizing skill prices, and policies that attempt to decrease human capital flight from low-wage areas via cash transfers are unlikely to be cost-effective.

Keywords: Intergenerational mobility, migration, human capital theory.

*Email: anstreicher@wisc.edu. I am heavily indebted to John Kennan and Ananth Seshadri for their insights and patience. The editor, James Heckman, and three anonymous referees provided feedback that substantially improved the paper. Thanks also to Manuel Amador, Dean Corbae, Chao Fu, Tim Kehoe, Rasmus Lentz, Joel McMurry, Jeff Smith, Sandra Spirovska, Joanna Venator, Chris Taber, Matt Wiswall, Anson Zhou, and seminar participants at Wisconsin labor workshops, EMCON, SOLE, and ASSA for helpful comments and discussions. This research was partially supported by the Summer Research Fellowship generously provided by the University of Wisconsin-Madison. Additionally, this material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1747503. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. Support was also provided by the Graduate School and the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison with funding from the Wisconsin Alumni Research Foundation. All remaining errors are my own.

1 Introduction

How do migration and migration opportunities influence the geography of intergenerational income mobility (IIM) in the United States? Seminal research on income mobility (Chetty et al., 2014) suggests that some of the most economically mobile parts of the country are located in the Great Plains and Mountain States[1], areas that generally lack high wages or large cities[2]. This is somewhat surprising — other things equal, one may expect that being born near a strong labor market and better-paying jobs would help a poor child escape poverty later in life.

However, the literature has predominantly focused on the importance of where somebody is *from* in influencing their later-life outcomes as opposed to where or whether they *go*. The same areas that appear to feature high levels of economic mobility (see Figure 1a for a visualization) also exhibit high rates of geographic mobility, or native children migrating elsewhere later in life (Figure 1b).[3] Migration into higher-wage locations may be important in explaining the relative success of children from these rural areas. Moreover, the opportunity to migrate in the future may provide an important incentive for human capital accumulation in places where local labor market opportunities are scarce (Becker, 1994).

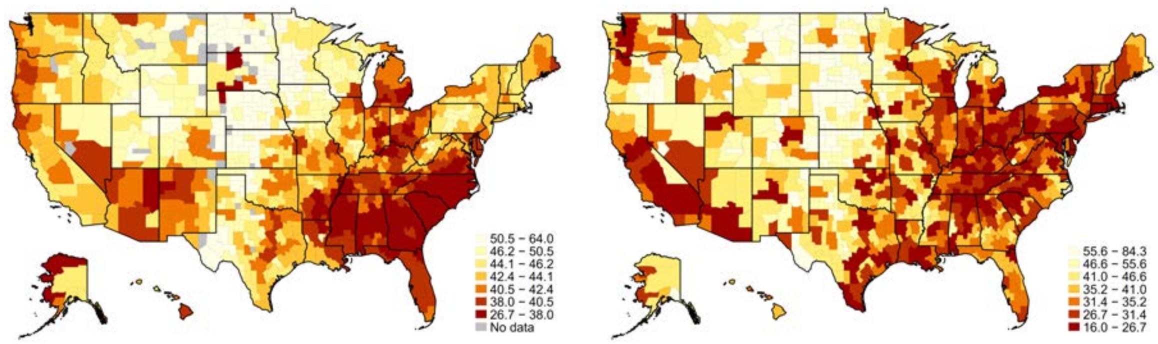
The goal of this paper is to study the role of migration and migration opportunities in influencing human capital investment decisions and income mobility in the United States. Investigating this relationship with data alone is challenging, both because of a lack of exogenous variation in people's ability to move within the U.S. and due to potential behavioral responses that would be difficult to capture empirically — that is, the option of migration in the future influencing human capital accumulation before migration decisions are actually made.

[1] Care needs to be taken when comparing locations in terms of income mobility (Mogstad et al., 2020), but the general trend of these areas enjoying an advantage in income mobility appears to be robust to uncertainty in location ranks. Additionally, whether these results reflect causal impacts of locations on outcomes or are generated by parental sorting on unobservables is a matter of ongoing debate Heckman and Landersø (2021).

[2] This relates to the inverse relationship between income inequality and income mobility observed both across and within countries (otherwise known as the "Great Gatsby Curve" (Durlauf and Seshadri, 2018; Heckman, 2013)), and is summarized also by Chetty et al. (2020): "...conditions that create greater upward mobility are not necessarily the same as those that lead to productive labor markets."

[3] Table C.5 demonstrates that this correlation is statistically significant after controlling for other factors related to IIM. For interpretation, a naive counterfactual would roughly say that reducing the typical Wyoming outflow rate of 57% to California's rate of 40% would result in the average national income percentile of a poor Wyoming native being about 2.38 points lower – this corresponds to a decrease in yearly income of roughly $1,500.

Figure 1: IIM and Native Outflow in U.S. Commuting Zones



(a) IIM (b) Native Outflow

Notes: IIM measured as the expected 2011-2012 family national income percentile of a child born in 1980-1982 to parents who were in exactly the 25th family national income percentile in 1996-2000. Native outflow rate defined as proportion of the same children who as adults live in a different CZ than when observed in 1996-2000. Commuting zone outflow rates and expected income rank for children with 25th-percentile parents taken from the Opportunity Atlas (https://opportunityinsights.org/data/).

To overcome these challenges, I construct and solve a model that follows the human capital investment, migration, and child-rearing decisions of agents over the life cycle. The model extends the classic Becker and Tomes (1979) framework of intergenerational human capital investment to a spatial context by incorporating local labor market conditions and moving opportunities. Agents are born in a home state to parents who endow them with ability and human capital investments. After childhood, the agent makes a sequence of human capital investment and moving decisions before potentially having offspring of their own. Locations differ across a variety of dimensions, including their returns to human capital, family structure, amenities, and government contributions to human capital development.

The main mechanism I capture in this framework resembles an intranational brain drain: if a given location features both low human capital returns and cheap human capital investments, natives may be motivated to heavily invest in their human capital before moving to a better labor market for human capital deployment. This enables areas with low human capital rental rates to have higher levels of IIM than high-rate locations. In counterfactuals that shut off migration in the model, I find that this channel is important in shaping adult outcomes among children from low-wage areas. As an example, I find that shutting migration off in the model results in the disparity in upward mobility between states in and out of the West North Central and Mountain Census divisions[4] shrinking by approximately half of the gap observed in the data. Failing to account for human capital investment behavioral responses, particularly those of parents investing in their children, in anticipation of future moving options would understate this result by 50 percent.

Next, I use the model to assess the importance of various factors in explaining interestate inequality in IIM. The model suggests that demographic differences across states, such as differences in racial compositions and family structure, remains the most important factor in generating cross-state disparities in child outcomes, with differences in school quality also playing a noteworthy role. However, equalizing skill prices across locations does little to nothing in reducing this inequality, consistent with the weak relationship observed between labor market productivity and upward mobility observed in the data.

While the intranational brain drain I document can be beneficial to individuals from low-wage states, many of these states have considered policies intended to reduce their outflow of talent. As an additional exercise, I consider a policy that attempts to increase a state's retention of individuals with a college degree through offering them cash transfers. I find that the offer of such payments typically does not elicit changes in migration behavior — as a result, the vast majority of these subsidies go to individuals who would have already chosen to live in the given state in the baseline world, and the policies would thus likely fail to be cost-effective. Finally, I find that equalizing public school characteristics across states does substantially more to reduce cross-state inequality in IIM than equalizing college tuition prices.

Related Literature

A vast literature exists on IIM and child human capital development (Todd and Wolpin, 2003; Cunha and Heckman, 2007; Cunha et al., 2010; Del Boca et al., 2014; Agostinelli and Wiswall, 2020), with Becker and Tomes (1979) constituting one of the first attempts to model it formally and many following papers enhancing their framework to consider issues such as borrowing constraints and policies related to education and childhood development (Abbott et al., 2019; Lee and Seshadri, 2019; Daruich, 2020; Caucutt and Lochner, 2020). However, this literature has largely ignored the role of geography, and the economic prospects of children may depend on where they are born and where/whether they move. Moreover, opportunities to migrate to different labor markets may have substantial impacts on human capital investment decisions.[5] In studying these issues, my model also contributes to the

[4] I.e. the Great Plains and Mountain States. See Appendix A for exact Census division definitions.

[5] Some empirical evidence of this can be found in the literature that studies international brain drain:

literature that studies optimal human capital development over the life cycle (Keane and Wolpin, 1997; Heckman et al., 1998; Huggett et al., 2011) through studying the role of geography in these decisions.

My paper's primary contribution comes from extending an intergenerational human capital theory model to a spatial context in order to allow the interaction of geographic and economic mobility to be studied more thoroughly. Most complementary to my paper are contemporaneous papers by Eckert and Kleineberg (2021) and Fogli and Guerrieri (2019), who develop general equilibrium models of residential and educational choice to study, respectively, the effects of school finance policy and segregation on income mobility. Human capital levels in the former paper are binary (based on college attainment), while locations are binary in the latter[6].

Relative to these papers, I allow for a combination of continuous human capital investment decisions (on top of a college decision) on the part of parents and a rich geographic structure in my model, as well as continuous human capital self-investments made on the part of agents before they have children of their own. Both of these features are meaningful: the geographic structure of my model allows my results to speak directly to actual locations in the United States, while continuous human capital allows my model to capture differences in earnings ability within educational attainment types that is likely correlated with parental socioeconomic status, along with endogenous wage growth after education decisions, which may have a spatial gradient. Moreover, continuous human capital prevents my model from constraining rich parents in how they invest in their children, since in the binary case the best they can do is pay for their child's college. To maintain tractability, however, I abstract away from general equilibrium concerns and conduct my exercises in partial equilibrium instead.

In addition to the theoretical literature, a new wave of descriptive evidence on IIM in the United States has emerged following Chetty et al. (2014) (henceforth CHKS). This work has studied numerous determinants of income mobility in the United States, such as racial disparities in IIM (Chetty and Hendren, 2018a), school quality (Rothstein, 2019), and neighborhood effects (Chetty and Hendren, 2018b; Chetty et al., 2020). However, while much has been done in this literature to demonstrate the importance of where somebody is from

Batista et al. (2012) find that increased emigration opportunities resulted in higher human capital investment in Cape Verde, and Shrestha (2017) and Spirovska (2021) find similar results in Nepal and Poland, respectively.

[6] See also Chyn and Daruich (2021) for a model with a similar structure to Fogli and Guerrieri (2019) in order to study the equilibrium effects of neighborhood-based interventions on child human capital. Bilal and Rossi-hansberg (2021) also consider a model with many locations and levels of skill but do not consider endogenous human capital accumulation or intergenerational issues.

in influencing their later-life outcomes, much less has been done in assessing the importance of later movements across labor markets. This may be in part because migration in the U.S. has been on a recent downward trend,[7] as well as because CHKS themselves appear to put the issue to rest. The authors find that their IIM estimates do not change meaningfully after limiting their sample to individuals who stay in their home CZ,[8] nor do they appear to be strongly correlated with net migration rates at the CZ level in 2004-2005.

But net migration rates in 2004-2005 say little specifically about the behavior of the individuals in the cohorts that CHKS actually use to form their IIM estimates, nor do they carry much information about whether those moving are natives leaving for the first time or are repeat movers. Limiting the sample to stayers is also insufficient to fully investigate the role of migration in forming the geography of U.S. income mobility because (as CHKS acknowledge) this sample is endogenously determined. In particular, if migration opportunities influence human capital accumulation decisions before the migration decisions actually take place, then a CZ that is highly mobile due to migration opportunities may continue to exhibit high levels of IIM even after the aforementioned sample restriction.[9] Furthermore, characteristics of a location that make migration more likely or profitable for its natives (such as higher-quality public schools) may also improve the outcomes of stayers. Another contribution of my paper comes from focusing on the impact of endogenous migration decisions made by the CHKS cohorts in adulthood on IIM in the U.S.

A similarly large literature also exists on movements across local labor markets and the migration decisions of both individuals (Kennan and Walker, 2011; Diamond, 2016; Ishimaru, 2022) and families (Mincer, 1978; Gemici, 2006; Venator, 2020). However, these papers focus on the effects of migration during adulthood on one's own earnings (or that of their spouse), not the future earnings or human capital of one's child. My paper's primary contribution to this literature comes from considering the interplay between such movements and intergenerational concerns. Individuals may move in part to provide opportunities for their future children (Bayer et al., 2007) — at the same time, the investments one's parents make in them as a child may have considerable bearing on their expected returns to migration as an adult. Overall, while the individual literatures on IIM and migration across labor

[7] Yearly interstate migration rates in the U.S. have been below 2% for much of the 21st century, a noticeable decline from the 1900s (Molloy et al., 2011; Kaplan and Schulhofer-Wohl, 2017).

[8] This restriction drops 38% of their original sample.

[9] See Mountford (1997) for a theoretical treatment of this possibility in an international context. A closely related thought experiment is to consider what would happen to IIM in the United States if those that *would* move are no longer allowed to. This is one of the key counterfactuals I evaluate in my paper, but doing so clearly requires a model.

markets in the United States are vast, attempts to synthesize the two are far less common.

The remainder of the paper is organized as follows. Section 2 introduces the model, and Sections 3 and 4 describe the data I use in model estimation along with my estimation strategy. Section 5 presents the results of counterfactual exercises, and Section 6 considers potential avenues for further research before concluding.

2 Model

While the relationship documented in Figure 1 may motivate the research question, the empirical correlation between out-migration and IIM is limited in that the role of migration in encouraging upward mobility is likely to be strongly heterogeneous across locations. Further, the data are silent on behavioral responses to migration opportunities — that is, we cannot observe a counterfactual state of the world in which people must stay where they are born to see if agent behavior and outcomes differ substantially from the status quo. I now turn to the economic model I use to study these questions.

2.1 Overview

I extend the Becker-Tomes framework to incorporate locations that differ in a variety of dimensions. The actors in the model start as children who receive human capital inputs from their parents and starting location. Children then consider how to invest in their own human capital and migrate before potentially having children of their own. Parents derive utility from their own consumption and the utility of their children and choose how much to invest in their offspring.

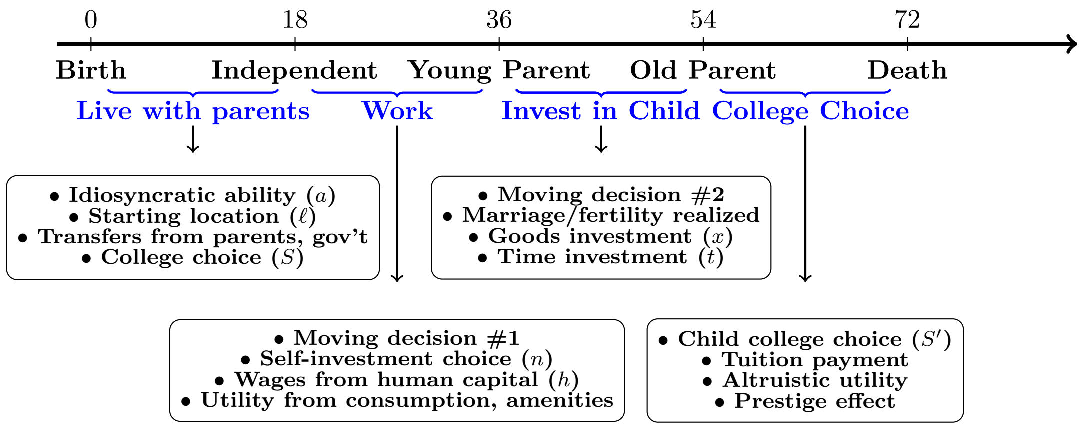
A period is 18 years, and agents live for four periods. Utility over consumption is assumed to be log.[10] The following is a description of the events that transpire and the choices that agents make in each period (see also Figure 2 for a visual representation):

1. **Period 1:** The agent as a child is endowed with an ability level and passively receives investments in their human capital from their parents and their local government. Following these investments, the parent-child pair makes a college decision.

2. **Period 2:** After emerging from childhood with a level of human capital, ability and

[10] This is something of a midpoint between typical human capital models with CRRA utility over consumption and migration models that often feature linear utility in income, e.g. Kennan and Walker (2011).

Figure 2: Model Timing



Notes: Figure presents timing of main decisions in model. See text for additional details.

schooling, the agent makes an initial moving decision before investing in their own human capital a la Ben-Porath.

3. **Period 3:** The agent has the choice to move again before observing marriage and fertility realizations based on stochastic functions of their schooling, human capital stock, and location as a young adult. If the agent becomes a parent, they balance consumption with providing expenditure and time inputs in the human capital of their child. The agent receives altruistic utility based on the expected happiness of their offspring.

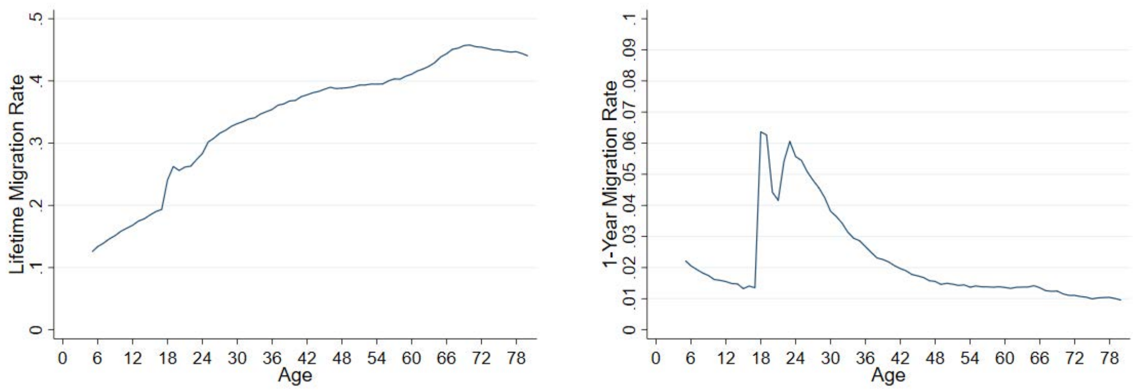
4. **Period 4:** The agent consumes the remainder of their resources (minus tuition should their child choose to go to college) and dies.

Locations (being the 50 states[11] in the U.S. and indexed by ℓ) differ in their costs of consumption/child inputs, amenities, family structure, levels of government child investment, college tuition prices, and rental rates of human capital (i.e. skill prices). The initial migration decision enables agents to move immediately after completing their desired level of schooling[12], and the second moving decision allows agents to potentially relocate to better

[11] I focus on states instead of CZs both for reasons of computational tractability and because lifetime cross-CZ migration rates are not publicly available. While state effects can account for over two-thirds of cross-CZ variation in IIM, a model that considers a more granular level of geography may be desirable.

[12] Note that the college decision here is assumed to be in-state, which for the vast majority of individuals captures the relevant college choice: while 20% of college students attend out-of-state, 94% of individuals

Figure 3: Lifetime and Yearly Interstate Migration Rates by Age



(a) Lifetime

(b) Yearly

Notes: Data from U.S. natives in the 2008-2012 American Community Survey. Lifetime migration defined as whether respondent currently resides in their state of birth. Yearly migration defined as whether respondent lives in different state than last year.

areas for raising children in anticipation of parenthood. In doing this, the model can capture agents moving in the most migratory period of the life cycle (Figures 3a and 3b show that both lifetime and yearly migration rates spike in the early 20s) while also allowing for multiple moves, which are a salient feature of the data (Kennan and Walker, 2011) and represent an additional contribution relative to Fogli and Guerrieri (2019) and Eckert and Kleineberg (2021).

Allowing for differences in marriage and fertility probabilities based on location enables the model to capture the large differences in family structure across different areas in the United States. The importance of doing so when considering income mobility is clear, both because the presence of children may detract from individual income and because the measure of IIM that CHKS report is at the family level. Imposing that these events be stochastic realizations eases the analysis greatly, but the model allows for agents to invest in their

either attend in-state or do not attend college at all. This paper is also particularly interested in lower-income children, and the corresponding statistic for children with parents in the bottom income quartile is 97%. Furthermore, while college attendance is an important driver of migration around age 18, the role it plays in lifetime migration is limited due to moves after college and in early adulthood: among individuals in their 30s living in a different state than where they lived at age 17, only 5% are college graduates living in the state where they first attended college. For all these reasons, extending the model to consider out-of-state options would be unlikely to change the main results. (author's calculations using National Longitudinal Survey of Youth 1997 geocode file). For a study that focuses more on migration and out-of-state college options, refer to Kennan (2020).

and their child's human capital with the knowledge that doing so will increase the odds of favorable realizations of marriage and fertility in the future.

2.2 Human Capital Development and Evolution

An agent's human capital stock determines their wages. At the beginning of life, agents are endowed with a level of ability that influences how effective they are at increasing their human capital. The distribution of child ability depends on the human capital of their parents, and the two are assumed to follow a joint log-normal distribution:

$$\begin{bmatrix} \log h_3 \\ \log a \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_h \\ \mu_a \end{bmatrix}, \begin{bmatrix} \sigma_h^2 & \rho_{ha}\sigma_h\sigma_a \\ \rho_{ha}\sigma_h\sigma_a & \sigma_a^2 \end{bmatrix} \right),$$

where a is the ability of the child and h_3 the human capital of the parent. Here ρ_{ha} captures the degree to which a child's ability is influenced by parent human capital and is assumed constant across states[13]. The mean and standard deviation of parent human capital will be obtained directly from observed wages in data after accounting for local human capital skill prices, leaving the parameters μ_a, σ_a, and ρ_{ha} to be estimated and allowing me to focus on the conditional distribution of a, denoted $G(a|h_3)$:

$$G(\log a|\log h_3) = N\left(\mu_a + \rho_{ha}\frac{\sigma_a}{\sigma_h}(\log h_3 - \mu_h), \sigma_a^2(1-\rho_{ha}^2) \right).$$

After being endowed with an ability level a, an agent enters period 2 with human capital formed by a Cobb-Douglas combination of time and good investments made by their parents and local government that resembles the specification used in Lee and Seshadri (2019):

$$h_2 = \xi a \left(t + \phi \frac{g^\ell}{s^\ell \cdot \exp\left(\mu_h + \frac{\sigma_h^2}{2}\right)} \right)^{\phi} \cdot \left(x + (1-\phi)g^\ell\right)^{(1-\phi)},$$

where x and t represent goods and time investments made by the parents, and g^ℓ represents real government expenditure on education in location ℓ, obtained by adjusting observed per-student expenditure by local price levels. The agent's ability multiplicatively alters the

[13] The joint distribution between parent income and child ability in the NLSY97 features quite comparable correlations across different Census regions, consistent with this assumption. I do, as detailed later, allow for geographic heterogeneity in mean ability levels.

effectiveness of the investments, and the parameter ξ is an anchor that governs the overall productivity of the process in forming adult human capital, which will be measured using wages. The parameter ϕ represents the weight of time inputs in forming child human capital and governs how parents choose to allocate total expenditure between time and good inputs when investing in the human capital of their children. While government expenditure may be spent on either good or time investments, viewing the exact ratio of this split in data is difficult. For lack of a better alternative, I follow Lee and Seshadri (2019) in assuming that public investments and parental inputs are perfect substitutes and that public investments are split between time and good investments in the same ratio as private parental inputs by imposing that proportion ϕ of public expenditures go to time inputs and $(1-\phi)$ to good inputs. Public time expenditures are additionally modified to be less effective in locations with higher normalized student-teacher ratios[14] s^ℓ and are then divided by the mean parent human capital level $\exp\left(\mu_h + \frac{\sigma_h^2}{2}\right)$ to be converted to a time measure[15].

As a young adult, human capital evolves according to a discrete-time Ben-Porath process that is standard in the empirical human capital literature [e.g. (Huggett et al., 2011; Lee and Seshadri, 2019)]:

$$h_3 = \varepsilon_2[a(h_2 n)^\kappa + h_2],\ \log \varepsilon_2 \sim N\left(-\frac{\sigma_{\varepsilon_2}^2}{2}, \sigma_{\varepsilon_2}^2\right) \equiv F(\varepsilon_2), \tag{1}$$

where $n \in [0,1]$ is the measure of self-investment that the agent commits to in period 2, κ the productivity of the Ben-Porath human capital process, and $\sigma_{\varepsilon_2}^2$ the spread of human capital shocks ε_2 agents are exposed to in early adulthood. Human capital is risky in that the agent receives a human capital shock after making their selection of n — human capital depreciation, however, is not a primary concern and so is assumed away by imposing that the mean of these shocks is unity.[16]

Finally, I assume parent human capital to evolve exogenously according to a shock ε_3:

$$h_4 = \varepsilon_3 h_3,\ \log \varepsilon_3 \sim N(\mu_{\varepsilon_{3,S}}, \sigma_{\varepsilon_{3,S}}^2) \equiv F(\varepsilon_{3,S}).$$

The mean and spread of the growth of parent human capital is allowed to vary depending on the parent's college attainment $S \in \{0,1\}$. The decision to allow exogenous evolution

[14]Specifically, student-teacher ratios across states are normalized to have a mean of 1.

[15]This follows since human capital in the model corresponds to the earnings an agent can make per one unit of time.

[16]Heckman et al. (1998) also assume away human capital depreciation.

of human capital in adulthood is made both to ease computation and because the most important determinants of human capital and inequality are realized in the early stages of the life cycle (Huggett et al., 2011). I allow for different distributions of human capital evolution shocks by period due to the length of the time periods in my model: while models with shorter time periods can draw from a single distribution of shocks for each age and estimate said distribution from the flat-point method (Heckman et al., 1998; Huggett et al., 2011; Bowlus and Robinson, 2012), 18-year periods are clearly too long for this method to be valid. The parameters of $F(\varepsilon_3)$ will be calibrated directly from the data, while κ and $\sigma^2_{\varepsilon_2}$ will be estimated via the simulated method of moments.

2.3 Recursive Formulation of Decisions

2.3.1 Period 2 — Independence:

The agent enters the second period as a newly independent adult with human capital h_2, ability a, and college attainment $S \in \{0, 1\}$. Given a location ℓ and a binary variable M indicating whether the agent has moved from their birth state[17], the agent solves a standard Ben-Porath problem with an added location decision that follows afterward:

$$V_2(a, h_2, S, \ell, M) = \max_{n \in [0, 1-\bar{T}S]} \{u(c_2) + \alpha N^\ell + \beta \mathbb{E}[v_2(a, h_2, S, \ell, M; n)]\},$$

$$s.t.\ p^\ell c_2 = e_2 = w^{\ell,S} h_2(1 - \bar{T}S - n)(1 - \tau^\ell).$$

Where n denotes the time spent producing additional human capital as opposed to working, with human capital evolving according to Equation 1. e_t denotes earnings in period t, which itself depends on $w^{\ell,S}$ — the price of human capital in location ℓ for education level S — as well as the amount of time spent investing in one's own human capital as opposed to working. State-specific taxes τ^ℓ are taken from the agent's earnings, and agents who go to college must spend four years obtaining a college degree as opposed to working, represented by $\bar{T} = 2/9$. States additionally differ in their costs of living p^ℓ, which determines how the agent's earnings map to their consumption.

In addition to consumption, agents derive utility from local amenities. To incorporate amenities into the model in an agnostic way, I assume that locations that are larger are

[17] This will be used to adjust probabilities of marriage and fertility realizations, described shortly. A richer model may store the home location in the state space instead, but this multiplies the state space by a factor of 50 instead of a factor of 2 and is computationally infeasible.

more amenity-rich: N^ℓ corresponds to the log of the population of state ℓ, normalized so that the smallest state (Wyoming) has an amenity value of zero. Given that amenities are typically measured at the city level, this specification captures that larger states — i.e. those that possess more large cities in the first place — are likely more amenity-rich than smaller states[18].

The agent optimizes their choice of n by weighing present consumption against their expected future happiness. Higher levels of n decrease their current earnings and consumption but raise their expected future human capital stock, which in turn facilitates migration and raises the likelihood of favorable realizations of marriage, fertility, and future earnings. After their selection of self-investment, the agent chooses whether and where to move:

$$v_2(a, h_2, S, \ell, M; n) = \max_{\ell'} \{\tilde{v}_2(a, h_2, S, \ell, M; n, \ell') + \zeta_{\ell'}\};$$

$$\tilde{v}_2(a, h_2, S, \ell, M; n, \ell') =$$
$$\int \left[\sum_{m,f} V_3(h_3, \ell', S, m, f; a') \mathbf{Pr}(m, f | h_3, S, \ell, M) - \Delta_2(h_3, S, \ell, \ell') \mathbb{1}\{\ell \neq \ell'\} \right] dG(a'|h_3) dF(\varepsilon_2),$$

where $\zeta_{\ell'}$ are a series of utility shocks drawn from the Type I Extreme Value distribution with location 0 and scale parameter σ_ζ. Combined with the amenity preferences, these shocks prevent my model from mechanically imposing that agents only move for pecuniary reasons, and the shocks in particular will be important in explaining moves from high-wage areas to low-wage areas observed in data. The distributional assumption also allows me to derive closed-form expressions of location choice probabilities and, conveniently, the expected value of v_2:

$$\mathbb{E}_\zeta[v_2(a, h_2, S, \ell, M; n)] = \bar{\gamma}\sigma_\zeta + \sigma_\zeta \log \left(\sum_{l'} \exp \left[\frac{1}{\sigma_\zeta} \tilde{v}_2(a, h_2, S, \ell, M; n, \ell') \right] \right),$$

where $\bar{\gamma}$ is the Euler-Mascheroni constant.

The probabilities of marriage and fertility are assumed to be stochastic functions of one's, schooling, human capital stock, and location that follow a probit process — with

[18] Among cities, those with higher college shares have been shown by Diamond (2016) to be more amenity-rich. At the state level, this relationship does not hold as well, since some locations are quite small and rural despite having high college shares, such as Vermont and New Hampshire. I also test whether the model's main predictions are sensitive to other notions of amenities in Appendix B.4 and find that they are not

$m = 1$ indicating the agent being married and $f = 1$ indicating the agent having a child in the upcoming period, I denote:

$$P(m = 1|h_3, S, \ell, M) = \Phi(\gamma_0^{\ell,S} + \gamma_1^{\ell,S} h_3 + \gamma_2^{\ell,S} h_3^2 + \gamma_3^{\ell,S} h_3^3 + \gamma_4^{\ell,S} M);$$

$$P(f = 1|h_3, \ell, m) = \Phi(\lambda^{\ell,m} + \lambda_1^{\ell,m} h_3 + \lambda_2^{\ell,m} h_3^2 + \lambda_3^{\ell,m} h_3^3),$$

where $\Phi()$ is the standard normal CDF. Marriage realizations are drawn first, which in turn influence the probability of the agent having a child. Marriage probability coefficients are computed separately for individuals based on their education level and location, and fertility probabilities are computed separately based on marital status and location. This allows the model to flexibly capture different family structures across locations while enabling agents to base their migration decisions in part on these differences. Marriage probabilities are adjusted further based on whether the agent moved from their home location, indicated by the binary variable M — this accounts for locations that may feature especially good marriage markets for movers (Compton and Pollak, 2007) while also adjusting probabilities for states such as Utah or Idaho, which are large outliers in terms of marriage rates but also feature certain cultural idiosyncrasies that one may worry apply more to natives than to movers.

Finally, while moving may allow the agent to locate in better places for human capital deployment or child-rearing, doing so at the end of period $t \in \{1, 2\}$ is associated with a utility cost $\Delta_t(h, S, \ell, \ell')$, paramaterized as:

$$\Delta_t(h, S, \ell', \ell) = \delta_t - \delta_3 h - \delta_4 C(\ell, \ell') - \delta_5 S - \delta_6 N^{\ell'}.$$

Thus, moving costs contain a fixed cost of moving that varies by period to allow the model to fit different rates of migration at different parts in the life cycle. Additionally, moving is less costly for individuals with higher human capital stocks and for those who have a college degree. Moving to a state is also more pleasant if the destination state is close by: $C(\ell, \ell')$ is a dummy function equal to 1 if states ℓ and ℓ' are either adjacent to one another or belong in the same Census division. Having moves to nearby states be less costly may be thought of as a way to account for resource costs or potential cultural attachments to certain parts of the country. I do not consider any further distance costs here, as additional resource costs required in moves to farther areas are trivial compared to earnings over an 18-year period. Finally, following Kennan and Walker (2011), I allow for larger-population states to be less costly to move to.

2.3.2 Period 3 — Investment in Children and Altruistic Utility:

In period 3, I assume that parents and children both enjoy consumption c_3, so a parent with a child enjoys an altruistic benefit from consumption, represented by θ. Denoting the unmarried state as $m = 0$, a single parent thus chooses consumption and child human capital investments in the form of expenditure and time (x and t, respectively), solving:

$$V_3(h_3, S, \ell, 0, 1, a') = \max_{x,t} \left\{ (1+\theta)u(\Lambda(c_3)) + \alpha N^\ell + \beta \left[\int \mathbb{E}[V_4^1(h_4, S, \ell, a'; h_2'(x, t, \ell))] dF(\varepsilon_3) \right] \right\}$$

$$s.t.\;\; p^\ell x + p^\ell c_3 = e_3 = w^{\ell,S} h_3 (1-t)(1-\tau^\ell),$$

where $\Lambda()$ represents the parent-child consumption equivalence scale. Following the assumptions made earlier in the model, I assume that parents cannot invest in their own human capital, but they can dedicate time inputs t for their child's human capital. Doing so, along with expenditure investments x, decreases current consumption but increases the child's future human capital h_2', which will confer an altruistic payoff to the parent in the future. If the single agent does not have a child, I assume them to inelastically supply labor before moving to the terminal period:

$$V_3(h_3, S, \ell, 0, 0; a') = u(c_3) + \alpha N^\ell + \beta \int V_4^0(h_4, S, \ell) dF(\varepsilon_3);\;\; p^\ell c_3 = e_3 = w^\ell h_3 (1-\tau^\ell).$$

A married parent differs from a single one only in that they additionally enjoy an altruistic benefit from sharing utility with their spouse: denoting the married state as $m = 1$, we have:

$$V_3(h_3, S, \ell, 1, f; a') = (1+\theta) V_3(h_3, S, \ell, 0, f; a'),$$

so married parents are assumed to have the same altruistic factor for each other as they do their children. Since being married increases individual utility monotonically, it does not affect optimal individual choices of x and t, so this specification effectively assumes that parents ignore one another's contributions when making child-rearing decisions and instead behave similar to how they would in a warm glow specification — as a result, children with married parents receive roughly twice the human capital inputs than those with single parents *ceteris paribus*. However, this specification also allows parents to adjust their behavioral margins to counterfactual scenarios that alter the expected payoffs to human capital for their children. The results of the model are robust to either discarding spousal altruism or allowing for a notion of parental coordination by modeling married parents as a single agent with double the available time. I choose this specification because the data suggest that the children of married parents indeed receive roughly twice the time inputs as their single-parent counterparts (see Table C.6b), and modeling parental coordination in labor supply and child investments can be exceptionally complicated. For a (much) more sophisticated treatment of these issues, refer to Gayle et al. (2014).

2.3.3 Period 4 — College Choice of Child, Final Consumption, and End of Life:

A childless agent in the final stage of the lifecycle simply consumes their remaining resources and expires:

$$V_4^0(h_4, S, \ell) = u(c_4) + \alpha N^\ell; \;\; p^\ell c_4 = e_4 = w^{\ell,S} h_4(1-\tau^\ell).$$

If the agent has a child, the parent-child pair make a binary college decision before the agent consumes their final resources and the child enters the young adult phase:

$$V_4^1(h_4, S, \ell, a'; h_2') = \max_{S' \in \{0,1\}} \Big\{ \tilde{V}_4^1(h_4, S, \ell, a', S') + (1+\theta)(\mathbb{E}[\tilde{V}_2(a', h_2', S', \ell)] + S'(\eta_1 + \eta_2 a' + \varepsilon_\eta)) \Big\},$$

where $\tilde{V}_4^1$ represents the parent's utility and $\tilde{V}_2$ the child's utility following the college choice. As before, the $(1+\theta)$ term represents the parent's altruistic benefit from the child's utility. The child's utility from college attendance includes a fixed non-pecuniary component η_1, similar to Lee and Seshadri (2019). I augment the child's college preferences further to include heterogeneity over ability η_2 and a preference shock $\varepsilon_\eta \sim N(0, \sigma_\eta^2)$ to prevent college attendance being deterministic based on parent and child characteristics. Immediately following their college decision, the child makes a moving decision that governs where they will start the young-adult phase of the model:

$$\tilde{V}_2(a', h_2', S', \ell) = \max_{\ell'} \Big\{ V_2(a', h_2', S', \ell', \mathbb{1}\{\ell' \neq \ell\}) + \Delta_1(h_2', S', \ell', \ell)\mathbb{1}\{\ell' \neq \ell\} + \zeta_{\ell'} \Big\},$$

where $V_2(\cdot)$, $\Delta_1(\cdot)$, and $\zeta_{\ell'}$ are the period-2 value function, moving costs, and location preference shocks discussed in Section 2.3.1. Given the distributional assumptions on $\zeta_{\ell'}$ and ε_η, the expectation of $\tilde{V}_2(\cdot)$ can be formed according to the standard Type I Extreme Value form, and the parent's expectation of $V_4^1(\cdot)$ can be computed by finding the threshold level of ε_η that governs college attendance and applying the standard normal CDF and inverse Mills ratio.

Finally, the parent's private utility is similar to the case where they have no children, except they potentially must pay tuition costs[19] and gain additional utility if both they and their child possess a college degree:

$$\tilde{V}_4^1(h_4, S, \ell, S', a') = u(c_4) + \alpha N^\ell + \eta_3 \mathbb{1}\{S = S' = 1\};$$

$$p^\ell c_4 = e_4 = w^{\ell,S} h_4(1-\tau^\ell) - S'(T^\ell - A(e_4, a')).$$

Here η_3 represents the prestige effect associated with the intergenerational transmission of college (Lee and Seshadri, 2019; Colas et al., 2021), and T^ℓ indicates the cost of tuition in location ℓ, which itself may be reduced by financial aid $A(e_4, a')$ available for low-income parents or especially high-ability children.

2.3.4 Model Solution

The altruistic payoff the parent gains from the child's expected utility in period 3 results in the problems the agents solve in the model being infinite horizon, so a single round of backward induction is insufficient in solving the model. Solving the model proceeds by guessing a value for V_2, after which a new value of V_2 may be produced via backward induction. The model is solved if the updated value of V_2 is sufficiently close to the provided guess. The distributions of the human capital shocks $F(\varepsilon_2)$, $F(\varepsilon_3)$ as well as the conditional distribution of child ability $G(a'|h_3)$ are discretized into five points according to the equal-mass approach (Kennan, 2006)[20]. Policy functions for x and t are computed via grid search, while policy functions for n are solved using Brent's method of optimization of a univariate function on a bounded interval. When solving for policy functions, I approximate value functions via linear interpolation over the human capital state.

Given the infinite horizon of the model, solving the model requires the assumption that location-specific characteristics (such as skill prices, taxes, tuition, etc.) remain stationary over time, as the infinite horizon of future values of these objects are unknown to the researcher. The simulation exercise of the paper will attempt to reproduce the outcomes of

[19]Given that the parent-child pair make the college decision together, who pays for college is not a pivotal assumption. However, having the child pay for college would require keeping their home state in the period-2 value function, which as mentioned is computationally infeasible.

[20]Some approximation of these distributions is necessary for tractability. Finer discretizations have little substantive impact on the main results.

the birth cohorts (1980-1982) studied in Chetty et al. (2014) — in reality, the economic conditions in the U.S changed between the time these cohorts were children and when they reached adulthood, a notable example being skill price shocks induced by the Great Recession. Allowing agents to respond to such shocks in some capacity may be important, so I obtain state-specific parameter values separately for the years 2000 and 2010 and solve the model using both sets of parameters. The child investment decisions of the initial parents, along with the initial schooling, moving, and self-investment decisions of the child are made according to the former set of value and policy functions, while the second move and child investment decisions of the now-adult children use the latter. To test the importance of this assumption, in Appendix B.4 I also conduct an exercise where agents only use year-2010 policy functions and obtain very similar results.

2.4 Discussion

The model presented attempts to nest a fairly straightforward model of intergenerational human capital investment into a model of migration with many locations. In multiple stages of the life cycle, agents trade off between consuming in the present and increasing their own expected future utility or that of their child. While geographic differences in factors such as family structure and school quality will mechanically inject heterogeneity in child outcomes across different states, differences in real returns to human capital introduce behavioral responses to migration opportunities that vary over space.

A state's given skill price in a vacuum does not necessarily impact human capital investment behavioral margins in a particular way, since while, for instance, a lower skill price decreases the opportunity cost of self-investing, it also results in one having to work more to achieve a given level of consumption. Thus, how skill prices affect investment behavior is sensitive to the curvature of utility over consumption and the productivity of human capital investments[21]. However, the presence of other locations with different returns to human capital has crucial implications for the expected *future* returns to human capital investment when the agent faces possible future migration. For an agent in a location with low human capital returns, the presence of future migration opportunities increases the present expected return to investment to their human capital or that of their child compared to a world in which agents are forced to stay put, consistent with evidence of brain drain in the interna-

[21]Note that while the model does not feature savings out of necessity for computational tractability, having utility satisfy the Inada conditions is needed to avoid corner solutions to the time allocation problems the agents solve.

tional literature (Batista et al., 2012; Shrestha, 2017; Spirovska, 2021). Importantly, this is true for both individuals who move and who ultimately choose to stay, which will allow the model to replicate a strong correlation between the outcomes of the overall sample and the stayer subsample across locations. However, the same is not true for an agent in a high-wage location — thus, the impact of migration opportunities on human capital investment will vary systematically across the geographic skill price gradient.

Another notable assumption is that moving costs are decreasing in human capital and education level. This enables agents in the model to invest in order to broaden migration opportunities for themselves or their children in the future, and it also allows the model to be consistent with higher rates of geographic mobility among college graduates observed in the data (Diamond, 2016; Kennan, 2020). While including human capital and schooling directly in the moving cost function achieves this, another modeling option would be to allow for different agent types that influence migration tastes, with more migration-inclined agents also being more inclined to attend school and accumulate human capital. The chosen specification, however, enables the model to capture behavioral responses to migration opportunities in an intuitive way — agents for whom migration is more rewarding will self-invest more in order to increase their probability of doing so. That higher human capital and schooling has a causal impact on migration costs is also not unreasonable: the process of human capital accumulation may make agents more open to experiencing a more diverse set of locations, and more skilled individuals may face smaller migration frictions through being better able to find jobs in other labor markets.

3 Data

I use a variety of data sources to estimate the model. I use the Panel Study of Income Dynamics (PSID) and the PSID Child Development Supplement (CDS) to obtain moments related to life-cycle earnings and child time allocations. I use the 2000 Decennial Census and the 2008-2012 waves of the American Community Survey (ACS) from Census Bureau and IPUMS (Ruggles et al., 2020) to obtain state-to-state migration flows as well as calibrate state-specific skill prices and realizations of marriage and fertility. Finally, I use the 1997 National Longitudinal Study of Youth (NLSY) to obtain moments related to college attainment over both the ability and parent income gradient.

3.1 PSID

I use the PSID 1968-2017 individual and family files to discipline the model parameters that govern earnings and earnings transitions over the life cycle. The PSID contains detailed socioeconomic information on a representative sample of American households. The sample started with 5,000 families and grew over time as children of the families left home and formed households of their own. In addition to annual hours worked and earnings, the PSID also contains information about the state in which its respondents reside. My sample restrictions largely follow Huggett et al. (2011) and Lee and Seshadri (2019). I first restrict my sample to household heads aged 18-72 and require that household heads older than 36 worked more than 520 hours and earned 1,500 1968 dollars or more and that household heads aged 18-36 worked and earned at least 260 hours and 1,000 dollars. The minimum hours restrictions for individuals older than 36 ensures that they supplied at least one quarter of full-time work hours, and the minimum earnings restriction is below the annual earnings of a full-time worker who earns the federal minimum wage. The earnings and hours requirements are relaxed for individuals aged 36 or younger to include individuals who may be working part-time while at school.

Observations that report having worked more than 5,820 hours per year are dropped, and top-coded earnings are multiplied by 1.5. Earnings are inflated to 2012 dollars using the PCE. After these restrictions, I am left with 178,839 person-year observations from 22,448 household heads. When computing moments for any 18-year age group, I require that household heads be observed in the age group for at least 6 years to keep the standard deviations of my earnings data reasonable — for the same reason, I also windsorize annual earnings at the 99th percentile. When computing these life-cycle earnings profiles, I also strip out time effects following the methodology of Huggett et al. (2011) to account for dramatic changes in the U.S. labor market between the middle of the 20th century and now. Sample weights are used when forming all moments from the PSID and the PSID CDS, described next.

3.2 CDS

To obtain information about how much time parents spend with their children, I use the PSID Child Development Supplement (CDS). In the years 1997, 2002, and 2007, the PSID collected information on time and expenditure investments in children and their outcomes for families with children aged 12 or below. The baseline sample contains information on approximately

3,500 children in 2,400 households. I refine this sample and time measurements following Del Boca et al. (2014) and Lee and Seshadri (2019). I merge information on adults in the CDS into the PSID using individual identifiers and keep only children who have at least one biological parent in the household. I use the same earnings/hours criteria for parents as listed above and exclude parent-child pairs with very small (<18 years) or large (>42) age gaps. These restrictions leave me with 4,402 observations over the three CDS waves.

The CDS contains detailed time diaries for each child that records whether or not a parent was present for a given activity. If so, the CDS also records whether the parent was actively participating in the given activity. Following Del Boca et al. (2014), such time is flagged as "active time" and is aggregated for each parent. Each child submits a diary for one weekday and one weekend day. To account for the possibility of specific weekdays or weekend days having different average levels of time use, I adjust hours so that average hours across weekdays and weekend days are equal across children of the same age. I then calculate weekly hours spent with children by multiplying weekday hours by 5 and weekend hours by 2 and summing the two.

3.3 2000 Census and ACS

While the PSID data are effective for capturing life-cycle earnings profiles in the United States, they contain too few observations to be effective in representing aggregate migration, fertility, and marriage patterns at the state level. To discipline the parameters that govern migration choices and stochastic realizations of marriage and fertility in the model, I make use of the 2008-2012 waves of the American Community Survey and limit my sample to household heads[22] born in the U.S. and aged between 36 and 54 (the age group corresponding to Period 3 in the model). I deflate earnings and limit the sample according to hours worked and earnings in an identical manner to how I handle the PSID, with the caveat that I restrict the sample to individuals who work at least 48 weeks per year due to only intervalled information on weeks worked per year being available in the ACS. These restrictions leave me with approximately 1.6 million observations that I use to compute marriage/fertility probabilities over human capital levels and location, state-level native outflow[23] rates, and state-to-state lifetime migration probabilities that are targeted when estimating my model. I also target the gap between average human capital levels of stayers and movers within

[22] As indicated by the "relationship to household head" variable in the surveys.

[23] With migration defined by whether individuals live in their state of birth, thus requiring the stipulation that individuals are born in the U.S.

educational levels observed in these data during estimation. I use a comparable sample from the 5% 2000 Census to obtain distributions of parent human capital, schooling and marriage at the state level used to form the initial condition of the model. To make this sample comparable to the parent sample used in Chetty et al. (2014), I also include authorized immigrants, and I use statistics of IIM for native-born children from the Opportunity Atlas instead of all children to further increase the consistency of my model's output with targeted moments. Sample weights are used in all calculations. These data are also used to calibrate state-level skill prices for high-school and college human capital $w^{\ell,S}$, described in depth later on.

3.4 NLSY 1997

The main source of data I use to obtain moments related to college attendance over the ability and parent income gradient is the NLSY 1997, a dataset that surveys 8,984 youths aged 12-16 as of December 31, 1996. The NLSY97 is divided into two subsamples: a nationally representative sample of 6,748 youths and an oversample of 2,236 minorities. A crucial feature of the NLSY97 is that it contains both measures of ability (an individual's ASVAB/AFQT score) and parental income and wealth.

I restrict the sample to individuals who completed a high school degree by age 20 in the data and have non-missing ASVAB scores and parent income. Late college-going and return college-going are salient features of these data (Kennan, 2020) but are not modeled explicitly in my framework — to account for these, I code maximum college attainment at up to age 29 as final completed schooling, similar to Ishimaru (2022). Individuals for whom final educational attainment is missing are dropped, resulting in a sample of 5,220 individuals. Longitudinal sample weights are used when computing all moments to target.

To assign ability levels to individuals in the NLSY97, I follow Dillon and Smith (2017) and use results from the ASVAB, a test designed for applicants to the U.S. military that most NLSY97 respondents took in 1997. The ASVAB has twelve separate component scores — I convert these scores to standard deviations within category and birth cohort (to account for individuals taking the test at different ages) before using the first principal component of the scores as my measure of ability. Consistent with Dillon and Smith (2017), I find that the first component explains roughly 60% of the total variance across the 12 sections and is strongly correlated with other test scores such as the ACT. I then compute quintiles of this measure to assign ability levels to individuals in the data that correspond to the five equal-mass levels of ability I discretize the ability distribution to when solving the model. I then compute quartiles[24] of parent income in 1997 to obtain moments of college attainment rates over both ability and parent income.

I also make use of other, more standard data sources when calibrating model parameters that warrant less commentary, as detailed in the following section.

4 Estimation

Estimation of the model proceeds in a two-step process: some parameters are taken from the preceding literature or calibrated outside the model directly from data, while the remainder of the model parameters are estimated via the simulated method of moments. More in-depth explanations may be found in the following sections.

4.1 Parameters Estimated Outside the Model

A summary of the parameters I obtain from data may be found in Table 1. The discount factor β is set to $0.96^{18} = 0.479$ to be consistent with an interest rate of 4%. The consumption equivalence scale for an adult with a child is set to $\Lambda(c) = \frac{c}{1.5}$ from the OECD standard. Cost of living levels p^ℓ are obtained from the American Chamber of Commerce Research Association's Cost of Living Index.[25] All values are normalized by the value of p^ℓ corresponding to Iowa.[26] State populations N^ℓ are taken from Census population estimates. I obtain student-teacher ratios s^ℓ (normalized by the mean value) and government expenditures on child human capital g^ℓ from public school statistics reported in the National Center for Education Statistics Common Core of Data Financial Surveys and follow Lafortune et al. (2018) in cleaning the data. State-level tuition rates T^ℓ are computed from enrollment-weighted average sticker prices of public-four year colleges for each state, and the financial

[24]This is chosen to reduce the number of very small cell sizes for certain parent income/ability combinations.

[25]The ACCRA index is a weighted average of costs of food, housing utilities, transportation, health care, and miscellaneous goods and services among different metro areas in the United States. The index is a standard measure for accounting for local costs of living, having been used for instance by both Kennan and Walker (2011) and Chetty et al. (2014). State-level indices have been published from 2016-onward by the ACCRA, and a state-level index constructed by Kennan and Walker (2011) for around 1980 is also available. Unsurprisingly, serial correlation in state-level costs of living is very strong (despite being separated by almost 40 years, the correlation of the two aforementioned sets of values is close to 0.8), so I simply take the midpoint of the two.

[26]The choice of normalizing state arises from home-state favoritism on the author's part.

Table 1: Parameters Estimated Outside the Model

Parameter		Value	Source
Discount rate	β	0.479	Literature; $0.96^{18} = 0.479$
Equivalence scale	$\Lambda(c)$	c/1.5	OECD
Costs of living	p^{ℓ}	Various	ACCRA Cost of Living Index
State populations	N^{ℓ}	Various	Census Population Estimates
Govt HC investment	g^{ℓ}	Various	NCES Financial Survey
Student-teacher ratios	s^{ℓ}	Various	NCES Financial Survey
Tuition Rates	T^{ℓ}	Various	IPEDS
Financial Aid	$A(e_4, a')$	Various	NPSAS
Taxes	τ^{ℓ}	Various	NBER TAXSIM
Parent HC Mean, Spread	μ_h, σ_h	0.902, 0.634	2000 Census
Skill prices	$w^{\ell,S}$	Various	Regressions on 2000 Census, ACS
Marriage probabilities	$\gamma_0^{\ell,S}, \gamma_1^{\ell,S}, \gamma_2^{\ell,S}, \gamma_3^{\ell,S}, \gamma_4^{\ell,S}$	Various	Probit Model in ACS
Fertility probabilities	$\lambda_0^{\ell,m}, \lambda_1^{\ell,m}, \lambda_2^{\ell,m}, \lambda_3^{\ell,m}$	Various	Probit Model in ACS
Period 3 shock means	$\mu_{\varepsilon_{3,0}}, \mu_{\varepsilon_{3,1}}$	0.02, 0.07	PSID
Period 3 shock SDs	$\sigma\varepsilon_{3,0}, \sigma\varepsilon_{3,1}$	0.24, 0.24	PSID

Notes: Table presents the values of parameters calibrated outside the model along with source material used in calibration. The leftmost columns describe the parameters and present their symbolic representation in the model. The third column presents the value of the parameter when possible, and the fourth column describes the source used to determine the parameter value.

aid schedule $A(e_4, a')$ is calibrated from published Federal Pell Grant schedules as well as the National Postsecondary Student Aid Study (NPSAS)[27], a nationally representative survey of 50,000 college students that contains detailed breakdowns of types of grant aid received by parent income, high school GPA, and type of institution attended. Taxes τ^{ℓ} are taken as the sum of state-level sales tax rates and combined average federal and state income tax rates as calculated by the NBER TAXSIM model[28]. The mean and standard deviation of parent human capital μ_h, σ_h that enter the joint distribution between parent human capital and child ability are set to 0.902 and 0.634, obtained from wage rates in the 2000 Census after correcting for local skill prices, the estimation of which I now turn to.

Skill Prices

The main simulation procedure will roughly attempt to reproduce the outcomes of the CHKS cohorts. Drawing child ability, however, requires knowledge of the underlying human capital of their parents. It is important to distinguish parental human capital from parental earnings

[27]See https://nces.ed.gov/surveys/npsas/.
[28]See http://users.nber.org/ taxsim/allyup/.

in the context of my model: for instance, one may be justifiably worried that two parents with identical earnings in a high-wage and a low-wage location still differ meaningfully in characteristics that may influence the ability and human capital of their child. Separating human capital from earnings is thus crucial in accounting for parental sorting (Heckman and Landersø, 2021) and credibly forming the initial condition of parents in my model, but doing so requires information about how the price of human capital differs across locations.

To approach this problem first consider an individual with a high school degree. Note that for any individual in the model we have that human capital is equal to total earnings divided by time spent working multiplied by the inverse of the local skill price for high school graduates, or:

$$h_3 = \frac{1}{w^{\ell,0}} \cdot \frac{e_3}{1-t}.$$

In words, the rightmost fraction $\frac{e_3}{1-t}$ is earnings over time spent working and is thus interpretable as a wage rate. This indicates that human capital levels may be inferred from observed wage rates in data if $w^{\ell,0}$ (location-specific skill prices) are known. Additionally, we have that

$$\frac{e_3}{1-t} = w^{\ell,0} h_3 \Longrightarrow \log\left(\frac{e_3}{1-t}\right) = \log(h_3) + \log(w^{\ell,0}), \tag{2}$$

so wages are log-linear in one's human capital stock and local skill price. This affords a strategy for estimating w^{ℓ} directly from data: in particular, I obtain location-specific skill prices w^{ℓ} from Mincer regressions with state dummies on the 2000 Census and 2008-2012 ACS. Year-2000 skill prices are used to form the parent initial condition, and year-2008-2012 skill prices are used to adjust child earnings when they reach the parent stage of the model. Computing skill prices for both sets of years allows the model to account for changes in returns to human capital across locations that may have transpired following events such as the Great Recession and the fracking boom.

For individuals with a college degree, note that if the human capital stock is held constant, the ratio of wages W between a college and high school graduate in a given location is equal to the ratio of that location's skill prices:

$$\frac{W^1}{W^0} = \frac{w^{\ell,1}}{w^{\ell,0}} \Longrightarrow \log\left(\frac{W^1}{W^0}\right) = \log(w^{\ell,1}) - \log(w^{\ell,0}),$$

implying that the exponential of college term in a regression on log wages in state ℓ can be multiplied by $w^{\ell,0}$ to obtain that state's college skill price $w^{\ell,1}$.

A key concern in this procedure is selective migration resulting in biased estimates of high school and especially college skill prices across states. When estimating state-specific college premia in log wages, I use the method described in Dahl (2002) to correct for selection. Further, I show that my estimates for high school skill prices are robust to tests that use only early labor market entrants or used a two-way fixed effects approach with the PSID. Moreover, I additionally demonstrate with the Dahl (2002) procedure that selection bias for high school skill prices appears to be trivial compared to college skill prices. For additional details on the procedure and these tests, refer to Appendix B.1.

Marriage and Fertility Realizations

The next step is to calibrate the parameters governing the stochastic marriage and fertility processes in the model, which I assume to be a function of one's state as a young adult, schooling, and human capital stock. With $w^{\ell,S}$ terms determined, human capital levels can be observed directly in the ACS by looking at hourly wages, which I compute by dividing total earnings by annual hours worked. Hourly wages are then adjusted by local skill prices obtained above and converted to human capital levels by being multiplied by 2,080 — in other words, by being transformed to the earnings the individual would have made had they worked 40 hours a week for 52 weeks. After having obtained human capital levels in the data, I sequentially limit my ACS sample to college graduates and non-graduates from each U.S. state aged 36-54 and run the probit model:

$$\mathbf{Pr}(m_i = 1) = \Phi(\gamma_0 + \gamma_1 h_i + \gamma_2 h_i^2 + \gamma_3 h_i^3 + \gamma_4 M_i + \varepsilon_i),$$

$$\mathbf{Pr}(f_i = 1) = \Phi(\lambda_0 + \lambda_1 h_i + \lambda_2 h_i^2 + \lambda_3 h_i^3 + \varepsilon_i),$$

where m_i and f_i are dummies for being married and having a child for individual i, and h_i is their level of human capital. Following the model, M_i is a dummy for individual i not living in their birth state. When estimating fertility probabilities I limit my ACS sample further to individuals aged 36-45 to prevent underestimating fertility from including parents whose children may have already left the household. Probability functions for fertility are estimated separately for married and single adults. The estimated probabilities for both outcomes are held constant past the level of human capital corresponding to the 99th percentile in the data to avoid Runge's phenomenon at the right tail of the human capital distribution. For visualizations of the marriage probabilities and an evaluation of how well they fit the data, refer to Appendix B.2.

Late Human Capital Shocks

Finally, $\mu_{\varepsilon_{3,S}}$ and $\sigma_{\varepsilon_{3,S}}$ are calibrated directly from data on older household heads in the PSID. With the assumption that parents do not invest in their own human capital and supply labor inelastically in the final stage of the life cycle, human capital growth in the later part of the life cycle becomes a function of only human capital shocks, or $\log h_4 - \log h_3 = \log \varepsilon_3$. Since I assume ε_3 to be iid across individuals within schooling groups, the mean and variance of the shock can be calibrated by looking at their sample analogues. In practice, I simply take the mean and variance of log hourly wage growth (adjusted for local skill prices) in the PSID from the 36-54 and 55-72 age ranges while excluding any person-year observations in which the individual is retired. Using wage rates as opposed to annual earnings circumvents the possibility of individuals tapering their work hours as they approach retirement. This results in a slightly positive estimate of $\mu_{\varepsilon_{3,S}}$ in contrast to Lee and Seshadri (2019) who instead look at annual earnings growth, but the main results of my paper are not sensitive to either specification.

4.2 Simulation

After the calibration described in the preceding section, I am left with 20 key parameters to estimate via the Simulated Method of Moments (SMM). These parameters are collected as

$$\Theta = [\theta, \rho_{ha}, \mu_a, \sigma_a, \xi, \phi, \kappa, \sigma_{\varepsilon_2}, \alpha, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \sigma_\zeta, \eta_1, \eta_2, \eta_3, \sigma_\eta].$$

The simulation procedure itself attempts to reproduce the outcomes of the same cohorts that CHKS study. I take the 2000 Census and limit my sample to individuals aged 36-54 who have at least one[29] child living in their household, after which I compute the distribution of human capital, parameterized as a log normal, and the joint distribution between human capital, education, and marital status in each state using the same methods as described above. I separate married individuals into three groups based on whether their spouse has a college degree, does not have a college degree, or does not work at all, and I nonparametrically estimate the joint distribution of household head education and spousal type by taking the relative frequencies of each type of head-spouse combination directly from the data. After

[29] The model only considers one child when evaluating the parent's decision problem. Limiting the data sample to individuals with exactly one child does not change the estimated distributions of parental human capital meaningfully for large states but does inject more noise into smaller states, which can have very small cell sizes even in the large data I use. As such, I include all parents in my baseline sample — Lee and Seshadri (2019) make a similar decision in their handling of the PSID.

conditioning on assortative matching on education, I find that underlying spousal human capital is only weakly correlated with that of the head's, so I draw human capital for working spouses independently. I conduct the same procedure when determining family income for children who reach the parent stage of the model with a spouse.

Using this as the distributions of parental characteristic for the CHKS cohorts, I then randomly draw 20,000 parents for each state, after which I draw the ability levels of their children[30] and simulate their migration, marriage, and earnings outcomes later in life. Spouses that do not work are assumed to provide a time investment into the child's human capital that I take directly from non-working parents in the PSID CDS but zero goods investments[31]. When computing moments from the simulated data, I weight by home state population sizes to ensure that the simulated data is representative of the U.S as a whole[32].

The values of the parameters in Θ are reported in Table 2, along with a description of the data moments used to discipline them (more on this in the next section). Denoting $M = [M_1, M_2, ..., M_N]$ as the vector of empirical moments I target in the simulation procedure, denote $g(\Theta)$ as the vector of percentage errors between the data moments and the simulated moments, so:

$$g(\Theta)_i = \left(\frac{M(\Theta)_i - M_i}{M_i}\right), \quad i = 1, ...N,$$

where $M(\Theta)_i$ is the ith moment simulated from the data given the parameter guess Θ. I find the point estimate $\hat{\Theta}$ by numerically solving:

$$\hat{\Theta} = \arg\min_{\Theta} g(\Theta)' W g(\Theta)$$

where $M(\Theta)$ are the simulated model moments and W is the diagonal of the variance-covariance matrix of the data moments, obtained from bootstrapping the various samples used to make the moments 1,000 times. Minimization of the objective function proceeds

[30] Ability levels are drawn according to the human capital of the household head. Drawing according to the mean of head and spouse (when available) human capital does not substantively impact the main takeaways.

[31] Note that since the spouse is unemployed, their human capital cannot be observed in the data, so applying the policy functions in the model is not an option. Recall from Section 3, though, that the data are limited to households with household heads that work a certain amount, so the human capital of the head is always observable.

[32] An alternative procedure would be to draw more parent-child pairs for more populous states. This yields similar results but considerably increases computational burdens out of needing to draw more people total to obtain a reasonable sample size for the smallest states.

Table 2: Parameters Estimated via SMM

Parameter		Value	SE	Targeted Moment
Preferences and Human Capital Technology				
Parental altruism	θ	0.566	(0.004)	Rank-rank IGE
Ability persistence	ρ_{ha}	0.552	(0.010)	Attendance by parent income
Ability mean	μ_a	-0.544	(0.006)	Life-cycle earnings means
Ability SD	σ_a	0.427	(0.004)	Life-cycle earnings SDs
Ben-Porath productivity	κ	0.381	(0.012)	Early % wage growth mean
Early HC shock SD	σ_{ε_2}	0.325	(0.004)	Early % wage growth variance
Child HC productivity	ξ	3.623	(0.008)	Young adult earnings mean
Child HC time share	ϕ	0.933	(0.002)	Time spent with children
Amenity preference	α	0.255	(0.023)	State-to-state migration flows
Moving Preferences				
Moving fixed cost, period 2	δ_1	14.258	(0.036)	High-school migration rate
Moving fixed cost, period 3	δ_2	15.650	(0.035)	Migration rate, period 2-3
Moving cost, HC component	δ_3	0.653	(0.024)	Mover-stayer HC difference
Moving cost, proximity component	δ_4	3.544	(0.037)	Share moves to nearby states
Moving cost, college component	δ_5	0.991	(0.031)	College migration rate
Moving cost, population component	δ_6	1.236	(0.041)	State in-migration rates
Location shock scale parameter	σ_ζ	1.988	(0.017)	Cross-state out-migration rate SD
College Preferences				
Attendance fixed cost	η_1	-1.805	(0.045)	Overall attendance
Attendance cost, ability component	η_2	0.841	(0.005)	Attendance by ability
Attendance cost, prestige component	η_3	0.650	(0.043)	Attendance by parent education
College shock SD	σ_η	1.096	(0.035)	Attendance by income within ability

Notes: Table reports descriptions of parameters and their symbolic representations in first two columns. Columns three and four report parameter estimates and standard errors, and column 5 describes data moments used in estimation.

via the Sbplx[33] routine. I compute standard errors by evaluating the numerical gradient[34]. of the objective function and applying the standard indirect inference formula (Gourieroux et al., 1993).

In the model estimation and simulation, I also allow for agents to differ by race (non-Hispanic White, Black, and Hispanic) and allow for mean levels of ability μ_a to vary over the four Census regions. Allowing for racial heterogeneity is important when attempting to account for high rates of economic mobility in rural states, given that these states also feature high levels of racial homogeneity. Among individuals of different races, I estimate

[33] A variant of the Subplex routine, which itself executes Nelder-Mead on a sequence of subspaces. See https://nlopt.readthedocs.io/en/latest/NLopt_Algorithm/.

[34] While minimization of the objective function relies on a gradient-free routine, the objective function appears almost completely smooth and convex in the neighborhood of the optimizer; see Figure C.1.

separate human capital productivity parameters ξ, college fixed costs η_1, and moving fixed costs $\delta_{1,2}$ to enable the model to fit racial heterogeneity in income, educational attainment, and migration patterns. The heterogeneity in child human capital productivity could also be thought to account for potential within-state racial disparities in government investment that is not explicitly modeled here. In addition to this, I allow for race to influence skill prices, marriage likelihoods, and fertility likelihoods across locations. Including separate means of ability across regions allow for some notion of peer group or neighborhood effects on child development while still allowing for within-region heterogeneity in school quality and family structure. Another option would be to allow for the correlation between parent human capital and child ability ρ_{ha} to vary across locations, but in the NLSY97 I find similar correlations between parent income and child AFQT scores across regions along with level differences across regions that are more consistent with mean shifts. Refer to Appendix B.3 for parameter estimates and the model's fit for these categories.

4.3 Identification

While the model is jointly identified in general, a conceptual argument for identification is as follows. The altruism parameter θ is tied to moments to do with intergenerational persistence in income and is thus estimated by targeting the rank-rank[35] intergenerational elasticity of family earnings of 0.341 as reported by CHKS. A higher level of persistence of learning abilities ρ_{ha} results in a larger proportion of high-ability children being born to richer parents. Because non-pecuniary costs of college attendance are decreasing in ability, this results in a sharper pattern of college-going over the parent income distribution. Thus, ρ_{ha} is chosen to match rates of college-going by parent income quartile as taken from the NLSY97.

The fixed utility cost, ability component, prestige component of college attendance, and college preference shock spread $\eta_1, \eta_2, \eta_3, \sigma_\eta$ are targeted to match overall college-going as well as college-going by ability and by parent educational attainment. I also include the complete set of attendance probabilities by parent income quartile, ability level, and parent educational attainment (40 moments total) to bolster the identification argument: since ability and parent income co-move in both the data and the model, targeting college attendance by ability *within* parent income quartiles allows for better identification of η_2, and differences in college attendance by parent educational attainment within given ability/income combi-

[35]The choice of income mobility measure to target follows Lee and Seshadri (2019). I assess the sensitivity of model results to other measures of income mobility in the Section 5.

nations identifies η_3. Since the concavity of utility over consumption results in richer parents being more willing to fund their child's education than poorer parents, and σ_η governs the magnitude of preference shocks relative to utility from consumption, I target σ_η to match the growth of college attendance over parent income *within* levels of ability and parental education[36]. I give these moments less weight when estimating the model to prevent the number of them from dominating the objective function. Additionally, since attendance for some of the cells is close to zero, I use the absolute error instead of percentage error to avoid low-attendance cells from being given undue weight.

Estimation of μ_a and σ_a starts with the observation that earnings means and standard deviations at any stage of the life cycle increase monotonically with higher μ_a and σ_a. Thus, I target the mean and standard deviation of normalized individual earnings[37] by education level in the PSID for the age ranges corresponding to Period 2 and Period 3 in the model to estimate the two parameters. Meanwhile, the parameters κ and σ_{ε_2} primarily govern *growth rates* of earnings as the agent transitions from Period 2 to Period 3. Thus, I target the mean and standard deviation of individual earnings growth rates between the same age ranges to estimate κ and σ_{ε_2} respectively. Period 2 earnings moments also assist in estimating the child investment productivity parameter ξ because higher values of a also result in faster earnings growth rates on average, thus restricting the values that μ_a can take.

The parameter ϕ governs how important time inputs are in forming a child's human capital, so a natural moment to target is the amount of time parents spend with their children. I obtain this moment from the PSID CDS sample described in Section 3. I compute the average amount of active time a child's parent(s) spend with them out of 168 hours in a week, resulting in a target of 0.18.[38]

The final group of parameters α, $\delta_1, \delta_2, \delta_3$, δ_4, δ_5, δ_6 and σ_ζ govern migration in the model. Moments to discipline these parameters are drawn from the ACS and the PSID. From the ACS, I obtain rates of native outflow and migrant inflow at the state level, overall migration rates for high school and college graduates, and gaps in human capital between high school and college movers and non-movers. The overall migration rate of high school graduates is targeted to estimate the period-2 fixed moving cost δ_1, and the college moving

[36]Note, though, that even if attendance by parent income/ability combinations are matched, fitting the attendance profile over the parent income gradient *alone* will still be contingent on having the joint distribution of parent income and child ability correct, which is governed exclusively by ρ_{ha}.

[37]All monetary units in the model are normalized by mean real individual PSID earnings of 47,961 2012 dollars.

[38]Lee and Seshadri (2019) target a value of 0.11, but their target is the average time an individual parent spends with their child and so does not distinguish between married and single parents.

cost component δ_5 is targeted to match the migration rate of college graduates, and the population component δ_6 is targeted to match rates of state in-migration. The human capital component of moving costs δ_3 is targeted to match the aforementioned human capital gaps between movers and stayers, and the proximity component of moving costs δ_4 is chosen to match the share of moves that are made to nearby states.

The period-3 fixed moving cost δ_2 governs the rate at which individuals move between period 2 and period 3. Since only current location and birth location are available in the ACS, I cannot use it to obtain information about moves made in early adulthood. Instead, I rely on the PSID to obtain moments that identify this parameter. I compute modal locations lived in for the age ranges 18-36 and 36-54 for PSID respondents and use these locations to compute rates of migration between period 2 and period 3 of the model, arriving at a rate of about 16 percent. The parameter α governs agent preferences for higher-amenity (larger) states and is chosen to maximize the correlation of state-level outflow and inflow rates between the data and the model. Finally, σ_ζ is targeted to match the cross-state standard deviation out-migration rates of 0.093, since smaller (larger) spreads of utility shocks will result in sharper (duller) out-migration patterns of individuals leaving low-wage states and staying in high-wage states. Since this parameter governs how agents in the model weigh consumption relative to idiosyncratic location preferences, I also verify in the counterfactuals section that the model predicts reasonable elasticities of state-level population growth to wage shocks compared to those estimated by Kennan and Walker (2011). To reduce risks of overfitting, I also include in the objective function the cross-state mean of out-migration rates and the overall correlation in state-level IIM rates between the data and the model.

4.4 Model Fit

Having estimated the model, I now evaluate its performance. Estimated parameter values can be found in Table 2, while Table 3 displays my model's performance in fitting its targeted moments. The parameters that govern parental altruism, ability inheritance, and human capital development are all well within the ranges of estimated values in other papers that use similar technologies — in particular, the value of ϕ is quite close to the values of time shares that Lee and Seshadri (2019) estimate from the PSID CDS. The model fits lifecycle earnings profiles for individuals with a high school degree well but overstates earnings growth for college graduates. The model also replicates a college attendance profile that increases over the parent income and ability gradients but does overstate attendance among the lowest-ability children and understates attendance among children with college-educated parents. The model fits the rank-rank IGE of earnings estimated by CHKS exactly — moreover, the model predicts a log-log IGE of 0.439, which is quite similar to CHKS's estimate of 0.413 when recoding cases of zero income to $1,000[39], indicating that the estimation results would not be particularly sensitive to the measure of intergenerational mobility used.

The moments regarding migration rates and gaps in human capital measured between stayers and movers are matched well. The estimates of parameters governing moving costs suggest that moving to a nearby state is about one quarter less costly than moving to a non-nearby state. However, the lack of an explicit preference for the individual's home state means that the model struggles to fit the (still quite low) rates of return migration observed in the data. Notably, the fixed costs for the two moves are similar: the model does not need a considerably higher fixed cost to rationalize lower rates of migration later in the life cycle since returns to migration when one is older are also lower. The estimates of the spread of location utility shocks and moving fixed costs are quite high — indeed, taking the estimate of δ_1 at face value would imply an implausibly high moving cost. However, given that the model abstracts away from things such as home attachment, location match quality, and search frictions, these numbers should be interpreted less as estimates of the cost of moving itself and more an indication that moving frictions and idiosyncratic preferences play an important role in governing state-to-state flows. Moreover, the amount that individuals pay to move after accounting for utility shocks and reductions in moving costs from college attainment, human capital, and destination state characteristics is substantially lower than the fixed cost alone.

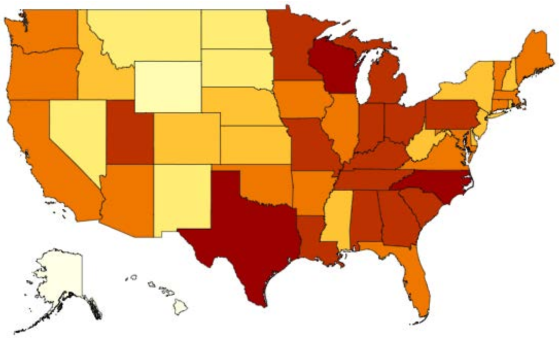
[39]This is most appropriate for my setting since the model does not feature unemployment, and so nobody has zero income. Minimum earnings among agents in my model are around $2,000.

Table 3: Model Fit

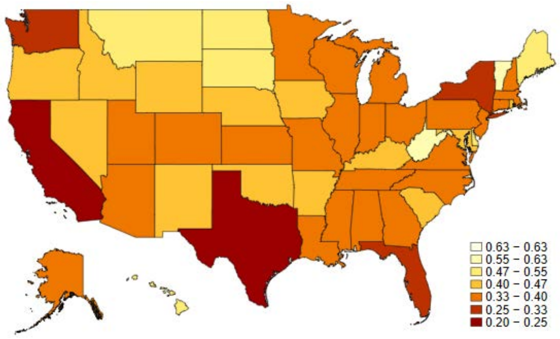
Moment	Data	Model	Source
Earnings and Human Capital			
Rank-rank IGE of family earnings	0.341	0.341	CHKS
Period 2 earnings mean, HS	0.687	0.669	PSID
Period 2 earnings SD, HS	0.336	0.268	PSID
Period 3 earnings mean, HS	0.954	0.909	PSID
Period 3 earnings SD, HS	0.495	0.513	PSID
Period 2-3 earnings % growth mean, HS	0.229	0.241	PSID
Period 2-3 earnings % growth SD, HS	0.357	0.346	PSID
Period 2 earnings mean, College	0.890	1.059	PSID
Period 2 earnings SD, College	0.406	0.380	PSID
Period 3 earnings mean, College	1.542	1.931	PSID
Period 3 earnings SD, College	0.736	0.998	PSID
Period 2-3 earnings % growth mean, College	0.408	0.541	PSID
Period 2-3 earnings % growth SD, College	0.358	0.380	PSID
Time spent with children	0.179	0.191	PSID CDS
Migration			
Overall Migration Rate, HS	0.340	0.340	ACS
Overall Migration Rate, College	0.492	0.485	ACS
Mover-stayer HC gap, HS	0.043	0.043	ACS
Mover-stayer HC gap, College	0.090	0.089	ACS
Share moves to nearby states	0.409	0.408	ACS
Period 2-3 Migration Rate	0.166	0.175	PSID
Period 2-3 Return Migration Rate	0.026	0.002	PSID
College Attendance			
Overall	0.344	0.338	NLSY97
Parent Income Quartile 1	0.162	0.112	NLSY97
Parent Income Quartile 2	0.265	0.290	NLSY97
Parent Income Quartile 3	0.402	0.413	NLSY97
Parent Income Quartile 4	0.548	0.528	NLSY97
Ability Level 1	0.069	0.173	NLSY97
Ability Level 2	0.182	0.250	NLSY97
Ability Level 3	0.300	0.328	NLSY97
Ability Level 4	0.473	0.425	NLSY97
Ability Level 5	0.648	0.589	NLSY97
Parents w/o College Degree	0.258	0.258	NLSY97
Parents w/ College Degree	0.636	0.507	NLSY97

Notes: Table presents the model fit by comparing moments obtained from data to moments simulated from the model. Column 1 describes the moment targeted, and columns 2 and 3 show data and model moment values. Column 4 documents the source of the moment. CHKS: Chetty et al. (2014). PSID: Panel Study of Income Dynamics. CDS: Child Development Supplement. ACS: American Community Survey. NLSY97: National Longitudinal Survey of Youth 1997. See text for details on sample construction.

Figure 4: Model Fit — State Outflow and Inflow Rates

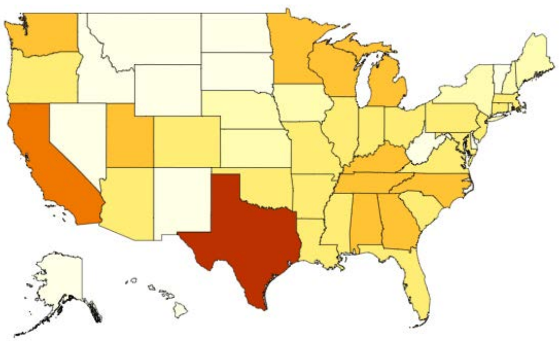


(a) Outflow (HS), Data

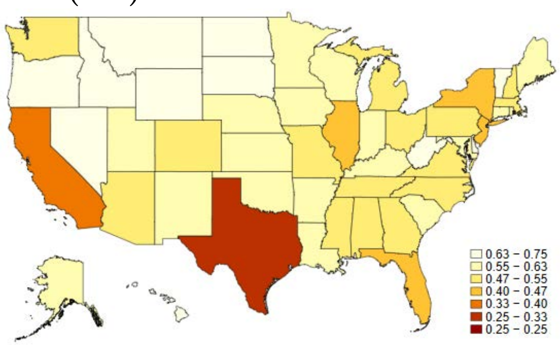


(b) Outflow (HS), Model

Outflow Correlation (HS): 0.40

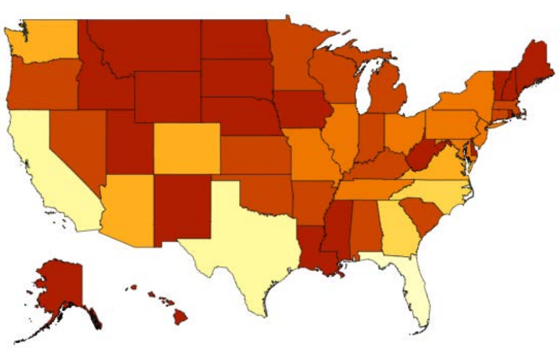


(c) Outflow (College), Data

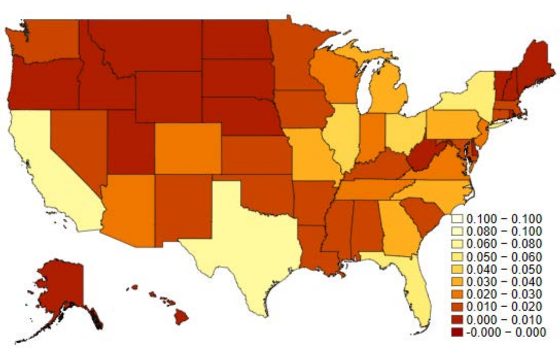


(d) Outflow (College), Model

Outflow Correlation (College): 0.67



(e) Inflow, Data

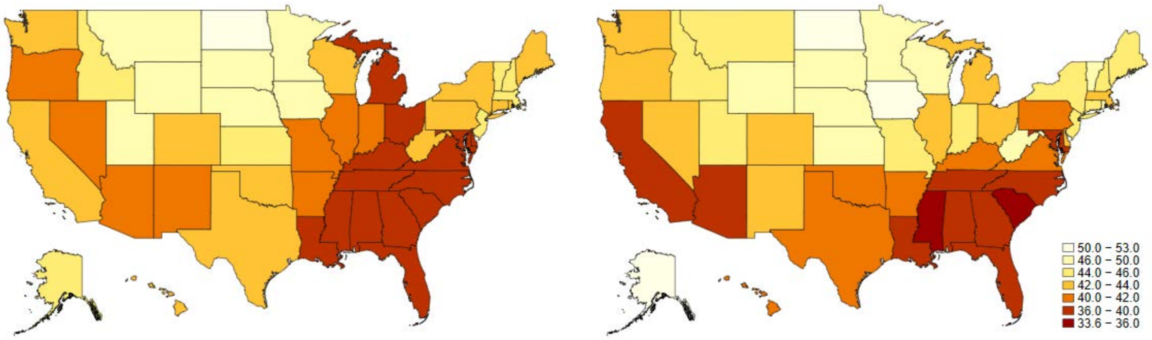


(f) Inflow, Model

Inflow Correlation: 0.89

Notes: Figures present rates of out- and in-migration as measured in the data and simulated in the model.

Figure 5: Model Fit — Upward Mobility by State



(a) Data (b) Model

Correlation: 0.81

Notes: Upward mobility measured as the expected family national income percentile of children born to parents in the 25th national income percentile in data and expected family income percentile of children born to below-median income parents in model.

While the average migration rates are matched almost exactly by the model, Figure 4 evaluates the model's performance in matching individual state outflow rates and migration destination probabilities. Generally speaking, the model does well, particularly for college graduates — the college graduate outflow rates at the state level predicted by the model and observed in the data are significantly correlated (coefficient 0.67, indicating that the model can account for nearly half the variation in cross-state college graduate migration patterns). Consistent with the data, the model predicts the Midwest and Mountain States to be highly migratory regions, with less migration being observed out of states in the Rust Belt and the Southeast. These qualitative patterns hold for the migration patterns of high school graduates as well, but here the model overpredicts migration out of the Rust Belt and the Southeast and understates migration out of the most populous states, such as California, Illinois, and New York. Another salient miss for high school graduates is that the model underpredicts migration out of the some of the more populous states, such as California, Illinois, and New York. In Appendix B.4 I assess the how the model's performance changes with alternate notions of amenities, finding comparable headline results across numerous specifications. The model does better in predicting rates at which states receive migrants, fitting this aspect of migration almost exactly.

I next evaluate how well my model reproduces the geography of intergenerational mobility in the United States as reported by CHKS. For each U.S. state, I take the average family income rank of children born to below-median income parents[40] to compute measures of absolute mobility in the model. Figure 5 juxtaposes the state-level IIM measures that CHKS find with the ones that my model predicts. The model's performance in replicating the geographic variation of upward mobility is respectable — the correlation between my estimates and those of CHKS is 0.81, indicating that my model can account for approximately 65% of the state-level variation in income mobility observed in the data. The model fits states in the Southeast and Midwest well but does slightly overpredict mobility in the Northeast and the Rust Belt. Additionally, the model underpredicts income mobility in some parts of the West Coast as well as the Southwest[41].

As an additional test, I evaluate whether the model can replicate the strong correlation in outcomes observed between the overall sample and the subsample of individuals who stay in their home state. If the model overstates selection into migration and the importance of migration for driving earnings growth for natives of low-wage states, the correlation between overall state-level rates of IIM and state-level rates among stayers should be low. However, Table 4 demonstrates that the model can reproduce this correlation quite well[42]. The model predicts a sensible difference in income mobility between the whole sample and for stayers and can replicate a strong correlation in outcomes between all individuals and stayers at the state level. As will be discussed in the next section, behavioral responses to migration opportunities play a key role in driving this result: since migration opportunities spur human capital investment for all individuals — not just those who eventually move — the model can rationalize favorable outcomes even for individuals that choose to remain in relatively low-wage states. Note additionally that this correlation is not explicitly targeted in the model estimation.

Table C.6 reports the model's fit for additional data moments. Table C.6a indicates

[40] As discussed in Chetty et al. (2014), this is approximately equal to the expected rank of a child with 25th-income-percentile parents. The measure of IIM obtained from the model is virtually identical if I average the expected outcomes of children from each parent income percentile in each state, indicating that the results are not driven by different income distributions of below-median parents across different locations.

[41] Part of this may come from the model not fully capturing differences in college attainment for natives of these states vs. elsewhere; see Figures C.2a and C.2b. Another consideration is heterogeneity in parent altruism: if parents in locations with low wages but good opportunities for child human capital development choose to live there because they are particularly altruistic toward their children, these same parents may invest more in their children than is currently captured in the model. Furthermore, if these especially altruistic parents invest more in their children so as to broaden their future migration opportunities, then the model may be understating the behavioral response of these parents to migration restrictions.

[42] See Figures C.2c and C.2d for a state-level visualization of this fit. Areas such as the South, the Rust Belt, and the Northeast are captured well, but the model does understate mobility for stayers in some Mountain and Plains States, especially Montana, Idaho, and South Dakota.

Table 4: IIM Statistics for Stayers

Moment	Data	Model
Absolute Mobility, Overall	42.43	43.34
Absolute Mobility, Stayers	39.87	40.63
Overall/Stayer Correlation	0.89	0.87

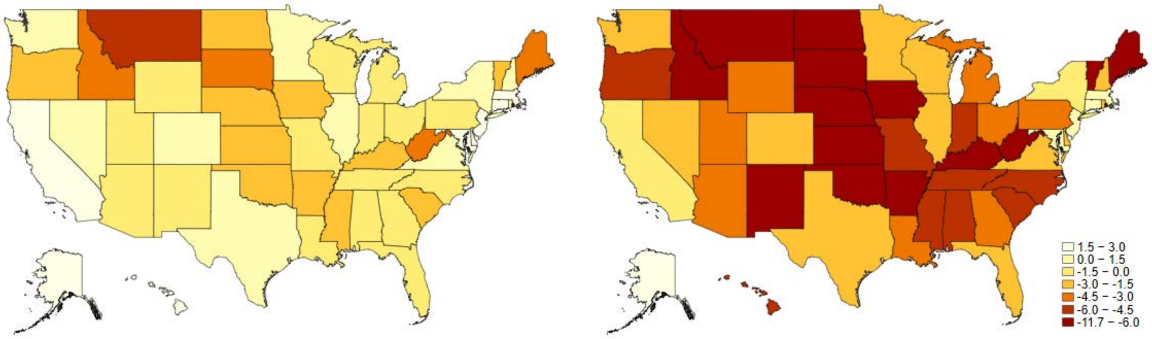
Notes: Table reports model fit in regards to outcomes of stayers vs. overall sample. The first row reports rates of IIM, measured as the expected family income percentile rank of children born to parents in 25th income percentile in the data and expected family income percentile of children born to below-median income parents in the model, averaged across all states. The second row reports the same statistic for individuals who stay in their home state. The third row reports the correlation of state-level IIM rates between the whole sample and the stayer subsample.

that the model does a reasonable job of fitting parent-child income quintile transitions, and Table C.6b indicates that the model slightly overstates the amount of time inputs received by children with married parents — however, the amount of time the typical individual parent spends with their child is close to the average time investment of 0.11 targeted by Lee and Seshadri (2019). Note further that the moments in Table C.6a and Table C.6b are not targeted in model estimation. Tables C.6c and C.6d display the model's fit for the full set of college attendance moments broken down by ability, parent income quartile, and parental educational attainment. In addition to overstating college attendance among the lowest-ability children, the model understates college-going for the poorest children without college-educated parents as well as children with below-median income parents who are college-educated. Otherwise, the fit is reasonable.

5 Decompositions and Counterfactual Exercises

I now use the model to perform decompositions of the model's mechanisms and evaluate the effects of counterfactual policies. Experiments that evaluate the impact of a counterfactual on IIM do so by comparing the average baseline income rank achieved by children with 25th-percentile-parents in a given state compared to the outcomes for the same group of individuals in the baseline model.

Figure 6: Counterfactual — Results of Migration Restrictions



(a) Change (no BR) (b) Change (full BR)

Notes: BR = Behavioral responses. Figure 6 plots the change in upward mobility from counterfactuals that restrict migration while ignoring or including behavioral responses. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile in the data and expected family income percentile of children born to below-median income parents in the model.

5.1 The Role of Migration

The main goal of this paper is to assess the importance of migration in influencing IIM in the United States. I approach this question from two directions — in one counterfactual, I run the model as before and then move all individuals back to their home states ex-post, and in another I simply set the moving fixed costs for periods 1 and 2 (δ_1, δ_2) to infinity, which eliminates any migration in the model entirely as well as any human capital accumulation incentives generated from migration opportunities. These two counterfactuals can be thought of as restricting migration while either ignoring or accounting for behavioral responses. While implementing migration restrictions in the real world would clearly have general equilibrium effects, I approach this exercise from a partial equilibrium point of view[43]: in essence, I consider how the expected outcomes of a *single* child born in a given state would change if that child alone were made unable to move.

Despite the model reproducing the strong correlation between stayer outcomes and overall outcomes, I find that the impacts of migration restrictions on earnings are large, and some of the largest effects appear in among the most income-mobile parts of the country. Figure 6 displays the geographic distribution of the changes in IIM induced by these counterfactuals.

[43]Indeed, historical evidence points to important general equilibrium reactions to large migration flows (Derenoncourt, 2021).

Agents generally gain little from these counterfactuals in terms of earnings — this is not surprising, as moves in the model are usually to higher-paying areas, so restricting migration ex-post typically reduces the earnings of movers while having no effect on stayers. States with the highest returns to human capital typically had low rates of out-migration to begin with and so stand to benefit little from a migration restriction. However, while average earnings in the majority of states suffer from the migration restriction, the effects are quite heterogeneous: with behavioral responses, the model predicts that barring a South Dakota native from moving later in life would result in their expected income percentile rank as an adult dropping by over 10 points (translating into a family income reduction of around $15,000); on the other hand, doing the same to a child from Connecticut would *increase* their expected rank by about 2.9 points (around $3,000)[44]. These effects are the result of a combination of mechanically moving people to higher or lower-paying areas (which is key in generating improved outcomes for the natives from Connecticut, the state with the highest skill prices) and behavioral responses of individuals decreasing their human capital investment and college-going following the removal of migration opportunities. These responses play an important role: the percentile shifts for South Dakota and Connecticut natives in the counterfactual that ignores behavioral responses are a 4.6 reduction and a 3.2 gain, respectively. More rural states are generally hit harder by the counterfactual, with particularly strong earnings effects observed among some states in the Great Plains and Appalachian areas.

This suggests that migration as well as opportunities to do so may be important in shaping adult outcomes for children from more remote areas — to frame the results differently, Table 5 summarizes the impacts of the counterfactual for states in and out of the West North Central and Mountain Census divisions, the two divisions with the highest levels of IIM in the United States. The effect is such that the gap in upward mobility measures between those states in the divisions of interest and those not decreases by approximately 1.40 points when ignoring the behavioral response (Row 3) and 2.23 points when including it fully (Row 6). The shift of 2.23 points with full behavioral responses is statistically significant and represents slightly more than half the gap in IIM between the two groups of states of 4.45, suggesting roughly that half the advantage these areas enjoy in measures of upward mobility[45] may be attributed

[44]Though note that individuals base their migration decisions both on pecuniary factors such as consumption but also on non-pecuniary factors such as amenities and preference shocks. As a result, the utility implications of these migration restrictions are always negative; see Figure C.4.

[45]I also consider gaps in earnings levels as an outcome and reach a similar, if not stronger conclusion: the gap in family earnings gaps between the two locations in the data is approximately $6,000, and the model

Table 5: Migration Restriction Impacts by State Group

Statistic	(1) West North Central/Mountain	(2) Not WNC/MO	(3) (1) − (2)
IIM (CHKS measure)	45.55 (1.03)	41.10 (0.45)	4.45 (0.96)
IIM (Model)	46.25 (0.93)	42.10 (0.68)	4.15 (1.21)
IIM Δ from model baseline (no BR)	-1.56 (0.46)	-0.16 (0.33)	-1.40 (0.59)
IIM Δ (period 2 BR only)	-0.28 (0.44)	1.12 (0.35)	-1.40 (0.60)
IIM Δ (period 2 and college BR only)	-2.72 (0.57)	-1.09 (0.44)	-1.70 (0.77)
IIM Δ (full BR)	-5.56 (0.67)	-3.36 (0.45)	-2.23 (0.82)

Notes: BR = Behavioral Response. Table 5 reports average impacts for states either in or out of the West North Central and Mountain Census divisions. Row 1 reports IIM as reported in CHKS, while Row 2 reports IIM as predicted by the model. Row 3 reports changes in IIM following a counterfactual that shuts down migration in the model, and Row 4 does the same while allowing for the agent to adjust their behavior in the young adult and adult stages of the model. Row 5 further allows the agent and their parent to adjust their college decision in response to the migration restriction, and Row 6 enables parents to adjust their child investment behavior as well. Standard errors of estimates in parentheses. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile in the data and expected family income percentile of children born to below-median income parents in the model. See Appendix A for division definitions.

to migration channels.

I further decompose the behavioral responses by simulating two alternate scenarios: in the first (Row 4), the children only adjust their behavioral margins when they reach the adult stages of the model (i.e. the self-investment decision in period 2 and child investment decision in period 3). In the second (Row 5), the children and their parents can also adjust their college decision in response to the migration restriction. This allows me to quantify which model mechanisms are most important in generating the result. I introduce behavioral responses backward in the lifecycle in this way to account for the fact that human capital in the model forms cumulatively from a series of investment decisions. Incorporating the young-adult self-investment behavioral response alone hardly changes the result at all, since migration and the returns to it are much lower later in the life cycle. In fact, since those from the West North Central and Mountain divisions are forced to live in areas with generally better marriage markets, they invest slightly more in their human capital to increase the odds of favorable family formation outcomes. The migration restrictions have much stronger impacts when the agent's college decision can adapt to the counterfactual world: college

predicts this gap to close by around $3,900 when behavioral responses are fully incorporated.

attendance falls by approximately 30%[46] in the model when nobody is allowed to move, resulting in considerable reductions to absolute mobility. However, these reductions do little to change the *difference* in outcomes between states in and out of the divisions of interest relative to the case where behavioral responses are ignored entirely, suggesting that the compounding nature of human capital and *parental* behavioral responses to the migration opportunities of their children play a crucial role in generating high upward mobility in low-wage states[47]. This decomposition also highlights the importance of considering continuous human capital and continuous human capital investments: were human capital limited to the binary college/non-college types (as in Eckert and Kleineberg (2021), for instance), behavioral responses in the model may be meaningfully understated.

5.2 Other Determinants of Intergenerational Mobility

While migration may be important in explaining salient aspects of income mobility in the United States, my model contains a rich set of cross-state heterogeneity that allows me to assess the role of other factors in explaining interstate inequality in economic mobility. In my next exercise, I separate these factors into three categories before equalizing them across states, one category at a time, and resimulating the model to observe how outcomes across states change. The categories I use include demographic attributes (include racial ratios, population size, family structure, and marriage/fertility probabilities), economic attributes (including skill prices and costs of living), and school attributes (including real government per-pupil expenditure, student-teacher ratios, and tuition prices).

Table 6 reports the results of this exercise. Among the different specifications, equalizing demographic attributes has the largest effect on reducing the cross-state variance in IIM, lowering the standard deviation by almost 40%. Equalizing schooling attributes also results in a meaningful reduction in inequality but not by as much. However, equalizing costs of living and skill prices slightly *increases* the cross-state standard deviation in income mobility, suggesting as in the data that local labor market conditions and income mobility may only be tenuously related to one another.

[46]While large, international evidence suggests that this magnitude may not be out of the question: Spirovska (2021) finds that college attendance rose by 30% in countries that were included in the 2004 enlargement of the European Union.

[47]See Figure C.3 for an alternate representation of these results as scatterplots plotting state-level IIM against predicted changes in IIM following the migration restrictions. The plots show a clear negative association between the two that sharpens when behavioral responses are considered compared to when they are ignored, indicating again that migration is particularly important in explaining upward mobility in some of the most income-mobile states in the U.S.

Table 6: Model Decomposition of Sources of Income Mobility

Specification	IIM Mean	IIM SD
Data	42.43	3.69
Baseline	43.34	4.33
Equalized demographic attributes	42.10	2.68
Equalized economic attributes	44.10	4.68
Equalized school attributes	43.08	3.87

Notes: Table presents mean and SD of cross-state IIM under several model specifications. Demographic attributes include racial ratios, population size, family structure, and marriage/fertility probabilities. Economic attributes include skill prices and costs of living. School attributes include real government per-pupil expenditure, student-teacher ratios, and tuition prices. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile in the data and expected family income percentile of children born to below-median income parents in the model.

5.3 Retention Policies

While migration is important in generating income mobility in low-wage parts of the country, several U.S. states have been or are concerned about the tendency of talented individuals to vacate them. As a result, these states have recently weighed legislation that would provide financial incentives for individuals with higher human capital (typically, college graduates) to locate in them.[48] Advocates of such bills argue that they would increase the retention of talent in the states and could help revitalize depressed local economies, perhaps through positive externalities generated by the presence of highly skilled individuals (Moretti, 2004). Critics argue that such subsidies are targeting the individuals that need them the least or are not on the margin of staying/leaving in the first place. Additionally, a natural economist's objection is that distorting the location choices of highly skilled individuals is unlikely to be efficient on a national level as well. Evaluating the global optimality of such policies would require an equilibrium analysis and more careful consideration of agglomeration economies and spillovers induced by high-human-capital individuals concentrating geographically. Such

[48] In 2018 New York introduced the Excelsior Scholarship, which provided free tuition for middle-class college students conditional on planning to live in New York following graduation. Montana considered but did not pass a measure that would have offered tax breaks for professionals to settle in rural areas in 2019. The Ohio legislature considered a bill to give monetary rewards to STEM graduates in 2017. The Mississippi house approved a measure to give tax breaks to college graduates in 2018, and Michigan considered a similar policy in 2013 that would give tax credits for student loan repayments. South Dakota and Nebraska have both introduced resolutions that at least formally recognize brain drain to be a problem while abstaining from prescribing any specific policy remedies.

Table 7: Counterfactual — State Retention Policies

Division	$10k Subsidy		$20k Subsidy		$50k Subsidy	
	Δ Coll Share	%Δ Rev	Δ Coll Share	%Δ Rev	Δ Coll Share	%Δ Rev
NE	0.37	-0.83	1.00	-1.00	2.49	-2.48
MA	0.46	2.20	0.97	1.84	2.49	0.75
ENC	0.28	1.10	0.67	0.87	1.75	-0.33
WNC	0.34	-1.10	0.66	-1.90	1.87	-4.71
SA	0.32	-0.81	0.67	-1.88	1.72	-4.68
ESC	0.37	-1.13	0.70	-2.01	1.80	-4.79
WSC	0.37	0.26	0.77	-0.39	1.85	-2.95
MO	0.33	-2.24	0.70	-3.48	1.70	-7.49
PA	0.66	-0.09	1.15	-0.71	2.29	-3.92

Notes: Table presents results of counterfactual policy that subsidizes individuals to live in specific states conditional on having a college degree as a young adult. Net revenue computed as taxes on earnings in periods 2-4 minus subsidy payouts. Tax rates computed as sales tax rates plus average state income tax rates from NBER TAXSIM. College share measured in period 3 after all moves in model have been made. Estimates account for a 1.6% and 0.4% spillover effect of a 1 percentage point increase in college share on high school and college graduate earnings, respectively. Results summarized at the divisional level; see Appendix A for division definitions.

issues are beyond the scope of this paper, so I instead focus on whether such a policy may be profitable from an individual state's point of view.

I use my model to assess the likely cost-effectiveness of these programs. Specifically, I consider three counterfactuals in which locations provide subsidies of $10,000, $20,000, and $50,000 to individuals who both choose to live in them as a young adult (that is, in period 2) and who have a college degree. These three policies are sequentially introduced in each individual state, one at a time, before re-solving the model and re-simulating data. I consider two impacts of the policies: the change in the end-period college share in the state after the introduction of the counterfactual as well as the percentage change in each state's net tax revenue[49].

While the mass of talented individuals will likely increase following the introduction of such a policy, the effect on state revenues is *a priori* ambiguous. Larger numbers of talented individuals will increase a state's tax base, but balances may fall if the income tax revenue cannot make up for the paid subsidies — additionally, a substantial proportion

[49] Revenue here is computed from tax rates on an individual's earnings in periods 2, 3 and 4. Losses come from the states having to pay out the subsidies to qualifying individuals. The analysis is not conducted for Alaska or New Hampshire, as both these states have income and sales tax rates of zero.

of the subsidies may be going toward individuals who would have stayed regardless. More individuals with high human capital stocks may also increase the tax base through increasing the income of other people via spillover effects; as a simple way to account for potential externalities of the presence of college graduates on the earnings of others, I allow for the earnings of high schoolers and college graduates in a state to increase by 1.6% and 0.4%, respectively, following a 1 percentage point increase in that state's college share.[50]

Table 7 presents the results of this exercise at the division level and indicates that the policies generally fail to be cost-effective. The responses of individuals to the policy is generally small — even in the counterfactual that offers a $50,000 subsidy, the typical state sees less than a 3 percentage point increase in the college share of their labor force. This happens because even $50,000 is negligible relative to lifetime earnings for highly skilled individuals, and utility shocks play an important role in both migration and college decisions. As a result, the majority of agents with a college degree are not sufficiently close to migration margins to respond to the policy, and the overwhelming majority[51] of the subsidies go to individuals who choose to both obtain a college degree and locate in the given state in the baseline model without the subsidy, which in turn renders the policy highly cost-ineffective. Moreover, the policies are the least cost-effective in the low-wage areas of the country that have considered implementing them the most.

5.4 Schooling Policies

As a final counterfactual exercise, I evaluate how changes to school characteristics at the state level influence cross-state inequality in social mobility. I first consider equalizations of schooling characteristics across each state: in one counterfactual, I set real government child expenditures and student-teacher ratios across all states to either the best values in the data (corresponding to the real expenditure of Vermont and the student-teacher ratios of Wyoming, respectively). In another, I set tuition prices for all states to the lowest value observed (that of Oklahoma's). I also run scenarios that set levels of school characteristics and tuition equal to average values instead of the best values. I resimulate the model in these counterfactual worlds and observe how cross-state means and spreads in absolute mobility compare to the baseline world[52].

[50] These are the spillover effects for high school and college graduates estimated by Moretti (2004). Agents are assumed to be unaware of these externalities when making migration decisions.

[51] 98.8.2%, 97.7%, and 94.5%, respectively, for the three policies.

[52] The caveat that these are partial equilibrium exercises that evaluate how the outcomes of a single child differ in expectation under the counterfactual policy bears repeating — in particular, general equilibrium

Table 8: Effects of Schooling Equalizations

Policy	IIM Mean	IIM SD	Utility
Data	42.43	3.69	—
Baseline	43.34	4.33	2.45
Equal School (Best)	48.97	3.52	2.78
Equal School (Average)	43.60	4.32	2.47
Equal Tuition (Best)	43.55	3.87	2.44
Equal Tuition (Average)	43.62	4.29	2.47

Notes: Table displays cross-state means and spreads of IIM from data, model baseline, and counterfactual simulations. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile in the data and expected family income percentile of children born to below-median income parents in the model. Equal School counterfactual raises real government expenditures and lowers student teacher ratios for all states to either levels (Vermont for government expenditure, Wyoming for student-teacher ratios) or average levels. Equal tuition counterfactual lowers tuition costs to either lowest level (Oklahoma) or average level for all states.

Table 8 reports the results of these policies and indicates that improvements to early school characteristics are considerably more impactful in both improving utility and equalizing outcomes than tuition reductions: exposing the average child to the best possible schooling environment improves their upward mobility in expectation approximately 30 times more than if they are offered the lowest tuition prices. Moreover, the cross-state spread in IIM falls by close to one third when school characteristics are equalized (this is true whether they are equalized according to best or average values), while equalized tuition hardly shifts the spread at all. This is not surprising, given that only around 19% of individuals with below-median income parents in the data (22% in the model) actually obtain a college degree in the first place, and other research indicates that interventions earlier in the life cycle are likely to be more potent in improving economic mobility (Heckman et al., 2010; Lee and Seshadri, 2019; Bailey et al., 2020).

I additionally run a counterfactual comparable to the skill price shocks where I simultaneously increase government expenditure on students g^{ℓ} by 10% and decrease student-teacher ratios s^{ℓ} by 10% for each state individually before resimulating the model and calculating population elasticities as before, which allows me to compare how individuals weigh school quality in the model when making migration decisions compared to skill prices. I recover an elasticity of about 0.3 for government expenditure shocks — as a point of comparison, I also

responses to school funding equalizations are important to consider (Eckert and Kleineberg, 2021).

compute this elasticity for a 10% wage shock in each location[53] and recover an elasticity of 0.9, similar to that estimated by Kennan and Walker (2011). This indicates that parents in my model do respond to child-raising environments when making migration decisions but to a lesser extent than they do economic conditions. Figure C.5 displays how IIM and utility in each state responds to the shocks to wages and school characteristics, indicating that the shocks generally have the least bite in the Great Plains and the Mountain States and are more effective in the Rust Belt, the Northeast, and the Southeast. The state-level effects between the two types of shocks are highly correlated, and the improvements of the schooling shocks on IIM are slightly higher — this is notable, since school characteristics are likely easier to influence via policy than state-wide wages.

6 Discussion and Conclusion

This paper extends a canonical model of intergenerational human capital investment to a geographic context in order to study the role of migration in determining optimal human capital accumulation and income mobility in the United States. The main result is that migration is considerably influential in shaping the high rates of economic mobility observed among children from low-wage areas. Roughly one half of the advantage some of the most rural areas in the country enjoy in measures of IIM can be attributed to natives from these states leaving them and earning more elsewhere. Behavioral responses are important to consider: natives from low-wage areas, along with their parents, invest in their human capital partly in anticipation of leaving, which helps motivate the weak relationship between labor market productivity and upward mobility observed in the data. Since migration opportunities may increase the expected returns to human capital investment before migration decisions are made, these behavioral responses result in improved outcomes for both stayers and movers. Policies designed to decrease the outflow of talented youth from low-wage areas via cash subsidies are unlikely to be effective due to the large majority of these transfers going to individuals who would have stayed regardless. Finally, attempts to equalize schooling resources will likely be more effective in reducing interstate inequality in income mobility if they are targeted earlier in the life cycle.

The main limitations of the model come from the combination of continuous human capi-

[53]This would correspond to an increase of about $4,800 for an individual who spends all their time working. To provide real-world analogues to these shocks, Michigan saw a 12.7% percent decrease in its high skill price between 2000 and 2010 following the crash of the auto industry. On the other hand, North Dakota saw a 9.7% percent increase in its high school skill price in the same period following the fracking boom.

tal and numerous locations forcing the decisions the agents make as well as the heterogeneity considered in the framework to be compressed to maintain computational tractability. Wages in the model are especially simplistically determined, and a more flexible specification of the wage process that considered factors such as location match effects (Kennan and Walker, 2011) or the role of geographic concentrations of occupations may be desirable. While the geographic specificity of occupational returns has declined over recent decades (Kaplan and Schulhofer-Wohl, 2017), the model's ignoring of any idiosyncratic match quality between individuals and places likely means that it is understating the role of migration in earnings growth.

Other extensions to my framework may open new and compelling avenues of study. A model that included multiple stages of childhood could consider how the effects of location on human capital growth vary over different stages of child development. Allowing a richer dynamic migration process could enable the model to capture the possibility of an agent explicitly moving back to their home location in anticipation of becoming a parent, perhaps due to a preference to raise a child where they grew up or to receive help in child rearing from grandparents. Another important limitation is the lack of equilibrium considerations — including such factors could allow the model to speak to whether the high rates of economic mobility in rural areas will be likely to persist as high-ability individuals increasingly sort themselves into high-wage areas in the United States (Diamond, 2016). While promising, I leave these issues to future research.

References

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2019). Education policy and intergenerational transfers in equilibrium. *Journal of Political Economy 127*(6), 2569–2624.

Agostinelli, F. and M. Wiswall (2020). Estimating the Technology of Children's Skill Formation.

Bailey, M., S. Sun, and B. Timpe (2020). Prep School for Poor Kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency. *American Economic Review Forthcoming*.

Batista, C., A. Lacuesta, and P. C. Vicente (2012). Testing the 'brain gain' hypothesis: Micro evidence from Cape Verde. *Journal of Development Economics 97*(1), 32–45.

Bayer, P., F. Ferreira, and R. McMillan (2007). A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy 115*(4), 588–638.

Becker, G. (1994). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education* (3rd ed.). University of Chicago Press.

Becker, G. S. and N. Tomes (1979). An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy 87*(6), 1153–1189.

Bilal, A. and E. Rossi-hansberg (2021). Location as an Asset. *Econometrica 89*(5), 2459–2495.

Bowlus, A. J. and C. Robinson (2012). Human capital prices, productivity, and growth. *American Economic Review 102*(7), 3483–3515.

Caucutt, E. M. and L. Lochner (2020). Early and late human capital investments, borrowing constraints, and the family. *Journal of Political Economy 128*(3), 1065–1147.

Chetty, R., J. N. Friedman, N. Hendren, M. R. Jones, and S. R. Porter (2020). The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility. *NBER Working Paper Series*, 25147.

Chetty, R. and N. Hendren (2018a). The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. *The Quarterly Journal of Economics 133*(3), 1107–1162.

Chetty, R. and N. Hendren (2018b). The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates. *The Quarterly Journal of Economics 133*(3), 1163–1228.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics 129*(4), 1553–1623.

Chyn, E. and D. Daruich (2021). An Equilibrium Analysis of the Effects of Neighborhood-based Interventions on Children.

Colas, M., S. Findeisen, and D. Sachs (2021). Optimal need-based financial aid. *Journal of Political Economy*.

Compton, J. and R. A. Pollak (2007). Why are power couples increasingly concentrated in large metropolitan areas? *Journal of Labor Economics 25*(3), 475–512.

Cunha, F. and J. Heckman (2007). The technology of skill formation. *American Economic Review 97*(2), 31–47.

Cunha, F., J. Heckman, and S. Schennach (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica 78*(3), 883–931.

Dahl, G. B. (2002). Mobility and the return to education: Testing a Roy model with multiple markets. *Econometrica 70*(6), 2367–2420.

Daruich, D. (2020). The Macroeconomic Consequences of Early Childhood Development Policies.

Del Boca, D., C. Flinn, and M. Wiswall (2014). Household choices and child development. *Review of Economic Studies 81*(1), 137–185.

Derenoncourt, E. (2021). Can you move to opportunity? Evidence from the Great Migration.

Diamond, R. (2016). The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill : 1980-2000. *American Economic Review 106*(3), 479–524.

Dillon, E. and J. Smith (2017). The Determinants of Mismatch Between Students and Colleges. *Journal of Labor Economics 35*(1), 45–66.

Durlauf, S. N. and A. Seshadri (2018). *Understanding the great gatsby curve*, Volume 32.

Eckert, F. and T. Kleineberg (2021). Saving the American Dream? Education Policies in Spatial General Equilibrium.

Fogli, A. and V. Guerrieri (2019). The End of the American Dream? Inequality and Segregation in US Cities.

Gayle, G.-L., L. Golan, and M. Soytas (2014). What Accounts for the Racial Gap in Time Allocation and Intergenerational Transmission of Human Capital?

Gemici, A. (2006). Family Migration and Labor Market Outcomes Family Migration and Labor Market Outcomes.

Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect inference. *Journal of Applied Econometrics 8*(1 S), S85–S118.

Heckman, J. and R. Landersø (2021). Lessons for Americans from Denmark about inequality and social mobility. *Labour Economics* (April).

Heckman, J., L. Lochner, and C. Taber (1998). Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents. *Review of Economic Dynamics 1*, 1–58.

Heckman, J. J. (2013). The Economics and the Econometrics of Human Development.

Heckman, J. J. and B. E. Honore (1990). The Empirical Content of the Roy Model. *Econometrica 58*(5), 1121–1149.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics 94*(1-2), 114–128.

Heckman, J. J. and R. Robb (1985). Alternative methods for evaluating the impact of interventions: An overview. *Journal of Econometrics 30*(1-2), 239–267.

Huggett, M., G. Ventura, and A. Yaron (2011). Sources of lifetime inequality. *American Economic Review 101*(7), 2923–2954.

Ishimaru, S. (2022). Geographic Mobility of Youth and Spatial Gaps in Local College and Labor Market Opportunities.

Kaplan, G. and S. Schulhofer-Wohl (2017). Understanding the Long-Run Decline in Interstate Migration. *International Economic Review 58*(1), 57–94.

Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *Journal of Political Economy 105*(3), 473–522.

Kennan, J. (2006). A Note on Discrete Approximations of Continuous Distributions.

Kennan, J. (2020). Spatial Variation in Higher Education Financing and the Supply of College Graduates.

Kennan, J. and J. R. Walker (2011). The Effect of Expected Income on Individual Migration Decisions. *Econometrica 79*(1), 211–251.

Lafortune, J., J. Rothstein, and D. W. Schanzenbach (2018). School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics 10*(2), 1–26.

Lee, S. and J. Lin (2017, 03). Natural Amenities, Neighbourhood Dynamics, and Persistence in the Spatial Distribution of Income. *The Review of Economic Studies 85*(1), 663–694.

Lee, S. Y. and A. Seshadri (2019). On the Intergenerational Transmission of Economic Status. *Journal of Political Economy 127*(2), 855–921.

Mincer, J. (1978). Family Migration Decisions. *Journal of Political Economy 86*(5), 749–773.

Mogstad, M., J. P. Romano, A. Shaikh, and D. Wilhelm (2020). Inference for Ranks with Applications to Mobility Across Neighborhoods and Academic Achievement Across Countries.

Molloy, R., C. L. Smith, and A. Wozniak (2011). Internal Migration in the United States. *Journal of Economic Perspectives 25*(3), 173–196.

Moretti, E. (2004). Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics 121*(1-2), 175–212.

Mountford, A. (1997). Can a brain drain be good for growth in the source economy? *Journal of Development Economics 53*(2), 287–303.

Rothstein, J. (2019). Inequality of educational opportunity? Schools as mediators of the intergenerational transmission of income. *Journal of Labor Economics 37*(S1), S85–S123.

Ruggles, S., S. Flood, R. Goeken, E. Meyer, J. Pacas, and M. Sobek (2020). IPUMS USA: Version 10.0 American Community Survey.

Shrestha, S. A. (2017). No Man Left Behind: Effects of Emigration Prospects on Educational and Labour Outcomes of Non-migrants. *Economic Journal 127*(600), 495–521.

Spirovska, S. (2021). Migration Opportunities, College Enrollment, and College Major Choice.

Todd, P. E. and K. I. Wolpin (2003). On the specification and estimation of the production function for cognitive achievement. *Economic Journal 113*(485), 3–33.

Venator, J. (2020). Dual-Earner Migration Decisions, Earnings, and Unemployment Insurance.

A Divisional Groupings of States

- **New England (NE):** Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont.
- **Mid-Atlantic (MA):** New Jersey, New York, Pennsylvania.
- **East North Central (ENC):** Illinois, Indiana, Michigan, Ohio, Wisconsin.
- **West North Central (WNC):** Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota.
- **South Atlantic (SA):** Delaware, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, District of Columbia, West Virginia.
- **East South Central (ESC):** Alabama, Kentucky, Mississippi, Tennessee.
- **West South Central (WSC):** Arkansas, Louisiana, Oklahoma, Texas.
- **Mountain (MO):** Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming.
- **Pacific (PA):** Alaska, California, Hawaii, Oregon, Washington.

B Estimation Appendix

B.1 Skill Prices

This section presents additional details on the Mincer wage regressions used to obtain skill prices w^ℓ from the 2000 Decennial Census and the 2008-2012 American Community Surveys.

High School Skill Prices

When estimating high school skill prices, I restrict my sample further following Eckert and Kleineberg (2021). I limit my sample to individuals aged 25 to 55 with exactly 12 years of education who work between 36 and 60 hours per week and also worked at least 48 weeks in the year preceding the interview. I then take reported wage income in the last year and divide by reported hours worked to arrive at an estimate for hourly wages for each observation in my sample. Exact hours worked in the previous year are available in the 2000 Census. For individuals in the ACS, I know that respondents worked at least 48 weeks in the previous year and see how many hours they worked per week. For lack of a better alternative, I compute annual hours for respondents in the ACS as though they worked all 52 weeks in the previous year. I then run the regression:

$$\log(w_{it}) = \beta_0 + \beta_1 \mathbf{X}_i + \beta_2 x_i + \beta_3 x_i^2 + \beta_4 x_i^3 + \beta_5 x_i^4 + \mathbf{D_{it}} + \varepsilon_{it},$$

where w_{it} is hourly wage for individual i (mapping to $\frac{e_3}{1-t}$ in equation (1)) and $\mathbf{X_i}$ is a vector of demographic characteristics (black, male, and hispanic dummies, along with dummies for having moved from one's home state) included to account for compositional differences across states.[54] This vector together with a quartic polynomial in years of potential experience serve as a collective proxy for h_3 in (1), and the vector $\mathbf{D_{it}}$ represents dummies for living in each state by time period (2000 or 2008-2012) and are what allow me to derive skill prices for high school graduates, computed as $w_t^{\ell,0} = \exp(D_{\ell,t})$. I omit the $\mathbf{D_{it}}$ dummy corresponding to Iowa in 2000 in the regression as a normalization.

Figure B.1 displays the geography of skill prices computed from this method for the two time periods as well as how these prices changed over time. These measures are presented both as they are obtained from the regression equation above and after adjusting for different cost-of-living levels across states. As one may expect, skill prices tend to be lower in states that are lacking in large cities, with particularly low returns for states in the Great Plains and

[54] Leaving out these demographic factors has no discernible impact on the estimates of $\mathbf{D_{it}}$.

Mountain regions. States with large cities, such as California, Illinois, and New York, feature considerably higher returns to human capital, though this attenuates when accounting for different costs of living.[55] Moreover, changes in skill prices observed between 2000 and 2010 intuitively reflect economic phenomena known to have happened in the 2000s: states in Appalachia generally see larger skill price reductions following struggles in the manufacturing sector, and Michigan experiences the single largest fall in skill price due to the collapse of the automotive industry there. Moreover, states such as North Dakota and Wyoming see increases in their skill prices following the fracking boom.

College Skill Prices

The next step is to compute college premia at the state level, after which college skill prices may be obtained by multiplying a state's high school skill price by its college premium. Bias from selective migration is a major concern when estimating college premia, so I use the semiparametric correction method described in Dahl (2002) to adjust my estimates. In particular, the paper presents a sample selection correction in a polychotomous choice Roy model that takes the form of an unknown function of the probability of the first-best (i.e. the observed) location choice[56], where probabilities are computed by observing the migration choices of individuals first categorized into cells, thereby allowing a distribution-free estimate of selection probabilities.

I begin by taking white men in the ACS and Census aged 25 to 54 who either have exactly a high school or exactly a college degree. Individuals living in group quarters are dropped, and I make similar hours and income restrictions compared to before. I then categorize individuals into cells based on birth state, education, marital status, and whether they moved from their birth state. Married stayers are split further according to whether their spouse works, whether they have children less than 5 years old in the household, and whether they have children aged between 5 and 18 in the household. Non-married stayers are separated by whether they are divorced and whether they live alone, with roommates, or with extended family. Married movers are split up by whether they have any children, and non-married movers by whether they live with roommates/extended family; the smaller sample

[55] The figures suggest that agents could as much as double their real earnings by moving from the lowest- to highest-ranked state. This is somewhat misleading as it is driven entirely by Hawaii, where costs of living are so high that the real skill price is adjusted to be very low. Moving from the second-lowest real wage state (Montana) to the highest (Michigan) in 2000 confers a real wage boost of around 30%, which is more reasonable.

[56] See Heckman and Robb (1985) for more on this approach and Heckman and Honore (1990) for details on the empirical content of the Roy model more generally.

of movers necessitates using coarser grids. As in Dahl (2002), the fraction of individuals in a cell who move from one state to another determines the probability that any individual in the given cell follows the same migration path, and the proportion of individuals in a cell who stay in their birth state gives retention probabilities for all individuals in the cell as well.

I then regress log wages on a cubic of experience; dummies for living in an urban area, marital status, and college; and the correction function derived from the migration probabilities computed above. I follow Dahl (2002) in using separate correction functions for stayers and movers: the regressions include the first-best probability for stayers and the first-best probability along with the retention probability for movers. The correction functions are quadratic polynomials of these probabilities; including higher-order terms has little effect on the results.

The results of this correction are shown in Figure B.3a and B.3b. Similar to Dahl (2002), the correction results in a statistically significant lowering of the college premium (typically around 10%) compared to raw OLS estimates.

Robustness

While the college premia I compute are corrected for selection bias, one may still be concerned that my estimates of high-school skill prices are biased from selection as well. I now demonstrate that several methods intended to reduce selection bias return very similar estimates to the high school skill prices that I use.

First, I run a specification that follows Kennan and Walker (2011) that attempts to limit selection from migration even further by limiting the sample to high-school educated males aged 18-20, the intuition being that focusing on new labor market entrants preempts the bulk of migration decisions. The numbers I use are strongly correlated with the output of this method (correlation >0.8), though after normalizing by Iowa's skill price the other estimates of $w^{\ell,0}$ are slightly lower, suggesting that Iowa may have relatively high early entrant wages. A juxtaposition of the high school skill prices obtained in this test vs. the ones I use are available in Figure B.4a.

Second, I run a specification that identifies skill prices exclusively from movers using a two-way fixed effects model with the PSID. The panel structure of the PSID allows me to observe the same individuals at multiple points in time, thus allowing me to include individual fixed effects to account for unobserved heterogeneity while using movers to get information about state-level skill prices. Specifically, I limit the sample to non-college-graduates and estimate:

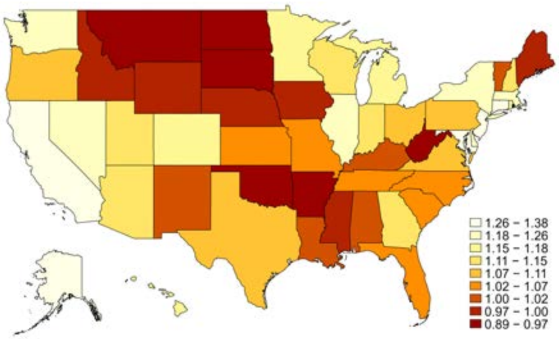
$$\log(w_{it}) = \beta_0 + \beta_1 x_{it} + \beta_2 x_{it}^2 + \beta_3 x_{it}^3 + \beta_4 x_{it}^4 + \beta_5 e_i + \delta_t + \gamma_{S(it)} + \lambda_i + \varepsilon_{it},$$

where w_{it} indicates wage and x_i years of potential experience as before, e_i education, and δ_t, $\gamma_{S(it)}$, and λ_i fixed effects for calendar year, state, and individuals respectively. Standard errors are clustered at the individual level. The γ terms correspond to skill prices and can be identified from wage changes among individuals who move from one state to another. Estimates from this procedure compared to my baseline estimates may be viewed in Figure B.4b. As before, the two sets of estimates are positively correlated, and only two of my baseline estimates fall outside the confidence interval for their corresponding estimate from this method.

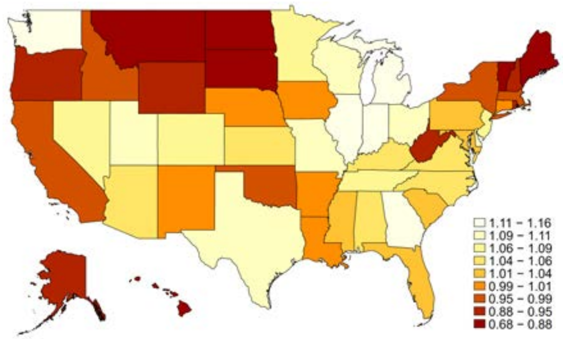
The key problem with both this and the previous test is that they result in very small sample sizes, particularly for low-population states. For any given survey year in the PSID, there are fewer than 10 individuals total in states such as Montana, Vermont, and Wyoming, rendering estimates of skill prices for these states extremely noisy. The early-entrant test features somewhat larger sample sizes, but still has fewer than 100 observations for some states in the 2000 Census. For this reason I prefer the estimates obtained from my baseline method.

The larger samples I use in the baseline method also allow for a final selection test. In particular, I run the same Dahl (2002) procedure to estimate selection corrections for state-specific *high school* earnings premia relative to high school dropouts. I follow the exact same procedure as before but limit the sample to either individuals with exactly a high school degree or high school non-graduates. Figures B.3c and B.3d display the raw and corrected high school earnings premia obtained after this procedure. In contrast to the college premia estimates, the selection correction barely changes the estimates of high school premia at all, providing evidence that selective migration for high school graduates is not nearly as large a concern as for college graduates.

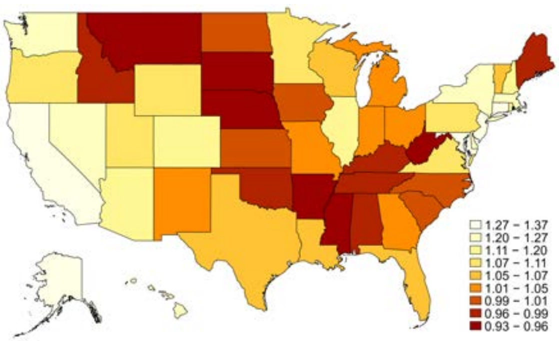
Figure B.1: High School Skill Prices



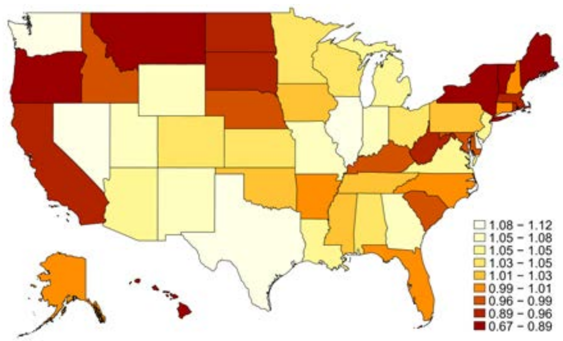
(a) 2000, Raw



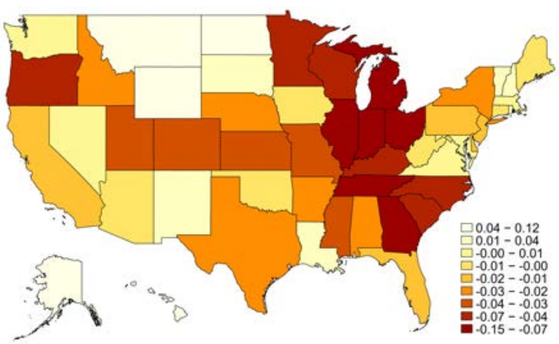
(b) 2000, Real



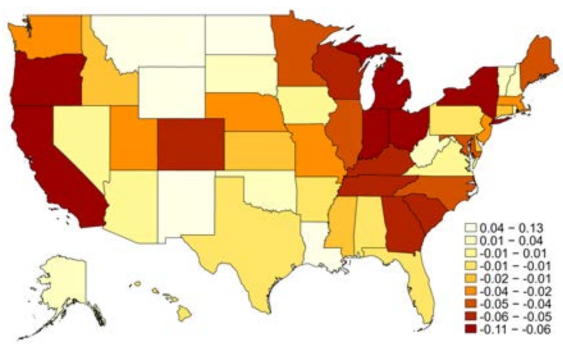
(c) 2010, Raw



(d) 2010, Real



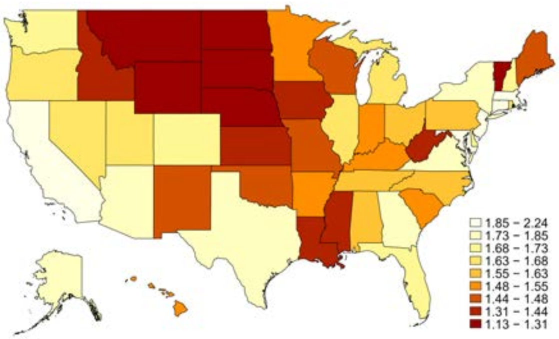
(e) 2000-2010 Change, Raw



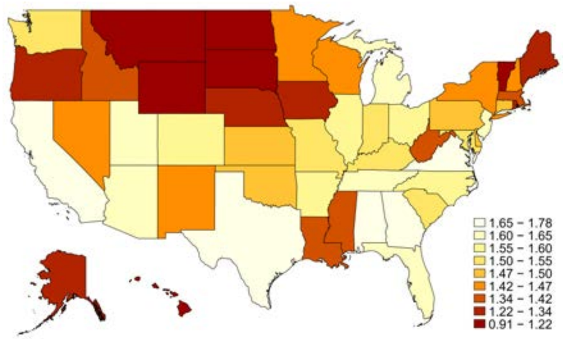
(f) 2000-2010 Change, Real

Notes: Figures visualize estimates of $w^{\ell,0}$. See Section 4.1 and Appendix B.1 for details on estimation procedure. Subfigures (a) and (b) present estimates for the year 2000, both raw and after adjusting for local skill prices, obtained from the 2000 Census. Subfigures (c) and (d) present the corresponding statistics for the year 2010, obtained from the 2008-2012 American Community Surveys. Subfigures (e) and (f) visualize changes in skill prices in the 2000-2010 period.

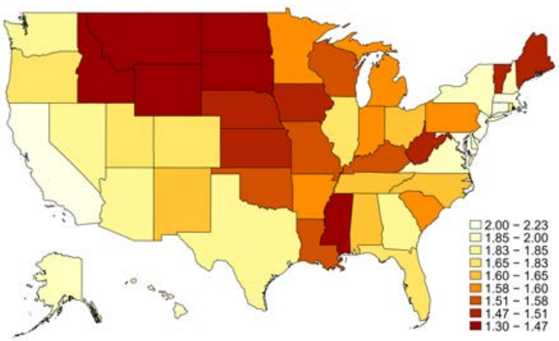
Figure B.2: College Skill Prices



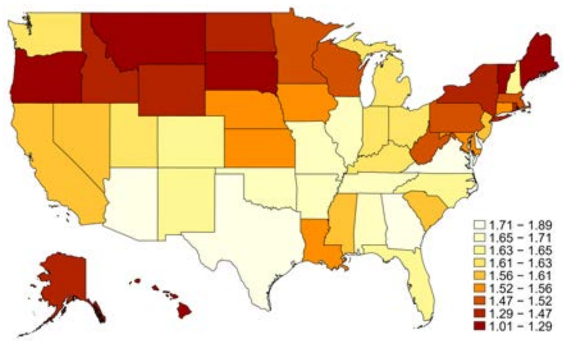
(a) 2000, Raw



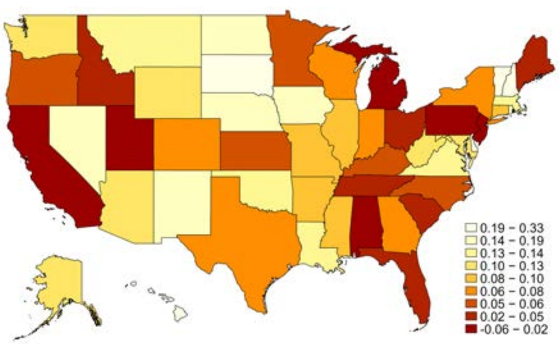
(b) 2000, Real



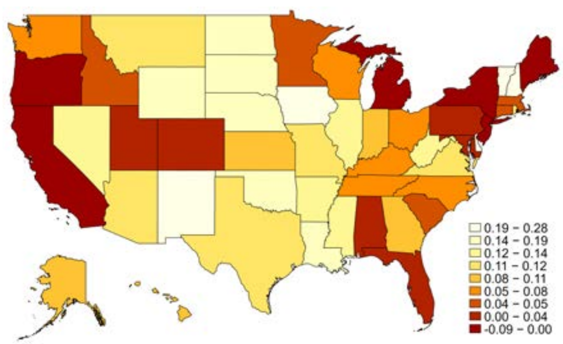
(c) 2010, Raw



(d) 2010, Real



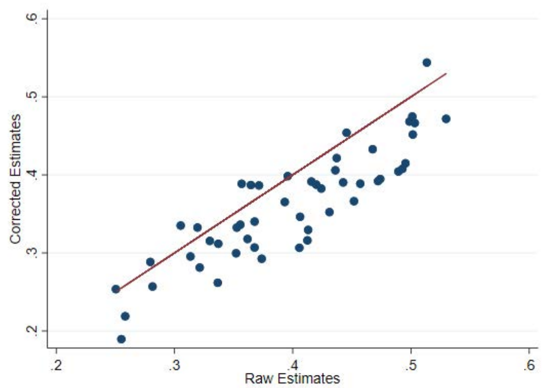
(e) 2000-2010 Change, Raw



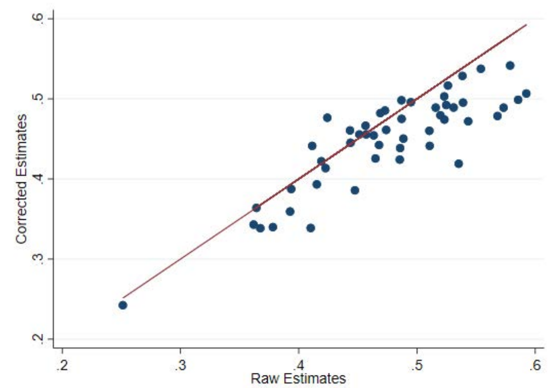
(f) 2000-2010 Change, Real

Notes: Figures visualize estimates of $w^{\ell,1}$. See Section 4.1 and Appendix B.1 for details on estimation procedure. Subfigures (a) and (b) present estimates for the year 2000, both raw and after adjusting for local skill prices, obtained from the 2000 Census. Subfigures (c) and (d) present the corresponding statistics for the year 2010, obtained from the 2008-2012 American Community Surveys. Subfigures (e) and (f) visualize changes in skill prices in the 2000-2010 period.

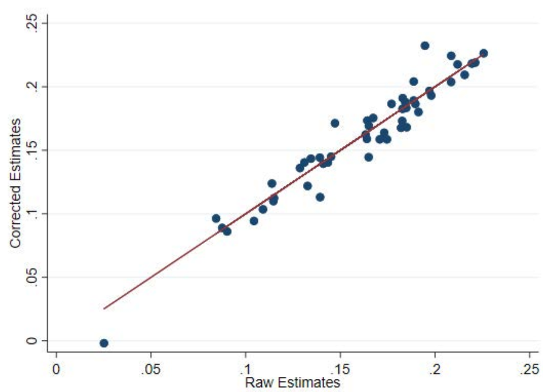
Figure B.3: Skill Price Selection Corrections



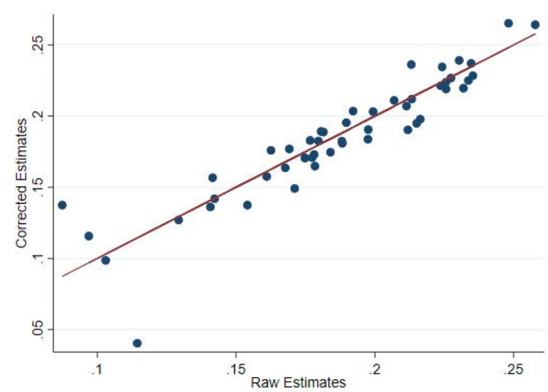
(a) 2000, College



(b) 2010, College



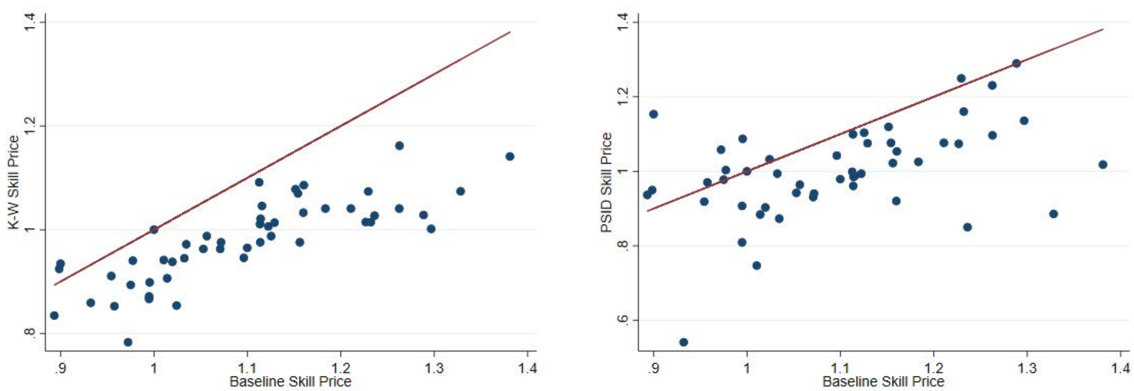
(c) 2000, High School



(d) 2010, High School

Notes: Figures display raw and selection-corrected estimates of $w^{\ell,S}$ using method described in Dahl (2002). Subfigures (a) and (b) juxtapose raw vs. corrected estimates of state-level college premia, relative to high school wages, in 2000 Census and 2008-2012 ACS. Subfigures (c) and (d) display raw vs. corrected estimates of state-level high school premia, relative to high school dropouts, in the same datasets.

Figure B.4: Robustness of High School Skill Price Estimates



(a) K-W (Early Entrants)

(b) PSID (Two-Way FE)

Notes: Figures display baseline estimates of $w^{\ell,0}$ juxtaposed with estimates obtained from alternative specifications. Subfigure (a) plots baseline estimates compared to estimates obtained among early labor market entrants following Kennan and Walker (2011). Subfigure (b) plots baseline estimates compared to estimates obtained using two-way fixed effects model in PSID.

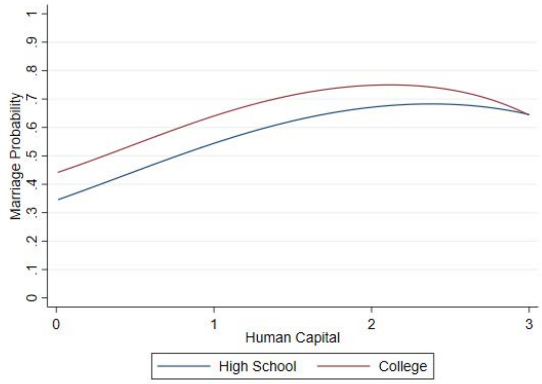
B.2 Marriage Realizations

Marriage probabilities are computed as probit functions of cubic polynomials of human capital, separated by education level and states. This section presents these estimated functions for a subset of states as well as the model's performance in fitting marriage rates by state of birth.

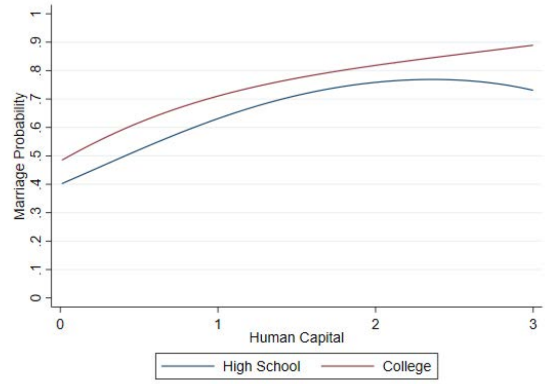
Figure B.5 displays how marriage probabilities evolve over human capital based on state of residence and education. I show probabilities for Mississippi (the state with lowest overall marriage rates), Utah (the highest), and for Iowa, California, New York and Texas. In most cases, there is a clear gap in probabilities between high school and college-educated individuals, and the probabilities of marriage increase steadily over the human capital distribution before eventually leveling off. I hold marriage probabilities constant after a human capital level of 3, which corresponds roughly to the top percentile, to prevent the curvature from making perverse predictions about marriage probabilities for extremely high-human capital individuals. While many states are comparable, considerable heterogeneity is present: note, for instance, that the marriage probabilities for Utah high-schoolers is never below 50%, while in Mississippi the probability for high schoolers starts at barely 20%.

Figure B.6 presents the model's fit of marriages rates for children with 25th-income-percentile parents by state of birth. The data for marriage rates of such children come from the Opportunity Atlas, while the model output corresponds to the average marriage rates of children with below-median income parents by state of birth. The fit is quite strong, with the model explaining more than 70% of the variation in state-level marriage rates, though the model underpredicts marriage rates across the board by a slight amount.

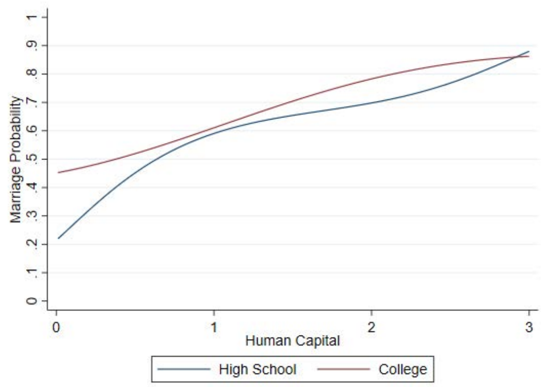
Figure B.5: Marriage Probabilities by State



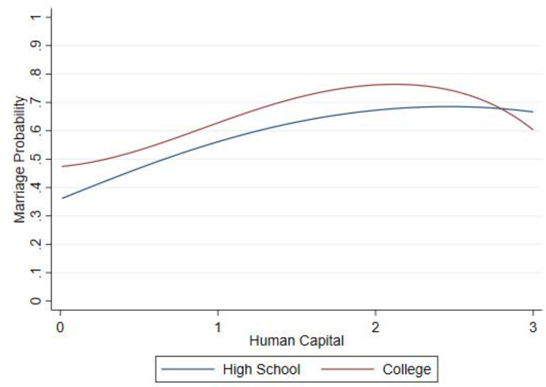
(a) California



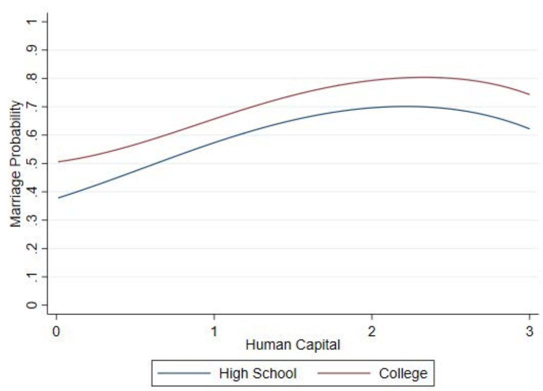
(b) Iowa



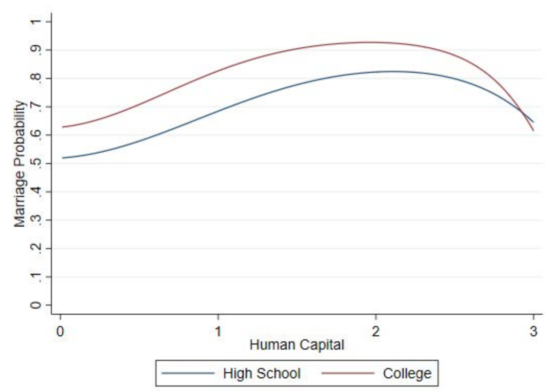
(c) Mississippi



(d) New York



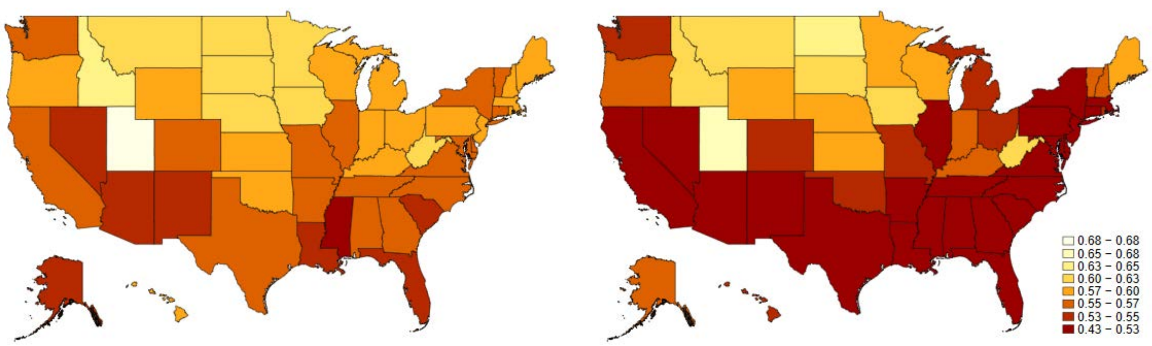
(e) Texas



(f) Utah

Notes: Figures present estimates of marriage probabilities over human capital, separated by education level and state. Probabilities computed as probit functions of human capital cubic and held constant after a human capital level of 3. See text for details and sample construction.

Figure B.6: Model Fit of Marriage Rates



(a) Data (b) Model

Correlation: 0.87

Notes: Figure displays model's fit of marriage rates by state of birth. Data on marriage rates for children with parents in 25th income percentile by state of birth from the Opportunity Atlas. Model output displays simulated average marriages rates for children with below-median income parents by state of birth.

Table B.1: Racial and Regional Heterogeneity Estimates

Parameter		Value	SE	Targeted Moment
Racial Heterogeneity				
Human capital productivity, Black	ξ_B	3.552	(0.012)	Black-White wage ratio
Human capital productivity, Hispanic	ξ_H	3.583	(0.072)	Black-White wage ratio
Migration preference modifier, Black	δ_B	0.112	(0.027)	Migration by race
Migration preference modifier, Hispanic	δ_H	-0.252	(0.045)	Migration by race
College fixed cost, Black	η_B	-1.577	(0.041)	College attainment by race
College fixed cost, Hispanic	η_H	-1.992	(0.063)	College attainment by race
Regional Heterogeneity				
Ability mean, region 1	$\mu_{a,1}$	-0.544	(0.006)	Attendance by region
Ability mean, region 1	$\mu_{a,2}$	-0.358	(0.021)	Attendance by Region
Ability mean, region 1	$\mu_{a,3}$	-0.496	(0.015)	Attendance by region
Ability mean, region 1	$\mu_{a,4}$	-0.481	(0.021)	Attendance by region

Notes: Table reports descriptions of parameters and their symbolic representations in first two columns. Columns three and four report parameter estimates and standard errors, and column 5 describes data moments used in estimation. Standard errors computed via indirect inference.

B.3 Racial and Regional Heterogeneity Estimates and Fit

In this section, I describe the enhancements made to the model to account for racial and regional heterogeneity that are not presented in the main model for ease of notation. I allow for three different races in the model: non-Hispanic White, Black, and Hispanic. I allow for races to influence skill prices and marriage/fertility probabilities by factors that are constant across states[57]. Further, I estimate separate parameters for human capital productivity (ξ), migration preferences (δ_1), and college preferences (η_1) across races to enable the model to fit racial heterogeneity in wages, migration, and educational attainment. I solve the model and compute policy functions for each race before the simulation. In the simulation, I account for state-level differences in racial compositions as well as different proportions of races represented in different types of families. In estimation, I target these parameters to match racial wage ratios (obtained from my ACS sample), rates of racial college attainment (from the NLSY97), and rates of migration across races (again from the ACS).

For regional heterogeneity, I assume that the ability mean μ_a varies across the four Census regions to allow for some flexibility in spatial distributions of talent while maintaining a reasonable number of parameters to estimate. I allow the mean ability μ_a to vary by

[57]While allowing heterogeneous racial effects across states would be more flexible, it is infeasible due to very small cell sizes for racial minorities in low-population states.

Table B.2: Model Fit in Racial and Regional Heterogeneity

Moment	Data	Model	Source
Racial Heterogeneity			
Black-White wage ratio	0.693	0.688	ACS
Hispanic-White wage ratio	0.816	0.825	ACS
Migration, Black	0.343	0.348	ACS
Migration, Hispanic	0.328	0.335	ACS
College Attainment, Black	0.225	0.223	NLSY97
College Attainment, Hispanic	0.217	0.214	NLSY97
College Attendance by Region			
College attainment, region 1	0.373	0.329	NLSY97
College attainment, region 2	0.379	0.318	NLSY97
College attainment, region 3	0.312	0.345	NLSY97
College attainment, region 4	0.311	0.355	NLSY97

Notes: Table presents the model fit by comparing moments obtained from data to moments simulated from the model. Column 1 describes the moment targeted, and columns 2 and 3 show data and model moment values. Column 4 documents the source of the moment. ACS: American Community Survey. NLSY97: National Longitudinal Survey of Youth 1997. See text for details on sample construction.

geography as opposed to the correlation between parent human capital and ability ρ_{ha} due to joint distributions between parent income and AFQT scores in the NLSY97 showcasing comparable correlations but noticeable mean shifts across the four Census regions. These parameters are targeted to improve the model's fit of economic and geographical mobility by state as well as by targeting regional rates of college attainment (from the NLSY97).

Table B.1 presents estimates and standard errors for the parameters governing racial and regional heterogeneity in the model. I estimate lower human capital productivity parameters for both Blacks and Hispanics, as well as higher migration costs for Blacks and higher college costs for Hispanics. Notably, I estimate lower migration costs for Hispanics and lower college costs for Blacks than Whites, indicating that racial differences in factors that influence human capital attainment, such as starting geography and family structure, play an important role in explaining disparities in certain outcomes that cannot be explained by preference heterogeneity alone. The model estimates higher levels of ability in the Midwest region, and lower ability levels in the West, South, and Northeast. The Northeast ability levels are estimated to be lower in part to temper the model's prediction of upward mobility from that region, which is lower than that of the Midwest despite higher wages, higher parental educational attainment, and comparable family structure.

Table B.2 presents the model's fit of salient aspects of additional racial and regional heterogeneity. The model fits Black and Hispanic wages, migration rates, and educational attainment rates quite well, but while the model predicts regional college attainment rates that are comparable to those observed in the data, it does not succeed in producing the qualitative pattern of higher attainment in the Northeast and Midwest and lower attainment in the South and West.

Table B.3: Correlation of Main Decomposition to Baseline under Alternate Specifications

Specification	Correlation with Baseline	Ratio of SDs
2010 Policy Functions	0.76	1.01
Additional amenities: weather and distance to shore	0.99	1.00
Additional amenities: crime and establishments	0.98	1.01
Additional amenities: pension debt and union power	0.99	0.99

Notes: Table presents correlation of impacts of migration restriction with full behavioral response under alternate model specifications to baseline estimates. First column describes alternate specification, second column reports correlation, and third column reports ratio of standard deviation of baseline effects to effects under alternate specification.

B.4 Robustness to Alternate Specifications

In this section, I evaluate the sensitivity of my main results to alternate model specifications. I first evaluate how sensitive the model's main results are to the time period I solve the model in order to speak to concerns regarding the stationarity assumptions that are necessary to impose for the model to be solved. I then investigate how the model's results and fit vary when including additional notions of location amenities. For each alternate specification, I run the paper's key decomposition of shutting off migration with full behavioral responses and report the correlation of state-level impacts on IIM predicted in the alternate specification to the ones in the baseline specification.

The altruistic factor of utility in the model results in it having an infinite horizon, requiring stationarity assumptions for the model to be solved. In the baseline exercise, I solve the model both in the year 2000 and in the year 2010 (essentially, pre- and post-recession) and then use year-2000 policy functions when simulating parent investment and college decisions and year-2010 policy functions when simulating self-investment decisions and final migration choices for the CHKS cohorts. One may be concerned that the parents' lack of knowledge of future economic conditions may alter my model's predictions, so as a simple test I also simulate a specification where I only use year-2010 policy functions, so that all agents in the model always behave as they would in the post-recession world. The first row of Table B.3 reports a high correlation (0.76) of the key predictions of this version of the model with my baseline results, providing some reassurance that the stationarity assumptions I employ are not key in driving my results.

The baseline version of the model also employs a simplistic treatment of location amenities, assuming that larger locations are higher-amenity due to the presence of larger cities.

Table B.4: Alternate Amenity Estimates and Model Fit

Specification	Parameter estimate
Environment	
Distance to shore	0.016 (0.090)
# Warm days	0.066 (0.019)
Value of objective function	2,817
Quality of Life	
Crime rates	-0.070 (0.028)
Establishments per capita	-0.219 (0.059)
Value of objective function	2,913
Political economy	
Union strength	0.045 (0.035)
Pension debt per capita	-0.053 (0.013)
Value of objective function	2,815

Notes: Table reports parameter estimates of additional amenity factors as well as the value of the objective function when including them. Baseline value of objective function: 2,933. Standard errors computed via indirect inference and are in parentheses. Distance to shore measure taken from Lee and Lin (2017). Crime rates measured as average of violent and property crime; statistics from FBI. Establishments per capita statistics from County Business Patterns.

As a second robustness test, I experiment with additional notion of amenities. Specifically, I run model specifications that include environmental factors (including average distance to a coast and number of warm days per year), quality of life factors (including crime rates and establishments per capita), and political economy factors (including union power, measured as whether the state is a right-to-work state, and pension debt per capita). In each case, I re-estimate the model with these additional amenities before performing the main decomposition again. Table B.4 reports parameter estimates for these other amenities and indicates that the inclusion of them typically does not meaningfully impact the model's fit. Moreover, the standard errors of the estimates indicate that the estimates are often imprecisely estimated or statistically indistinguishable from zero. Additionally, Table B.3 again indicates that the model's key predictions are not sensitive to these alternate specifications.

C Supplementary Figures and Tables

Table C.5: OLS Estimates for Various Correlates on CZ-Level IIM

VARIABLES	(1) IIM	(2) IIM	(3) IIM	(4) IIM
Share Single Mothers	-0.409	-0.466	-0.459	-0.490
	(0.0500)	(0.0510)	(0.0511)	(0.0538)
LFP Rate		-0.119	-0.0889	-0.0625
		(0.0241)	(0.0243)	(0.0234)
Student-Teacher Ratio			-0.335	-0.134
			(0.0550)	(0.0544)
Native Outflow				0.139
				(0.0143)
Constant	61.06	68.23	66.50	49.97
	(5.072)	(5.257)	(5.234)	(5.601)
Observations	709	709	680	680
R-squared	0.706	0.716	0.735	0.773

Notes: Robust standard errors in parentheses. IIM measured as the expected 2011-2012 family national income percentile of a child born in 1980-1982 to parents who were in exactly the 25th family national income percentile in 1996-2000. All specifications also include controls for share Black; Theil segregation index; high school graduation rate, college graduation rate, crime, and marriage rates; and Gini coefficient.

Table C.6: Additional Moments

	Parent Quintile				
Child Quintile	1	2	3	4	5
1	34/40%	24/22%	18/16%	13/13%	11/9%
2	28/24%	24/23%	20/20%	16/18%	12/16%
3	18/18%	22/21%	22/21%	21/21%	17/19%
4	12/12%	18/19%	22/22%	24/23%	24/24%
5	8/6%	12/15%	18/21%	25/26%	37/33%

(a) Income Quintile Transtions (Data/Model)

	Married		Unmarried	
Statistic	Data	Model	Data	Model
Total Inputs	0.19	0.21	0.11	0.10
Individual Parent Inputs	0.10	0.11	—	—

(b) Time Investments

	Child Ability				
Parent Quartile	1	2	3	4	5
1	3/6%	13/9%	18/13%	31/17%	40/27%
2	4/19%	15/21%	16/28%	34/25%	48/48%
3	11/22%	15/26%	33/32%	40/41	56/58%
4	17/19%	24/26%	32/34%	46/43	66/60%

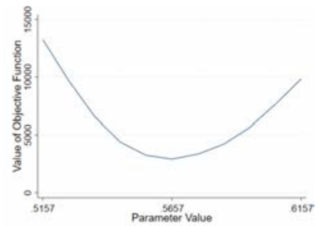
(c) College Attendance, Parents w/o Degree (Data/Model)

	Child Ability				
Parent Quartile	1	2	3	4	5
1	3*/5%	19/6%	34*/8%	42/10%	64*/13%
2	39*/21%	31*/23%	32/31%	65/36%	81/46%
3	3*/38%	43/43%	50/51%	72/59%	77/72%
4	24*/43%	38/51%	61/58%	73/67%	85/79%

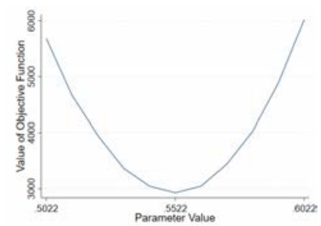
(d) College Attendance, Parents w/ Degree (Data/Model)

Notes: Table C.6a reports income quintile transition probabilities between parents and children. Data moments from Table II of CHKS. Table C.6b reports both total and individual parent time inputs for the children of married or unmarried parents. Data moments from PSID Child Development Supplement; see text for sample construction. Tables C.6c and C.6d report rates of college attendance by parent income quartile and ability quintile for kids with parents without and with a college degree. A star indicates fewer than 25 observations being in the cell and the moment not being used in estimation. Data from NLSY97; see text for sample construction.

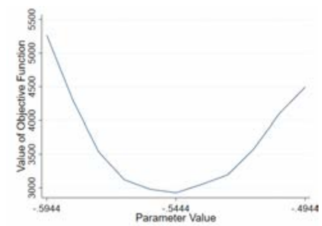
Figure C.1: Behavior of Objective Function



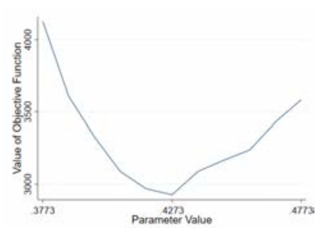
(a) θ



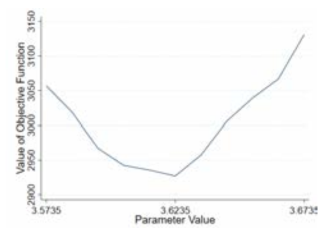
(b) ρ_{ha}



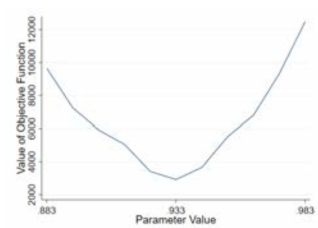
(c) μ_a



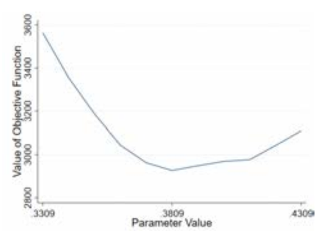
(d) σ_a



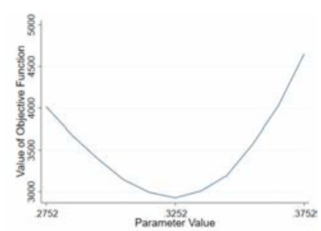
(e) ξ



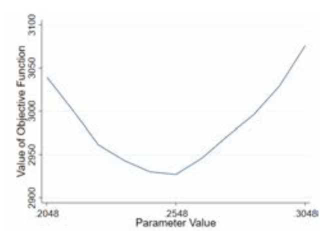
(f) ϕ



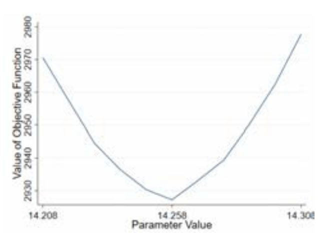
(g) κ



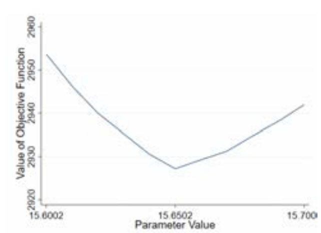
(h) σ_{ε_2}



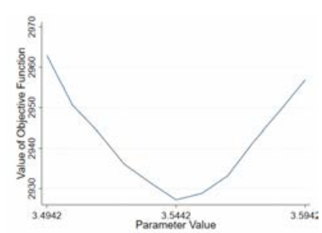
(i) α



(j) δ_1



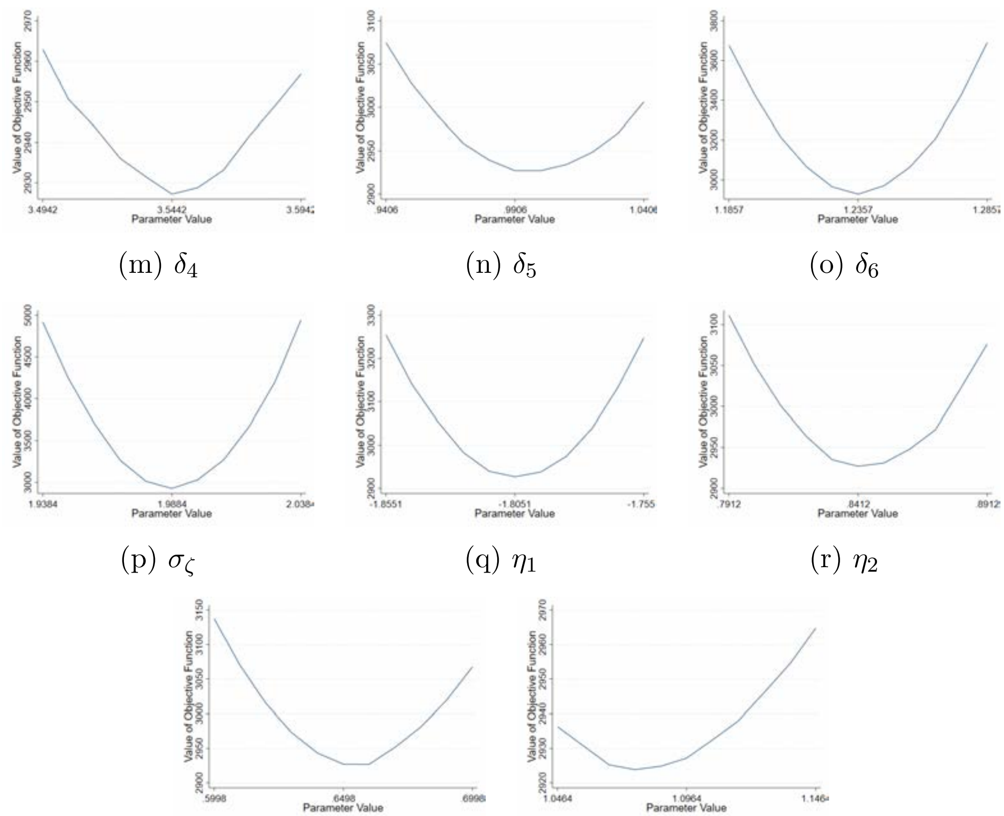
(k) δ_2



(l) δ_3

Notes: Figures plot value of objective function while varying single parameter value indicated by caption and holding all other parameters constant.

Figure C.1: Behavior of Objective Function (continued)



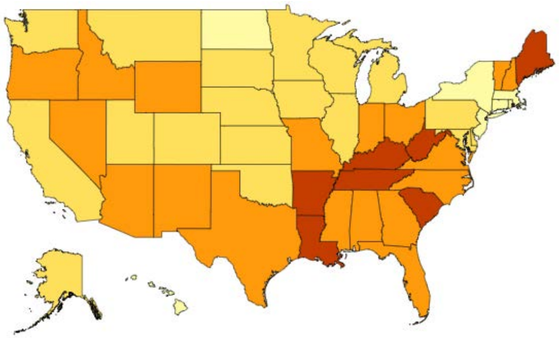
(m) δ_4 (n) δ_5 (o) δ_6

(p) σ_ζ (q) η_1 (r) η_2

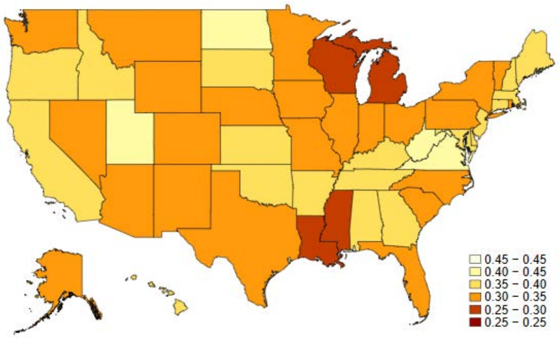
(s) η_3 (t) σ_η

Notes: Figures plot value of objective function while varying single parameter value indicated by caption and holding all other parameters constant.

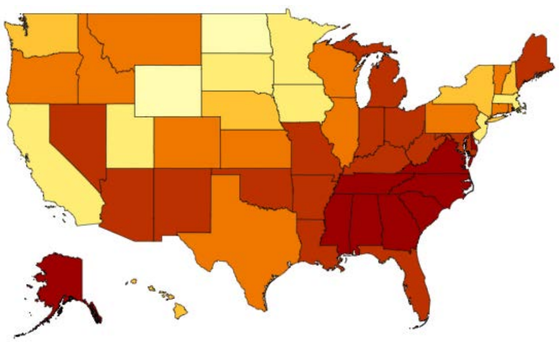
Figure C.2: Additional Model Fit Visualizations



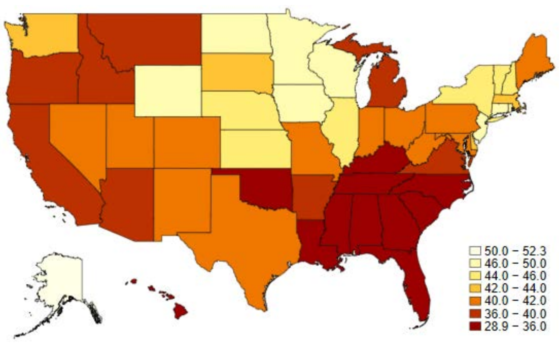
(a) College Attainment, Data



(b) College Attainment, Model



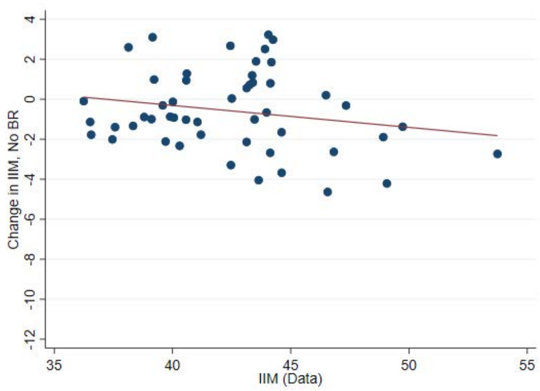
(c) IIM (Stayers) Data



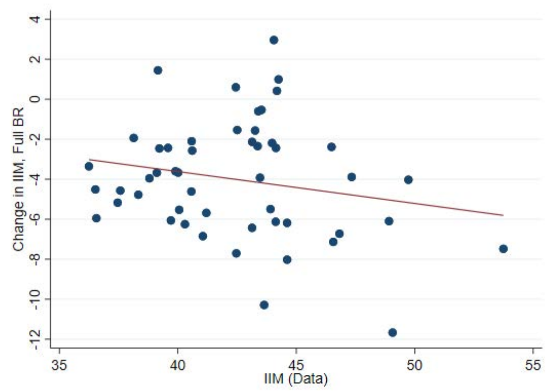
(d) IIM (Stayers), Model

Notes: Figures present rates of college attainment and IIM among stayers as measured in the data and simulated in the model. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile in the data and expected family income percentile of children born to below-median income parents in the model.

Figure C.3: IIM Effects of Migration Restrictions, Scatterplot



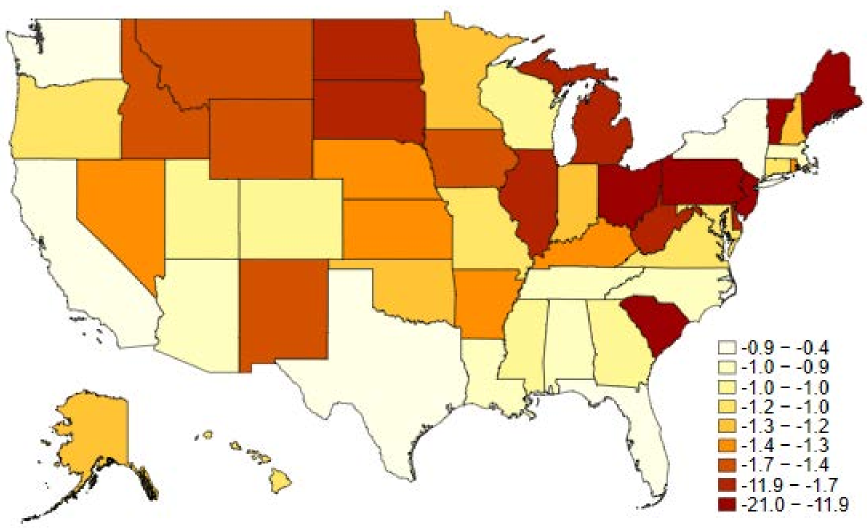
(a) No BR



(b) Full BR

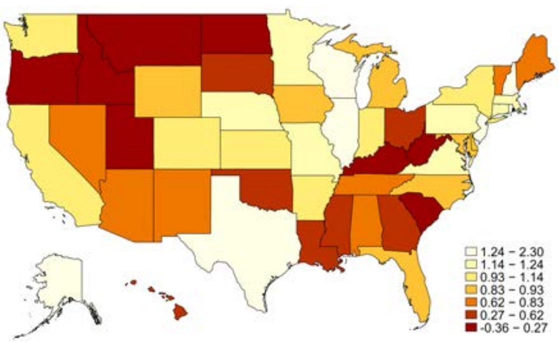
Notes: BR = Behavioral responses. Figure C.3 plots the change in upward mobility from counterfactuals that restrict migration while ignoring or including behavioral responses. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile in the data and expected family income percentile of children born to below-median income parents in the model. X-axis reports state-level IIM rates as measured by Chetty et al. (2014), while Y-axis reports model-predicted changes in IIM following the counterfactuals.

Figure C.4: Utility Effects of Migration Restrictions

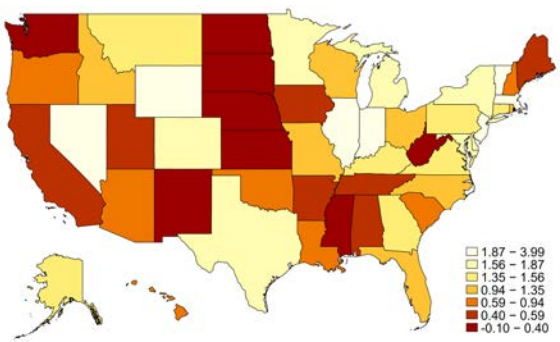


Notes: Figure C.3 plots the change in utiles from counterfactuals that restrict migration while including behavioral responses.

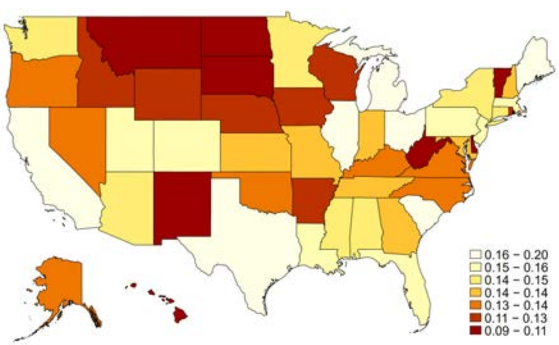
Figure C.5: IIM and Utility Effects of Wage and Schooling Shocks



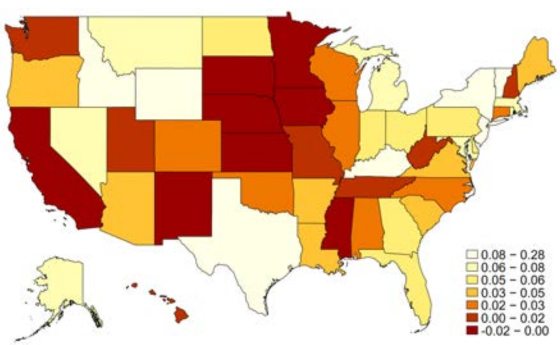
(a) Wage Shocks (IIM)



(b) Schooling Shocks (IIM)



(c) Wage Shocks (Utility)



(d) Schooling Shocks (Utility)

Notes: Figure C.5 plots the change in upward mobility or utility from counterfactuals that either raise skill prices by 10% in a state or raise government school expenditure and decrease student-teacher ratios by 10% in a state. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile in the data and expected family income percentile of children born to below-median income parents in the model.