Spatial Influences in Upward Mobility*

Garrett Anstreicher
University of Wisconsin-Madison

December 5, 2020

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. Accounting for human capital investment behavioral responses to future migration opportunities is important, and these behavioral responses result in skill price shocks having highly varied effects on the earnings of natives based on the location in which they occur. 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. Thanks also to Manuel Amador, Dean Corbae, Chao Fu, Tim Kehoe, Joel McMurry, Jeff Smith, Sandra Spirovska, Joanna Venator, Chris Taber, Matt Wiswall, Anson Zhou, and seminar participants at both Wisconsin labor workshops and EMCON for helpful comments and feedback. 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

A persistent result in the recent literature on income mobility in the United States is that conventional measures of local economic quality, such as labor force participation rates or wage growth, are either weakly or negatively correlated with an area's level of intergenerational income mobility (IIM).[1] 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. Related to this result is that the most income-mobile commuting zones (CZs) in the United States are primarily located in states in the West North Central and Mountain Census divisions,[2] areas that largely lack high wages or productive cities.

However, these rural labor markets with high levels of economic mobility (see Figure 1a for a visualization) also observe 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 isolated 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.

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 difficult, 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 cannot be captured empirically — that is, the option of migration in the future influencing human capital accumulation before migration decisions are actually made. 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.

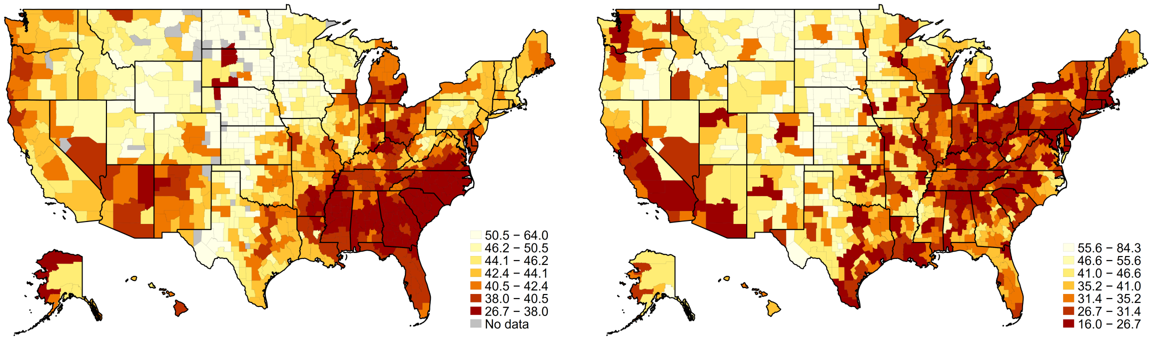
The model extends the seminal Becker and Tomes (1979) framework of intergenerational

[1]Measured as Chetty et al. (2014), or 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.

[2]I.e. the Great Plains and Mountain States. Exact divisional groupings of states available in Appendix A. 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.

[3]Table D.1 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 for children with 25th-percentile parents taken from the Opportunity Atlas (https://opportunityinsights.org/data/).

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 of a certain income level who also endow them with ability and human capital investments. After childhood, the agent invests in their own skills and chooses whether to stay or move to a new location, after which they potentially have offspring of their own. Locations differ across their returns to human capital, family structure, costs of living, and government contributions to child human capital development.

The main mechanism I capture in this framework resembles an intranational brain drain: if a given location has 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 labor market with larger rewards for a high human capital stock. Thus, areas with low human capital rental rates can 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 one third of the gap in upward mobility between states in and out of the West North Central and Mountain Census divisions is attributable to migration. Failing to account for human capital investment in anticipation of future moving options would understate this result by one third.

Next, I use the model to assess the incidence of counterfactual shocks and policies. Increasing skill prices in a given location in the model induces dual effects on native earnings: higher wages improve the outcomes of stayers but also depress incentives for agents to leave and potentially earn higher incomes elsewhere. As a result, my model predicts that positive shocks to skill prices have highly varied impacts on future earnings based on the location in which they occur — for instance, a skill price shock in Texas has more than four times the impact of increasing native adult earnings than an identical shock in North Dakota.

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 a final exercise, I consider a policy that attempts to increase a state's retention of high-skilled individuals 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.

Related Literature

A vast literature exists on IIM and child human capital development (Cunha and Heckman, 2007; Del Boca et al., 2014), with Becker and Tomes (1979) constituting one of the first attempts to model it formally. Lee and Seshadri (2019) incorporate multi-period decisions into the Becker-Tomes model, alleviating some of the issues that arose from compressing the life cycle into two periods. However, the economic prospects of children may depend on where they are born and where/whether they move, and opportunities to migrate to different labor markets may have substantial impacts on optimal human capital accumulation decisions.[4]

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 Eckert and Kleineberg (2019) and Fogli and Guerrieri (2019), who develop general equilibrium models of residential and educational choice to study the effects of school finance policy and segregation on income mobility. Human capital levels in the former paper are binary, while locations are binary in the latter. Relative to these papers, I allow for a combination of

[4]Some empirical evidence of this can be found in the literature that studies international brain drain: Batista et al. (2012) find that increased emigration opportunities resulted in higher human capital investment in Cape Verde, and Shrestha (2017) and Spirovska (2020) find similar results in Nepal and Poland, respectively.

continuous human capital investment decisions 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. This enables my model to precisely capture how migration opportunities influence human capital investment decisions at multiple stages of the life cycle. To maintain tractability, however, I abstract away from general equilibrium concerns and conduct my exercises in partial equilibrium instead. In allowing agents to invest in their own human capital in addition to the human capital of their children, my model also speaks to the literature that studies optimal human capital development over the life cycle (Heckman et al., 1998; Huggett et al., 2011) through studying how geography influences optimal human capital attainment.

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* in influencing their later-life outcomes, much less has been done in assessing the importance of where (or whether) somebody *goes*. This may be in part because migration in the U.S. has been on a recent downward trend,[5] 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,[6] 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 specific 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 decision actually takes 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.[7] Furthermore,

[5]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).

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

[7]See Mountford (1997) for a theoretical treatment of this possibility in an international context. A closely

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) and families (Mincer, 1978; Gemici, 2006). 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 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 markets in the United States are vast, attempts to combine the two are much 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 and detail my estimation strategy. Section 5 evaluates the fit of the model, and Section 6 presents the results of counterfactual simulations. Finally, Section 7 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 differs substantially from the status quo. I now turn to the economic model I use to study these issues.

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.

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.[8] The following is a description of the events that transpire and the choices that agents make in each period:

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.

2. **Period 2:** After emerging from childhood with a level of human capital and ability, the agent invests in their own human capital and decides whether and where to move.

3. **Period 3:** The agent potentially marries and has children based on stochastic functions of their human capital stock and birth location. 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.

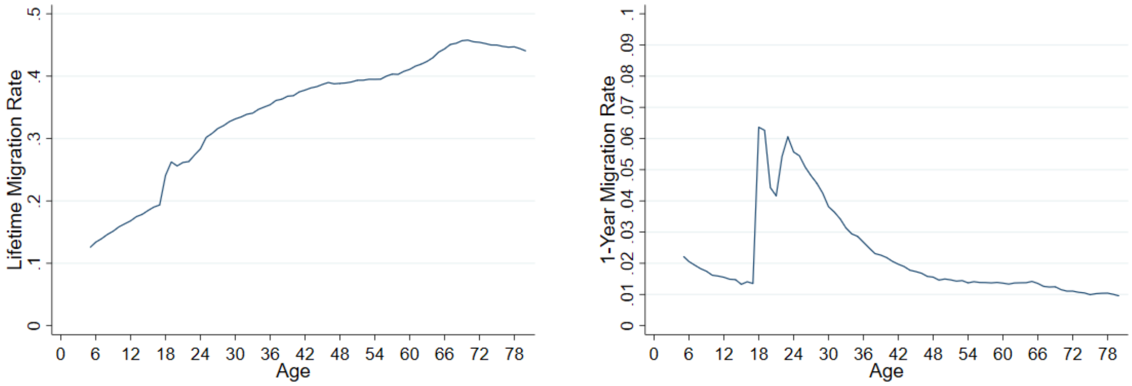
4. **Period 4:** The agent consumes the remainder of their resources and dies.

Locations (being the 50 states[9] in the U.S. and indexed by ℓ) differ in their costs of consumption/child inputs, family structure, levels of government child investment, and skill prices. The migration process in this model is suppressed to a single decision made in the second period of the model — while repeated moves are a salient feature of migration in the United States (Kennan and Walker, 2011), they are not crucial in addressing my research question, and abstracting away from dynamic migration considerably reduces computational burdens. Moreover, data suggests that this age range covered in the second period is by far

[8] This departs somewhat from typical human capital models and is justified in Appendix B.

[9] I focus on states instead of commuting zones 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.

Figure 2: Lifetime and Yearly Interstate Migration Rates by Age



(a) Lifetime

(b) Yearly

Notes: Data from U.S. natives in the 2001-2010 American Community Survey.

the most migratory period for the typical American — lifetime interstate migration rates see the sharpest increase in this period (Figure 2a), and 1-year migration rates are the highest in this age range by a considerable margin (Figure 2b).

Allowing for differences in marriage and fertility probabilities based on state of birth enables the model to capture the large differences in family structure across different locations 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 and their child's human capital with the knowledge that doing so will increase the odds of favorable realizations of marriage and offspring 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 parent(s), 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. 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 good 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 parameter ϕ will govern 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[10] s^ℓ and are then divided by the mean parent human capital level

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

$\exp\left(\mu_h + \frac{\sigma_h^2}{2}\right)$ to be converted to time units. The agent's ability multiplicatively alters the 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.

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)]:

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

where $n \in [0,1]$ is the measure of self-investment that the agent commits to in period 2 and could be thought of as a proxy for schooling. Human capital is also 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.[11] I additionally abstract away from an explicit college choice in my model, both because CHKS do not find college graduation rates to be a significant predictor of IIM and because other research indicates that wage differentials across education levels appear to be generated more by different human capital stocks than they do different human capital prices (Bowlus and Robinson, 2012).

I assume parent human capital to evolve exogenously:

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

The decision to allow exogenous evolution 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 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 from data, while κ and $\sigma_{\varepsilon_2}^2$ will be estimated via the simulated method of moments.

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

2.3 Recursive Formulation of Decisions

2.3.1 Period 2 — Independence:

As a newly independent adult, the agent solves a standard Ben-Porath problem with an added location decision that follows afterward:

$$V_2(a, h_2, \ell) = \max_n \{u(c_2) + \alpha \overline{S}^\ell + \beta \mathbb{E}[v_2(a, h_2, \ell, n)]\}, \quad s.t.\ p^\ell c_2 = e_2 = w^\ell h_2(1 - n).$$

Here e_t denotes earnings in period t, which itself depends on w^ℓ — the price of human capital in location ℓ — as well as the amount of time spent investing in one's own human capital as opposed to working. The agent optimizes their choice of n by weighing present consumption against their expected future happiness. In addition to consumption, agents derive utility from location amenities. To incorporate amenities into the model in an agnostic way,[12] I assume that locations that on average attract more people are more enjoyable to live in: $\overline{S}^\ell$ refers to the average share of natives across all states who choose to reside in state ℓ as adults, taken from observed migration flows in the American Community Survey. This essentially represents a revealed preference argument and borrows from insights in the quantitative spatial economics literature.

After their selection of self-investment, the agent chooses whether and where to move:

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

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

In addition to increasing earnings, the likelihood of marriage and children, and the expected ability of the agent's child a', higher human capital also decreases moving costs. Moving costs are parameterized as:

$$\Delta(h_3, \ell', \ell) = \Delta_1 - \Delta_2 h_3 - \Delta_3 C(\ell, \ell').$$

[12] One challenge with amenities is the substantial room for disagreement about what amenities are — those who enjoy coastlines and warm weather may wish to move to Florida, but Idaho becomes far more attractive if one prefers air quality and low crime instead.

Thus, moving costs contain a fixed cost of moving and a variable component that makes moving less costly for agents with higher human capital stocks. 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.

Having moving costs be decreasing in human capital allows the model to be consistent with empirical observations that more educated individuals tend to be more geographically mobile (Diamond, 2016; Kennan, 2020). While including human capital 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 better at growing their human capital stock. The chosen parameterization, 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 has a causal impact on migration costs is also not unreasonable: the process of human capital accumulation may make agents more open to experiencing other cultures, and more skilled individuals may face smaller migration frictions through being better able to find jobs in other labor markets.

While some individuals may primarily base their migration decisions on pecuniary concerns, other work (Kennan and Walker, 2011) suggests that a substantial number of people move (or do not move) across states for reasons that are entirely orthogonal to money. To account for this, I include utility shocks $\zeta_{\ell'}$ associated with moving to a particular location that are drawn from the Type I Extreme Value distribution with location 0 and scale parameter σ_ζ, which itself will be estimated. 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:

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

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

Finally, the probabilities of marriage and fertility are assumed to be stochastic functions of one's human capital stock and place of birth that follow a probit process — with $m = 1$ indicating the agent being married and $f = 1$ indicating the agent having a child in the upcoming period, I denote:

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

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

where $\Phi()$ is the standard normal CDF. Marriage realizations are drawn first, which in turn influence the probability of the agent having a child. So, agents may self-invest in part to improve their chances of marriage and fertility. I estimate separate coefficients for each state of birth and, in the case of fertility, each marital status, which allows my model to flexibly capture lingering effects of one's state of origin on their future family structure. Large cross-state differences in marital rates and rates of single parenthood over the education gradient suggest that such effects are important to consider.

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, denoted with 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:

$$V_3(h_3, \ell, 0, 1; a') = \max_{x,t} \left\{ (1+\theta)u(c_3) + \alpha \overline{S}^\ell + \beta \left[\int V_4(h_4, \ell) dF(\varepsilon_3) + \theta V_2(a', h_2', \ell) \right] \right\}$$

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

The final term of the value function reflects the parent's altruistic payoff from their child's next-period utility.[13] 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. If the single agent does not have a child, I assume them to

[13]The $(1+\theta)$ object in the first term of V_3 comes from the parent-child pair solving an optimal sharing problem over consumption, which also yields a constant that is here discarded.

inelastically supply labor before moving to the terminal period:

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

Marriage is assumed to be perfectly assortative in human capital levels, and a married parent differs from a single one only in that they additionally enjoy altruistic utility from the welfare of their spouse: denoting the married state as $m = 1$, we have:

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

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 by a monotonic factor, it does not affect optimal individual choices of x and t, so this specification effectively assumes that parents do not coordinate in child-raising decisions — as a result, children with married parents receive twice the human capital inputs than those with single parents *ceteris paribus*.[14]

2.3.3 Period 4 — Final Consumption and End of Life:

With their child having reached independence, the agent simply consumes their remaining resources and perishes:

$$V_4(h_4, \ell) = u(c_4) + \alpha \overline{S}^\ell; \;\; p^\ell c_4 = e_4 = w^\ell h_4.$$

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 5 points according to the equal-mass approach (Kennan, 2006). Policy functions for x and t are computed via grid search,

[14] The results of the model are robust to either discarding spousal altruism or allowing parental coordination in child-raising decisions. I choose this specification because it avoids concerns of consumption re-scaling and turns out to be fairly consistent with data on time inputs received by children with one vs. two parents (see Table 4c). For a (much) more sophisticated treatment of these issues, refer to Gayle et al. (2014)

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 bilinear interpolation over the human capital and ability state variables.

3 Data

The three sources of data I primarily use to estimate the model are the Panel Study of Income Dynamics (PSID), the PSID Child Development Supplement (CDS), and the 2000 Decennial Census and American Community Survey (ACS) from Census Bureau and IPUMS (Ruggles et al., 2020).

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 GDP PCE deflator. 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.

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.

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 2010-2016 waves of the American Community Survey and limit my sample to household heads 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 birth location, state-level native outflow 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 observed in these data during estimation. I use a comparable sample from the 5% 2000 Census to obtain distributions of parent human capital and marriage at the state level used to form the initial condition of the model.

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 from Data

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%. Cost of living levels p^ℓ are obtained from the American Chamber of Commerce Research Association's Cost of Living Index.[15] All values are normalized by the value of p^ℓ corresponding to Iowa.[16] Migration flows $\bar{S}^\ell$ are obtained from migration flows from birth states observed in the ACS. 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 2000-2001 Financial Survey and follow Lafortune et al. (2018) in cleaning the data.

Skill Prices

The main simulation procedure will roughly attempt to reproduce the outcomes of the CHKS

[15]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.

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

Table 1: Parameters Estimated from Data

Parameter	Description	Value	Source
β	Discount rate	0.479	Literature; $0.96^{18} = 0.479$
p^ℓ	Costs of living	Various	ACCRA Cost of Living Index
$\overline{S}^\ell$	Average migration shares	Various	ACS
g^ℓ	Govt HC investment	Various	NCES 2000-2001 Financial Survey
s^ℓ	Student-teacher ratios	Various	NCES 2000-2001 Financial Survey
w^ℓ	Skill prices	Various	Mincer Regression on 2000 Census
$\gamma_0^\ell, \gamma_1^\ell, \gamma_2^\ell, \gamma_3^\ell$	Marriage probabilities	Various	Probit Model in ACS
$\lambda_0^{\ell,m}, \lambda_1^{\ell,m}, \lambda_2^{\ell,m}, \lambda_3^{\ell,m}$	Fertility probabilities	Various	Probit Model in ACS
μ_{ε_3}	Period 3 shock mean	0.03	PSID
σ_{ε_3}	Period 3 shock SD	0.20	PSID

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 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 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, note that for any individual in the parent stage of the model we have:

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

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^ℓ (location-specific skill prices) are known. Additionally, we have that

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

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 2011-2012 ACS. Year-2000 skill prices are used to form the parent initial condition, and year-2011/2012 skill prices are used to adjust child earnings when they reach the parent stage of the model.

2011 and 2012 were the years in which the incomes of the CHKS cohorts were measured, and computing skill prices for these years as well allows the model to account for changes in returns to human capital across locations that may have transpired following the Great Recession. For additional details on this procedure, refer to Appendix C.

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 birth state and human capital stock. With w^ℓ 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 natives from each U.S. state aged 36-54 and run the probit model:

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

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

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. When estimating fertility probabilities I limit my ACS sample further to individuals aged 36-45 to prevent underestimating fertility from including parents whose children 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.

Late Human Capital Shocks

Finally, μ_{ε_3} and σ_{ε_3} 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, 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 μ_{ε_3} 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 13 parameters to estimate via the Simulated Method of Moments (SMM). These parameters are collected as $\Theta = [\theta, \rho_{ha}, \mu_a, \sigma_a, \xi, \phi, \kappa, \sigma_{\varepsilon_2}, \alpha, \Delta_1, \Delta_2, \Delta_3, \sigma_\zeta]$.

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 child living in their household, after which I compute the distribution of human capital and the joint distribution between human capital and marital status in each state using the same methods as described above. Interpreting these figures as the distributions of parental characteristics for the CHKS cohorts, I then randomly draw 20,000[17] parents for each state, after which I draw the ability levels of their children and simulate their migration, marriage, and earnings outcomes later in life.

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, I find the point estimate $\hat{\Theta}$ by numerically solving:

$$\hat{\Theta} = \arg\min_{\Theta} \sum_{i=1}^{N} \left(\frac{M_i(\Theta) - M_i}{M_i} \right)^2$$

where $M_i(\Theta)$ are the simulated model moments. All moments are weighted equally except the correlation between marital rates and IIM at the state level. This correlation does not receive as much weight as other moments because CHKS find that children with married parents

[17]An alternate procedure would be to draw more parents for more populous states. This does not meaningfully change where natives of each state fall in the income distribution but increases computational load due to having to draw more individuals in order to obtain a reasonable number of people from the least populous states. When computing migration moments from the ACS, all states receive equal weight to be consistent with the simulation procedure.

Table 2: Parameters Estimated via SMM

Parameter	Value	Description	Targeted Moment
θ	0.543	Parental Altruism	Rank-rank IGE
ρ_{ha}	0.232	Ability persistence	IGE/Parent marriage corr
μ_a	0.690	Ability mean	Life-cycle earnings means
σ_a	0.312	Ability SD	Life-cycle earnings SDs
κ	0.383	Ben-Porath HC accumulation	Early % wage growth mean
σ_{ε_2}	0.348	Early HC shock SD	Early % wage growth variance
ξ	5.371	Child investment productivity	Young adult earnings mean
ϕ	0.803	Child HC Time Share	Time spent with children
α	0.305	Amenity preference	State-to-state migration flows
Δ_1	8.443	Moving cost, component 1	Out-migration rate mean
Δ_2	0.542	Moving cost HC component	Mover-stayer HC difference
Δ_3	2.993	Moving cost proximity component	Share to nearby states
σ_ζ	1.888	Location shock scale parameter	Out-migration rate SD

Notes: This table reports the estimated values of the model parameters in the second column, along with the moments most important in estimating specific model parameters in the fourth column. Ability mean μ_a reported in level as opposed to log.

still do poorly if they hail from areas with high rates of family instability, indicating that parent marriage rates influence child outcomes beyond the direct mechanism of single parents investing less in their own children. My model currently cannot account for these potential neighborhood or peer effects, so I instead use this correlation to discipline θ and ρ_{ha} while ensuring that the model hits the other targeted moments as well as possible. Minimization of the objective function proceeds via the Nelder-Mead downhill simplex routine.

4.3 Identification

While the model is jointly identified in general, a conceptual argument for identification is as follows. The altruism parameter θ and persistence of learning abilities ρ_{ha} are tied to moments to do with intergenerational persistence in income and are thus estimated by targeting the rank-rank intergenerational elasticity of family earnings of 0.34 as reported by CHKS. As mentioned above, the correlation between 25th-percentile parent marriage rates and child IIM of 0.61 (obtained from the Opportunity Atlas) is also targeted in order to separately identify the two parameters, because a single parent will have a higher level of human capital than married parents who have identical family earnings. A higher level of ρ_{ha} thus results in lower-ability children being born to married parents than single parents of the same income level, which will reduce the correlation of interest. Thus, θ and ρ_{ha} are effectively estimated by making the simulated correlation as large as possible while demanding that other targeted moments are hit.

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[18] 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.[19]

Finally, α, $\Delta_1, \Delta_2, \Delta_3$, and σ_ζ govern the dynamics and returns to migration in the model. From the ACS, I calculate rates of native outflow at the state level and state-to-state migration probabilities. I target the average state outflow rate of 0.43 to estimate Δ_1 and target the cross-state standard deviation of outflow rates of 0.09 to estimate σ_ζ. The parameter Δ_2 governs how quickly moving costs decline in human capital levels, so I estimate it by targeting the average gap in human capital between movers and stayers observed in my ACS sample. Leveraging the degree to which movers select to move to either adjacent states or states in the same Census division as their birth state allows for straightforward identification of Δ_3, and α is estimated by targeting the full set of state-to-state migration flows.[20]

[18]All targeted earnings moments are converted to real terms by adjusting for state-specific skill prices to account for possible oversampling of certain states in the PSID. All monetary units in the model are normalized by mean real individual PSID earnings of 37,483 2012 dollars.

[19]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.

[20]This set of moments is weighted so as to receive the same total weight as any other individual moment. Given that many state-to-state migration flows are quite close to zero, I compute level differences instead of percentage differences between my model moments and these moments.

Table 3: Model Fit — Targeted Moments

Moment	Source	Data	Model
Rank-rank IGE of family earnings	CHKS	0.34	0.34
Parent marriage / IIM correlation	Opp Atlas	0.61	0.36
Period 2 individual earnings mean	PSID	0.98	0.94
Period 2 individual earnings SD	PSID	0.46	0.40
Period 3 individual earnings mean	PSID	1.33	1.33
Period 3 individual earnings SD	PSID	0.77	0.80
Period 2-3 individual earnings % growth mean	PSID	0.27	0.27
Period 2-3 individual earnings % growth SD	PSID	0.37	0.36
Outflow rate mean	ACS	0.43	0.49
Outflow rate SD	ACS	0.09	0.10
Share moves to nearby states	ACS	0.41	0.42
Mover-stayer HC gap	ACS	0.20	0.20
Time spent with children	PSID CDS	0.18	0.18

I also target the overall mean and spread of average adult income rank of children born to parents in the 25th income percentile across states to reduce risks of overfitting. To supplement the arguments in support of identification, Figure D.1 plots the objective function's value as individual parameters are moved from their estimated value with all others held constant. The objective function reaches local minima at the estimated parameter values in all cases and increases monotonically (and generally convexly) as the guesses move further away from the estimates.

5 Model Fit

Having estimated the model, I now evaluate its performance.

5.1 Targeted Moments

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 value of ρ_{ha} is on the lower end of the values reported for the comparable parameter in

Kotera and Seshadri (2017), which likely stems from the estimation procedure driving the value down so as to maximize the correlation between parental marriage and child outcomes while maintaining the model's fit elsewhere. The estimation procedure does not send ρ_{ha} down to zero, as with sufficiently low levels of ρ_{ha} even higher levels of altruism θ are not adequate in enabling the model to reproduce the rank-rank IGE targeted in estimation.

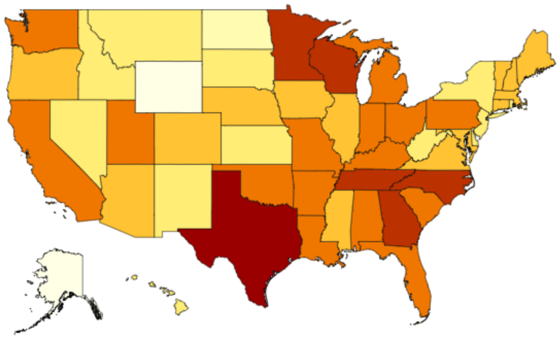
Unsurprisingly, the model is unable to produce the strength of the aforementioned correlation, but all other targeted aggregate moments are fit either exactly or almost exactly.[21] The model slightly underestimates the spread of earnings observed in the young-adult stage of the model but matches the adult earnings distribution quite well — additionally, the model is successful in capturing the intergenerational earnings elasticity observed in data. The model also slightly overstates the overall migration rate but matches the share of moves to nearby states and mover-stayer differences in human capital almost exactly.

The estimates of parameters governing moving costs suggest that moving to a nearby state is about one third less costly than moving to a non-nearby state. 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 a very important role in governing state-to-state flows.

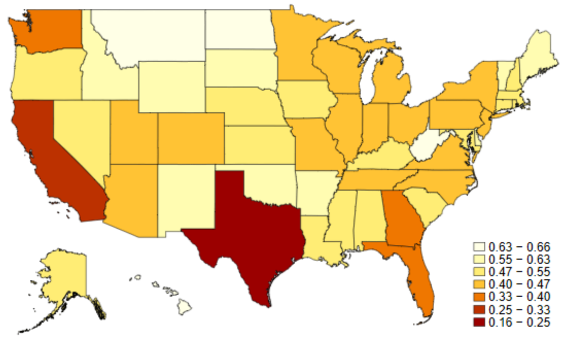
While the average native outflow rate is matched almost exactly by the model, Figure 3 evaluates the model's performance in matching individual state outflow rates and migration destination probabilities. Generally speaking, the model does well — the native outflow rates at the state level predicted by the model and observed in data are positively and significantly correlated (coefficient 0.64, indicating that the model can account for roughly two fifths of cross-state variation in native out-migration rates). 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. The most salient miss is that the model over-predicts migration out of the Southeast. The model does better in predicting

[21]The correlation of 0.36 I obtain can actually be argued to be at least somewhat reasonable for a model that does not currently account for peer/neighborhood effects. At the CZ level, CHKS report a correlation of -0.76 between rates of single parenthood and IIM and a correlation of -0.66 for the same variables for the subsample of their data born to married parents. The difference in R^2 of 0.142 between the two specifications could be loosely interpreted as the share of the variance in IIM across CZs explained by family structure influences that are unrelated to neighborhood effects, which would translate to a correlation of 0.38

Figure 3: Model Fit — State Outflow and Inflow Rates

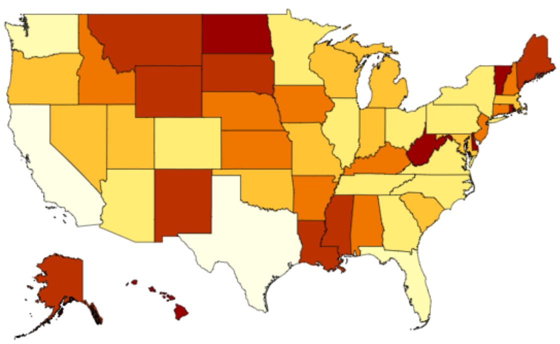


(a) Outflow, Data

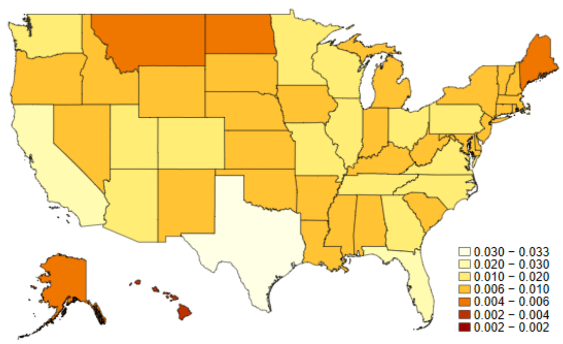


(b) Outflow, Model

Outflow Correlation: 0.64



(c) Inflow, Data



(d) Inflow, Model

Inflow Correlation: 0.74

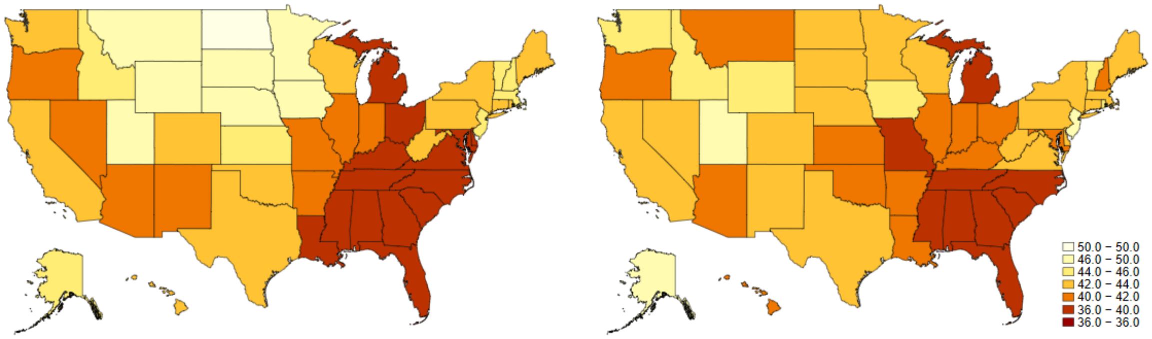
rates at which states receive migrants: it can account for over half the variation in cross-state inflow rates though struggles somewhat to reproduce the full extent of the spread observed in the data, instead allocating migrants more evenly across the country.

5.2 Untargeted Moments

As the main model validation test, I next evaluate how well my model reproduces the geography of intergenerational mobility in the United States as reported by CHKS, as this is not explicitly targeted in the estimation procedure. For every U.S. state, I simulate the outcomes of 20,000 children born to families in the 25th income percentile[22] and compute

[22]Since time investments and resultant earnings are subject to the random draw of the child's ability, it is difficult to impose that the simulated families here have income in exactly the national 25th percentile. As a workaround, I instead simulate families with the mean human capital levels for either single or married families in the 25th national income percentile in the relevant state. In practice, the large majority of parents

Figure 4: Model Fit — Upward Mobility by State



(a) Data (b) Model

Correlation: 0.64

Notes: Upward mobility measured as the expected family national income percentile of children born to parents in the 25th national income percentile.

the expectation of their income rank when they enter the parent stage of the model. The proportion of children who are born to married or unmarried parents in every state is taken directly from the Opportunity Atlas, but the marriage outcomes of children are left to the same stochastic process as before.

Figure 4 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 0.64, indicating that my model can account for two fifths of the state-level variation in income mobility observed in the data. The model fits states in the Northeast, Rust Belt, and the South well, but there is a downward mean shift in predicted income mobility compared to data for states west of Illinois: in general, the model cannot quite capture the full extent of IIM observed in the Midwest and the Mountain States.

Nonetheless, the model does predict states in these regions to be quite income-mobile relative to other locations. While the inclusion of marriage allows the model to attain a respectable fit for states such as Utah (which has poor measures for public school quality but exceptionally high rates of marriage for parents and children), further extending the model to consider neighborhood effects or racial disparities may be helpful in bolstering the predicted income mobility in more remote parts of the country.

simulated this way are in the 25th percentile, and all of them are in the 24-26 range.

Table 4: Additional Untargeted Moments

Statistic	Data	Model
Corr w/ Baseline	0.94	0.81

(a) IIM, Stayer Sub-Sample

	Parent Quintile				
Child Quintile	1	2	3	4	5
1	34/35%	24/25%	18/18%	13/13%	11/9%
2	28/25%	24/23%	20/21%	16/18%	12/14%
3	18/19%	22/21%	22/21%	21/20%	17/18%
4	12/14%	18/18%	22/21%	24/23%	24/23%
5	8/7%	12/12%	18/19%	25/26%	37/37%

(b) Income Quintile Transtions (Data/Model)

	Married		Unmarried	
Statistic	Data	Model	Data	Model
Total Inputs	0.19	0.20	0.10	0.10
Individual Parent Inputs	0.10	0.10	—	—

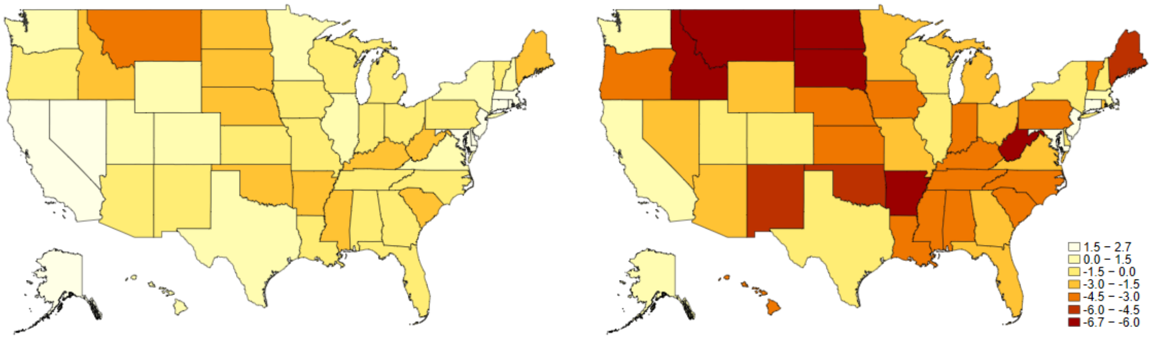
(c) Time Investments

Notes: Table 4a reports the correlation of state-level IIM between the baseline sample and the subsample of individuals who did not move. Data moments are at the CZ level and are reported in Appendix Table VII of CHKS. Table 4b reports income quintile transition probabilities between parents and children. Data moments from Table II of CHKS. Table 4c 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.

Table 4 reports the model's fit for additional untargeted data moments. Notably, the model can reproduce the finding in CHKS that the geographic variation in IIM in the sub-sample of individuals who stay in their state of birth is strongly correlated with the overall estimates. As reported in Table 4a, the correlation of state-level IIM between the stayer and overall samples is 0.81, compared to 0.94 as reported by CHKS.[23] Lastly, Tables 4b and 4c report the model's ability to reproduce rates of parent-child income quintile transitions and differences in time investments received by the children of single and married parents. The model's fit in both of these categories is reasonable.

[23]These moments are not directly comparable in that CHKS report results at the commuting zone level, but the finding is encouraging nonetheless.

Figure 5: Counterfactual — Results of Migration Restrictions



(a) Change (no BR) (b) Change (with BR)

Table 5: Migration Restriction Impacts by State Group

Statistic	West North Central/Mountain	Not WNC/MO	Δ
IIM (CHKS measure)	45.55	41.10	4.45
IIM (Model Baseline)	42.85	41.88	0.97
IIM Δ (no BR)	-0.88	0.09	0.99
IIM Δ (with BR)	-3.66	-2.12	1.54

Notes: BR = Behavioral responses. Figure 5 plots the change in upward mobility from counterfactuals that restrict migration while ignoring or including behavioral responses. 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 behavioral responses in human capital investment decisions. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile. See Appendix A for division definitions.

6 Counterfactual Exercises

I now use the model to evaluate the effects of three sets of counterfactual policies. Experiments that evaluate the impact of a counterfactual on IIM do so by comparing the average 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.

6.1 Migration Restrictions

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 $\Delta_1 = \infty$, 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: 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 5 displays the geographic distribution of the changes in IIM induced by these counterfactuals. 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 Montana native from moving later in life would result in their expected income percentile rank as an adult dropping by over 6 points; on the other hand, doing the same to a child from New Jersey would *increase* their expected rank by about 3 points. 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 shifts by approximately 0.99 points when ignoring the behavioral response and 1.54 points when including it. Thus, behavioral responses to migration opportunities disproportionately spur human capital accumulation in more rural parts of the country and are important to consider when evaluating the effects of migration restrictions. The shift of 1.54 points with behavioral responses more than covers the gap in income mobility between the two groups of states predicted by the model, but there is no reason *a priori* to expect that the size of the effect would scale up if the model were to fit IIM levels in the West North Central and Mountain divisions better. Since the gap in IIM between the two groups is approximately 4.45 points as reported by CHKS, a conservative interpretation of the results is that roughly one third of the advantage that these areas enjoy in measures of upward mobility may be attributed to migration channels, and behavioral responses play an important role in generating this result.

6.2 Effects of Local Economic Shocks

Next, I investigate how local shocks to economic conditions impact the expected earnings of natives. I do this by increasing w^ℓ by 0.10 points[24] for each state, one at a time, before resolving the model and resimulating outcomes for 20,000 children born to 25th-percentile parents in the state of interest. Conceptually, simulating a local labor market improvement by increasing the human capital rental rate of an area induces dual effects on the future earnings of natives: stronger labor markets improve the outcomes of stayers but also depress incentives for agents to leave and potentially earn more elsewhere. These competing effects may be particularly at odds in areas with lower human capital rental rates and high native outflow.

Figure 6 displays the level changes in adult family earnings[25] induced by this counterfactual for each state, while Table 6 reports average state-level results at the Census division level. The increase in local skill prices results in the average state's lifetime migration rate for poor children falling by approximately 0.04 points, from 0.48 to 0.44. However, the results of reduced migration on native earnings is ambiguous and depends on the characteristics of the starting location — as Figure 6 shows, the impacts of these economic shocks vary widely by location. While the increases in skill prices never result in lower average outcomes in any state, their effect on expected family earnings levels later in life for poor children range from

[24] This would correspond to an increase of about $3,750 for an individual who spends all their time working. Note that this is a level increase as opposed to a percentage increase in the skill price, as a 10% increase in the skill price would have greater or smaller earnings implications depending on the base skill price of the state in which it occurred.

[25] Note that the increase of w^ℓ of 0.1 points can increase adult earnings by more than 0.1 because adult earnings are measured at the family as opposed to individual level.

Figure 6: Counterfactual — Wage Shocks

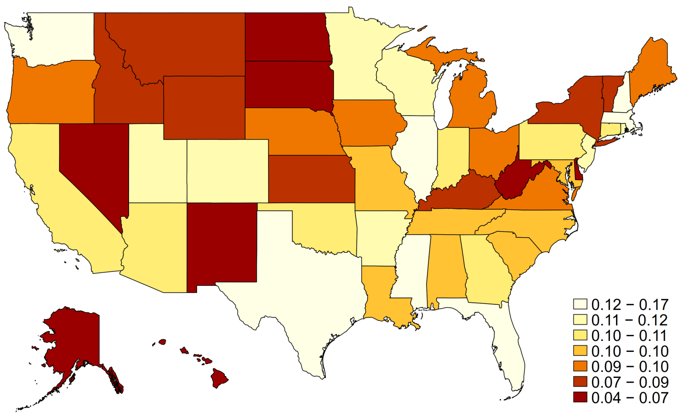


Table 6: Δ by Division

Division	Mean Δ
NE	.113
MA	.105
ENC	.101
WNC	.084
SA	.100
ESC	.109
WSC	.124
MO	.093
PA	.087

Notes: Figure 6 displays state-level changes in period-3 native family earnings following a counterfactual that increases the skill price in each state by 0.1 points, while Table 6 reports mean effects across states in each Census division. Earnings normalized by 37,483 2012 dollars. See Appendix A for division definitions.

an increase of barely 0.04 (roughly $1,600) to 0.17 depending on the state in which they occur (for instance, the results suggest that an increase in skill prices in North Dakota would have less than one fourth of the effectiveness in increasing earnings for natives compared to such a change happening in Texas). Many of the states in the Midwest and Mountain areas appear to have the least to gain from such economic windfalls, along with some states in the Appalachian area and many of the more rural states in the Northeast. These results suggest that migration is an important consideration when both designing and evaluating policies that impact specific locations more than others. These results may also provide some motivation for the ambiguous relationships observed between labor market conditions and income mobility observed in data.

6.3 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.[26] 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.[27]

Despite college not being modeled explicitly in this paper, my model can still provide some insight into the likely efficacy 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 adults and who have sufficiently high levels of human capital.[28] 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 "college share" in the state after the introduction of the counterfactual as well as the percentage change in each state's net revenue.[29]

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 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

[26]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.

[27]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. I take no stand on this argument here and instead focus on whether such a policy may be a good idea from an individual state's point of view.

[28]This level of human capital is set to 1.71, which is the average level of human capital observed among college graduates aged 36-54 in 2010-2016 American Community Survey. Distributionally, this level of human capital is at around the 70th percentile.

[29]Revenue here is computed crudely by assuming a 5 percent income tax rate on an individual's period 3-4 earnings that does not distort labor supply incentives. Losses come from the states having to pay out the subsidies to qualifying individuals. I have estimated a version of the model where agent earnings are taxed according to the sum of federal and state average tax rates in 2000 as calculated by the NBER TAXSIM model (see http://users.nber.org/ taxsim/allyup/). This modification does not substantively change the results of the paper.

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	0.11	-1.09	0.22	-2.23	0.69	-5.35
MA	0.04	-1.32	0.16	-2.49	0.55	-5.79
ENC	0.09	-1.45	0.24	-2.72	0.53	-7.12
WNC	0.10	-1.53	0.29	-2.83	0.62	-7.36
SA	0.11	-1.36	0.21	-2.71	0.55	-6.92
ESC	0.06	-1.70	0.18	-3.28	0.48	-8.23
WSC	0.01	-1.74	0.13	-3.17	0.50	-7.55
MO	0.11	-1.37	0.22	-2.75	0.69	-6.73
PA	0.19	-0.89	0.42	-1.75	0.91	-5.11

Notes: Table presents results of counterfactual policy that subsidizes individuals to live in specific states conditional on achieving a human capital level of at least 1.71 (corresponding to approximately the 70th percentile). Revenue computed as 5 percent of individual period 3-4 earnings minus subsidy payouts. Estimates account for a 1.9% spillover effect of a 1 percentage point increase in college share on individual earnings. Results summarized at the divisional level; see Appendix A for division definitions.

externalities of the presence of college graduates on the earnings of others, I allow for the earnings of all individuals in a state to increase by 1.9% following a 1 percentage point increase in that state's college share.[30]

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 highly skilled individuals to the policy is generally small — even in the counterfactual that offers a $50,000 subsidy, the typical state sees less than a 1 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. As a result, the majority of agents with high human capital are not sufficiently close to migration margins to respond to the policy, and the overwhelming majority[31] of the subsidies go to individuals who choose to locate in the given state in the baseline model without the subsidy, which in turn renders the policy highly cost-ineffective.

[30]This is the upper bound of spillover effects estimated by Moretti (2004). Agents are assumed to be unaware of these externalities when making migration decisions.

[31]99.2%, 98.5%, and 96.4%, respectively, for the three policies.

7 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 third 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 self-invest partly in anticipation of leaving them, which in turn has important implications for the incidence of local economic shocks. Since migration opportunities 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.

The main limitations of the model come from the combination of continuous human capital and numerous locations forcing the decisions the agents make as well as the life cycle to be compressed to maintain computational tractability. A model that included multiple stages of childhood could consider how the effects of location on human capital development differ over stages of child development. Allowing multiple migration decisions would enable the model to capture the possibility of an agent 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. These issues may offer promising avenues for future research.

References

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

Black, D., N. Kolesnikova, and L. Taylor (2009). Earnings functions when wages and prices vary by location. *Journal of Labor Economics 27*(1), 21–47.

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

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, M. R. Jones, and S. R. Porter (2020). Race and Economic Opportunity in the United States: An Intergenerational Perspective. *The Quarterly Journal of Economics 135*(2), 711–783.

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.

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

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

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

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.

Eckert, F. and T. Kleineberg (2019). Can we save the American Dream? A Dynamic General Equilibrium Analysis of the Effects of School Financing on Local Opportunities.

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.

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.

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

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

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.

Kotera, T. and A. Seshadri (2017). Educational policy and intergenerational mobility. *Review of Economic Dynamics 25*, 187–207.

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. 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. (2020). The effect of migration opportunities on college enrollment and college major choice.

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, Montaha, Nevada, New Mexico, Utah, Wyoming.
- **Pacific (PA):** Alaska, California, Hawaii, Oregon, Washington.

B The Choice of Utility Function

The standard utility function over consumption in human capital theory is CRRA utility $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$, with γ typically around 2. This is problematic in a migration model because of its predictions to do with the geographic sorting of individuals across different levels of human capital. To illustrate this, consider a simple two-location choice problem faced by an agent with human capital stock h. The price of human capital in their current location is 1, and the price of human capital in the other location is $w > 1$. Ignoring moving costs and preference shocks, the value of migrating to the agent is then

$$v = \frac{(wh)^{1-\gamma}}{1-\gamma} - \frac{h^{1-\gamma}}{1-\gamma}.$$

The trouble arises when considering the derivative of v with respect to h:

$$\frac{\partial v}{\partial h} = \frac{1}{h^{\gamma}} \left(\frac{1}{w^{\gamma-1}} - 1 \right);$$

$$\frac{\partial v}{\partial h} < 0 \Longleftrightarrow \gamma > 1.$$

In words, $\gamma > 1$ implies that individuals with lower human capital stocks respond more strongly to financial incentives than individuals with higher human capital stocks when making migration choices, which is roundly rejected in data (see for instance Diamond (2016)). While one can make migrating less costly for individuals with higher human capital stocks, a $\gamma > 1$ will still result in geographic sorting patterns among highly skilled individuals that do not match well-established trends in data whatsoever. I thus set $\gamma = 1$ (yielding log utility) so that the value of migration does not change with one's human capital stock while allowing migration to be less costly for individuals with higher stocks. This modeling decision can be thought of as something of a midpoint between standard human capital theory utility function specifications and those often seen in migration models (utility is linear in income in Kennan and Walker (2011), for instance).

C Additional Details on Skill Price Estimation

This section presents additional details on the Mincer wage regressions used to obtain skill prices w^ℓ from the 2000 Decennial Census and the 2011-2012 American Community Surveys.

When estimating skill prices, I restrict my sample further following Eckert and Kleineberg (2019). I limit my sample to individuals aged 25 to 55 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 agents 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. In order to reduce concerns about selection into migration by individuals with high human capital levels biasing my estimates, I restrict my sample to individuals who have exactly 12 years of education.[32] I then run the regression:

$$\log(w_{it}) = \beta_0 + \beta_1 \mathbf{X}_{it} + \beta_3 x_{it} + \beta_4 x_{it}^2 + \beta_5 x_{it}^3 + \beta_6 x_{it}^4 + \mathbf{D_{it}} + \varepsilon_{it},$$

where w_{it} is hourly wage (mapping to $\frac{e_3}{1-t}$ in equation (1)) and $\mathbf{X_{it}}$ is a vector of demographic characteristics (black, male, and hispanic dummies) included to account for compositional differences across states.[33] This vector together with a quartic polynomial in years of potential experience serve a 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 2011-2012) and are what allow me to derive skill prices, computed as $w_t^\ell = \exp(D_{\ell,t})$. I omit the $\mathbf{D_{it}}$ dummy corresponding to Iowa in 2000 in the regression as a normalization.

Figure C.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

[32]Including other education levels returns similar skill price estimates but does require the assumption that returns to education are equivalent across states after stripping out state fixed effects on wages. A related but model-inconsistent exercise would be to drop the fixed effects and instead estimate state-specific returns to education in the spirit of Black et al. (2009). These state-specific returns, after normalizing by Iowa's return, are also quite similar to the values I use.

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

that are lacking in large cities, with particularly low returns for states in the Great Plains and 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.[34]

I conduct two additional tests to assess the robustness of my skill price estimates. 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 >0.7): in particular, the year-2000 skill prices obtained from this method are virtually identical to the baseline estimates. The key difference comes from the 2011-2012 ACS, where the new entrant sample restriction results in very small cell sizes and considerably more variable estimates of w^ℓ for low-population states.

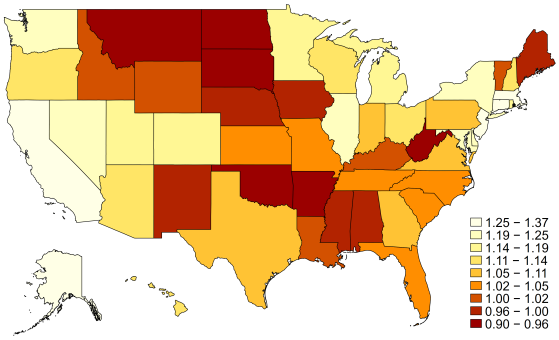
Second, I use the method described by Dahl (2002), where the selection correction takes the form of a function of the first-best probability of location choices. In particular, individuals are categorized into cells based on observable characteristics, and the probability that individuals in a given cell move from location j to location k is calculated in order to obtain an estimate of the selection probability. The first-best probability is then included in the regression using a flexible functional form, such as a polynomial approximation. I group individuals into cells based on their birth state, marital status (single, married with working spouse, married with non-working spouse), 5-year age bin, presence of any children, sex, and race (non-Hispanic white or not) and drop any cells with fewer than 50 observations. I then include a quartic polynomial of the first-best probability of choosing a location for one's cell in the Mincer regression. This method also does not meaningfully change my estimates: the correlation between the skill prices obtained via this procedure and the ones I use is greater than 0.9.

Given that I estimate skill prices for 2000 and 2011-2012, a natural question that follows is how I should allow agents to respond to future changes in human capital rental rates in the model when making migration decisions. Full foresight is impossible, as doing so would entail knowledge of the infinite horizon of human capital rental rates across locations, which

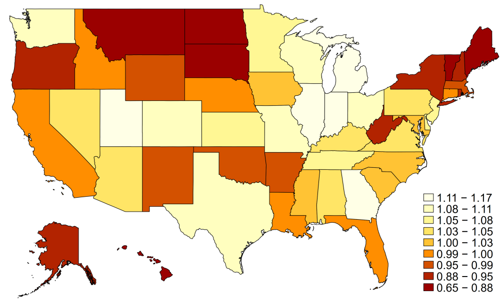
[34] The figures suggest that agents could as much as double their real earnings by moving from the lowest to highest ranked state. This is somewhat misleading as it is driven entirely by Hawaii, where costs of living are so high that the real skill price is adjusted to be very low. Moving from the second-lowest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

is unknown to the researcher. I have estimated a version of the model with limited foresight, in which the initial parents do not take next-period changes in state-specific parameters (including costs of living, skill prices, and school characteristics) into account, but the children do when making their self-investment and migration decisions. The paper's main results are robust to this specification, but the model's fit regarding migration moments is slightly worse. The largest changes in skill prices between 2000 and 2012 likely came from the Great Recession, which transpired when the CHKS cohorts were in their late 20s. Given that this is slightly after the most migratory years of the life cycle, these cohorts likely made a sizable portion of their migration decisions before the Great Recession occurred. The specification of the model presented here thus imposes that agents have no foresight of changes in real wages or other economic characteristics across locations when making early-life migration decisions.

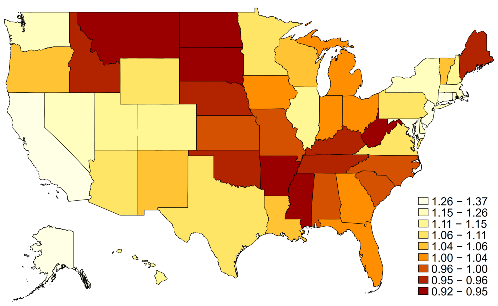
Figure C.1: Skill Prices



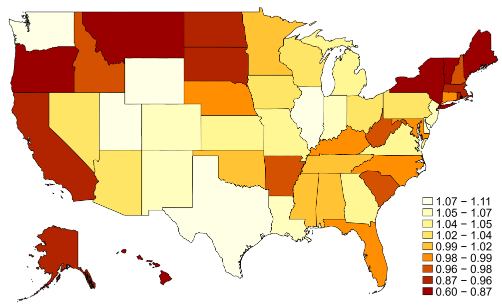
(a) 2000, Raw



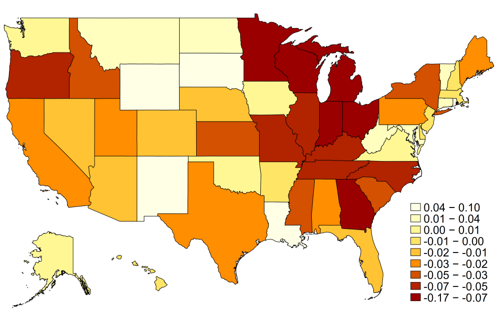
(b) 2000, Real



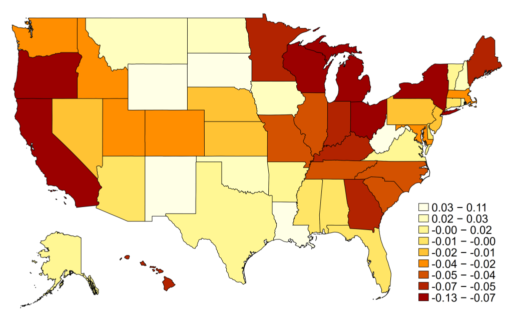
(c) 2011/2012, Raw



(d) 2011/2012, Real



(e) 2000-2011/2012 Change, Raw



(f) 2000-2011/2012 Change, Real

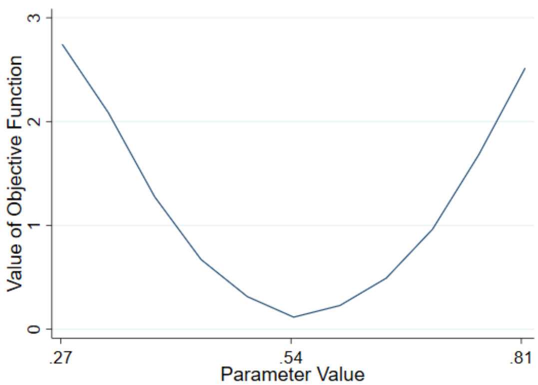
D Supplementary Figures and Tables

Table D.1: OLS Estimates for Various Correlates on CZ-Level IIM

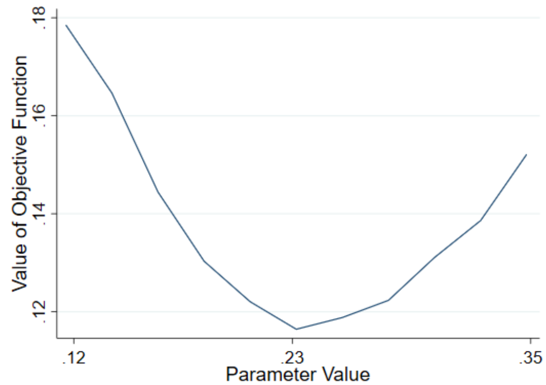
VARIABLES	(1) IIM	(2) IIM	(3) IIM	(4) IIM
Native Outflow				0.139
				(0.0143)
Student-Teacher Ratio			-0.335	-0.134
			(0.0550)	(0.0544)
LFP Rate		-0.119	-0.0889	-0.0625
		(0.0241)	(0.0243)	(0.0234)
Share Single Mothers	-0.409	-0.466	-0.459	-0.490
	(0.0500)	(0.0510)	(0.0511)	(0.0538)
Constant	61.06	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.

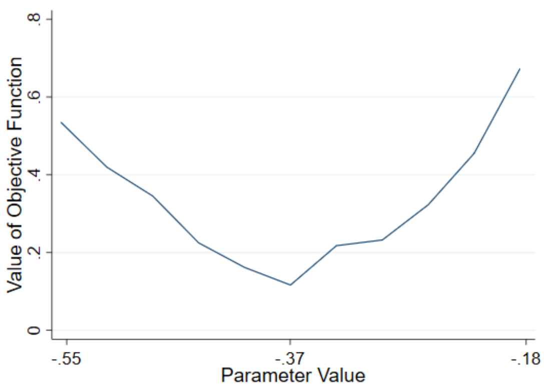
Figure D.1: Behavior of Objective Function



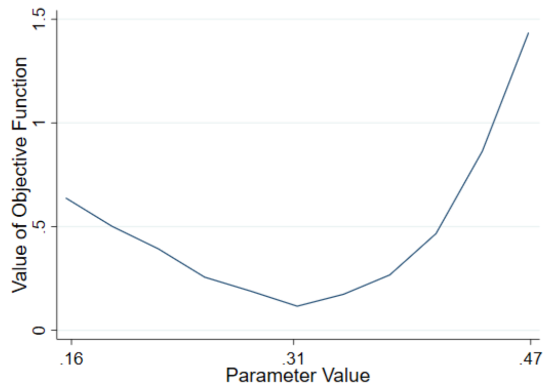
(a) θ



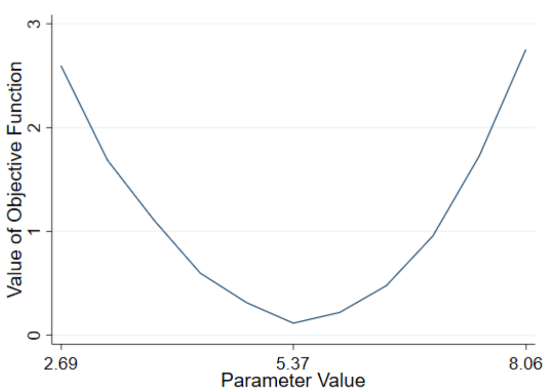
(b) ρ_{ha}



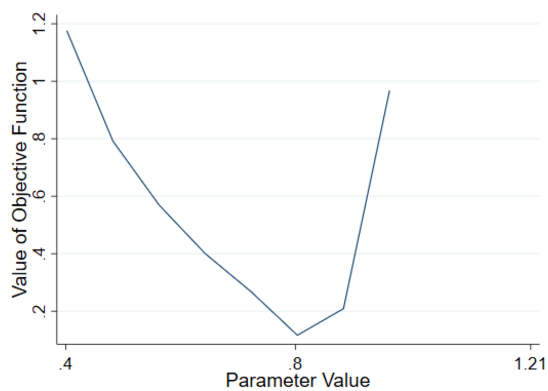
(c) μ_a



(d) σ_a



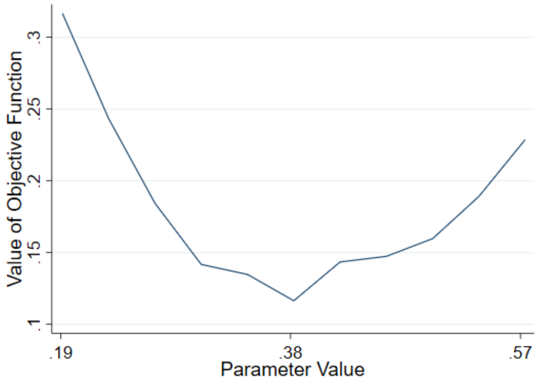
(e) ξ



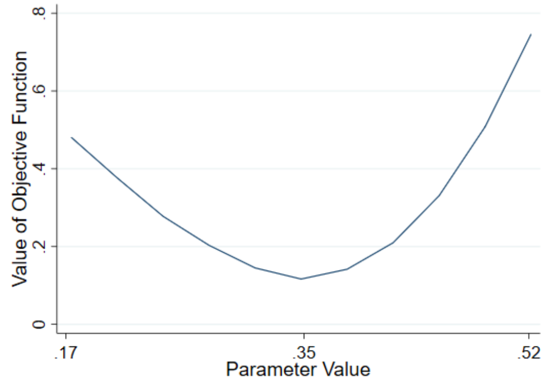
(f) ϕ

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

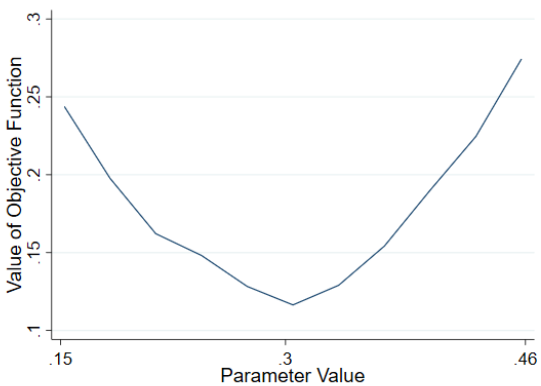
Figure D.1: Behavior of Objective Function (continued)



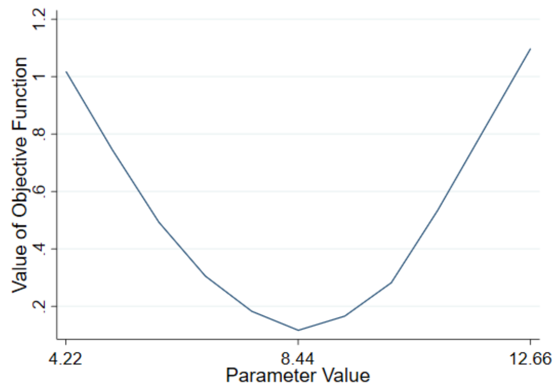
(g) κ



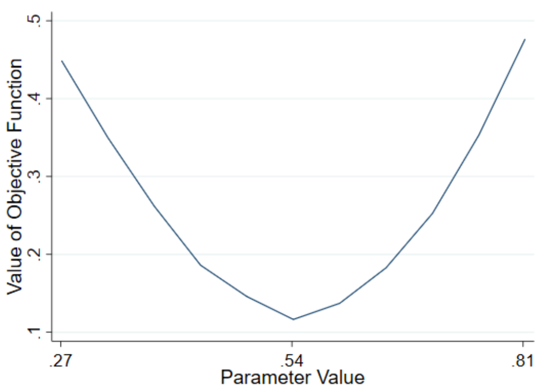
(h) σ_{ε_2}



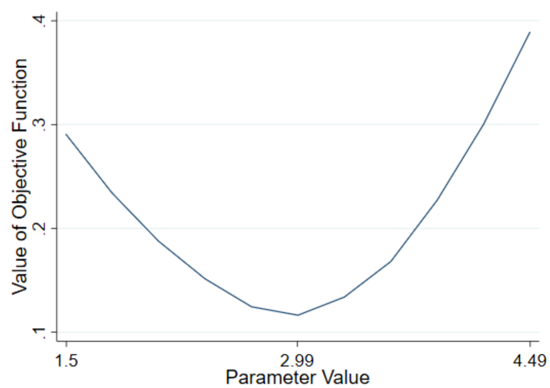
(i) α



(j) Δ_1



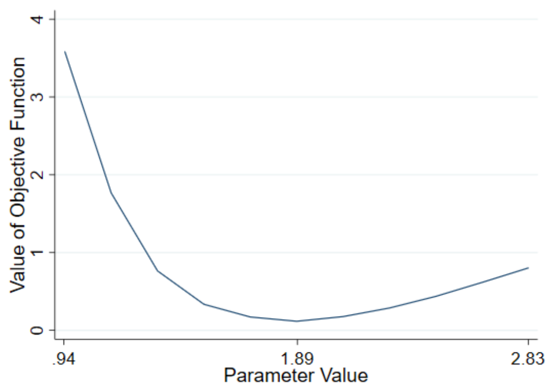
(k) Δ_2



(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 D.1: Behavior of Objective Function (continued)



(m) σ_ζ

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