Essays on Human Capital, Geography, and the Family

by

Garrett Anstreicher

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Economics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2023

Date of final oral examination: 04/03/2023

The dissertation is approved by the following members of the Final Oral Committee:

John Kennan, Professor, Economics
Christopher Taber, Professor, Economics
Jeffrey Smith, Professor Economics
Dean Corbae, Professor, Economics and Finance

To Rowan.

Acknowledgements

I am deeply grateful to my advisor, John Kennan, whose insights and patience has led to my substantial growth as an economist. Chris Taber, Jeff Smith, Dean Corbae, Ananth Seshadri, and many other faculty have been very generous with their time and feedback over the course of my graduate study. Martha Bailey, my first boss as a predoctoral research assistant, has been magnanimous in sharing advice and data, which along with the skills I learned under her have had a considerable effect on my career. Thanks are also due to the Economic department's support staff, with special thanks to Kim Grocholski for answering innumerable questions from me over the past six years. Bob Thomas and Charles Fiss have also been invaluable resources in applying for and working with restricted-use data.

My dissertation's focus on human capital and the family make it particularly apt to acknowledge my parents, Kurt Anstreicher and Jane Van Voorhis, in the incredible job they have done in developing my own human capital and supporting me in all the chapters of my life, including during some very difficult times. It is my ardent hope to provide to my offspring a childhood remotely comparable to the one they gave to me. Among extended family, Helene Donta deserves a very special thanks for providing childcare throughout the spring of 2021, without which much of this work would have been unlikely to be completed.

The cohorts above my own have been remarkable in their willingness to provide feedback on my work for the past six years. Among them, thanks to Moshi Alam, Elan Segarra, Shoya Ishimaru, Dennis McWeeny, Amrita Kulka, Sandra Spirovska, Arpita Patnaik and Nico Badarraco for much helpful feedback. Special thanks go to Joel McMurry and Joanna

Venator for their consistent mentorship, friendship, and in the latter's case co-authorship.

The University of Wisconsin-Madison Department of Economics has been exceptionally generous in their financial support to me, giving me two years and three summers of fellowship support. I additionally acknowledge the National Science Foundation for three years of financial support. The U.S. Census Bureau has also been pivotal in allowing access to the data used in the first chapter of this dissertation.

Lastly, I would like to acknowledge the enormous support provided by Sophie Donta throughout my entire Ph.D. From financial support to copy-editing my papers, she has done at least as much to make the contents of the following pages possible as anyone else.

Disclaimers

The work in this dissertation was produced with support from the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1747503. This dissertation comes with the disclaimer that 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. Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1849. (CBDRB-FY22-P1849-R9537, CBDRB-FY22-P1849-R9383, and CBDRB-FY23-P1849-R10122).

Abstract

In this dissertation, I study the interplay of familial and geographic factors in influencing human capital development and economic mobility in the United States.

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

The second chapter, joint with Joanna Venator, studies how childcare costs, the location of extended family, and fertility events influence both the labor force attachment and labor mobility of women in the United States. We begin by empirically documenting strong patterns of women returning to their home locations in anticipation of fertility events, indicating that the desire for intergenerational time transfers is an important motivator of home migration. Moreover, women who reside in their parent's location experience a substantial long-run reduction in their child earnings penalty. Next, we build a dynamic model of labor force participation and migration to assess the incidence of counterfactual scenarios

and childcare policies. We find that childcare subsidies increase lifetime earnings and labor mobility for women, with particularly strong effects for women who are ever single mothers and Blacks. Ignoring migration understates these benefits by a meaningful extent.

The third chapter, joint with Owen Thompson and Jason Fletcher, studies the long-run impacts of court-ordered desegregation. Court ordered desegregation plans were implemented in hundreds of US school districts nationwide from the 1960s through the 1980s, and were arguably the most substantive national attempt to improve educational access for African American children in modern American history. Using large Census samples that are linked to Social Security records containing county of birth, we implement event studies that estimate the long run effects of exposure to desegregation orders on human capital and labor market outcomes. We find that African Americans who were relatively young when a desegregation order was implemented in their county of birth, and therefore had more exposure to integrated schools, experienced large improvements in adult human capital and labor market outcomes relative to Blacks who were older when a court order was locally implemented. There are no comparable changes in outcomes among whites in counties undergoing an order, or among Blacks who were beyond school ages when a local order was implemented. These effects are strongly concentrated in the South, with largely null findings in other regions. Our data and methodology provide the most comprehensive national assessment to date on the impacts of court ordered desegregation, and strongly indicate that these policies were in fact highly effective at improving the long run socioeconomic outcomes of many Black students.

Contents

Chapter 1

Spatial Influences in Upward Mobility

Chapter Summary

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

1.1 Introduction

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

However, the literature has predominantly focused on the importance of where somebody is *from* in influencing their later-life outcomes as opposed to where or whether they *go*. The same areas that appear to feature high levels of economic mobility (see Figure 1.1a for a visualization) also exhibit high rates of geographic mobility, or native children migrating elsewhere later in life (Figure 1.1b).[3] Migration into higher-wage locations may be important in explaining the relative success of children from these rural areas. Moreover, the opportunity to migrate in the future may provide an important incentive for human capital accumulation in places where local labor market opportunities are scarce (Becker, 1994).

The goal of this paper is to study the role of migration and migration opportunities in influencing human capital investment decisions and income mobility in the United States.

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

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

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

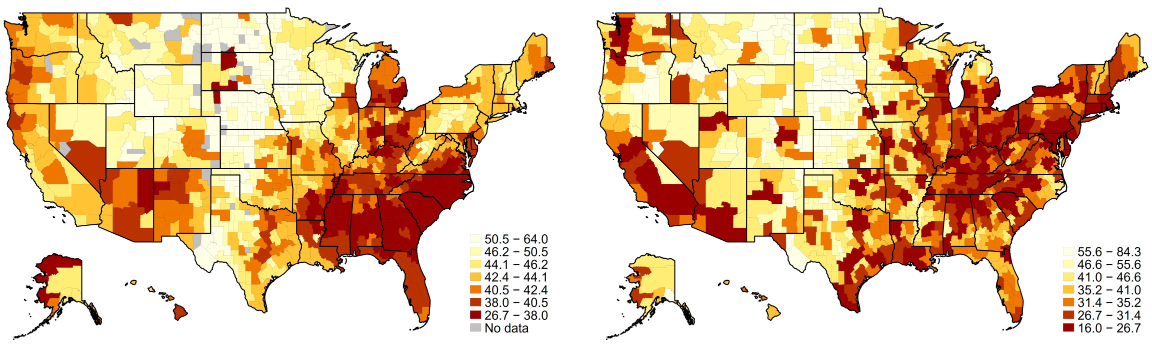
Investigating this relationship with data alone is challenging, both because of a lack of exogenous variation in people's ability to move within the U.S. and due to potential behavioral responses that would be difficult to capture empirically — that is, the option of migration in the future influencing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

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

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

[4]I.e. the Great Plains and Mountain States. See Appendix 1.A for exact Census division definitions.

Figure 1.1: IIM and Native Outflow in U.S. Commuting Zones



(a) IIM (b) Native Outflow

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

moving options would understate this result by 50 percent.

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

While the intranational brain drain I document can be beneficial to individuals from low-wage states, many of these states have considered policies intended to reduce their outflow of talent. As an additional exercise, I consider a policy that attempts to increase a state's retention of individuals with a college degree through offering them cash transfers. I find that the offer of such payments typically does not elicit changes in migration behavior —

as a result, the vast majority of these subsidies go to individuals who would have already chosen to live in the given state in the baseline world, and the policies would thus likely fail to be cost-effective. Finally, I find that equalizing public school characteristics across states does substantially more to reduce cross-state inequality in IIM than equalizing college tuition prices.

Related Literature

A vast literature exists on IIM and child human capital development (Todd and Wolpin, 2003; Cunha and Heckman, 2007; Cunha et al., 2010; Del Boca et al., 2014; Agostinelli and Wiswall, 2020), with Becker and Tomes (1979) constituting one of the first attempts to model it formally and many following papers enhancing their framework to consider issues such as borrowing constraints and policies related to education and childhood development (Abbott et al., 2019; Lee and Seshadri, 2019; Daruich, 2020; Caucutt and Lochner, 2020). However, this literature has largely ignored the role of geography, and the economic prospects of children may depend on where they are born and where/whether they move. Moreover, opportunities to migrate to different labor markets may have substantial impacts on human capital investment decisions.[5] In studying these issues, my model also contributes to the literature that studies optimal human capital development over the life cycle (Keane and Wolpin, 1997; Heckman et al., 1998; Huggett et al., 2011) through studying the role of geography in these decisions.

My paper's primary contribution comes from extending an intergenerational human capital theory model to a spatial context in order to allow the interaction of geographic and economic mobility to be studied more thoroughly. Most complementary to my paper are contemporaneous papers by Eckert and Kleineberg (2021) and Fogli and Guerrieri (2019),

[5]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 (2021) find similar results in Nepal and Poland, respectively.

who develop general equilibrium models of residential and educational choice to study, respectively, the effects of school finance policy and segregation on income mobility. Human capital levels in the former paper are binary (based on college attainment), while locations are binary in the latter[6].

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

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

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

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

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

A similarly large literature also exists on movements across local labor markets and the migration decisions of both individuals (Kennan and Walker, 2011; Diamond, 2016; Ishimaru, 2022) and families (Mincer, 1978; Gemici, 2006; Venator, 2022). However, these papers focus on the effects of migration during adulthood on one's own earnings (or that of

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

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

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

their spouse), not the future earnings or human capital of one's child. My paper's primary contribution to this literature comes from considering the interplay between such movements and intergenerational concerns. Individuals may move in part to provide opportunities for their future children (Bayer et al., 2007) — at the same time, the investments one's parents make in them as a child may have considerable bearing on their expected returns to migration as an adult. Overall, while the individual literatures on IIM and migration across labor markets in the United States are vast, attempts to synthesize the two are far less common.

The remainder of the paper is organized as follows. Section 1.2 introduces the model, and Sections 1.3 and 1.4 describe the data I use in model estimation along with my estimation strategy. Section 1.5 presents the results of counterfactual exercises, and Section 1.6 considers potential avenues for further research before concluding.

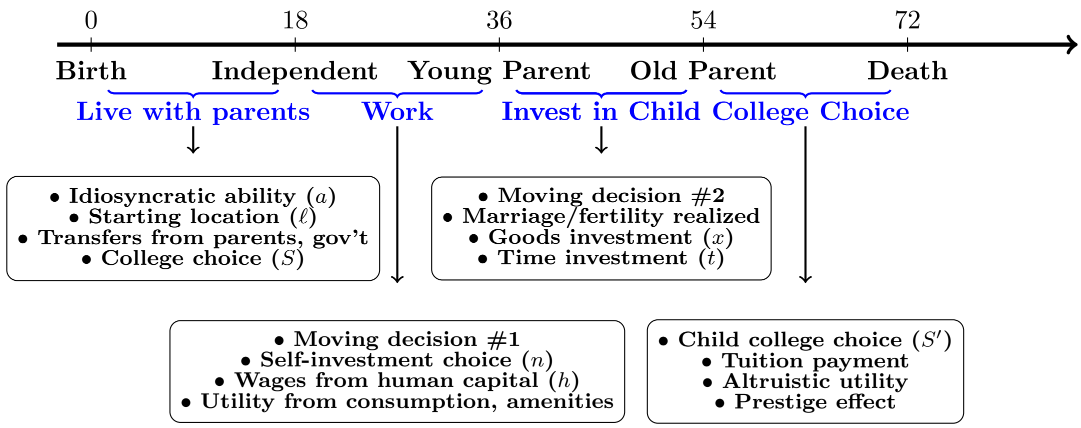
1.2 Model

While the relationship documented in Figure 1.1 may motivate the research question, the empirical correlation between out-migration and IIM is limited in that the role of migration in encouraging upward mobility is likely to be strongly heterogeneous across locations. Further, the data are silent on behavioral responses to migration opportunities — that is, we cannot observe a counterfactual state of the world in which people must stay where they are born to see if agent behavior and outcomes differ substantially from the status quo. I now turn to the economic model I use to study these questions.

1.2.1 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

Figure 1.2: Model Timing



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

human capital and migrate before potentially having children of their own. Parents derive utility from their own consumption and the utility of their children and choose how much to invest in their offspring.

A period is 18 years, and agents live for four periods. Utility over consumption is assumed to be log.[10] The following is a description of the events that transpire and the choices that agents make in each period (see also Figure 1.2 for a visual representation):

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

2. **Period 2:** After emerging from childhood with a level of human capital, ability and schooling, the agent makes an initial moving decision before investing in their own human capital a la Ben-Porath.

3. **Period 3:** The agent has the choice to move again before observing marriage and

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

fertility realizations based on stochastic functions of their schooling, human capital stock, and location as a young adult. If the agent becomes a parent, they balance consumption with providing expenditure and time inputs in the human capital of their child. The agent receives altruistic utility based on the expected happiness of their offspring.

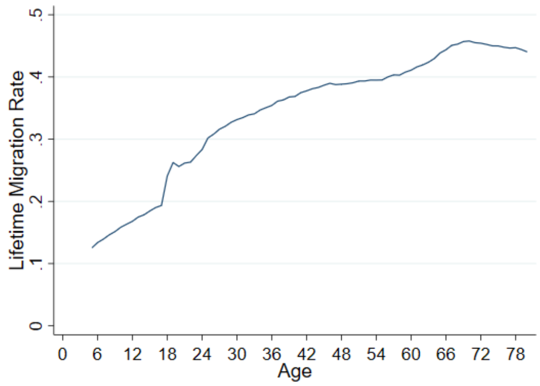
4. **Period 4:** The agent consumes the remainder of their resources (minus tuition should their child choose to go to college) and dies.

Locations (being the 50 states[11] in the U.S. and indexed by ℓ) differ in their costs of consumption/child inputs, amenities, family structure, levels of government child investment, college tuition prices, and rental rates of human capital (i.e. skill prices). The initial migration decision enables agents to move immediately after completing their desired level of schooling[12], and the second moving decision allows agents to potentially relocate to better areas for raising children in anticipation of parenthood. In doing this, the model can capture agents moving in the most migratory period of the life cycle (Figures 1.3a and 1.3b show that both lifetime and yearly migration rates spike in the early 20s) while also allowing for multiple moves, which are a salient feature of the data (Kennan and Walker, 2011) and represent an additional contribution relative to Fogli and Guerrieri (2019) and Eckert and Kleineberg (2021), wherein migration is a one-shot decision.

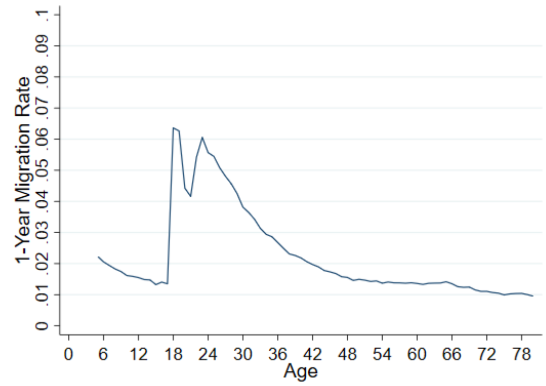
[11] I focus on states instead of CZs both for reasons of computational tractability and because lifetime cross-CZ migration rates are not publicly available. While state effects can account for over two-thirds of cross-CZ variation in IIM, a model that considers a more granular level of geography may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

[12] Note that the college decision here is assumed to be in-state, which for the vast majority of individuals captures the relevant college choice: while 20% of college students attend out-of-state, 94% of individuals either attend in-state or do not attend college at all. This paper is also particularly interested in lower-income children, and the corresponding statistic for children with parents in the bottom income quartile is 97%. Furthermore, while college attendance is an important driver of migration around age 18, the role it plays in lifetime migration is limited due to moves after college and in early adulthood: among individuals in their 30s living in a different state than where they lived at age 17, only 5% are college graduates living in the state where they first attended college. For all these reasons, extending the model to consider out-of-state options would be unlikely to change the main results. (author's calculations using National Longitudinal Survey of Youth 1997 geocode file). For a study that focuses more on migration and out-of-state college options, refer to Kennan (2020).

Figure 1.3: Lifetime and Yearly Interstate Migration Rates by Age



(a) Lifetime



(b) Yearly

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

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

1.2.2 Human Capital Development and Evolution

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

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

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

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

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

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

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

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

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

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

$$h_3 = \varepsilon_2[a(h_2 n)^{\kappa} + h_2],\ \log \varepsilon_2 \sim N\left(-\frac{\sigma_{\varepsilon_2}^2}{2}, \sigma_{\varepsilon_2}^2\right) \equiv F(\varepsilon_2), \tag{1.1}$$

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

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

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

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

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

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

The mean and spread of the growth of parent human capital is allowed to vary depending on the parent's college attainment $S \in \{0, 1\}$. The decision to allow exogenous evolution of human capital in adulthood is made both to ease computation and because the most important determinants of human capital and inequality are realized in the early stages of the life cycle (Huggett et al., 2011). I allow for different distributions of human capital evolution shocks by period due to the length of the time periods in my model: while models with shorter time periods can draw from a single distribution of shocks for each age and estimate said distribution from the flat-point method (Heckman et al., 1998; Huggett et al., 2011; Bowlus and Robinson, 2012), 18-year periods are clearly too long for this method to be valid. The parameters of $F(\varepsilon_3)$ will be calibrated directly from the data, while κ and $\sigma^2_{\varepsilon_2}$ will be estimated via the simulated method of moments.

1.2.3 Recursive Formulation of Decisions

Period 2 — Independence:

The agent makes no decision in period 1, so I begin discussing the decisions agents solve in depth in period 2. The agent enters the second period as a newly independent adult with human capital h_2, ability a, and college attainment $S \in \{0, 1\}$. Given a location ℓ and a binary variable M indicating whether the agent has moved from their birth state[17], the agent solves a standard Ben-Porath problem with an added location decision that follows

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

afterward:

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

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

Where n denotes the time spent producing additional human capital as opposed to working, with human capital evolving according to Equation 1.1. e_t denotes earnings in period t, which itself depends on $w^{\ell,S}$ — the price of human capital in location ℓ for education level S — as well as the amount of time spent investing in one's own human capital as opposed to working. State-specific taxes τ^ℓ are taken from the agent's earnings, and agents who go to college must spend four years obtaining a college degree as opposed to working, represented by $\bar{T} = 2/9$. States additionally differ in their costs of living p^ℓ, which determines how the agent's earnings map to their consumption.

In addition to consumption, agents derive utility from local amenities. To incorporate amenities into the model in an agnostic way, I assume that locations that are larger are more amenity-rich: N^ℓ corresponds to the log of the population of state ℓ, normalized so that the smallest state (Wyoming) has an amenity value of zero. Given that amenities are typically measured at the city level, this specification captures that larger states — i.e. those that possess more large cities in the first place — are likely more amenity-rich than smaller states[18].

The agent optimizes their choice of n by weighing present consumption against their expected future happiness. Higher levels of n decrease their current earnings and consumption but raise their expected future human capital stock, which in turn facilitates migration and

[18] Among cities, those with higher college shares have been shown by Diamond (2016) to be more amenity-rich. At the state level, this relationship does not hold as well, since some locations are quite small and rural despite having high college shares, such as Vermont and New Hampshire. I also test whether the model's main predictions are sensitive to other notions of amenities in Appendix 1.B.4 and find that they are not

raises the likelihood of favorable realizations of marriage, fertility, and future earnings. After their selection of self-investment, the agent chooses whether and where to move:

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

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

where $\zeta_{\ell'}$ are a series of utility shocks drawn from the Type I Extreme Value distribution with location 0 and scale parameter σ_ζ[19]. Combined with the amenity preferences, these shocks prevent my model from mechanically imposing that agents only move for pecuniary reasons, and the shocks in particular will be important in explaining moves from high-wage areas to low-wage areas observed in data. The distributional assumption also allows me to derive closed-form expressions of location choice probabilities and, conveniently, the expected value of v_2:

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

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

The probabilities of marriage and fertility are assumed to be stochastic functions of one's, schooling, human capital stock, and location that follow a probit process — with $m = 1$ indicating the agent being married and $f = 1$ indicating the agent having a child in

[19]So, period 2 could be thought of as comprising two sub-periods containing the self-investment and location decision respectively.

the upcoming period, I denote:

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

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

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

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

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

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

Period 3 — Investment in Children and Altruistic Utility:

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

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

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

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

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

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

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

so married parents are assumed to have the same altruistic factor for each other as they do their children. Since being married increases individual utility monotonically, it does not affect optimal individual choices of x and t, so this specification effectively assumes that parents ignore one another's contributions when making child-rearing decisions and instead behave similar to how they would in a warm glow specification — as a result, children with married parents receive roughly twice the human capital inputs than those with single parents *ceteris paribus*. However, this specification also allows parents to adjust their behavioral margins to counterfactual scenarios that alter the expected payoffs to human capital for their children. The results of the model are robust to either discarding spousal altruism or allowing for a notion of parental coordination by modeling married parents as a single agent with double the available time. I choose this specification because the data suggest that the children of married parents indeed receive roughly twice the time inputs as their single-parent counterparts (see Table 1.C.6b), and modeling parental coordination in labor supply and child investments can be exceptionally complicated. For a (much) more sophisticated treatment of these issues, refer to Gayle et al. (2014).

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

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

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

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

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

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

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

where $V_2(\cdot)$, $\Delta_1(\cdot)$, and $\zeta_{\ell'}$ are the period-2 value function, moving costs, and location preference shocks discussed in Section 1.2.3. Given the distributional assumptions on $\zeta_{\ell'}$ and ε_η, the expectation of $\tilde{V}_2(\cdot)$ can be formed according to the standard Type I Extreme Value form, and the parent's expectation of $V_4^1(\cdot)$ can be computed by finding the threshold level of ε_η that governs college attendance and applying the standard normal CDF and inverse Mills ratio.

Finally, the parent's private utility is similar to the case where they have no children, except they potentially must pay tuition costs[20] and gain additional utility if both they and their child possess a college degree:

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

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

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

Model Solution

The altruistic payoff the parent gains from the child's expected utility in period 3 results in the problems the agents solve in the model being infinite horizon, so a single round of backward induction is insufficient in solving the model. Solving the model proceeds by guessing a value for V_2, after which a new value of V_2 may be produced via backward induction. The model is solved if the updated value of V_2 is sufficiently close to the provided guess. The distributions of the human capital shocks $F(\varepsilon_2)$, $F(\varepsilon_3)$ as well as the conditional distribution of child ability $G(a'|h_3)$ are discretized into five points according to the equal-mass approach (Kennan, 2006)[21]. Policy functions for x and t are computed via grid search,

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

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

while policy functions for n are solved using Brent's method of optimization of a univariate function on a bounded interval. Since the continuous nature of human capital investment decisions and shocks can result in human capital evolving to levels not explicitly included when discretizing the human capital grid, I approximate continuation values when solving for policy functions via linear interpolation over the human capital state[22].

Given the infinite horizon of the model, solving the model requires the assumption that location-specific characteristics (such as skill prices, taxes, tuition, etc.) remain stationary over time, as the infinite horizon of future values of these objects are unknown to the researcher. The simulation exercise of the paper will attempt to reproduce the outcomes of the birth cohorts (1980-1982) studied in Chetty et al. (2014) — in reality, the economic conditions in the U.S changed between the time these cohorts were children and when they reached adulthood, a notable example being skill price shocks induced by the Great Recession. Allowing agents to respond to such shocks in some capacity may be important, so I obtain state-specific parameter values separately for the years 2000 and 2010 and solve the model using both sets of parameters. The child investment decisions of the initial parents, along with the initial schooling, moving, and self-investment decisions of the child are made according to the former set of value and policy functions, while the second move and child investment decisions of the now-adult children use the latter. To test the importance of this assumption, in Appendix 1.B.4 I also conduct an exercise where agents only use year-2010 policy functions and obtain very similar results.

1.2.4 Discussion

The model presented attempts to nest a fairly straightforward model of intergenerational human capital investment into a model of migration with many locations. In multiple stages

[22]The baseline model uses 25 human capital points. Using a cubic spline interpolation to account for potential curvature for very low human capital levels has no substantive effect on the result, nor does increasing the fineness of the human capital grid.

of the life cycle, agents trade off between consuming in the present and increasing their own expected future utility or that of their child. While geographic differences in factors such as family structure and school quality will mechanically inject heterogeneity in child outcomes across different states, differences in real returns to human capital introduce behavioral responses to migration opportunities that vary over space.

A state's given skill price in a vacuum does not necessarily impact human capital investment behavioral margins in a particular way, since while, for instance, a lower skill price decreases the opportunity cost of self-investing, it also results in one having to work more to achieve a given level of consumption. Thus, how skill prices affect investment behavior is sensitive to the curvature of utility over consumption and the productivity of human capital investments[23]. However, the presence of other locations with different returns to human capital has crucial implications for the expected *future* returns to human capital investment when the agent faces possible future migration. For an agent in a location with low human capital returns, the presence of future migration opportunities increases the present expected return to investment to their human capital or that of their child compared to a world in which agents are forced to stay put, consistent with evidence of brain drain in the international literature (Batista et al., 2012; Shrestha, 2017; Spirovska, 2021). Importantly, this is true for both individuals who move and who ultimately choose to stay, which will allow the model to replicate a strong correlation between the outcomes of the overall sample and the stayer subsample across locations. However, the same is not true for an agent in a high-wage location — thus, the impact of migration opportunities on human capital investment will vary systematically across the geographic skill price gradient.

Another notable assumption is that moving costs are decreasing in human capital and education level. This enables agents in the model to invest in order to broaden migration

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

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

1.3 Data

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

1.3.1 PSID

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

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

1.3.2 CDS

To obtain information about how much time parents spend with their children, I use the PSID Child Development Supplement (CDS). In the years 1997, 2002, and 2007, the PSID collected information on time and expenditure investments in children and their outcomes for families with children aged 12 or below. The baseline sample contains information on approximately 3,500 children in 2,400 households. I refine this sample and time measurements following Del Boca et al. (2014) and Lee and Seshadri (2019). I merge information on adults in the CDS into the PSID using individual identifiers and keep only children who have at least one biological parent in the household. I use the same earnings/hours criteria for parents as listed above and exclude parent-child pairs with very small (<18 years) or large (>42) age gaps. These restrictions leave me with 4,402 observations over the three CDS waves.

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

1.3.3 2000 Census and ACS

While the PSID data are effective for capturing life-cycle earnings profiles in the United States, they contain too few observations to be effective in representing aggregate migration, fertility, and marriage patterns at the state level. To discipline the parameters that govern migration choices and stochastic realizations of marriage and fertility in the model, I make use of the 2008-2012 waves of the American Community Survey and limit my sample to household heads[24] born in the U.S. and aged between 36 and 54 (the age group corresponding to Period 3 in the model). I deflate earnings and limit the sample according to hours worked and earnings in an identical manner to how I handle the PSID, with the caveat that I restrict the sample to individuals who work at least 48 weeks per year due to only intervalled information on weeks worked per year being available in the ACS. These restrictions leave me with approximately 1.6 million observations that I use to compute marriage/fertility probabilities over human capital levels and location, state-level native outflow[25] rates, and state-to-state lifetime migration probabilities that are targeted when estimating my model. I also target the gap between average human capital levels of stayers and movers within educational levels observed in these data during estimation. I use a comparable sample from the 5% 2000 Census to obtain distributions of parent human capital, schooling and marriage at the state level used to form the initial condition of the model. To make this sample comparable to the parent sample used in Chetty et al. (2014), I also include authorized immigrants, and I use statistics of IIM for native-born children from the Opportunity Atlas instead of all children to further increase the consistency of my model's output with targeted moments. Sample weights are used in all calculations. These data are also used to calibrate state-level skill prices for high-school and college human capital $w^{\ell,S}$, described in depth

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

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

later on.

1.3.4 NLSY 1997

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

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

To assign ability levels to individuals in the NLSY97, I follow Dillon and Smith (2017) and use results from the ASVAB, a test designed for applicants to the U.S. military that most NLSY97 respondents took in 1997. The ASVAB has twelve separate component scores — I convert these scores to standard deviations within category and birth cohort (to account for individuals taking the test at different ages) before using the first principal component of the scores as my measure of ability. Consistent with Dillon and Smith (2017), I find that the first component explains roughly 60% of the total variance across the 12 sections and is strongly correlated with other test scores such as the ACT. I then compute quintiles of this measure to assign ability levels to individuals in the data that correspond to the five equal-mass levels of ability I discretize the ability distribution to when solving the model. I then compute quartiles[26] of parent income in 1997 to obtain moments of college attainment rates over both ability and parent income.

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

1.4 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.

1.4.1 Parameters Estimated Outside the Model

A summary of the parameters I obtain from data may be found in Table 1.1. The discount factor β is set to $0.96^{18} = 0.479$ to be consistent with an interest rate of 4%. The consumption equivalence scale for an adult with a child is set to $\Lambda(c) = \frac{c}{1.5}$ from the OECD standard. Cost of living levels p^{ℓ} are obtained from the American Chamber of Commerce Research Association's Cost of Living Index.[27] All values are normalized by the value of p^{ℓ} corresponding to Iowa.[28] State populations N^{ℓ} are taken from Census population estimates.

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

[27] 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.

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

I obtain student-teacher ratios s^ℓ (normalized by the mean value) and government expenditures on child human capital g^ℓ from public school statistics reported in the National Center for Education Statistics Common Core of Data Financial Surveys and follow Lafortune et al. (2018) in cleaning the data. State-level tuition rates T^ℓ are computed from enrollment-weighted average sticker prices of public-four year colleges for each state, and the financial aid schedule $A(e_4, a')$ is calibrated from published Federal Pell Grant schedules as well as the National Postsecondary Student Aid Study (NPSAS)[29], a nationally representative survey of 50,000 college students that contains detailed breakdowns of types of grant aid received by parent income, high school GPA, and type of institution attended. Taxes τ^ℓ are taken as the sum of state-level sales tax rates and combined average federal and state income tax rates as calculated by the NBER TAXSIM model[30]. The mean and standard deviation of parent human capital μ_h, σ_h that enter the joint distribution between parent human capital and child ability are set to 0.902 and 0.634, obtained from wage rates in the 2000 Census after correcting for local skill prices, the estimation of which I now turn to.

Skill Prices

The main simulation procedure will roughly attempt to reproduce the outcomes of the CHKS cohorts. Drawing child ability, however, requires knowledge of the underlying human capital of their parents. It is important to distinguish parental human capital from parental earnings in the context of my model: for instance, one may be justifiably worried that two parents with identical earnings in a high-wage and a low-wage location still differ meaningfully in characteristics that may influence the ability and human capital of their child. Separating human capital from earnings is thus crucial in accounting for parental sorting (Heckman and Landersø, 2021) and credibly forming the initial condition of parents in my model, but doing so requires information about how the price of human capital differs across locations.

[29] See https://nces.ed.gov/surveys/npsas/.

[30] See http://users.nber.org/ taxsim/allyup/.

Table 1.1: Parameters Estimated Outside the Model

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

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

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

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

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

$$\frac{e_3}{1-t} = w^{\ell,0} h_3 \Longrightarrow \log\left(\frac{e_3}{1-t}\right) = \log(h_3) + \log(w^{\ell,0}), \tag{1.2}$$

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

College skill prices are then estimated via computing state-specific college premia. Note that if a college and high school graduate have identical underlying human capital, the ratio of their wages in a given location is equal to the ratio of that location's skill prices:

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

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

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

Marriage and Fertility Realizations

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

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

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

where m_i and f_i are dummies for being married and having a child for individual i, and h_i is their level of human capital. Following the model, M_i is a dummy for individual i not living in their birth state. When estimating fertility probabilities I limit my ACS sample further to individuals aged 36-45 to prevent underestimating fertility from including parents whose children may have already left the household. Probability functions for fertility are estimated separately for married and single adults. The estimated probabilities for both outcomes are held constant past the level of human capital corresponding to the 99th percentile in the data to avoid Runge's phenomenon at the right tail of the human capital distribution. For visualizations of the marriage probabilities and an evaluation of how well they fit the data, refer to Appendix 1.B.2.

Late Human Capital Shocks

Finally, $\mu_{\varepsilon_{3,S}}$ and $\sigma_{\varepsilon_{3,S}}$ are calibrated directly from data on older household heads in the PSID.

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

1.4.2 Simulation

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

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

The simulation procedure itself attempts to reproduce the outcomes of the same cohorts that CHKS study. I take the 2000 Census and limit my sample to individuals aged 36-54 who have at least one[31] child living in their household, after which I compute the distribution of human capital, parameterized as a log normal, and the joint distribution between human

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

capital, education, and marital status in each state using the same methods as described above. I separate married individuals into three groups based on whether their spouse has a college degree, does not have a college degree, or does not work at all, and I nonparametrically estimate the joint distribution of household head education and spousal type by taking the relative frequencies of each type of head-spouse combination directly from the data. After conditioning on assortative matching on education, I find that underlying spousal human capital is only weakly correlated with that of the head's, so I draw human capital for working spouses independently. I conduct the same procedure when determining family income for children who reach the parent stage of the model with a spouse.

Using this as the distributions of parental characteristic for the CHKS cohorts, I then randomly draw 20,000 parents for each state, after which I draw the ability levels of their children[32] and simulate their migration, marriage, and earnings outcomes later in life. Spouses that do not work are assumed to provide a time investment into the child's human capital that I take directly from non-working parents in the PSID CDS but zero goods investments[33]. When computing moments from the simulated data, I weight by home state population sizes to ensure that the simulated data is representative of the U.S as a whole[34].

The values of the parameters in Θ are reported in Table 1.2, along with a description of the data moments used to discipline them (more on this in the next section). Denoting $M = [M_1, M_2, ..., M_N]$ as the vector of empirical moments I target in the simulation procedure, denote $g(\Theta)$ as the vector of percentage errors between the data moments and the simulated

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

[33] Note that since the spouse is unemployed, their human capital cannot be observed in the data, so applying the policy functions in the model is not an option. Recall from Section 1.3, though, that the data are limited to households with household heads that work a certain amount, so the human capital of the head is always observable.

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

moments, so:

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

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

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

where $M(\Theta)$ are the simulated model moments and W is the diagonal of the variance-covariance matrix of the data moments, obtained from bootstrapping the various samples used to make the moments 1,000 times. Minimization of the objective function proceeds via the Sbplx[35] routine. I compute standard errors by evaluating the numerical gradient[36]. of the objective function and applying the standard indirect inference formula (Gourieroux et al., 1993).

In the model estimation and simulation, I also allow for agents to differ by race (non-Hispanic White, Black, and Hispanic) and allow for mean levels of ability μ_a to vary over the four Census regions. Allowing for racial heterogeneity is important when attempting to account for high rates of economic mobility in rural states, given that these states also feature high levels of racial homogeneity. Among individuals of different races, I estimate separate human capital productivity parameters ξ, college fixed costs η_1, and moving fixed costs $\delta_{1,2}$ to enable the model to fit racial heterogeneity in income, educational attainment, and migration patterns. The heterogeneity in child human capital productivity could also be thought to account for potential within-state racial disparities in government investment that is not explicitly modeled here. In addition to this, I allow for race to influence skill

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

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

Table 1.2: Parameters Estimated via SMM

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

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

prices, marriage likelihoods, and fertility likelihoods across locations. Including separate means of ability across regions allow for some notion of peer group or neighborhood effects on child development while still allowing for within-region heterogeneity in school quality and family structure. Another option would be to allow for the correlation between parent human capital and child ability ρ_{ha} to vary across locations, but in the NLSY97 I find similar correlations between parent income and child AFQT scores across regions along with level differences across regions that are more consistent with mean shifts. Refer to Appendix 1.B.3 for parameter estimates and the model's fit for these categories.

1.4.3 Identification

While the model is jointly identified in general, a conceptual argument for identification is as follows. The altruism parameter θ is tied to moments to do with intergenerational persistence in income and is thus estimated by targeting the rank-rank[37] intergenerational elasticity of family earnings of 0.341 as reported by CHKS. A higher level of persistence of learning abilities ρ_{ha} results in a larger proportion of high-ability children being born to richer parents. Because non-pecuniary costs of college attendance are decreasing in ability, this results in a sharper pattern of college-going over the parent income distribution. Thus, ρ_{ha} is chosen to match rates of college-going by parent income quartile as taken from the NLSY97.

The fixed utility cost, ability component, prestige component of college attendance, and college preference shock spread $\eta_1, \eta_2, \eta_3, \sigma_\eta$ are targeted to match overall college-going as well as college-going by ability and by parent educational attainment. I also include the complete set of attendance probabilities by parent income quartile, ability level, and parent educational attainment (40 moments total) to bolster the identification argument: since ability and parent income co-move in both the data and the model, targeting college attendance by ability *within* parent income quartiles allows for better identification of η_2, and differences in college attendance by parent educational attainment within given ability/income combinations identifies η_3. Since the concavity of utility over consumption results in richer parents being more willing to fund their child's education than poorer parents, and σ_η governs the magnitude of preference shocks relative to utility from consumption, I target σ_η to match the growth of college attendance over parent income *within* levels of ability and parental education[38]. I give these moments less weight when estimating the model to prevent the

[37]The choice of income mobility measure to target follows Lee and Seshadri (2019). I assess the sensitivity of model results to other measures of income mobility in the Section 1.5.

[38]Note, though, that even if attendance by parent income/ability combinations are matched, fitting the attendance profile over the parent income gradient *alone* will still be contingent on having the joint

number of them from dominating the objective function. Additionally, since attendance for some of the cells is close to zero, I use the absolute error instead of percentage error to avoid low-attendance cells from being given undue weight.

Estimation of μ_a and σ_a starts with the observation that earnings means and standard deviations at any stage of the life cycle increase monotonically with higher μ_a and σ_a. Thus, I target the mean and standard deviation of normalized individual earnings[39] by education level in the PSID for the age ranges corresponding to Period 2 and Period 3 in the model to estimate the two parameters. Meanwhile, the parameters κ and σ_{ε_2} primarily govern *growth rates* of earnings as the agent transitions from Period 2 to Period 3. Thus, I target the mean and standard deviation of individual earnings growth rates between the same age ranges to estimate κ and σ_{ε_2} respectively. Period 2 earnings moments also assist in estimating the child investment productivity parameter ξ because higher values of a also result in faster earnings growth rates on average, thus restricting the values that μ_a can take.

The parameter ϕ governs how important time inputs are in forming a child's human capital, so a natural moment to target is the amount of time parents spend with their children. I obtain this moment from the PSID CDS sample described in Section 1.3. 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.[40]

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

distribution of parent income and child ability correct, which is governed exclusively by ρ_{ha}.

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

[40]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.

graduates is targeted to estimate the period-2 fixed moving cost δ_1, and the college moving cost component δ_5 is targeted to match the migration rate of college graduates, and the population component δ_6 is targeted to match rates of state in-migration. The human capital component of moving costs δ_3 is targeted to match the aforementioned human capital gaps between movers and stayers, and the proximity component of moving costs δ_4 is chosen to match the share of moves that are made to nearby states.

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

1.4.4 Model Fit

Having estimated the model, I now evaluate its performance. Estimated parameter values can be found in Table 1.2, while Table 1.3 displays my model's performance in fitting its targeted moments. The parameters that govern parental altruism, ability inheritance, and human capital development are all well within the ranges of estimated values in other papers that use similar technologies — in particular, the value of ϕ is quite close to the values of time shares that Lee and Seshadri (2019) estimate from the PSID CDS. The model fits lifecycle earnings profiles for individuals with a high school degree well but overstates earnings growth for college graduates. The model also replicates a college attendance profile that increases over the parent income and ability gradients but does overstate attendance among the lowest-ability children and understates attendance among children with college-educated parents. The model fits the rank-rank IGE of earnings estimated by CHKS exactly — moreover, the model predicts a log-log IGE of 0.439, which is quite similar to CHKS's estimate of 0.413 when recoding cases of zero income to $1,000[41], indicating that the estimation results would not be particularly sensitive to the measure of intergenerational mobility used.

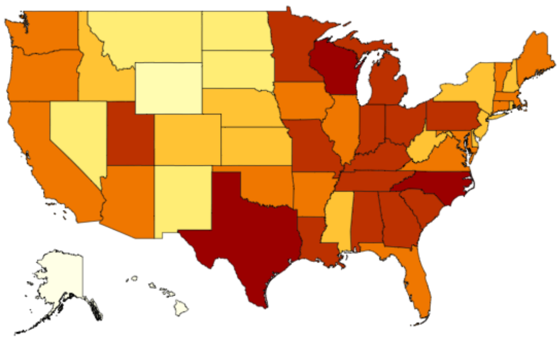
[41] This is most appropriate for my setting since the model does not feature unemployment, and so nobody has zero income. Minimum earnings among agents in my model are around $2,000.

Table 1.3: Model Fit

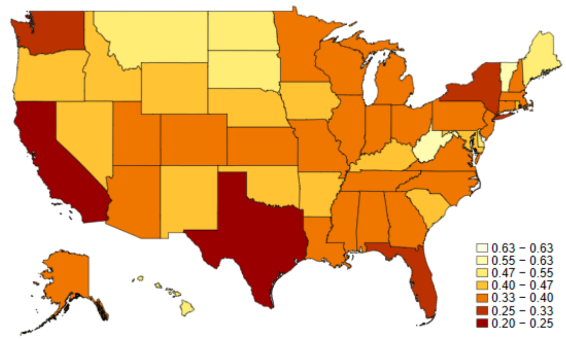
Moment	Data	Model	Source
Earnings and Human Capital			
Rank-rank IGE of family earnings	0.341	0.341	CHKS
Period 2 earnings mean, HS	0.687	0.669	PSID
Period 2 earnings SD, HS	0.336	0.268	PSID
Period 3 earnings mean, HS	0.954	0.909	PSID
Period 3 earnings SD, HS	0.495	0.513	PSID
Period 2-3 earnings % growth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

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

Figure 1.4: Model Fit — State Outflow and Inflow Rates

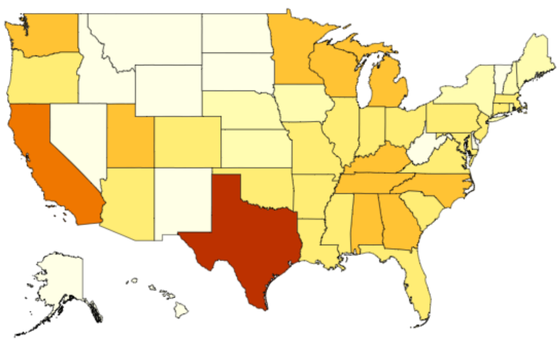


(a) Outflow (HS), Data

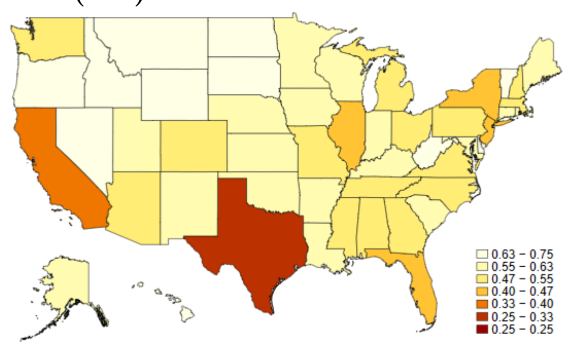


(b) Outflow (HS), Model

Outflow Correlation (HS): 0.40

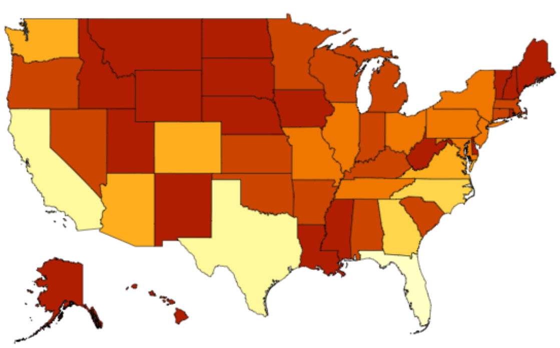


(c) Outflow (College), Data

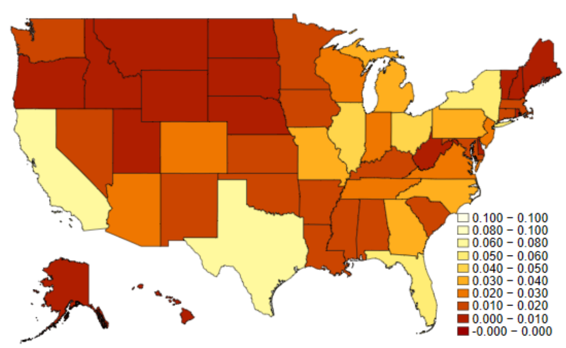


(d) Outflow (College), Model

Outflow Correlation (College): 0.67



(e) Inflow, Data



(f) Inflow, Model

Inflow Correlation: 0.89

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

The moments regarding migration rates and gaps in human capital measured between stayers and movers are matched well. The estimates of parameters governing moving costs suggest that moving to a nearby state is about one quarter less costly than moving to a non-nearby state. Moreover, the population component of moving costs suggest that a move to the most populous state (California) is roughly one-third less costly than a move to the least populous state (Wyoming), all else equal. However, the lack of an explicit preference for the individual's home state means that the model struggles to fit the (still quite low) rates of return migration observed in the data. Notably, the fixed costs for the two periods are similar: the model does not need a considerably higher fixed cost to rationalize lower rates of migration later in the life cycle since returns to migration when one is older are also lower.

Because utility for consumption is non-linear, I can only express moving costs as a function of base consumption as opposed to an exact dollar amount. The log form of utility implies that any utility cost, denoted X, is equal to Y dollars lost on a base consumption of C_0:

$$\ln(C_0 + Y) = \ln(C_0) + X \Longrightarrow Y = C_0 \left(\exp(X) - 1\right).$$

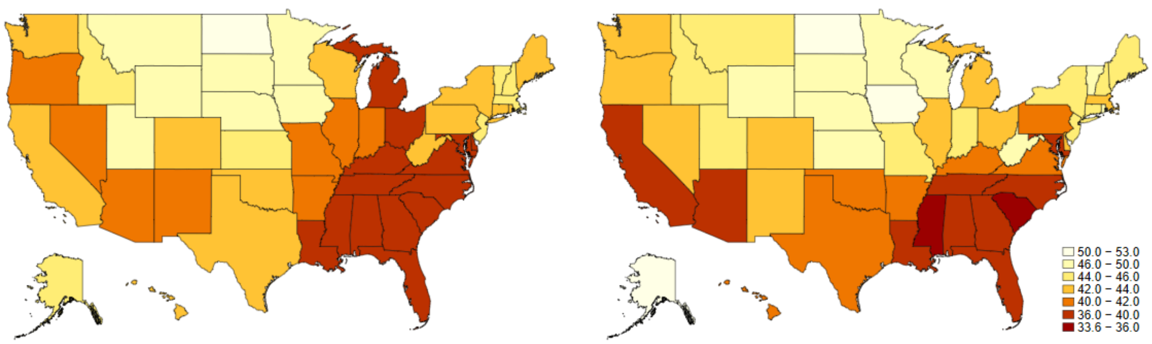
Because my estimation involves simulating moving decisions and the draws of moving utility shocks, I can conduct a similar exercise to Kennan and Walker (2011) and compute average utility costs individuals experience conditional on moving, which depend on the δ parameters, the utility shock draw ζ_l for the chosen location, and the utility shock draw ζ_0 for the starting location (since moving requires exchanging ζ_1 for ζ_0). Taking these factors into account, I recover average utility costs of 0.977 for moves made before period 2 and -1.4622 for moves made before period 3 — using average individual consumption levels for period 2 and 3 of 0.557 and 0.681 as bases and applying the monetary scaling of 47,961 2012 dollars, this translates to monetary moving costs of $43,731 and -$25,051 on average for the two periods.

The negative moving cost for moves preceding period 3 in the model suggest that the moves are being driven by idiosyncratic utility shocks as opposed to higher utility flows from the destination location, consistent with estimates in other entries of the structural migration literature (Kennan and Walker, 2011; ?). However, the positive average cost for moves preceding period 2 suggests that the prospects of higher utility in certain locations may be a compelling driver of migration earlier in the life cycle.

While the average migration rates are matched almost exactly by the model, Figure 1.4 evaluates the model's performance in matching individual state outflow rates and migration destination probabilities. Generally speaking, the model does well, particularly for college graduates — the college graduate outflow rates at the state level predicted by the model and observed in the data are significantly correlated (coefficient 0.67, indicating that the model can account for nearly half the variation in cross-state college graduate migration patterns). Consistent with the data, the model predicts the Midwest and Mountain States to be highly migratory regions, with less migration being observed out of states in the Rust Belt and the Southeast. These qualitative patterns hold for the migration patterns of high school graduates as well, but here the model overpredicts migration out of the Rust Belt and the Southeast and understates migration out of the most populous states, such as California, Illinois, and New York. Another salient miss for high school graduates is that the model underpredicts migration out of the some of the more populous states, such as California, Illinois, and New York. In Appendix 1.B.4 I assess the how the model's performance changes with alternate notions of amenities, finding comparable headline results across numerous specifications. The model does better in predicting rates at which states receive migrants, fitting this aspect of migration almost exactly.

I next evaluate how well my model reproduces the geography of intergenerational mobility in the United States as reported by CHKS. For each U.S. state, I take the average family

Figure 1.5: Model Fit — Upward Mobility by State



(a) Data (b) Model

Correlation: 0.81

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

income rank of children born to below-median income parents[42] to compute measures of absolute mobility in the model. Figure 1.5 juxtaposes the state-level IIM measures that CHKS find with the ones that my model predicts. The model's performance in replicating the geographic variation of upward mobility is respectable — the correlation between my estimates and those of CHKS is 0.81, indicating that my model can account for approximately 65% of the state-level variation in income mobility observed in the data. The model fits states in the Southeast and Midwest well but does slightly overpredict mobility in the Northeast and the Rust Belt. Additionally, the model underpredicts income mobility in some parts of the West Coast as well as the Southwest[43].

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

[43] Part of this may come from the model not fully capturing differences in college attainment for natives of these states vs. elsewhere; see Figures 1.C.2a and 1.C.2b. Another consideration is heterogeneity in parent altruism: if parents in locations with low wages but good opportunities for child human capital development choose to live there because they are particularly altruistic toward their children, these same parents may invest more in their children than is currently captured in the model. Furthermore, if these especially altruistic parents invest more in their children so as to broaden their future migration opportunities, then

Table 1.4: IIM Statistics for Stayers

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

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

As an additional test, I evaluate whether the model can replicate the strong correlation in outcomes observed between the overall sample and the subsample of individuals who stay in their home state. If the model overstates selection into migration and the importance of migration for driving earnings growth for natives of low-wage states, the correlation between overall state-level rates of IIM and state-level rates among stayers should be low. However, Table 1.4 demonstrates that the model can reproduce this correlation quite well[44]. The model predicts a sensible difference in income mobility between the whole sample and for stayers and can replicate a strong correlation in outcomes between all individuals and stayers at the state level. As will be discussed in the next section, behavioral responses to migration opportunities play a key role in driving this result: since migration opportunities spur human capital investment for all individuals — not just those who eventually move — the model can rationalize favorable outcomes even for individuals that choose to remain in relatively low-wage states. Note additionally that this correlation is not explicitly targeted in the model estimation.

the model may be understating the behavioral response of these parents to migration restrictions.

[44]See Figures 1.C.2c and 1.C.2d for a state-level visualization of this fit. Areas such as the South, the Rust Belt, and the Northeast are captured well, but the model does understate mobility for stayers in some Mountain and Plains States, especially Montana, Idaho, and South Dakota.

Table 1.C.6 reports the model's fit for additional data moments. Table 1.C.6a indicates that the model does a reasonable job of fitting parent-child income quintile transitions, and Table 1.C.6b indicates that the model slightly overstates the amount of time inputs received by children with married parents — however, the amount of time the typical individual parent spends with their child is close to the average time investment of 0.11 targeted by Lee and Seshadri (2019). Note further that the moments in Table 1.C.6a and Table 1.C.6b are not targeted in model estimation. Tables 1.C.6c and 1.C.6d display the model's fit for the full set of college attendance moments broken down by ability, parent income quartile, and parental educational attainment. In addition to overstating college attendance among the lowest-ability children, the model understates college-going for the poorest children without college-educated parents as well as children with below-median income parents who are college-educated. Otherwise, the fit is reasonable.

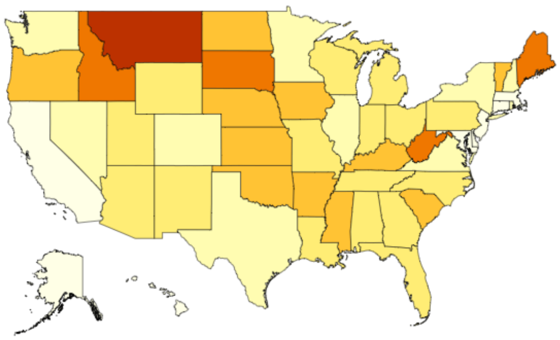
1.5 Decompositions and Counterfactual Exercises

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

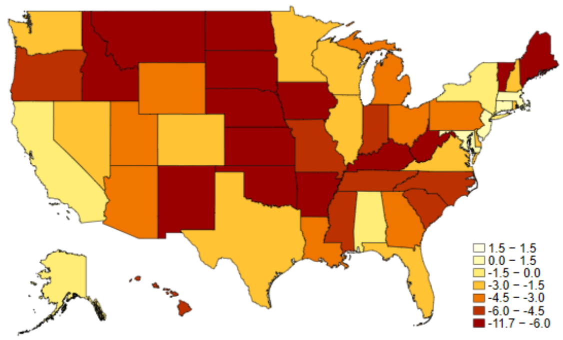
1.5.1 The Role of Migration

The main goal of this paper is to assess the importance of migration in influencing IIM in the United States. I approach this question from two directions — in one counterfactual, I run the model as before and then move all individuals back to their home states ex-post, and in another I simply set the moving fixed costs for periods 1 and 2 (δ_1, δ_2) to infinity, which

Figure 1.6: Counterfactual — Results of Migration Restrictions



(a) Change (no BR)



(b) Change (full BR)

Notes: BR = Behavioral responses. Figure 1.6 plots the change in upward mobility from counterfactuals that restrict migration while ignoring or including behavioral responses. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile in the data and expected family income percentile of children born to below-median income parents in the model.

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[45]: 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 1.6 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

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

surprising, as moves in the model are usually to higher-paying areas, so restricting migration ex-post typically reduces the earnings of movers while having no effect on stayers. States with the highest returns to human capital typically had low rates of out-migration to begin with and so stand to benefit little from a migration restriction. However, while average earnings in the majority of states suffer from the migration restriction, the effects are quite heterogeneous: with behavioral responses, the model predicts that barring a South Dakota native from moving later in life would result in their expected income percentile rank as an adult dropping by over 10 points (translating into a family income reduction of around \$15,000); on the other hand, doing the same to a child from Connecticut would *increase* their expected rank by about 2.9 points (around \$3,000)[46]. These effects are the result of a combination of mechanically moving people to higher or lower-paying areas (which is key in generating improved outcomes for the natives from Connecticut, the state with the highest skill prices) and behavioral responses of individuals decreasing their human capital investment and college-going following the removal of migration opportunities. These responses play an important role: the percentile shifts for South Dakota and Connecticut natives in the counterfactual that ignores behavioral responses are a 4.6 reduction and a 3.2 gain, respectively. More rural states are generally hit harder by the counterfactual, with particularly strong earnings effects observed among some states in the Great Plains and Appalachian areas.

This suggests that migration as well as opportunities to do so may be important in shaping adult outcomes for children from more remote areas — to frame the results differently, Table 1.5 summarizes the impacts of the counterfactual for states in and out of the West

[46]Though note that individuals base their migration decisions both on pecuniary factors such as consumption but also on non-pecuniary factors such as amenities and preference shocks. As a result, the utility implications of these migration restrictions are always negative; see Figure 1.C.3. The geographic profile of utility changes closely resembles changes to IIM: natives of states that had high rates of out-migration due to low wages or amenities experience the greatest utility losses from migration opportunities being taken away. Since natives of these same states self-invest in part to enable out-migration, these are also the states that experienced some of the largest changes in IIM, especially with behavioral responses included.

Table 1.5: Migration Restriction Impacts by State Group

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

Notes: BR = Behavioral Response. Table 1.5 reports average impacts for states either in or out of the West North Central and Mountain Census divisions. Row 1 reports IIM as reported in CHKS, while Row 2 reports IIM as predicted by the model. Row 3 reports changes in IIM following a counterfactual that shuts down migration in the model, and Row 4 does the same while allowing for the agent to adjust their behavior in the young adult and adult stages of the model. Row 5 further allows the agent and their parent to adjust their college decision in response to the migration restriction, and Row 6 enables parents to adjust their child investment behavior as well. Standard errors of estimates in parentheses. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile in the data and expected family income percentile of children born to below-median income parents in the model. See Appendix 1.A for division definitions.

North Central and Mountain Census divisions, the two divisions with the highest levels of IIM in the United States. The effect is such that the gap in upward mobility measures between those states in the divisions of interest and those not decreases by approximately 1.40 points when ignoring the behavioral response (Row 3) and 2.23 points when including it fully (Row 6). The shift of 2.23 points with full behavioral responses is statistically significant and represents slightly more than half the gap in IIM between the two groups of states of 4.45, suggesting roughly that half the advantage these areas enjoy in measures of upward mobility[47] may be attributed to migration channels.

I further decompose the behavioral responses by simulating two alternate scenarios: in the first (Row 4), the children only adjust their behavioral margins when they reach the adult stages of the model (i.e. the self-investment decision in period 2 and child investment

[47] I also consider gaps in earnings levels as an outcome and reach a similar, if not stronger conclusion: the gap in family earnings gaps between the two locations in the data is approximately \$6,000, and the model predicts this gap to close by around \$3,900 when behavioral responses are fully incorporated.

decision in period 3). In the second (Row 5), the children and their parents can also adjust their college decision in response to the migration restriction. This allows me to quantify which model mechanisms are most important in generating the result. I introduce behavioral responses backward in the lifecycle in this way to account for the fact that human capital in the model forms cumulatively from a series of investment decisions. Incorporating the young-adult self-investment behavioral response alone hardly changes the result at all, since migration and the returns to it are much lower later in the life cycle. In fact, since those from the West North Central and Mountain divisions are forced to live in areas with generally better marriage markets, they invest slightly more in their human capital to increase the odds of favorable family formation outcomes. The migration restrictions have much stronger impacts when the agent's college decision can adapt to the counterfactual world: college attendance falls by approximately 30%[48] in the model when nobody is allowed to move, resulting in considerable reductions to absolute mobility.

Moreover, allowing the college decision to respond to the migration shutdown begins to further shrink the gap in outcomes between the West North Central/Mountain divisions and elsewhere, since the decision to go to college is more heavily influenced by opportunities to out-migrate in the former states than the latter. The gap shrinks further when parental behavioral responses are accounted for as well, suggesting that the compounding nature of human capital and parental responses to the migration opportunities of their children play a crucial role in generating high upward mobility in low-wage states[49]. This decomposition also

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

[49] See Figure 1.C.4 for an alternate representation of these results as scatterplots plotting state-level IIM against predicted changes in IIM following the migration restrictions. The plots show a negative association between the two that sharpens when behavioral responses are considered compared to when they are ignored, indicating again that migration is particularly important in explaining upward mobility in some of the most income-mobile states in the U.S. At the same time, it is important to note that there are other highly income-mobile states that are affected little by the migration restriction (particularly ones in the Northeast), which dampens the overall correlation.

highlights the importance of considering continuous human capital and continuous human capital investments: were human capital limited to the binary college/non-college types (as in Eckert and Kleineberg (2021), for instance), behavioral responses in the model may be meaningfully understated. Nonetheless, states in the Midwest and the Mountain states are still meaningfully more income-mobile than other states even with the full set of behavioral responses, suggesting that state-level differences in demographic, educational, or economic factors likely still play an important role in shaping the outcomes of their natives. I turn to evaluating the importance of these attributes next.

1.5.2 Other Determinants of Economic and Geographic Mobility

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

Table 1.6 reports the results of this exercise. Among the different specifications, equalizing demographic attributes has the largest effect on reducing the cross-state variance in IIM, lowering the standard deviation by almost 40%. Equalizing schooling attributes also results in a meaningful reduction in inequality but not by as much. However, equalizing costs of living and skill prices slightly *increases* the cross-state standard deviation in income mobility, suggesting as in the data that local labor market conditions and income mobility

Table 1.6: Model Decomposition of Sources of Income Mobility

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

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

may only be tenuously related to one another.

I also estimate spatial labor supply elasticities in a similar fashion to Kennan and Walker (2011) by shocking characteristics of an individual state, resimulating the model, and observing how the population size of the state responds to the shock. Population is measured as the number of individuals in the CHKS cohort living in a given state in the parent stage of the model (period 3) — for example, the elasticity of a wage shock in a given state is computed as $\frac{\Delta N}{\Delta w} \cdot \frac{w}{N}$, where N is the number of individuals in the CHKS cohort living in the given state as adults and ΔN is the difference between this and the number of individuals in the state as adults in the counterfactual simulation.

I estimate such elasticities for shocks to either economic or educational conditions for each state individually. In the first set of experiments, I introduce a 10% wage shock in each location for both education levels[50]. In the second set, I simultaneously increase government expenditure on students g^{ℓ} by 10% and decrease student-teacher ratios s^{ℓ} by 10%. The

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

average labor supply elasticity in response to the wage shocks is approximately 0.9, which is similar to the long-term elasticity estimated by Kennan and Walker (2011) and suggests strong migration responses to spatial wage differences despite large moving costs and moving utility shocks. In contrast, the average labor supply elasticity in response to the education shocks is around 0.3, despite the schooling shocks having slightly larger effects on the income mobility and utility of the natives of the affected states (1.C.5)[51]. That the educational shocks have comparable IIM effects as wage shocks is notable, given that school characteristics are likely easier to influence via policy than state-wide wages.

1.5.3 Retention Policies

While migration is important in generating income mobility in low-wage parts of the country, several U.S. states have been or are concerned about the tendency of talented individuals to vacate them. As a result, these states have recently weighed legislation that would provide financial incentives for individuals with higher human capital (typically, college graduates) to locate in them.[52] 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

[51] In general, the states that benefit the most from these shocks in terms of either IIM or utility are those with low rates of out-migration to begin with, since more of their natives will stay and enjoy the improved conditions, potentially for multiple generations. This, along with the shocks dampening incentives to accumulate human capital in order to leave certain places, results in the profile of which states observe the largest improvements to IIM or utility looking quite similar to the profile of which states saw the smallest reductions in IIM following restrictions to migration.

[52] 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.

not on the margin of staying/leaving in the first place. Additionally, a natural economist's objection is that distorting the location choices of highly skilled individuals is unlikely to be efficient on a national level as well. Evaluating the global optimality of such policies would require an equilibrium analysis and more careful consideration of agglomeration economies and spillovers induced by high-human-capital individuals concentrating geographically. Such issues are beyond the scope of this paper, so I instead focus on whether such a policy may be profitable from an individual state's point of view.

I use my model to assess the likely cost-effectiveness of these programs. Specifically, I consider three counterfactuals in which locations provide subsidies of $10,000, $20,000, and $50,000 to individuals who both choose to live in them as a young adult (that is, in period 2) and who have a college degree. These three policies are sequentially introduced in each individual state, one at a time, before re-solving the model and re-simulating data. I consider two impacts of the policies: the change in the end-period college share in the state after the introduction of the counterfactual as well as the percentage change in each state's net tax revenue[53].

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 externalities of the presence of college graduates on the earnings of others, I allow for the earnings of high schoolers and college graduates in a state to increase by 1.6% and 0.4%,

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

Table 1.7: Counterfactual — State Retention Policies

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

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

respectively, following a 1 percentage point increase in that state's college share.[54]

Table 1.7 presents the results of this exercise at the division level and indicates that the policies generally fail to be cost-effective. The responses of individuals to the policy is generally small — even in the counterfactual that offers a $50,000 subsidy, the typical state sees less than a 3 percentage point increase in the college share of their labor force. This happens because even $50,000 is negligible relative to lifetime earnings for highly skilled individuals, and utility shocks play an important role in both migration and college decisions. As a result, the majority of agents with a college degree are not sufficiently close to migration margins to respond to the policy, and the overwhelming majority[55] of the subsidies go to individuals who choose to both obtain a college degree and locate in the given state in the baseline model without the subsidy, which in turn renders the policy highly cost-ineffective.

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

[55] 98.8.2%, 97.7%, and 94.5%, respectively, for the three policies.

Moreover, the policies are the least cost-effective in the low-wage areas of the country that have considered implementing them the most.

1.5.4 Schooling Policies

As a final counterfactual exercise, I evaluate how changes to school characteristics at the state level influence cross-state inequality in social mobility. I first consider equalizations of schooling characteristics across each state: in one counterfactual, I set real government child expenditures and student-teacher ratios across all states to either the best values in the data (corresponding to the real expenditure of Vermont and the student-teacher ratios of Wyoming, respectively). In another, I set tuition prices for all states to the lowest value observed (that of Oklahoma's). I also run scenarios that set levels of school characteristics and tuition equal to average values instead of the best values. I resimulate the model in these counterfactual worlds and observe how cross-state means and spreads in absolute mobility compare to the baseline world[56].

Table 1.8 reports the results of these policies and indicates that improvements to early school characteristics are considerably more impactful in both improving utility and equalizing outcomes than tuition reductions: exposing the average child to the best possible schooling environment improves their upward mobility in expectation approximately 30 times more than if they are offered the lowest tuition prices. Moreover, the cross-state spread in IIM falls by close to one third when school characteristics are equalized (this is true whether they are equalized according to best or average values), while equalized tuition hardly shifts the spread at all. This is not surprising, given that only around 19% of individuals with below-median income parents in the data (22% in the model) actually obtain a college degree in the first place, and other research indicates that interventions earlier in the life cycle are

[56]The caveat that these are partial equilibrium exercises that evaluate how the outcomes of a single child differ in expectation under the counterfactual policy bears repeating — in particular, general equilibrium responses to school funding equalizations are important to consider (Eckert and Kleineberg, 2021).

Table 1.8: Effects of Schooling Equalizations

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

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

likely to be more potent in improving economic mobility (Heckman et al., 2010; Lee and Seshadri, 2019; Bailey et al., 2020).

1.6 Discussion and Conclusion

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

While a substantial portion of the upward mobility in the Midwest and Mountain States is attributable to migration, it is important to note that these parts of the country remain meaningfully more income-mobile than elsewhere even in the decomposition that fully incorporates behavioral responses to the migration restriction. Other factors specific to the locations themselves identified by Chetty et al. (2014) and others, such as school quality and family structure, remain influential in driving economic mobility. Thus, where individuals are from as well as where they go are both important in understanding salient features of income mobility in the United States.

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

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

1.A Divisional Groupings of States

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

1.B Estimation Appendix

1.B.1 Skill Prices

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

High School Skill Prices

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

$$\log(w_{it}) = \beta_0 + \beta_1 \mathbf{X}_i + \beta_2 x_i + \beta_3 x_i^2 + \beta_4 x_i^3 + \beta_5 x_i^4 + \mathbf{D_{it}} + \varepsilon_{it}, \tag{1.3}$$

where w_{it} is hourly wage for individual i (mapping to $\frac{e_3}{1-t}$ in equation (1)) and $\mathbf{X_i}$ is a vector of demographic characteristics (black, male, and hispanic dummies, along with dummies for having moved from one's home state) included to account for compositional differences across states.[57] This vector together with a quartic polynomial in years of potential experience serve as a collective proxy for h_3 in (1), and the vector $\mathbf{D_{it}}$ represents dummies for living in each state by time period (2000 or 2008-2012) and are what allow me to derive skill prices for

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

high school graduates, computed as $w_t^{\ell,0} = \exp(D_{\ell,t})$. I omit the $\mathbf{D_{it}}$ dummy corresponding to Iowa in 2000 in the regression as a normalization.

Figure 1.B.1 displays the geography of skill prices computed from this method for the two time periods as well as how these prices changed over time. These measures are presented both as they are obtained from the regression equation above and after adjusting for different cost-of-living levels across states. As one may expect, skill prices tend to be lower in states that are lacking in large cities, with particularly low returns for states in the Great Plains and 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.[58] Moreover, changes in skill prices observed between 2000 and 2010 intuitively reflect economic phenomena known to have happened in the 2000s: states in Appalachia generally see larger skill price reductions following struggles in the manufacturing sector, and Michigan experiences the single largest fall in skill price due to the collapse of the automotive industry there. Moreover, states such as North Dakota and Wyoming see increases in their skill prices following the fracking boom.

College Skill Prices

The next step is to compute college premia at the state level, after which college skill prices may be obtained by multiplying a state's high school skill price by its college premium. Bias from selective migration is a major concern when estimating college premia, so I use the semiparametric correction method described in Dahl (2002) to adjust my estimates. In particular, the paper presents a sample selection correction in a polychotomous choice Roy model that takes the form of an unknown function of the probability of the first-best (i.e.

[58]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.

the observed) location choice[59], where probabilities are computed by observing the migration choices of individuals first categorized into cells, thereby allowing a distribution-free estimate of selection probabilities.

I begin by taking white men in the ACS and Census aged 25 to 54 who either have exactly a high school or exactly a college degree. Individuals living in group quarters are dropped, and I make similar hours and income restrictions compared to before. I then categorize individuals into cells based on birth state, education, marital status, and whether they moved from their birth state. Married stayers are split further according to whether their spouse works, whether they have children less than 5 years old in the household, and whether they have children aged between 5 and 18 in the household. Non-married stayers are separated by whether they are divorced and whether they live alone, with roommates, or with extended family. Married movers are split up by whether they have any children, and non-married movers by whether they live with roommates/extended family; the smaller sample of movers necessitates using coarser grids. As in Dahl (2002), the fraction of individuals in a cell who move from one state to another determines the probability that any individual in the given cell follows the same migration path, and the proportion of individuals in a cell who stay in their birth state gives retention probabilities for all individuals in the cell as well.

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

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

the results.

The results of this correction are shown in Figure 1.B.3a and 1.B.3b. Similar to Dahl (2002), the correction results in a statistically significant lowering of the college premium (typically around 10%) compared to raw OLS estimates.

Robustness

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

First, I run a specification that follows Kennan and Walker (2011) that attempts to limit selection from migration even further by limiting the sample to high-school educated males aged 18-20, the intuition being that focusing on new labor market entrants preempts the bulk of migration decisions. The numbers I use are strongly correlated with the output of this method (correlation >0.8), though after normalizing by Iowa's skill price the other estimates of $w^{\ell,0}$ are slightly lower, suggesting that Iowa may have relatively high early entrant wages. A juxtaposition of the high school skill prices obtained in this test vs. the ones I use are available in Figure 1.B.4a.

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

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

where w_{it} indicates wage and x_i years of potential experience as before, e_i education, and δ_t, $\gamma_{S(it)}$, and λ_i fixed effects for calendar year, state, and individuals respectively. Standard errors are clustered at the individual level. The γ terms correspond to skill prices and can be identified from wage changes among individuals who move from one state to another. Estimates from this procedure compared to my baseline estimates may be viewed in Figure 1.B.4b. As before, the two sets of estimates are positively correlated, and only two of my baseline estimates fall outside the confidence interval for their corresponding estimate from this method.

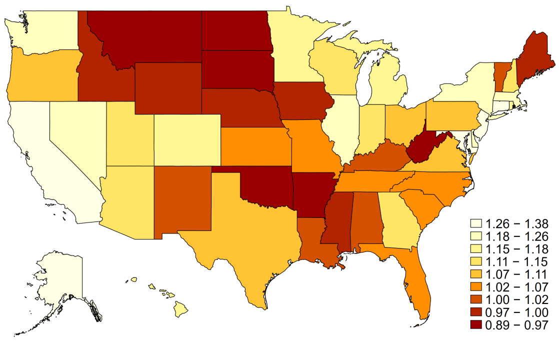
The key problem with both this and the previous test is that they result in very small sample sizes, particularly for low-population states. For any given survey year in the PSID, there are fewer than 10 individuals total in states such as Montana, Vermont, and Wyoming, rendering estimates of skill prices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

The larger samples I use in the baseline method also allow for a final selection test for the high school skill prices. In particular, I run the same Dahl (2002) procedure to estimate selection corrections for state-specific *high school* earnings premia relative to high school dropouts. I follow the exact same procedure as before but limit the sample to either individuals with exactly a high school degree or high school non-graduates. Figures 1.B.3c and 1.B.3d display the raw and corrected high school earnings premia obtained after this procedure. In contrast to the college premia estimates, the selection correction barely changes the estimates of high school premia at all, providing evidence that selective migration for high school graduates is not nearly as large a concern as for college graduates.

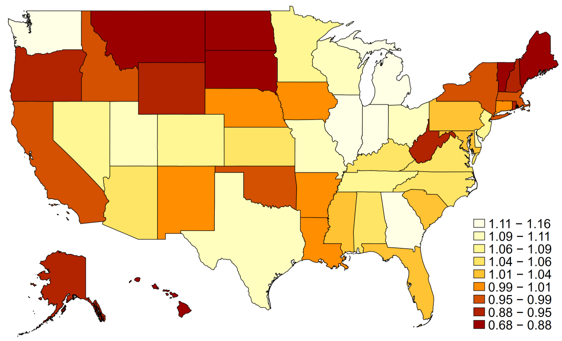
One notable limitation with the Dahl (2002) procedure is that it requires the assumption that one's birth state does not affect their earnings, which may be at odds with the model's structure that attempts to capture heterogeneous child human capital formation across locations. To gauge how problematic this assumption may be, I run an alternate version of Equation (1.3) that simultaneously estimates state-level skill prices by educational status and time period while fully interacting the experience quartic with one's state of birth. Conceptually, this allows for one's level of experience to matter differently for their earnings based on their state of origin in a flexible way. The skill prices recovered by this procedure are compared with the ones I use in my model baseline in Figure 1.B.5. Within any time period and educational category, the skill prices used in the model baseline are strongly correlated ($\rho > 0.9$) with the alternate skill prices. However, the lack of a selection correction for college skill prices is clearly reflected in the larger magnitudes of the estimates of the alternate procedure.

Notably, the skill prices for high school graduates are generally higher when interacting experience with birth place as well — since skill prices are normalized by 2000 Iowa high school graduates, this indicates that this procedure results in relatively lower skill prices for Iowa and places like it. This is consistent with the theme of the model: favorable demographic and economic conditions in the Great Plains and Mountain States may result in natives from these areas having a relatively higher level of underlying human capital given a fixed level of labor market experience, so the skill prices of the states themselves must be lower to rationalize the low wages observed in the data. Since relatively low wages result in the option to migrate having a larger effect on the expected returns of human capital investment, this means that, if anything, the assumption required in the Dahl (2002) selection correction may understate the role of migration predicted by the model. Given the strong correlation between these estimates and the ones I use in my model's baseline, however, moving to the alternate specification has little effect on the paper's key takeaways, and I prefer the baseline skill price estimates so as to be able to apply the selection correction in the first place.

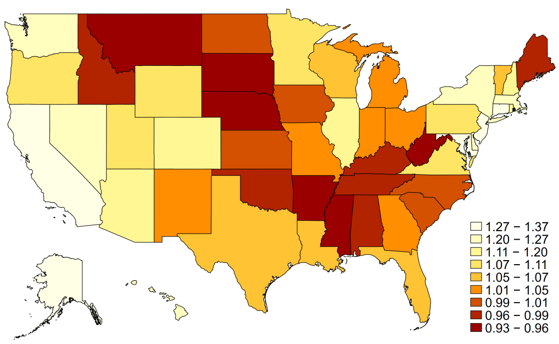
Figure 1.B.1: High School Skill Prices



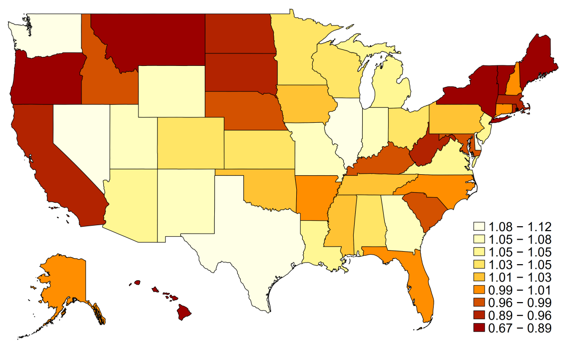
(a) 2000, Raw



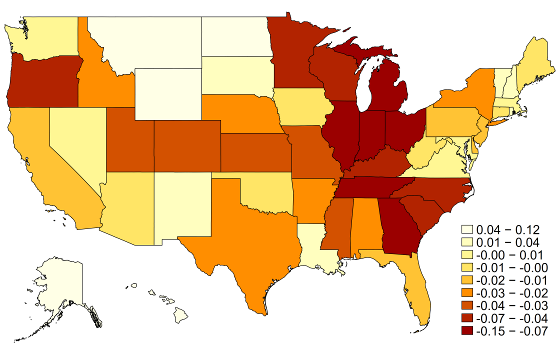
(b) 2000, Real



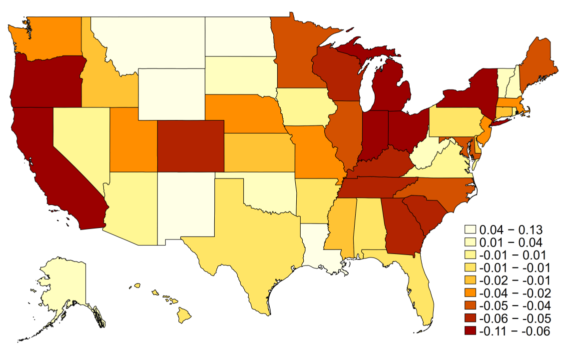
(c) 2010, Raw



(d) 2010, Real



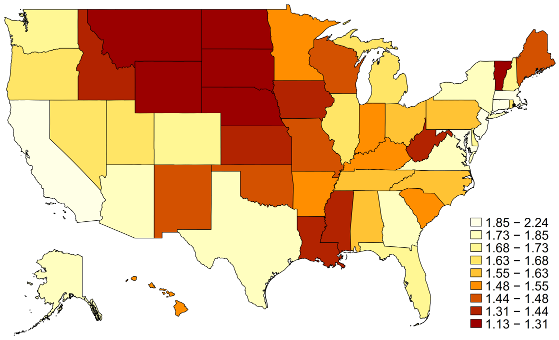
(e) 2000-2010 Change, Raw



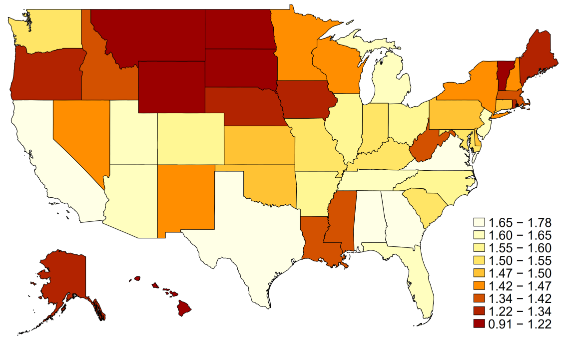
(f) 2000-2010 Change, Real

Notes: Figures visualize estimates of $w^{\ell,0}$. See Section 1.4.1 and Appendix 1.B.1 for details on estimation procedure. Subfigures (a) and (b) present estimates for the year 2000, both raw and after adjusting for local skill prices, obtained from the 2000 Census. Subfigures (c) and (d) present the corresponding statistics for the year 2010, obtained from the 2008-2012 American Community Surveys. Subfigures (e) and (f) visualize changes in skill prices in the 2000-2010 period.

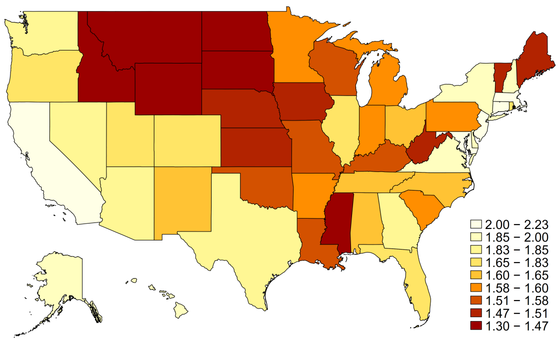
Figure 1.B.2: College Skill Prices



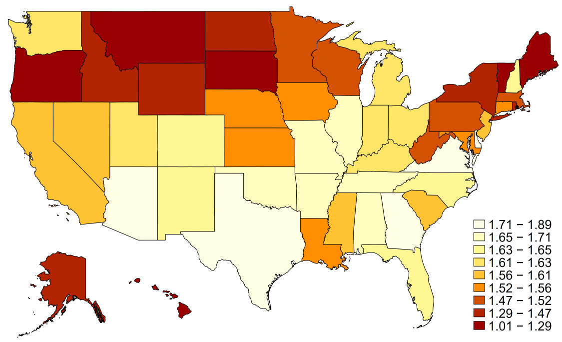
(a) 2000, Raw



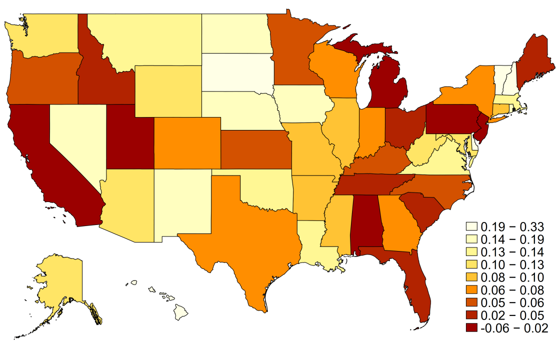
(b) 2000, Real



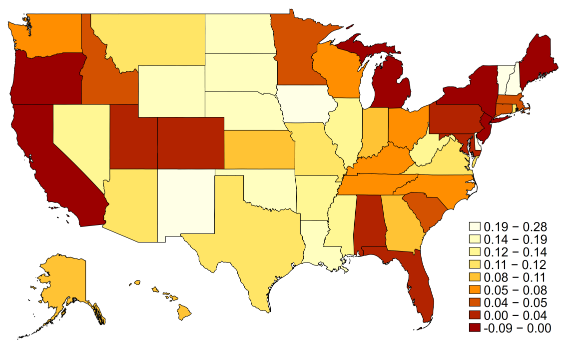
(c) 2010, Raw



(d) 2010, Real



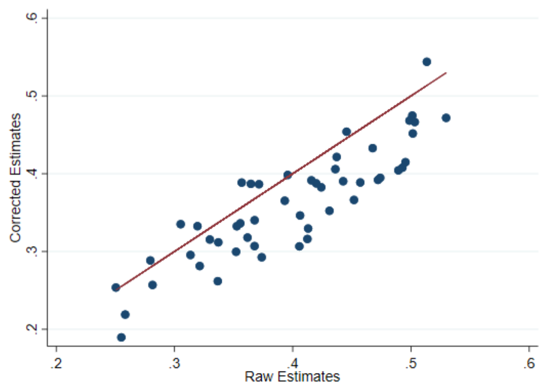
(e) 2000-2010 Change, Raw



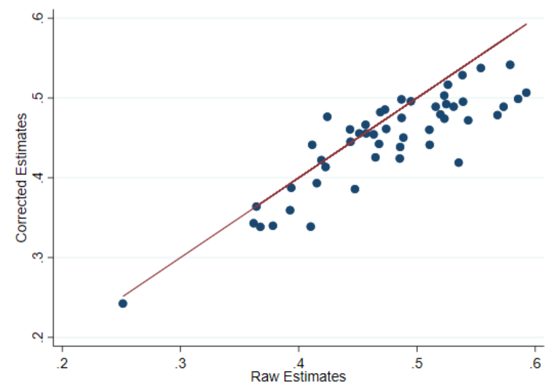
(f) 2000-2010 Change, Real

Notes: Figures visualize estimates of $w^{\ell,1}$. See Section 1.4.1 and Appendix 1.B.1 for details on estimation procedure. Subfigures (a) and (b) present estimates for the year 2000, both raw and after adjusting for local skill prices, obtained from the 2000 Census. Subfigures (c) and (d) present the corresponding statistics for the year 2010, obtained from the 2008-2012 American Community Surveys. Subfigures (e) and (f) visualize changes in skill prices in the 2000-2010 period.

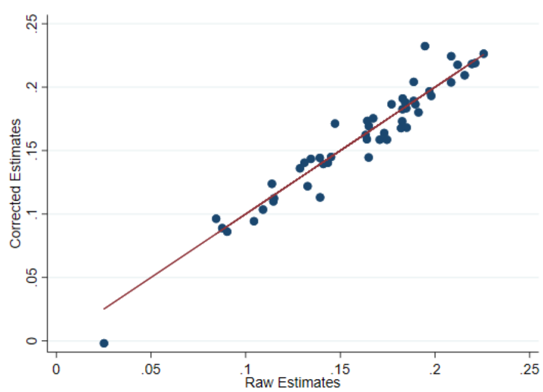
Figure 1.B.3: Skill Price Selection Corrections



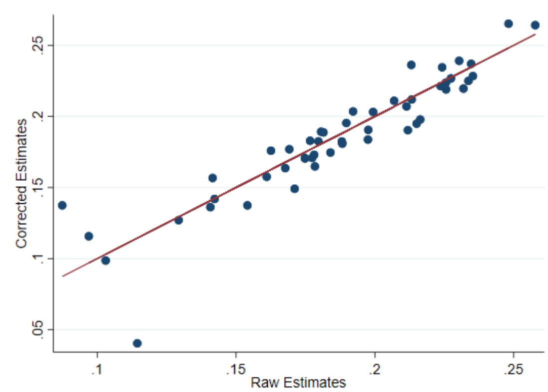
(a) 2000, College



(b) 2010, College



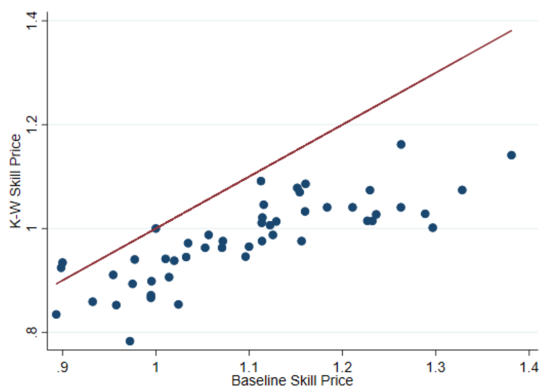
(c) 2000, High School



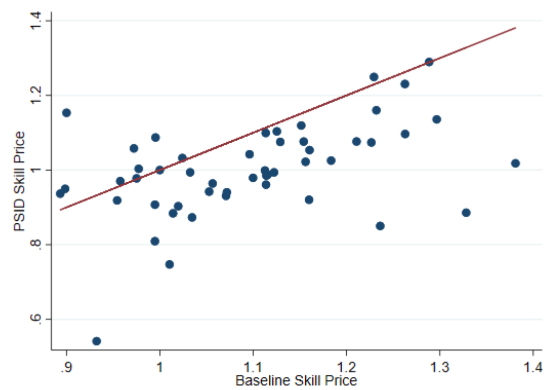
(d) 2010, High School

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

Figure 1.B.4: Robustness of High School Skill Price Estimates



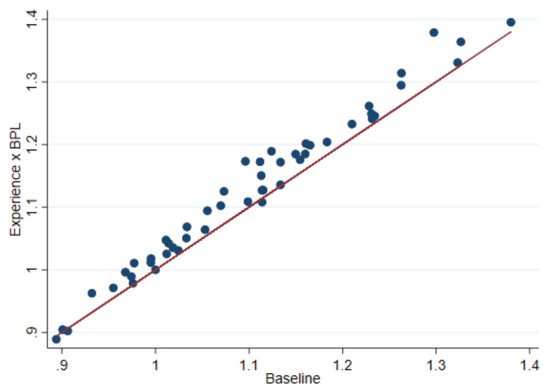
(a) K-W (Early Entrants)



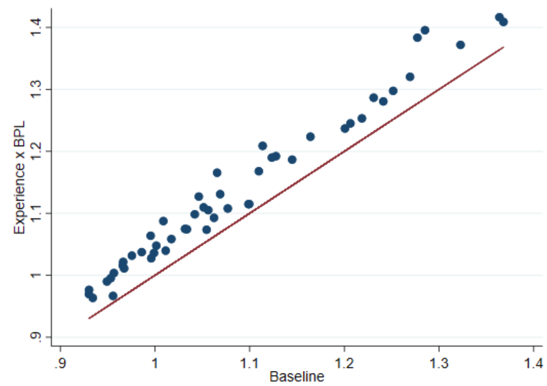
(b) PSID (Two-Way FE)

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

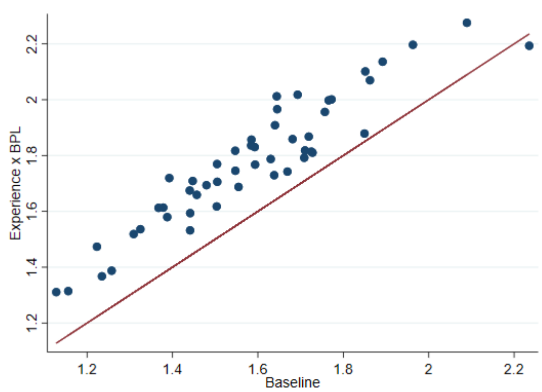
Figure 1.B.5: Skill Prices with Experience-Birth State Interaction



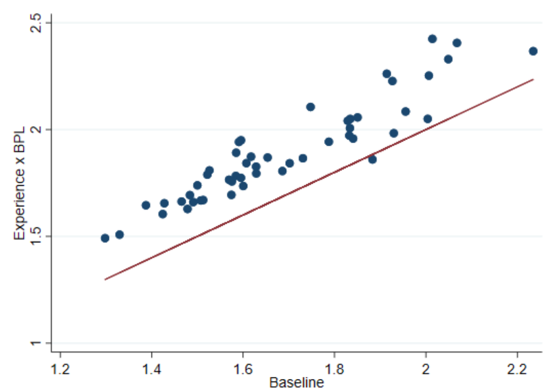
(a) 2000, High School



(b) 2010, High School



(c) 2000, College



(d) 2010, College

Notes: Figures display skill prices used in model baseline and skill prices estimated via Mincerian wage regression with a full interaction between quartic experience polynomial and state of birth. College skill prices in model baseline corrected for selection using method described in Dahl (2002). Subfigures (a) and (b) juxtapose baseline vs. alternate high school skill prices for years 2000 and 2010, and Subfigures (c) and (d) do the same for college skill prices.

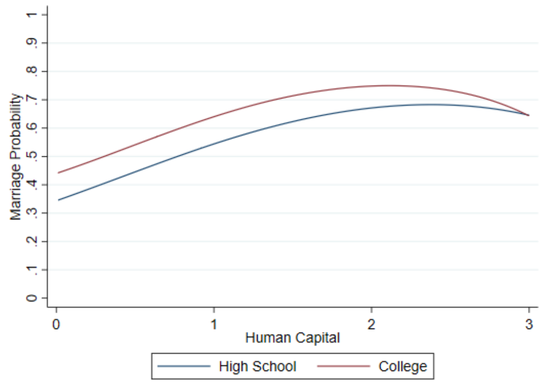
1.B.2 Marriage Realizations

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

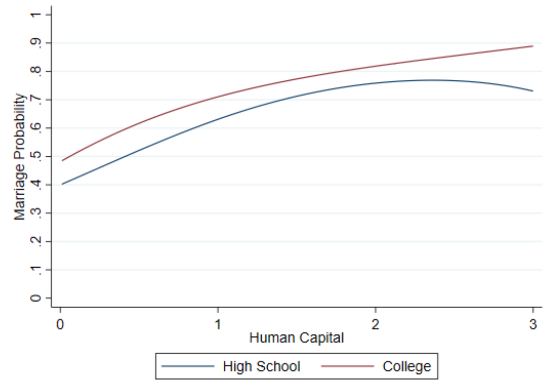
Figure 1.B.6 displays how marriage probabilities evolve over human capital based on state of residence and education. I show probabilities for Mississippi (the state with lowest overall marriage rates), Utah (the highest), and for Iowa, California, New York and Texas. In most cases, there is a clear gap in probabilities between high school and college-educated individuals, and the probabilities of marriage increase steadily over the human capital distribution before eventually leveling off. I hold marriage probabilities constant after a human capital level of 3, which corresponds roughly to the top percentile, to prevent the curvature from making perverse predictions about marriage probabilities for extremely high-human capital individuals. While many states are comparable, considerable heterogeneity is present: note, for instance, that the marriage probabilities for Utah high-schoolers is never below 50%, while in Mississippi the probability for high schoolers starts at barely 20%.

Figure 1.B.7 presents the model's fit of marriages rates for children with 25th-income-percentile parents by state of birth. The data for marriage rates of such children come from the Opportunity Atlas, while the model output corresponds to the average marriage rates of children with below-median income parents by state of birth. The fit is quite strong, with the model explaining more than 70% of the variation in state-level marriage rates, though the model underpredicts marriage rates across the board by a slight amount.

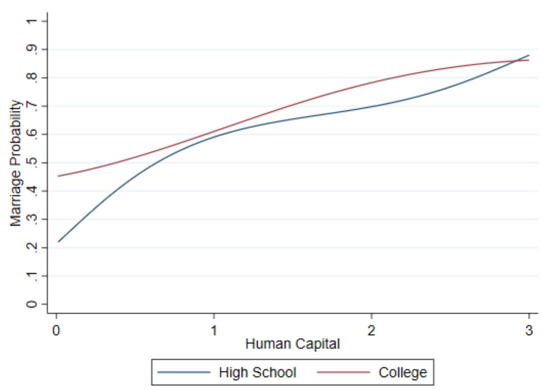
Figure 1.B.6: Marriage Probabilities by State



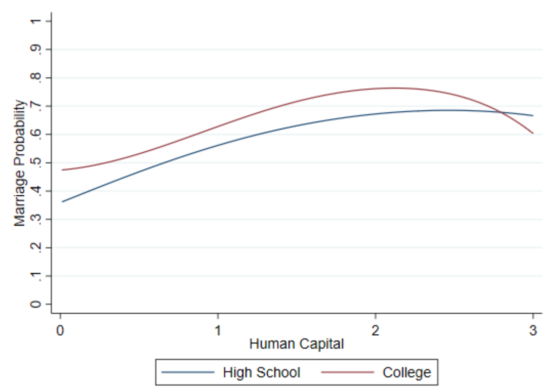
(a) California



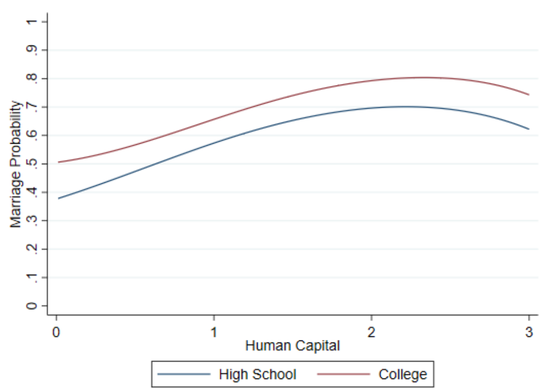
(b) Iowa



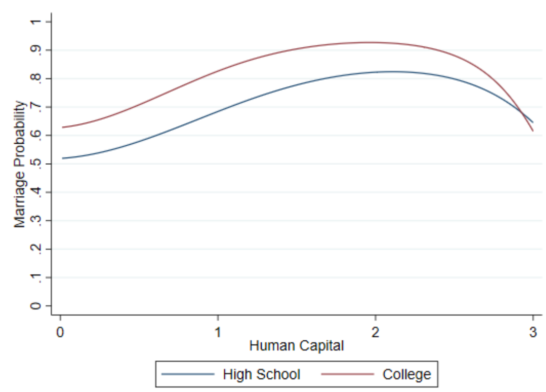
(c) Mississippi



(d) New York



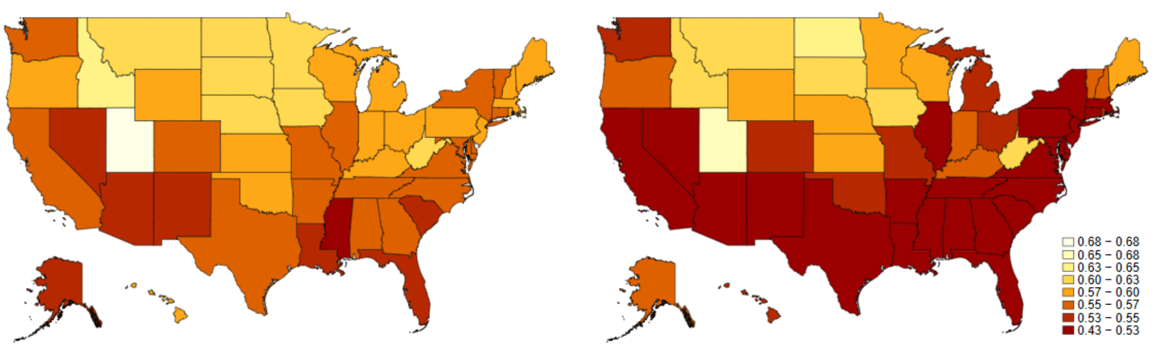
(e) Texas



(f) Utah

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

Figure 1.B.7: Model Fit of Marriage Rates



(a) Data (b) Model

Correlation: 0.87

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

1.B.3 Racial and Regional Heterogeneity Estimates and Fit

In this section, I describe the enhancements made to the model to account for racial and regional heterogeneity that are not presented in the main model for ease of notation. I allow for three different races in the model: non-Hispanic White, Black, and Hispanic. I allow for races to influence skill prices and marriage/fertility probabilities by factors that are constant across states[60]. Further, I estimate separate parameters for human capital productivity (ξ), migration preferences (δ_1), and college preferences (η_1) across races to enable the model to fit racial heterogeneity in wages, migration, and educational attainment. I solve the model and compute policy functions for each race before the simulation. In the simulation, I account for state-level differences in racial compositions as well as different proportions of races represented in different types of families. In estimation, I target these parameters to match racial wage ratios (obtained from my ACS sample), rates of racial college attainment (from the NLSY97), and rates of migration across races (again from the ACS).

For regional heterogeneity, I assume that the ability mean μ_a varies across the four Census regions to allow for some flexibility in spatial distributions of talent while maintaining a reasonable number of parameters to estimate. I allow the mean ability μ_a to vary by geography as opposed to the correlation between parent human capital and ability ρ_{ha} due to joint distributions between parent income and AFQT scores in the NLSY97 showcasing comparable correlations but noticeable mean shifts across the four Census regions. These parameters are targeted to improve the model's fit of economic and geographical mobility by state as well as by targeting regional rates of college attainment (from the NLSY97).

Table 1.B.1 presents estimates and standard errors for the parameters governing racial and regional heterogeneity in the model. I estimate lower human capital productivity parameters for both Blacks and Hispanics, as well as higher migration costs for Blacks and

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

Table 1.B.1: Racial and Regional Heterogeneity Estimates

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

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

Table 1.B.2: Model Fit in Racial and Regional Heterogeneity

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

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

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

Table 1.B.2 presents the model's fit of salient aspects of additional racial and regional heterogeneity. The model fits Black and Hispanic wages, migration rates, and educational attainment rates quite well, but while the model predicts regional college attainment rates that are comparable to those observed in the data, it does not succeed in producing the qualitative pattern of higher attainment in the Northeast and Midwest and lower attainment in the South and West.

Table 1.B.3: Correlation of Main Decomposition to Baseline under Alternate Specifications

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

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

1.B.4 Robustness to Alternate Specifications

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

The altruistic factor of utility in the model results in it having an infinite horizon, requiring stationarity assumptions for the model to be solved. In the baseline exercise, I solve the model both in the year 2000 and in the year 2010 (essentially, pre- and post-recession) and then use year-2000 policy functions when simulating parent investment and college decisions and year-2010 policy functions when simulating self-investment decisions and final migration choices for the CHKS cohorts. One may be concerned that the parents' lack of knowledge of future economic conditions may alter my model's predictions, so as a simple test I also simulate a specification where I only use year-2010 policy functions, so that all agents in the model always behave as they would in the post-recession world. The first row of Table 1.B.3

reports a high correlation (0.76) of the key predictions of this version of the model with my baseline results, providing some reassurance that the stationarity assumptions I employ are not key in driving my results.

The baseline version of the model also employs a simplistic treatment of location amenities, assuming that larger locations are higher-amenity due to the presence of larger cities. As a second robustness test, I experiment with additional notions of amenities. Specifically, I run model specifications that include environmental factors (including average distance to a coast and number of warm days per year), quality of life factors (including crime rates and establishments per capita), and political economy factors (including union power, measured as whether the state is a right-to-work state, and pension debt per capita). In each case, I re-estimate the model with these additional amenities before performing the main decomposition again. Table 1.B.4 reports parameter estimates for these other amenities and indicates that the inclusion of them typically does not meaningfully impact the model's fit. Moreover, the standard errors of the estimates indicate that the estimates are often imprecisely estimated or statistically indistinguishable from zero. Additionally, Table 1.B.3 again indicates that the model's key predictions are not sensitive to these alternate specifications.

Table 1.B.4: Alternate Amenity Estimates and Model Fit

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

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

1.C Supplementary Figures and Tables

Table 1.C.5: OLS Estimates for Various Correlates on CZ-Level IIM

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

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

Table 1.C.6: Additional Moments

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

(a) Income Quintile Transtions (Data/Model)

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

(b) Time Investments

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

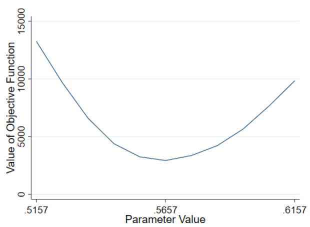
(c) College Attendance, Parents w/o Degree (Data/Model)

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

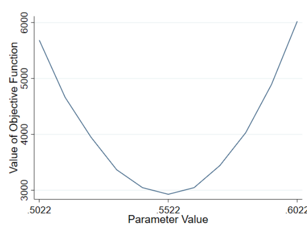
(d) College Attendance, Parents w/ Degree (Data/Model)

Notes: Table 1.C.6a reports income quintile transition probabilities between parents and children. Data moments from Table II of CHKS. Table 1.C.6b reports both total and individual parent time inputs for the children of married or unmarried parents. Data moments from PSID Child Development Supplement; see text for sample construction. Tables 1.C.6c and 1.C.6d report rates of college attendance by parent income quartile and ability quintile for kids with parents without and with a college degree. A star indicates fewer than 25 observations being in the cell and the moment not being used in estimation. Data from NLSY97; see text for sample construction.

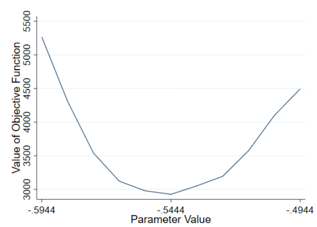
Figure 1.C.1: Behavior of Objective Function



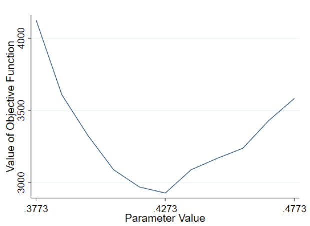
(a) θ



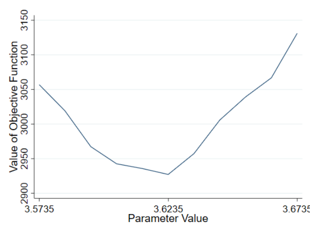
(b) ρ_{ha}



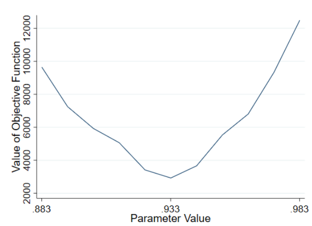
(c) μ_a



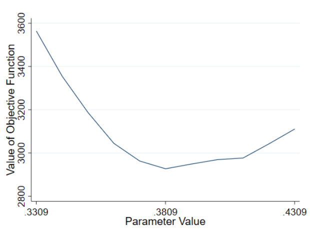
(d) σ_a



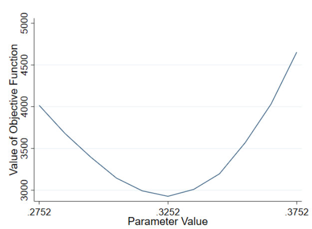
(e) ξ



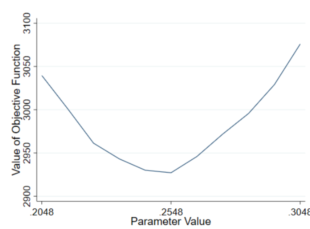
(f) ϕ



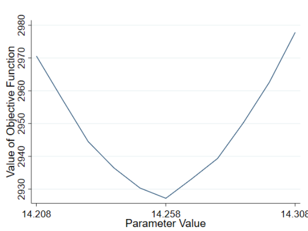
(g) κ



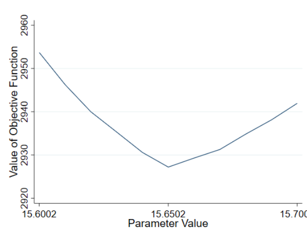
(h) σ_{ε_2}



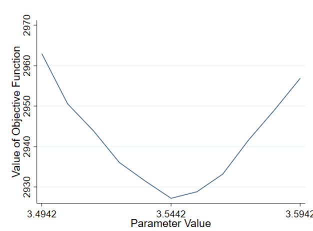
(i) α



(j) δ_1



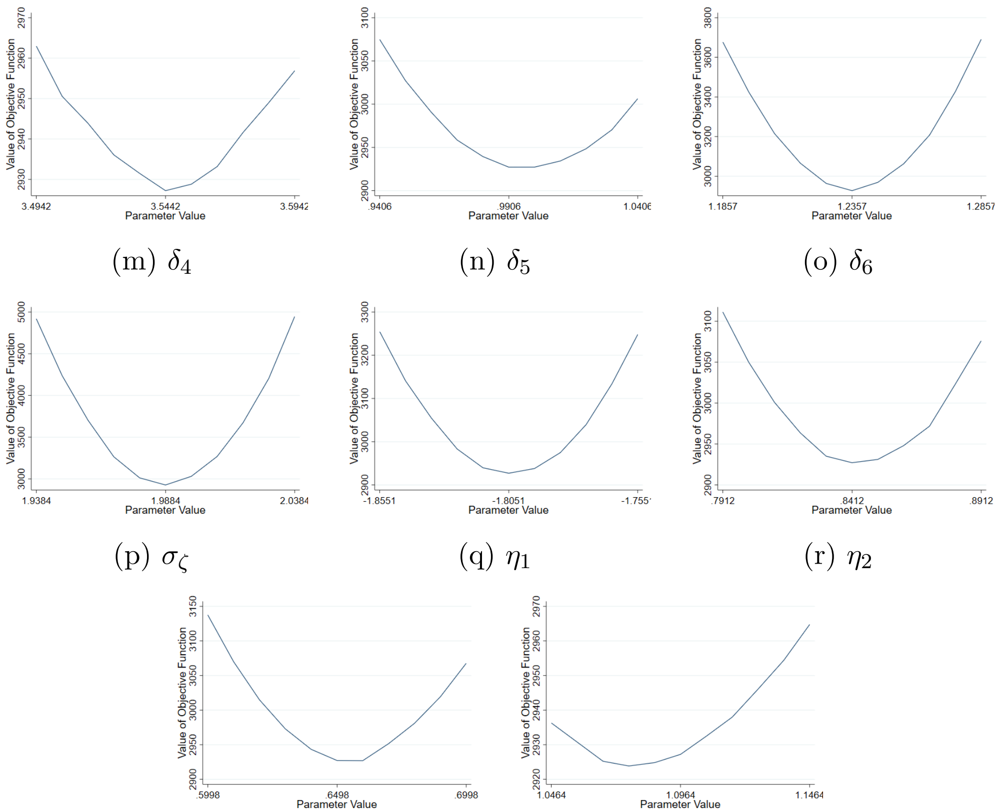
(k) δ_2



(l) δ_3

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

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



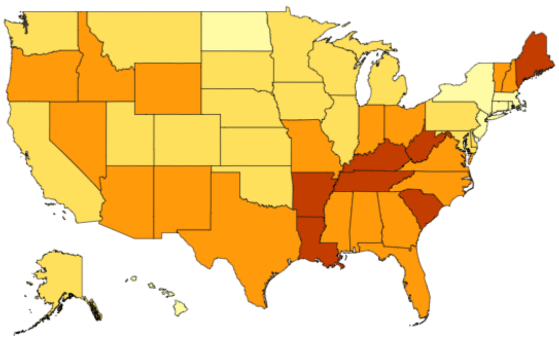
(m) δ_4 (n) δ_5 (o) δ_6

(p) σ_ζ (q) η_1 (r) η_2

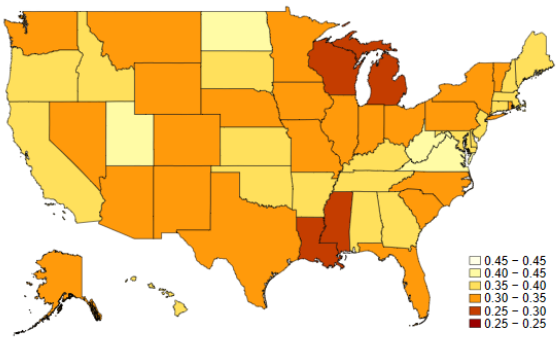
(s) η_3 (t) σ_η

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

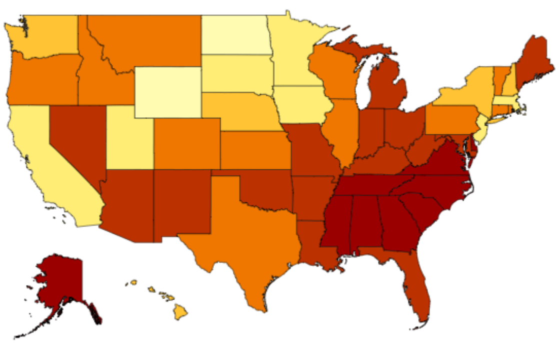
Figure 1.C.2: Additional Model Fit Visualizations



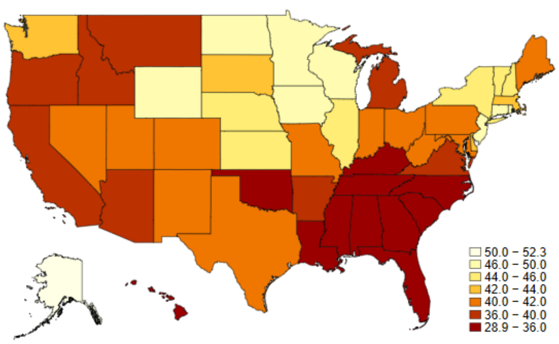
(a) College Attainment, Data



(b) College Attainment, Model



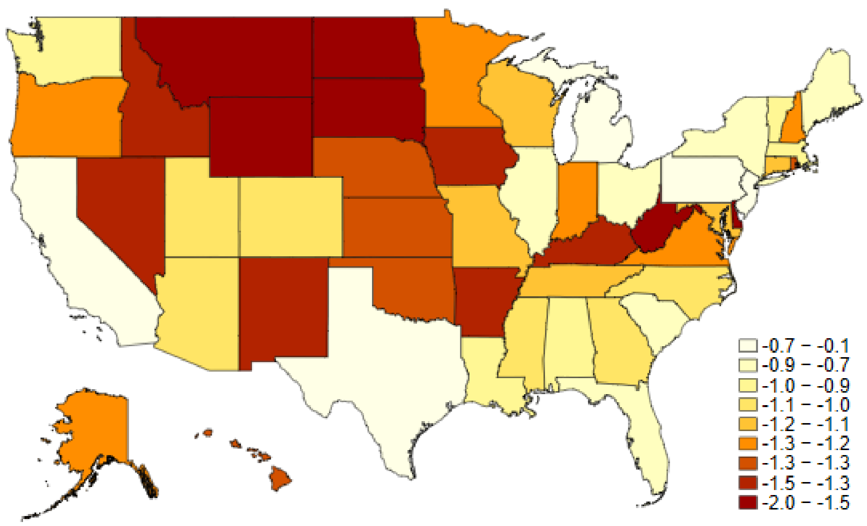
(c) IIM (Stayers) Data



(d) IIM (Stayers), Model

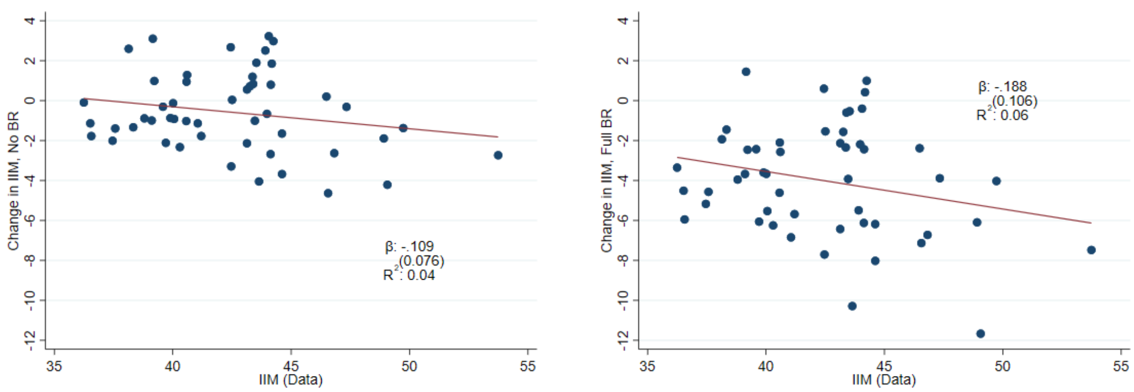
Notes: Figures present rates of college attainment and IIM among stayers as measured in the data and simulated in the model. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile in the data and expected family income percentile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

Figure 1.C.3: Utility Effects of Migration Restrictions



Notes: Figure 1.C.4 plots the change in utiles from counterfactuals that restrict migration while including behavioral responses.

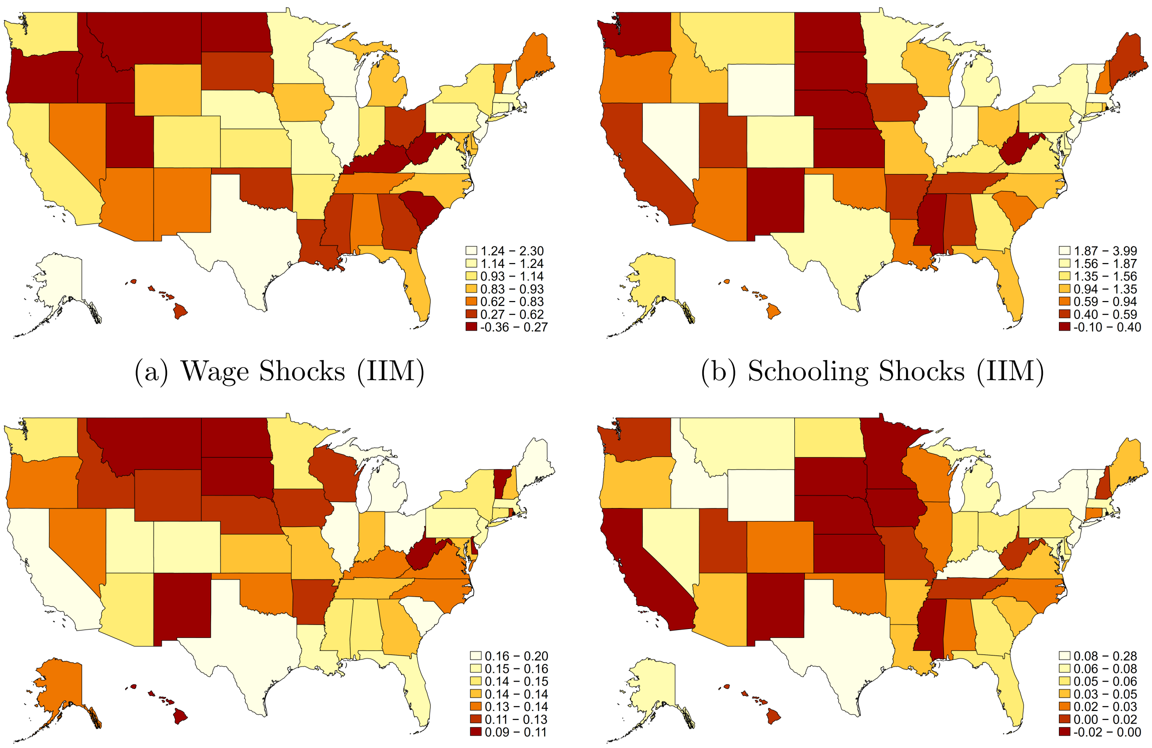
Figure 1.C.4: IIM Effects of Migration Restrictions, Scatterplot



(a) No BR (b) Full BR

Notes: BR = Behavioral responses. Figure 1.C.4 plots the change in upward mobility from counterfactuals that restrict migration while ignoring or including behavioral responses. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile in the data and expected family income percentile of children born to below-median income parents in the model. X-axis reports state-level IIM rates as measured by Chetty et al. (2014), while Y-axis reports model-predicted changes in IIM following the counterfactuals.

Figure 1.C.5: IIM and Utility Effects of Wage and Schooling Shocks



(a) Wage Shocks (IIM)

(b) Schooling Shocks (IIM)

(c) Wage Shocks (Utility)

(d) Schooling Shocks (Utility)

Notes: Figure 1.C.5 plots the change in upward mobility or utility from counterfactuals that either raise skill prices by 10% in a state or raise government school expenditure and decrease student-teacher ratios by 10% in a state. IIM measured as the expected family national income percentile of children born to parents in the 25th national income percentile in the data and expected family income percentile of children born to below-median income parents in the model.

Chapter 2

To Grandmother's House we Go: Childcare Time Transfers and Female Labor Mobility

Joint with Joanna Venator

Chapter Summary

Women in the United States frequently rely on childcare from extended family but can only do so if they live in the same location as them. This paper studies how childcare costs, the location of extended family, and fertility events influence both the labor force attachment and labor mobility of women in the United States. We begin by empirically documenting strong patterns of women returning to their home locations in anticipation of fertility events, indicating that the desire for intergenerational time transfers is an important motivator of home migration. Moreover, women who reside in their parent's location experience a substantial long-run reduction in their child earnings penalty. Next, we build a dynamic model of labor force participation and migration to assess the incidence of counterfactual

scenarios and childcare policies. We find that childcare subsidies increase lifetime earnings and labor mobility for women, with particularly strong effects for women who are ever single mothers and Blacks. Ignoring migration can understate the welfare benefits of these policies by a meaningful extent.

2.1 Introduction

How does childcare availability influence the labor force attachment and migration behavior of women in the United States? The cost of childcare in the United States is widely acknowledged to be a non-trivial financial hardship for many families: recent surveys indicate that the average cost of center-based infant care exceeds 27 percent of median income for single parents (Child Care Aware, 2017). One option that families may choose in response to these high costs is living near their own parents to use take advantage of cheap or free child-care — informal care is commonly used in the U.S.[1] but can only be used if relatives are nearby. Thus, childcare needs may constrain both the labor force *participation* and the labor *mobility* of U.S. women.

The central goal of this paper is study how migration choices are constrained by childcare needs and the implications for these location constraints for women's earnings. We begin with an empirical analysis that documents strong patterns of women moving to their birth state in anticipation of fertility events, suggesting that childcare assistance indeed plays a role in motivating home migration. Migratory mothers who are not moving back to their home location additionally appear to prefer states with lower childcare costs. We also find that women with children exhibit considerably stronger labor force attachment when living in their home state or in states with lower childcare costs. Lastly, we use panel data from the Panel Study of Income Dynamics to explore the impacts of a birth on women's earnings using an event study style design (a la (Kleven et al., 2019)) and analyze how this "child penalty" varies based on the mother's proximity to her own mother and on the local childcare costs. We show that women who give birth in the same state as their own parents experience a substantially smaller child earnings penalty than women who have a child elsewhere, suggesting that these immediate responses have long-run implications. In

[1]Roughly 20% of families with young children report use relative-provided childcare McMurry (2021).

all these cases, the effects are stronger among unmarried women.

These descriptive facts motivate the question of how subsidized childcare would alter the migration and working decisions of women in the U.S. Typically, analyses of the impacts of childcare subsidies focus on the direct impacts of such subsidies on labor force participation and human capital accumulation as the primary mechanism through which such subsides impact women's earnings.[2] However, if the high costs of non-subsidized childcare prevent households from moving far from their parents and thus from optimally sorting across labor markets, we might expect that there is a secondary effect on earnings and welfare stemming from reduced frictions in labor mobility.

To explore this mechanism, we construct and estimate a model that nests a canonical model of dynamic labor force participation in a model of dynamic migration. Women receive shocks to their fertility and marriage status at the beginning of every period, after which they choose their labor force attachment and whether/where to move. The women in the model with children must balance the trade-off between building experience (Eckstein and Wolpin, 1989) through labor force participation and paying more in childcare costs, though they retain the option to move to their parent's location so as to receive their assistance. Results from the model suggest that fully subsidizing childcare would increase lifetime wages for women who are ever single parents by over 9.5 percent on average and over 5.8 percent for women who are never single parents. These subsidies also encourage labor mobility for single mothers substantially, raising lifetime labor mobility by 3 percent. Ignoring migration when estimating the welfare gains of the policies results in understating them by a considerable extent, particularly for single mothers. Moreover, we find that labor mobility increases in counterfactual settings where children are born only to married women, suggesting that the recent increases in the share of single-parent families may have played a role in concurrent

[2]For an overview of the literature on the elasticity of women's labor supply to childcare costs, see (Del Boca, 2015).

declines in labor mobility. We also estimate our model separately for Black and white women and find considerably stronger effects for the former group.

Our paper expands upon and ties together three broad areas of research: the literature on childcare costs and women's labor force participation, the literature on the determinants of migration, and the literature on the implications of family-based ties for labor market outcomes.

First, our paper introduces a new mechanism that contributes to the 'child penalty' faced by mothers: increased job mobility frictions caused by location-specific childcare access. A long existing literature has documented the fact that women experience large dips in earnings following the birth of a child (Kleven et al., 2019, Cortes and Pan, 2020, Goldin and Mitchell, 2017, Budig and England, 2001, Angrist et al., 1998). One explanation for mother's dip in earnings is that the lack of affordable childcare forces women either out of the labor force or into part-time work, resulting in periods of low or zero earnings and lower earnings growth over time due to slower human capital accumulation. Analyses of programs which provide free or subsidized child-care/early childhood education in Canada (Baker et al., 2008, Lefebvre and Merrigan, 2008), Europe (Bauernschuster and Schlotter, 2015, Bettendorf et al., 2015, Havnes and Mogstad, 2011, Lundin et al., 2008), and the United States (Cascio, 2009, Tekin, 2007, Blau and Tekin, 2007) indicate that subsidized childcare increase the likelihood that women work, while also crowding out their use of informal care. This crowd-out is usually discussed in context of changing the quality of care received by children or as reducing the elasticity of women's labor supply with respect to care subsidies. We argue, however, that this substitution away from informal care is also a mechanism through which these subsidies may improve women's labor market prospects. By allowing women to no longer rely on relative care, they can be more mobile and potentially achieve welfare gains by moving to a more productive labor market than their parents live in.

By incorporating this mechanism into a dynamic model of labor force participation and fertility, we are building upon a long-standing strand of the women's labor participation literature which considers how childcare would change women's labor force participation decisions throughout the life cycle rather than just in the immediate aftermath of policy implementation. Our model builds directly on the frameworks of dynamic labor supply in presence of fertility seen in Eckstein and Wolpin (1989), Francesconi (2002), Haan and Wrohlich (2011), and Bick (2016). The latter two do incorporate private childcare costs, but we extend these models by incorporating informal care from spouses and grandparents as well as adding in a migration component. By incorporating these components, we aim to consider how childcare subsidies may change women's welfare not only through increasing their labor force participation, but by changing the labor markets in which they supply labor.

This approach is complementary to the analysis seen in Adda et al. (2017) which considers how women make decisions about both labor force participation and the occupation in which to work. Their model suggests that while three-quarters of the career costs of children are attributable to reduced labor supply, part of the loss is attributable to occupation choices in anticipation of fertility that reduce earnings. Our model considers a different margin — location, rather than occupation — through which women may be adjusting their labor supply to account for childcare needs. Our results similarly suggest that while the majority of the wage gains women might accrue from childcare subsidies are from increased participation, 7.5% of the increase in wages are attributable to women being able to sort into better labor migration through migration.

Second, our model also contributes to our understanding of the factors influencing return migration and home-biases in location choices. Older work has studied repeated and return migration (Davanzo, 1983; Dierx, 1988) with the view that such moves are driven entirely by monetary influences. Some more recent work (Diamond, 2016; Kennan and Walker, 2011; Bishop, 2008) considers non-monetary factors agents weigh when making repeated moving and location choices, but these papers typically condense preferences for living in one's home location into a single utility premium. A small literature has documented the role of emotional attachment to places' characteristics and the role of concentration of extended family in location decisions. (Boyd et al., 2005; Spilimbergo and Ubeda, 2004; Zabek, 2019; Spring et al., 2017). Through focusing on fertility as a new driver of home migration, we aim to contribute to the endeavor to unpack the specific determinants of return migration and add to the literature that studies how individuals balance pecuniary and non-pecuniary factors when making migration decisions in the United States.

In particular, these analyses may help to understand the factors underpinning recent declines in long-distance migration (Molloy et al., 2011). Recent research (Johnson and Schulhofer-Wohl, 2019) suggests that declines in the long-distance migration rate in recent generations is primarily a consequence of a decline in return migration. Johnson and Schulhofer-Wohl focus, however, on a different definition of return migration than the current paper – a move to any location one once lived in, rather than a move to the location one was raised in. Nonetheless, our results point to recent declines in U.S. fertility rates as a potential component of this observed drop in return migration.

Lastly, by focusing on home-based return migration, our results also marry the literature on migration with a growing literature on the implications of family-based ties for labor market outcomes. Proximity to family can mitigate child or elder care needs, allowing greater attachment to the labor force. Geographic distance from one's mother or mother-in-law is associated with a greater likelihood of childcare transfers, allowing for higher labor force participation for women (Compton and Pollak, 2015, 2014; Chan and Ermisch, 2015). To identify the effects of access to grandparent care, past research has used variation in pension generosity and retirement age (Dimova and Wolff, 2011; Aparicio-Fenoll and Vidal-Fernandez, 2015; Zamarro, 2020; Bratti et al., 2018; Posadas and Vidal-Fernandez, 2013) and the death of grandparents (Arpino et al., 2014; McMurry, 2021) to show that larger grandparent time transfers are associated with higher labor force participation and earnings for mothers. Beyond the realm of childcare, co-location near parents acts as a buffer against earnings losses for adult children following a job displacement (Coate et al., 2017; Kaplan, 2012). Conversely, care needs may flow in the opposite direction, with adult children living near parents having greater care responsibilities for aging or ill parents, resulting in worse economic outcomes (Charles and Sevak, 2005; Konrad et al., 2002; Rainer and Siedler, 2009).

To our knowledge, the only other paper that directly assess the role of informal childcare in influencing the migration choices of women is García-Morán and Kuehn (2017), who build a model of residence choice, fertility decisions, and female labor force participation in the context of Germany. Our contribution relative to their paper comes from our focus on dynamics: the authors model migration, working, and fertility decisions as one-shot choices the agent solves at the start of the model. However, labor force participation and migration are dynamic processes[3], which means our dynamic framework will better capture the life-cycle implications of childcare availability and policies. Our model of migration decisions is most complementary to Coate (2013), who considers a dynamic model of migration where agents take the location of their parents into account and are willing to accept lower wages in exchange for closer proximity to their parents. His model, however, focuses on early adulthood migration decisions and does not take fertility into consideration.

The paper is organized as follows: Section 2.2 motivates our research question by providing descriptive evidence regarding the timing of home migration and fertility events observed in U.S. data. Section 2.3 details our model, and Section 2.4 describes our estimation procedure. Section 2.5 presents model estimates and evaluates the model's fit, while Section 2.6 presents the results of counterfactual simulations. Finally, Section 2.7 considers potential avenues for future research before concluding.

[3]Multiple moves and return migration are salient features of the data (Kennan and Walker, 2011).

2.2 Motivation

In this section, we present empirical evidence to suggest that U.S. women respond to the incentives discussed in the introduction. We begin by showing in the American Community Survey (Ruggles et al., 2020) that U.S women frequently return to the birth state (which we take as a proxy for their parent's location for lack of a better alternative) in anticipation of fertility events and that those who have children in their home state exhibit markedly higher labor force attachment than those who live elsewhere. Next, to further motivate our focus on dynamics, we construct event-study representations of the child earnings penalty in the style of Kleven et al. (2019) using the Panel Study of Income Dynamics (PSID) and show that women who live in the same state as their grandparents experience a considerably smaller long-run child penalty than those who do not.

2.2.1 Fertility and Return Migration among U.S. Women

With how costly childcare is in the United States, one may expect women with small children to make different location and working decisions than those without. In particular, we may expect women with young children to be more likely to move back to their parent's location to take advantage of familial support in raising young and for women with children to work more hours if their parents are in their same location than if not. Women with small children should also be more reluctant to move to locations with higher childcare costs than those without, other things equal.

To test these hypotheses, we use data from the 2005-2017 waves of the American Community Survey (ACS). Each year of the ACS contains a 1 percent sample of the entire United States' population, providing a large number of observations. Additionally, the 2005-onward waves of the ACS also contain information on one-year migration histories (or, the state that respondents were in the year before their interview). While somewhat limited, this information, coupled with extensive demographic information and state-level measures of childcare costs, will allow us to observe some simple margins of behavior that the presence of young children influences.

We restrict our ACS sample to women aged 22-35 who were born in the United States. We drop individuals that did not complete at least one year of high school education and exclude observations that either report working more than 75 hours per week on average or who have negative income. The women in our sample are limited to those who are coded as household heads, spouses of household heads, or children/children-in-laws of household heads (to allow for the possibility of "boomerang migration," or individuals moving back into their parents' home). The ACS additional records the youngest own child in for all respondents, allowing us to distinguish women who have young children from those who do not. We exclude observations whose age and age of youngest child imply a birth before the respondent was age 14.

We first investigate whether women are more likely to move home in response to fertility events. We restrict our sample to women who were not living in their state of birth in the year before the interview and then run the linear probability model:

$$h_{it} = \beta_0 + \boldsymbol{\beta}_1 \boldsymbol{X}_{it} + \beta_3 f_{it} + \tau_t + \varepsilon_{it},$$

where h_{it} indicates whether individual i moved back to their birth state in year t.[4] $\boldsymbol{X}_{it}$ contains a vector of demographic controls, while f_{it} indicates individual i's first fertility status in year t, defined by presence of a child belonging to the respondent that is less than 1 year old while also being the only child of the respondent in the household. The term τ_t contains year fixed effects, while ε_{it} is an error term. Standard errors are corrected for heteroskedasticity, and regressions are weighted using sampling weights provided by the

[4] The variable is scaled to be either 0 or 100 — thus, regression coefficients can be interpreted simply as percentage point changes to the likelihood of a home move.

Table 2.1: Effects of First Pregnancy on Home Migration Probability (HMP)

	(1)	(2)	(3)	(4)
VARIABLES	HMP	HMP	HMP	HMP
Mean Dep Var	4.03	4.44	3.58	4.44
First Pregnancy (FP)	0.608	2.237	0.259	2.076
	(0.198)	(0.525)	(0.204)	(0.687)
FP × High Childcare Costs				0.361
				(1.059)
Age	-1.573	-1.817	-1.749	-1.816
	(0.141)	(0.194)	(0.233)	(0.194)
High School Degree	0.0123	-0.0538	0.0639	-0.0530
	(0.195)	(0.253)	(0.304)	(0.253)
College Degree	0.863	1.525	0.219	1.524
	(0.194)	(0.255)	(0.302)	(0.255)
Sample	All	Non-Married	Married	Non-Married
Observations	572,964	279,471	293,493	279,471
R-squared	0.008	0.010	0.008	0.010

Notes: Robust standard errors in parentheses. Sample is US-native women aged 22-35 in the 2005-2017 ACS who completed at least one year of high school and were not located in birth state the previous year. Additional controls include fixed effects for birth state and calendar year, a quadratic in age, an indicator for some college attained, amenity measures for state lived in last year (college share, unemployment rate, rates of violent and property crime, population, per-capital government student expenditure, student-teacher ratios, and share of days warmer than 70 degrees) and Black and Hispanic indicators. First pregnancy indicator defined by presence of a child less than one year old while being the only own child of the respondent in the household. Regressions weighted by sampling weights.

ACS. We focus on the first pregnancy because the presence of additional children may make migration more cumbersome — thus, women may be more likely to move home in response to their first fertility event than subsequent ones. We also run our specification for all women as well as for non-married and married women separately, as the additional spousal financial support available to married women may make them less likely to migrate in response to fertility than single women.

Tables 2.1 and 2.2 report the results of this exercise for two different specifications of f_{it}. In Table 2.1, f_{it} is a dummy variable equal to 1 if the respondent had their first child in the previous year, which may be observed by the presence of a child of the respondent's that is

Table 2.2: Effects of Young Child on Home Migration Probability (HMP)

	(1)	(2)	(3)	(4)
VARIABLES	HMP	HMP	HMP	HMP
Mean Dep Var	4.03	4.44	3.58	4.44
Young Child (YC)	-0.0212	0.174	0.0240	0.156
	(0.0743)	(0.134)	(0.0889)	(0.175)
YC × High Childcare Costs				0.0398
				(0.260)
Age	-1.570	-1.844	-1.752	-1.843
	(0.141)	(0.195)	(0.233)	(0.195)
High School Degree	0.0132	-0.0259	0.0676	-0.0260
	(0.195)	(0.253)	(0.304)	(0.253)
College Degree	0.867	1.558	0.234	1.556
	(0.194)	(0.258)	(0.301)	(0.258)
Sample	All	Non-Married	Married	Non-Married
Observations	572,964	279,471	293,493	279,471
R-squared	0.008	0.010	0.008	0.010

Notes: Robust standard errors in parentheses. Sample is US-native women aged 22-35 in the 2005-2017 ACS who completed at least one year of high school and were not located in birth state the previous year. Additional controls include fixed effects for birth state and calendar year, a quadratic in age, an indicator for some college attained, amenity measures for state lived in last year (college share, unemployment rate, rates of violent and property crime, population, per-capital government student expenditure, student-teacher ratios, and share of days warmer than 70 degrees), and Black and Hispanic indicators. Young child defined as presence of own child aged at most 4 in household. Regressions weighted by sampling weights.

less than 1 year old while also being the only child of the respondent in the household. In Table 2.2, f_{it} is a dummy variable equal to 1 if the respondent has any children four years old or younger in the household. Conceptually, we may the presence of children to make migration more cumbersome and costly, so women may be more likely to move home in response to their first fertility event than subsequent ones. At the same time, the additional spousal financial support available to married women may make them less likely to migrate in response to fertility than single women.

These predictions are well born-out in the data. While Table 2.2 indicates that the presence of small children does not meaningfully influence the likelihood of a home move,

Table 2.1 suggests that initial fertility events make women noticeably more likely to home-migrate. These effects are also much stronger for single women than married women — indeed, the subgroup analysis indicates that married women respond to initial fertility events by migrating quite little. However, with a base home rate of migration of around 4 percent in the data, initial fertility events make single women roughly 50% more likely to move home compared to the rest of the sample. We also find that having previously lived in a state with high childcare costs[5] is associated with a slightly higher likelihood of moving back home, though this effect is not statistically significant.

While higher childcare costs do not seem to substantially influence the extensive margin of probability of a move itself, they may still have intensive margin impacts in that they could distort the location choices of women conditional on moving in the first place. We next investigate whether the presence of young children distort the location choices of women who choose to migrate. We limit our ACS sample to women who are observed to have moved from their previous-year state and have not moved to their state of birth. We then test whether the presence of young children result in women being less likely to locate in high childcare-cost states, defined as being above-median as before. Table 2.3 presents the results of this test and affirms the hypothesis — moving women with young children on average choose to locate to states with lower childcare costs than those without, with the effects again being noticeably stronger for single women than married ones.

Finally, we investigate how location and the presence of children influence the labor force attachment of women in the ACS. The ACS records usual hours worked per week for all employed respondents — for unemployed respondents or respondents not in the labor force, we code usual hours worked per week as zero. We then regress usual hours worked per week

[5]Defined as having above median costs, with numbers coming from Child Care Aware (2017), who survey state childcare resource and referral networks to obtain average prices for full time childcare centers for three age groups in each U.S. state. For a visual representation of average full-time infant childcare expenses across U.S. states, refer to Figure 2.A.1a.

Table 2.3: Effects of Children on Probability of Moving to High CCC State (HCS)

	(1)	(2)	(3)	(4)	(5)
VARIABLES	HCS (0-100)	HCS (0-100)	HCS (0-100)	HCS (0-100)	HCS (0-100)
Mean Dep Var	50.48	52.86	47.96	42.86	47.96
Young Child	-1.693	-4.816	-0.412	-4.647	-0.387
	(0.654)	(1.271)	(0.763)	(1.258)	(0.760)
High School Degree	3.895	2.590	5.240	2.153	5.519
	(1.897)	(2.521)	(2.831)	(2.518)	(2.796)
College Degree	13.52	10.88	15.07	10.61	15.28
	(1.853)	(2.478)	(2.786)	(2.481)	(2.751)
Sample	All	Non-Married	Married	Non-Married	Married
Home State FE	YES	YES	YES	YES	YES
Previous State FE	NO	NO	NO	YES	YES
Observations	54,233	26,582	27,651	26,582	27,651
R-squared	0.069	0.088	0.052	0.106	0.063

Notes: Robust standard errors in parentheses. Sample is US-native women aged 22-35 in the 2005-2017 ACS who completed at least one year of high school and who moved in the previous year and not to their state of birth. Additional controls include fixed effects for birth state and calendar year, a quadratic in age, an indicator for some college attained, amenity measures for state lived in last year (college share, unemployment rate, rates of violent and property crime, population, per-capital government student expenditure, student-teacher ratios, and share of days warmer than 70 degrees), and Black and Hispanic indicators. Young child defined as presence of own child aged at most 4 in household. Regressions weighted by sampling weights.

on a variety of covariates to do with the presence of children, location, childcare costs, and marital status. Intuitively, higher childcare costs ought to decrease hours worked by women because it makes working relatively more expensive. Being proximal to parents ought to increase labor force attachment if parents primarily provide time transfers in child-rearing, but effects of marital status on labor force attachment are a-priori ambiguous. Women with husbands may exhibit higher labor force attachment if their husbands also provide time assistance in raising their progeny, but if a husband's primarily role is to slacken budget constraints by providing supplementary income, then we ought to see married women with children work less than single women, other things equal.

Table 2.4 presents the results of this exercise, with many of the above predictions clearly manifesting in the data. The presence of children decreases usual weekly hours worked by

Table 2.4: Effects of Children and Location on Hours Worked

	(1)	(2)	(3)	(4)
VARIABLES	Hours	Hours	Hours	Hours
Mean Dep Var	27.27	27.27	27.80	26.64
Married	0.291	0.292		
	(0.0403)	(0.0403)		
Child Present (CP)	-2.640	-4.436	-2.902	-9.411
	(0.0574)	(0.0875)	(0.128)	(0.0960)
In Home State	0.609	-0.324	-1.058	1.270
	(0.0364)	(0.0430)	(0.0548)	(0.0692)
CP × Married	-5.577	-5.475		
	(0.0723)	(0.0723)		
CP × High Childcare Costs		-0.502	-0.640	-0.683
		(0.0665)	(0.112)	(0.0882)
CP × In Home State		2.726	1.036	2.260
		(0.0756)	(0.133)	(0.0996)
Sample	All	All	Non-Married	Married
Observations	2,056,614	2,056,614	1,014,536	1,042,078
R-squared	0.117	0.118	0.116	0.128

Notes: Robust standard errors in parentheses. Sample is US-native women aged 22-35 in the 2005-2017 ACS who completed at least one year of high school. Additional controls include fixed effects for birth state and calendar year, a quadratic in age, an indicator for some college attained, amenity measures for state lived in last year (college share, unemployment rate, rates of violent and property crime, population, per-capital government student expenditure, student-teacher ratios, and share of days warmer than 70 degrees), and Black and Hispanic indicators. Young child defined as presence of own child aged at most 4 in household. Regressions weighted by sampling weights.

women substantially, and the effects are noticeably stronger for married women. However, women who have children in their birth state work more than women who do not, while women with children in states with higher childcare costs also work relatively less.

2.2.2 Grandparent Proximity and the Child Penalty

While our analyses using ACS data provide a snapshot of the impacts of young children on household location choices, cross-sectional data cannot tell us the long-term impacts of living near relatives or living in low child-care cost regions on women's lifetime earnings trajectory. It is well-documented that women experience a decline in earnings following births, often referred to as the 'child penalty,' which persists for up to ten years post birth (Kleven et al., 2019), This child penalty is common across a number of European and North American countries (Kleven et al., 2019), ranging from a 20% decline in women's earnings relative to pre-birth earnings in Scandinavian countries to 44% in the US or 60% of pre-birth earnings in Germanic countries. A large factor in the decline in earnings is women's withdrawal from the labor market. Therefore, we might expect that having access to cheaper or free childcare would allow women to work more hours and reduce the child penalty.

To test this, we use data from the Panel Study of Income Dynamics (PSID) to estimate the size of the child penalty for women living near or far from grandparent care and for women living in high vs. low child-care cost regions. We adopt a modified form of the event study specification first proposed by Kleven et al. (2019). Because the authors show that there is no child penalty for men, we focus solely on women's first births. For each mother in the data, we define event time (t) based on the year of their first child's birth. Our outcome of interest is woman i's earnings Y_{ist} in year s and at event time t. The regression is as follows:

$$Y_{ist} = \sum_{j \neq -1} \alpha_j \mathbb{1}[j = t] + \sum_k \beta_k \mathbb{1} k = age_{is}] + \sum_n \gamma_n \mathbb{1}[n = s] = \epsilon_{ist}. \tag{2.1}$$

The regression contains event-time dummies with α coefficients, age dummies with β coefficients to control for life-cycle trends, and year dummies with γ coefficients to control for time trends. Event-time $t = -1$ is omitted, so all estimates are relative to the year just prior to birth. As noted in Kleven et al. (2019), we are able to identify effects of all three sets of dummies because of the variation in the age at which women have children.[6]

The parameters of interest are the α parameters, but they will represent differences in levels. To transform them into percent changes, we calculate $P_t = \frac{\hat{\alpha}_t}{\hat{\mathbb{E}(Y_{ist}|t)}}$, where the bottom

[6]For more details on the identification assumptions needed to assume this is the causal impacts of child birth, see Kleven et al. (2019).

of the fraction is the predicted outcome omitting the contribution of the event-time dummies.

We estimate this regression for all women living in the same state as the mother of the mother (henceforth the grandmother), all woman living in different states than the grandmother, women living in states in the top half of the childcare cost distribution, and women living in states in the bottom half of the childcare cost distribution. For these comparisons of the child penalty across groups, our figure of interest is the child penalty gap:

$$\frac{\hat{\alpha}_t^1}{\hat{\mathbb{E}}(Y_{ist}^1|t)} - \frac{\hat{\alpha}_t^2}{\mathbb{E}(\hat{Y}_{ist}^2|t)}.$$

We use the Delta method to calculate standard errors of this gap and then test whether we can reject the null that the child penalty is equal for those living in the same state as the grandmother and those living in different states.

Note that these estimates should not be interpreted as the causal impact of living near a grandparent or in a childcare cost region on the child penalty. We expect that women are sorting across these locations in part based on their attachment to the labor force; women who want to continue working after a birth for reasons unobservable to us as econometricians are more likely to settle in places with affordable childcare, whether that be relative care or cheaper private care options. We cannot separate these indirect selection effects from the direct effects of having cheaper childcare available. Nonetheless, these patterns will provide suggestive evidence of whether childcare cost factors are meaningfully related to the long-term child penalty women face following their first birth.

Data

For this analysis, we need a panel of income data for women in the years prior to and the years following their first birth. To create this, we use the PSID's full retrospective history of births and adoptions, which provides the full history of births for those interviewed in the years 1985 onward. Using this data set, we create a sample of all PSID women who have at least one birth, the year their first birth occurred, their age at that birth, and whether they were married at the time of that birth. We then combine this data with information from the PSID family files on earned income in each year of the women's life, the US state they live in in each year, and the US state that their parents live in each year. Earned income is defined as the reported total income including wages and other income.[7]

Following the restrictions used by Kleven et al. (2019), the panel of years includes five years pre-birth and ten years post-birth. Women are excluded from the sample if they are missing more than 8 years in this period, missing all years pre-birth, or all years post-birth. We also restrict the data to be from 1985 onward in part to reduce measurement error from retrospective birth histories and in part to match the data sample cleaned for the estimation sample, which only contains locations from 1985 onward.

We look at differences in the child penalty across three categories of mothers, as well as the interaction between these categories:

1. Near vs. Far to Grandma: A woman is near to Grandma if the mother is in the same state as her own mother in the year of her first birth.

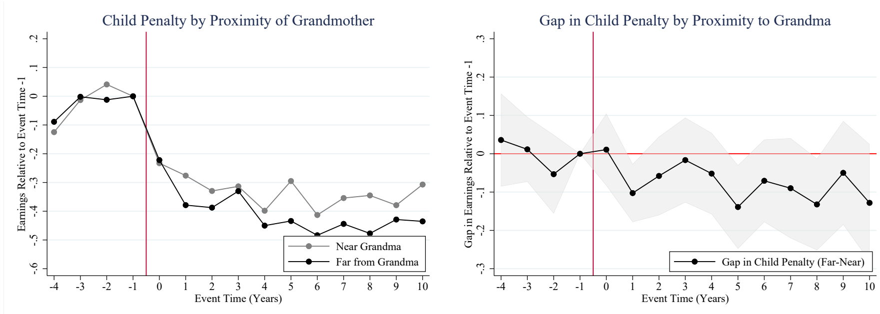
2. High vs. Low Childcare Costs: A woman is a high childcare cost type if she lives in a state that has childcare costs above the median of our CC cost index in the year of her first birth.

3. Married vs. Unmarried: A woman is married if she was married in the year of her first birth.

[7]For women who were household heads, this is based on reported income from 'wages and other income". For women who were spouses in the data, the measurement of income changes in 1993 when they begin separating out business and farm income. Due to this change, we create total income for spouses by adding together the total income excluding business and farms and total income from businesses for all years post 1993. Income is then assigned by the sex of the head of household.

Figure 2.1: Child Penalty for Women Living Near or Far from Grandmother

Panel A   Panel B

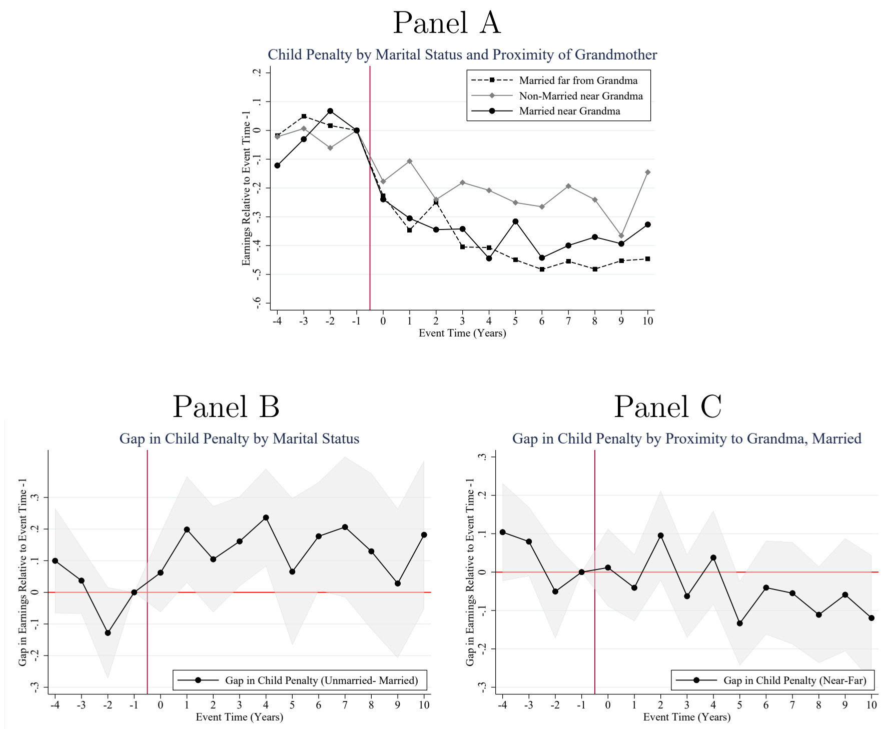


Notes: Figure 2.1A (left) plots coefficients from event studies of earnings on indicators for years surrounding a woman's first birth for both women who live in the same state as the grandmother (near) or different states (far). The unit are percent changes (0 to 1) in earnings relative to the year prior to birth. The regression includes controls for age of mother at first birth and year of birth. Figure 2.1B (right) calculates the gap for those near vs. far and reports 10% confidence intervals for a test of the null that this gap is equal to zero.

Results

Figure 2.1 plots the coefficients from the event studies described in (2.1) with panel A plotting the coefficients separately for mothers living in the same state or in different states from the child's grandmother and panel B plotting the size of the gap between these groups. While both types of mothers experience a large child penalty, those living distant from the grandmother experience a 10 percentage point larger child penalty that persists for up to 10 years following the child's birth. When we split the results by marital status (Figure 2.2), we see that the effects are driven primarily by smaller earnings losses for single mothers living near the child's grandmother rather than married mothers.

Figure 2.2: Child Penalty for Women Living Near or Far from Grandmother, by Marital Status



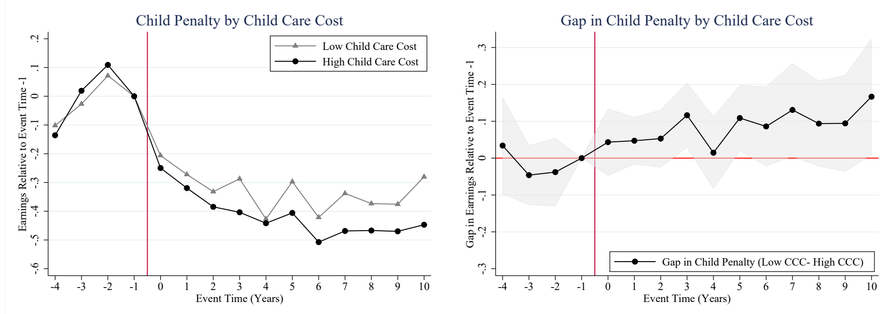
Notes: Figure 2.2A (top) plots coefficients from event studies of earnings on indicators for years surrounding a woman's first birth for married women who live in the same state as the grandmother, married women who live in different states, and unmarried women who live in the same state as the grandmother. The unit are percent changes (0 to 1) in earnings relative to the year prior to birth. The regression includes controls for age of mother at first birth and year of birth. Figure 2.2B (bottom left) calculates the gap in the percent decline for those near the grandmother who are married vs. unmarried. Figure 2.2C (bottom right) calculates the gap in the percent decline for those married near vs. far from the grandmother. Both report 10% confidence intervals for a test of the null that this gap is equal to zero.

Table 2.5 reports the results of a regression which aggregates the coefficients into pre-period (excluding the year prior to birth), year of birth, and post-period and interacts these periods with an indicator for living near the grandmother. Here, we see that the effects of living near the grandmother are statistically significant in the full sample and recoups about $2,700, or 22% of the child penalty faced by mothers. In contrast, for single mothers the effects are much larger, around $8,400 or 66% of the child penalty faced by single mothers. The benefit of living near grandmothers is smaller and not statistically significant for married mothers.

We next do a similar exercise for those living in high or low childcare cost states. Figure 2.3 reports the coefficients for event studies of earnings on indicators for years surrounding a woman's first birth for both women who live in the low or high childcare cost states. We see that the child penalty is larger in states with high childcare costs. The difference across childcare cost regions is of similar magnitude to the difference in the child penalty for those near vs. far from the child's grandmother. Interestingly, the effects of childcare on the child penalty seem to primarily occur for married mothers, as shown in Table 2.6, which reports the aggregated post-birth effects of a child by childcare cost region for the full sample (column 1), unmarried mothers (column 2), and married mothers (column 3). While the child penalty is unaffected by childcare costs for unmarried mothers, married mothers' child penalty is approximately $4000 larger in a high childcare cost state.

Figure 2.3: Child Penalty for Women Living in High or Low Childcare Cost States

Panel A Panel B



Notes: Figure 2.3A (left) plots coefficients from event studies of earnings on indicators for years surrounding a woman's first birth for both women who live in the low or high childcare cost states. The unit are percent changes (0 to 1) in earnings relative to the year prior to birth. The regression includes controls for age of mother at first birth and year of birth. Figure 2.3B (right) calculates the gap in the percent decline for those in high cost relative to those in low cost states and reports 10% confidence intervals for a test of the null that this gap is equal to zero.

Table 2.5: Aggregated Child Penalty, by Distance to Grandmother

	(1) Full Sample	(2) Unmarried	(3) Married
Pre-period	-430.3	-7218.8**	493.8
	(1096.6)	(2714.4)	(1174.1)
Year of birth	-5113.4***	-5644.1	-5094.2***
	(1262.3)	(4771.6)	(1279.9)
Post-period	-12068.9***	-12836.3***	-12454.1***
	(1042.8)	(3124.9)	(1135.1)
Near Grandma × Pre-period	19.38	6885.0*	-851.5
	(1331.3)	(2849.9)	(1459.2)
Near Grandma × Year of birth	-169.1	2541.9	-662.7
	(1462.7)	(4703.6)	(1572.4)
Near Grandma × Post-period	2721.8*	8418.1**	1927.6
	(1203.7)	(3174.8)	(1381.3)
Women-Year Obs.	13530	2201	11329

Note. This table reports the coefficients of a regression of earnings on indicators for years surrounding a woman's first birth, collapsed into the pre-period (2 to 5 years pre-birth), year of the birth, and post-period (1 to 10 years post-birth). The year prior to birth is omitted. All indicators are interacted with an indicator for is the woman is living in the same state as her own mother (Near Grandma). Controls for year of survey and age of mother are also included. Column 1 is the full sample; column 2 are unmarried at year of birth; column 3 are married at year of birth. Standard errors clustered at the individual level in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.6: Aggregated Child Penalty, by Childcare Costs

	(1) Full Sample	(2) Unmarried	(3) Married
Pre-period	-293.2	-1731.5	427.3
	(676.3)	(999.8)	(774.7)
Year of Birth	-4208.6***	-1194.3	-4762.2***
	(915.3)	(1984.0)	(1194.6)
Post-period	-8435.2***	-6255.9	-9075.2***
	(733.6)	(3361.0)	(1158.3)
High CCC × Pre-period	-858.9	1517.6	-1564.4
	(1226.2)	(2362.3)	(1324.1)
High CCC × Year of Birth	-1924.4	-1593.0	-1465.7
	(1403.0)	(2990.3)	(1659.4)
High CCC × Post-period	-3738.6*	235.0	-3955.1*
	(1514.4)	(4570.2)	(1858.1)
N	9568	1674	7894
Women-Year Obs.	9568	1674	7894

Note. This table reports the coefficients of a regression of earnings on indicators for years surrounding a woman's first birth, collapsed into the pre-period (2 to 5 years pre-birth), year of the birth, and post-period (1 to 10 years post-birth). The year prior to birth is omitted. All indicators are interacted with an indicator for is the woman is living in the top half of the state childcare cost distribution (High CCC). Controls for year of survey and age of mother are also included. Column 1 is the full sample; column 2 are unmarried at year of birth; column 3 are married at year of birth. Standard errors clustered at the individual level in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.3 Model

Taken together, these analyses demonstrate the importance of geographic proximity to affordable childcare for women's long-run labor market outcomes– whether it be informal care from a grandparent or less expensive private childcare. However, in both the analyses, we are not fully accounting for the joint selection process of location, fertility, and labor force participation. For example, when we observe that mothers living in high childcare cost regions earn less than those low childcare cost regions, it may be that the mothers in low-cost regions were motivated to select into those regions due to higher ability or attachment to the labor force that is known to them but unobserved to us as econometricians. Therefore, a model that places some assumptions on the selection process will be required to account for the endogeneity of migration decisions and to evaluate the impact of policy counterfactuals.

In particular, we are interested in how policies that may substitute for intergenerational time transfers (such as subsidized childcare) would influence the migration decisions and subsequent earnings of women who might otherwise rely on their parents to assist in child-rearing. Using, we will be able to explore the effectiveness of such policies in improving welfare for different types of parents (e.g., single vs. married), as well as decompose any effects on earnings into a direct effect of changes in attachment to the labor force due to childcare policies versus the secondary effects of the policies such as allowing households to sort into better paying labor markets.

Lastly, we plan to estimate the model separately by race to explore heterogeneity in the value of these policies for Black mothers relative to White mothers. The frictions associated with childcare access may be particularly important in explaining racial gaps in migration rates and wages, as single motherhood is more common for Black mothers. Our reduced form analysis suggests that single mothers are more dependent on geographic proximity of family for access to care. The model will allow us to precisely quantify the extent to which fertility events drive migration across demographic groups in the United States and speak to the extent to which recent changes in family structure in the U.S. may be related to ongoing changes in labor mobility.

2.3.1 Setup and Timing of Decisions

Our model adapts the dynamic migration labor force participation of Eckstein and Wolpin (1989) and nests it in a simple framework of dynamic migration (Kennan and Walker, 2011) while incorporating multiple dimensions of family structure. The model is a dynamic discrete choice model that follows the labor force participation and migration decisions of women. We focus on women due to their stronger geographical attachment to their children compared to men and due to the wealth of evidence that points to fertility events being more influential on female labor force attachment than male.

A period is one year. Agents enter the model at age 22 and are at risk of pregnancy until age 35. Between ages 35 to 40, though agents cannot get pregnant, they may either have young children or have no children. After age 35, we additionally assume that the agent's current marital status remains fixed for the rest of the lifecycle. Agents choose whether to supply labor and, afterward, whether and where to move until making a final labor force decision at age 65, after which they accrue no further utility[8]. We select age 22 as the starting point to allow the bulk of higher education choices to be made while pre-empting the prime fertility years of U.S. women[9].

At the beginning of each period, the women in our model observe the location of their parents and stochastic realizations of marriage and fertility[10]. The women then choose whether

[8]So, the final migration decision is made at age 64.

[9]In our estimation sample, 75% of individuals had the same educational attainment at age 35 as they did at age 22. Moreover, while this choice prevents us from being able to account for teenage pregnancies, we observe that in 2006, 90% of first-time mothers were aged at least 23 (Authors' calculations, 2006 ACS).

[10]Fully endogenizing marriage and fertility in the model would be impossibly complicated. We make this assumption to allow us to focus on the influence of childcare availability on female labor force participation

to participate in the labor force[11], weighing increased utility from consumption[12] should they choose to vs. preferences for leisure and savings on childcare expenditures should they not. Participation also increases future expected earnings through accumulating work experience. The women then choose whether or where to move — in particular, their migration options include staying in their current location, moving to the state of their parents, or moving anywhere else, which we subsume into the nine Census divisions[13]. Following their migration decision, women enter the subsequent period.

2.3.2 State Variables and Value Functions

Table 2.7 presents a complete summary of state variables and notation in the model, which are described in more detail in this section. Locations are indexed by ℓ, with and ℓ^P denoting an agent's parent location. The other locations represent the nine Census divisions, each of which are vary by childcare costs δ^ℓ, wage effects η^ℓ, and cost of living κ^ℓ. One's college attainment is indexed by $e \in \{0, 1\}$, years of experience by x, and age by a. College attainment here is assumed to be static, but experience will be allowed to grow endogenously over time.

We now turn to describing notation for family structure. Marital status is denoted by

and mobility. We assess the robustness of our policy counterfactuals to allowing a simple fertility elasticity effect later in the paper.

[11]We focus on participation vs. non-participation both for simplicity and because we observe in the ACS that the share of women working part-time appears invariant to the presence of small children in the household, suggesting that outright participation is the relevant behavioral margin to consider.

[12]We scale consumption so that one unit corresponds to \$2,080, following from full-time work involving working 40 hours per week, 52 weeks per year. This normalization can also be thought of as normalizing the agent to have a single hour of time.

[13]See Appendix 2.C for division definitions. Agents are allowed to live in the same division of their parents while not being in the parent location, as well, so that they do not receive parental childcare time transfers. Gemici (2011) uses the same geographic structure. 72% of cross-state moves observed in the data involve cross-divisional moves as well, so our framework allows us to capture the bulk of labor mobility activity. We have estimated a version of the model that instead uses the 48 mainland U.S. states as geography — in addition to dramatically increased computational load, the behavior of the likelihood is somewhat more erratic in trying to rationalize very low-probability moves. However, the main counterfactual results are virtually unchanged.

Table 2.7: Model Notation

Description		Values
Locations	ℓ	$\ell^P, \{1, ..., 9\}$
Location Daycare Cost Types	δ	$\delta^\ell, \ell \in \{1, ..., 9\}$
Location Wage Effects	η	$\eta^\ell, \ell \in \{1, ..., 9\}$
Location Costs of Living	κ	$\kappa^\ell, \ell \in \{1, ..., 9\}$
College Attainment	e	0,1
Age	a	[22,65]
Age of Youngest Child	a_c	$\emptyset$, [0,4]
Years of Experience	x	[0,40]
Marital Status	m	0,1
Fertility Status	f	0,1
Spouse Wage FE	μ_S	μ_S^L, μ_S^H
Previous LFP Status	p	0,1
Hours	h	0,1
Time Transfers from Spouse/Parents	τ	τ^S, $\tau^{P,m}$

Notes: Table presents model notation. Description of variables and symbolic representations are contained in the first two columns of the table, while the potential values the variables can take are presented in the third column.

$m \in \{0, 1\}$ and is assumed to evolve entirely stochastically, depending on other elements of the state space. Men make no decisions in our framework and are assumed to inelastically provide monetary and childcare time transfers to their wives. The variable a_c captures the age of the youngest child in the household, provided that they are less than 5 years old.[14] The state $a_c = \emptyset$ stands for when the household has no children aged 5 or younger.

Meanwhile, the variable f captures the fertility status of the woman: $f = 1$ indicates pregnancy, i.e., if $f = 1$ in year t then $a_c = 0$ in year $t + 1$ with certainty. Having pregnancy be a known state allows our women to make migration and labor force participation decisions *in anticipation* of fertility events. Conception, meanwhile, is entirely stochastic and depends

[14]We currently do not keep track of the number of young children and instead focus on the presence of any at all. Rosenzweig and Wolpin (1980) study the effects of twins on labor force participation and find that women with twins exhibit a labor force participation rate 0.371 pp lower than women without. While these effects are significant, we view their magnitude as small enough to permit the omission for the time being. This almost certainly means that we are understating the costs of childcare and the potential effects of subsidies to them in terms of labor force participation and wages.

on the other elements of the state space.[15] Women are allowed to have multiple children in that their f state may equal 1 even if the household currently contains a young child, in which case a_c will be reset to zero in the subsequent period. We shut down fertility events at age 35, meaning that when women leave the model at age 40 all children have aged out of early childhood. For more details on how the stochastic processes that govern marriage and fertility realizations are determined, refer to Section 2.4.

Women are assumed to be endowed with a single unit of time and may choose to work full time ($h = 1$) or not at all ($h = 0$). Spouses are also assumed to be endowed with a fixed effect μ_S that affects their earnings potential. Subsuming all the state variables outside of the agent's current location into the vector Ω, the value function for a woman without young children in the model is as follows:

$$
\begin{aligned}
V(\Omega, \ell) = \max_h \Bigg\{ & \alpha_1(c) + (1-h)(\alpha_2 + \alpha_e e + \alpha_x + \alpha_c c) + \alpha_3 \mathbb{1}(h \neq p) + \alpha_4 \mathbb{1}(\ell = \ell^P) \\
& + \boldsymbol{\alpha_\Gamma \Gamma} + \mathbb{E}_{\zeta_{\ell'}}[V'(\Omega, \ell; h)] \Bigg\}; \qquad (2.2)
\end{aligned}
$$

$$
\kappa^\ell c = w_S \mathbb{1}(m = 1) + wh.
$$

Thus, α_1 rescales utility over consumption in dollars to util terms, and α_2 represents a preference for leisure. Preferences for leisure are further modified based on experience (α_x) or if the agent has a college degree (α_e), and α_c represents a consumption-leisure complementarity that makes married women less likely to work. The parameter α_3 constitutes a penalty borne from changing one's labor force participation status (i.e., $p = 0$ and $h = 1$, or $p = 1$

[15]In 2008, 54% of births among unmarried women aged 20–29 were unintended, compared with 31% of births to married women in the age group (Zolna and Lindberg, 2012). Moreover, even when planned the timing is not always in the control of women; in PRAMS surveys of women who gave birth, about 18% reported that they would have preferred to have had the birth sooner (Maddow-Zimet and Kost, 2020). Additionally, a preponderance of women cite non-economic reasons as the drivers of the choice to conceive (Edin and Kefalas, 2011).

and $h = 0$), allowing the model to account for frictions individuals face in moving in and out of the labor force. Utility premia for currently being in one's parent's location is captured by α_4. Locations differ in amenities $\mathbf{\Gamma}$ that include average distance to shore (taken from Lee and Lin (2017)), average number of warm-weather days in a calendar year (taken from Kennan and Walker (2011)), and an index of other amenities related to government provisions and quality of life taken from Diamond (2016).

Consumption here is given by the wages of the woman's spouse (assumed to be supplied inelastically and equal to zero if the woman is unmarried) and the earnings of the woman herself. Log wages of the woman and her spouse are given by the following equations:

$$\log(w_S) = \beta_{S,0} + \boldsymbol{\beta}_S \boldsymbol{X}_S + \mu_S + \eta^{\ell};$$

$$\log(w) = \beta_0 + \boldsymbol{\beta X} + \eta^{\ell} + \varepsilon + \xi;$$

$$\varepsilon \sim N(0, \sigma_\varepsilon) \text{ i.i.d; } \ \xi \sim N(0, \sigma_\xi) \text{ i.i.d.}$$

The vector of observables of the spouse $\boldsymbol{X}_S$ contain a college dummy and a quadratic in experience, while the agent's observables $\boldsymbol{X}$ contain the same standard Mincerian combination along with dummies for having a child aged 0-1 or a child aged 2-4.[16] With the assumption that husbands supply labor inelastically, the terms of the husband's wage equations can be uncovered directly from data if we assume husbands to be identical to their wives in age and schooling level. Meanwhile, the components of the woman's wage process will be parameters to be estimated. Location fixed effects, η are also assumed to be constant across time and equal for men and women, which with the assumption of exogenous male labor supply will allow us to estimate values for η outside the model using male wages. Wages offers for women are additionally shocked by a transient component ε that will be the key factor in

[16]While we abstract away from an explicit part-time choice, this allows the model to be consistent with women potentially preferring more flexible and possibly lower-paying jobs when parenting a small child.

determining whether a woman works in a given state and are measured with error ξ assumed uncorrelated with ε.

The final term of (2.2), $\mathbb{E}_{\zeta_{\ell'}}[V'(\Omega, \ell; h)]$, represents the expected continuation value given the woman's labor force participation decision. Following her choice of h, the woman receives a series of location preference shocks that will determine whether and where she moves:

$$V'(\Omega, \ell; h) = \max_{\ell'} \left\{ \beta \sum_{\Omega'} \mathbb{E}_{\varepsilon}[V(\Omega', \ell')]\mathbf{Pr}(\Omega' | \Omega, h, \ell') - \Delta(\Omega, \ell')\mathbb{1}\{\ell' \neq \ell\} + \zeta_{\ell'} \right\}.$$

The agent takes into account possible state transitions Ω' and expected next-period utility after solving her optimal hours decision problem and optimizes their choice of next-period location following a series of location preference shocks $\zeta_{\ell'}$ distributed Type 1 Extreme Value with location 0 and the scale parameter normalized to 1. Fertility and marriage transitions are governed by stochastic functions that we calibrate directly from the data. We assume that the woman can no longer become pregnant at age 35 and that their marriage state at age 35 carries on for the remainder of the life cycle. The agent's next value of p (past-period labor force participation) depends on her selection of h. The agent's experience x increments by 1 should they choose to work and 0 if they do not, and the agent's age a increments by 1 with certainty. Next-period utility is discounted by the factor β.

The parameter $\Delta(\Omega, \ell')$ captures moving costs that the agent faces should they have chosen to do so, which itself depends on other elements of the state space. If a woman moves across locations in a period, she must incur moving costs given by

$$\Delta(\Omega, \ell') = \gamma_0 + \gamma_1 e + \gamma_2 \mathbb{1}\{a_c \neq \emptyset\} + \gamma_3 m + \gamma_4 N^{\ell'}.$$

Moving costs involve a fixed cost, are potentially smaller for college graduates, and are assumed to be larger for married agents and for agents that already have young children.

Furthermore, we allow for moves to larger locations (N^ℓ represents the population of division ℓ in tens of millions) to less costly as in Kennan and Walker (2011).

Finally, a woman with young children enjoys utility:

$$V(\Omega, \ell) = \max_h \Bigg\{ \alpha_5(c) + (1-h)(\alpha_6 + \alpha_e e + \alpha_x x + \alpha_c c) + \alpha_3 \mathbb{1}(h \neq p) + \alpha_7 \mathbb{1}(\ell = \ell^P) \\ + \boldsymbol{\alpha}_{\boldsymbol{\Gamma}} \boldsymbol{\Gamma} + \mathbb{E}_{\zeta_{\ell'}}[V'(\Omega, \ell; h)] \Bigg\}; \quad (2.3)$$

$$\kappa^\ell c = w_S \mathbb{1}(m=1) + wh - \delta^\ell \cdot \max\Big\{0, h - \tau^S \mathbb{1}(m=1) - \tau^{P,m} \mathbb{1}(\ell = \ell^P)\Big\}.$$

The specification thus flexibly allows women to have different preferences for consumption, leisure, and location based on the presence of young children in the household. When young children are present, the agent must also either dedicate time for caring for their children or absorb childcare costs, which depend on their current location, the current location's type, and the woman's marital status. The specification ensures that women never pay for childcare costs if they do not work ($h = 0$), and spouses and grandparents are assumed to contribute fixed time transfers to childcare (τ^S and $\tau^{P,m}$, respectively) if the woman is either married or living in her parent's location. The grandparents' contribution is allowed to vary based on the marital status of the woman. Furthermore, we allow for unobserved heterogeneity in grandparent helpfulness, such that with probability P_τ the agent's parents will provide time transfers of zero regardless of marital status[17].

2.3.3 Model Solution

The model is solved via backward induction. In each point of the state space, labor force participation is governed by whether the transient component of the wage offer ε is suffi-

[17]We have also tried including unobserved heterogeneity in grandparents while estimating the actual transfers provided by unhelpful grandparents and estimated transfers of zero directly.

ciently high. We compute cutoff values of ε for each element in the state space, after which continuation values can be computed by applying the usual type-1 extreme value formula and using the cutoff values in conjunction with properties of the normal distribution to solve for an agent's expected flow utility in the next period. A more detailed description of the procedure is as follows:

1. Solve for cutoff values of ε that govern labor force participation for the terminal age-65 period, where continuation values are zero by construction.

2. Using properties of the normal distribution, solve for *expected* utility $\mathbb{E}_\varepsilon[V_{65}(\Omega, \ell)]$ following the optimal hours decision in the age-65 state space.

3. Apply the type-1 extreme value formula to construct the agent's expected utility from choosing their optimal next-period location at age 64:

   $$\mathbb{E}_{\zeta_{\ell'}}[V'_{64}(\Omega, \ell; h)] = \bar{\gamma} + \log\left(\sum_{\ell'} \exp\left(\beta \sum_{\Omega'} \mathbb{E}_\varepsilon[V_{65}(\Omega', \ell')]\mathbf{Pr}(\Omega'|\Omega, h, \ell') - \Delta(\Omega, \ell')\mathbb{1}\{\ell' \neq \ell\}\right)\right)$$

   where $\bar{\gamma}$ is the Euler-Mascheroni constant. This gives continuation values for all possible combinations of state space and labor supply decisions for the age-64 period.

4. With continuation values in hand, compute cutoff values of ε for the age-64 period.

5. Repeat steps 2-4 through ages 63 to 22, at which point the model is solved.

Algebraic details for solving cutoff values of ε and expected utility from hours decisions can be found in in Appendix 2.B.

2.3.4 Discussion

We have presented a tractable model of dynamic labor force participation and migration aimed to capture the geographic constraints imposed by childcare costs and grandparent locations that U.S. women face. The presence of grandparents reduces childcare costs thus reservation wages, allowing women to maintain their participation in the labor force and to continue building work experience that will subsequently raise wages for the remainder of their life. However, these benefits only apply if women are located in the same place as their grandparents, which is a notable constraint given the extent to which migration plays a role in wage growth (Kennan and Walker, 2011). We do not impose that location decisions are one-shot as in García-Morán and Kuehn (2017), however: women may leave their parent's location and then move back when they know that a fertility event is imminent.

The assumption that women only have one child at a time means that our model almost certainly understates childcare costs and the potential effects of childcare subsidy policies on labor force participation, experience, and wages. Moreover, the discretization of geography into nine locations means that we may be suppressing the role that geography plays in wages, which may also have implications for our counterfactual policy predictions. Allowing for a richer geographic structure and potentially an urban/rural distinction while retaining computational tractability in the model may be desirable. Extending the model to account for additional unobserved heterogeneity in wages such that the agents as well as their spouses differ in fixed effects and, possibly, location-specific match effects may also be worthwhile (a la Kennan and Walker (2011)).

2.4 Estimation

2.4.1 Data

We use data on women aged 22-35 in the 2000-onward PSID. All women must be observed at age 22 to be included in our sample. The PSID shifted to a bi-annual schedule starting in 1997 — however, in years following 2000, respondents were asked of their income and hours worked for both the preceding year and the year before. Furthermore, if the respondent had moved across states since their most recent interview, they were asked in which year the move was made. This information, combined with detailed marital and childbirth histories for all respondents, allows us to construct yearly data from the biennial survey with minimal assumptions. If a respondent is observed to be living in a different state since their last interview but does not report the year in which they moved, we assume they moved in the same year as their previous interview.[18]

Importantly, the PSID additionally allows for intergenerational linkages, through which we can track the location of the parents of the respondent. When coding the location of one's parents, we use the state of both parents if both parents are living in the same state (which is the case the overwhelming majority of the time). If the parents are living in different states or if the father's location is missing, we use the location of the mother, and we use the location of the father if the mother's location is missing. Parents are assumed to be living in the agent's home location if the location of both the mother and the father are missing.

The PSID also provides information on the year of birth of the first child of all respondents, as well as the birth years of their four youngest children. We use these years to code fertility events for our sample. If a child is born to a woman in year t, it is assumed that the woman was aware of the impending birth in year $t-1$ — in other words, $f = 1$ in year $t-1$. We limit our sample to women who are coded as either household heads or the spouses of household heads — thus, information about marital transitions and spousal earnings can be easily obtained from household head information for women labeled as spouses.

We categorize the educational attainment of our sample based on their college status at age 35[19]. Earnings in the data are deflated to real 2012 dollars using the PCE deflator. To

[18]This happens quite infrequently.

[19]For 75% of our sample, college attainment at age 22 was the same as college attainment at age 35. To account for delayed graduation, we exclude college graduates age younger than 25 when evaluating the likelihood function.

Table 2.8: Observations by Age

Age	# Observations
22	909
23	909
24	909
25	909
26	813
27	741
28	631
29	554
30	479
31	434
32	362
33	297
34	233
35	174

Notes: Table presents number of individuals observed at each age in PSID analysis sample. See text for details on sample construction.

constrain the measurement error for wages in the data to reasonable levels, we winsorize hourly wages at the bottom at $7.25 per hour and at the top at the 95th percentile. Observations that report positive hours and zero income are dropped. Observations that reported working 30 hours per week or more are coded as full-time workers, while individuals coded as working less than 30 hours per week are coded as non-participants[20]. Individuals that report working more than 5,820 hours in a year are dropped. Finally, we limit our sample to individuals who are observed continuously in the data for at least 4 years.

These restrictions leave us with a sample of 909 women and 8,354 person-year observations. The median woman in our sample is observed for seven years (i.e. up through age 28), but we have 479 women observed through age 30 and 174 observed through age 35 (see Table 2.8 for a complete tabulation of ages in our analysis sample). Table 2.9a presents

[20]Given substantial bunching at 0 or 40 hours per week, alternate thresholds for determining labor force participation have little substantive effect on our results.

descriptive demographic and economic statistics broken down by age ranges and college attainment, while Table 2.9b presents migration statistics in our estimation sample with additional breakdowns by college attainment[21]. Women who have earned a college degree by age 35 have children and marry later, work more, and earn more than their non-college-educated counterparts. College-educated women are unsurprisingly also more migratory, with close to twice the share of college-educated women moving at least once in our data compared to women without a college degree. However, women without a college degree appear more inclined to move back to their parent's location than women with a college degree. The presence of spouses and young children appear to migration rates.

2.4.2 Parameters Estimated Outside the Model

We assume a discount rate of $\beta = 0.95$. Childcare costs levels δ for an hour of care are at the division level following our data from Childcare Aware by averaging across states within a division with population weights. Costs of living κ^{ℓ} are taken from the American Chamber of Commerce Research Association's Cost of Living Index.[22] The parameters governing spousal wages are taken a comparable PSID sample to our analysis sample. With the assumption that husbands supply labor exogenously and are of the same age and education level as their wives, these parameters can be estimated directly from Mincerian wage regressions. Since we assume location wage effects to be equal between men and women, this also allows for the recovery of location wage effects η^{ℓ}, which are again grouped at the division level[23].

[21]We do not use sample weights when creating these statistics or when estimating our model. Including longitudinal sample weights available in the PSID does little to change our parameter estimates.

[22]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. 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 sets of values is close to 0.8), so we simply take the midpoint of the two while normalizing the cost-of-living level of Iowa to be zero before averaging by division with population weights.

[23]See Figure 2.A.1 for representations of state-level childcare costs, wage effects, and living costs.

Table 2.9: Summary Statistics of PSID Estimation Sample

Sample	All		College		Non-College	
Age	22/23	34/35	22/23	34/35	22/23	34/35
LFP Rate	50.00	46.93	51.61	55.13	49.28	41.83
	(50.01)	(49.97)	(50.02)	(49.90)	(50.01)	(49.43)
Years of Experience	3.01	8.92	0.23	7.44	4.25	9.84
	(1.91)	(4.15)	(0.42)	(3.88)	(0.43)	(4.05)
Hourly Wage	13.26	17.90	16.07	20.87	11.95	15.47
	(5.56)	(7.35)	(5.95)	(7.57)	(4.85)	(6.21)
Share Married	43.40	57.00	45.00	73.08	42.69	47.01
	(49.58)	(49.57)	(49.79)	(44.50)	(49.48)	(50.01)
Young Child Present	50.99	30.47	26.79	38.46	61.76	25.50
	(50.00)	(46.08)	(44.32)	(48.81)	(48.62)	(43.67)
Observations	1818	407	560	156	1258	251

(a) Demographic and Economic Statistics

Sample	All	College	Non-College
Annual Migration Rate	4.35	6.63	3.2
	(20.39)	(24.88)	(17.61)
With Children	3.38	4.83	2.83
	(18.08)	(21.45)	(16.58)
If Married	3.66	5.73	2.35
	(18.77)	(23.25)	(15.17)
Ever Migrated	25.33	40.32	17.81
	(43.49)	(49.06)	(38.26)
Share of Moves to ℓ^P	29.01	26.06	32.08
	(45.45)	(44.03)	(46.82)
N	8,354	2,765	5,589

(b) Migration Statistics

Notes: Standard deviations in parentheses. Data from 2001-2017 biennial waves of PSID. Table 2.9a presents demographic statistics for analysis sample, broken down by educational attainment and age at observation. Table 2.9b presents migration statistics for the estimation sample, broken down by educational attainment. See text for details on sample restrictions.

Table 2.10: Parameters Estimated Outside the Model

Parameter		Value
Discount rate	β	0.95
Childcare cost levels	δ^{ℓ}	Various
Location wage effects	η^{ℓ}	Various
Location living costs	κ^{ℓ}	Various
Location populations	N^{ℓ}	Various
Spouse wage, constant	$\beta_{S,0}$	2.234
Spouse wage, education	$\beta_{S,1}$	0.571
Spouse wage, experience (linear)	$\beta_{S,2}$	0.047
Spouse wage, experience (quadratic)	$\beta_{S,3}$	-0.0007
Spouse wage, fixed effects	μ_S^L, μ_S^H	-0.39, 0.39

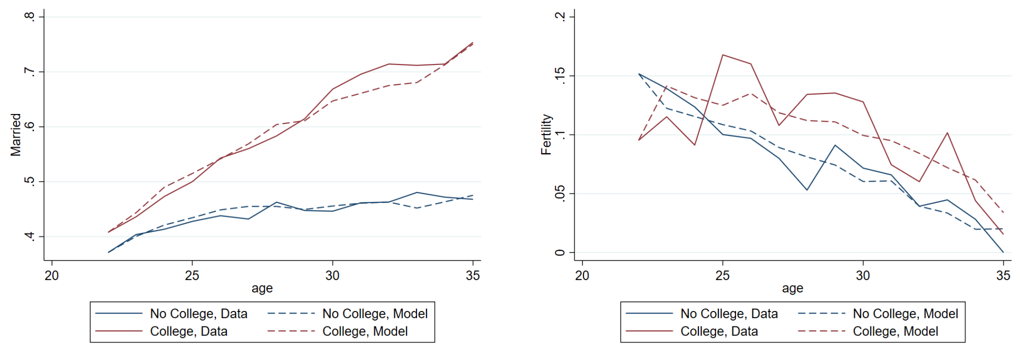
Notes: Table reports values of parameters that are estimated outside the model. Columns 1 and 2 describe the parameters and presents their symbolic representation. Column 3 reports parameter values. See text for details on model and sample construction. See Figure 2.A.1 for representations of state-level childcare costs, wage effects, and living costs.

Marriage, divorce, and conception probabilities are estimated via linear probability models that admit as inputs whether the agent is currently married, pregnant, or a parent to young children, as well as a cubic polynomial in age. Probabilities of marital dissolution and formation are also allowed to vary over spousal wage type μ_S. All probabilities are also calculated separately for women with and without a college degree. These linear probability models are estimated directly using our estimation sample. Figure 2.4 presents the fit of our model with regards to life-cycle profiles of marriage and fertility rates for women with and without a college degree and indicates that our model fits salient features of the data well.

2.4.3 The Likelihood Function

We use maximum likelihood to estimate the remaining parameters of our model. If a woman is never observed to work in the data, we assume that she is a never-working type that always chooses not to work with probability 1. The joint likelihood function for labor force participation, wages, and migration for the N women in our sample, each observed for T_i

Figure 2.4: Model Fit — Marriage and Fertility Life-Cycle Profiles



(a) Marriage (b) Fertility

Notes: Figure presents model fit of marriage and fertility rates over lifecycle for women in PSID analysis sample and in data simulated from model. Probabilities estimated separately for women with and without a college degree and depend on marital status, pregnancy, presence of young children, a cubic in age, and spouse wage type. See text for details on sample construction.

periods, is given by:

$$L = \prod_{i}^{N} \sum_{\tau} \mathbf{Pr}(\tau) \prod_{t=1}^{T_i} \mathbf{Pr}(h = h_{it} | \Omega_{it}, \ell_{it}) \cdot \mathbf{Pr}(w = w_{it} | \Omega_{it}, \ell_{it}, h_{it}) \cdot \mathbf{Pr}(l' = l'_{it} | \Omega_{it}, \ell_{it}, h_{it}).$$

We employ a mixture model over unobserved heterogeneity in grandparent transfers, letting $\mathbf{Pr}(\tau)$ denote the probability of the agent being unobserved type τ. The probability of observing wages and hour decisions joint with a location are separable using the assumption that the next-period location shocks are independently distributed from labor supply shocks in each period. For any given element in the state space (Ω_{it}, ℓ_{it}), a reservation value of the transient component of the wage offer $\varepsilon^{**}(\Omega_{it}, \ell_{it})$ can be found that governs whether the woman supplies labor in the period[24]. Recall further that wages are measured with error:

$$\log(w) = \beta_0 + \boldsymbol{\beta X} + \eta^{\ell} + \varepsilon + \xi;$$

[24]For details on deriving these reservation values, refer to Appendix 2.B.

with $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ and $\xi \sim N(0, \sigma_\xi^2)$ distributed both i.i.d. and independently from one another. With this assumption, following Eckstein and Wolpin (1989) the first two components of the likelihood function corresponding to labor supply decisions and wages can be defined as

$$\begin{aligned} L &= \prod_i^N \sum_\tau \mathbf{Pr}(\tau) \prod_{t=1}^{T_i} \left[\Phi\left(\frac{\varepsilon^{**}(\Omega_{it}, \ell_{it})}{\sigma_\varepsilon} \right) \right]^{1-h_{it}} \cdot \\ & \left[\left(1 - \Phi\left(\frac{\varepsilon^{**}(\Omega_{it}, \ell_{it}) - \rho \frac{\sigma_\varepsilon}{\sigma_\nu} \nu_{it}}{\sigma_\varepsilon \sqrt{1-\rho^2}} \right) \right) \frac{1}{\sigma_\nu} \phi\left(\frac{\nu_{it}}{\sigma_\nu} \right) \right]^{h_{it}} \cdot \\ & \mathbf{Pr}(l' = l'_{it} | \Omega_{it}, \ell_{it}, h_{it}), \end{aligned}$$

where ϕ and Φ are the standard normal density and cumulative, respectively, $\nu_{it} = \varepsilon_{it} + \xi_{it}$, $\rho = \sigma_\varepsilon / \sigma_\nu$, and $\sigma_\nu = \sqrt{\sigma_\varepsilon^2 + \sigma_\xi^2}$, leading to $1 - \rho^2$ having the interpretation of the fraction of the wage variance attributable to measurement error. The third component of the likelihood function $\mathbf{Pr}(l' = l'_{it} | \Omega_{it}, \ell_{it}, h_{it})$ can be derived easily following the assumption that the location shocks $\zeta_{\ell'}$ are distributed type-1 extreme value. Denote $V(\Omega, \ell, h, \ell')$ as the expected utility gained from selecting location ℓ' following labor supply decision h after starting in state (Ω, ℓ), so:

$$V(\Omega, \ell, h, \ell') = \beta \sum_{\Omega'} \mathbb{E}_\varepsilon[V(\Omega', \ell')] \mathbf{Pr}(\Omega' | \Omega, h, \ell') - \Delta \mathbb{1}\{\ell' \neq \ell\}.$$

Recall that $\mathbb{E}_\varepsilon[V(\Omega', \ell')]$ represents the expected value of $V(\Omega', \ell')$ after optimizing over the labor supply decision given ε. The method for deriving closed-form expressions of these values is presented in Appendix 2.B, but their recursive nature renders it infeasible to write them out fully. With this, we can now derive the following final representation of the likelihood:

$$L = \prod_i^N \sum_\tau \mathbf{Pr}(\tau) \prod_{t=1}^{T_i} \left[\Phi\left(\frac{\varepsilon^{**}(\Omega_{it}, \ell_{it})}{\sigma_\varepsilon} \right) \right]^{1-h_{it}} \cdot$$

$$\left[\left(1 - \Phi\left(\frac{\varepsilon^{**}(\Omega_{it}, \ell_{it}) - \rho \frac{\sigma_\varepsilon}{\sigma_\nu} \nu_{it}}{\sigma_\varepsilon \sqrt{1-\rho^2}} \right) \right) \frac{1}{\sigma_\nu} \phi\left(\frac{\nu_{it}}{\sigma_\nu} \right) \right]^{h_{it}} \cdot$$

$$\frac{\exp(V(\Omega_{it}, \ell_{it}, h_{it}, \ell'_{it}))}{\sum_{l'} \exp(V(\Omega_{it}, \ell_{it}, h_{it}, \ell')}.$$

2.4.4 Model Assumptions and Identification

The relationship between labor force participation and migration decisions in our model are identified from jointly observing participation, earnings, and location choices for women, conditional on demographic characteristics and location of grandparents.

First, we assume that the shocks drawn in the model – location preferences, earnings shocks, fertility realization, marriage realization – are all independently and identically distributed across individuals and time. While this may seem a strong assumption at first, we do allow the likelihood of pregnancy and marriage to vary on observable characteristics, including many of the factors that contribute to a woman having a higher or lower earnings potential. This means that this assumption relies only on the weaker assumption that pregnancies and marriages are not correlated across time with the component of earnings that varies idiosyncratically across time. Due to the high rates of unintended and mistimed births in the US (Zolna and Lindberg, 2012), we believe this is a more reasonable assumption. The assumption of independence of location preference shocks and earnings shocks is a stronger assumption: we assume that wage differences across time/individuals are not place-specific and not correlated with amenities in a location in a given year. While the inclusion of amenities along with a reasonably rich wage process in the model helps justify this assumption, allowing for additional heterogeneity in idiosyncratic wage match effects may be helpful as well.

With these distributional assumptions in place, identification of the structural parameters falls out cleanly from the maximum likelihood estimation equation. Following Eckstein and Wolpin (1989) and using equations for reservation wages in Appendix 2.B, the reservation wages ε^*, the wage parameters ($\beta_0, \boldsymbol{\beta}$, and μ), and σ_ε and σ_ξ are all identified from data on participation and wages. We can then use the identified ε^* and our equation for the definition of the reservation wage described in Appendix 2.B to identify $\frac{\alpha_2}{\alpha_1}, \frac{\alpha_e}{\alpha_1}, \frac{\alpha_\mu}{\alpha_1}, \frac{\alpha_c}{\alpha_1}$, and $\frac{\alpha_3}{\alpha_1}$. Based on the similar equation for women with young children, we can identify $\frac{\alpha_6}{\alpha_5}, \frac{\alpha_e}{\alpha_5}, \frac{\alpha_\mu}{\alpha_5}, \frac{\alpha_c}{\alpha_5}, \frac{\alpha_3}{\alpha_5}, \tau_{p1}, \tau_{p0}$, and τ_s. Using any combination of pairs in which the leisure parameter is the same across the presence of children (e.g., $\frac{\alpha_e}{\alpha_1}, \frac{\alpha_\mu}{\alpha_1}, \frac{\alpha_e}{\alpha_5}, \frac{\alpha_\mu}{\alpha_5}$) would allow us to separately identify α_1 and α_5 and thus separately identify all α parameters governing leisure.

The remaining parameters include the parameters governing preferences for the parent's location, amenities, and the moving cost parameters. These are identified off the observation of location choices conditional on demographic type and participation in that period. Specifically, we can identify the parent's location preference parameter off the difference in the likelihood of moving to the parent's location ℓ^p from some location k and the likelihood of moving to a non-parent's location from that same location k for agents who are similar on all demographic characteristics. The same logic applies for identifying the amenity utility parameters $\boldsymbol{\alpha_\Gamma}$. The moving cost parameters are identified off the differences in likelihood of moving from location j to k versus staying in location j by demographic group. The parameter on population is identified off of the relative likelihood of moving from a small division to a large division vs. from a large division to a small division.

2.5 Results

Parameter estimates are presented in Table 2.11. When evaluating the likelihood, we exclude women age less than 25 who have a college degree to account for a non-trivial number of women with a college degree who finished their degree in their mid-20s. Standard errors are computed via inverting the numerical Hessian of the likelihood function and taking its diagonal. The estimation recovers preferences for consumption and leisure that increase and decrease respectively over the presence of a small child. The disutility associated from changing one's labor force participation status is substantial, and we also estimate considerable leisure preferences for women with high earnings potential, which is necessary to rationalize their rates of labor force participation that, while higher than low-earning women, are still considerably lower than those of men. The leisure-consumption complementary α_c is positive, reflecting women being less likely to work with higher-earning spouses, all else hold equal. The estimates of wage returns to a college degree and experience are all in line with previous estimates in the literature. We also estimate a meaningful reduction in wages associated with having a child between 0 and 1 year old, but the effect of older children on wages is statistically insignificant.

The estimates of time transfers τ suggest that spouses and grandparents considerably offset the direct cost of childcare for women with children — indeed, helpful grandparents cover virtually all childcare needs for unmarried women. However, we find substantial heterogeneity in grandparent time transfers based on whether a woman is married, with the married grandparent transfer $\tau^{P,1}$ being quite close 0.4. We note that this is consistent with the heterogeneity in child penalties we estimated in Section 2.2.2. Moreover, we find that Blacks are more likely to have helpful grandparents than Whites.

Table 2.11: Parameters Estimated via Maximum Likelihood

Parameter		$\hat{\theta}$	$\hat{\sigma}_\theta$	$\hat{\theta}$	$\hat{\sigma}_\theta$	$\hat{\theta}$	$\hat{\sigma}_\theta$
Utility							
Consumption, no children	α_1	0.103	0.011	0.097	0.014	0.112	0.029
Leisure, no children	α_2	1.173	0.127	1.105	0.174	1.243	0.257
LFP switch penalty	α_3	-0.127	0.015	-0.138	0.020	-0.104	0.028
Parent preference, no children	α_4	-0.402	0.024	-0.408	0.033	-0.409	0.031
Consumption, with children	α_5	0.089	0.010	0.087	0.012	0.084	0.015
Leisure, with children	α_6	0.765	0.086	0.742	0.127	0.724	0.050
Parent preference, with children	α_7	-0.405	0.104	-0.046	0.204	-0.634	0.120
Consumption/leisure complementarity	α_c	0.002	0.001	0.004	0.001	0.000	0.001
College leisure preference modifier	α_e	0.503	0.060	0.493	0.094	0.501	0.018
Experience leisure preference modifier	α_x	-0.002	0.003	-0.003	0.004	-0.003	0.004
Amenity preference: distance to shore	$\alpha_{\Gamma,1}$	-0.007	0.010	-0.009	0.017	-0.016	0.015
Amenity preference: amenity index	$\alpha_{\Gamma,2}$	0.045	0.057	0.059	0.097	0.022	0.063
Amenity preference: warm days	$\alpha_{\Gamma,3}$	0.149	0.057	0.076	0.112	0.182	0.057
Time Transfers							
Spouse time transfer	τ^S	0.229	0.046	0.209	0.066	0.180	0.101
Parent time transfer, unmarried	$\tau^{P,0}$	0.997	0.085	0.999	0.151	0.938	0.160
Parent time transfer, married	$\tau^{P,1}$	0.388	0.077	0.394	0.105	0.460	0.204
Probability of $\tau^P = 0$	P_τ	0.687	0.032	0.708	0.054	0.658	0.056
Wages							
Wage intercept	β_0	1.972	0.020	2.000	0.031	1.928	0.007
College effect	β_1	0.458	0.016	0.440	0.022	0.488	0.024
Experience effect, linear	β_2	0.058	0.002	0.062	0.003	0.059	0.003
Experience effect, quadratic	β_3	-0.002	0.000	-0.002	0.000	-0.002	0.000
Child aged 0-1	β_4	-0.085	0.016	-0.082	0.021	-0.069	0.028
Child aged 2-4	β_5	-0.028	0.015	-0.026	0.021	-0.028	0.024
Wage shock SD	σ_ε	0.265	0.010	0.274	0.013	0.251	0.017
Wage measurement error	σ_ξ	0.356	0.006	0.355	0.008	0.351	0.010
Moving Costs							
Fixed cost	γ_0	3.840	0.208	3.804	0.265	3.906	0.326
College effect	γ_1	0.153	0.115	0.096	0.142	0.340	0.191
Child effect	γ_2	0.277	0.138	0.261	0.173	0.422	0.250
Marriage effect	γ_3	0.480	0.134	0.496	0.165	0.250	0.295
Population effect	γ_4	0.007	0.054	-0.022	0.068	0.042	0.092
Sample		All		Whites		Blacks	
N		8,354		4,837		2,964	
Individuals		909		519		324	
Log Likelihood		-7,429		-4,227		-2,750	

Notes: Table presents estimates and standard errors of parameters estimated via maximum likelihood. Data from PSID. See text for details on sample construction and formation of likelihood function.

Of note is that preferences for home location α_4, α_7 are estimated to be negative. Since most individuals in our estimation sample start in their parent's location, parameters such as the moving fixed cost γ_1 are most directly identified from rates of migration out of parent locations. On the other hand, the parent location preference parameters are identified from the rates at which women with and without children *return* to their parent's location. Since parent time transfers are substantial, the model estimates negative values for these preferences to rationalize why we do not see a larger proportion of women moving back to their parents than we would if childcare time transfers were the only factor that influenced the utility from doing so. Amenity preference estimates indicate that the agents in our model prefer higher levels of the Diamond (2016) amenity index, shorter distances to shores, and warmer weather, but only the last of these factors is estimated to be statistically significant.

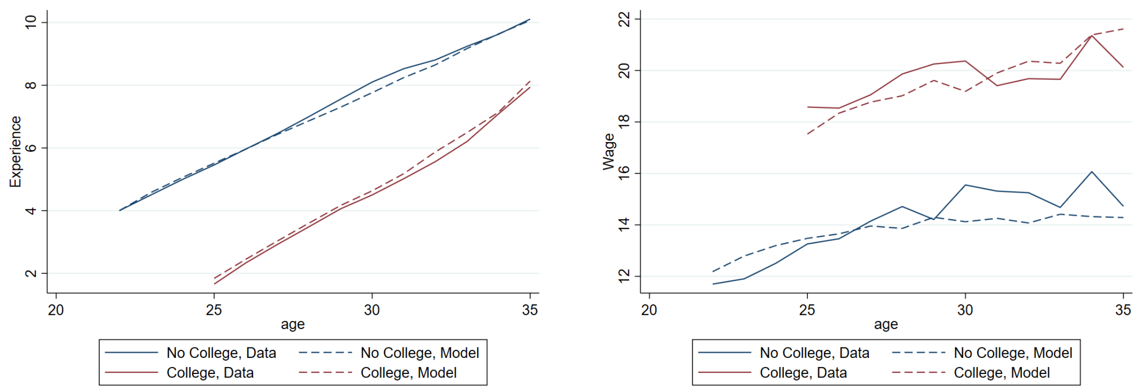
Because utility is linear in consumption, we can convert the moving parameters into dollars by dividing by the consumption scaling parameter and then multiplying by the consumption equivalence unit (i.e., one unit of consumption equal \$2000). For the "average" mover, the moving cost is about \$44,000 ignoring the value of the payoffs shocks.[25] For comparison, a woman's life earnings gain would be \$97,0000 if, holding all other behavior constant, she moved from the lowest paid region to the highest paid region in age 22 and then stayed in that region for the remainder of her life. Though this is an extreme example of the potential earnings gains from a move, it demonstrates that our moving costs net of payoff shocks are lower than, but of similar scale to the potential earnings gains. However, we will note that these moving costs are the estimated costs for a hypothetical move to an arbitrary location, whereas in the model people will only choose to move to high pay-off locations. Thus, these average costs are higher than the costs that households which actually choose to move will face once pay-off shocks are accounted for.[26]

[25]To calculate, we sum Δ for all individuals who move, discounted by the relevant consumption scaling. That is $\bar{\Delta} = 2000 \times \frac{1}{N_{move}} \sum_{i=1}^{N_{move}} \left[\left(\frac{\mathbb{1}(a_c \neq \emptyset)_i}{\alpha_5} + \frac{\mathbb{1}(a_c = \emptyset)_i}{\alpha_1} \right) \times (\gamma_0 + \gamma_1 e_i + \gamma_2 \mathbb{1}(a_c \neq \emptyset)_i + \gamma_3 m_i + \gamma_4 N_i^{\ell'}) \right]$.

[26]See (Kennan and Walker, 2011) for further discussion of the distinction between average moving costs

We estimate our model with our entire analysis sample as well as separately with Blacks and whites, allowing us to compare estimates between the two groups. We find some heterogeneity in grandparent helpfulness across races, along with children impacting migration patterns more for Blacks while marital status has a greater effect for whites. Broadly speaking, though, the extent of the racial heterogeneity we find is limited.

Figure 2.5: Model Fit — Labor Force Lifecycle Profiles



(a) Experience (b) Hourly Wage

Notes: Data from PSID. Figures compare life-cycle trends of experience and wages for women with and without a college degree in estimation sample and data simulated from model. Fit reported for all ages for women with a high school degree and ages 25-onward for women with a college degree. See text for details on sample construction.

2.5.1 Goodness of Fit

To assess our model's ability in approximating the true data generating process, we randomly simulate the outcomes of each woman in our estimation sample ten times, starting at age 22 and ending at the final age the given woman is observed in the data, using Bayes' rule to

versus average moving costs conditional on moving. (Kennan and Walker, 2011) show that while the moving costs for households that choose to move to their home location are large, moving costs to non-home locations are actually negative representing the fact that these moves are moves with a large expected future payoffs for the households who make these moves.

draw unobserved types. We then compare key moments in the estimation sample to those in our simulated data.

Figure 2.5 presents our model's fit in terms of lifecycle profiles of labor market outcomes separately for women with and without a college degree. In general, the model fits the data well, reproducing profiles of wages and experience accumulation that look very similar to the data. The model slightly understates earnings for women with a college degree early in the lifecycle, but by the end of the lifecycle the wage fit is reasonable for both education categories. and slightly overstates rates of migration for women of both levels of education at the beginning and end of the simulation period. The fit of experience suggests that the model does a good job in reproducing the total years worked among women in our sample and overall higher labor force participation rates of college-educated women.

Table 2.12: Model Fit — Labor Force Participation by Location, Marital Status, and Fertility

Panel A: Data

Marital Status	No Kids	Pregnant	Kids	No Kids, $\ell = \ell^P$	Pregnant, $\ell = \ell^P$	Kids, $\ell = \ell^P$	No Kids, $\ell \neq \ell^P$	Pregnant,$\ell \neq \ell^P$	Kids, $\ell \neq \ell^P$
All	0.617	0.645	0.395	0.602	0.636	0.405	0.672	0.674	0.350
$m = 0$	0.626	0.626	0.435	0.608	0.618	0.436	0.698	0.667	0.425
$m = 1$	0.568	0.558	0.386	0.556	0.543	0.398	0.624	0.609	0.332

Panel B: Model

Marital Status	No Kids	Pregnant	Kids	No Kids, $\ell = \ell^P$	Pregnant, $\ell = \ell^P$	Kids, $\ell = \ell^P$	No Kids, $\ell \neq \ell^P$	Pregnant,$\ell \neq \ell^P$	Kids, $\ell \neq \ell^P$
All	0.633	0.515	0.380	0.628	0.518	0.397	0.654	0.503	0.293
$m = 0$	0.661	0.546	0.381	0.658	0.550	0.402	0.676	0.529	0.247
$m = 1$	0.615	0.521	0.396	0.608	0.526	0.413	0.650	0.500	0.306

Notes: Data from PSID. Table compares labor force participation rates for women in estimation sample and data simulated from model. Pregnancy corresponds to woman being pregnant with their first child. See text for details on sample construction.

Table 2.13: Model Fit — Migration by Fertility

Panel A: Data

Direction	All	No Kids	Pregnant	Kids
ℓ^p Out-Migration Rate	2.06	2.69	2.69	1.52
ℓ^p In-Migration Rate	4.90	4.96	4.35	4.90

Panel B: Out of Sample (ACS)

Direction	All	No Kids	Pregnant	Kids
ℓ^p Out-Migration Rate	1.80	2.03	1.64	1.31
ℓ^p In-Migration Rate	4.04	4.15	4.52	3.45

Panel C: Model

Direction	All	No Kids	Pregnant	Kids
ℓ^p Out-Migration Rate	1.86	2.39	1.91	1.39
ℓ^p In-Migration Rate	4.47	4.70	5.19	4.14

Notes: Data from PSID and ACS. Table compares migration rates for women in estimation sample and data simulated from model. Pregnant corresponds to being pregnant with one's first child. See text for details on sample construction.

We evaluate the model's fit of labor force participation in more detail in Table 2.12 by breaking up labor force participation by fertility status (no kids, pregnant with first child, has young children), marital status, and location (in or out of parent's location). Qualitatively, the model can reproduce patterns of women with young children supplying labor markedly less than those without and women with spouses also being less inclined to work than married women. Across all women, the profile of labor force participation the model outputs over different locations and fertility statuses is reasonable. However, the model does slightly understate the gap in participation between married and single women and understates participation for pregnant women as a whole. Notably, the model also understates participation for single mothers who live outside their parents' location.

Next, we assess the model's fit of migration decisions by breaking down moves according to sending location, destination, and fertility status in Table 2.13. Since the PSID can have

Table 2.14: Model Fit — Childcare Time Transfers Received

Relative Care Received	Data	Model
Some	14.1	15.7
All	9.20	8.44

Notes: Data from PSID CDS. Table reports share of individuals who either use relatives for some or all of their childcare. Sample includes women with at least one child aged 0-4. See text for details on sample construction.

very small samples of movers (particularly pregnant movers), we supplement the table with statistics from the ACS sample used in Section 2.2 as well. Among all women and women without children, the model predicts sensible rates of migration both out and into the parent location. Moreover, the model is able to qualitatively match the pattern observed in the ACS of pregnant women moving back to their parents' location more frequently, while such behavior is not observed for women with young children in general.

Finally, we evaluate the frequency and quantity of grandparent childcare time transfers observed in our simulated data and compare it to comparable statistics computed from the PSID's Child Development Supplement (JOANNA: Maybe do a bit of describing of this sample and how you cleaned it here? Doesn't have to be that long.). In Table 2.14, we report the share of individuals who either get some childcare (i.e. time transfers that do not amount to full-time care) from relatives or all their childcare (i.e. full-time care) from relatives. While the parents in the CDS are not necessarily the same women in our data, the model predicts very sensible rates of childcare receipt overall and can reproduce the data pattern that a substantial proportion of those who receive any childcare from relatives in fact receive all their childcare from them.

2.6 Counterfactual Analysis

2.6.1 Migration and the Family

We begin by predicting lifecycle earnings and migration profiles under alternate demographic scenarios. In a first experiment, we impose that children are only born to married women, that is $P(f = 1|m = 0) = 0$, allowing our model to speak to how recent changes in out-of-wedlock births in the U.S. may have impacted female labor mobility. In a second counterfactual, we evaluate the role of grandparents in wage formation by removing them from the model entirely, setting parent location preferences parameters α_4, α_7 as well as grandparent time transfers $\tau^{P,m}$ to zero. Conceptually, the presence of grandparents on wages is ambiguous, since residing with them may increase labor supply and experience in the short run but may also impact wages negatively by discouraging moving to higher-paying locations. In all counterfactuals, we evaluate impacts in the age span of 22 to 55.

Table 2.15 presents the results of these exercises. For each counterfactual, we estimate impacts in terms of averages changes in lifetime wages, years of experience, number of lifetime moves, and share of time spent with parents. We also calculate a willingness to pay metric (WTP) by taking the change in utility resulting from the counterfactual scenario and dividing it by α_1, the utility scaling parameter for consumption for women without children. We conduct demographic heterogeneity analyses by assessing impacts for women by grandparent helpfulness type and who were or were not ever single mothers in the baseline simulation. We additionally conduct racial heterogeneity analyses for estimating the impacts for whites and Blacks separately while using the separate parameters estimated for them in Table 2.11. We also compute changes in wages, experience, and WTP in a scenario for moving costs are infinite ($\gamma_1 = \infty$) to assess the importance of accounting for migration when making counterfactual predictions.

Table 2.15: Effects of Alternate Demographic Scenarios

Panel A: Impacts of Children only Being Born to Married Women

Sample	Wages	Wages, $\gamma_1 = \infty$	Years x	Years x, $\gamma_1 = \infty$	WTP	WTP, $\gamma_1 = \infty$	# Moves	Share Time with Parents
All	0.11	-0.32	0.02	-0.01	23.20	15.34	0.32	-0.03
$\tau^P \neq 0$	-8.18	-8.16	-0.50	-0.51	16.99	7.28	0.37	-0.04
$\tau^P = 0$	3.66	3.03	0.24	0.21	25.92	18.74	0.30	-0.03
Never SM	-4.96	-6.42	-0.24	-0.30	4.37	0.00	0.17	-0.02
Ever SM	4.66	5.14	0.25	0.25	40.10	29.03	0.46	-0.05
Whites	-1.00	-2.86	-0.02	-0.09	12.68	4.54	0.22	-0.03
Blacks	1.53	0.79	0.01	-0.03	44.29	36.43	0.43	-0.04

Panel B: Impacts of Removal of Grandparents

Sample	Wages	Wages, $\gamma_1 = \infty$	Years x	Years x, $\gamma_1 = \infty$	WTP	WTP, $\gamma_1 = \infty$	# Moves	Share Time with Parents
All	-8.37	-9.88	-0.49	-0.56	-5.34	-7.18	0.03	-0.01
$\tau^P \neq 0$	-27.95	-32.99	-1.63	-1.87	-17.67	-23.88	0.11	-0.02
$\tau^P = 0$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Never SM	-5.64	-7.63	-0.31	-0.40	-2.82	-4.47	0.02	-0.01
Ever SM	-10.81	-11.89	-0.65	-0.71	-7.57	-9.51	0.04	-0.01
Whites	-6.68	-7.24	-0.37	-0.41	-4.95	-5.67	0.01	0.00
Blacks	-11.03	-12.00	-0.75	-0.82	-6.52	-8.75	0.04	-0.01

Notes: SM = Single Mother. Table presents average impacts of counterfactual scenarios on lifetime wages, years of experience, number of moves, and fraction of time spent in parent's location. Second columns for wage and experience changes and WTP conduct counterfactual experiment when moving costs are infinite. Monetary units scaled by $2,080. Results for whites and Blacks computed using separate parameter estimates shown in Table 2.11. See text for details on estimation sample and procedure.

The effects of children being born to only married women on wages and experience are generally quite limited. Surprisingly, the effects on work experience are slightly negative, which happens because women acquire experience in part to insure against the state of being a single mother — when the probability of this occurring vanishes, incentives to work decline slightly. While the effects on wages are generally small, we do observe that the counterfactual increases the number of moves made by individuals over the lifecycle by 0.32 on average, with stronger effects for Blacks and women who are ever single mothers, resulting from women who would have had moves encumbered by the presence of children facing smaller moving costs. On a base rate of approximately 1.75 moves over the life cycle in our simulated data, this constitutes an increase of roughly 25%, suggesting that recent increases in single parenthood may have been a contributor to concurrent declines in female labor mobility.

The removal of grandparents is associated with more substantial reductions in wages and experience for women who are ever single parents and Blacks. For example, we see that the existence of grandparents as a potential source of childcare is associated with approximately half a year of experience and subsequently earnings, which corresponds to a percentage increase of about 2.3%. To put these effects in context with the reduced form estimates earlier in the paper, the child penalty gap was about 10 p.p. smaller for mothers living near their grandparents. Removing grandmothers is thus able to account for about 25% of the child penalty we documented, which is reasonable given that not all mothers in our model are living near the grandparents even when they are an available option (and are thus presumably unaffected by the removal of grandmothers) and not all grandparents actually provide childcare assistance.

Additionally, these results demonstrate that parents' mobility is not only influenced by grandparents but also by regional costs for childcare: ignoring migration results in overstating the effects of the counterfactual, since when moving is allowed the affected individuals can migrate to areas with lower childcare costs or higher wages as a means of insurance. As such, the utility cost of the counterfactual is greater when moving is prohibited, and the difference in utility between the world where moving is allowed is larger for more-affected groups. Eliminating the pull of the parent location, however, does result in increased migration for the same groups of women who see the largest declines in earnings, suggesting that they are substituting from staying in their home location towards either higher paying or lower childcare cost locations when they can no longer take advantage of free relative care in their parent location.

2.6.2 Childcare Subsidies

As a final exercise, we conduct counterfactuals where we halve and then remove childcare costs entirely, while breaking down our counterfactual effects by demographics, race, and migration cost scenarios as before. The results of this exercise can be found in Table 2.16.

In all cases, these policies increase years of experience, labor mobility, and lifetime wages, with particularly strong effects for women who are ever single parents and larger effects for the complete removal of childcare costs than halving them. Full subsidizing childcare increases women's lifetime earnings by about 7.5%. For comparison, in the reduced form estimates, we saw that the child penalty for women living in low childcare cost states was about 10% lower than for women in high-cost states.

Table 2.16: Effects of Childcare Subsidies

Panel A: Impacts of Halving Childcare Costs

Sample	Wages	Wages, $\gamma_1 = \infty$	Years x	Years x, $\gamma_1 = \infty$	WTP	WTP, $\gamma_1 = \infty$	# Moves	Share Time with Parents
All	19.96	22.06	1.21	1.29	16.80	16.12	0.01	0.00
$\tau^P \neq 0$	7.77	9.40	0.47	0.55	8.64	6.80	0.04	-0.01
$\tau^P = 0$	25.17	27.46	1.52	1.61	20.29	20.10	0.00	0.00
Never SM	14.87	17.22	0.88	0.99	12.52	11.84	0.00	0.00
Ever SM	24.52	26.39	1.50	1.56	20.58	20.00	0.03	0.00
Whites	20.79	25.73	1.14	1.36	18.14	17.73	0.01	0.00 2
Blacks	18.89	19.10	1.26	1.26	14.02	12.86	0.03	0.00

Panel B: Impacts of Removing Childcare Costs

Sample	Wages	Wages, $\gamma_1 = \infty$	Years x	Years x, $\gamma_1 = \infty$	WTP	WTP, $\gamma_1 = \infty$	# Moves	Share Time with Parents
All	26.82	29.93	1.72	1.86	28.35	26.60	0.03	-0.01
$\tau^P \neq 0$	11.39	13.08	0.72	0.79	17.96	11.84	0.10	-0.03
$\tau^P = 0$	33.41	37.13	2.15	2.31	32.72	33.01	0.00	0.00
Never SM	20.28	23.69	1.27	1.42	21.55	19.90	0.01	-0.01
Ever SM	32.67	35.51	2.12	2.24	34.47	32.62	0.05	-0.01
Whites	28.49	36.50	1.64	1.97	30.72	29.90	0.02	-0.01
Blacks	25.26	25.59	1.79	1.79	23.57	21.07	0.07	-0.01

Notes: SM = Single Mother. Table presents average impacts of counterfactual scenarios on lifetime wages, years of experience, number of moves, and fraction of time spent in parent's location. Second columns for wage and experience changes and WTP conduct counterfactual experiment when moving costs are infinite. Monetary units scaled by $2,080. Results for whites and Blacks computed using separate parameter estimates shown in Table 2.11. See text for details on estimation sample and procedure.

Among all women in our sample, the complete removal of childcare costs raises lifetime moves by 0.03, or roughly 2 percent. Similar to before, the effects on earnings are concentrated among women who are at any point single mothers, and the migration effects are particularly strong for single mothers and Blacks. The effects on earnings and wages are stronger for women who do not have helpful grandparents, since grandparent childcare does crowd out the labor force effects of the policies. However, the migration effects are largest for women who *do* have helpful grandparents, since it is these women for whom the geographic constraint induced by grandparent childcare applies.

A key feature of our results is that they demonstrate that ignoring labor mobility may misstate certain effects of childcare policies. Notably, the effects of the policies on experience and wages are typically larger in the version of the model where moving is prohibited — while this may seem counterintuitive, this happens because women in the no-moving world choose to move to locations that induce labor force participation, such as higher-paying locations or locations with grandparents. This depresses labor force participation in the baseline world with no moving, leaving more room for improvement from the counterfactual policies. Overall, though, the wage effects are fairly comparable regardless of whether moving is allowed.

However, ignoring migration results in the welfare effects of the policies being consistently understated. Across all individuals, the average willingness to pay for the full removal of childcare costs is approximately $\$28 \times 2080\$58,240$, which compared to the average cost of full-time childcare for five years $\$25 \times 2080 = \$52,00$ suggests that the policies on a whole may be welfare improving. Unsurprisingly, the willingness to pay for the policies is considerably higher for women that benefit more from them, such as single mothers, Whites (who, recall, have fewer helpful grandparents than Blacks), and women with parents who do not provide childcare. At the same time, the migration mechanism matters more for the groups that more often receive grandparent assistance — compared to the WTP in the world without migration, allowing for migration increases WTP by over 50% for women with helpful grandparents, while it decreases WTP by about 1% for women without helpful grandparents. The same qualitative pattern holds for Blacks and Whites.

While fertility in our model is exogenous, we also assess the sensitivity of these results to allowing a simple fertility elasticity in response to the policies — in particular, we assume that fertility increases by 10% following the halving of childcare costs and by 20% with the full removal, following an elasticity of fertility to reductions in childcare costs estimated by Haan and Wrohlich (2011)[27]. Table 2.A.1 presents the results of this robustness check, reassuringly finding comparable effects of the policies on experience and wages (often larger than in our baseline, since the heightened probability of children in the future appears to induce agents to work more in the present). The additional children result in decreased migration, which results in the WTP for the policies being lower when migration is allowed compared to when not, but this is a mechanical result that should be interpreted with caution, since it would clearly not apply if fertility were endogenous.

2.7 Conclusion

This paper studies how childcare costs, the location of extended family, and fertility events influence both the labor force attachment and labor mobility of women in the United States. Using both empirical evidence and a dynamic structural model, we argue that the draw of intergenerational time transfers results in substantial geographic constraints for women with children, thus resulting in children impacting both whether and where women work.

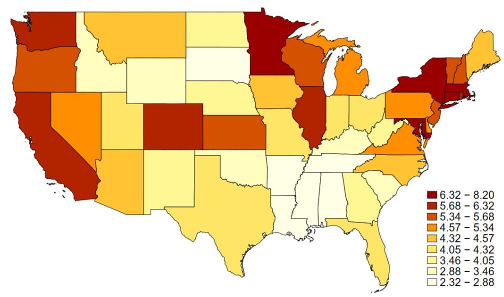
[27]Though note that other entries in the literature do not estimate such an elasticity — Bick (2016), for instance, finds that such subsidies do not increase fertility at all. Guner et al. (2020) also model fertility as exogenous, arguing: "We doubt that the inclusion of endogenous parental choices in the analysis could change our quantitative findings in a significant way. Specifically on fertility, child related policies that lead to higher participation rates are unlikely to alter parental decisions. There are countervailing effects that are expected to cancel each other out. Childcare costs are only a small fraction of the lifetime costs of raising children, and a reduction in these costs is balanced by increases in tax rates needed to finance the expansion of childcare subsidies."

As a result, focusing only on labor force participation is likely to understate the potential impacts of childcare policies on wages and welfare of U.S. women. Heterogeneity analyses in our model suggest that accounting for migration is particularly important for women who are ever single parents and for Blacks. With the COVID-19 crisis introducing substantial upheaval into childcare markets, these geographic constraints may have markedly increased in their importance in recent years.

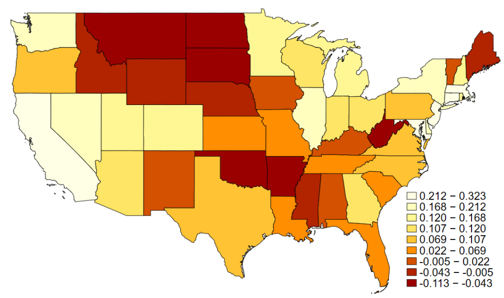
While the dimensionality reductions in our model result in a high degree of tractability, they do involve considerably suppressing geographic heterogeneity in childcare costs and wage effects. Enriching the geography of the model would be highly desirable in that it would allow us to better capture the extent to which women forego opportunities in stronger labor markets in the face of fertility events. Investigating how these decisions differ among women in and out of urban environments would also be interesting. Further enrichening the wage process in the model to account for additional unobserved heterogeneity would also be highly desirable. Another important limitation in our framework is our assumption that fertility is exogenous. Allowing for partial or complete fertility control among women in our framework would allow us to study additional behavioral responses to childcare policies, though robustness exercises suggest that such responses may be small. However, these issues may offer promising avenues for future research.

2.A Supplementary Figures and Tables

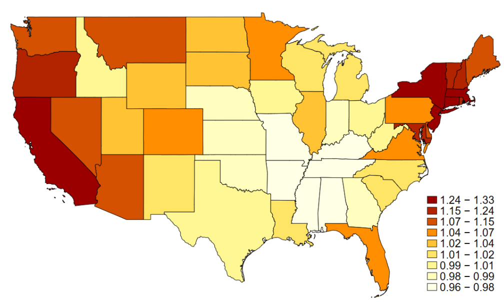
Figure 2.A.1: State Characteristics



(a) Childcare Costs



(b) Wage Effects



(c) Costs of Living

Notes: Data on childcare costs from Childcare Aware 2017 report. Units measured in 1000s of 2012 dollars. Wage effects from Mincerian regressions for men in American Community Survey, and costs of living from ACCRA. Wage effect and living cost for Iowa normalized to 0 and 1.

Table 2.A.1: Effects of Childcare Subsidies with Fertility Response

Panel A: Impacts of Halving Childcare Costs

Sample	Wages	Wages, $\gamma_1 = \infty$	Years x	Years x, $\gamma_1 = \infty$	WTP	WTP, $\gamma_1 = \infty$	# Moves	Share Time with Parents
All	28.82	32.32	1.75	1.90	9.51	15.83	-0.24	0.02
$\tau^P \neq 0$	16.04	19.01	0.95	1.09	0.49	6.31	-0.20	0.02
$\tau^P = 0$	34.28	38.00	2.09	2.25	13.30	19.90	-0.25	0.03
Never SM	25.50	29.55	1.52	1.71	5.15	10.87	-0.25	0.02
Ever SM	31.79	34.79	1.95	2.08	13.30	20.29	-0.22	0.03
Whites	31.52	45.40	1.73	2.36	25.46	78.25	-0.10	-0.03
Blacks	26.22	25.83	1.67	1.72	2.95	36.25	-0.22	0.04

Panel B: Impacts of Removing Childcare Costs

Sample	Wages	Wages, $\gamma_1 = \infty$	Years x	Years x, $\gamma_1 = \infty$	WTP	WTP, $\gamma_1 = \infty$	# Moves	Share Time with Parents
All	33.82	37.10	2.16	2.30	18.74	24.37	-0.26	0.02
$\tau^P \neq 0$	19.27	21.41	1.20	1.29	7.96	9.61	-0.18	0.00
$\tau^P = 0$	40.03	43.81	2.57	2.74	23.40	30.58	-0.29	0.02
Never SM	29.99	33.75	1.87	2.05	11.75	16.80	-0.30	0.02
Ever SM	37.24	40.11	2.41	2.54	25.05	31.07	-0.23	0.02
Whites	37.25	51.06	2.16	2.80	38.97	91.75	-0.13	-0.05
Blacks	30.83	30.39	2.08	2.13	8.39	41.61	-0.23	0.04

Notes: SM = Single Mother. Table presents average impacts of counterfactual scenarios on lifetime wages, years of experience, number of moves, and fraction of time spent in parent's location. Second columns for wage and experience changes and WTP conduct counterfactual experiment when moving costs are infinite. Monetary units scaled by \$2,080. Results for whites and Blacks computed using separate parameter estimates shown in Table 2.11. Fertility probability increases by 10% in halved childcare cost counterfactual and 20% in full subsidy counterfactual. See text for details on estimation sample and procedure.

2.B Model Solution Details

This section details the procedure for computing reservation levels of transient wage components ε and expected value functions when solving the model using the backward induction method described in Section 2.3.3. There are three stages of life in which households are making decisions: the post-children period (41-65), the post-fertility period (36-40), and the fertility period. The agents in our model can move at any point and the marital state at age 35 is assumed to be maintained for the remainder of the life cycle. We focus on reservation wages for age 65 and 64 here to build intuition about decisions without young children and for age 39 for decisions with young children; the procedure for earlier ages is identical after accounting for uncertainty over realizations of marriage and fertility shocks.

In the post-children period, the households no longer ever have young children (i.e., $a_c = \emptyset$). The agent makes a decision of whether they should work or not work, which will depend on the realizations of the ε shock. At age 65, given the other elements of the state space Ω, flow utility from working and not working u^1_{65}, u^0_{65} is given by

$$u^1_{65}(\Omega, \ell) = \alpha_1(w_S(\Omega, \ell)) + \alpha_1 \exp(\beta_0 + \beta_1 e + \beta_2 x + \beta_3 x^2 + \mu + \eta^\ell + \varepsilon) + \alpha_3 \mathbb{1}(p = 0) + \alpha_4 \mathbb{1}(\ell = \ell^P)$$

$$u^0_{65}(\Omega, \ell) = (\alpha_1 + \alpha_c)(w_S(\Omega, \ell)) + \alpha_2 + \alpha_e e + \alpha_x x + \alpha_\mu \mathbb{1}\{\mu = \mu^H\} + \alpha_3 \mathbb{1}(p = 1) + \alpha_4 \mathbb{1}(\ell = \ell^P)$$

$w_S(\Omega, \ell)$ is the realization of spousal income, conditional on spousal characteristics contained within (Ω, ℓ). p is the participation decision the previous period which determines if the person receives the switching cost α_3. We ignore costs of living differences and amenities in this formulation, but these can easily be accounted for by dividing α_1 and α_c by the relevant κ^ℓ or by adding the relevant $\boldsymbol{\alpha_\Gamma}$.

Because agents are finitely lived, $V_{66} = 0$ and the value function in the terminal period is then

$$V_{65}(\Omega_{65}, \ell) = \max\left\{u^1_{65}(\Omega, \ell), u^0_{65}(\Omega, \ell)\right\}$$

Participation is governed by the following:

$h = 1$ if

$$\varepsilon_{65} > \log\left(\frac{\alpha_2 + \alpha_e e + \alpha_x x + \alpha_c w_S(\Omega, \ell) + \alpha_3\big(\mathbb{1}(p = 1) - \mathbb{1}(p = 0)\big)}{\alpha_1}\right) - \underbrace{(\beta_0 + \beta_1 e + \beta_2 x + \beta_3 x^2 + \eta^\ell)}_{G_{65}(\Omega,\ell)} \equiv \varepsilon_{65}^{**}(\Omega, \ell)$$

$h = 0$ otherwise

We can then use this decision rule to calculate the expected utility following the optimal age-65 labor supply choice. That is,

$$\begin{aligned}
\mathbb{E}_{\varepsilon}[V_{65}(\Omega,\ell)] &= \mathbf{Pr}\big(\varepsilon_{65} > \varepsilon_{65}^{**}\big)\mathbb{E}[u_{65}^{1}(\Omega,\ell)|\varepsilon_{65} > \varepsilon_{65}^{**}] + \mathbf{Pr}\big(\varepsilon_{65} < \varepsilon_{65}^{**}\big)\mathbb{E}[u_{65}^{0}(\Omega,\ell)|\varepsilon_{65} < \varepsilon_{65}^{**}] \\
&= \alpha_1 w_S + \alpha_4 \mathbb{1}(\ell = \ell^P) \\
&\quad + \mathbf{Pr}\big(\varepsilon_{65} > \varepsilon_{65}^{**}\big)\alpha_3 \mathbb{1}[p=0] \\
&\quad + \mathbf{Pr}\big(e^{\varepsilon_{65}} > e^{\varepsilon_{65}^{**}}\big)\alpha_1 \mathbb{E}(e^{\varepsilon}|e^{\varepsilon_{65}} > e^{\varepsilon_{65}^{**}})\exp(G_{65}(\Omega,\ell)) \\
&\quad + \mathbf{Pr}\big(\varepsilon_{65} < \varepsilon_{65}^{**}\big)\big(\alpha_2 + \alpha_e e + \alpha_x x + \alpha_c w_S(\Omega,\ell) + \alpha_3 \mathbb{1}[p=1]\big) \\
&= \alpha_1 w_S + \alpha_4 \mathbb{1}(\ell = \ell^P) \\
&\quad + \left[1 - \Phi\left(\frac{\varepsilon_{65}^{**}}{\sigma_{\varepsilon}}\right)\right]\alpha_3 \mathbb{1}[p=0] \\
&\quad + \left[1 - \Phi\left(\frac{\varepsilon_{65}^{**} - \sigma_{\varepsilon}^2}{\sigma_{\varepsilon}}\right)\right]\alpha_1 e^{0.5\sigma_{\varepsilon}^2 + G_{65}(\Omega,\ell)} \\
&\quad + \Phi\left(\frac{\varepsilon_{65}^{**}}{\sigma_{\varepsilon}}\right)\big(\alpha_2 + \alpha_e e + \alpha_x x + \alpha_c w_S(\Omega,\ell) + \alpha_3 \mathbb{1}[p=1]\big).
\end{aligned}$$

Moving back to period 64, the agent will end the period by realizing their location preference shocks and choosing their optimal age-64 location, ℓ', conditional on their current state (Ω), the participation decision made at the beginning of period 64 (h), and their current location (ℓ):

$$\ell' = \arg\max_{k \in \mathbb{N}^{\ell}} \left(\mathbb{1}(k \neq \ell) \times \underbrace{\left(\gamma_0 + \gamma_1 e + \gamma_3 m + \gamma_4 N^k\right)}_{\Delta_k} + \beta \left(\mathbb{E}_{\varepsilon} \sum_{\Omega'} [V_{65}(\Omega', k)] \mathbf{Pr}(\Omega' | \Omega, h, k) \right) + \zeta_k \right)$$

With the assumption that these shocks are drawn from the type-1 extreme value location with a variance normalized to 1, we can calculate the probability of choosing location ℓ':

$$Pr(\ell_{64} = \ell' | \Omega, \ell, h) = \frac{\exp\left(\mathbb{1}(\ell' \neq \ell) \times \Delta_{\ell'} + \beta\left(\mathbb{E}_\varepsilon \sum_{\Omega'} [V_{65}(\Omega', \ell')] \mathbf{Pr}(\Omega' | \Omega, h, \ell')\right)\right)}{\sum_k \exp\left(\mathbb{1}(k \neq \ell) \times \Delta_k + \beta\left(\mathbb{E}_\varepsilon \sum_{\Omega'} [V_{65}(\Omega', k)] \mathbf{Pr}(\Omega' | \Omega, h, k)\right)\right)}$$

and the expected utility following the optimal decision as:

$$\mathbb{E}_{\zeta_{\ell'}}[V'_{64}(\Omega, \ell; h)] = \bar{\gamma} + \log\left(\sum_{\ell'} \exp\left(\beta \sum_{\Omega'} \mathbb{E}_\varepsilon [V_{65}(\Omega', \ell')] \mathbf{Pr}(\Omega' | \Omega, h, \ell') - \Delta_{\ell'} \mathbb{1}\{\ell' \neq \ell\}\right)\right).$$

This then allows us to express the age-64 value function as:

$$V_{64}(\Omega, \ell) = \max\left\{u^1_{64}(\Omega, \ell) + \mathbb{E}_{\zeta_{\ell'}}[V'_{64}(\Omega, \ell; 1)], u^0_{64}(\Omega, \ell) + \mathbb{E}_{\zeta_{\ell'}}[V'_{64}(\Omega, \ell; 0)]\right\},$$

where $u^1_{64}(\Omega, \ell)$ and $u^0_{64}(\Omega, \ell)$ are defined comparably to their age-65 counterparts. The decision rule for working given ε_{64} is then given by:

$h = 1$ if

$$\varepsilon_{64} > \log\left(\frac{\alpha_2}{\alpha_1} + \frac{\alpha_e}{\alpha_1} e + \frac{\alpha_x}{\alpha_1} x + \frac{\alpha_c}{\alpha_1} w_S(\Omega, \ell) + \frac{\alpha_3}{\alpha_1}\left(\mathbb{1}(p = 1) - \mathbb{1}(p = 0)\right)\right.$$
$$\left. + \frac{1}{\alpha_1} \mathbb{E}_{\zeta_{\ell'}}\left([V'_{64}(\Omega, \ell; 0)] - \mathbb{E}_{\zeta_{\ell'}}[V'_{64}(\Omega, \ell; 1)]\right)\right) - G_{64}(\Omega, \ell) \equiv \varepsilon^{**}_{64}(\Omega, \ell)$$

$h = 0$ otherwise.

This then allows us to express the expected utility following the age-64 labor supply choice as follows:

$$\begin{aligned}
\mathbb{E}_{\varepsilon}[V_{64}(\Omega,\ell)] &= \alpha_1 w_S + \alpha_4 \mathbb{1}(\ell = \ell^P) \\
&+ \left[1 - \Phi\left(\frac{\varepsilon_{64}^{**}}{\sigma_\varepsilon}\right)\right]\left(\alpha_3 \mathbb{1}[p=0] + \mathbb{E}_{\zeta_{\ell'}}[V'_{64}(\Omega,\ell;1)]\right) \\
&+ \left[1 - \Phi\left(\frac{\varepsilon_{64}^{**} - \sigma_\varepsilon^2}{\sigma_\varepsilon}\right)\right]\alpha_1 e^{0.5\sigma_\varepsilon^2 + G_{64}(\Omega,\ell)} \\
&+ \Phi\left(\frac{\varepsilon_{64}^{**}}{\sigma_\varepsilon}\right)\left(\alpha_2 + \alpha_e e + \alpha_x x + \alpha_c w_S(\Omega,\ell) + \alpha_3 \mathbb{1}[p=1] + \mathbb{E}_{\zeta_{\ell'}}[V'_{64}(\Omega,\ell;0)]\right),
\end{aligned}$$

which in turn allows us to compute age-63 continuation values, and so on. This continues recursively in the same fashion until we reach age 39, which is the last year in which an agent may have a young child. For those without children at 39, the decision process is unchanged. For those with a child, they now have the costs of childcare to consider in their hours decision and the location of parents as a source of cheaper care to consider in their location decision.

For a person with a young child, given the other elements of the state space Ω, flow utility from working and not working u^1, u^0 is given by:

$$\begin{aligned}
u^1(\Omega,\ell,a_c \neq \emptyset) &= \alpha_5\Big(w_S(\Omega,\ell) + \exp\left(\beta_0 + \beta_1 e + \beta_2 x + \beta_3 x^2 + \eta^\ell + \varepsilon\right) \\
&\quad -\delta_\ell\big(1 - \tau_{pm}\mathbb{1}(\ell = \ell^P) - \tau_s m\big)\Big) + \alpha_3 \mathbb{1}(p=0) + \alpha_7 \mathbb{1}(\ell = \ell^P); \\
u^0(\Omega,\ell,a_c \neq \emptyset) &= (\alpha_5 + \alpha_c)(w_S(\Omega,\ell)) + \alpha_6 + \alpha_e e + \alpha_x x + \alpha_3 \mathbb{1}(p=1) + \alpha_7 \mathbb{1}(\ell = \ell^P).
\end{aligned}$$

At age 39, the expected utility following the optimal location decision, $\mathbb{E}_{\zeta_{\ell'}}[V'_{39}(\Omega,\ell;h)]$ follows the same general form as the previously described expected utility in period 64. This, combined with the flow utility, gives us the following participation decision rule:

$h = 1$ if

$$\varepsilon_{39} > \log\left(\frac{\alpha_6}{\alpha_5} + \frac{\alpha_e}{\alpha_5}e + \frac{\alpha_x}{\alpha_5}x + \frac{\alpha_c}{\alpha_5}w_S(\Omega,\ell) + \frac{\alpha_3}{\alpha_5}\big(\mathbb{1}(p=1) - \mathbb{1}(p=0)\big)\right.$$
$$\left. + \frac{1}{\alpha_5}\mathbb{E}_{\zeta_{\ell'}}\left([V'_{39}(\Omega,\ell;0)] - \mathbb{E}_{\zeta_{\ell'}}[V'_{39}(\Omega,\ell;1)]\right) + \delta_\ell\big(1 - \tau_{pm}\mathbb{1}(\ell = \ell^P) - \tau_s m\big)\Big)\right) - G_{39}(\Omega,\ell) \equiv \varepsilon^{**}_{39}(\Omega,\ell);$$

$h = 0$ otherwise.

There are two notable differences in the reservation wage for women with children relative to those without. First, the parameters governing valuation of consumption (α_5) and the value of leisure (α_6) differ, potentially raising the reservation wage relative to non-mothers if α_6 is higher than α_2 or lowering the reservation wage if α_5 is higher than α_1. Second, the added cost of childcare δ^ℓ will raise the reservation wage for mothers.

This then allows us to express the expected utility following the age-64 labor supply choice as follows:

$$\begin{aligned}
\mathbb{E}_\varepsilon[V_{39}(\Omega,\ell)] &= \alpha_5 w_S + \alpha_7\mathbb{1}(\ell = \ell^P) \\
&+\left[1 - \Phi\left(\frac{\varepsilon^{**}_{39}}{\sigma_\varepsilon}\right)\right]\left(\alpha_3\mathbb{1}[p=0] - \alpha_5\delta_\ell\big(1 - \tau_{pm}\mathbb{1}(\ell=\ell^P) - \tau_s m\big) + \mathbb{E}_{\zeta_{\ell'}}[V'_{39}(\Omega,\ell;1)]\right) \\
&+\left[1 - \Phi\left(\frac{\varepsilon^{**}_{29} - \sigma^2_\varepsilon}{\sigma_\varepsilon}\right)\right]\alpha_5 e^{0.5\sigma^2_\varepsilon + G_{39}(\Omega,\ell)} \\
&+\Phi\left(\frac{\varepsilon^{**}_{39}}{\sigma_\varepsilon}\right)\left(\alpha_6 + \alpha_e e + \alpha_x x + \alpha_c w_S(\Omega,\ell) + \alpha_3\mathbb{1}[p=1] + \mathbb{E}_{\zeta_{\ell'}}[V'_{39}(\Omega,\ell;0)]\right),
\end{aligned}$$

which in turn allows us to compute age-38 continuation values. The location choice decision at age 38 takes the same form as previously, though with the addition of a component of the moving cost for parents with a young child and with an additional component of expected future utility (i.e., the future childcare costs) varying by location. The process continues in this manner until age 35, at which fertility shocks and marriage shocks enter the model. As these shocks merely impact the probability of being in a given state Ω in the following period, we omit discussion of life-period one decision solutions; the remainder of the model can thus be solved in a similar fashion until reaching age 22.

2.C Divisional Groupings of States

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

Chapter 3

The Long-Run Impacts of Court-Ordered Desegregation

Joint with Jason Fletcher and Owen Thompson

Chapter Summary

Court-ordered desegregation plans were implemented in hundreds of US school districts nationwide from the 1960s through the 1980s, and were arguably the most ambitious national attempt to improve educational access for African American children in modern American history. Using large Census samples that are linked to Social Security records containing county of birth, we implement event studies that estimate the long-run effects of exposure to desegregation orders on human capital and labor market outcomes. Within the South, we find that African Americans who were relatively young when a desegregation order was implemented in their county of birth, and therefore had more exposure to integrated schools, experienced large improvements in adult human capital and labor market outcomes relative to Blacks who were older when a court order was locally implemented. There are no comparable changes in outcomes among whites in southern counties undergoing an order. In contrast, outside of the South we find no evidence that desegregation orders impacted the

adult outcomes of Black or white students. Our data and methodology provide the most comprehensive national assessment to date on the impacts of court-ordered desegregation, and indicate that these policies were in fact highly effective at improving the long-run socioeconomic outcomes of Black students, but only in regions with a history of state sponsored de-jure segregation.

3.1 Introduction

Beginning with the 1954 *Brown v. Board of Education* ruling and continuing through the 1990s, most large school districts in the US were placed under court orders requiring them to reduce student segregation by race. These orders were extremely controversial and often faced fierce local resistance, but were nonetheless substantively implemented and were almost always followed by significant increases in racial integration (Welch and Light, 1987; Kluger, 2011).

Court-ordered school desegregation arguably constituted the most ambitious attempt in modern US history to reduce racial inequality in educational access, and a full understanding of its impacts is of clear importance from both a research and policy perspective. This paper provides new evidence on what is perhaps the central question regarding the efficacy of these orders: whether they improved the long-run socioeconomic outcomes of the minority students they were intended to benefit.

Key to our contribution is the use of Census and ACS samples that have recently been matched to respondent's counties of birth using the Social Security Administration's Numident file. This makes it possible to observe both childhood geographic locations and adult human capital and labor market outcomes for several million individuals who were attending school in the period when major desegregation orders were being rolled out. As discussed below, previous studies on this topic have had to either rely on much smaller samples like the PSID to observe both childhood location and adult outcomes (e.g. Johnson (2011)), or have used larger Census samples but have only been able to estimate contemporaneous or short run effects (e.g. Guryan (2004)).

Using this novel data source and exploiting the staggered timing of court orders, we estimate event study specifications that compare the long-run outcomes of individuals who were relatively young when an order was implemented in their county of birth, and therefore had more exposure to the post-integration educational environment, to individuals who were older when an order was locally implemented and therefore had fewer years of post-integration schooling. In conjunction with the large sample sizes available in our data, this approach allows us to precisely estimate non-parametric "dose-response" relationships between years of exposure to the court orders and adult outcomes, rather than relying on binary or linear treatment measures.

Integration orders plausibly impacted multiple inputs of the educational production function simultaneously, including peers and school resources, and potentially had subtle psycho-social benefits as well by promoting more generally equitable and inclusive education systems.[1] To gain a better sense of what the treatment actually consisted of, we begin by reporting "first stage" estimates of how the court orders in our sample impacted contemporaneous measures of racial segregation and school characteristics. These analyses indicate that the court orders we study led to increased per-pupil funding levels in districts with more Black students, and also caused Black students to be exposed to more white peers, although we do observe white enrollment declines after court orders that make the change in peers smaller than they otherwise would have been. Notably, these first stage estimates indicate that court orders had much larger effects on both peer composition and school resources in the South than in other regions, while white flight was strongest outside of the South. We may expect any psycho-social benefits of the orders to be greater in the South as well, given the region's history of overt state-sponsored discrimination and de-jure school segregation.

We then report our main reduced form estimates of how childhood exposure to desegregation orders affected a variety of adult socioeconomic outcomes. These analyses indicate that greater exposure to post-integration educational environments did indeed improve the

[1] While difficult to quantify, the psychological benefits of integrated education weighed substantially in key judicial rulings, with Chief Justice Warren famously writing in the *Brown* decision that segregation "generates a feeling of inferiority as to [minority students'] status in the community that may affect their hearts and minds in a way unlikely to ever be undone."

human capital and labor market outcomes of African Americans in adulthood. Mirroring the first stage estimates, we find that these effects were concentrated in the South, with no substantive effects outside of the South. Among Southern Blacks, having a desegregation order implemented in an individual's county of birth prior to age 5 improves an index of their adult human capital by over .3 standard deviations relative to having an order implemented at age 17, and increases an index of their economic self-sufficiency by approximately 0.5 standard deviations. Significant effects are also observed across a variety of disaggregated adult outcomes variables, including high school completion, years of educational attainment, labor force attachment, and poverty levels.

Our estimation approach has the desirable feature of building in placebo tests that help establish the key identifying assumption of parallel trends. Most importantly, we do not expect to observe any trends in adult outcomes across individuals who were over age 17 when an order was implemented, since none of these individuals had any exposure to court-ordered desegregation.[2] Additionally, while the expected effects of court orders on white students are not necessarily zero, and while we consider evaluating whether court-ordered integration had any negative impacts on white students as one important contribution of our study, large effects among whites would raise concerns that unobserved factors correlated with the desegregation orders may be driving the effects we observe among Blacks, and whites can therefore serve as a useful secondary comparison group.

Reassuringly, we find no significant trends across individuals ages 17-24 at the time of a local order, and we consistently find little or no effect on whites, patterns that are consistent with the parallel trends assumption holding. Our key findings are also robust to a variety of modeling and specification choices, and to implementing a version of the procedure suggested by Sun and Abraham (2021) to address potential biases that can occur in two-way fixed

[2]As we discuss below, expected trends across younger age-at-order values are more ambiguous, so we consider ages 17-24 to be the primary test of "pre-trends."

effects specifications in settings like ours where treatment timing varies across units, which is the topic of an influential recent methodological literature.

Our paper contributes to several important literatures. Most directly, we build on a number of previous studies evaluating the effects of school desegregation on student outcomes, primarily but not exclusively educational attainment (Guryan, 2004; Angrist and Lang, 2004; Rivkin and Welch, 2006; Ashenfelter et al., 2006; Card and Rothstein, 2007; Reber, 2010; Lutz, 2011; Johnson, 2011; Bergman, 2018; Tuttle, 2019; Angrist et al., 2022).

Of these, our work most closely resembles Guryan (2004) and Johnson (2011).[3] Guryan (2004) uses Census data to estimate whether the high school drop-out rates of young African Americans differentially improved in counties where desegregation rulings were implemented between 1970 and 1980, and finds significant relative reductions in Black dropout rates in counties undergoing orders. Because Guryan (2004) uses Census samples containing only *current* county of residence among young adults, rather than geographic locations during childhood, he focuses on contemporaneous high school completion as the outcome measure and relies primarily on binary difference-in-difference models across the 1970 and 1980 Censuses rather than dynamic event study specifications. In contrast, Johnson (2011) uses data from the Panel Study of Income Dynamics (PSID) that does contain childhood location, as well as a wide variety of adult outcomes, and estimates event study models broadly similar to our own preferred specifications below. However, after basic restrictions, his PSID sample contains fewer than 4,500 Black respondents spread across over 600 school districts, which can sometimes make the estimates imprecise or unstable.[4]

We build on these studies by using comprehensive exposure and long-run outcome mea-

[3] Johnson (2011) is a working paper, with many of its key findings published in book form for a popular audience in Johnson (2019).

[4] Angrist and Lang (2004), Bergman (2018) and Tuttle (2019) also study the effects of desegregation on student outcomes, but are distinct from our work and that of Guryan (2004) and Johnson (2011) in that they are each case studies of a single desegregation plan and use idiosyncratic features of those plan's assignment rules to provide compelling identification but less comprehensive scope. Additionally, all three of these case studies are of desegregation plans that spanned school multiple districts, which were in general very rare.

sures within a national sample that is sufficiently large and representative to produce precise estimates and to evaluate heterogeneity across basic characteristics like region and race. Our combination of data and research design allow us to provide, in our view, the most comprehensive national assessment of the long-run impacts of court-ordered desegregation to date.

Our work is also related to a broader literature that evaluates the effects of school desegregation on a wide variety of outcomes including white flight (Reber, 2005; Baum-Snow and Lutz, 2011), school finances (Cascio et al., 2010; Reber, 2011), teacher labor markets (Jackson, 2009; Thompson, ming), and crime (Weiner et al., 2009). Our findings additionally have implications for the extensive literature on how factors like school resources and teacher characteristics influence student outcomes (Hanushek, 1986, 2003; Jackson et al., 2016; Hyman, 2017; Lafortune et al., 2018; Cascio et al., 2013; Card and Krueger, 1992b,a; Chetty et al., 2011). Many of these studies find significant heterogeneity by race, and our analysis provides new evidence that "schools matter" for shaping long-run adult outcomes as well as influencing racial inequality.

Finally, from a methodological perspective our work builds on a number of recent papers that have used large national data sets newly linked with county-level exposure variation to evaluate the long-run impacts of various policies and shocks. These include Community Health Centers (Bailey and Goodman-Bacon, 2015), Head Start (Bailey et al., 2021), the Food Stamp Program (Hoynes et al., 2016; Bailey et al., 2020), air pollution (Isen et al., 2017) and recessions (Stuart, 2022).

3.2 Background

The landmark 1954 *Brown v. Board of Education of Topeka Kanses* decision ruled that the de-jure segregation of schools was unconstitutional, overturning the "separate but equal" doctrine of *Plessy v. Ferguson* that had prevailed since 1896. While some meaningful school desegregation did occur after *Brown* in "border" areas like Washington DC and West Virginia, the ruling lacked strong enforcement mechanisms, and even a full decade after *Brown* fewer than 5% of Black students in the eleven states of the Former Confederacy were attending integrated schools (Cascio et al., 2010).

The pace of school integration accelerated dramatically after passage of the 1964 Civil Rights Act, which authorized the US attorney general to bring suits against districts failing to desegregate (Title IV) and allowed federal agencies to withhold funding to non-compliant state and local governments (Title VI). The latter provision was given greater bite by the 1965 passage of the Elementary and Secondary Education Act, which greatly expanded federal educational funding and increased the opportunity cost for schools to not comply with integration requirements (Cascio et al., 2013).

An additional key judicial ruling in 1968 was *Green v. County School Board of New Kent County.* Prior to *Green* there was significant legal ambiguity as to what constituted compliance with *Brown* and with the Civil Rights Act, especially the status of "freedom of choice" plans, which technically allowed minority students to enroll in historically white schools but typically led to only token desegregation. *Green* provided a far more specific and stringent set of criteria than previous rulings, requiring integration in every facet of school operations from staffing to transportation to extracurricular activities. Another landmark case from this period, *Swann v. Charlotte-Mecklenburg* (1971) explicitly sanctioned the use of district-wide busing to achieve desegregation. The *Green* and *Swann* rulings serve as the judicial basis for most of the desegregation orders that we study here.

While early judicial and legislative integration policy was heavily focused on de-jure school segregation in the South, in 1973 the Supreme Court ruled in *Keyes v. School District No. 1, Denver* that the forms of de-facto school segregation common in many large cities outside of the South were also unconstitutional. This paved the way for desegregation orders nationwide during the 1970s and into the early 1980s, including high profile cases in cities like Boston, Chicago, Los Angeles and Detroit that are included in our analysis.

The specific methods used by the courts to achieve racial integration were highly varied, and many districts adopted multiple plans over time that utilized different techniques. The broad approaches used in the court orders that we study include majority-to-minority transfer plans, which allow either all students or sometimes only minority students to voluntarily transfer into schools where their racial group was under-represented; the establishment of magnet schools with district wide catchment and race-specific enrollment preferences or quotas; a wide array of "rezoning" approaches that redrew attendance zones, often non-contiguously, to increase integration levels; and "pairing and clustering" approaches that paired together a predominantly white and a predominantly minority school then reassigned students across the schools, often through restructuring by grade level. These approaches proved far more effective at substantively reducing segregation than "freedom of choice" plans or other earlier approaches.

The ambitious scope of the described efforts bears emphasis. For a brief but critical period, a series of landmark judicial rulings and federal legislative efforts forcefully obligated US school systems to take robust actions to reduce racial inequality, even in the face of strong and sometimes violent local resistance. There have arguably not been equally substantive efforts to increase educational access for African American students before or since. Understanding the long-run impacts of these ambitious initiatives is of clear importance.

3.3 Data

3.3.1 Court Orders

We focus on the set of all districts where, as of 1968, there were at least 15,000 total students and where 10-90% of the students were Black, yielding a national sample of 187 medium to large districts with racially diverse student bodies.[5] This sample is substantially broader than the 125 district sample that was constructed by Welch and Light (1987) and used in numerous subsequent studies, while also maintaining focus on districts that are sufficiently large and diverse that the basic features of court-ordered desegregation policies are applicable.[6] The 187 districts in our sample are listed in the Online Appendix, and included 60% of all Black students attending public US schools in 1968.[7] For each of these districts, we then gathered data on desegregation plans from a variety of sources.

Many of the districts were included in the data collection efforts of Welch and Light (1987), and where available we use the year of the plan implementation listed in their Table A3.[8]

For districts not included in the Welch and Light (1987) sample, we use databases of court orders compiled by ProPublica (2014) and by the American Communities Project (Logan, 2021). In many cases these sources do not specify the date of plan implementation,

[5]Enrollments data are from the 1968 Office of Civil Rights school survey, which was generously made available by Sarah Reber.

[6]For instance smaller districts frequently operate only one high school, which will by construction have the same racial composition as the district's overall population of high school students, and race-based school reassignments are moot in districts where virtually all the students are white or Black.

[7]We focus on desegregation order's impacts among African Americans, rather than a broader set of racial and ethnic minorities, primarily because African Americans were the main focus of segregation related litigation efforts, although we do note that *Keyes* extended desegregation remedy requirements to Hispanics, and Antman and Cortes (2021) find significant benefits of pre-*Brown* integration rulings among Mexican-Americans in California. An additional, more practical consideration is that we are only able to observe location of birth for individuals with social security numbers, which is a more significant limitation among Hispanics and other minority populations with larger amounts of recent immigration than among Blacks.

[8]Many districts were subject to multiple court order over the course of several years, especially in the South. In these cases we use the year of the order that corresponded to the largest change in the dissimilarity index as listed in Table A3 of Welch and Light (1987).

or have conflicting dates and limited information about the substance of the relevant orders. In these instances, we use contemporaneous newspaper accounts, government reports, and other sources to determine the year in which court orders were substantively implemented in each of the studied districts. Detailed documentation of the sources used to date each court order, as well as notes related to district mergers, cross district desegregation plans, and other idiosyncratic features of specific orders, are provided in the Online Appendix.[9]

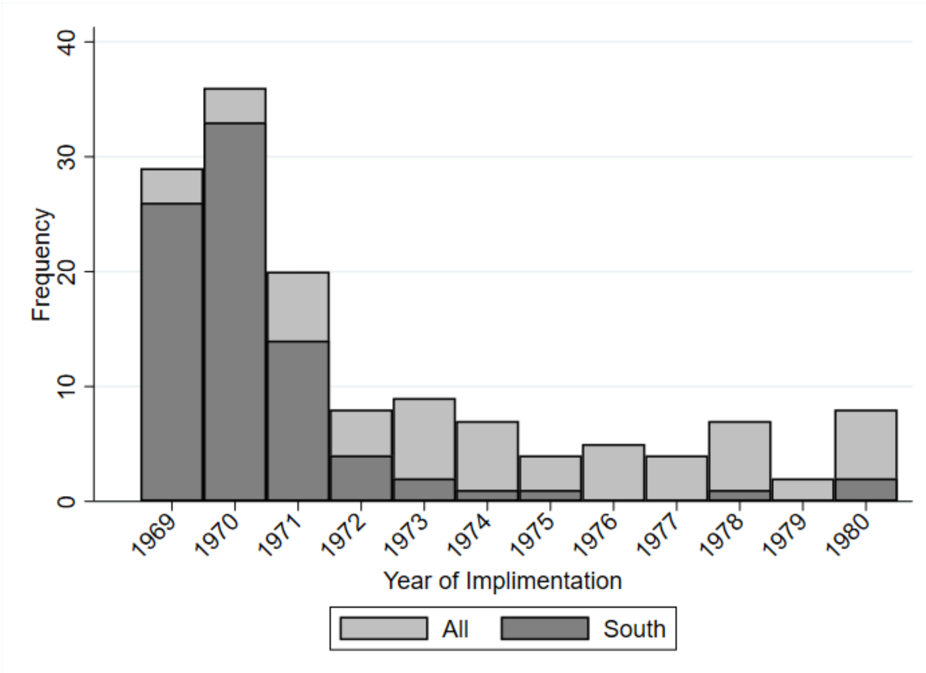
Districts that met the sample criteria but were never placed under a court order are retained and are used as controls in some specifications, as discussed in the methodology section below. Approximately 10% of our sample consists of these "untreated" districts. In practice, all southern districts in our sample were placed under desegregation orders at some point in the 1960s and 1970s, such that all of the untreated districts are located in the North, which as we discuss below has some implications for our empirical strategy, and more generally provides an additional reason for it being appropriate to disaggregate the analysis by region.

Both cities and counties are commonly used as school district boundaries in the US, whereas the Numident file we use for linking to childhood geographic information only consistently contains *county* of birth. In our baseline estimates we simply assign municipal districts to the county in which they were located, and as a robustness check we estimate our preferred models using the subset of districts that were the sole district within the county where they were located, where assignment of individuals to school districts is unambiguous. We also note that any measurement error due to the misassignment of individuals in city school districts to counties would likely bias our estimates towards zero.

Figure 3.1 shows the distribution of plan implementation years within the sample that

[9]Our use of the years that orders were *implemented* follows Guryan (2004) but differs from Johnson (2011), who uses the year that the orders were *ruled*, which he argues is more exogenous than implementation year. Because there were often lags of many years or even decades between rulings and implementations, and there were typically no significant reductions in segregation in these intermediate periods, we feel that the use of implementation dates is the preferable approach.

Figure 3.1: Desegregation Orders by Year and Region



Notes: Figure shows the years that the orders in our estimation sample were implemented, by region.

we use to estimate our preferred event study specifications.[10] The figure indicates that the most intensive period of court-ordered desegregation occurred in 1969, 1970 and 1971 and that these order were predominantly, but not exclusively, in the South. A steady flow of approximately five major rulings per year was then implemented throughout the 1970s, and these later rulings were concentrated outside of the South.

3.3.2 Outcome Data

We measure long-run outcomes using restricted versions of the 2000 Long-Form Census and the 2001-2015 American Community Survey (ACS). These samples are multiple orders of magnitude larger than available longitudinal samples like the PSID or NLSY, jointly covering

[10]As discussed below, to construct an estimation sample that is balanced in event time we restrict our main analysis to events occurring from 1969 through 1980.

over 20% of the US population. Critically, we are able to link these large Census and ACS samples to the SSA Numident file via Protected Identity Keys (PIKs; essentially scrambled Social Security Numbers) to obtain information on respondents' counties of birth.[11]

After basic restrictions, our working sample contains over 5.1 million individuals. These individuals identified as being non-Hispanic white or African American, and were born between 1945 and 1985. Following Bailey et al. (2021), we also limit the sample to individuals who are aged 25 to 54 and exclude observations that had allocated or missing values for any outcomes of interest, and also drop any individuals who are linked to the same PIK and who are assigned more than one possible county of birth. We collapse our data by birth year, survey year, county of birth, race, and sex in our regression analyses to reduce computation and weight all models by the number of people represented (that is, the sum of individual survey weights) in each cell (Solon et al., 2015).

Our primary outcomes of interest are summary indices of human capital and economic self-sufficiency, and we follow the outcome variable constructions used by Bailey et al. (2021). The human capital index includes binary indicators for attainment of a high school degree, some college, a four-year college degree, and an advanced/professional postgraduate degree, continuous years of schooling, and a indicator for working in a professional occupation. The economic self-sufficiency index includes dummy variables for employment, poverty status, income from public assistance, non-zero family income, and non-zero income from other non-governmental sources, continuous measures of weeks and hours worked, and the logs of labor income, income from non-governmental sources, and the ratio of family income to the poverty threshold. The poverty status and public assistance indicators are reverse-coded so that all positive subcomponent values indicate improvements. We convert all subcomponents

[11]Public-use versions of the long-form Census and ACS contain *state* of birth and in some cases *current* county of residence, but not county of birth. The Numident data provides a string variable recording place of birth as written onto Social Security card applications, and we use the crosswalk created by Taylor et al. (2016) in Census project 1248 to map string values in the Numident data to consistent birth counties.

into z-scores and average them while weighting each subcomponent equally. For transparency and completeness, we report estimates for each subcomponent in addition to the summary indices.[12]

Table 3.1 presents summary statistics for our working sample. The typical individual is exposed to a court desegregation order at approximately age 8, and is approximately age 41 when observed in the data. Education, labor force attachment, earnings and other outcome measures display commonly observed gaps with respect to race and sex.

[12]We also investigated as outcome variables incarceration, marital status, homeownership, and disability status and did not find evidence for exposure to school desegregation orders significantly impacting any of them. These results were not disclosed to reduce the number of disclosures submitted for review to Census Bureau.

Table 3.1: Summary Statistics

Variable	White Men	White Women	Black Men	Black Women
High School Degree	0.93	0.95	0.87	0.89
	(0.25)	(0.22)	(0.33)	(0.31)
Some College	0.61	0.48	0.50	0.50
	(0.49)	(0.48)	(0.50)	(0.50)
Four-Year Degree	0.36	0.38	0.18	0.23
	(0.48)	(0.49)	(0.39)	(0.42)
Advanced Degree	0.04	0.03	0.01	0.01
	(0.20)	(0.17)	(0.11)	(0.12)
Years of Schooling	13.96	14.10	12.98	13.37
	(2.60)	(2.49)	(2.24)	(2.31)
Professional Occupation	0.33	0.35	0.19	0.26
	(0.47)	(0.48)	(0.39)	(0.44)
Employed	0.87	0.75	0.72	0.73
	(0.33)	(0.44)	(0.45)	(0.45)
In Poverty	0.06	0.08	0.20	0.23
	(0.24)	(0.27)	(0.40)	(0.42)
Weeks Worked	45.06	37.73	37.27	37.26
	(15.54)	(21.00)	(21.63)	(21.35)
Hours per Week	41.28	30.63	33.80	31.49
	(15.40)	(18.01)	(19.75)	(17.73)
Wage Earnings (2012 $1,000)	59.74	33.39	33.77	28.07
	(59.27)	(38.12)	(36.42)	(28.81)
Hourly Wage	26.37	17.91	16.54	15.25
	(22.92)	(17.89)	(16.54)	(15.25)
Received Public Assistance Income	0.01	0.02	0.02	0.06
	(0.09)	(0.13)	(0.13)	(0.23)
Family Income-Poverty Ratio	4.79	4.58	3.52	2.95
	(2.65)	(2.70)	(2.36)	(2.30)
Married	0.65	0.65	0.46	0.34
	(0.48)	(0.48)	(0.50)	(0.47)
Incarcerated	0.01	0.00	0.07	0.01
	(0.10)	(0.05)	(0.25)	(0.08)
Homeowner	0.65	0.68	0.38	0.38
	(0.48)	(0.47)	(0.48)	(0.49)
Disabled	0.09	0.10	0.15	0.15
	(0.29)	(0.29)	(0.36)	(0.36)
Age when Treated	8.67	8.66	7.17	7.11
	(7.79)	(7.79)	(7.49)	(7.56)
Age	41.15	41.13	40.67	40.38
	(7.73)	(7.77)	(7.63)	(7.73)
Year of Birth	1965	1965	1966	1966
	(7.91)	(7.93)	(7.63)	(7.74)
N	2252000	2350000	238000	292000

Notes: Standard deviations in parentheses. Data from 2000 Long-Form Census and 2001-2015 American Community Survey. See text for details on sample restrictions. Observation counts rounded for disclosure avoidance purposes.

3.4 First Stage Effects

Before evaluating the effects of desegregation orders on the long-term outcomes of exposed children, we provide "first-stage" estimates that quantify how the orders impacted several contemporaneous school district characteristics. As noted, desegregation orders plausibly impacted the long-term outcomes of Black students both by changing the peers they were exposed to and by changing the resource levels at the schools they attended, so we study changes in school characteristics related to both of these potential mechanisms.

Specifically in Figure 3.2 we report the results of models that simply regress relevant school characteristics onto indicators for the number of years relative to an order (event time) as well as district and year fixed effects. The estimates from these models provide a more concrete picture of what the "treatment" generated by the orders actually consisted of, and also provide some empirically based guidance on which settings the largest effects on children's long-term outcomes could be expected in.

We begin by studying changes in peer composition using school level data on student racial compositions that was collected in Office of Civil Rights (OCR) surveys. This data is available annually from 1967-1974 and biennially from 1976-1980, for a total of 11 school years.[13] Because the OCR data is available for a relatively narrow range of years, and because our main goal in this section is simply to characterize the changes occurring after the orders were implemented, we prioritize including more treatment cohorts and observing a longer post-period over evaluating pre-trends over many years, and in particular we estimate event time coefficients ranging from -2 to +5, which allows us to use a panel that is balanced in event time and to retain all court orders occurring between 1969 and 1975.[14]

[13] The 1967 data was transcribed by the authors while subsequent years were generously digitized and provided by Sarah Reber. Of the 160 districts in our working sample that underwent desegregation orders, school level integration data is available in all 11 years for 149 of them.

[14] In our primary results for long-term outcomes, we estimate considerably more pre-treatment event time coefficients to help evaluate the parallel trends assumption.

The top panel of Figure 3.2 reports estimates that use the Exposure Index (of Black students to white) as the dependent variable. Values of this index can be interpreted as the share of the average Black student's schoolmates who are white in a given district-year.[15] The results are reported separately by region, and have two key features.

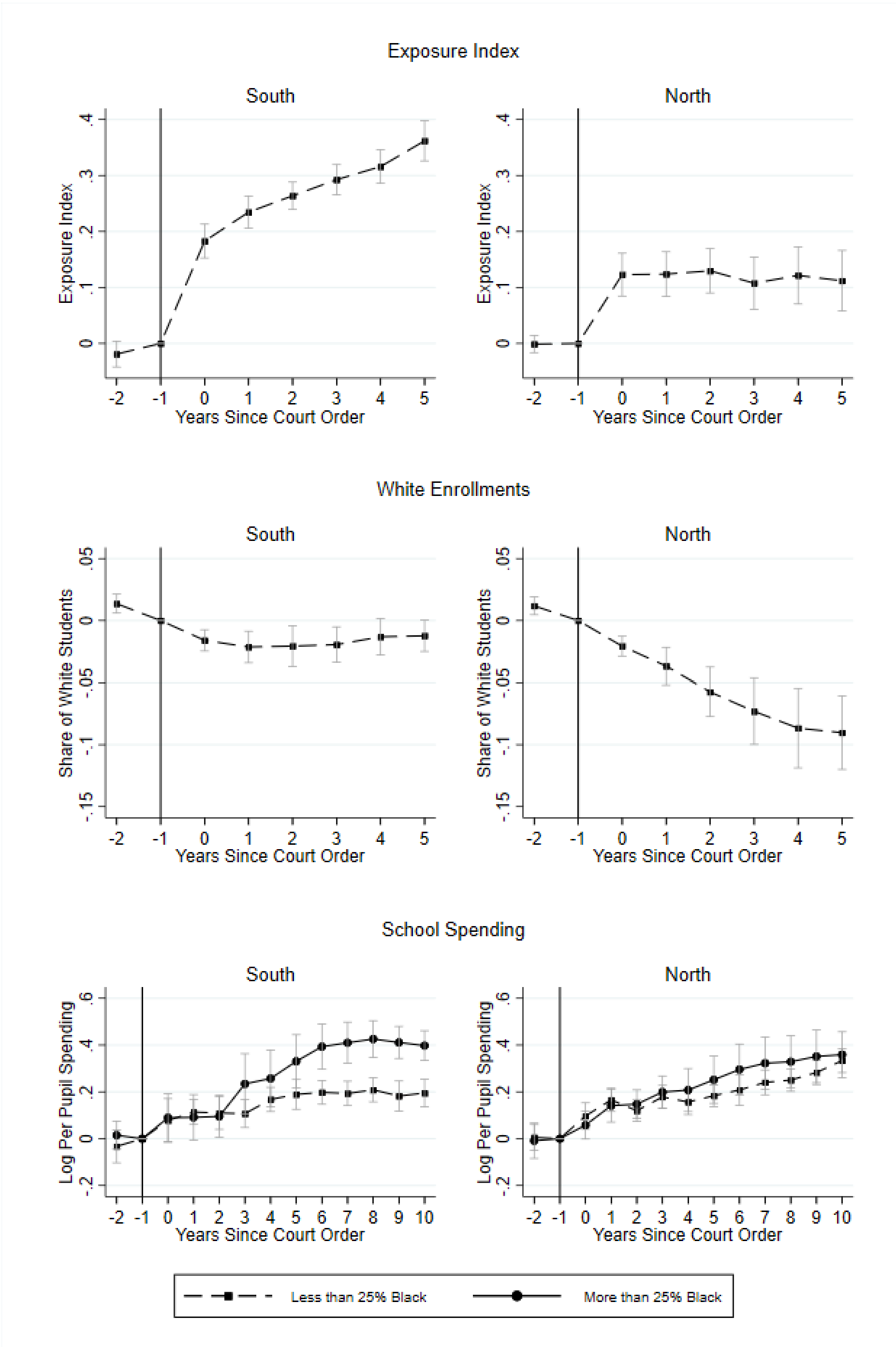
First, the orders we study resulted in qualitatively large increases in the number of white peers that Black students were exposed to in both regions, but these increases were more than 3 times larger in the South than in the North. Specifically, five years after an order, the Exposure Index had risen by approximately .35 in southern districts and by approximately .10 in northern districts. This suggests that larger effects on long-term outcomes could also reasonably be expected within the South.

Second, there are large increases in the exposure of Black students to white immediately in the year of the order, but within the South these are followed by more gradual additional desegregation in subsequent years, and the Exposure Index is still trending positively after five years. This gradual phasing in of integration is likely due to comparatively weak initial integration plans being renegotiated or replaced by more stringent plans, and suggests that we would expect to begin observing treatment effects among individuals who were younger than school-entry age at the time of the order. For instance, two individuals who were respectively ages 0 and 5 at the time of a local desegregation order would both have attended post-order schools for the entirety of their education, but the patterns in Figure 3.2 suggest that the individual who was zero at the time of the order would have attended more thoroughly integrated schools, and may have experienced better long-term outcomes as a result.[16]

[15]See Massey and Denton (1988) and O'Flaherty (2015) for a complete definition and discussion of the Exposure Index. Patterns in the Dissimilarity Index, also widely used to study segregation levels, are reported in the Appendix and are very similar to those in Figure 3.2.

[16]In addition to the phase-in of actual integration, there may be negative effects of being exposed to the very earliest years of court-ordered integration plans, since as noted these plans were extraordinarily controversial and were often fiercely and even violently resisted by local white populations.

Figure 3.2: Changes in District Characteristics after Court Orders, by Region



Notes: Figure plots coefficients from regressing the outcome listed in the subtitle of each panel onto indicators for the number of years relative to a court order (event time) as well as district and year fixed effects. See text for data description and sources. Bands show 90% confidence intervals calculated with standard errors clustered at the school district level. Each district-year is given equal weight.

In contrast to this phasing-in pattern in the South, the Exposure Index for northern districts jumps in the year following an order, but then stagnates or even declines modestly in subsequent years. One explanation for these patterns is that the number of white families withdrawing from districts undergoing desegregation orders was greater in the North, perhaps due to more widely available suburban or private school alternatives. White flight of this kind is of interest in part because it will mechanically reduce the exposure of Black students to white peers (indeed the maximum possible value of the Exposure Index is simply the overall share of white students in a district), and more generally white disenrollment constitutes an important effect of the court orders. (Derenoncourt, 2022; Boustan, 2010)

The middle panel of Figure 3.2 assesses the extent of white flight directly by using the share of each district's student body that was white as the dependent variable. The results indicate that the studied integration orders were followed by white enrollment declines, and that these declines were substantially stronger in the North (approximately 10 percentage points after five years) than in the South (approximately 1 percentage point after five years).[17] This greater degree of white flight in the North would both limit exposure of Black students to white peers in the aggregate, as was observed in the top panel of Figure 3.2, and may generate other types of peer effects as well if white students from more socioeconomically advantaged backgrounds or of higher ability were more likely to exit integrating districts, as seems plausible. Additionally, white disenrollments would effectively limit the extent to which white students themselves were actually exposed to the schools placed under orders, which could attenuate any estimated effects of the orders on their long-term outcomes.

In addition to changing peer composition, previous studies have found that court-ordered integration increased the material resources available at the schools attended by Black students ((Reber, 2010; Johnson, 2011)). In practice it is more difficult to maintain large school

[17]In the relatively brief two year pre-period that we observe, there is a negative pre-trend in white enrollment, which may reflect anticipatory withdrawals of white families after orders were ruled but before they were fully implemented.

quality disparities by race when Black and white students attend the same schools, and furthermore some federal school funding streams were explicitly conditioned on districts complying with court orders and some integration plans explicitly included funding-related clauses.

To help assess the extent to which the court orders in our sample affected school resources, we estimate models that use per-student funding as the dependent variable. Data on per-student funding is drawn from Census Bureau's Annual Survey of State and Local Government Finances, as harmonized by Pierson et al. (2015), which contains data for 105 of the treated districts in our sample and is available in 1967 and then annually from 1970 onward. This data availability range allows us to estimate specifications with event times ranging from -2 to +10 while still retaining a balanced panel of all districts undergoing orders from 1969-1980.

While ideally we would observe the per-pupil funding levels of the specific schools attended by Black and white students, school finance data is only available at the district level. Given this data constraint, we estimate separate specifications for districts where the pre-treatment share of Black students was less than 25% versus greater than 25%, and interpret any relative increases in funding levels at schools with larger Black student populations at baseline as strong suggestive evidence that the orders increased the relative resources available to the average Black student.[18] This interpretation seems reasonable given that the orders themselves decreased racial sorting of students across schools within districts and decreased overall white student shares.

The bottom panel of Figure 3.2 reports the results of these models, which indicate that within the South, per-student funding grew by approximately 20% in the decade following an integration order among districts where less than 25% of the students were Black at baseline,

[18]The median Black share in this sample was 26.9%, so that this threshold divides the districts approximately in half.

but grew by approximately 40% within districts that had a baseline share of Black students greater than or equal to 25%. This faster funding growth in districts with more Black students would lead to significant improvements in the relative financial resources available to African American students in the period following a court order, and are consistent with the findings of Cascio et al. (2010) and Reber (2010). The analogous estimates for the North show that while northern districts did experience increasing per-student funding in the decade following an integration order, these increases did not significantly differ by baseline racial composition, such that there would not have been differential improvements in the financial resources available to northern African American students.

In summary, the results in Figure 3.2 suggest that the court orders we study had a qualitatively large effect on peer racial composition, especially in the South; were followed by large declines in white enrollments, especially in the North; and that southern districts with more Black students experienced large relative increases in per-pupil funding. These basic patterns inform the estimation and interpretation of our main results for long-term outcomes in the next section, and in particular suggest that we may expect stronger effects in the South as well as impacts that phase in across cohorts that are younger than school entry age.

3.5 Empirical Strategy and Results

3.5.1 Empirical Strategy

Our empirical approach relies on county-level differences in the years that school desegregation orders were implemented, which led to individual differences in age at the time that a local order was introduced. Specifically, we estimate the following flexible event study specification:

$$Y_{bcst} = \theta_c + \alpha_{s(c)b} + \gamma_t + \lambda_s + \boldsymbol{X}_c\boldsymbol{\beta} \cdot b + \sum_{\tau=-5}^{24} \delta_\tau \mathbb{1}\{T_c^* - b = \tau\} + \varepsilon_{bcst}, \tag{3.1}$$

where Y_{bcst} represents a mean outcome for individuals from birth cohort b, born in county c, of sex s, and observed in calendar year t. To be as flexible as possible in assessing racial and regional heterogeneity in effects, we estimate this specification separately for whites and Blacks as well as for individuals born in the South versus the North.[19] Fixed effects for county of birth, survey year and sex (θ_c, γ_t, and λ_s, respectively) account for time-invariant differences in outcome variable means across counties, national-level changes affecting all cohorts in a given year, and mean-level sex differences in outcomes. State-of-birth-by-cohort fixed effects, $\alpha_{s(c)b}$, are also included, and account for any general cohort trends that occur at the national or state level, including for example contemporaneous federal or state policy changes or macroeconomic conditions, among many other factors. Interacting the cohort fixed effects with state indicators is more flexible, but also leads our specification to largely rely on differences in the county-level timing of school desegregation orders *within* states. This will exclude variation from states that only had school desegregation orders implemented in one particular year,[20] so as a robustness check we report results that instead use Census-division-by-birth-cohort fixed effects, and do not find any substantive changes in our results.

Following the literature (Bailey and Goodman-Bacon, 2015; Hoynes et al., 2016; Bailey et al., 2021), we also include a vector of 1960 county characteristics $\boldsymbol{X}_c$ and interact them with a linear trend in birth year. The characteristics in this vector include the 1960 poverty rate, log county population, population share over age 65, under age 5, living in an urban setting, and non-white. We also include vote shares for Strom Thurmond in the 1948 presi-

[19]Our baseline results define the South as the eleven states of the former confederacy and the North as all other states, and below we demonstrate robustness to alternative regional definitions.

[20]Among former confederate states, this includes Arkansas and South Carolina.

dential election (who ran on an explicitly segregationist platform) as a proxy for prejudicial attitudes and preferences for racial segregation. While this relatively extensive set of covariates improves precision and makes the identifying assumptions more likely to hold, below we also demonstrate that our key findings are robust to a more parsimonious specification that controls only for county, cohort and sex fixed effects, which provides more transparency in terms of the conditional sources of identifying variation.

The key parameters of interest are the event study coefficients, δ_τ. T_c^* denotes the year that county c experienced a school desegregation court order, and δ_τ are therefore difference-in-differences estimates that track the effect of a school desegregation order being implementation at age τ, relative to ages 17-18, which is the omitted category. We include event-time dummies for treatment from ages -5 to 24, where this range was chosen to be sufficiently wide to observe both a phase-in period of exposure among individuals who were very young (or not yet born) when a local integration order was implemented, as well as the period where individuals had already completed schooling at the time of the order. We group event times into two-year bins to increase precision and decrease the number of disclosed estimates for Census Bureau review, and standard errors are clustered at the county level to account for arbitrary serial correlation of error terms within counties (Bertrand et al., 2004).

We restrict our working sample to be balanced in event time, and because our outcome data contains the 1945-1985 cohorts, constructing a balanced panel requires restricting the analysis to orders occurring from 1969 through 1980. There are 16 court orders that occurred outside of this range, such that imposing balanced event time retains 92% of the districts in our sampling frame.

The key identifying assumption is a parallel trends assumption, which in this application holds that the timing of local court order implementations, and by extension the age of a given individual at the time of an order, was conditionally unrelated to the counterfactual trends in outcomes. This assumption could be violated if, for instance, desegregation orders were first sought and implemented in counties where the potential outcomes of African American children were trending positively even in the absence of school integration orders, or if desegregation orders were implemented concurrently with other county level policies that had positive effects on the long-run trajectories of Black children.

In our view the most compelling evidence on whether this assumption holds is provided by the placebo tests built into our event study design as discussed above, most importantly whether we observe null pretrends beyond age 17 - that is, insignificant estimates of $\delta_\tau,\ \tau \in [18, 24]$ - and secondarily whether we do not observe large effect estimates among whites, who were not the primary target of the desegregation orders we study.

To provide additional evidence on the nature of the timing of desegregation orders, Table 3.2 reports the results of regressing the year that an integration order was implemented in each county in our working sample onto the 1960 county characteristics enumerated above. Table 3.2 first reports estimates of the association between the year of a court order and each of these county characteristics individually, and then in the final column includes all of these county characteristics simultaneously. Because desegregation order timing varied strongly by region (see Figure 3.1) we condition on a South dummy in each specification, but include no other controls.

Table 3.2: Correlations Between County Characteristics and Treatment Timing

Variable	Year of Order	Year of Order	Year of Order	Year of Order	Year of Order	Year of Order	Year of Order	Year of Order	Year of Order
South	-4.119***	-4.217***	-3.647***	-4.031***	-3.919***	-3.552***	-4.215***	-4.273***	-3.757***
	(0.515)	(0.489)	(0.504)	(0.612)	(0.643)	(0.634)	(0.500)	(0.492)	(0.758)
Total Population (10,000s)	0.001								-0.001
	(0.004)								(0.004)
1960 % Under 4		0.066							0.242
		(0.156)							(0.179)
1960 % Urban			0.034***						0.052**
			(0.013)						(0.022)
1960 % Nonwhite				-0.010					-0.029
				(0.027)					(0.035)
1959 Median Family Income					0.000				-0.001
					(0.000)				(0.001)
1959 % with Family Income Under $3k						-0.044			-0.046
						(0.030)			(0.128)
1960 Median Education							-0.061		0.638
							(0.201)		(0.921)
1960 % with 12+ Years of Education								-0.023	-0.158
								(0.029)	(0.134)
Observations	134	134	134	134	134	134	134	134	134

Notes: Table presents OLS estimates of regressing the year of a court order in each county onto 1960 county characteristics. See text for details on determination of court order timings. Robust standard errors reported in parentheses. $*$, $**$, and $***$ indicate statistical significance at the 10%, 5%, and 1% level, respectively.

The estimates in Table 3.2 indicate that more urban counties were placed under desegregation orders somewhat later than more rural counties, which is consistent with the findings of (Cascio et al., 2008), but that no other pre-treatment county characteristics had statistically or economically significant associations with treatment timing, conditional on region. On balance we feel that the patterns shown in Table 3.2 are consistent with the validity of our key identifying assumption that *timing* of orders across counties was quasi-random.

Finally, we note that previous studies on this topic have argued that the NAACP Legal Defense Fund, the main organization litigating school integration cases in this period, strategically brought suits in localities where they believed that they had the highest probability of winning, and therefore establishing favorable legal precedents, rather than localities where the expected benefits of integration would be greatest (Guryan, 2004; Johnson, 2011). This legal strategy decreases the likelihood that the timing of integration cases was partially a function of time varying determinants of African American children's long-term outcomes, which would lend further support to the parallel trends assumption.

3.5.2 Baseline Findings

In Figure 3.3 we report the results of estimating Equation 3.1 using the human capital and economic self-sufficiency indices as the dependent variables. Results are shown separately for Blacks and whites, and for counties in and out of the South.

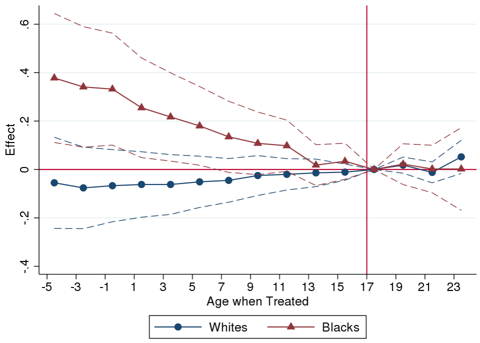
Figure 3.3a indicates that among southern African Americans, shown with red triangles, earlier exposure to school integration orders had large positive effects on human capital outcomes. Specifically, being born five years prior to an order is estimated to increase the human capital index by almost .4 standard deviations relative to being age 17 at the time of the order. This effect declines modestly between ages -5 and 0 at the time of the order, and then shows a more rapid monotonic decline as age at the time of a local order increases from

0 to 17. Figure 3.3b reports results for the economic self-sufficiency index within the South, and shows slightly larger treatment effects among Blacks but very similar overall patterns. The fact that effects begin phasing in before age five, the typical age for school entry, is likely attributable to the court orders themselves often taking five or more years to be fully implemented, as documented above.

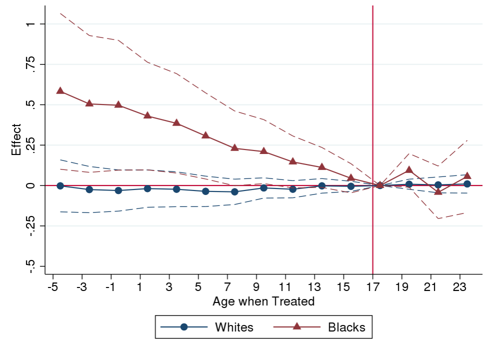
Critically from an identification perspective, neither Figure 3.3a nor Figure 3.3b indicate any *additional* declines in outcomes across African Americans who were ages 17-24 at the time of a local order. This is analogous to a flat 'pre-trend' in a typical event study design, and is reassuring given that individuals in this range did not have differential exposure to the court orders.

Figure 3.3a and Figure 3.3b also find no economically or statistically significant effects for either outcome among southern whites, shown with blue squares. This is itself an important finding, as it suggests that gains among southern Blacks did not come at the expense of local white students. The fact that white enrollment declines following court orders in the South were generally modest, as shown in Figure 3.2, suggests that the null effects among southern whites was not simply because they avoided attending schools undergoing court orders in large numbers. Additionally, from an identification perspective, we interpret the paucity of significant effects among whites as reassuring evidence that the estimates for southern Blacks do not reflect some unobserved factor that affected all individuals from a particular county and birth cohort. In other words, any unobserved time varying factor that may be biasing our main estimates would need to be race-specific, which narrows that scope of potential confounders.

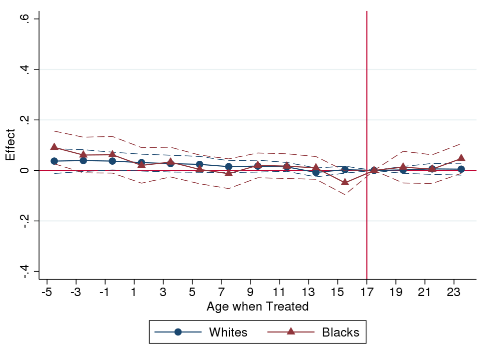
Figure 3.3: Baseline Results



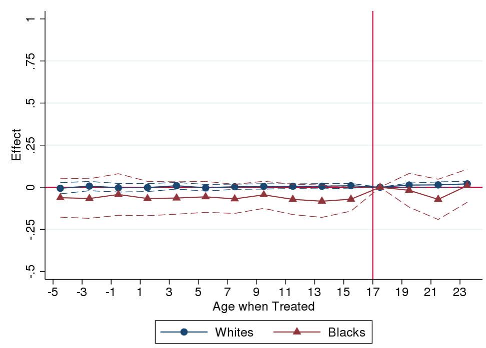
(a) HC Index, South



(b) ESS Index, South



(c) HC Index, North



(d) ESS Index, North

Notes: Figures report estimates of δ_τ from Equation 3.1. Sample is balanced in event time. Dotted lines indicate 95% confidence intervals, constructed using standard errors clustered at the county level. The sample contains 5.1 million observations collapsed to the county, survey year, cohort, race, and sex level. Weights equal to the sum of individual survey weights in each cell are applied. Controls include 1960 county characteristics interacted with linear cohort trends, as well as fixed effects for county, survey year, sex, and birth cohort-by-state of birth.

Figures 3.3c and 3.3d report analogous results for the North, and in strong contrast to the patterns in the South, no significant exposure effects are observed for either outcome index or racial group. This is perhaps especially notable given that many of the most politically contentious integration fights occurred in large northern cities like Boston, Detroit and Chicago. The extent to which the benefits of court-ordered integration appear to have been concentrated in the South is essential for a full understanding of court-ordered desegregation's effects, and we note that this regional heterogeneity would have been difficult to detect without the very large sample sizes provided by our data. There are a number of plausible explanations for these strong regional differences, and we investigate these in detail and provide additional relevant empirical findings in Section 3.6.

3.5.3 Estimates for Index Subcomponents

To explore the specific outcomes driving the baseline results in Figure 3.3, we next report results for each subcomponent of the human capital and economic self-sufficiency indices. The analyses in this section focus on southern Blacks, since the baseline results identify this as the population that was substantively impacted by court-ordered desegregation. The unreported analogous estimates for northern Blacks and for whites in all regions are uniformly small and statistically insignificant.

Figure 3.4 reports results for the human capital outcomes. The strongest effects are for high school attainment and years of schooling: full exposure to court-ordered integration is estimated to increase the probability of high school completion by approximately 15 percentage points, and to increase total years of schooling by approximately one full year. These effects are broadly consistent with those obtained by Johnson (2011), who estimates approximately a 25 percentage point improvement in high school graduation and a one year increase in educational attainment.[21] Likewise Guryan (2004) estimates that contemporaneous Black

[21]There are several factors that make our estimates not directly comparable to Johnson (2011). For

high school dropout rates declined by 3.8 percentage points over the ten years between the 1970 and 1980 Censuses within districts that implemented school desegregation orders over this span. This magnitude is highly consistent with our estimates of the first 10 years of exposure (that is, treatment at ages 7 to 17) in Figure 3.4a.

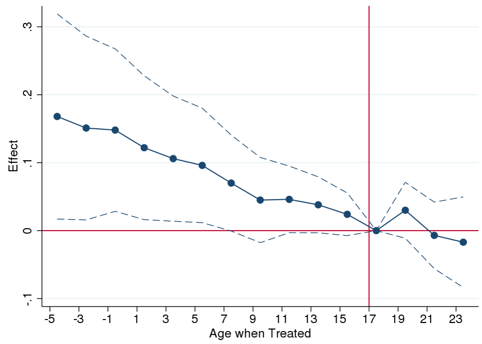
The other results in Figure 3.4 indicate that these increases in high school completion did not consistently translate into higher rates of post-secondary schooling, with positive but imprecise estimates for attaining "some college" and no clear relationship between exposure to the desegregation orders and Black attainment of four-year or advanced degrees.

Figure 3.5 reports results for the subcomponents of the economic self-sufficiency index, primarily labor market outcomes, and finds more even results over the subcomponents. Earlier exposure to desegregation orders is associated with stronger labor force attachment, whether measured as employment, hours worked, or weeks worked. Individuals treated at earlier ages also experience lower rates of poverty and public assistance receipt, as well as approximately 30% higher annual wage earnings.[22]

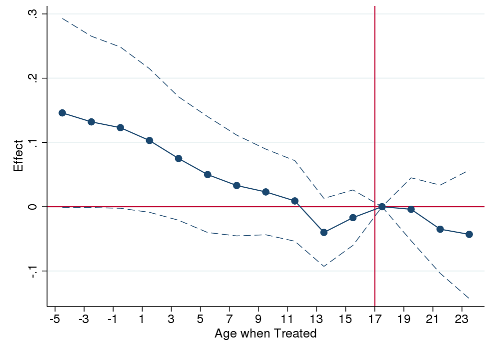
instance Figure 3.4 uses a southern sample while Johnson (2011) uses a national sample, and our samples differ meaningfully in salient characteristics, with for instance a high school completion rate in the relevant PSID sample of 77% vs. 88% in our working sample, which may be attributable to the relatively small sample sizes or over-representation of low income African Americans in the PSID.

[22]In the online appendix, we show results for additional earnings measures, including the log of total earnings +1 and total earnings measured in levels, which both retain zero earners in the sample, as well as for hourly wage measures. These analyses continue to find strong effects on total earnings when zero earners are included, but less clear impacts on hourly wages, suggesting that the effects of school integration on earnings worked primarily though extensive margins.

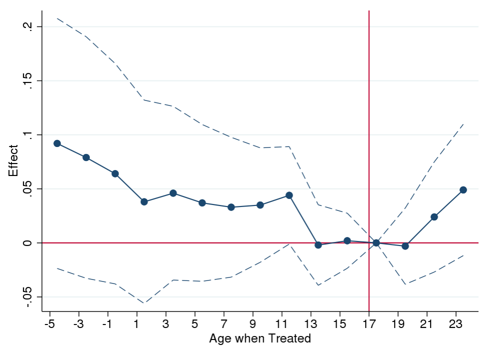
Figure 3.4: Results for HC Index Components



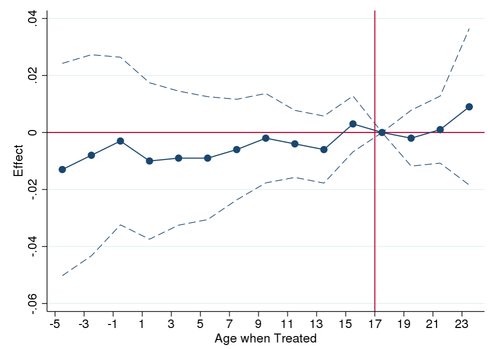
(a) High School Degree



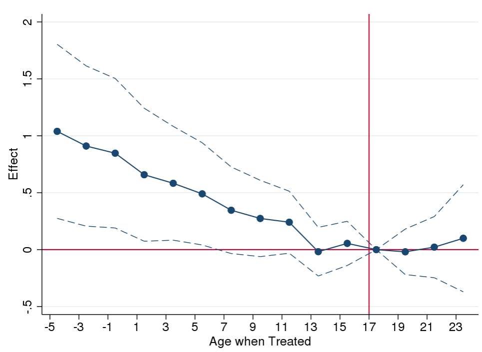
(b) Some College



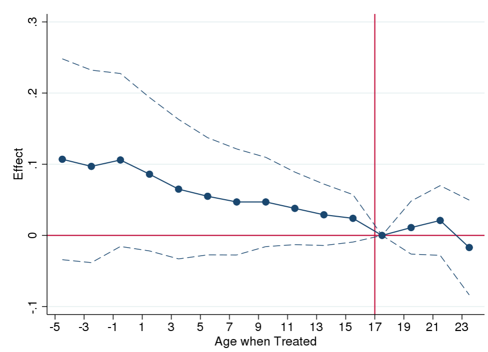
(c) Four-Year Degree



(d) Advanced Degree



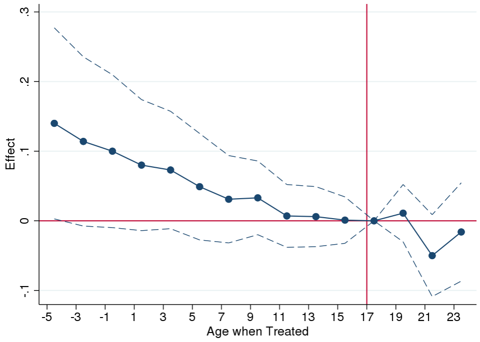
(e) Years of Schooling



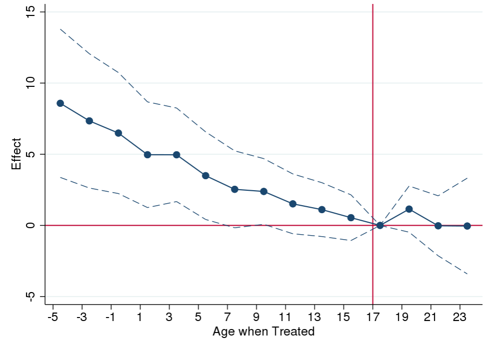
(f) Professional Occupation

Notes: Figures report estimates of δ_τ from Equation 3.1 with the indicated dependent variable and only within the southern African American sample. Sample is balanced in event time. Dotted lines indicate 95% confidence intervals, constructed using standard errors clustered at the county level. The full sample contains 5.1 million observations collapsed to the county, survey year, cohort, race, and sex level. Weights equal to the sum of individual survey weights in each cell are applied. Controls include 1960 county characteristics interacted with linear cohort trends, as well as fixed effects for county, survey year, sex, and birth cohort-by-state of birth.

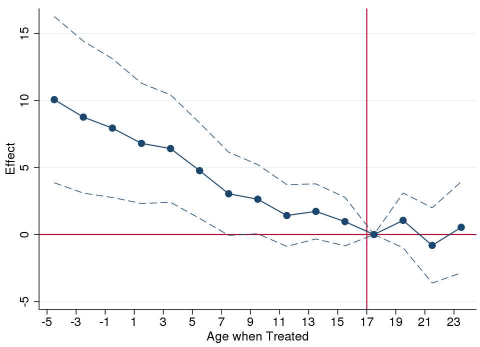
Figure 3.5: Results for ESS Index Components



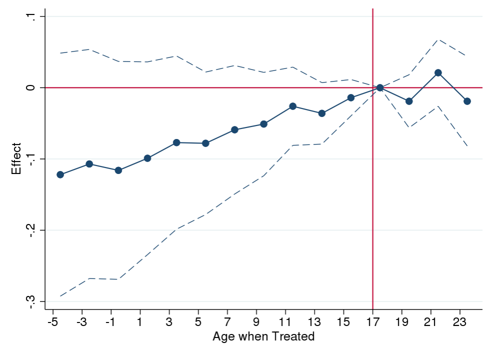
(a) Employed



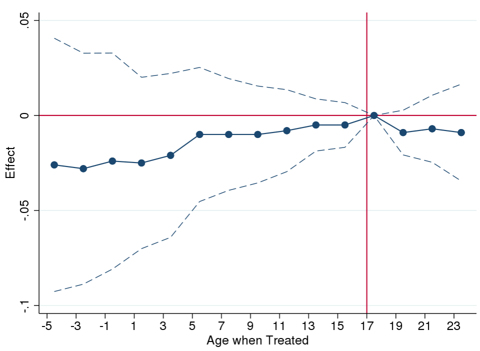
(b) Weekly Hours



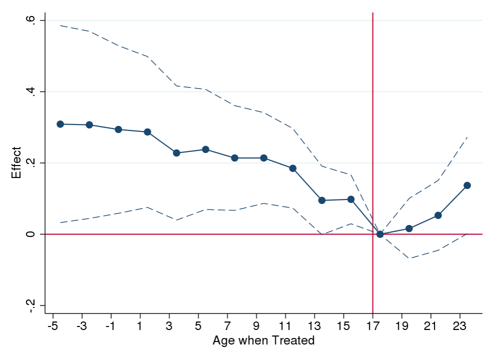
(c) Weeks Worked



(d) In Poverty



(e) Received Public Assistance Income



(f) Log Wage Earnings

Notes: Figures report estimates of δ_τ from Equation 3.1 with the indicated dependent variable and only within the southern African American sample. Sample is balanced in event time. Dotted lines indicate 95% confidence intervals, constructed using standard errors clustered at the county level. The full sample contains 5.1 million observations collapsed to the county, survey year, cohort, race, and sex level. Weights equal to the sum of individual survey weights in each cell are applied. Controls include 1960 county characteristics interacted with linear cohort trends, as well as fixed effects for county, survey year, sex, and birth cohort-by-state of birth.

3.5.4 Robustness

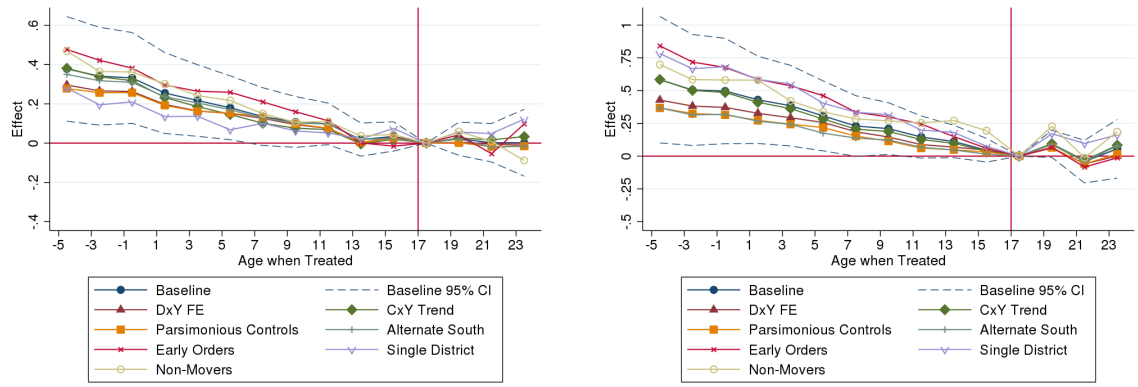
Figure 3.6 assesses the sensitivity of our main results to a number of alternate specifications, again focusing on the southern African American sample.

We first assess the robustness of our findings to a variety of basic modeling and specification choices: we replace our state-by-cohort fixed effects with division-by-cohort fixed effects, which allows us to use variation from states that only had counties treated in one particular year;[23] we replace the interacted 1960s county characteristics with county-specific linear cohort trends to account for potential pre-treatment county characteristics that we do not control for in our baseline specification; we estimate a model with minimal controls that includes only sex and county and cohort fixed effects; we restrict the sample to court orders occurring in 1969-1971, since these early orders may have differed in content or implementation compared to later orders; and we include all states belonging to the Southern Census region instead of former confederate states (thus adding Oklahoma, Arkansas, West Virginia, Kentucky, Maryland, and Delaware). Figure 3.6 shows that in all of these cases and for both outcomes indices, the estimates of δ_τ are substantively and statistically indistinguishable from the results of our preferred baseline specification.

Figure 3.6 also reports the results of two checks that address the potential for mismeasurement of exposure to court-ordered integration. First, we report results that use the sub-sample of respondents from locations where there is a single county-operated school district that covers the entire county, dropping counties that contain multiple districts, most commonly one or more city-operated districts. Because substantial numbers of school districts are operated at the municipal level or as independent local government agencies, but we are only able to match individuals to their counties of birth, exposure is potentially measured more accurately for individuals from locations with a single county-operated school

[23] In the south, this adds Arkansas and South Carolina.

Figure 3.6: Robustness



(a) HC Index (b) ESS Index

Notes: Figures present estimates of δ_τ from Equation 3.1 for southern Blacks. Relative to the baseline specification, DxY FE replaces the state-by-birth year fixed effects with division-by-birth year fixed effects; CxY Trend replaces the interactions between 1960 county characteristics and survey year with a county-specific linear survey year trend; Parsimonious Controls conditions only on sex, county and cohort fixed effects; Alternate South defines the South as the states in the southern Census Region rather than the states of former Confederacy; Early Orders limits to counties with orders implemented in 1969-1971; Single District restricts to counties containing a single school district; and Non-Movers excludes observations who were not residing in their county of birth at the time they were surveyed.

system.[24] Second, because exposure would be mismeasured for individuals who moved away from their county of birth prior to completing schooling, we report results that use the sub-sample of individuals who were still residing in their county of birth when they were enumerated in the Census or ACS. While it is possible for individuals to leave their county of birth during school aged years but then return there as an adult and be enumerated in the Census or ACS, and treatment itself may affect migration propensities, patterns within this sub-sample are at a minimum less likely to be biased by mismeasured exposure due to childhood migration.

[24]We note that county operated districts may be less subject to white flight, tend to be more rural, and likely vary along other dimensions as well, so that it is not possible to reliably distinguish the importance of these other factors from improved exposure measurement.

The relevant results in Figure 3.6 show that for both outcome indices, the magnitudes of the treatment effects are moderately stronger within the sub-sample of non-movers, consistent with exposure being better measured in this population. Restricting the analysis to full-county districts makes the magnitudes of the effects for the ESS index substantially stronger, while the magnitudes of the HC index estimates fall modestly, which is somewhat unexpected, but the estimates in this sample remain economically large and statistically significant. On balance, these patterns suggest that our baseline findings are if anything attenuated by childhood migration or by inaccurately matching individuals to municipal versus county-operated school districts, and these measurement issues are very unlikely to be driving our key positive findings.

Next, we address the potential estimation issues that can occur in two-way fixed effects specifications like Equation 1 when treatment turns on at different times across units, as highlighted in a recent methodological literature (de Chaisemartin and D'Haultfœuille, 2020; Sun and Abraham, 2021; Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021).

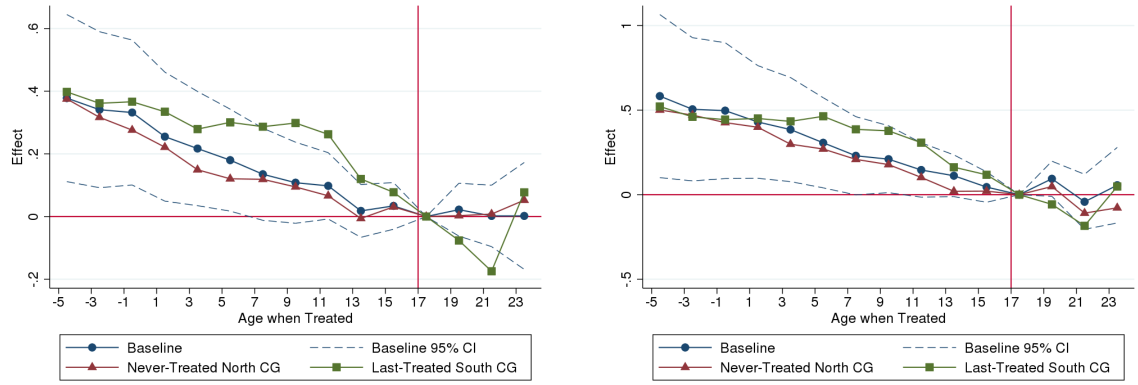
The key conceptual issue raised in this literature is that estimates from typical two-way fixed effects specifications are partially based on comparisons of units that were treated in later periods to units that were already treated in earlier periods. Such comparisons between sets of units that were both already treated are often problematic, since trends in early-treated units are potentially impacted by treatment itself and do not provide a valid counterfactual for later-treated units. These issues are particularly acute when treatment effects are heterogeneous across units or with time since treatment, and in many applications can lead to severe biases including reversing the sign of difference-in-difference estimates (Goodman-Bacon, 2021) or producing wholly spurious apparent pre-trends in event studies (Sun and Abraham, 2021).

Several approaches to resolving this issue have been proposed in the literature, typically through some type of re-weighting. These alternative estimators all have the common feature of focusing on "good comparisons" between the outcomes of already-treated observations and the outcomes of not-yet-treated or never-treated observations, and not on "bad comparisons" that use already-treated observations as controls. We specifically implement the estimator recommended by Sun and Abraham (2021), which is well-suited for our setting because it is specifically designed for event study research designs, rather than two way fixed effects estimators with a binary treatment. Sun and Abraham (2021) propose that researchers first estimate separate event-study coefficients for the sets of units that were treated in each period, which the authors refer to as "treatment cohorts." In the current application, the treatment cohorts are simply the sets of counties that were placed under a desegregation order in *1969, 1970,...,1980*. By construction, each set of cohort-specific event time coefficients do not rely on any "bad comparisons" of later-treated units to earlier-treated units, since they each include only one treatment cohort. The Sun and Abraham (2021) procedure then simply takes the weighted average of each set of cohort-specific event time coefficients, using the share of treated units from each treatment cohort as weights.

Estimating separate event time coefficients for each treatment cohort of course requires a set of control units, and the Sun and Abraham (2021) procedure uses as controls either never-treated units or units from the last-treated cohort. In the current application, neither type of control group is available in an ideal form: all of the untreated counties were from outside of the South, and within the South there were very few late-treated counties to use as controls (see Figure 3.1).

Given these features of the treatment patterns in our data, we estimate two separate versions of the Sun and Abraham (2021) estimator. The first uses northern counties that did not undergo court orders during the study period as never-treated controls. The second uses late-treated southern counties as controls, specifically the three southern counties undergoing orders in 1978-1980. The former set of controls is larger but comes from another region with potentially different counterfactual outcome trends, while the latter set of controls is from

Figure 3.7: Results using Sun and Abraham (2021) Estimator



(a) HC Index

(b) ESS Index

Notes: Figures present estimates of δ_τ from Equation 3.1 for Southern Blacks while implementing the Sun and Abraham (2021) estimator using either never-treated northern counties or late-treated southern counties as control groups, as indicated. See text and Sun and Abraham (2021) for more detailed descriptions of the estimator.

the South but is small and potentially idiosyncratic.[25]

Figure 3.7 reports the results using the HC and ESS indices as outcome measures, as well as the baseline estimates for reference. Reassuringly, for both outcomes and both sets of control units the alternative estimates are quite similar to the baseline results, although the coefficient estimates for the human capital index when using late-treated southern counties as controls become somewhat erratic. As noted, neither set of control units is ideal, but taken together we do believe that the Sun and Abraham (2021) estimator using these two sets of control counties provides reasonably strong evidence that our main findings are not driven by comparisons that use early-treated counties as controls for later-treated counties.

[25]The last-treated southern units were Lubbock County, TX; Dougherty County GA, which contains Albany; and Travis County, TX which contains Austin.

3.6 Discussion

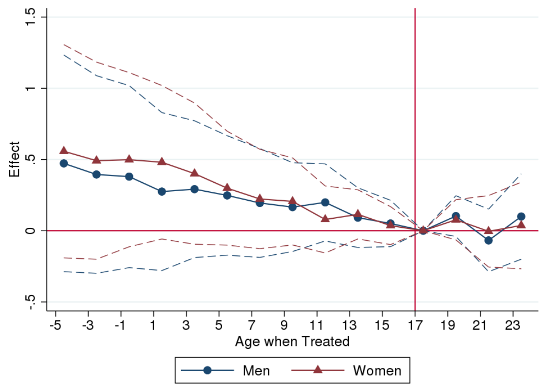
3.6.1 Heterogeneity

One of the key advantages of the large sample sizes available in our data is that we are able to generate reasonably precise estimates within sub-populations of potential interest. We have already shown that disaggregating by region and race is critical for a complete understanding of court-ordered integration's long-term impacts, and in this section we extend our analysis of heterogeneity to several additional dimensions. The results in this section focus on the southern Black sample, where our baseline estimates were concentrated, but in the next section we also use several of the heterogeneity patterns we find for the South to better understand the paucity of effects in the North.

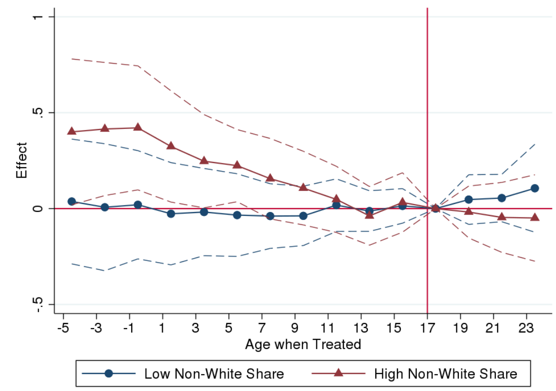
We estimate models that allow the effects of court orders to vary with respect to five characteristics: (1) sex; (2) the pre-treatment white share in each county; (3) the level of *residential* segregation in the county containing each district; (4) the total number of school districts in the county containing each district; and (5) the share of the vote received by Strom Thurmond in the 1948 presidential election.

The results are reported in Figure 3.8 for the Human Capital Index and in Figure 3.9 for the Economic Self Sufficiency Index, and discussed in turn below. Since we are using simple cross sectional variation in these characteristics, we consider these estimates of how treatment effects varied across different types of counties to be descriptive, but still believe that systematic descriptive evidence on which types of counties experienced larger treatment effects provides useful information.

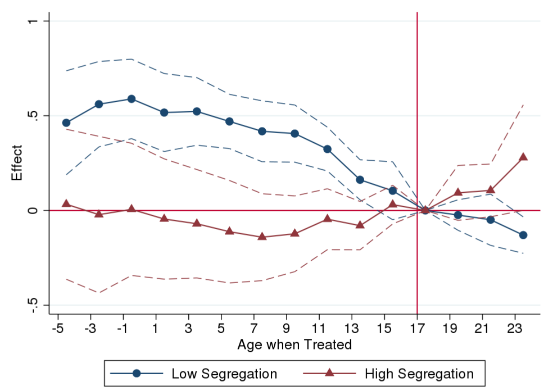
Figure 3.8: Heterogeneity, HC Index



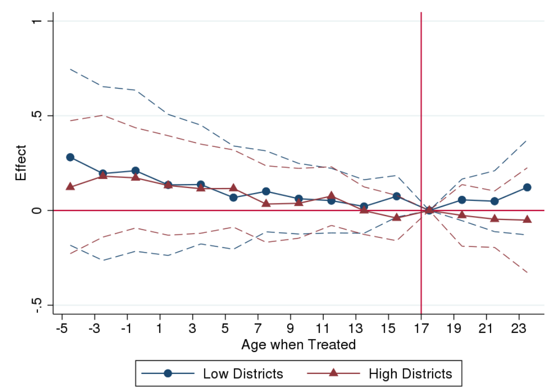
(a) Sex



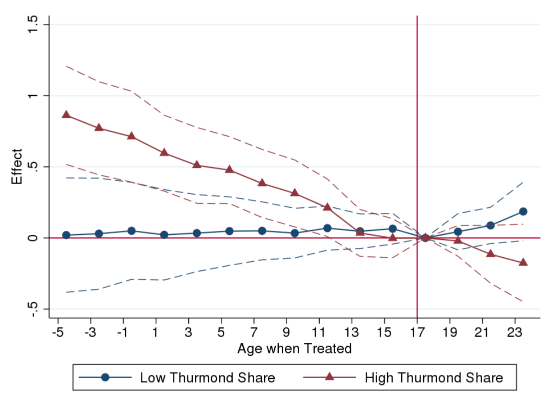
(b) Non-White Share



(c) Residential Segregation



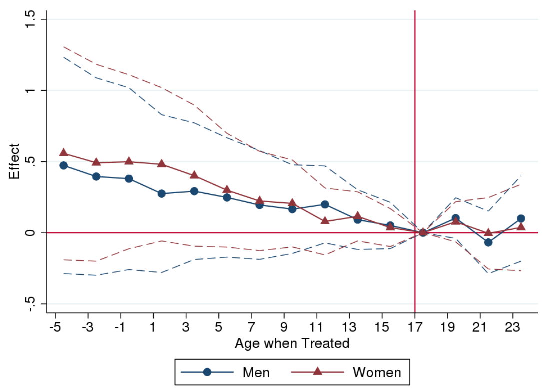
(d) Number of School Districts



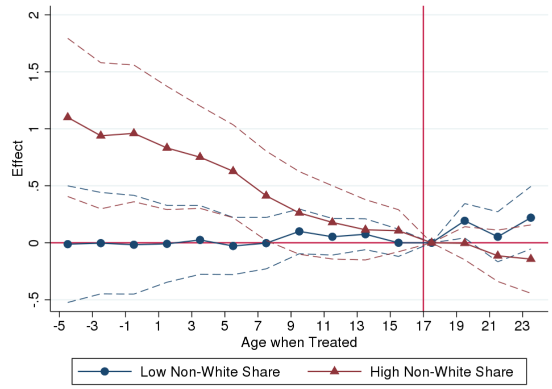
(e) Thurmond Vote Share

Notes: Figures report estimates of δ_τ from Equation 3.1 estimated within the indicated sub-populations, which are described in detail in the text. All regressions restricted to African Americans in the South.

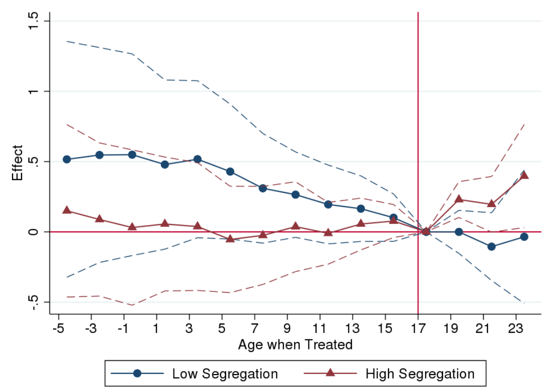
Figure 3.9: Heterogeneity, ESS Index



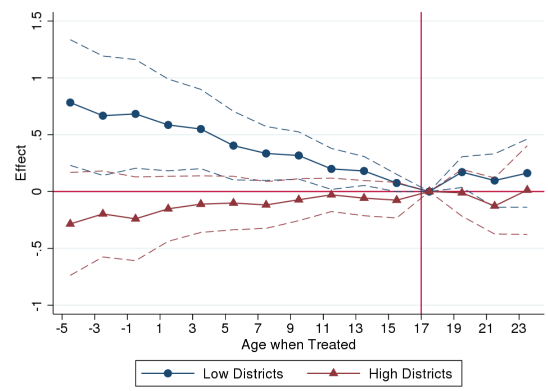
(a) Sex



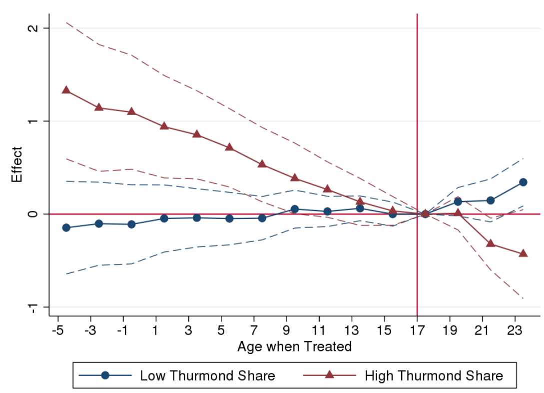
(b) Non-White Share



(c) Residential Segregation



(d) Number of School Districts



(e) Thurmond Vote Share

Notes: Figures report estimates of δ_τ from Equation 3.1 estimated within the indicated sub-populations, which are described in detail in the text. All regressions restricted to African Americans in the South.

Beginning with sex, studies of educational programs frequently find differential effects for boys and girls. The first entry of Figures 3.8 and 3.9 show the results of estimating our baseline specification separately for men and women. Perhaps surprisingly, heterogeneity by sex appears quite limited, with Southern African American men and women benefiting comparably from exposure to school desegregation orders with respect to both human capital acquisition and economic self sufficiency.

We next turn to heterogeneity by the non-white population share in each county. Changes in exposure to white peers after integration will be mechanically larger in settings with more whites, simply because there are more white students to be exposed to. But as emphasized by Reber (2010), districts with greater non-white enrollment shares usually experienced larger reductions in school resource gaps, since these gaps were historically larger in districts with more Black students but were then "leveled up" after integration. Since these factors likely push in opposite directions, the net impact of baseline racial composition on treatment effect magnitudes is ambiguous.

The second entry of Figures 3.8 and 3.9 report estimates of our baseline specification separately for counties with 1960 non-white population shares above and below the median. For both outcome indices, the effect size magnitudes are much larger in less white counties. This is consistent with the results in Reber (2010), who also finds stronger positive effects from integration in Louisiana districts with more Black students. One interpretation of these patterns, following the reasoning of Reber (2010), is that changes in school resources dominate changes in peer effects in this setting, although other reasonable interpretations exist as well.

We next estimate models separately by the level of racial residential segregation, measured using county-level Dissimilarity Indices from the 1960 Census constructed by Cutler et al. (1999).[26] School desegregation orders that redrew enrollment zones, which was a very

[26]As constructed here, the Dissimilarity Index measures the share of Black (or white) families that would

common approach, may have been more effective at influencing peer composition in counties where Black and white families lived closer to one another, effectively leading to a stronger "first stage" effect on the exposure of Black students to white. Alternatively, less residential segregation may allow integration orders to accomplish similar changes in peer composition while busing students shorter distances, reducing the time and resources that needed to be devoted to student transportation. The third entry of Figures 3.8 and 3.9 finds that our estimates are indeed stronger in counties with less residential segregation. This is in line with expectations, and suggests the presence of important interactions between segregation in schools and neighborhoods.

Another potentially important dimension of heterogeneity is the total number of school districts in each county. This can be viewed as a proxy for the number of potentially competing districts, and more competing districts may in turn lead to greater white flight and reduce changes in the exposure of Black students to white peers. Greater numbers of potentially competing proximate districts may also limit how ambitious the scope of the implemented orders were.[27] The fourth entry of Figures 3.8 and 3.9 show that the estimated impacts on the ESS Index were concentrated almost wholly among orders occurring in counties with relatively few potentially competing districts, while the effect sizes for the HC index are comparable across the two sets of counties. The ESS Index results are consistent with court orders having less impact in settings where a greater number of alternative school districts are accessible, although the lack of heterogeneous treatment effects for human capital outcomes prevent any broad conclusions along these lines.

The final entry of Figures 3.8 and 3.9 report results split by the share of the vote received

need to move across Census Tracts in order for each tract to have the same racial composition as the overall county.

[27] A critical 1974 Supreme Court ruling originating in the Detroit metro, *Milliken v. Bradley*, effectively prohibited desegregation plans from spanning multiple school districts within a metro area. In undisclosed results we have estimated models that use the total number of school districts in each district's commuting zone, rather than county, and observe qualitatively similar patterns.

by Strom Thurmond in the 1948 presidential election. Since Thurmond ran on an explicitly segregationist platform, we view his vote share as a proxy for the local level of overtly segregationist preferences and overall racial animus. While plausible arguments can be made for such attitudes causing the impacts of court-ordered integration to be both weaker and stronger, we believe that on net the orders would likely have had stronger effects in areas with higher levels of racial animus and segregationist preferences, since the system of unequal education that the orders were attempting to rectify was likely stronger is such areas. Consistent with this reasoning, the results in Figures 3.8 and 3.9 show that the effects of court-ordered integration were indeed strongly concentrated in counties with above-median Thurmond vote shares.

We again note that the heterogeneous patterns in Figures 3.8 and 3.9 are descriptive in nature: many of the the studied characteristics are strongly correlated with each other, and are likely correlated with various potentially relevant unobserved factors as well. But the strong differences in effect size magnitudes across different types of counties still provide useful suggestive evidence on where and why the orders were most effective.

3.6.2 Potential Explanations for Regional Differences

A key feature of our baseline results was that court-ordered integration only had significant effects on long-term outcomes within the South. A partial explanation for these strong regional differences was evident in the first stage results reported in Section 3.4, which showed that the effects of the studied orders on school resources and on Black students' exposure to white peers were strongly concentrated in the South, mirroring the estimates for long-run outcomes. Closely related, Section 3.4 also showed that the extent of white flight was greater in the North, which could have attenuated potentially beneficial peer effects, especially if withdrawing white students were positively selected with respect to academic

Table 3.3: Mean Levels of County Characteristics, by Region

Variable	North	South
Strom Thurmond Vote Share	-	26.92
Dissimilarity Index	0.82	0.74
Percent Non-White (1960)	9.95	22.60
Number of Districts	27.24	6.29

Notes: Table displays average values of county characteristics, separated by counties in former confederate states (South) and not (North). See text for details on sample construction and data sources.

achievement.

We also believe that the observed regional heterogeneity is generally consistent with previous studies. For instance Angrist and Lang (2004) and Angrist et al. (2022) both find little or no impact of integration plans on the test scores of minority students in New York and Boston, while Billings et al. (2014), Reber (2010), and Tuttle (2019) all find positive impacts of integration for Black students in southern settings (North Carolina, Louisiana, and Kentucky, respectively). The two previous national evaluations of court-ordered integration, Guryan (2004) and Johnson (2011), did not estimate separate models by region, such that the positive national estimates from these studies may be driven by southern orders, although we note that in Online Appendix Table G3 of Johnson (2011) there are models that interact exposure to court orders with a South indicator, and this interaction is not statistically significant. This result is in some tension with our finding of strong regional heterogeneity, although the standard error on this interaction term is relatively large, such that non-negligible regional differences cannot be ruled out, and it is also potentially relevant that we estimate separate models by region rather than interacting treatment with a South dummy, since the latter approach restricts the effects of covariates, including year effects, to be equal across regions.

Although stronger effects of integration orders on long-term outcomes in the South is *con-*

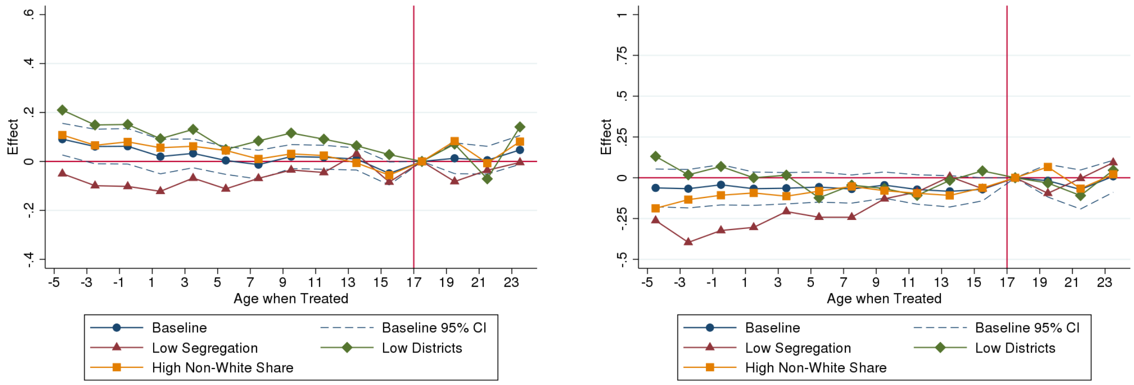
sistent with the first stage estimates and with the previous literature, neither of these points directly provides an actual *explanation* for these differences. An opportunity to evaluate some potential explanations for the observed regional differences in an empirically grounded fashion is provided by the heterogeneity results for the South reported in the previous subsection. In particular, we documented that the effects within the South were often concentrated in counties with a larger share of non-whites, with less residential segregation, with fewer potentially competing school districts, and with a larger Strom Thurmond vote share. To the extent that the North differed with respect to these characteristics, that may help explain the paucity of observed effects there.

Table 3.3 reports the average levels of each characteristic by region, and shows that the North did indeed differ from the South along several of these potentially relevant dimensions. Specifically Table 3.3 shows that average non-white population shares were higher in the South than the North (22.60% versus 9.95%), that the southern counties in our sample were less residentially segregated than the northern counties (with average Dissimilarity Indices of .74 and .82, respectively) and also that the number of potentially competing districts in the same county were much larger in the North (an average of 27 districts in the North and only 6 districts in the South).[28]

To investigate whether these regional differences in contextual characteristics help to explain the lack of significant estimated program impacts in the North, Figure 3.10 reports results for subsets of northern districts with relative high (above median) values of the characteristics that the previous section found were positively correlated with the efficacy of desegregation orders. The figure shows that we continue to observe almost uniformly null estimates across all of the subsets of northern districts where the heterogeneity results for the South suggested there may be larger impacts from court-ordered integration. A partial

[28]The differences in residential segregation by region may themselves reflect the history of de-jure school segregation in the South, since southern white families could be assured of all-white schools regardless of their residential location choices.

Figure 3.10: Treatment Effects in Potentially High Impact Northern Counties



(a) HC Index (b) ESS Index

Notes: Figures report estimates of δ_τ from Equation 3.1 estimated within the indicated sub-populations, which are described in detail in the text. All regressions restricted to African Americans in the North.

exception is that we observe significant positive impacts for the HC index among northern Blacks from counties with relatively few competing school districts, although null effects are still observed for the ESS index within this set of northern counties.

One characteristics that we are not able to investigate in Figure 3.10 is the 1948 Strom Thurmond vote share, since with few exceptions Thurmond did not appear on the ballot in northern states. While on the one hand this can be viewed simply as a data limitation, it can alternatively be viewed as conveying relevant information.

In particular Thurmond's omission from northern ballots, as the Democratic Party nominee or even as a third party candidate, broadly reflected the lack of support for overt racial segregation in the North. This follows from the basic observation that despite strong defacto segregation and widespread racial animus, northern school districts simply did not share the South's history of overt state-sponsored racial discrimination in education, such that court-ordered integration was a less direct challenge to the status quo of northern education systems. This suggests that the small or non-existent impacts of the orders on long-run outcomes in the North may be because the substantive content of court-ordered integration as a policy was qualitatively different in a region where the education system lacked a history of institutionalized state sponsored racial discrimination. This would be broadly consistent with the finding in Figures 3.8 and 3.9 that there were much smaller effects from integration orders in southern settings with low levels of support for Thurmond.

3.7 Conclusion

Virtually all large US school districts were compelled by the judicial system to increase racial integration in the 1970s and 1980s, and this era of court-ordered desegregation was a far more wide-reaching and proactive attempt to equate educational access across racial groups than any set of educational policies implemented before or after. With the benefit of newly available Census and ACS data linked to childhood geographic locations via Social Security Administration records, this paper has provided what we believe to be the most authoritative national evidence to date on whether court-ordered integration had positive long-term impacts on the educational and labor market outcomes of the minority students it was designed to benefit.

Our primary finding was that court-ordered integration did indeed have positive impacts for southern African American students, and that these effects were qualitatively quite large. For instance full exposure was estimated to have increased high school graduation rates by approximately 15 percentage points, increased employment rates by approximately 10 percentage points, and increased annual wages by approximately 30%. However, these impacts apply only to southern school districts, and effects for African Americans in the North were indistinguishable from zero.

In terms of establishing the historical record on the efficacy of court-ordered integration, our results suggest that the most impactful legacy of these policies lies in their systematic dismantling of the overtly segregated educational systems that prevailed in the Jim Crow South. The large estimated effects on concrete measurable outcomes like adult educational attainment and poverty rates strongly indicate that this effort was not merely symbolic in nature, but was rather a generational achievement that tangibly improved the long-term well-being of southern African American children. The null effects among southern whites further suggests that these gains among Black students did not come at the expense of their white peers.

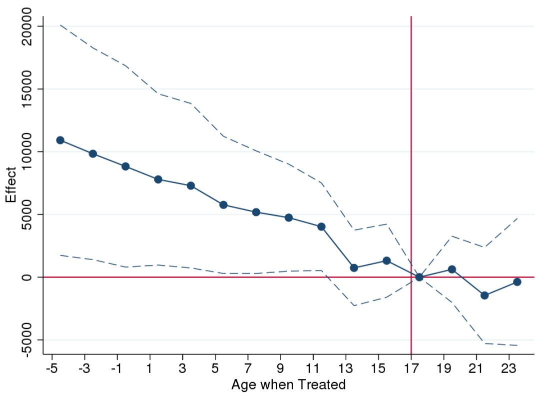
Our null results for northern school districts, however, do highlight potential limitations of even strongly implemented school desegregation initiatives in settings where they are not part of a transformative dismantling of overtly discriminatory education systems.

3.A Alternate Earnings Measures

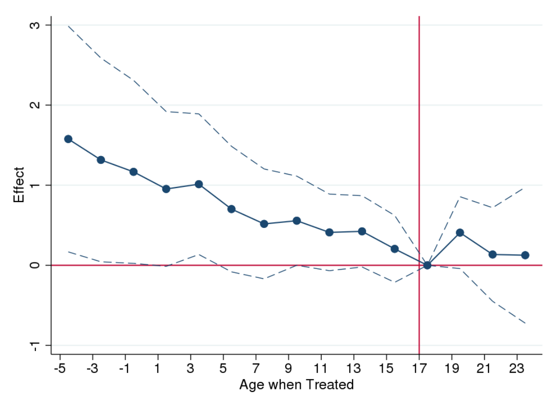
One concern with using log wage earnings is that our estimates may be attenuated by desegregation inducing negative selection into working at the extensive margin.[29] To assess the extent to which this may be happening, we also evaluate school desegregation order impacts on the log of wage earnings +1 as well as wage earnings in levels. Results from these alternative measures are presented in Figures 3.11a and 3.11b and indicate that treatment at age 5 compared to age 17 is associated with an additional $5,000 of earnings (in constant 2012 dollars) and a 1 point increase in the log of wage earnings +1. We also study effects for hourly wages and log hourly wages to see if our effects for wage earnings are coming solely from increased labor force attachment or a combination of this with improved job quality. Figures 3.11c and 3.11d indicate that the former story is most likely the one at play: effects for either measure of hourly wages are minimal, suggesting that while exposure to desegregation orders was successful in increasing labor force participation at both the extensive and intensive margin, it had limited impacts in improving the actual jobs that treated individuals matched to.

[29] Bailey et al. (2021) contend with a similar issue.

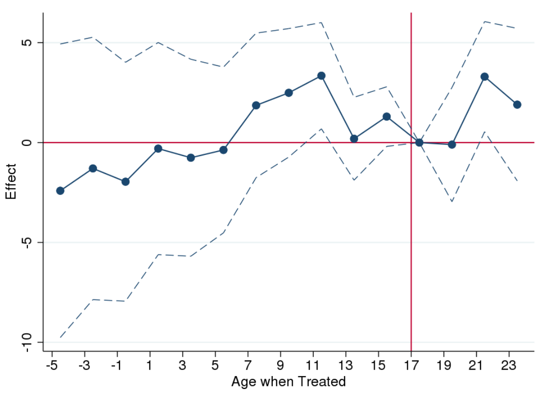
Figure 3.11: Results for Alternate Earnings Measures



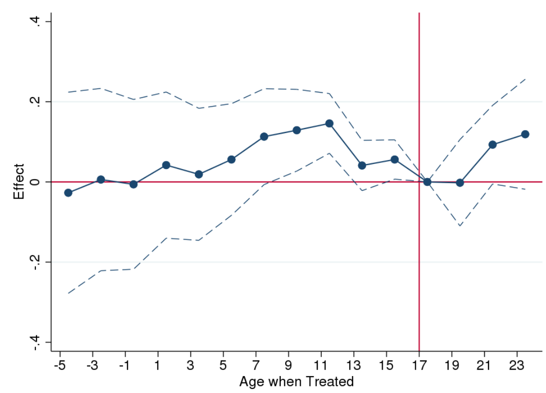
(a) Wage Earnings, Levels



(b) log(Wage Earnings + 1)



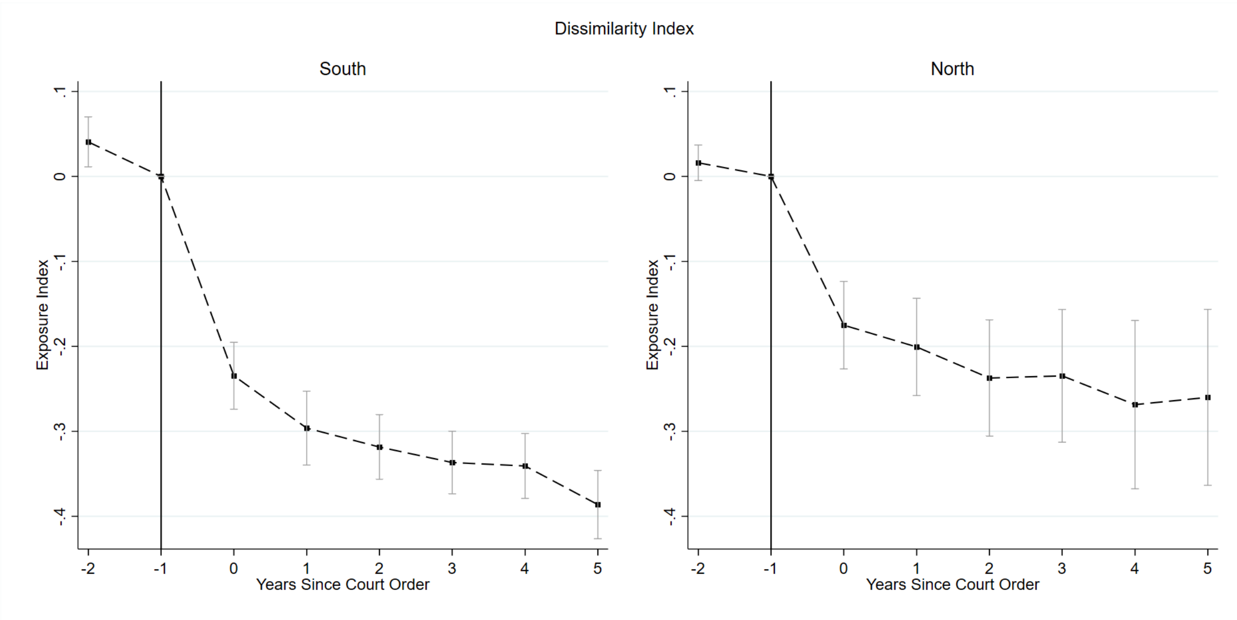
(c) Hourly Wage



(d) log(Hourly Wage)

Notes: Figures report estimates of δ_τ from Equation 3.1 for Southern Blacks. Sample is balanced in event time. Dotted lines indicate 95% confidence intervals, constructed using standard errors clustered at the county level. The sample contains 5.1 million observations collapsed to the county, survey year, cohort, race, and sex level. Weights equal to the sum of individual survey weights in each cell are applied. Controls include 1960 county characteristics interacted with linear cohort trends, as well as fixed effects for county, survey year, sex, and birth cohort-by-state of birth.

Figure 3.12: First Stage Effects on Dissimilarity Index



Notes: Figure plots coefficients from regressing the Dissimilarity Index onto indicators for the number of years relative to a court order (event time) as well as district and year fixed effects. Data come from school-level racial compositions collected by Office of Civil Rights Surveys, digitized and shared by Sarah Reber. Bands show 90% confidence intervals calculated with standard errors clustered at the school district level. Each district-year is given equal weight.

3.B Dissimilarity Index First Stage

Figure 3.12 reports first stage effects for the Dissimilarity Index, which are very similar to those for the Exposure Index reported in the main paper.

Bibliography

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

Adda, J., C. Dustmann, and K. Stevens (2017). The career costs of children. *Journal of Political Economy 125*(2), 293–337.

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

Angrist, J., G. Gray-Lobe, C. M. Idoux, and P. A. Pathak (2022). Still Worth the Trip? School Busing Effects in Boston and New York. Technical report, National Bureau of Economic Research.

Angrist, J. D., W. N. Evans, et al. (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *American Economic Review 88*(3), 450–477.

Angrist, J. D. and K. Lang (2004). Does school integration generate peer effects? Evidence from Boston's Metco Program. *American Economic Review 94*(5), 1613–1634.

Antman, F. M. and K. Cortes (2021). The Long-Run Impacts of Mexican-American School Desegregation. *NBER Working Paper 29200*.

Aparicio-Fenoll, A. and M. Vidal-Fernandez (2015). Working women and fertility: the role of grandmothers' labor force participation. *CESifo Economic Studies 61*(1), 123–147.

Arpino, B., C. D. Pronzato, and L. P. Tavares (2014). The effect of grandparental support on mothers' labour market participation: An instrumental variable approach. *European Journal of Population 30*(4), 369–390.

Ashenfelter, O., W. J. Collins, and A. Yoon (2006). Evaluating the Role of Brown v. Board of Education in School Equalization, Desegregation, and the Income of African Americans. *American Law and Economics Review 8*(2), 213–248.

Bailey, M., S. Sun, and B. Timpe (2021). Prep School for Poor Kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency. *American Economic Review 111*(12), 3963–4001.

Bailey, M. J. and A. Goodman-Bacon (2015). The War on Poverty's Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans. *American Economic Review 105*(3), 1067–1104.

Bailey, M. J., H. W. Hoynes, M. Rossin-Slater, and R. Walker (2020). Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence from the Food Stamps Program. *Review of Economic Studies Forthcoming.*

Baker, M., J. Gruber, and K. Milligan (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy 116*(4), 709–745.

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.

Bauernschuster, S. and M. Schlotter (2015). Public child care and mothers' labor supply—evidence from two quasi-experiments. *Journal of Public Economics 123*, 1–16.

Baum-Snow, N. and B. Lutz (2011). School Desegregation, School Choice, and Changes in Residential Location Patterns by Race. *American Economic Review 101*(December), 3019–3046.

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

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

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

Bergman, P. (2018). The risks and benefits of school integration for participating students: Evidence from a randomized desegregation program. *IZA Discussion Paper 11602*.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we Trust Difference-in-Differences Estimates? *The Quarterly Journal of Economics 119*(1), 249–275.

Bettendorf, L. J., E. L. Jongen, and P. Muller (2015). Childcare subsidies and labour supply—evidence from a large dutch reform. *Labour Economics 36*, 112–123.

Bick, A. (2016). The quantitative role of child care for female labor force participation and fertility. *Journal of the European Economic Association 14*(3), 639–668.

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

Billings, S. B., D. J. Deming, and J. Rockoff (2014). School Segregation, Educational Attainment, and Crime: Evidence from the end of Busing in Charlotte-Mecklenburg. *The Quarterly Journal of Economics 129*(1), 435–476.

Bishop, K. (2008). A dynamic model of location choice and hedonic valuation.

Blau, D. and E. Tekin (2007). The determinants and consequences of child care subsidies for single mothers in the usa. *Journal of Population Economics 20*(4), 719–741.

Boustan, L. P. (2010). Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration. *Quarterly Journal of Economics 125*(1), 417–443.

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

Boyd, D., H. Lankford, S. Loeb, and J. Wyckoff (2005). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management 24*(1), 113–132.

Bratti, M., T. Frattini, and F. Scervini (2018). Grandparental availability for child care and maternal labor force participation: pension reform evidence from italy. *Journal of Population Economics 31*(4), 1239–1277.

Budig, M. J. and P. England (2001). The wage penalty for motherhood. *American sociological review*, 204–225.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics 225*(2), 200–230.

Card, D. and A. B. Krueger (1992a). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *Journal of Political Economy 100*(1), 1–40.

Card, D. and A. B. Krueger (1992b). School Quality and Black-White Relative Earnings: A Direct Assessment. *The Quarterly Journal of Economics 107*(1), 151–200.

Card, D. and J. Rothstein (2007). Racial segregation and the black–white test score gap. *Journal of Public Economics 91*(11-12), 2158–2184.

Cascio, E., N. Gordon, E. Lewis, and S. Reber (2008). From Brown to Busing. *Journal of Urban Economics 64*(2), 296–325.

Cascio, E., N. Gordon, E. Lewis, and S. Reber (2010). Paying for Progress: Conditional Grants and the Desegregation of Southern Schools. *The Quarterly Journal of Economics 125*(1), 445–482.

Cascio, E. U. (2009). Maternal labor supply and the introduction of kindergartens into american public schools. *Journal of Human resources 44*(1), 140–170.

Cascio, E. U., N. Gordon, and S. Reber (2013). Local Responses to Federal Grants: Evidence from the Introduction of Title I in the South. *American Economic Journal: Economic Policy 5*(3), 126–159.

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

Chan, T. W. and J. Ermisch (2015). Proximity of Couples to Parents: Influences of Gender, Labor Market, and Family. *Demography 52*(2), 379–399.

Charles, K. K. and P. Sevak (2005, Nov). Can family caregiving substitute for nursing home care? *Journal of Health Economics 24*(6), 1174–1190.

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

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How does your Kindergarten Classroom Fffect your Earnings? Evidence from Project Star. *The Quarterly Journal of Economics 126*(4), 1593–1660.

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

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

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

Child Care Aware (2017). Parents and the high cost of child care. Technical report.

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

Coate, P. (2013). Parental Influence on Labor Market Outcomes and Location Decisions of Young Workers.

Coate, P., P. M. Krolikowski, and M. Zabek (2017). Parental Proximity and Earnings after Job Displacements.

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

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

Compton, J. and R. A. Pollak (2014). Family proximity, childcare, and women's labor force attachment. *Journal of Urban Economics 79*, 72–90.

Compton, J. and R. A. Pollak (2015). Proximity and Coresidence of Adult Children and their Parents in the United States: Description and Correlates. *Annals of Economics and Statistics* (117-118), 91–114.

Cortes, P. and J. Pan (2020). Children and the remaining gender gaps in the labor market. Technical report, National Bureau of Economic Research.

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

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

Cutler, D. M., E. L. Glaeser, and J. L. Vigdor (1999). The rise and decline of the american ghetto. *Journal of Political Economy 107*(3), 455–506.

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

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

Davanzo, J. (1983). Repeat Migration in the United States: Who Moves Back and Who Moves On? *The Review of Economics and Statistics 65*(4), 552–559.

de Chaisemartin, C. and X. D'Haultfœuille (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review 110*(9), 2964–2996.

Del Boca, D. (2015). Child care arrangements and labor supply. Technical report, IDB Working Paper Series.

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

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

Derenoncourt, E. (2022). Can you move to opportunity? Evidence from the Great Migration. *American Economic Review 112*(2), 369–408.

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.

Dierx, A. (1988). A Life-Cycle Model of Repeat Migration. *Regional Science and Urban Economics 18*, 383–397.

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

Dimova, R. and F.-C. Wolff (2011). Do downward private transfers enhance maternal labor supply? evidence from around europe. *Journal of Population Economics 24*(3), 911–933.

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

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

Eckstein, Z. and K. I. Wolpin (1989). Dynamic labour force participation of married women and endogenous work experience. *Review of Economic Studies 56*(3), 375–390.

Edin, K. and M. Kefalas (2011). *Promises I can keep: Why poor women put motherhood before marriage.* Univ of California Press.

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

Francesconi, M. (2002). A joint dynamic model of fertility and work of married women. *Journal of Labor Economics 20*(2), 336–380.

García-Morán, E. and Z. Kuehn (2017). With strings attached: Grandparent-provided child care and female labor market outcomes. *Review of Economic Dynamics 23*, 80–98.

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.

Gemici, A. (2011). Family Migration and Labor Market Outcomes.

Goldin, C. and J. Mitchell (2017). The new life cycle of women's employment: Disappearing humps, sagging middles, expanding tops. *Journal of Economic Perspectives 31*(1), 161–82.

Goodman-Bacon, A. (2021). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics 225*(2), 254–277.

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

Guner, N., R. Kaygusuz, and G. Ventura (2020, 03). Child-Related Transfers, Household Labour Supply, and Welfare. *The Review of Economic Studies 87*(5), 2290–2321.

Guryan, J. (2004). Desegregation and Black Dropout Rates. *American Economic Review 94*(4), 919–943.

Haan, P. and K. Wrohlich (2011). Can child care policy encourage employment and fertility?: Evidence from a structural model. *Labour Economics 18*(4), 498–512.

Hanushek, E. (1986). The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature 24*(3), 1141–1177.

Hanushek, E. (2003). The Failure of Input-based Schooling Policies. *The Economic Journal 113*, F64–F98.

Havnes, T. and M. Mogstad (2011). Money for nothing? universal child care and maternal employment. *Journal of Public Economics 95*(11-12), 1455–1465.

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

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

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

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

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

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

Hoynes, H., D. W. Schanzenbach, and D. Almond (2016). Long-Run Impacts of Childhood Access to the Safety Net. *American Economic Review 106*(4), 903–934.

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

Hyman, J. (2017). Does Money Matter in the Long Run? Effects of School Spending on Educational Attainment. *American Economic Journal: Economic Policy 9*(4), 256–280.

Isen, A., M. Rossin-Slater, and W. R. Walker (2017). Every Breath You Take—Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970. *Journal of Political Economy 125*(3), 848–902.

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

Jackson, C. K. (2009). Student Demographics, Teacher Sorting, and Teacher Quality: Evidence from the End of School Desegregation. *Journal of Labor Economics 27*(2), 213–256.

Jackson, C. K., R. C. Johnson, and C. Persico (2016). The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. *The Quarterly Journal of Economics 131*(1), 157–218.

Johnson, J. E. and S. Schulhofer-Wohl (2019). Changing Patterns of Geographic Mobility and the Labor Market for Young Adults. *Journal of Labor Economics 37*(S1), S199–S241.

Johnson, R. C. (2011). Long-Run Impacts of School Desegregation and School Quality. *NBER Working Paper 16664*.

Johnson, R. C. (2019). *Children of the dream: Why school integration works.* Basic Books.

Kaplan, G. (2012). Moving Back Home : Insurance against Labor Market Risk. *Journal of Political Economy 120*(3), 446–512.

Kaplan, G. and S. Schulhofer-Wohl (2017). Understanding the long-run decline in interstate migration. *International Economic Review 58*(1), 57–94.

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

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

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

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

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2019). Child Penalties across Countries: Evidence and Explanations. *AEA Papers and Proceedings 109*, 122–126.

Kleven, H., C. Landais, and J. E. Søgaard (2019). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics 11*(4), 181–209.

Kluger, R. (2011). *Simple justice: The history of Brown v. Board of Education and Black America's struggle for equality*. Vintage.

Konrad, K. A., H. Kunemund, K. E. Lommerud, and J. R. Robledo (2002, Sept). Geography of the family. *American Economic Review 92*(4), 981–998.

Lafortune, J., J. Rothstein, and D. W. Schanzenbach (2018). School Finance Reform and the Distribution of Student Achievement. *American Economic Journal: Applied Economics 10*(2), 1–26.

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

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

Lefebvre, P. and P. Merrigan (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from canada. *Journal of Labor Economics 26*(3), 519–548.

Logan, J. (2021). American Communities Project Desegregation Case Data, Brown University . *https://s4.ad.brown.edu/Projects/USSchools/SchoolDLoad.htm*.

Lundin, D., E. Mörk, and B. Öckert (2008). How far can reduced childcare prices push female labour supply? *Labour Economics 15*(4), 647–659.

Lutz, B. (2011). The end of court-ordered desegregation. *American Economic Journal: Economic Policy 3*(2), 130–68.

Maddow-Zimet, I. and K. Kost (2020). Effect of changes in response options on reported pregnancy intentions: A natural experiment in the united states. *Public Health Reports 135*(3), 354–363.

Massey, D. S. and N. A. Denton (1988). The dimensions of residential segregation. *Social Forces 67*(2), 281–315.

McMurry, J. (2021). Child Care Policy and Informal Care.

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.

O'Flaherty, B. (2015). *The economics of race in the United States*. Harvard University Press.

Pierson, K., M. L. Hand, and F. Thompson (2015). The government finance database: A common resource for quantitative research in public financial analysis. *PloS one 10*(6), e0130119.

Posadas, J. and M. Vidal-Fernandez (2013). Grandparents' childcare and female labor force participation. *IZA Journal of Labor Policy 2*(1), 1–20.

ProPublica (2014). A national survey of school desegregation orders. *https://projects.propublica.org/graphics/desegregation-orders*.

Rainer, H. and T. Siedler (2009). O brother, where art thou? The effects of having a sibling on geographic mobility and labour market outcomes. *Economica 76*(303), 528–556.

Reber, S. J. (2005). Court-Ordered Desegregation: Successes and Failures Integrating American Schools since Brown versus Board of Education. *Journal of Human Resources 40*(3), 559–590.

Reber, S. J. (2010). School Desegregation and Educational Attainment for Blacks. *Journal of Human Resources 45*(4), 893–914.

Reber, S. J. (2011). From Separate and Unequal to Integrated and Equal? School Desegregation and School Finance in Louisiana. *Review of Economics and Statistics 93*(2), 404–415.

Rivkin, S. and F. Welch (2006). Has School Desegregation Improved Academic and Economic Outcomes for Blacks? In E. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 2, pp. 1019–1049. Elsevier.

Rosenzweig, M. and K. I. Wolpin (1980). Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. *Econometrica 48*(1), 227–240.

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.

Solon, G., S. J. Haider, and J. M. Wooldridge (2015). What are we Weighting for? *Journal of Human Resources 50*(2), 301–316.

Spilimbergo, A. and L. Ubeda (2004). Family attachment and the decision to move by race. *Journal of Urban Economics 55*(3), 478–497.

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

Spring, A., E. Ackert, K. Crowder, and S. J. South (2017). Influence of Proximity to Kin on Residential Mobility and Destination Choice: Examining Local Movers in Metropolitan Areas. *Demography 54*(4), 1277–1304.

Stuart, B. A. (2022). The Long-Run Effects of Recessions on Education and Income. *American Economic Journal: Applied Economics 14*(1), 42–74.

Sun, L. and S. Abraham (2021). Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics 225*(2), 175–199.

Taylor, E. J., B. A. Stuart, and M. J. Bailey (2016). Summary of Procedure to Match NUMIDENT Place of Birth County to GNIS Places. Technical report, Census Bureau.

Tekin, E. (2007). Childcare subsidies, wages, and employment of single mothers. *Journal of Human Resources 42*(2), 453–487.

Thompson, O. (Forthcoming). School Desegregation and Black Teacher Employment. *Review of Economics and Statistics*.

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

Tuttle, C. (2019). The Long-run Economic Effects of School Desegregation.

Venator, J. (2022). Dual-Earner Migration Decisions, Earnings, and Unemployment Insurance.

Weiner, D. A., B. F. Lutz, and J. Ludwig (2009). The Effects of School Desegregation on Crime. *NBER Working Paper 15380*.

Welch, F. and A. Light (1987). *New Evidence on School Desegregation.* US Commission on Civil Rights.

Zabek, M. (2019). Local Ties in Spatial Equilibrium. Technical report.

Zamarro, G. (2020). Family labor participation and child care decisions: the role of grannies. *SERIEs 11*(3), 287–312.

Zolna, M. and L. D. Lindberg (2012). Unintended pregnancy: Incidence and outcomes among young adult unmarried women in the united states, 2001 and 2008. Technical report, Alan Guttmacher Institute New York.