The Impact of Racial Segregation on College Attainment in Spatial Equilibrium*

Victoria Gregory — FRB of St. Louis

Julian Kozlowski — FRB of St. Louis

Hannah Rubinton — FRB of St. Louis

October 13, 2022

Abstract

We incorporate race into an overlapping-generations spatial-equilibrium model with neighborhood spillovers. Race matters in two ways: (i) the Black-White wage gap and (ii) homophily—the preferences of individuals over the racial composition of their neighborhood. We find that these two forces generate a Black-White college gap of 22 percentage points, explaining about 80% of the college gap in the data for the St. Louis metro area. Counterfactual exercises show that the wage gap and homophily explain 7 and 18 percentage points of the college gap, respectively. A policy of equalizing school funding across neighborhoods reduces the college gap by almost 10 percentage points and generates large welfare gains.

JEL Classification: J15, J24, O18

Keywords: Racial disparities, neighborhood segregation, education, income inequality.

---

*We thank our discussants Gregorio Caetano, Marcus Casey, David Schonholzer, and Curtis Simon. We also thank Milena Almagro, Diego Daruich, Vikram Maheshri, Charly Porcher, Richard Rogerson, and seminar participants at the Urban Economics Association meeting, DePaul University, Online Spatial & Urban Seminar, Society of Economics Dynamics, McMaster, and CIREQ Macro Conference for comments and suggestions. We thank Maggie Isaacson, Joel Steinberg, and Elisabeth Harding for their valuable assistance with the data. The views do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis. This work employed resources from the BigTex High Performance Computing Group at the Federal Reserve Bank of Dallas.

1 Introduction

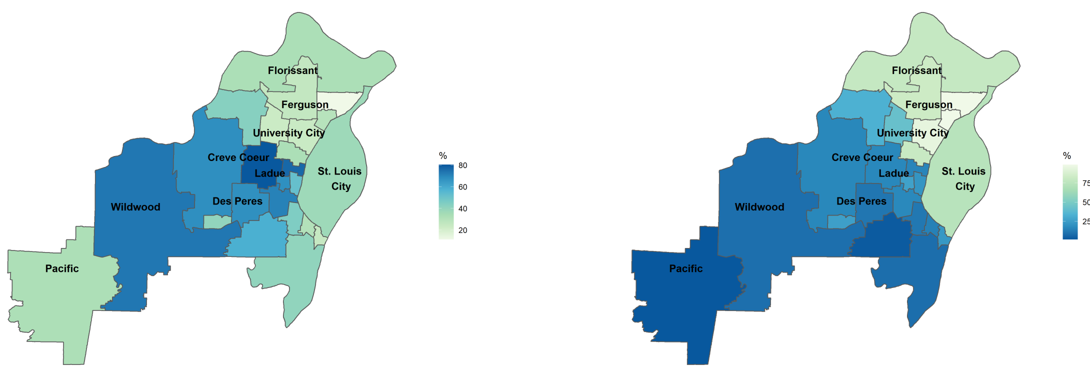
A growing body of research shows that the neighborhood a child grows up in profoundly impacts adult outcomes, such as college attainment and intergenerational mobility (e.g., Chetty et al., 2018; Chetty and Hendren, 2018). However, segregation by race is a predominant feature of American cities, and as a result, there is substantial racial inequality in exposure to advantageous neighborhoods (e.g., Bayer et al., 2021). This paper explores how racial differences in neighborhood sorting can account for the Black-White gap in college attainment.

For example, in St. Louis, one of the most segregated cities in the country and the focus of this paper, there is a Black-White gap in college attainment of 28 percentage points. Of the White children who grow up in St. Louis, 47% of them will earn a college degree while only 19% of the Black children will (Chetty et al., 2018). At the same time, there is substantial neighborhood segregation by race. To see this, Figure 1 presents two maps of the school districts in St. Louis City and County. The left panel shows the proportion of the district's students who enroll in a four-year college degree program, while the right panel shows the share of the district's students who are Black. The figures show a striking correlation between the two, suggesting a relationship between the city's segregation and its racial gap in college attainment. While previous authors have empirically demonstrated a causal relationship between segregation and college attainment (Ananat, 2011; Cutler and Glaeser, 1997), the underlying mechanisms which generate this relationship are not yet fully understood.

This paper builds an overlapping-generations spatial-equilibrium model to study the effect of the Black-White wage gap and homophily—defined as the preference to live with neighbors of your own race. We find that these two forces account for 22 percentage points of the Black-White gap in college attainment, explaining about 80% of the 28 percentage points college gap in the data. Both features are essential for generating segregation and the racial gap in educational attainment, with homophily being the primary driver of both. Even when the wage gap is closed, homophily generates racial differences in exposure to neighborhoods with high local spillovers and well-funded schools, which maintains the gap in educational attainment. Thus, incorporating both features is essential for understanding segregation and its impact on opportunity and intergenera-

Figure 1: Neighborhood Segregation in St. Louis County

College Attainment Black Share



Notes: The left panel shows the share of students who go on to attend a 4-year college in each school district. The right panel shows the share of Black students who attend that school district. Source: School-district level data for Missouri in 2020 from the National Center for Education Statistics.

tional mobility. Finally, we consider a policy reform that equalizes school funding across neighborhoods. The intervention reduces the college attainment gap by almost 10 percentage points.

Section 2 extends a standard overlapping-generations spatial-equilibrium model of a city to include race. Families choose the neighborhood where they live, taking into consideration local spillovers and differences in school funding which, in addition to innate ability and parental private investment, affect their child's future education and income. We model school funding as responding endogenously to local tax revenue, which increases with neighborhood rent. Local spillovers are a function of the share of adults with a college education, as suggested by Fogli and Guerrieri (2019). The model incorporates race in only two ways. First, race affects workers' wages, with Black workers earning less than White workers conditional on education and skills—the Black-White wage gap. Second, race affects household preferences over the racial composition of their neighborhood—known as homophily. While there are other ways in which race might affect housing and education choices, we incorporate only these two aspects because a substantial body of empirical evidence supports them. In addition, we find that these two differences account for a large share of the college gap.

Section 3 derives the empirical estimates needed to calibrate the model and match critical features of the St. Louis MSA. We empirically document three facts that serve as essential inputs to the calibration. First, it is well known that there is a race penalty in wage equation estimates from Mincer (1974) regressions (e.g., Heckman et al., 2006; Thompson, 2021); conditional on test scores, we document a racial wage gap of about 8%. Second, we calibrate the model to reflect neighborhood segregation patterns in the data. We cluster the St. Louis MSA Census Tracts into three groups: a) Black and low-income, b) White and low-income, and c) White and high-income. This grouping parsimoniously captures the most prominent patterns in the data on household sorting across neighborhoods along the lines of race and income. Third, using data from the 1997 National Longitudinal Survey of Youth (NLSY97), we document the relationship between (i) college attainment and (ii) ability and parental transfers. These factors positively affect college attainment, and these estimates help us to inform their importance on the probability of college attainment in the model.

Section 4 uses the empirical estimates to take the model to the data. We then carry out three validation tests of the key mechanisms in the model. We compare the model with estimates from the literature on (1) the causal effect of neighborhoods on college attainment estimated by Chetty et al. (2016); (2) the causal effect of segregation on educational attainment estimated by Ananat (2011); (3) the causal effect of a neighborhood's Black share on neighborhood choice estimated by Caetano and Maheshri (2019). In each case, we find that the model is consistent with the estimates from the literature.

Two forces allow the model to match the neighborhood segregation patterns in the data.[1] First, high-income households are more willing to pay for neighborhood amenities than low-income households, implying that neighborhoods are segregated by income.[2] Because of the Black-White wage gap, segregation by income also implies segregation by race. Second, segregation is generated by homophily.[3] Following Banzhaf and Walsh (2013), White and Black households prefer

[1]While the underlying causes of the segregation are many-fold, in this project, we include economic factors and preferences abstracting from historical barriers enacted to limit integration (such as redlining). Cutler et al. (1999) find that while redlining and other institutional factors were more important before the 1970s, today, the decentralized choices of individuals are more important for generating segregation.

[2]Following Couture et al. (2019) our utility function includes unit-demand for housing. This generates a non-homotheticity such that high-income households spend a lower share of their income on housing and, therefore, are more willing to pay for neighborhood amenities.

[3]Among others, Aliprantis et al. (2019); Bayer et al. (2017); Caetano and Maheshri (2019) estimate neighborhood choice models and find empirical evidence of homophily. Farley et al. (1997) interview households in four cities and

neighborhoods with a specific racial composition, with Black families preferring more integration than White households. The two forces interact; high-income White households sort into higher-amenity neighborhoods; consequently, these neighborhoods are less appealing to Black families.

Section 5 shows that the presence of (i) the wage gap—which we estimate directly from the data—and (ii) homophily—which we calibrate to generate the racial composition of neighborhoods—generate a college gap of 22 percentage points and explain about 80% of the college gap in the data. Notably, the college gap is not a target in the calibration. The other 20% that the model does not capture is likely coming from other forces that are not present in the model (e.g., differential access to credit).

We use the model to decompose how much of the college attainment gap is due to the wage gap and homophily through counterfactual exercises. On the one hand, removing the wage gap increases the educational attainment of Black children. This is because Black households now have more resources to invest in their children's education. Because of homophily, neighborhoods remain substantially segregated, with White households segregating by income across two neighborhoods and Black households clustering in one neighborhood. As a result, there is still a racial gap in educational attainment because the school funding and the neighborhood spillovers are smaller in the Black neighborhood than in the high-income White neighborhood.

On the other hand, removing homophily increases neighborhood integration even when the Black-White wage gap remains in place. Some racial segregation remains because households sort by income, and Black households earn less than otherwise similar White households. The racial education gap decreases because Black households move to neighborhoods with better-funded schools and more spillovers, but a gap in parental investment remains due to wage differences.

Finally, Section 6 explores the effects of a policy reform. In the baseline economy, school funding in the poorest neighborhood is about 2 times larger than in the richest neighborhood due to differences in the tax revenue collected from local property taxes. We explore the effects of a policy reform that equalizes school funding across neighborhoods. The college gap reduces by almost 10 percentage points due to an increase in the college attainment of Blacks and a decrease in the college attainment of Whites. Interestingly, these effects are reduced by one-half if you

find that racial composition is an important component of their neighborhood preference.

only consider the partial equilibrium effect and do not take into account households' equilibrium response to the policy. The main reason is that in the new equilibrium, college-educated parents endogenously move to neighborhoods that receive a boost in school funding, which equalizes spillovers across neighborhoods and helps to reduce the college attendance gap.

Related literature. This paper builds on results from several strands of empirical literature. This includes work on the underlying causes of segregation (Boustan, 2013; Card et al., 2008; Cutler et al., 1999; Dawkins, 2005; Echenique and Fryer, 2007; Monarrez and Schonholzer, 2021; Sethi and Somanathan, 2009) and the consequences of segregation (Ananat, 2011; Billings et al., 2013; Cutler and Glaeser, 1997; Derenoncourt, 2022; Johnson, 2011).

Besides the empirical literature, this paper builds on two separate strands of structural literature. On the one hand, there is a literature which examines racial segregation in spatial equilibrium models first pioneered in Schelling (1969, 1971). Several papers (e.g. Banzhaf and Walsh, 2013; Bayer and McMillan, 2005; Bayer et al., 2004; Caetano and Maheshri, 2019; Sethi and Somanathan, 2004) examine racial segregation in models with homophily, Black-White wage gaps and exogenous neighborhood amenities, but do not consider the impact on human capital accumulation.

On the other hand, several papers (e.g. Aliprantis and Carroll, 2018; Chyn and Daruich, 2022; Eckert and Kleineberg, 2019; Fogli and Guerrieri, 2019; Zheng and Graham, 2022) examine human-capital spillovers in quantitative spatial equilibrium models and several papers (e.g. Almagro and Domínguez-Iino, 2020; Couture et al., 2019; Hoelzlein, 2020) examine the interaction between endogenous amenities and sorting by income, but do not consider race.[4] This second set of papers is built on the literature on discrete choice models with local spillovers (e.g., Benabou, 1994; Brock and Durlauf, 1995; Fernandez and Rogerson, 1996), which, while seemingly motivated by racial inequalities, do not specifically model race. One exception is Badel (2015) who presents a model of racial segregation and human capital accumulation. He shows that the model has multiple equilibrium, including an equilibrium in which White households earn more than Black ones due to differences in human capital accumulation.

[4] An exception is Aliprantis and Carroll (2018) who calibrate a two-neighborhood model to mimic a predominantly Black neighborhood and a predominantly White neighborhood in Chicago. However, they do not model race directly so cannot consider the effect of homophily or the Black-White wage gap.

Our model builds on previous work in two ways. First we incorporate two ways in which race affects neighborhood segregation: inequalities in the labor market and homophily. Second, we incorporate two ways in which segregation endogenously affects educational attainment and intergenerational mobility: local neighborhood spillovers and school funding. Our quantitative analysis shows that capturing each of these mechanisms and their interactions is important for studying the interplay between racial segregation, income segregation, and human capital accumulation.

The rest of this paper proceeds as follows. Section 2 presents the model. Then, in Section 3, we present facts on the Black-White wage gap and segregation in the St. Louis MSA. While these facts are well known, we reproduce them in our data as they are important inputs to our quantitative analysis. Section 4 shows how we take the model to the data. Section 5 presents the main counterfactual exercises removing homophily and closing the Black-White wage gap. In Section 6 we show the policy exercise of equalizing school funding. Finally, Section 7 concludes.

2 Model

We extend a standard overlapping-generation spatial-equilibrium model to incorporate race. We model a single metro area where families choose a neighborhood to live in, considering differences in local spillovers and school funding that affect their children's future income and education. The model incorporates two mechanisms through which the individual's race affects the neighborhood choice: racial disparities in the labor market, which are reflected in their income, and homophily—the preference to live with neighbors of the same race.

2.1 Set up

The economy is populated by overlapping generations of agents who live for two periods. Agents are of race $r \in \{r_B, r_W\}$, which is a permanent characteristic of the dynasty. In the first period, the agent is young and acquires education. In the second period, the agent is an adult with an income that depends on their education, skills, and race.[5] Labor is perfectly mobile across neighborhoods,

[5]The second period represents the entire working life, which means that there are complete markets in this stage. Hence, the paper abstracts from borrowing constraints and the potential differential parental wealth and transfers across

so wages do not depend on the neighborhood in which a household lives.

There are 3 neighborhoods, denoted by $n \in \{A, B, C\}$. All houses are of the same size and quality, and the rent in neighborhood n is denoted by p_n. Black and White households pay the same rental price, meaning we do not explicitly include discrimination in the housing market.[6] Housing is supplied elastically according to $S_n = \eta_n p_n^{\psi}$, where $\psi > 0$ is the price elasticity of housing supply and η_n reflects land availability.

There are two educational levels, $e \in \left\{e^L, e^H\right\}$. Agents choose whether to have low or high education. Five key characteristics shape the education choice. First, the education choice depends on the agent's race, as wages are race-specific. Second, two individual inputs affect the cost of education: the innate ability of the agent and the level of parental investment. Third, two additional inputs depend on the neighborhood where the child grows up in: local spillovers and school funding. The size of the local spillover effect in neighborhood n is a function of the share of households with parents with high education in that neighborhood, X_n. School funding, Q_n, is determined by local property tax revenue and funding provided by state and federal sources. Because local tax revenues depend on neighborhood house prices, the component of school funding due to local funding is neighborhood-specific, $Q_n = \tau p_n + \bar{Q}$. $\bar{Q}$ is federal and state funding, and τ is the property tax rate.

2.2 Adult's Problem

For an adult of race r, innate ability a, skills s, and education level e who was born in neighborhood n_0, the value of living in neighborhood n is

$$V(r, a, s, e, n_0, n) = \max_{c,i} \log(h) + \log(c) + \beta \mathbb{E}\left[\mathcal{V}\left(r, a', s', e', n\right)\right]$$

races.

[6] The literature has found that, at least since 1990, housing market discrimination is less important for generating segregation than the preferences of White households to self-segregate (Cutler and Glaeser, 1997). Estimates of the gap in prices paid for the same unit of housing tend to be small. For example, Bayer et al. (2017) estimates that Black and Hispanic home-buyers pay a premium of 2% relative to White households for the same house.

subject to

$$
\begin{aligned}
c+i+(1+\tau)\, p_n+m\mathbb{I}_{\{n\neq n_0\}} &= y(r,e,s)\\
\log s' &= F^s(a',i,Q_n,X_n)\\
P(e'=e^H) &= G^e\left(r,a',s',n\right)\\
a' &\sim \Gamma(a)\\
h &= A(n,r,S_{r,n})
\end{aligned}
$$

where β is the altruistic discount factor, that is, the extent to which parents care about the utility of their offspring. The cash-on-hand available for adults to spend is comprised solely of their labor income $y(r,e,s)$. Income is a function the race, education, and skills s. They split their budget between consumption, investments into their children's education, rent, and moving costs. In order to live in a neighborhood, agents must consume one unit of housing services at rental price p_n. They must also pay the property taxes in their chosen neighborhood and a moving cost m if they decide to move to a neighborhood different from the one they grew up in.

The function F^s determines the skills of the children, which depends on the child's ability, the parent's investment, and the neighborhood characteristics in which the children grow up—Q_n and X_n. We describe the functional form for skills in the calibration. The education probability is endogenous and depends on the optimization of the child, which we summarize by the policy function G^e, which depends on the state variables of the child. The child's innate ability a' is drawn from a distribution that depends on the parent's innate ability. Thus, innate ability is imperfectly transmitted across generations but does not depend on race. Finally, we model homophily as adults receiving utility from a neighborhood amenity, h, which depends on the racial composition of the neighborhood and its race. In the next Section, we describe its functional form and calibration.

Given the value from living in each neighborhood, an adult of race r, innate ability a, skills s, education e, and initial neighborhood n_0 chooses a neighborhood in which to live

$$
\mathscr{V}\left(r,a,s,e,n_0\right)=\mathbb{E}_\varepsilon\left[\max_n\left\{V\left(r,a,s,e,n_0,n\right)+\varepsilon^n\right\}\right].
$$

Households have preference shocks ε^n that are independently, identically distributed, and drawn from an extreme value distribution with shape parameter κ. Therefore, the probability that a household of type (r,a,s,e,n_0) chooses to live in neighborhood n is

$$\lambda\left(r,a,s,e,n_0,n\right)=\frac{\exp\left(\frac{1}{\kappa}V\left(r,a,s,e,n_0,n\right)\right)}{\sum_{n\in N}\exp\left(\frac{1}{\kappa}V\left(r,a,s,e,n_0,n\right)\right)},$$

and the expected value function is

$$\mathscr{V}\left(r,a,s,e,n_0\right)=\kappa\ln\left(\sum_{n\in N}\exp\left(\frac{V\left(r,a,s,e,n_0,n\right)}{\kappa}\right)\right).$$

2.3 Child's Problem

A child of race r, innate ability a, skills s, growing up in neighborhood n, with parental investment i chooses the education level $e\in\{e^L,e^H\}$ such that

$$e=\underset{\{e^L,e^H\}}{\operatorname{argmax}}\{\mathscr{V}\left(r,a,s,e^L,n\right)+\sigma^L,\mathscr{V}\left(r,a,s,e^H,n\right)-C(s)+\sigma^H\}.$$

Children have a preference shocks for education, σ^L and σ^H, that are independently, identically distributed, and drawn from an extreme value distribution with shape parameter σ. $C(s)$ is a utility cost of acquiring education which is decreasing on the skills of the child: $C(s)=\bar{c}-s$. Finally, the probability that the child chooses high education is

$$G^e\left(r,a,s,n\right)=\frac{1}{1+exp\left(-\frac{1}{\sigma}\left[\mathscr{V}\left(r,a,s,e^H,n\right)-C(s)-\mathscr{V}\left(r,a,s,e^L,n\right)\right]\right)}.$$

Intergenerational transmission The model captures three channels of intergenerational linkages. First, ability is imperfectly transferred from parent to child; while there is mean reversion in innate ability, a high-ability adult is likely to have a high-ability child. Second, we explicitly model investment as inter-vivos transfers between parents and children. Investment by parents leads to higher-skilled children, which implies higher incomes and higher education probabilities. The third intergenerational linkage is through the neighborhood in which a child is born. The neighborhood captures two forces. On the one hand, living in a high quality neighborhood is a complementary way of investing in the skills of a child, increasing income and educational attainment. On the other hand, there is persistence in neighborhood choices due to the moving cost. A child who is born in a high-quality neighborhood does not need to pay the moving cost to live there as an adult, while a child born in a low-quality neighborhood will face an additional barrier to upgrading their neighborhood quality. While we do not explicitly model an inherited-wealth gap, the model the intergenerational persistence of income and education through these other channels.

2.4 Equilibrium

Definition: A Recursive Competitive Equilibrium is characterized policy functions for residential choice $n(r,a,s,e,n_0)$, consumption, $c(r,a,s,e,n_0,n)$, education choice $e(r,a,s,e,n)$, investment $i(r,a,s,e,n_0,n)$, value functions $V(r,a,s,e,n_0,n)$, house prices $\{p_n\}_{n=1}^N$, local spillovers $\{X_n\}_{n=1}^N$, school qualities $\{Q_n\}_{n=1}^N$, and an ergodic distribution $F(r,a,s,e,n_0,n)$ of population over race, ability, skills, education, birth neighborhood and neighborhood choice such that:

(i) Household optimization: the policy functions n,c,e,i solve the adult and child's problem.

(ii) Housing market clearing

$$\eta_n p_n^{\psi} = S_n = \int F(dr,da,ds,de,dn_0,n)$$

for all $n = 1,\ldots,N$

(iii) Spillover consistency:

$$X_n = \frac{\int F\left(dr,da,ds,e^H,dn_0,n\right)}{\int F(dr,da,ds,de,dn_0,n)}$$

for all $n = 1,\ldots,N$

(iv) School funding consistency:

$$Q_n = \tau p_n + \bar{Q}$$

for all $n = 1,\ldots,N$

(v) Location consistency:

$$S_{\tilde{r},n} = \frac{\int F(\tilde{r}, da, ds, de, dn_0, n)}{\int F(dr, da, ds, de, dn_0, n)}$$

for all $n = 1, \ldots, N$ and $r = \{b, w\}$.

3 Functional Forms and Empirical Estimates

Before taking the model to the data, we explain how we model homophily and derive critical empirical estimates that inform the calibration in Section 4. We focus on St. Louis MSA. First, we describe how we model homophily. Second, we use a k-means clustering to divide neighborhoods into three groups. Third, we estimate wage equations to quantify the Black-White wage gap. Fourth, we look at the relationship between college graduation, skill, and parental investment to provide empirical discipline to the skill function and the cost of education.

3.1 Homophily

We model homophily as a penalty on amenities as in Banzhaf and Walsh (2013). The amenities are as follows:

$$A(n, r, S_{r,n}) = A_n \left(1 - \varphi_r (S_{r,n} - \gamma_r)^2\right). \tag{1}$$

The amenity an individual enjoys is made up of two parts. First is an exogenous component, A_n, representing fixed neighborhood characteristics such as proximity to downtown. Second, the amenity includes an endogenous component that depends on the neighborhood's racial composition, $S_{r,n}$. We follow Banzhaf and Walsh (2013) and assume that households have a "bliss point" for the degree of racial integration in their neighborhoods. The "bliss point" γ_r depends on the race. As the racial composition of the neighborhood deviates from their bliss point, households benefit less from the exogenous amenity by a factor $\left(1 - \varphi_r (S_{r,n} - \gamma_r)^2\right)$. As a result, homophily captures an explicit preference to live with neighbors of your race. An alternative interpretation is that agents fear being discriminated against when living in a neighborhood in which the neighborhood racial composition deviates from their “bliss point”. Explicitly, if an agent worries about facing discrimination in a public park, they enjoy the amenity less than an individual who does not worry about discrimination.

We provide two sources of empirical evidence on homophily in housing markets. First, survey evidence from Farley et al. (1978), and Krysan and Farley (2002) show that Black households prefer a neighborhood mix that is 50% White and 50% Black, while White households prefer a mix that is 90% White and 10% Black. Hence, we set $\gamma_W = 0.9$ and $\gamma_B = 0.5$, a standard assumption in the literature (e.g., Banzhaf and Walsh, 2013). Interestingly, the survey evidence also shows that the main reasons for these choices are based on racial characteristics, independent of other neighborhood characteristics and amenities. Second, Caetano and Maheshri (2019) isolates the causal effect of race from other neighborhood characteristics. In particular, it finds that the responses for the racial composition are more significant than the responses for income, consistent with the survey evidence mentioned above. We use these moments as part of our validation exercises in Section 4.[7]

3.2 Neighborhood Segregation

The model has three neighborhoods, so we need to map the data to three types of places. We use data on Census Tracts from the 2000 Census, and Chetty et al. (2018) and group them according to their socioeconomic characteristics. Specifically, we cluster Census Tracts into three groups using a k-means clustering algorithm. We focus on four different characteristics: the median household income, the fraction of adults over 25 years with at least a bachelor's degree, the share of the population that is Black, and the median house price.[8]

Table 1 shows the results for the St. Louis MSA.[9] In the entire MSA, the share of Black households is 20%. The clustering algorithm makes two predominantly White neighborhoods with 9 and 7 percent Black share, respectively. The smaller of the two has the highest income, house

[7] There are several other empirical papers finding similar results such as Bayer et al. (2017); Bayer and McMillan (2005); Bayer et al. (2004); Boustan (2013); Card and Rothstein (2007).

[8] To compare different variables, we normalize each variable by the z-score. We exclude Census Tracts with missing values in characteristics.

[9] Appendix A.1 shows the results for other MSAs such as New York, Chicago, and Los Angeles.

Table 1: Neighborhood Characteristics in St Louis

	All	Cluster A	Cluster B	Cluster C
Population Share	1.00	0.17	0.62	0.21
Black Share	0.20	0.78	0.09	0.07
Income ($)	57,835	33,273	55,405	84,749
College Share of Adults	0.28	0.15	0.23	0.53
Median House Price ($)	171,749	82,699	150,060	307,244

Notes: K-means clustering for STL MSA.

prices, and share of college graduates. The third neighborhood has a larger Black population with a 78% Black share. The Black neighborhood is also relatively low-income, houses are cheaper, and it has a lower share of college graduates. Hence, a good description of the data is that there is one predominantly White and high-income cluster, one predominantly White and low-income cluster, and one predominantly Black and low-income cluster. We perform robustness exercises extending to four or five clusters. With four clusters, the neighborhoods look similar, but cluster B is split into two groups. With five clusters, neighborhood C is also split in two. Hence, we believe that focusing on three clusters is enough to capture the features of the data relevant to this paper while also helping to keep the model quantitatively tractable.

3.3 Black-White Wage Gap

Household income $y(r,e,s)$ depends on race, education, and skills. To discipline how income varies with its inputs we estimate versions of the Mincer (1974) equation. We use data from the 1997 National Longitudinal Survey of Youth (NLSY97), which contains wage and salary income, race, gender, age, usual hours and weeks worked, labor force participation, and skills measures. We use individuals' total wage and salary income at age 34 or 35.[10] We measure the skills with the Armed Services Vocational Aptitude Battery (ASVAB) test score. ASVAB maps well to the notion of skill in our model as both skill measures are measured after childhood inputs such as school quality and parental investment. This notion of skill is distinct from innate ability, which we model as being passed imperfectly from parent to child and as an input to skill production, but it is unobservable in the data.

[10] Respondents in the NLSY97 are interviewed every two years.

Table 2: Black-White College gap

	Wage
Below college	
White	1.00
Black	0.92
College or above	
White	1.71
Black	1.58
Return to skill	
χ	0.18

Notes: Estimated Black-White wage gap and college premium from Mincer regressions.

We estimate wage regressions of the type

$$\log(\text{wage}_i) = \alpha \text{ race}_i + \beta \text{ college}_i + \chi \log(\text{ASVAB}_i) + \delta \mathbf{X}_i + \varepsilon_i, \tag{2}$$

where race is either White or Black, college indicates if the education level is bachelor's degree or above, and ASVAB is the Armed Services Vocational Aptitude Battery score which we normalize, so it has a mean of one, and $\mathbf{X}_i$ is a control for gender.

Table 2 shows the predicted average hourly wage by race and education level for a man with average skills. We normalize the wage of White no college degree to one. First, White college graduates earn a wage premium relative to White non-graduates of 71%. Second, there is a sizable race penalty for non-college wages of 8%. Third, there is also a race penalty for college graduates. By construction, the race penalty for graduates is similar to non-graduates—around 8%—because we constrain the college premium to be the same for Blacks and Whites. Finally, the return to skills χ equals 0.18. This implies that we model $y(r,e,s) = w(r,e)\, s^{\chi}$. We call $w(r,e)$ the wage conditional on race and education per unit of skills and take it from Table 2. These estimates are in line with the empirical estimates in the literature (e.g., Heckman et al., 2006; Neal and Johnson, 1996).

Table 3: College attainment and skills

	Bachelor's Degree or More
Log(skills)	0.2071
	(0.009)
Constant	0.4901
	(0.010)
Observations	2,879
R^2	0.1741

3.4 College Graduates

We next study the relationship between college graduation, skills, and parental investment in the NLSY97. These relationships give us the information we use in the next section to inform the calibration of the skills production function, F^s, and the dispersion of taste shock for college.

As described above, we measure skill using the ASVAB score from the NLSY97. In the model, skill is distinct from innate ability. We map skills to the data on test scores from the NLSY97 and treat innate ability as unobservable. For parental investment, we look at the amount of money given to their children, including an imputed value for rent, following Abbott et al. (2019), which calculates intervivos transfers from questions on income transfers and allowances from parents from NLSY97.

First, we study the relationship between skills and college attainment. Table 3 shows the results of a regression of a dummy for whether the individual obtains at least a bachelor's degree in log skill, where log(ASVAB) is standardized so that both the mean and the standard deviation are 1. We find a positive relation between college attainment and skills, with a coefficient of 0.2071. Both the slope and the R-squared give information on how predictive skill is on college attainment, which we use to inform the calibration of the model.

Finally, we look at the covariance in the data between log ASVAB and log parental investment, which equals 0.1475. This moment provides information about the importance of parental investment, and we also use it as a target in the calibration.

4 Quantitative Evaluation

We now solve and take the model to the data. We solve the model globally using the equations derived in Section 2.[11] We use the estimates from Section 3 as inputs to calibrate the model. To validate the model, we show that it replicates several causal empirical estimates. In particular, the model is consistent with three studies regarding the causal effect of (i) growing in better neighborhoods, (ii) segregation on education, and (iii) neighborhood demographics on neighborhood choice.

4.1 Calibration

We calibrate the model using a simulated method of moments to match data moments for St. Louis MSA. Some parameters can be estimated "externally," while others must be estimated "internally" from the simulation of the model. Table 4 shows the parameters that are set externally based on the literature or on estimates from Section 3. Table 5 shows the internal parameters' values and the corresponding moments that were targeted in the data and the model. Although the calibration is done jointly, we briefly discuss how each moment helps us to identify the parameters of interest.

Preliminaries Agents live for two periods, so we set each period length equal to 40 years. We set the discount factor $\beta = 0.97^{40}$. For wages and returns from skills, we use the estimates from Table 2.

Ability While innate ability, a, is unobserved, we use data on test scores from the NLSY79 to measure the inter-generational persistence of skill. We estimate an AR(1) process for skills using the measure of mother and child's skills from the NLSY79 (see Appendix A.5). We assume that innate ability in the model follows an AR(1) process. We internally choose the persistence parameter, ρ_a, and the standard deviation of the ability process, σ_a, to match the persistence and the R-squared of the regression between parent and child's skills.

[11] We quantitatively show that the equilibrium seems to be unique despite not having a theoretical result.

Homophily There are some exogenous and some internally calibrated parameters. First, we set the bliss points on racial composition to $\gamma_B = 0.5$ and $\gamma_W = 0.9$, as in Banzhaf and Walsh (2013) and described in Section 3. Second, we calibrate φ_B and φ_W, which controls the importance of homophily in neighborhood choice, to match the share of Blacks who live in each neighborhood A, B, and C from Table 1.[12]

Housing We set the housing supply elasticity using the estimate for St. Louis from Saiz (2010), $\psi = 2.36$. Then, using data from ACS, we estimate the average annual property tax in St. Louis of \$2,474 while the average house value is \$201,688. We assume a return of 5% so the rent property tax is equal to $2,474/(201,688 \times 0.05) = 0.245$. Then, we internally calibrate the federal funding $\bar{Q}$ to match the mean local funding ratio of 0.58 from NCES data.[13]

Finally, we normalize the exogenous amenity $A_A = 1$ and internally calibrate A_B and A_C and the housing supply elasticities, η_A, η_B, and η_C, to match population and rent in each neighborhood from Table 1.

Moving For moving costs, we set m in the model equivalent to \$2,660 from the experimental evidence in Bergman et al. (2019) about movers in Seattle.[14] Finally, we internally calibrate the shape parameter κ to match the share of people who live in a different neighborhood than where they were born.[15]

Skills We choose the following functional form for skills formation

$$\log s = F^s(a, i, Q_n, X_n) = \theta_s + \theta_a \log(a) + \theta_i \log(i) + \theta_Q \log(Q_n) + \theta_X \log(X_n). \tag{3}$$

There are four inputs into the skills formation, two at the individual level and two at the neighborhood level. First skills are increasing in individual innate ability a and parental investment i with

[12] Note that by construction, the model gets the total share of Black households in the economy, so one of these moments is implied by the combination of other targeted moments (i.e., one of them is redundant).

[13] Local revenue share is calculated as local revenue divided by the sum of local, state, and federal revenue. We average the local revenue share across districts in the St. Louis MSA, weighting by the number of students.

[14] In the model, we normalize $w(W,L) = 1$, so 1 dollar is equivalent to \$57912, the average annual income of a White low-educated worker.

[15] See Appendix A.4 for the details on this estimation.

elasticities θ_a and θ_i, respectively. Second, skills are increasing with neighborhood spillovers X_n and neighborhood school funding Q_n with elasticities θ_X and θ_Q, respectively.

We set θ_s to normalize mean skills to one, both the model and the data. This normalization is crucial so that average wages are the same in the model and the data. For the individual variables of ability and investment, we look at cross-sectional moments across individuals. For θ_a we target the regression coefficient of education on skills from Table 3. For θ_i, we target the covariance between skills and parental investment. For neighborhood externalities, we look at moments at the neighborhood level. For θ_Q, we target the causal effect of school funding on educational attainment. Hyman (2017) finds that a 10% increase in school funding increases college attainment by 2.3 percentage points. We replicate the exercise in the model and target this moment. For θ_X, we target the ratio of the share of college graduates in neighborhood C relative to neighborhood A. This moment provides information on the relative spillovers in the richest neighborhood relative to the poorest one, which is valuable information regarding the role of externalities and the incentives for high-education adults to cluster in a given neighborhood.

Education We calibrate $\bar{c}$ to target the aggregate education level of 42%—the share of children born between 1978-1983 who grow up in St. Louis that complete at least a bachelor's degree (Chetty et al., 2018). This number is distinct from the 28% of *adults* living in St. Louis who have completed a college degree (Column 1 of Table 1). These numbers differ for two reasons: first, college attainment is higher for the younger cohort of adults than previous cohorts; second, the share of adults with a college degree is affected by in- and out-migration. Because we solve our model in the steady state, we impose that the share of children who go to college is equal to that of adults with high education. In the data, we scale up the college share of adults so that the aggregate level of college attainment is equal to 42%; this amounts to multiplying the college share from Table 1 by a factor of 1.49. Thus, neighborhoods A, B, and C in our targets have a college share of 22%, 34%, and 79%, respectively.

Note that we do not target the college share by neighborhood but the ratio of the college share in C to A, nor do we target the education level by race. Section 5 shows that the model matches these non-targeted moments well. Finally, we calibrate the shape parameter σ to match the R-squared

Table 4: Externally calibrated parameters

Parameter	Description	Value	Source
β	Discount factor	0.97^{40}	
γ_B	Bliss points for Black	0.50	Banzhaf and Walsh (2013)
γ_W	Bliss points for White	0.90	Banzhaf and Walsh (2013)
$w(B,L)$	Relative wage of Black, low education	0.92	Mincer regressions
$w(B,H)$	Relative wage of Black, high education	1.57	Mincer regressions
$w(W,H)$	Relative wage of White, high education	1.71	Mincer regressions
ψ	Housing supply elasticity	2.36	Saiz (2010)
τ	Rent property tax	0.21	ACS
m	Moving cost, money	$2660/34443.96$	Bergman et. al. (2019)

from the regression of education of skills from Table 3.

Table 5 shows that the model can replicate the 19 targeted moments. We next evaluate the performance of the model in matching non-calibrated moments.

4.2 Validation: Causal estimates in the data and model

We validate the model using credible estimates of causal effects from the literature to test the influence of race and homophily on college and neighborhood choices. The first two exercises show how college attainment depends on the neighborhood in which children grow up and segregation. The third exercise examines the role of the Black share in neighborhood choice. The model is consistent with these three non-targeted causal effects that are behind the key mechanisms of the model.

Moving to Opportunity First, we validate the model using estimates from the literature to test the influence of neighborhoods on college attainment. Chetty et al. (2016) studied the Moving to Opportunity (MTO) experiment, which provided housing vouchers to low-income families living in public housing in low-income neighborhoods in Baltimore, Boston, Chicago, Los Angeles, and New York. Families were randomized into two groups. Those in the experimental group received housing vouchers that could be used to subsidize rent for private market housing units located in Census tracts with poverty rates below 10 percent. Members of the control group received no vouchers through this experiment. Chetty et al. (2016) finds that moving through MTO increased

Table 5: Internally calibrated parameters

Parameter	Description	Value	Moment	Data	Model
Neighborhoods					
A_A	Amenity in A	1.0000	Population neighborhood A	0.1700	0.1778
A_B	Amenity in B	1.1812	Population neighborhood B	0.6200	0.5608
A_C	Amenity in C	1.4461	Population neighborhood C	0.2100	0.2614
η_A	Housing supply in A	25.3023	Rent neighborhood A	0.1200	0.1224
η_B	Housing supply in B	22.6165	Rent neighborhood B	0.2178	0.2088
η_C	Housing supply in C	1.4118	Rent neighborhood C	0.4460	0.4894
κ	Shape parameter for location	0.1630	Neighborhood flows	0.3500	0.3409
Homophily: Importance of bliss point					
φ_W	White	0.5641	Black share neighborhood A	0.7780	0.6805
φ_B	Black	0.9560	Black share neighborhood B	0.0850	0.1032
			Black share neighborhood C	0.0710	0.0799
Skill production					
θ_c	Constant term	1.0526	Mean skills	1.0000	1.1600
θ_a	Ability	0.2948	Reg coef of educ on skills	0.2100	0.2284
θ_i	Investment	0.1172	Covariance (s,i)	0.1475	0.1012
θ_X	Spillovers	0.2575	Ratio college share neighborhood C to A	3.5333	4.8472
θ_Q	School quality	0.3305	Causal effect school funding	0.0230	0.0189
Ability					
ρ_a	Persistence ability	0.5487	Reg coef child on parent ability	0.5100	0.5156
σ_a	Std. dev. ability	1.2348	R^2 child on parent ability	0.2600	0.2659
Education					
$\bar{c}$	Education cost level	2.0811	Educational probability	0.4166	0.4169
$\bar{Q}$	Federal and state funding	0.0366	Local funding ratio	0.5800	0.6072
σ	Shape parameter education	0.3722	R^2 education choice	0.1740	0.2100

college attainment and earnings.

We simulate a policy similar to the MTO voucher program in our model. From the steady state, we evaluate a scenario in which the government provides low-education families that live in the lowest-income neighborhood (neighborhood A) with a voucher that subsidizes rent for housing if they move to either B or C. The subsidy in our simulation covers 50 percent of the rent differential between the previous and the new neighborhood. Note that this validation exercise also assumes that rental prices and other equilibrium quantities (such as school funding or the share of college graduates) do not change. These assumptions align with the idea that relatively few families move in a small-scale RCT such as MTO, implying that neighborhood characteristics are unaffected.

Voucher-eligible families make two critical choices in our model. First, they must decide whether to take up the voucher and relocate to the more advantaged neighborhood. Panel I of Table 6 shows that 39% of households opt for the voucher in our simulation, while in the data, it is a bit higher, between 46 and 49%. Second, households also change their investment and education

Table 6: Validation: Replicating empirical causal effects

	Data	Model
I. Moving to Opportunity		
Takeup rate (%)	[46.0 , 49.3]	39.2
Δ College attainment, treatment-on-the-treated (%)	[2.9 , 7.6]	13.1
Δ College attainment, intent-to-treat (%)	[1.4 , 3.7]	5.1
II. Segregation		
Δ College attainment, White	[-.290, .012]	0.112
Δ College attainment, Black	[-.520, -.078]	-0.278
III. Homophily		
Δ Probability for White, college graduate	[-0.36, -0.27]	-0.13
Δ Probability for White, non-college graduate	[-0.43, -0.29]	-0.43
Δ Probability for Black, college graduate	[0.74, 0.95]	0.79
Δ Probability for Black, non-college graduate	[0.31, 0.43]	0.33

Notes: Data estimates from Ananat (2011); Caetano and Maheshri (2019); Chetty et al. (2016). Data shows one standard deviation confidence intervals.

choice. We find that for college graduation, the treatment-on-the-treated (TOT) is 13.1 percentage points, and the intent-to-treat (ITT) is 5.1 percentage points, meaning that college attainment increases by 5.1 percentage points for families offered the voucher regardless of whether they used it. Note that the TOT and the ITT are close to the upper bound of the data estimates' one standard deviation confidence interval. Reassuringly, we see that the simulation generates very close results to the MTO results from Chetty et al. (2016).

Segregation Second, the model is consistent with estimates of the causal effect of segregation on educational attainment. Ananat (2011) uses exogenous variation in a city's susceptibility to segregation (measured with dissimilarity index) from the historical layout of train tracks to measure the causal impact of segregation on college attainment. In the model, to test the impact of segregation, we set the parameters governing homophily to zero, $\varphi_B = \varphi_W = 0$, and solve the general equilibrium. Then, we calculate the change in educational attainment per change in the dissimilarity index, as in the data.

Panel III of Table 6 shows that the model is consistent with the reduced-form causal estimates: more segregation implies lower educational attainment for Blacks and higher for Whites, with more substantial effects for Blacks than for Whites.[16]

Homophily Third, we show that the model is consistent with empirical estimates of the causal effect of the neighborhood's Black share on neighborhood choice. Caetano and Maheshri (2019) estimate a dynamic discrete choice model of neighborhood choice. This study provides evidence of the marginal effect of an increase in a neighborhood's Black share on the valuation of the neighborhood for Black and White households. We compare the change in the probability group g (i.e., Black or White, and college or non-college graduates) chooses neighborhood j (conditional on moving) in response to a 1 percentage point increase in Black share.[17]

The bottom panel of Table 6 shows that the model is consistent with the causal effects of the Black share on neighborhood choices. For White households, the model predicts a response slightly smaller (in absolute terms) or close to the lower bound of the one standard deviation confidence interval. For Black households, the model's prediction falls within the one standard deviation confidence interval.

5 Sources of College Attainment Gap: Wage gap and Homophily

In the model, college attainment is 46% and 24% for White and Black households, respectively. In the data, college attainment is 47% and 19% for White and Black households, respectively. Hence, the presence of (i) the wage gap—which we estimate directly from the data—and (ii) homophily—which we calibrate to generate the racial composition of neighborhoods—generate a college gap of 22 percentage points, explaining about 80% of the college gap in the data. Notably, the calibration targets the aggregate college attainment but not college attainment by race. The other 20% that the

[16] In robustness exercises, we verify that the model predictions are similar if we reduce φ_B and φ_W by 50% instead of 100%.

[17] For the data target, in Caetano and Maheshri (2019), the marginal effect of the Black share is given by the equation $\frac{\partial P_{gj}}{\partial s_j} = \beta_g P_{gj}(1-P_{gj})$. We use their estimates of β_g combined with the neighborhood choice probabilities (for movers) for each neighborhood from our model. In the model, we exogenously increase the Black share in the neighborhood j by 1 percentage point; resolve the value functions holding spillover, school funding, and rents constant; compute the change in the probability that a mover chooses neighborhood j; and repeat for each neighborhood, 1 at a time. The model moment corresponds to the average over the three neighborhoods. In robustness exercises, we study the models' asymmetries and non-linearities by increasing or decreasing the black share by different amounts.

model does not capture is likely coming from other forces that are not present in the model (e.g., differential access to credit).

This section studies the drivers of the college gap. First, we show how education and neighborhood choices work in the model. Then, we examine the importance of the wage gap and homophily for the college attainment gap through a series of counterfactual exercises.

5.1 Segregation and Education Attainment

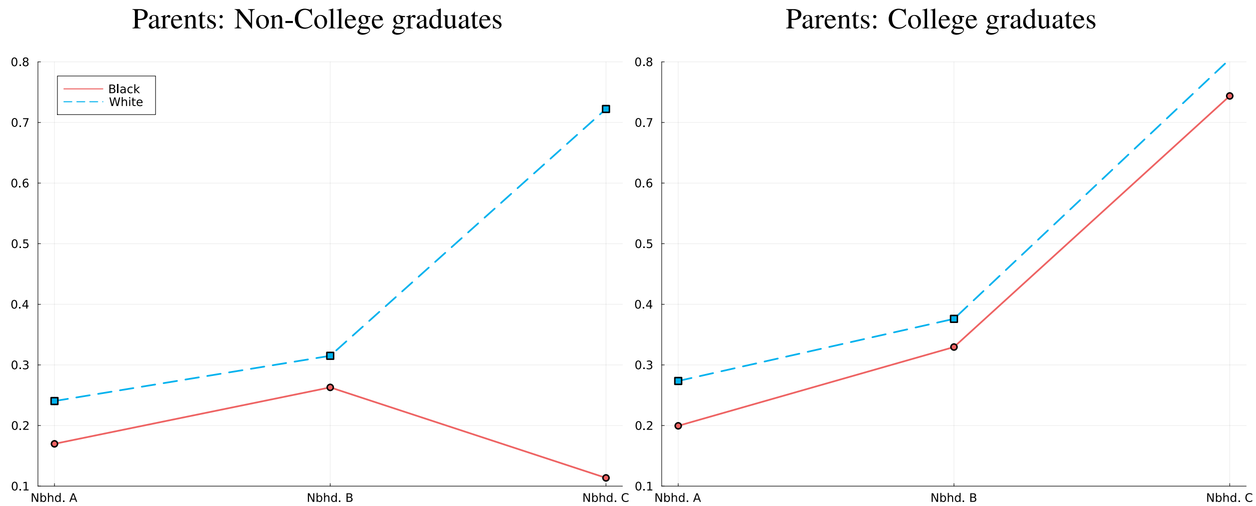
We first examine education and neighborhood choices in the baseline model. Figure 2 shows the probability of becoming a college graduate for a child with median innate ability and skills as a function of the other state variables: parental education, race, and neighborhood. The solid red line in the left panel shows the education probability of a Black child with non-college parents as a function of the neighborhood in which they grow up. The blue dashed line shows the education probability for a White child with non-college parents. There is a striking difference in college attainment across races, although we are comparing children with the same innate ability. While White children have a college attainment probability between 25 and 70%, it is less than 25% for Black children.[18]

The right panel shows the college attainment probability for Black and White children of college-graduate parents. Again, White children have a higher probability of going to college than Black ones. Having college parents increases the probability of college attainment for Black and White children. Interestingly, the college attainment gap by race narrows for children of college parents.

A breakdown of the race and education composition of each neighborhood's residents is in Table 7. Most neighborhood A residents are Black, and the vast majority are non-college graduates. Residents of Neighborhoods B and C are primarily White households with a higher share of college households, particularly for neighborhood C. Overall, the basic features of our three neighborhoods match up well with the three clusters we identify in the data for St. Louis. Importantly,

[18]For children of Non-college Black parents in Neighborhood C the probability of going to college is very small. The reason is that these poor families choose to live in Neighborhood C primarily due to their preference shock, and as a result they are left with very low resources for investment which is an important input for skills and education.

Figure 2: College attainment



Notes: College attainment for a child's with median innate ability and skills, as a function of parent's education, race, and neighborhood.

neighborhood C has the highest college share meaning it also has the highest spillover effect for children who grow up there.

Table 7: Neighborhood Demographics

	Neighborhood A	Neighborhood B	Neighborhood C
Black non-college	0.58	0.07	0.03
Black college	0.10	0.03	0.05
White non-college	0.24	0.64	0.10
White college	0.07	0.25	0.82
Total	1.00	1.00	1.00

Notes: The table shows the composition of each neighborhood by race and education level.

Neighborhood choices are also very different for Black and White households. Figure 3 shows the probability of going to each neighborhood for an agent with median innate ability and skills as a function of the initial neighborhood, race, and education. Examining the figures reveals two patterns. First, Black households have a much higher probability than White households of living in neighborhood A, and this probability is more significant for children of non-college parents than for children of college-graduate parents. Second, the probability of going to either B or C is larger for White than for Black households. The probability of going to C is almost zero for Black non-college households, although neighborhood C has the highest spillover and school quality.

5.2 Removing all differences across races

In the model, there are two reasons why we observe differences by race: (i) the wage gap and (ii) homophily. We now remove both differences. We give low- and high-educated Black households the same wage as their White counterparts, conditional on skill. In the model, income is a function of skill and the wage, which depends on race and education, $y(r,e,s) = w(r,e)\ s^{\chi}$. When we equalize wages by race, we set $w(B,e) = w(W,e)$ for each education level, low or high. This means average income y could still differ by race if the average skill differs. Equalizing wages entails increasing wages, $w(B,l)$, for Blacks without college from 0.92 to 1.00 and increasing wages, $w(B,h)$, from 1.57 to 1.71 for Blacks with college. We also set the parameters governing the importance of homophily to zero, $\varphi_B = \varphi_W = 0$.

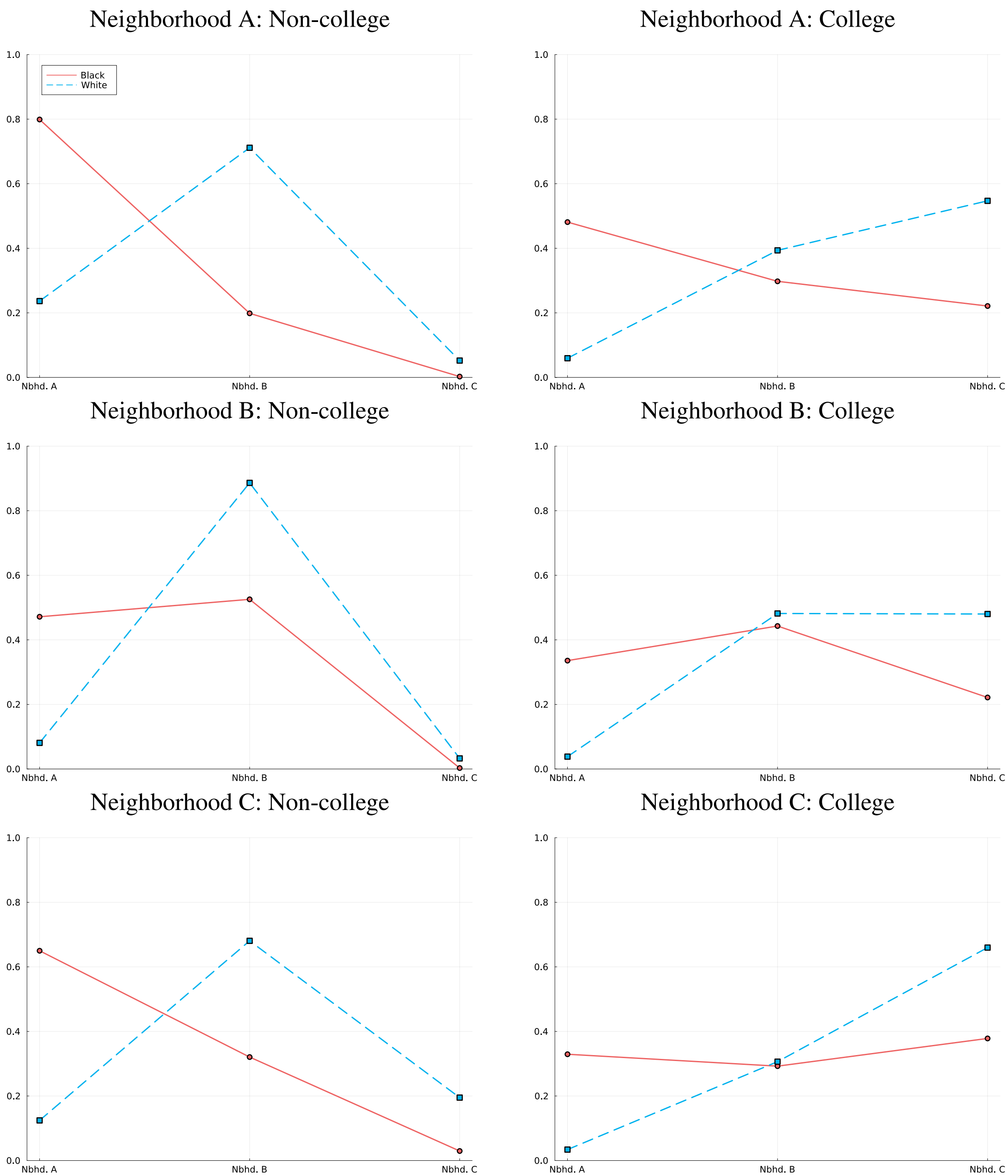
In this scenario, we have removed all differences between Black and White households, so unsurprisingly, the college attainment gap is equal to zero: Black and White households have the same college attainment of 40% (Table 8, column 5). The neighborhood racial shares are identical, and the remaining disparities in the educational composition of neighborhoods come only from higher-income households sorting into neighborhoods with better exogenous amenities. Moreover, parents of both races make the same investment decisions, which means that their children face the same probability of going to college—any differences in intergenerational mobility come only from the parent's education and skills but do not depend on race.

Next, we evaluate the role of the wage gap and homophily in isolation and their economic consequences.

5.3 Removing the wage gap

We now assess the role of the Black-White wage gap. We give low- and high-educated Black households the same wage as their White counterparts, conditional on education and skill. We compare the new steady-state equilibrium with the baseline in Table 8.

Figure 3: Neighborhood choice



Notes: Neighborhood choice for median innate ability and skill child as a function of parent's education, race, and neighborhood.

Table 8: The Role of the Wage Gap and Homophily

	Benchmark (1)	No wage gap (2)	Equal spillovers (3)	No homophily (4)	Both (5)
College attainment:					
All	0.42	0.41	0.39	0.40	0.40
White	0.46	0.44	0.39	0.41	0.40
Black	0.24	0.29	0.39	0.37	0.40
College gap	0.22	0.15	0.00	0.04	0.00
Δ gap p.p.	–	0.07	0.22	0.18	0.22
Segregation Index	0.53	0.42	0.02	0.07	0.00

Notes: "No wage gap" here means that wages for black and white are set equal conditional on education and skill. The column equal spillovers also has equal wages, and the column equal amenities also has both equal wages and equal spillovers. No homophily means $\varphi_B = \varphi_W = 0$.

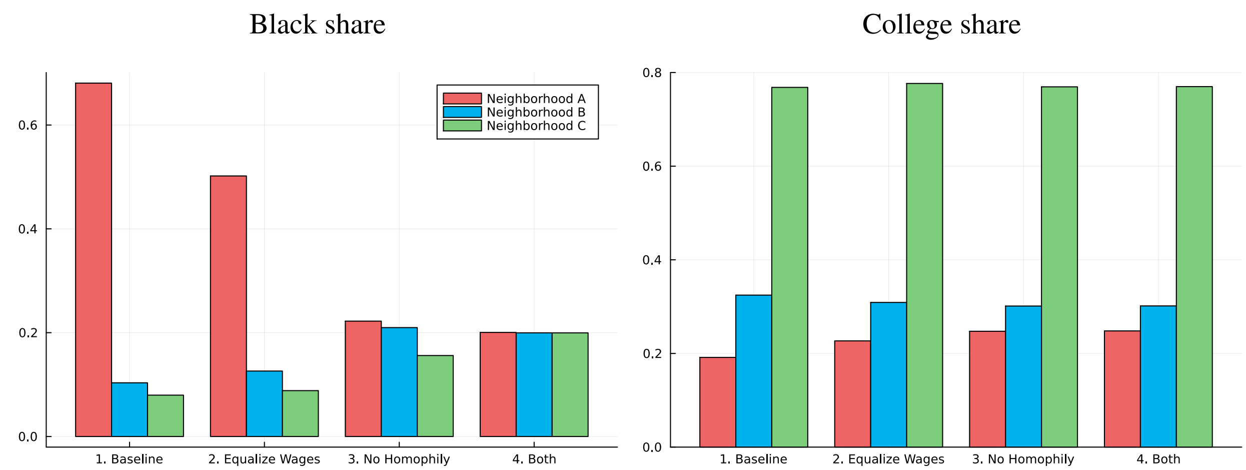
The college gap reduces by 7 percentage points (from 22% to 15%) when we remove the wage gap (Table 8, column 2). The education of White households is almost unchanged—it reduces from 46% to 44%. Instead, for Black households, college attainment increased from 24% to 29%.

Interestingly, a significant gap remains when the only difference across races is the homophily term. We next evaluate different endogenous equilibrium outcomes of the model to understand the sources and effects of this counterfactual exercise.

Figure 4 shows who lives in each neighborhood under each counterfactual experiment. Equalizing wages affect the racial composition of the neighborhoods, with a decrease in the Black share of neighborhood A. It also affects the college share of neighborhoods, primarily increasing the college share of neighborhood A. According to the right panel, there was a 68.8 percentage points gap in the share of college graduate workers between Neighborhoods A and C, which closes slightly to 67.2 percentage points when we remove the wage gap. This comes from the fact that more Black workers go to college in the new equilibrium. These workers comprise many of the new residents of Neighborhood C. Thus, we find that overall removing the wage gap mildly affect segregation by race.

If closing the wage gap has modest impacts on neighborhood choice, then what do Black households do with their higher wages? Instead of using them to move to more expensive neigh-

Figure 4: Black and College share: comparison among counterfactual economies



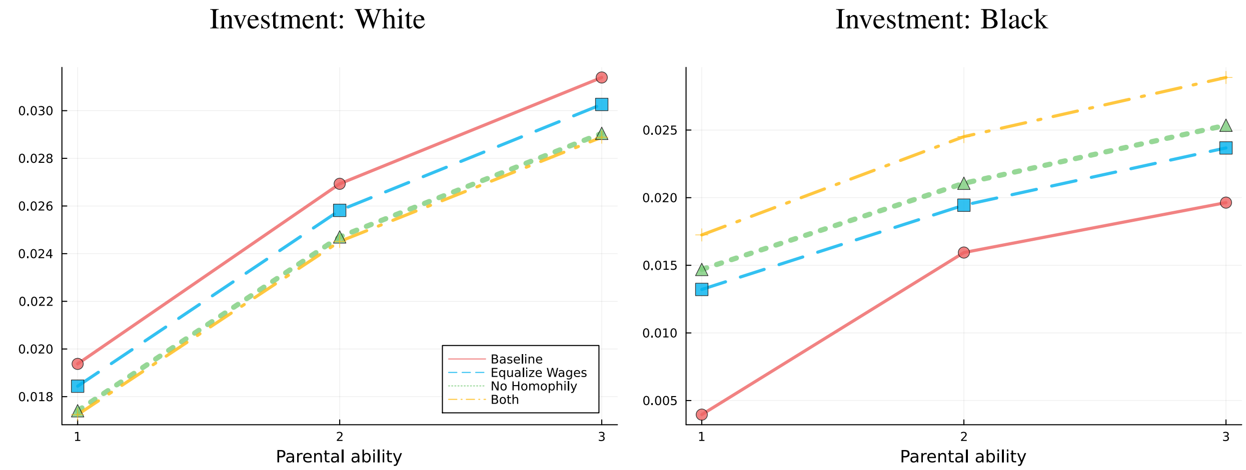
Notes: Black and college share by neighborhoods in the baseline and counterfactual economies.

borhoods, they invest the money into their children. In Figure 5, we plot the average investment in the model conditional on race and innate ability, for average parental education and skills. Moving from the baseline in the solid red line to the equalized wages counterfactual in the dashed blue line increases Black investment. Therefore, these households make up for the fact that the neighborhood spillovers do not change by investing in their children. In contrast, White households do not significantly change their investment because their wages have not changed, and their neighborhoods have barely changed. If anything, White households reduce a bit their investment.

This increased investment drives the improved educational attainment for Black workers in this counterfactual. Figure 6 compares intergenerational mobility under each of our scenarios. The middle panel indicates considerable improvements in intergenerational mobility among Black children: the gap between the red and the blue bars decreases from 21.8 percentage points in the benchmark to 16.2 percentage points for college workers. For non-college, the gap reduces from 13.3 percentage points to 8.4 percentage points. Hence, we see an equalization of intergenerational mobility across races both for children of non-college and college graduates.

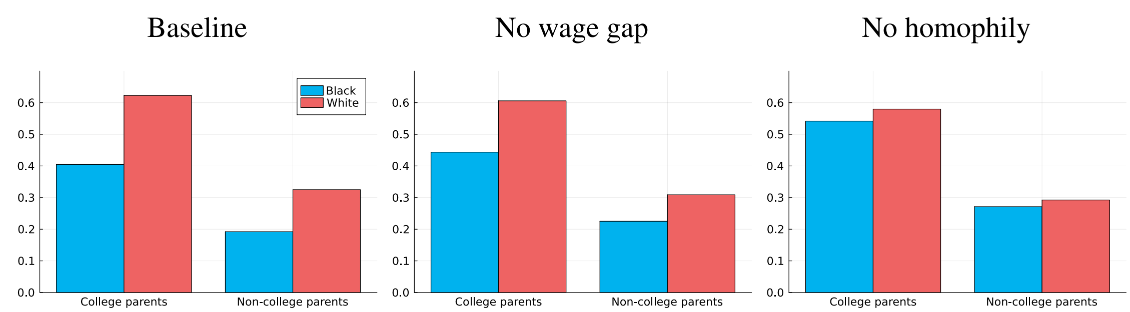
Spillovers Even when we equalize wages by race, a substantial amount of the college attainment gap remains due to neighborhood effects. In the third column of Table 8 we equalize wages and spillovers across neighborhoods. Specifically, we set school funding, Q_n, and spillovers, X_n, equal

Figure 5: Parental investment: comparison among counterfactuals



Notes: Parental investment for each race and innate ability, and average parental education and skills, of the baseline and counterfactual economies.

Figure 6: Intergenerational mobility: comparison among counterfactuals



Notes: Each bar shows the probability of going to college conditional on parental education and race for different economies.

to the average across the three neighborhoods. We find that the college gap completely closes when we close the racial wage gap and equalize the different spillovers across neighborhoods. Note that in this counterfactual, we fully close the college gap, although households still care about the neighborhood's racial composition.

These results imply that even when wages are equalized, substantial racial inequality remains in school funding and spillovers due to homophily, which maintains segregation and the gap in college attainment. However, the college gap would be fully closed if the spillovers were also equalized across neighborhoods.

5.4 Race-blind Counterfactual

The racial gap that remains when we equalize wages is a consequence of household preferences—White households, who tend to be college educated, cluster in Neighborhoods B and C, resulting in lower spillovers in the majority Black Neighborhood A. In this next counterfactual, we perform a "race-blind" counterfactual, in which households are unresponsive to the racial composition of their neighbors, but we leave the Black-White wage gap in place. We remove homophily from preferences by setting $\varphi_B = \varphi_W = 0$. This makes the racial makeup of the neighborhood irrelevant in utility.

The college attainment gap reduces by 18 percentage points when both White and Black households are race-blind (Table 8, column 4). This reduction results from two forces: (a) White households reduce their education—from 46% to 41%, and (b) Black households increase their education—from 24% to 37% percent.

Figure 4 shows that without homophily, the neighborhoods become more similar, particularly in terms of the racial composition. Nevertheless, some amount of segregation by race remains. This is due to households segregating themselves by income which, in turn, differs by race. High-income households are more willing to pay for the exogenously given amenities that are unequal across neighborhoods. Since White households have higher incomes than Black households, they are more likely to choose to live in neighborhood C, the high-rent and high-amenity neighborhood.

Homophily also affects educational attainment for Black children through an increase in investment of Black parents, as in the wage-gap counterfactual (see in Figure 5). Black parents now expect their children to want to live in high-rent neighborhoods as adults to enjoy the exogenous amenities. As such, they will want to be college graduates to afford the more expensive rent while maintaining their level of consumption.

Moreover, because many Black households now live in better neighborhoods than before, their chances of going to college increase, as seen in the third panel of Figure 6. The gap in intergenerational mobility closes for both non-college and college households. It goes from 21.8 percentage points to 3.8 percentage points for college households. For non-college households, the gap goes from 13.3 percentage points to 2.1 percentage points. Note that this change is more significant than the equalization that occurred when the wage gap was removed.

The closing of the intergenerational mobility gap in this experiment is also partially driven by a slight decline in the probability of going to college for White children, due to equilibrium changes in rents and spillovers. In the race-blind counterfactual, White households are more willing to live in Neighborhood A that has lower rents, so there is a lower demand for college for White households. Therefore, our takeaway from this experiment is that although there are considerable differences in neighborhood characteristics for both Black and White children, most of the gains from removing homophily accrue to Black children, particularly those whose parents do not have college degrees.

6 Policy: Equalize School Funding

In the baseline economy, school funding in neighborhood C is more than 2 times larger than in neighborhood A. In this section we explore the effects of equalizing school funding across neighborhoods. In particular, we examine the effects of a policy change in which the government collects property taxes in each neighborhood, but then distributes them equally across districts.

First, we study the partial equilibrium effects of equalizing school funding. We keep the household neighborhood, investment decisions, and prices constant as in the baseline economy. The change in college attainment is driven only by the direct effect of the change in school funding. Table 9 shows that in this scenario, Black households increase their college attainment while White households decrease it very slightly. As a result, the college gap reduces from 22 to 18 percentage points.

We then consider the policy in general equilibrium, allowing households to re-optimize and prices to adjust to clear markets. In this scenario, the increase in the college attainment of Blacks and the decrease in the college attainment of Whites are more pronounced than in partial equilibrium, with the college gap reduced by more than 10 percentage points. There are two reasons for this effect. First, in general equilibrium, the spillovers increase in neighborhood A, the predominantly Black neighborhood, and decrease in neighborhood C, the predominantly White neighborhood, see Table 10. As a result of spillover equalization across neighborhoods, Black households increase their college attainment while White households decrease theirs. Second, there is a significant increase in investment of Black parents, particularly in neighborhood A, which further increases their college attainment. We conclude that the general equilibrium response of parents will amplify the direct effect of the change in school funding on the Black-White college attainment gap.

Table 9: Equalize School Funding

	College Attainment		
	Black	White	Gap
Baseline	23.98	46.11	22.13
Partial Equilibrium	26.92	45.25	18.33
General Equilibrium	29.63	41.56	11.93

Notes: The table shows the effects on college attainment of equalizing school funding.

Finally, in Table 11, we evaluate the welfare consequences of this policy intervention. The effects vary by household, with some winners and some losers depending on their characteristics. First, in neighborhood A, everyone benefits from better schools and the decline in the Black share as it becomes closer to the bliss point for both Black and White households. Second, in neighborhood B, Black households like the increase in the Black share, although the change is very small, and they like the increase in school funding and college share. Quantitatively, as these changes are very small, everyone is better off, but the welfare gains are modest. Third, in neighborhood C, everyone dislikes the decline in spillovers (about 5 percentage points decline in college share), but everyone benefits from the decline in rents. As a result, the welfare effect of the intervention in neighborhood C depends on the race and education status of the household, with some winners (e.g., the White non-college) and some losers (e.g., the White college households).

These results imply that the policy reform is not a Pareto improvement. While the majority of households prefer the new equilibrium, there are a few households that prefer the baseline. Given that most households prefer the new equilibrium, a social planner using a typical welfare aggregation function (e.g., Utilitarian) would recommend the reform.

Table 10: Equalize School Funding: Neighborhoods effects

	Nbhd A	Nbhd B	Nbhd C
School Quality			
baseline	0.0666	0.0877	0.1565
equal funding	0.0963	0.0963	0.0963
Population			
baseline	0.1778	0.5608	0.2614
equal funding	0.2390	0.5433	0.2177
College Share			
baseline	0.1790	0.2842	0.8667
equal funding	0.2329	0.2947	0.8095
Black Share			
baseline	0.6807	0.1034	0.0797
equal funding	0.5300	0.1094	0.0632
Rent			
baseline	0.1224	0.2088	0.4894
equal funding	0.1387	0.2060	0.4529
Investment, Black			
baseline	0.0132	0.0179	0.0187
equal funding	0.0177	0.0204	0.0293
Investment, White			
baseline	0.0215	0.0244	0.0324
equal funding	0.0249	0.0252	0.0343

Notes: The table shows the effects on neighborhood characteristics of equalizing school funding.

7 Conclusion

There is growing empirical evidence that the neighborhood in which a child grows up substantially impacts a range of adult outcomes, including college attainment. At the same time, there is ample empirical evidence of a Black-White wage gap and the impact of race on neighborhood choice. These empirical patterns suggest that exposure to neighborhoods with good schools and significant spillovers may systematically differ by race and be partly responsible for the racial gap in adult outcomes.

To examine these issues, we develop an overlapping-generation spatial-equilibrium model that incorporates race in two ways. First, Black households are subject to a Black-White wage gap. Second, households have preferences over the racial composition of the neighborhood in which

Table 11: Equalize School Funding: Welfare effects

	Black		White	
Cons. equivalent, (pct.)	Non-College	College	Non-College	College
Average	2.45	2.10	1.64	-0.11
Nbhd A	2.60	4.13	8.85	9.94
Nbhd B	1.88	1.77	1.32	1.19
Nbhd C	4.51	-1.37	2.76	-1.02

Notes: The table shows the welfare effects in consumption equivalent units of equalizing school funding for stayers, i.e., those with $n_0 = n$.

they live. Households in the model choose where to live and how much to invest in their child's education, affecting whether the child goes to college and receives higher wages as an adult.

We calibrate the model to match the neighborhood characteristics of the St. Louis metro area. We find that the presence of the Black-White wage gap and homophily generate a college gap of 22 percentage points—about 80% of the college gap in the data. We also find that removing the racial wage gap helps improve Black workers' educational attainment. However, homophily in preferences is also essential for reducing neighborhood segregation and improving access to neighborhoods with high spillovers and school funding, both necessary inputs into skill and college attainment. A policy experiment equalizing school funding across neighborhoods reduces the college attainment gap by 10 percentage points.

References

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

Aliprantis, D. and Carroll, D. R. (2018). Neighborhood dynamics and the distribution of opportunity. *Quantitative Economics*, 9(1):247–303.

Aliprantis, D., Carroll, D. R., and Young, E. R. (2019). What Explains Neighborhood Sorting by Income and Race? Working Papers 201808R, Federal Reserve Bank of Cleveland.

Almagro, M. and Domínguez-Iino, T. (2020). *Location Sorting and Endogenous Amenities: Evidence from Amsterdam*. PhD thesis, New York University.

Ananat, E. O. (2011). The wrong side(s) of the tracks: The causal effects of racial segregation on urban poverty and inequality. *American Economic Journal: Applied Economics*, 3(2):34–66.

Badel, A. (2015). A Racial Inequality Trap. Working papers, Federal Reserve Bank of St. Louis.

Banzhaf, H. S. and Walsh, R. P. (2013). Segregation and tiebout sorting: The link between place-based investments and neighborhood tipping. *Journal of Urban Economics*, 74:83–98.

Bayer, P., Casey, M., Ferreira, F., and McMillan, R. (2017). Racial and ethnic price differentials in the housing market. *Journal of Urban Economics*, 102(C):91–105.

Bayer, P., Charles, K. K., and Park, J. (2021). Separate and unequal: Race and the geography of the american housing market. Technical report, Mimeo., Duke University.

Bayer, P. and McMillan, R. (2005). Racial Sorting and Neighborhood Quality. NBER Working Papers 11813, National Bureau of Economic Research, Inc.

Bayer, P., McMillan, R., and Rueben, K. (2004). Residential Segregation in General Equilibrium. Working Papers 885, Economic Growth Center, Yale University.

Benabou, R. (1994). Education, Income Distribution and Growth: The Local Connection. NBER Working Papers 4798, National Bureau of Economic Research, Inc.

Bergman, P., Chetty, R., DeLuca, S., Hendren, N., Katz, L. F., and Palmer, C. (2019). Creating moves to opportunity: Experimental evidence on barriers to neighborhood choice. Working Paper 26164, National Bureau of Economic Research.

Billings, S. B., Deming, D. J., and Rockoff, J. (2013). School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg *. *The Quarterly Journal of Economics*, 129(1):435–476.

Boustan, L. P. (2013). Racial residential segregation in american cities. Technical report, National Bureau of Economic Research.

Brock, W. A. and Durlauf, S. N. (1995). Discrete Choice with Social Interactions I: Theory. NBER Working Papers 5291, National Bureau of Economic Research, Inc.

Caetano, G. and Maheshri, V. (2019). A unified empirical framework to study segregation. Technical report, Working paper.

Card, D., Mas, A., and Rothstein, J. (2008). Tipping and the Dynamics of Segregation. *The Quarterly Journal of Economics*, 123(1):177–218.

Card, D. and Rothstein, J. (2007). Racial segregation and the black–white test score gap. *Journal of Public Economics*, 91(11):2158–2184.

Chetty, R., Friedman, J. N., Hendren, N., Jones, M. R., and Porter, S. R. (2018). The opportunity atlas: Mapping the childhood roots of social mobility. Working Paper 25147, National Bureau of Economic Research.

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

Chetty, R., Hendren, N., and Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *The American Economic Review*, 106(4):855–902.

Chyn, E. and Daruich, D. (2022). An equilibrium analysis of the effects of neighborhood-based interventions on children. Technical report, National Bureau of Economic Research.

Couture, V., Gaubert, C., Handbury, J., and Hurst, E. (2019). Income Growth and the Distributional Effects of Urban Spatial Sorting. NBER Working Papers 26142, National Bureau of Economic Research, Inc.

Cutler, D. M. and Glaeser, E. L. (1997). Are Ghettos Good or Bad?*. *The Quarterly Journal of Economics*, 112(3):827–872.

Cutler, D. M., Glaeser, E. L., and Vigdor, J. L. (1999). The Rise and Decline of the American Ghetto. *Journal of Political Economy*, 107(3):455–506.

Dawkins, C. J. (2005). Tiebout choice and residential segregation by race in us metropolitan areas, 1980–2000. *Regional Science and Urban Economics*, 35(6):734–755.

Derenoncourt, E. (2022). Can you move to opportunity? evidence from the great migration. *American Economic Review*, 112(2):369–408.

Echenique, F. and Fryer, Roland G., J. (2007). A Measure of Segregation Based on Social Interactions*. *The Quarterly Journal of Economics*, 122(2):441–485.

Eckert, F. and Kleineberg, T. (2019). Can We Save the American Dream? A Dynamic General Equilibrium Analysis of the Effects of School Financing on Local Opportunities. 2019 Meeting Papers 1197, Society for Economic Dynamics.

Farley, R., Fielding, E. L., and Krysan, M. (1997). The residential preferences of blacks and whites: A four-metropolis analysis. *Housing Policy Debate*, 8:763–800.

Farley, R., Schuman, H., Bianchi, S., Colasanto, D., and Hatchett, S. (1978). “chocolate city, vanilla suburbs:” will the trend toward racially separate communities continue? *Social Science Research*, 7(4):319–344.

Fernandez, R. and Rogerson, R. (1996). Income distribution, communities, and the quality of public education. *The Quarterly Journal of Economics*, 111(1):135–164.

Fogli, A. and Guerrieri, V. (2019). The end of the american dream? inequality and segregation in us cities. Technical report, National Bureau of Economic Research.

Heckman, J. J., Lochner, L. J., and Todd, P. E. (2006). Earnings functions, rates of return and treatment effects: The mincer equation and beyond. *Handbook of the Economics of Education*, 1:307–458.

Hoelzlein, M. (2020). *Two-Sided Sorting and Spatial Inequality in Cities*. PhD thesis, University of California, Berkeley.

Hyman, J. (2017). Does money matter in the long run? effects of school spending on educational attainment. *American Economic Journal: Economic Policy*, 9(4):256–80.

Johnson, R. C. (2011). Long-run impacts of school desegregation & school quality on adult attainments. Working Paper 16664, National Bureau of Economic Research.

Krysan, M. and Farley, R. (2002). The residential preferences of blacks: Do they explain persistent segregation? *Social forces*, 80(3):937–980.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. National Bureau of Economic Research, Inc.

Monarrez, T. and Schonholzer, D. (2021). Dividing lines: Racial segregation and local government boundaries. Technical report, Working paper.

Neal, D. A. and Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. *Journal of Political Economy*, 104(5):869–895.

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296.

Schelling, T. C. (1969). Models of Segregation. *American Economic Review*, 59(2):488–493.

Schelling, T. C. (1971). Dynamic models of segregation. *The Journal of Mathematical Sociology*, 1(2):143–186.

Sethi, R. and Somanathan, R. (2004). Inequality and segregation. *Journal of Political Economy*, 112(6):1296–1321.

Sethi, R. and Somanathan, R. (2009). Racial inequality and segregation measures: Some evidence from the 2000 census. *The Review of Black Political Economy*, 36(2):79–91.

Thompson, O. (2021). Human capital and black-white earnings gaps, 1996–2017. Working paper, National Bureau of Economic Research.

Zheng, A. and Graham, J. (2022). Public education inequality and intergenerational mobility. *American Economic Journal: Macroeconomics*, 14(3):250–82.

Appendix

A Data

A.1 Neighborhoods

Table 12 shows the neighborhood clustering for St. Louis, Chicago, NY, and LA.

Table 12: Neighborhood Characteristics in different MSA

	All	Cluster 1	Cluster 2	Cluster 3
St Louis				
Population Share	1.00	0.17	0.54	0.29
black Share	0.20	0.79	0.08	0.08
Income	57,835.32	33,328.16	53,304.46	79,888.26
College Graduates	0.28	0.15	0.20	0.48
Rent	841.81	801.63	753.55	1025.72
Chicago				
Population Share	1.00	0.14	0.55	0.31
black Share	0.18	0.83	0.08	0.06
Income	63,277.11	36,322.53	55,379.09	89,654.80
College Graduates	0.31	0.16	0.22	0.54
Rent	1091.76	899.61	966.72	1402.19
NY				
Population Share	1.00	0.44	0.12	0.44
black Share	0.23	0.45	0.05	0.05
Income	58,444.49	43,389.66	105,275.60	61,224.71
College Graduates	0.30	0.18	0.72	0.32
Rent	1406.77	1193.68	2164.33	1420.95
LA				
Population Share	1.00	0.50	0.32	0.18
black Share	0.09	0.12	0.07	0.04
Income	58,243.79	42,001.56	64,782.33	91,596.32
College Graduates	0.25	0.11	0.31	0.56
Rent	1387.74	1173.93	1426.30	1910.55

Notes: .

A.2 Black-White wage gap by education, controlling for occupation

Table 13 shows the Mincer regressions for different decades, controlling for occupation.

Table 13: Black-White wage gap by education, controlling for occupation

	1968-1979	1980-1989	1990-1999	2000-2009	2010-2019
Below college					
white	1.000	1.000	1.000	1.000	1.000
black	0.885	0.900	0.908	0.935	0.941
College or above					
white	1.239	1.274	1.320	1.377	1.395
black	1.201	1.200	1.216	1.267	1.249

Notes: .

A.3 College and individual and neighborhood characteristics

Table 14 shows the regression of having a bachelor's degree or above on ability, parental transfers, share of college graduates in the neighborhood, and expenditures per student. These four variables jointly explain 21% of the variance of college graduation.

Table 14: Education Choice

	Bachelor's Degree or More
Log(ability)	0.1890***
	(0.008)
Log(High Skill Share)	0.0709**
	(0.031)
Log(Parental Transfers)	0.1103 ***
	(0.010)
Log(Expenditure per Student)	0.0311
	(0.036)
Constant	-0.5620 ***
	(0.131)
Observations	4,350
R^2	0.2081

Notes: Statistics in parentheses.

A.4 Neighborhood flows

First, we cluster census tracts at the national level. We use a K–means clustering algorithm on race share, income and housing prices. Second, we impute the cluster where people live in the NLSY79 using the probability that each a person in a county lives a certain cluster by race, household income of parents and compute the probability of moving between clusters between age 17 and 30. We find that 47% move across clusters. This estimate is robust to constricting the sample to counties which map to certain cluster with probability over 50%.

A.5 NLSY: Ability

This appendix describes the estimation of the AR(1) process for skills. We use data from the Children of the NLSY79' Survey via Abbott et al. (2019) which contains information on the ability of the child matched to the ability of the mother. For the mother, the data includes the AFQT score which is the typical measure of ability used in this literature. For the child, we do not have AFQT score so, following Abbott et al. (2019), we use the first principal component of their PIAT math, reading recognition and reading comprehension scores. For comparability, we transform both mother and child's scores into Z-scores.

We assume the inter-generational transmission of skills follows an AR(1) process and estimate its parameters with a regression of child's ability on mother's ability. The persistence parameter of the AR(1) is equal to the coefficient (0.51) and the R2 is equal to 0.26. We use these numbers as targets in the internal calibration of the AR(1) process for innate ability. We note that the mean of the ability is a normalization and will not affect the results (we set it to 2). Finally, we discretize the AR(1) process using Tauchen's method to estimate a 3-point grid and transition matrix.