Neighborhood Dynamics and the Distribution of Opportunity

Dionissi Aliprantis Daniel Carroll*

Federal Reserve Bank of Cleveland

April 5, 2017

Abstract: This paper studies neighborhood effects using a dynamic general equilibrium model. Households choose where to live and how much to invest in their child's human capital. The return on parents' investment is determined in part by their child's ability and in part by a neighborhood externality. We calibrate the model using data from Chicago in 1960 assuming that in previous decades households were randomly allocated to, and then could not move from, neighborhoods with different TFP. This restriction on neighborhood choice allows us to overcome the fundamental problem of endogenous neighborhood selection. We use the calibrated model to study Wilson (1987)'s hypothesis that racial equality under the law need not ensure equality of opportunity due to neighborhood dynamics. We examine the consequences of allowing for mobility, equalizing TFP, or both. In line with Wilson (1987), sorting can lead to persistent inequality of opportunity across locations if initial conditions are unequal. Our results highlight the importance of forward-looking agents.

*Address: Federal Reserve Bank of Cleveland, Research Department, PO Box 6387, Cleveland, OH 44101-1387, USA, e-mails: *dionissi.aliprantis@clev.frb.org*, *daniel.carroll@clev.frb.org*.

We are grateful to Alejandro Badel, Gerhard Glomm, Jim MacGee, Julia Thomas, Aubhik Khan, Daniel Hartley, Kyle Fee, Pedro Silos, Eric Young, Bob Becker, Roberto Pinheiro, Mahmoud Elamin, Stephen Ross, Víctor Ríos-Rull, Randy Walsh, James Heckman, anonymous referees, and seminar participants at the Cleveland Fed, Atlanta Fed, Kent State, Northeast Ohio Economics Workshop, 2014 Urban Economics Association Conference, and 2013 Latin American Meeting of the Econometric Society for helpful comments. We also thank Nelson Oliver and Mary Zenker for their research assistance. The views stated herein are those of the authors and are not necessarily those of the Federal Reserve Bank of Cleveland or the Board of Governors of the Federal Reserve System.

1 Introduction

This paper studies how the rise and fall of neighborhoods contributes to the rise and fall of families. We use a heterogeneous agents dynamic stochastic general equilibrium model in the spirit of Bewley (1986), Aiyagari (1994), Huggett (1996), and Krusell and Smith (1998) with three key features: residential sorting, neighborhood externalities, and intergenerational human capital accumulation. In the model, households choose where to live and how much to invest toward the production of their child's human capital. The return on parents' investment is determined in part by the child's ability, and in part by an externality from the average human capital in their neighborhood.[1] Opportunity is defined in our model as the productivity of parents' investments conditional on their child's ability.

In order to take a model with residential sorting, neighborhood externalities, and dynamics to the data, researchers have typically had to abstract from at least one of these mechanisms. The literature on the Moving to Opportunity (MTO) housing mobility experiment, for example, is either focused entirely on sorting (Galiani et al. (2015)), or else adopts stylized, static models of sorting in order to identify neighborhood externalities (Kling et al. (2007), Aliprantis and Richter (2016), Aliprantis (2017), Pinto (2014)).[2] Quantitative analyses of more theoretical models are typically constrained by computational limitations to adopt dynamics in the form of repeated static decisions[3].

We calibrate our model to a time and place without the endogenous neighborhood selection that plagues the neighborhood effects literature. This avoids the identification problem in which multiple parameterizations generate the same moments. Using data from Chicago in 1960, we first divide the city into a "black" and "white" neighborhood (N1 and N2, respectively). While allowing for neighborhood-specific technologies for the accumulation of human capital, we then calibrate the model without mobility to match the 1960 income distributions in each neighborhood.[4]

Our calibration uses race to initially constrain households' mobility and technology for accumulating human capital.[5] Beyond determining initial conditions, however, race is not

[1]Thus the technology for the intergenerational accumulation of human capital lies somewhere between Becker and Tomes (1979) and Calvó-Armengol and Jackson (2009).

[2]Rich microeconometric models of residential sorting are rarely specified and estimated jointly with outcomes (Ioannides (2010), Bayer et al. (2007)), even in the rare case that they do include both sorting and dynamics (Bayer et al. (2016)).

[3]Some related papers include Lundberg and Startz (1998), Bénabou (1996), Bénabou (1993), Glomm and Ravikumar (1992), Durlauf (1996), Epple et al. (2012), and Fernandez and Rogerson (1998). Badel (2010) presents a quantitative analysis with forward-looking agents.

[4]Neighborhood-specific technologies are driven by formal institutions like schools and public safety, as well as the informal arrangements determining social control/collective efficacy (Joseph et al. (2007), Sampson (2012)).

[5]See Johnson (2014) and Fuchs-Schuendeln and Hassan (2015) for related identification strategies.

the source of any other heterogeneity in our model, such as ability or preferences. Thus, households could have been allocated to different technologies based on any arbitrary rule, like eye color or passports, as in the cases of North versus South Korea or East versus West Germany. In this sense, our analysis is less about racial discrimination, and more fundamentally about how residential sorting and neighborhood effects drive the distribution of human capital over time.

This is precisely the subject of Wilson (1987)'s seminal work that launched a literature on neighborhood effects. After studying the concentration of poverty in Chicago between 1970 and 1980, Wilson concluded that racial equality under the law would not by itself ensure equality of opportunity since residential sorting could change neighborhood effects over time.

We use our calibrated model to study Wilson's hypothesis, and to evaluate how two counterfactual policies would have changed opportunity and welfare across neighborhoods. For each policy change we find a transition path from the original steady state to the new steady state. We document changes in opportunity and welfare for the transition.

Our first numerical experiment allows for residential choice, which we interpret as the counterfactual resulting from eliminating legal racial discrimination. This is one of the central thought experiments suggested by Wilson (1987)'s ex-post analysis of Chicago, interpreting legislative victories of the Civil Rights Movement like the Fair Housing Act of 1968 as a discrete change to residential sorting rules. In line with Wilson (1987), we find that the calibrated model predicts a rapid and complete depopulation of N1. Only very poor households would choose to live in N1 in this world; so poor, in fact, that such levels are never visited by agents in the model.

A key feature of our model is that parents internalize the utility of their descendants, and therefore they take into account the entire future path of prices. Due to the difficulty of incorporating such dynamics, in most related quantitative exercises parents get warm glow utility from endowing their child with income. Since this makes information about the future irrelevant for parents' decisions, these models can be characterized as a series of repeated static problems. To demonstrate the importance of forward-looking behavior, we also study the dynamics of our model when agents are myopic. In the myopic case, parents believe that future human capital externalities and house prices will remain at their current levels forever.

We find vastly different dynamics under myopia than we do under rational expectations. With forward-looking agents, households anticipate the decline of the poorer neighborhood after allowing for mobility. As a result, N1 depopulates and only very poor households would choose to live there. In contrast, myopic agents do not anticipate a decline of the poorer neighborhood in response to allowing for mobility. The result is that N1 remains stable and

actually has slightly higher income after allowing for mobility.

In a second numerical experiment, we examine what would happen if the restriction on residence were maintained while equalizing technologies across neighborhoods. We interpret this counterfactual as Malcolm X's ex-ante vision of separation.[6] In this counterfactual, N1 makes a smooth transition to a human capital distribution like N2's. Since high income households stay in N1, all residents benefit from the resulting buildup in the neighborhood's externality. Nevertheless, the X policy is not the preferred policy for residents in N1, since a high income location is immediately accessible when sorting is allowed, but takes time to emerge in the X transition.

Finally, we allow for residential choice while also equalizing technologies across neighborhoods. We interpret this counterfactual as Martin Luther King, Jr.'s ex-ante vision for the integration of Chicago.[7] Although there are large welfare gains to households initially in N1 under this policy, the neighborhoods still have permanently unequal incomes in this experiment. These long run differences persist even without moving frictions or racial preferences.

We interpret our numerical experiments as support for Wilson's hypothesis. While stylized theoretical models have shown the possibility for residential sorting and neighborhood effects to generate permanent income inequality across neighborhoods, our analysis has shown that this scenario is empirically relevant for the neighborhood dynamics of late 20th-century Chicago.

The remainder of the paper is structured as follows: Section 2 presents four stylized facts that are used to motivate the model. Section 3 presents a dynamic general equilibrium model of neighborhood dynamics and human capital accumulation. Section 4 presents the calibration of the model to data from Chicago in 1960, and Section 5 presents the results of the numerical experiments we implement with this model. Section 6 concludes.

2 Stylized Facts

The central issue in the neighborhood effects literature is understanding what generates spatial correlations in outcomes. Is the local environment a primary cause of individuals' outcomes, or do people with similar outcomes simply choose to live near each other? In nearly all contexts, the endogeneity of neighborhood selection has represented a fundamental obstacle to identifying neighborhood effects and distinguishing between these explanations.

[6] See X's definition of black nationalism in X (1990) or the related discussion in O'Flaherty (2015).

[7] While King is often remembered in terms of his work for open housing, integrating schools was also a primary focus of his work in Chicago, and improving the general conditions in N1 was another major goal. See Chapter 28 of King (1998) for a description of King's work in and vision for Chicago.

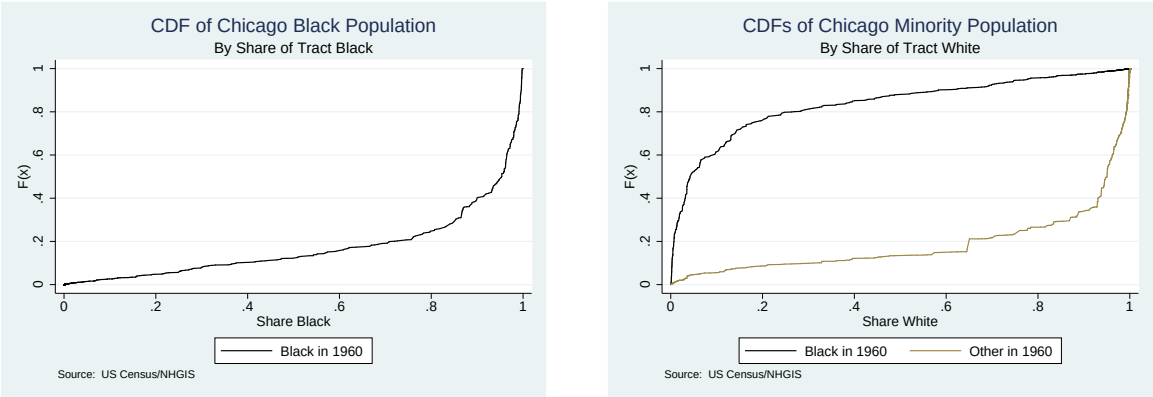
This analysis offers insight into the broad question of how neighborhood externalities impact income inequality, using Chicago in the 20th century as a circumstance restricting the endogeneity of neighborhood selection. Here we establish four stylized facts about the decades before 1960 to justify the key features of our model as we apply it to Chicago: There were two neighborhoods in the city; they were defined by race; they were unequal; and there was no mobility between the two neighborhoods for decades.

2.1 Stylized Fact 1: Black Residents of Chicago Lived in Black Areas

The black ghetto in the US was born between 1890 and 1940 and grew between 1940 and 1970 (Cutler et al. (1999)). Massey and Denton (1993) note that blacks and whites were not particularly segregated before 1900. This changed in the first decades of the 20th century in response to the Great Migration, in which large numbers of African Americans moved to Northern cities from the South. By 1930, in most urban areas in the US the boundaries within which blacks were allowed to live had been established through violence, collective anti-black action, racially restrictive covenants, and discriminatory real estate practices (Massey and Denton (1993)).

Focusing on Chicago, in 1930 two-thirds of all black residents lived in census tracts that were 90 percent black or more, and by 1940 this had grown to over three-quarters (Hirsch (1998), p 4). In 1960 the median black person in Chicago lived in a neighborhood that was 95 percent black (Figure 1a), which actually *increased* to 98 percent by 1990. In our empirical analysis we define neighborhood N1 as all census tracts in which 80 percent or more of the residents were black in 1960, and under this definition a full 75 percent of African Americans in Chicago lived in N1 in 1960.[8] Figure 1b shows that the level of segregation experienced in African American neighborhoods was unlike that of the immigrant enclaves experienced by other minority groups (See also Massey and Denton (1993) on this point, especially Chapter 2.).

[8]These data are all consistent with the national data presented in Cutler et al. (1999); see especially Table 4.



(a) Blacks in Chicago

(b) Minorities in Chicago

Figure 1: Segregation of Minorities in Chicago in 1960 and 1990

2.2 Stylized Fact 2: Limited Black Mobility

Violence was a key factor restricting mobility from black neighborhoods to the rest of Chicago. Between 1945 and 1950 alone Chicago experienced 357 "incidents" related to housing (Hirsch (1998)). Meyer (2000) discusses several of these incidents, such as the complete razing of a house purchased by an African American woman located just two blocks outside of the ghetto, or the firebombing of a house that killed two children (p 89). Rubinowitz and Perry (2002) conclude that racial crimes "around housing conflicts... became the norm in Chicago the way other forms of racial violence, such as lynchings and church bombings, became commonplace in the South" (p 347). This environment had not changed much by the time Martin Luther King, Jr. led a march in Chicago for open housing in 1966: His group was met by such violent resistance that he was led to conclude "The people of Mississippi ought to come to Chicago to learn how to hate" (Polikoff (2006), p 41).

Legal roadblocks also restricted blacks from moving into white neighborhoods. For example, in 1924 the National Association of Real Estate Brokers' code of ethics adopted the statement that "a Realtor should never be instrumental in introducing into a neighborhood... members of any race or nationality... whose presence will clearly be detrimental to property values in that neighborhood" (Massey and Denton (1993), p 37). This provision remained in effect until 1950.

Recognized as a spokesperson for the African American experience of the mid-twentieth century (Polsgrove (2001)), the writer James Baldwin was challenged in a debate over his use of the word ghetto to describe black neighborhoods: "There is no law in America or indeed no practice in America that makes rich Negroes live in the ... as-you-call-it 'ghetto.' " Famously careful with his words, Baldwin reacted strongly that, "I stick to the word

ghetto... [because] There is no way for any black man to move out of it... I say ghetto, and I say ghetto because you can't move out..." (Baldwin (1989), pp 115-116).

2.3 Stylized Fact 3: Separate and Unequal Neighborhood Externalities

Separation would not necessarily be a problem for economic outcomes if blacks and whites lived in separate but equal neighborhoods (Cutler and Glaeser (1997), Borjas (1995), X (1963)). But racial discrimination precluded this possibility in the decades before the Civil Rights Movement: N1 and N2 were not equal in important ways related to the intergenerational transmission of human capital.

Schools in Chicago were segregated in the decades prior to 1960. In 1945 the president of the NAACP branch serving Chicago stated: "We have segregated schools outright... They are as much segregated as the schools in Savannah, Georgia, or Vicksburg, Mississippi" (Homel (1984), p 27). Chicago's school boards adjusted attendance-area boundaries to segregate students in schools along the same lines they were segregated geographically (Homel (1984)). In 1964, the first time the Chicago Board of Education published racial statistics, 67 percent of black students attended high schools that were (more than 90 percent) black, and 89 percent of black elementary school students attended black schools (Neckerman (2007), p 95).

School segregation impacted individual-level experiences because black schools did not have the same resources as white schools. Black schools faced overcrowding, resulting in limited instruction time with odd schedules, difficulty staffing teachers, and fewer resources for things like facilities relative to white schools (See Chapter 4 of Neckerman (2007), especially pages 88-97.). While it is hard to find historical data on measures of school quality by race for Northern schools since they were not explicitly segregated (Collins and Margo (2006)), these data from Chicago are consistent with evidence from the South that teachers' pay was lower in black schools relative to white schools (Collins and Margo (2006)), class sizes were generally larger and the length of the school year was shorter (Collins and Margo (2006), Orazem (1987)), and other inputs were lower (Margo (1986)).

African Americans residing in N1 faced discrimination in other important processes such as redlining practices that decreased the family and community resources that could have been devoted towards the transmission of human capital to children (Squires (1997), President's National Advisory Panel on Insurance in Riot-Affected Areas (1968)), as well as discrimination in the justice (Blackmon (2008)) and health care systems (Washington (2006))

2.4 Stylized Fact 4: There Were *Two* Neighborhoods

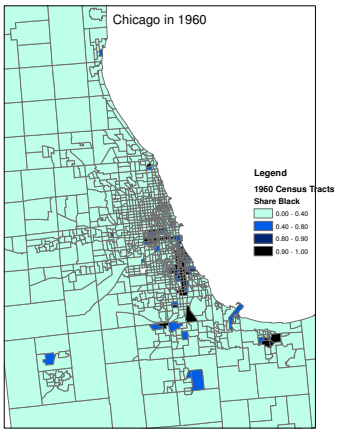
Defining the word "neighborhood" is crucial to determining whether a two-neighborhood model is a useful lens through which to look at Chicago in 1960. The literature does not give much guidance on this topic: Durlauf (2004) notes that nearly all empirical studies in the neighborhood effects literature take a particular neighborhood structure as known *ex ante*, despite the centrality of this definition. The appropriate definition of neighborhood is likely to depend on the specific analysis under consideration (See Sampson (2012), especially Chapter 3, or page 37 of Lucas (1988).).

The salience of race in Chicago justifies defining neighborhoods in terms of racial composition. Social interactions and resources were distributed in Chicago according to geographically determined racial lines, whether they pertained to schooling (Neckerman (2007)), housing (Polikoff (2006)), or broader political processes (Sampson (2012), pages 40-42). Even in 1980 and 1990, Conley and Topa (2002) find that racial/ethnic composition was by far the most important predictor of spatial correlation in unemployment across census tracts in Chicago.

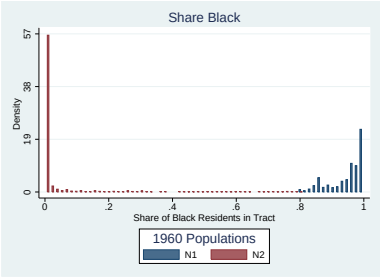
Furthermore, the definition of race in the US justifies a model with precisely *two* neighborhoods. The one-drop rule categorizing individuals with *any* African heritage as being African American has generated a binary definition of race that is quite different from the broader spectrum experienced in other locations (Hickman (1997), Arthur (1999)). For a striking example, consider that President Barack Obama classified himself as black, and black alone, on the 2010 US Census (Roberts and Baker (2010)).

The data also provide justification for viewing Chicago in 1960 as two neighborhoods defined in terms of racial composition. Almost all of N1 is spatially connected (Figure 2), and spatial proximity is considered to be a key determinant of neighborhood externalities (Sampson (2012), Bayer et al. (2008)).

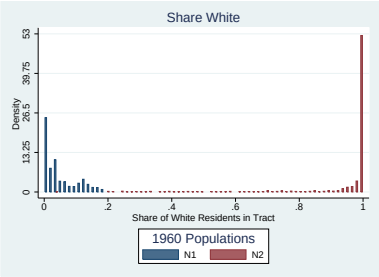
Figure 2: Racial Composition of Neighborhoods in Chicago, 1960



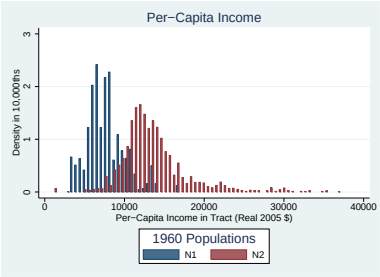
Additionally, Figure 3 shows that N1 and N2 were fundamentally different in 1960 according to several measures of human capital. Thus while one could imagine there being important variation in the externalities experienced by residents within each neighborhood (Pattillo (2003)), our two-neighborhood division is a useful abstraction. In terms of racial composition, Figures 3a and 3b illustrate that N1 and N2 were racially homogenous in a way suggesting two distinct externalities. The neighborhood externality in our model operates through income, which could involve mechanisms operating through other outcomes like employment or educational attainment. One could easily interpret the census tracts in N1 and N2 as coming from two distributions for these mediators.



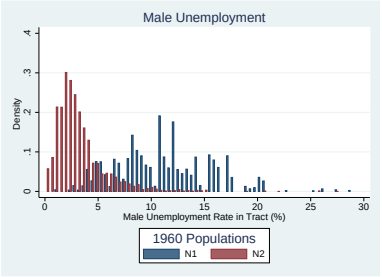
(a) Share Black



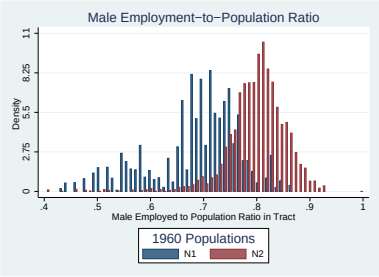
(b) Share White



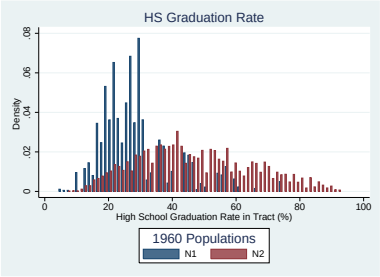
(c) Income



(d) Male UR



(e) Male EPR



(f) High School Graduation Rate

Figure 3: Distributions of Census Tract Characteristics by Neighborhood in 1960

We emphasize that because N1 is the focus of our analysis, we are most concerned that its residents experienced a "uniform" neighborhood externality. Future research can relax our abstraction from the variation in the externality experienced by residents in the much larger area of N2.

3 A Model of Neighborhood Dynamics and Human Capital Accumulation

We now present a dynamic general equilibrium model of the intergenerational accumulation of human capital. The model has three key features: Residential sorting, location-specific inputs to production (ie, a neighborhood externality), and forward-looking agents. We expand on the roles of these mechanisms where they appear in the model description below.

3.1 Households

A unit continuum of overlapping generations households lives in a city that is divided into two neighborhoods. Each household consists of two individuals, a parent and a child, and all individuals live for two periods: At the end of each period adults die, children become adults, and new children are born. Adults receive utility from consuming housing services whose units are ordered according to a single housing quality index ($s \in \mathbb{R}_+$), a non-housing good ($c \in \mathbb{R}_+$), and the discounted expected utility of their offspring. Children receive no utility until they become adults, however parents are altruistic; therefore, a household is functionally identical to an infinitely-lived dynasty.[9] Preferences for a dynasty take the form

$$U(c, s) = E_0 \sum_{t=0}^{\infty} \beta^t u(c_t, s_t)$$

where $\beta \in (0, 1)$ is the discount factor between a parent and its offspring.

Each household is characterized by its state vector (h_t, a_t, n_t), where $h_t \in \mathbb{H} = [\underline{h}, \overline{h}] \subset \mathbb{R}_+$ is the human capital level of its adult, $a_t \in \mathbb{A} \equiv \{\underline{a} = a_1, a_2, \ldots, a_n = \overline{a}\} \subset \mathbb{R}_+$ is the ability of its child to produce human capital, and $n_t \in \{N1, N2\}$ is the neighborhood in which the household ends the period. We assume that (h_t, a_t, n_t) is a random vector whose joint distribution μ_t has density function $f_t : \mathbb{H} \times \mathbb{A} \times \{N1, N2\} \to \mathbb{R}$ defined by $f_t(h, a, n)$, and we sometimes refer to the conditional density $f_{N1,t} = \frac{f_t(h,a,N1)}{\sum_{a \in \mathbb{A}} \int_{\mathbb{H}} f_t(h,a,N1)dh}$, with $f_{N2,t}$ defined

[9] Because this paper focuses on the effects of forces external to the household (i.e., the neighborhood), we abstract away from the distributions of consumption and housing services across household members.

analogously. A neighborhood is a joint density of human capital and ability $f_{n,t}$, a per-unit price of housing services $p_{n,t}$, an externality level $\chi_{n,t}$, and a share of the citywide population.

We define human capital as the skills and knowledge that generate labor income. We think of human capital not only in terms of the skills acquired through formal education, like those measured by the AFQT, but also in terms of any of the other factors that help to determine labor income, like personality traits and social skills (Borghans et al. (2008)). The child's human capital is determined by a function $G : \mathbb{R} \times \mathbb{A} \times \mathbb{H} \times \mathbb{H} \to \mathbb{H}$ defined by

$$h_{t+1} = G\left(Z_n, a_t, i_t, \chi_{n,t}\right). \tag{1}$$

Under this specification, human capital is produced by the combination of four sources: three factors of production (innate ability a_t, private investment i_t, and a public good $\chi_{n,t}$), and a neighborhood-specific technology for combining these inputs summarized by a Total Factor Productivity (TFP) parameter (Z_n). We think of ability as immutable characteristics, including cognitive and non-cognitive abilities. Private investment is consumption foregone for the sake of endowing one's child with human capital. This might be time spent with the child (ie, on homework after school, providing healthy meals, safe transportation to and from school), or money spent on the child (ie, tutors, extracurricular activities, and summer camps).[10]

The public good is meant to capture a wide range of spatially-determined mechanisms, like schools and safety. We think of the externality level χ_n in terms of resources devoted to things like teachers and police, and the TFP parameter Z_n as capturing institutional differences across neighborhoods in the productivity of these resources. This production allows for identical levels of ability and private and public investment to produce different levels of human capital. As an example, similar tax revenues devoted to hiring more teachers or police (captured in χ_n) could be inputs to institutions with different levels of productivity (Z_n). One can additionally interpret the externality level χ_n as summarizing the social interactions one typically has, as determined by peers and role-models in the neighborhoods, under the assumption that peer quality is positively correlated with parents' human capital. G is assumed to be strictly concave in each argument.

The timing of decisions and updating is shown below in Figure 4. There are two sub-periods. In the first, neighborhood distributions change because households sort across neighborhoods. In the second, neighborhood distributions change because of the evolution of human capital across generations.

[10]Note that we have restricted the support of h_t to be strictly positive because our utility function is not bounded below for some parameterizations. Since ability is strictly positive, this restriction is equivalent to there being a positive minimum amount of investment required on the part of parents.

More specifically, at the beginning of period t, an adult with human capital h_t resides in one of the two neighborhoods, the location n_{t-1} chosen by their parent. In the first subperiod, the adult has a child, observes its ability a_t, and chooses whether or not to move. The initial distribution of households $\widehat{\mu}_t$ is updated according to the law of motion $\mu_t = \widetilde{\Psi}(\widehat{\mu}_t)$, and the first subperiod ends. Taking as given the price and externality in the chosen neighborhood, in the second subperiod the household chooses consumption c_t, housing services s_t, and investment in its child's human capital i_t. Each household's human capital is updated according to $h_{t+1} = G(Z_n, a_t, i_t, \chi_t)$, adults die, children become adults, a new child is born with ability a_{t+1}, and the second subperiod ends.

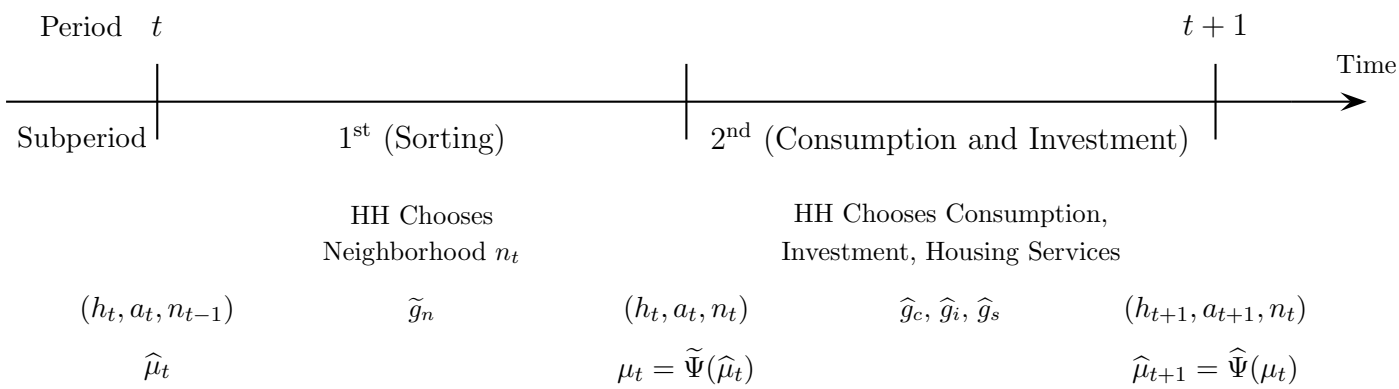


Figure 4: Updating of Households and Neighborhoods over Time

We assume rational expectations, meaning that households know the sequence of neighborhood externalities and housing prices, so that households are able to solve a well-posed problem. This requires that households' expectations about the neighborhoods are consistent with the neighborhoods realized by the moving and investment decisions of all households. This implies that households know both the law of motion determined by moving decisions $\left(\mu_t = \widetilde{\Psi}(\widehat{\mu}_t)\right)$ and the law of motion determined by investment decisions/the production technology and the ability process $\left(\widehat{\mu}_{t+1} = \widehat{\Psi}(\mu_t)\right)$ so that households have perfect foresight over the full sequence of relevant human capital externalities and housing prices. Under segregation, this is simply the sequence in the neighborhood in which the household resides, either $\{\chi_{N1,t}, p_{N1,t}\}_{t=0}^{\infty}$ or $\{\chi_{N2,t}, p_{N2,t}\}_{t=0}^{\infty}$. When moving is allowed, perfect foresight means knowledge of the full sequence of intra-temporal human capital externalities and housing prices for both neighborhoods, $\{\chi_{N1,t}, \chi_{N2,t}, p_{N1,t}, p_{N2,t}\}_{t=0}^{\infty}$.

Under rational expectations, conditional on choosing a location, each household has a well-defined budget constraint

$$c_t + i_t + p_{n,t}\, s_t \leq \omega_t h_t.$$

A key feature of our model is the distinction between the intra-temporal updating rule $\widetilde{\Psi}$ and the inter-temporal updating rule $\widehat{\Psi}$, since the wealth distribution typically only changes *intertemporally*. In similar incomplete markets models of physical capital accumulation with transitional dynamics (e.g., Ríos-Rull (1999)), the rule Ψ would typically only capture end of period changes to the household state vector from the idiosyncratic shock process and optimal investment decisions. Here, though, the wealth distribution in neighborhood n can change both intertemporally, as human capital changes across generations due to investment decisions, and *intratemporally* due to migration decisions.

This distinction matters because the composite rule for updating distributions between time periods $\Psi = \widehat{\Psi}\left(\widetilde{\Psi}(\cdot)\right)$ changes depending upon the sorting rules $\widetilde{\Psi}$ permitted. If no sorting is permitted, so that $\widetilde{\Psi}\left(\widehat{\mu}_t\right) = \left(\widehat{\mu}_t\right)$, then the central assumption that production is a function of neighborhood-specific human capital implies that any differences in neighborhood steady states can only exist if neighborhoods differ in either household preferences, the ability process, or the human capital production function.[11] Our model assumes the final explanation. These differences could arise from many sources like racial discrimination, political economy for public services like schooling (Ichino et al. (2010), Glomm et al. (2011)), crime and personal security (Anderson (1999), Aliprantis (2016)), or social efficacy (Sampson et al. (1999)).

3.2 The Firm

We assume that non-housing goods are produced in a national market. For simplicity, we assume that this market wage rate is 1, and we assume that labor is perfectly mobile so that the city-wide wage ω is also equal to 1.

Housing services Q are produced by a price-taking firm using labor according to a constant returns to scale function of effective labor, H, and land, L:

$$Q = H^{\alpha} L^{1-\alpha}, \qquad 0 < \alpha < 1.$$

Taking the wage rate as given, the firm supplies units at the neighborhood-specific price p_n. Solving the firm's maximization problem and imposing that supply equals demand in each neighborhood, i.e. $Q_n = S_n = \sum_{a \in \mathbb{A}} \int_{\mathbb{H}} g_s(h, a, n) f(h, a, n) dh$ for housing services in both locations returns the pricing equations.

[11] See Kremer (1997) for a related model in which sorting has negligible implications for steady state inequality when it is assumed there is a constant technology across neighborhoods.

In equilibrium, there are rents to land equal to

$$\frac{1-\alpha}{\alpha}\left[\left(\frac{pop_1}{L_1}\right)^{\frac{1}{\alpha}}\bar{S}_1^{\frac{1}{\alpha}}L_1+\left(\frac{pop_2}{L_2}\right)^{\frac{1}{\alpha}}\bar{S}_2^{\frac{1}{\alpha}}L_2\right]. \tag{2}$$

We assume that these rents go to the absentee landlord.

$$\begin{aligned} p_n &= \frac{1}{\alpha}\left(\frac{S_n}{L_n}\right)^{\frac{1-\alpha}{\alpha}}, \qquad n=1,2 \\ &= \frac{1}{\alpha}\left(\frac{pop_n}{L_n}\bar{S}_n\right)^{\frac{1-\alpha}{\alpha}}. \end{aligned}$$

The last expression decomposes the market clearing price into the product of average housing demanded in the neighborhood $\bar{S}_n$ and the ratio of population residing in n to the land available, which we refer to as the congestion ratio. $\frac{\alpha}{1-\alpha}$ is the price elasticity of supply. As α approaches unity, the price becomes extremely sensitive to changes in either input. In our calibration we set L_n equal to the initial population share in each neighborhood. Note that because of our assumption of a national labor market, in our model Chicago is a small open economy for labor. This implies that labor will be paid the same to produce both the non-housing good and housing services.[12]

Notice that the price of housing acts like a congestion cost. The more that households sort into the same neighborhood, the higher the cost is to everyone. At the same time the price in the other neighborhood decreases, encouraging migration back. Without congestion costs, corner solutions where one neighborhood is empty are more likely to arise.

3.3 Sorting Rules

The recursive formulation and equilibrium concepts of the model depend crucially on the types of neighborhood mobility permitted. To show these distinctions explicitly, especially as they pertain to our empirical analysis, we will state the recursive problems solved by households and define a competitive equilibrium under each sorting policy separately. The broad point helpful for interpreting the remainder of the model description is that because we are studying Chicago in the mid- to late-twentieth century, we allow different sorting rules depending on the time period under consideration. We assume that up to 1960 households

[12]Because we do not model race, we are unable to account for racial discrimination in the labor market. The focus of this analysis is to quantify the impact of neighborhood externalites and sorting on outcomes with a general model that abstracts from legal racial discrimination.

are prohibited from moving across neighborhoods (i.e., $n_{t+1} = n_t$). In this case, the model is of two economies that do not interact with each other.

We then interpret the legal victories of the Civil Rights Movement as a change to a new model in which households are allowed to move across neighborhoods. We assume that the prohibition on sorting is removed between 1960 and 1970, and that the new two-neighborhood model allowing for sorting characterizes Chicago in subsequent decades. Neighborhoods N1 and N2 become interconnected in this second model, as intra-period migration flows change the price of housing and the return to investment in each neighborhood.

3.4 Recursive Formulation and Equilibrium

A household's problem can be described recursively by a nested value function:

$$V\left(h, a, n_{-}\right) = \max_{n} \left\{ \max_{c,i,s} \; u\left(c, s\right) + \beta EV\left(h', a', n\right) \right\} \tag{3}$$

subject to:

$$c + i + p_n \, s \leq h, \tag{4}$$

$$h' = G(Z_n, a, i, \chi_n). \tag{5}$$

Solving 3 subject to 4 and 5 returns the value function V and decision rules $\widetilde{g}_n$, $\widehat{g}_c$, $\widehat{g}_i$, and $\widehat{g}_s$ for location, consumption, investment, and housing services, respectively.

One of the distinguishing features of our model is the forward-looking behavior of households. The continuation value $\beta EV\left(h', a', n\right)$ makes the parent's utility a function of the entire sequence of their descendants' utilities. Related models in the literature on intergenerational mobility typically assume that the parent's utility is a function only of current period variables. This might include the size of bequests to their children (Glomm and Ravikumar (1992)) or the education/income level of their children (Fernandez and Rogerson (1998)).

Forward-looking behavior leads to very different choices for households in our model, as we show in 5.3. Because parents care about their children's *utility* in our model, their decisions will take into account the future trajectories of neighborhoods. For instance, if a transition between steady-states implies that a neighborhood will decline over time, forward-looking households will move sooner than households that only care about current-period neighborhood characteristics. This has major implications for the rise and fall of neighborhoods, and by implication, intergenerational mobility (Becker and Tomes (1986)).

We now state our equilibrium concept:

Definition 1. *A* ***steady-state recursive competitive equilibrium with moving (MRCE)*** *is a set of neighborhoods, a value function* $V(h, a, n_-)$*, policy functions* $\widetilde{g}_n(h, a, n_-)$*,* $\widehat{g}_c(h, a, \widetilde{g}_n(h, a, n_-))$*,* $\widehat{g}_i(h, a, \widetilde{g}_n(h, a, n_-))$*,* $\widehat{g}_s(h, a, \widetilde{g}_n(h, a, n_-))$*, and laws of motion* $\widehat{\Psi}$ *and* $\widetilde{\Psi}$ *such that*

1. Given prices and the laws of motion, V, $\widetilde{g}_n$, $\widehat{g}_c$, $\widehat{g}_i$, and $\widehat{g}_s$ solve the household problem.

2. The housing market clears in each neighborhood:

$$S_n = \sum_{a \in \mathbb{A}} \int_{\mathbb{H}} g_s(h, a, n) f(h, a, n) dh \quad \text{for} \quad n = N1, N2$$

3. Neighborhood externality depends on its residents, $\chi_n = X(\mu)$.

4. $\widetilde{\Psi}$ is consistent with the moving decisions $\widetilde{g}_n$ of households in neighborhoods N1 and N2.

5. The law of motion $\widehat{\Psi}$ is consistent with human capital decisions $g_{h'}(h, a, n_-) = G(Z_n, a, g_i(h, a, n_-), \chi_n)$ and the ability process.

6. The joint distribution of human capital and ability is stationary $\widehat{\mu}' = \widehat{\Psi}\left(\widetilde{\Psi}\left(\widehat{\mu}\right)\right) = \widehat{\Psi}\left(\mu\right) = \widehat{\mu}$.

3.4.1 Equilibrium in the Model with Segregation (SRCE)

Definition 2. *A* ***steady-state recursive competitive equilibrium under segregation (SRCE)*** *is an MRCE under the following restriction:*

SRCE-a $n = \widetilde{g}_n(h, a, n_-) = n_-$

Note that an implication of the SRCE-a restriction is that the law of motion for sorting is trivial: $\widetilde{\Psi}\left(\widehat{\mu}\right) = \widehat{\mu}$. In other words, since there is no location decision in the model under segregation; nothing happens in the 1st sub-period (See Figure 4.).

3.4.2 Existence and Characterization of Equilibria in These Models

We show in Appendix A that the household problem under segregation can be expressed recursively, and furthermore prove the existence of an SRCE. Appendix B discusses how one might prove the existence of an MRCE, as well as the condition in such a proof that is difficult to show analytically, and some intuition of how this condition could be met.

4 Model Specification and Parameterization

4.1 Sorting Equilibria and Production

It is worth considering the types of sorting patterns that can give rise to stable asymmetric equilibria in this model, since they influence the specification of several functional forms and the calibration of some of the important parameters. In the absence of binding moving constraints (e.g., fixed cost of moving, moving opportunity shock), a little reflection makes clear that an MRCE must be one of two types. Either the neighborhoods are identical (same prices, same externality levels, and same wealth distribution) or they are asymmetric where one neighborhood has a higher externality value and higher housing price than the other. It would be inconsistent with optimizing behavior for one neighborhood to have a low externality and high house prices since households would choose to move away from that location, which in turn would induce the firm to lower its housing price.

In order for an asymmetric equilibrium to be sustained, the moving decisions $g_n(h, a)$ must have a particular ordering over h. Without loss of generality, label the low externality/low price location $N1$ and the other $N2$. The required ordering in the moving decision rule is summarized in the following condition:

Sorting Condition (sorting by h) : *Given neighborhood prices $p_{N1} < p_{N2}$ and externalities $\chi_{N1} < \chi_{N2}$, for any $a_i \in \mathbb{A}$, if $g_n(h_1, a_i) = N2$, then $g_n(h_2, a_i) = N2$ for all $h_2 > h_1$.*

The Sorting Condition says that given ability, high human capital households are willing to pay more for a high externality than are low human capital households.[13] The intuition for this condition is illustrated by considering how sorting changes N2 in response to a price increase. The Sorting Condition ensures that all else constant, a price increase in N2 will induce an outflow of *below average* (in N2) human capital households, increasing the externality in N2. This rise in the externality compensates households who remain and pay the higher housing price, allowing for a higher-income, higher-price neighborhood to exist in a stable equilibrium under the Sorting Condition.

Suppose that in contradiction to the Sorting Condition, high human capital households were the first to move in response to a price increase. Then the implied sorting would reduce the externality in N2, penalizing the remaining households. This would push more above average human capital households to move, decreasing the externality in N2 still further, illustrating how a higher-income, higher-price neighborhood cannot exist in a stable equilibrium without sorting rules satisfying the Sorting Condition.

[13] Bénabou (1993) makes an analogous assumption in terms of the cost of skill acquisition (A2).

Theoretically, inputs of production must be complements in order to satisfy the Sorting Condition. If private investment and the externality are substitutes, then high-income households are capable of offsetting a low externality by spending more privately. Since in this case the neighborhood externality is not as important to these high-income households, they are attracted to the low price neighborhood where they can afford more housing. Thus if inputs are substitutes, high-income households will not sort into the high-price, high-externality neighborhood (N2). A similar equilibrium failure results from the "chasing problem" discussed in Durlauf (1996).

In contrast, if private investment and the neighborhood externality are complements, then high-income households cannot easily offset a low externality by spending more privately. Thus, high-income households could be willing to pay a higher price for housing in exchange for a higher externality. Since households have a desire to smooth consumption over time, a high human wealth household in particular has a strong motivation to endow its child with a high level of human capital. The increasing marginal cost of producing human capital gives these households an incentive to locate in a high externality location.

4.2 Production Function Specification

Recalling Equation 1, we specify that h' units of human capital are produced next period according to the Constant Elasticity of Substitution (CES) production function

$$\begin{aligned} h' &= G\left(Z_n, a, i, \chi_n\right) \\ &= Z_n\Big(\left(\frac{a}{3}\right)^{\frac{\gamma-1}{\gamma}} + \left(\frac{i}{3}\right)^{\frac{\gamma-1}{\gamma}} + \left(\frac{\chi_n}{3}\right)^{\frac{\gamma-1}{\gamma}}\Big)^{\frac{\gamma}{\gamma-1}}. \end{aligned} \tag{6}$$

where the externality level χ_n is determined as:

$$\chi_n = H_n = \sum_{a\in\mathbb{A}} \int_{\mathbb{H}} h f_n(h,a)dh,$$

or the intra-period average human capital in neighborhood n.

The CES functional form adopted in Equation (6) allows for flexibility in parameterizing the factors of production either as substitutes or complements (Uzawa (1962)). We restrict the technology so that inputs are complements in production (ie, $\gamma < 1$) for both empirical and theoretical reasons. Empirically, the best available evidence indicates that parental investments (i) and investments in public schools (part of χ_n in our model) are likely to be complements (Grawe (2010)). Beyond that insight, the empirical literature offers little guidance on specifying the production function, including even what are the factors of

production.[14]

Theoretically, complementarity between inputs is necessary for the existence of a general equilibrium. When moving is allowed, complementarity is needed for decision rules to satisfy the Sorting Condition just discussed. When moving is not allowed, complementarity of inputs makes the marginal cost to producing h' become infinite at some point. Since a has a finite upper bound, this ensures that h will be bounded above by a maximum sustainable human capital level (and hence we are assured that an equilibrium will exist - see the Appendix for further discussion). Complementarity guarantees that there exists some x_h such that for any χ_n in $\mathcal{H}$, all households with human capital above x_h will choose a lower level for their child. Without this restriction or a similar one, it would be possible for a sufficient mass of households to have human capital above some high level that would generate a large enough externality for h to grow for all households. In such a case the mass of households above x_h will be even larger the next period, and so too will be the externality, generating explosive dynamics.

We exploit the parsimonious specification in Equation (6) for identification. Under this specification, the Z_n and γ parameters each have clear and distinct implications for inequality across and within neighborhoods.

The level of income in a neighborhood is greatly dependent on Z_n. This is because the role of the neighborhood-specific TFP parameter, Z_n, is straight-forward: it scales production.

The degree of inequality within neighborhoods is greatly dependent upon γ. The parameter $\gamma \in [0, \infty)$ is the elasticity of substitution between inputs, and changes the concavity of output as a function of investment for a given level of ability and the externality.[15]

[14] In the related Education Production Function (EPF) literature, for example, "there is a remarkable lack of consensus over which inputs increase children's achievement and to what extent" (Todd and Wolpin (2007), p F4). In the neighborhood effects literature, "Perhaps disappointingly there remains substantial scope to conduct studies whose primary aim is simply to test for the presence, and measure the magnitude of, neighborhood effects. There is not yet even a loose disciplinary consensus on the rough importance of neighborhoods on life outcomes" (Graham (2016), p 53).

[15] Inputs are complements when $\gamma < 1$ and substitutes when $\gamma > 1$.

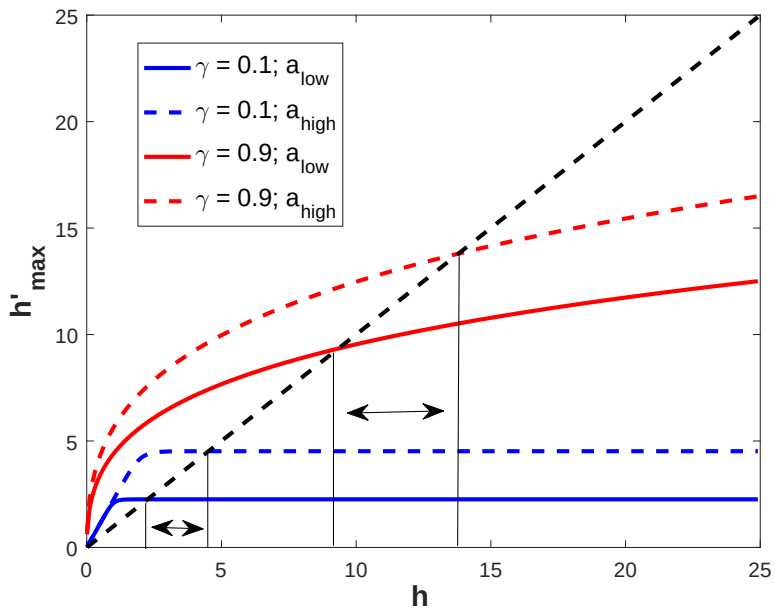


Figure 5: Maximum attainable human capital in steady state

Figure 5 shows how the elasticity of substitution changes the steady state distribution of wealth. The figure plots next period's human capital, h', in a given neighborhood for a low and a high ability child given that the parent invests all its human capital (i.e., $i = h$). A higher elasticity creates a wider distance between the steady state wealth levels of the two households.

4.3 Utility and the Stochastic Process for Child's Ability

Period utility is assumed to be separable in housing services and non-housing goods as follows:

$$u\left(c_t, s_t\right) = \frac{c^{1-\nu_c}}{1-\nu_c} + \theta \frac{s^{1-\nu_s}}{1-\nu_s},$$

where $\frac{1}{\nu_c}$ is the intertemporal elasticity of substitution in non-housing consumption and ν_s is the curvature of utility with respect to housing. The ratio $\frac{\nu_c}{\nu_s}$ is the elasticity of substitution between housing and non-housing goods.

We restrict $\nu_s > \nu_c$, which implies that households' demand for housing services relative to non-housing services declines in income. This is consistent with data on housing expenditures from the 1973 and 1991 American Housing Surveys. Whether measured in terms of rent or housing costs, households in both surveys spent a smaller fraction of income on housing as income increased.[16] This restriction helps the model to satisfy the Sorting Condition, by

[16] See Figure 1 of HUD (1976), as well as Tables 3-20 and 4-20, respectively, of HUD (1993).

ensuring that richer households place relatively more weight on the neighborhood externality than house prices.[17]

We assume that the stochastic process for child's ability is a stationary Markov chain with transition probabilities denoted by $\pi(a_i|a_j)$.

4.4 Calibrating a SRCE to 1960 Data from Chicago

4.4.1 Data and Variables

Most of the data used in the calibration exercise are tract-level decennial census data for 1960 from the National Historical Geographic Information System (NHGIS, Minnesota Population Center (2004)). The first variable is the share of African-American residents in each census tract, which we use to define neighborhoods N1 and N2. This variable is the number of African Americans in each tract divided by the total number of residents.

N1 is defined as all Census tracts with a share black greater than or equal to 0.80 in 1960, and N2 is defined as all remaining census tracts in the city. Census tracts are part of N1 in subsequent years if they are contained within 1960's N1. Figure 2 shows the share black in Chicago census tracts in 1960. We can see that N1 contains Chicago's "Black Belt," the segregated area in which most of the city's African Americans lived.

Parameters are also calibrated to match moments from data on per-capita earnings, which we use to measure human capital. This variable is created as the aggregate income in each census tract divided by the total number of residents, where aggregate income is created from variables on the income of families and unrelated individuals, and then converted to 2005 dollars using the appropriate BEA GDP price deflator.

4.4.2 Calibration Results

Four model parameters are jointly calibrated to match four inter-neighborhood and intra-neighborhood moments. In addition to moments from the US Census data from Chicago in 1960, we also calibrate the model to match moments from the literature and the National Income and Product Accounts (NIPA). The moments targeted in the model calibration are displayed below in Table 2.

From the NIPA we use the Personal Consumption Expenditures data (Table 2.5.5) to calculate the ratio

$$\frac{\text{Housing Service Expenditures}}{\text{Non-Housing Expenditures}}$$

[17] In related models, Fisher and Gervais (2011) report difficulties matching important features of the data unless $\nu_s > \nu_c$, while Chambers et al. (2009) set $\nu_s < \nu_c$.

where Housing Service Expenditures are Housing + Utility + Fuels and Non-Housing Expenditures are Total Expenditures (PCE) - [Housing +Utility + Fuels] - Consumption of Non-Profit Institutions.

With respect to the intergenerational elasticity (IGE) of earnings, Solon (1999)'s survey concludes that the correlation among American brothers in the permanent component of their log earnings is somewhere around 0.4, and that most of the estimates of the IGE in the literature fall in a range between 0.3 and 0.5. While there is evidence that the IGE is higher (Mazumder (2005)) or lower (Behrman and Taubman (1985)), we target 0.4 in part because of Aaronson and Mazumder (2008)'s estimate of a 0.43 time-invariant IGE between 1950 and 2000.

We must set several additional model parameters in order to calibrate the model. Some parameters are set to values within the plausible ranges found in the literature, like $\nu_c = 1.5$ and $\nu_s = 2.0$.[18] We set $\beta = 0.67$ so that the complete-market annualized interest rate equivalent in our model is 3 percent for 15-year periods. The total factor productivity (TFP) parameter Z_{N2} is an arbitrary scaling factor, which we set to 5. The housing production technology parameter α is set to generate a price elasticity of supply of 1.77.

We specify the Markov chain stochastic process for child's ability using the Rouwenhorst method (Kopecky and Suen (2010)). This method approximates an AR(1) process. We set the number of ability states to 9 and assume draws are i.i.d. over time.[19] This leaves σ_a as the one parameter to be calibrated for the ability process.

Table 1 lists the values of the parameters of the calibrated model.

[18] While our utility function is similar in form to the ones used in Fernandez and Rogerson (1998) and Badel (2010), our parameterization implies that housing and non-housing are substitutes with an elasticity of substitution of 1.33. The elasiticy in Fernandez and Rogerson (1998) is –0.6 (complements), and in Badel (2010) it is 1.0.

[19] Originally, ρ was also calibrated but was found to be very close to zero (less than 0.01). We set ρ to zero to get sharper identification. For evidence of very low ability persistence across generations see (Black et al. (2015)).

Table 1: Model Parameters

	Parameter	Value	Identification
Preferences: $u(c,s) = \frac{c^{1-\nu_c}}{1-\nu_c} + \theta \frac{s^{1-\nu_s}}{1-\nu_s}$			
Utility Function (Consumption)	ν_c	1.5	Set by the Authors
Utility Function (Housing Services)	ν_s	2.0	Set by the Authors
Utility Function (C v S)	θ	0.09	Calibrated
Time Preference	β	0.67	Set by the Authors
Production Function: $h' = Z_n\Big(\left(\frac{a}{3}\right)^{\frac{\gamma-1}{\gamma}} + \left(\frac{i}{3}\right)^{\frac{\gamma-1}{\gamma}} + \left(\frac{\chi_n}{3}\right)^{\frac{\gamma-1}{\gamma}} \Big)^{\frac{\gamma}{\gamma-1}}$			
TFP	Z_{N1}	4.24	Calibrated
TFP	Z_{N2}	5.00	Set by the Authors
Elasticity of Substitution	γ	0.91	Calibrated
Ability Process: $\pi(a_i \vert a_j)$			
Standard Deviation of Shocks	σ_a	0.59	Calibrated
Firm's Production Function: $Y = Y(H,L) = H^{\alpha}L^{1-\alpha}$			
Technology Parameter	α	0.64	Set by the Authors

The targeted moments in the data are displayed in Table 2 alongside the moments generated by the calibrated model.[20] The model moments are very close to their data counterparts, with a maximum percentage deviation from the target of 4.9×10^{-4}. Given the relatively small number of adjustable parameters, the model does a good job of capturing inequality in both neighborhoods. The distribution of per-capita income for each neighborhood in the 1960 is shown in Figure 6 with the distribution from the model's steady state.

Table 2: Moments Used to Calibrate the Model

Moment	Source	Data 1960	Model Steady State
H_{N1}/H_{N2}	1960 US Census	0.56	0.56
$CORR\,(ln(h), ln(h'))$ in N2	Aaronson and Mazumder (2008), Solon (1999)	0.40	0.40
$Q^h(0.75)/Q^h(0.50)$ in N1	1960 US Census	1.18	1.18
pS/C in N2	1960 NIPA	0.22	0.22

[20] $Q^X(q)$ represents the qth quantile of the distribution of the random variable X.

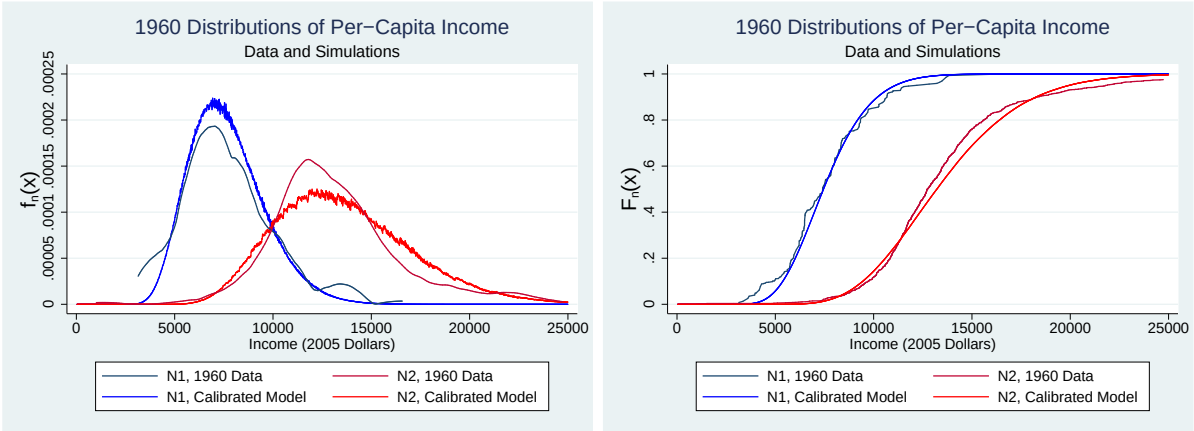


Figure 6: Distributions in 1960, Data and Simulations

5 Numerical Experiments

To learn about sorting and neighborhood externalities, we conduct three numerical experiments in which we either improve N1's production technology, allow for neighborhood choice, or both. For each experiment we find a new steady state and transition path. We associate these experiments with three prominent figures:

X : This counterfactual maintains the restriction on mobility between N1 and N2 while equalizing their production technologies, which we associate with Malcolm X's ex-ante vision of separation. X's goal for black nationalism in X (1990) can be interpreted as equalizing production technologies while maintaining the status quo for residential choice.

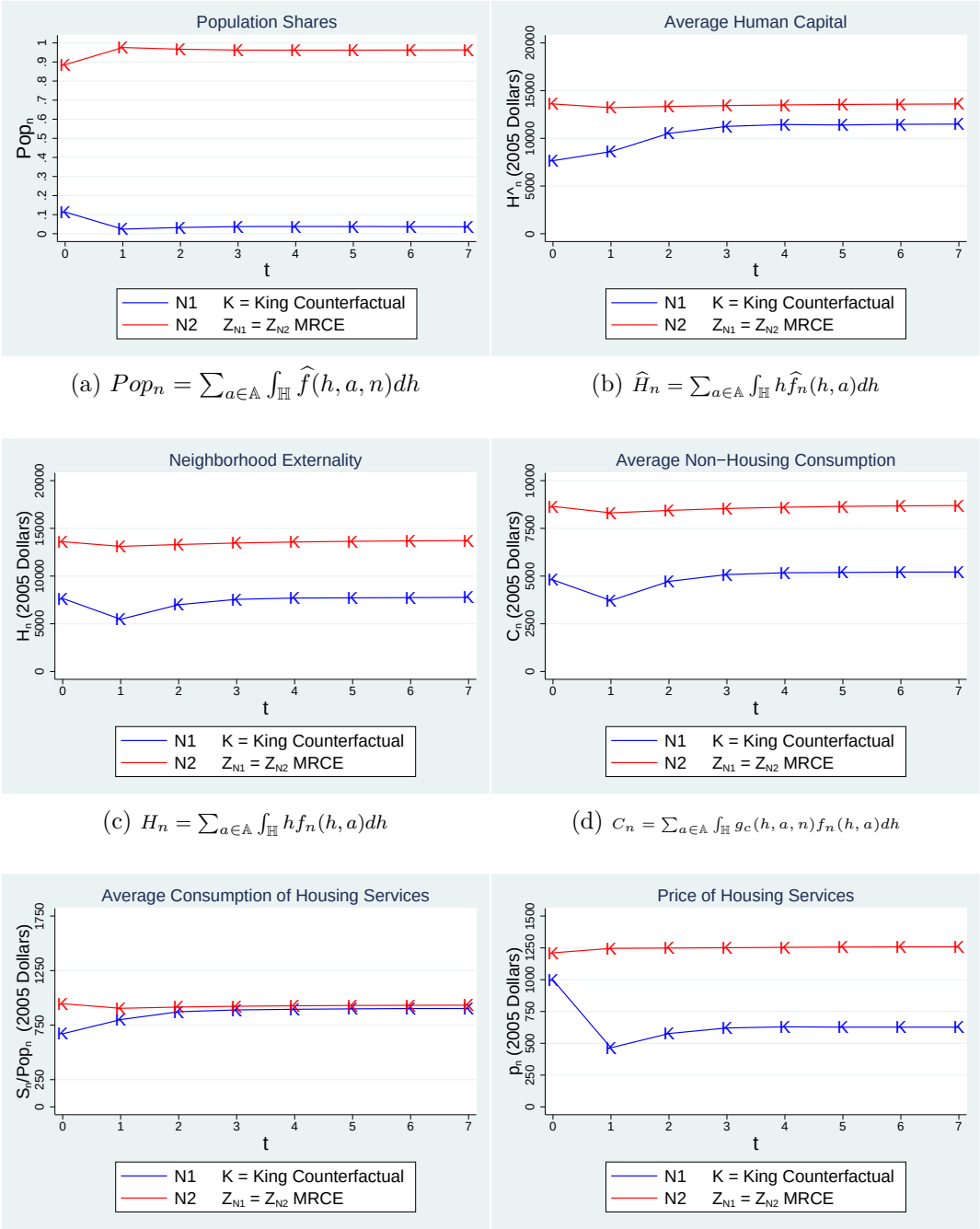
Wilson : This counterfactual allows for mobility with unequal production technologies in N1 and N2, and can be interpreted as the counterfactual resulting from eliminating legal racial discrimination in housing. This is one of the central thought experiments suggested by Wilson (1987)'s ex-post analysis of how residential sorting and neighborhood externalities contributed to outcomes in Chicago between 1970 and 1980. See pages 46-62 of Wilson (1987).

King : This counterfactual allows for neighborhood choice while also equalizing production technologies across neighborhoods. We interpret this counterfactual as Martin Luther King, Jr.'s ex-ante vision for the integration of Chicago.[21] While King is often remembered in terms of his work for open housing, integrating schools was also a primary

[21] A location-specific production technology might also be interpreted as an exogenous location-specific public good. Banzhaf and Walsh (2013) study how changing such a public good impacts segregation in a model with race preferences.

focus of his work in Chicago, and improving the general conditions in N1 was another major goal. See Chapter 28 of King (1998) for a description of how King's work in and vision for Chicago can be interpreted as equalizing production technologies across neighborhoods in addition to allowing for residential choice.

5.1 The Transition Paths



(a) $Pop_n = \sum_{a\in\mathbb{A}} \int_{\mathbb{H}} \widehat{f}(h,a,n)dh$

(b) $\widehat{H}_n = \sum_{a\in\mathbb{A}} \int_{\mathbb{H}} h\widehat{f}_n(h,a)dh$

(c) $H_n = \sum_{a\in\mathbb{A}} \int_{\mathbb{H}} hf_n(h,a)dh$

(d) $C_n = \sum_{a\in\mathbb{A}} \int_{\mathbb{H}} g_c(h,a,n)f_n(h,a)dh$

(e) $\overline{S}_n = \sum_{a\in\mathbb{A}} \int_{\mathbb{H}} g_s(h,a,n)f_n(h,a)dh$

(f) $p_n = \frac{1}{\alpha} S_n^{\frac{1-\alpha}{\alpha}}$

Figure 7: The Transition Path after Moving Allowed and Technologies Equalized

Figure 7 shows the transition path to the King steady state. Population flows out of N1 immediately, reducing its city population share to 2.4 percent (Figure 7a). With equalized production technologies, the externality in N1 improves over time and some households move

into N1 in response to lower housing services prices. In the final steady state, 3.9 percent of the city lives in N1; and there is significant amount of sorting across neighborhoods, with migrants into N1 having low- to moderate- human capital.[22] Figure 7c shows that the externality in N1 decreases slightly in period 1, as the right tail of N1 exits. Over time, the outflow of human capital is partially offset as households in N1 increase their investment in response to the change in technology. As a result, N1 average human capital rises (Figure 7b). The increase in investment early in the transition causes consumption of non-housing goods to dip in the early periods of the transition before returning to its initial value later, as income rises. Despite an increase in average housing demand, the price in N1 falls after allowing sorting because the congestion ratio decreases (Figure 7f).

N2 experiences little change after sorting. This is because the measure of immigrants to N2 make up a small fraction of the total N2 population, and also because many of these migrants are near the N2 average. Figure 8 displays how human capital is distributed across migrant households. With the exception of a small overlap, the distribution of migrants to N1 are poorer than those to N2. The average human capital of entrants into N1 is approximately the steady state average in N1, while the average human capital for households moving to N2 is 11 percent lower than the steady state mean in N2. The housing services price in N2 increases by 4 percent due to increased congestion.

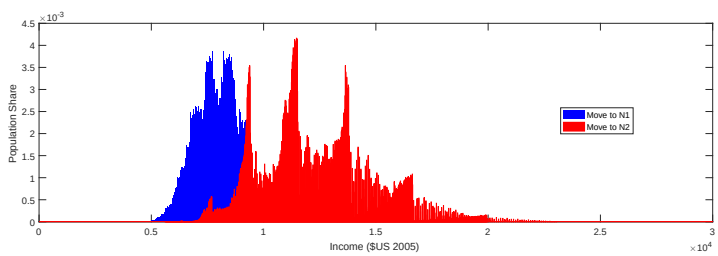
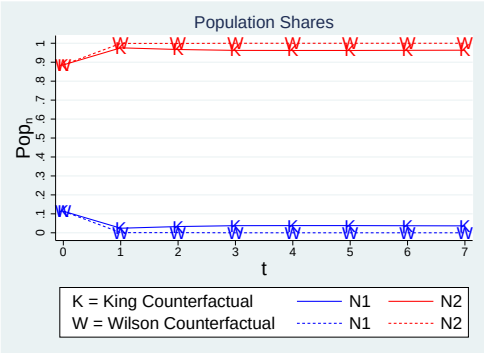


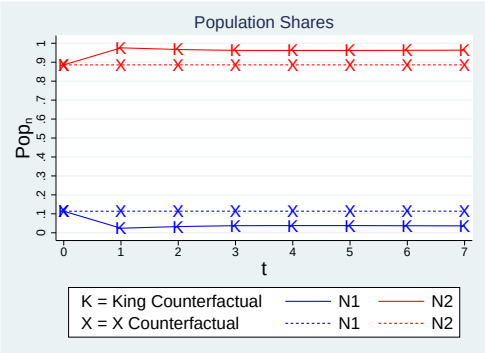
Figure 8: Distribution of migrants' wealth in King steady state

Figure 9 compares the King transition to the X and Wilson transitions. The X policy of equalizing production technologies leads to a gradual increase in N1's human capital as residents respond to the improved technology (Figure 9d). Eventually, the transition reaches a new steady state where neighborhoods are completely equal (with the exception of population share). This steady state is equivalent to a single neighborhood with land mass $L_1 + L_2$. Naturally, N2 is unaffected because the production technology is the same in either case.

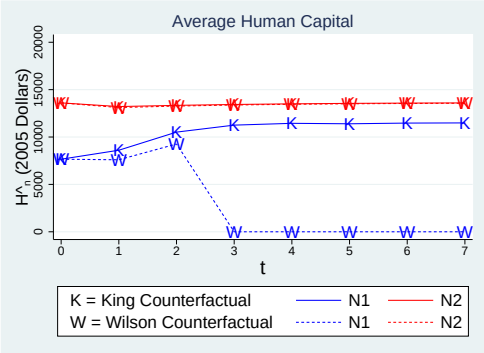
[22]Moving decisions in the transition are close to the steady state decisions, which are shown in Section 5.2.



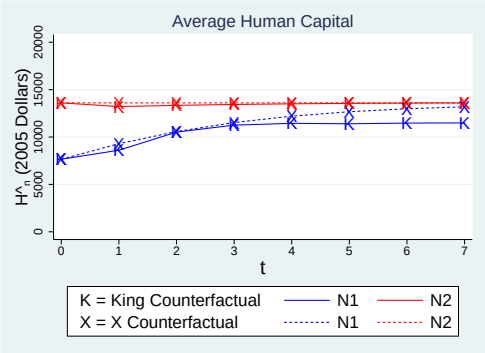
(a) Moving Allowed



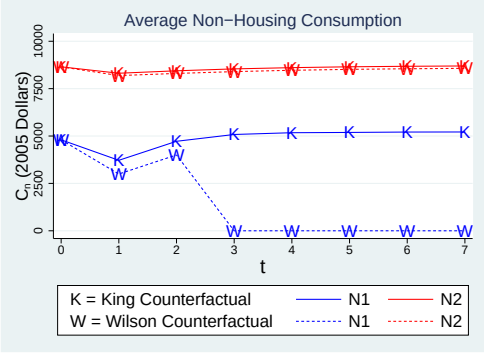
(b) Tech Equalized



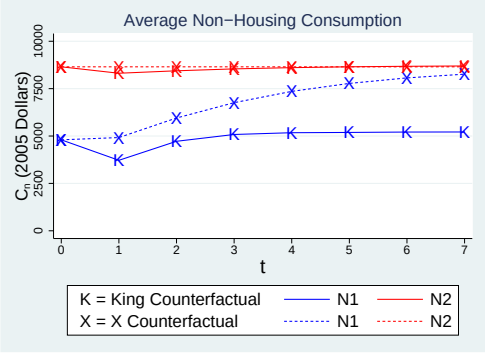
(c) Moving Allowed



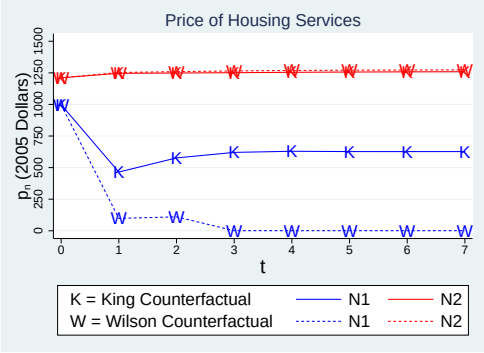
(d) Tech Equalized



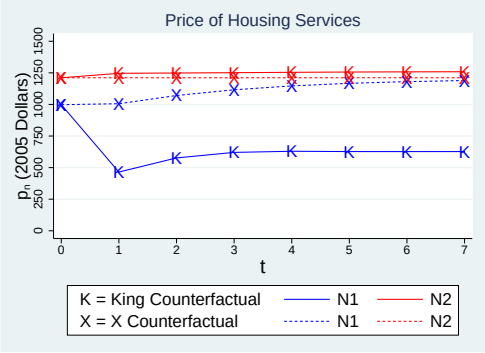
(e) Moving Allowed



(f) Tech Equalized



(g) Moving Allowed



(h) Tech Equalized

Figure 9: The Transition Path after Moving Allowed or Technologies Equalized

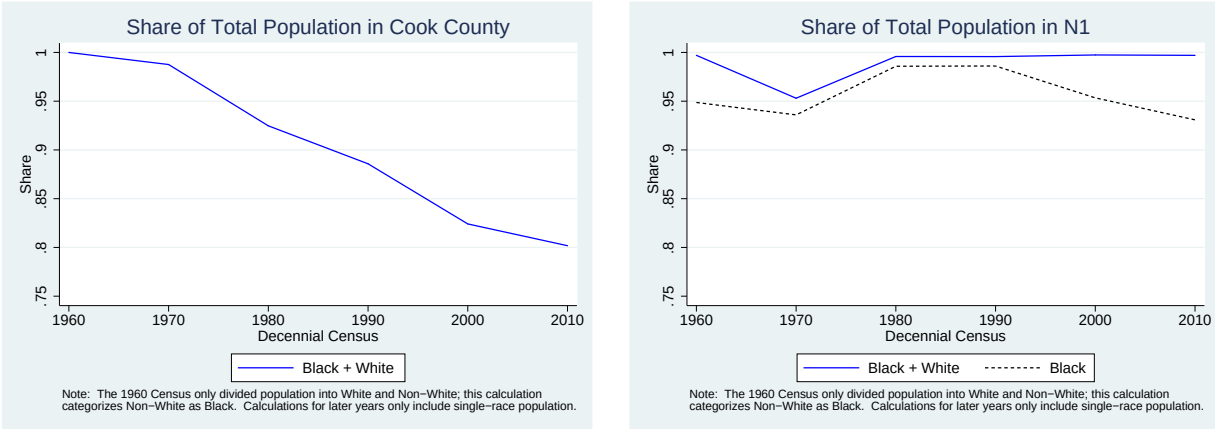
Figure 9d shows how sorting positively impacts N2 at the expense of N1. Although N1 average income rises when technologies are equalized and households have a choice over location, it rises more when they do not have a location choice. This is due to the the positive contribution to the production externality from high-income N1 residents. If moving is allowed, those households go to N2 early in the transition, stunting the income prospects for N1.

N1's fate is worst when moving is allowed but the technology in N1 is unchanged. If TFP is left unequal across locations while mobility is allowed, N1 is rapidly abandoned, and only very poor households would choose to live there. The population share in N1 declines from 11.4 to 0.0 percent in three periods (Figure 9a). In the end, this experiment also effectively results in a single neighborhood, but one in which the distribution of households across neighborhoods is less efficient than in the X steady state.

To put these numerical experiments in context, Table 3 compares the transition paths with data from 1960-2010. We think our model is most applicable until around 1990. After that date, Hispanic migration into Chicago changes N2 (Figure 10a) and gentrification changes N1 (Figure 10b).

Table 3: Comparing Transition Paths with the Data

Year (Data)	1960	1970	1980	1990	2000	2010
t (Model)	0	1		2	3	
Population in N1 (%)						
Data	11.4	8.3	5.8	4.1	3.8	7.2
Wilson	11.4	0.1		0.1	0.0	
King	11.4	2.4		3.3	3.8	
X	11.4	11.4		11.4	11.4	
Human Capital (H_{N1}/H_{N2})						
Data	0.56	0.54	0.49	0.41	0.52	0.78
Wilson	0.56	0.58		0.70	h_{min}	
King	0.56	0.65		0.79	0.84	
X	0.56	0.68		0.78	0.85	

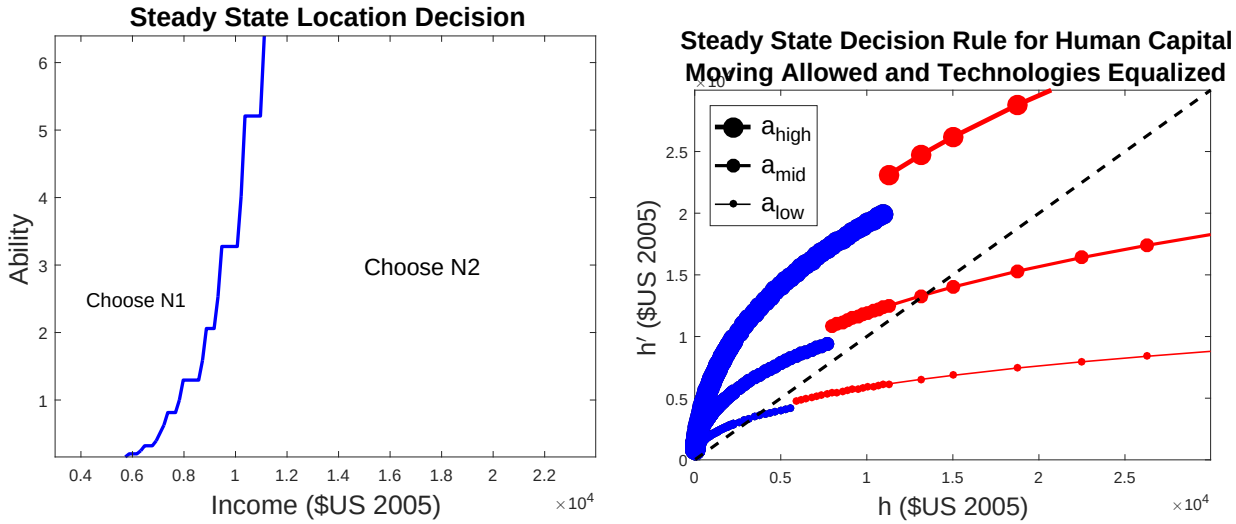


(a) Cook County (b) N1

Figure 10: City and Neighborhood Racial/Ethnic Composition

5.2 The New Steady States after Allowing for Moving

Figures 11a and 11b show the residential location and human capital decision rules in the King steady state. Figure 11a illustrates that conditional on ability, there is some human capital level h_a for which households below h_a choose to live in N1, while households with $h > h_a$ live in N2. This figure also shows that high-ability households move to N2 at a higher income level, while low-ability households will move to N2 with a lower income.



(a) $\widetilde{g}_n$ (b) $\widehat{g}_{h'}$

Figure 11: King Counterfactual: Moving Allowed and Technologies Equalized

The human capital decision rules in Figure 11b help to understand how households cycle through the distribution of h, a, and n in the King steady state. Low-h households choose

to reside in N1. The only way to leave N1 is to draw a string of sufficiently high ability shocks, so as to increase the household's h. Households in N2 remain there, cycling around the human capital distribution in response to ability shocks, until they draw a string of sufficiently low ability shocks.

The role of ability is especially stark when looking at the decision rule of the lowest ability households in Figure 11b. Receiving the lowest ability draw drastically limits the amount of h' passed on to the next generation. Figure 12 plots h', h'', and h''' as function of initial income for households starting that receive a string of the lowest realizations of ability. The horizontal line marks the income threshold below which households with the lowest ability live in N1. Low income households choose to live in N1 upon realizing a bad ability shock, while richer households choose N2. Although some high income households stay above the threshold for several periods, three consecutive lowest ability draws is sufficient to guarantee that a household lives in N1.

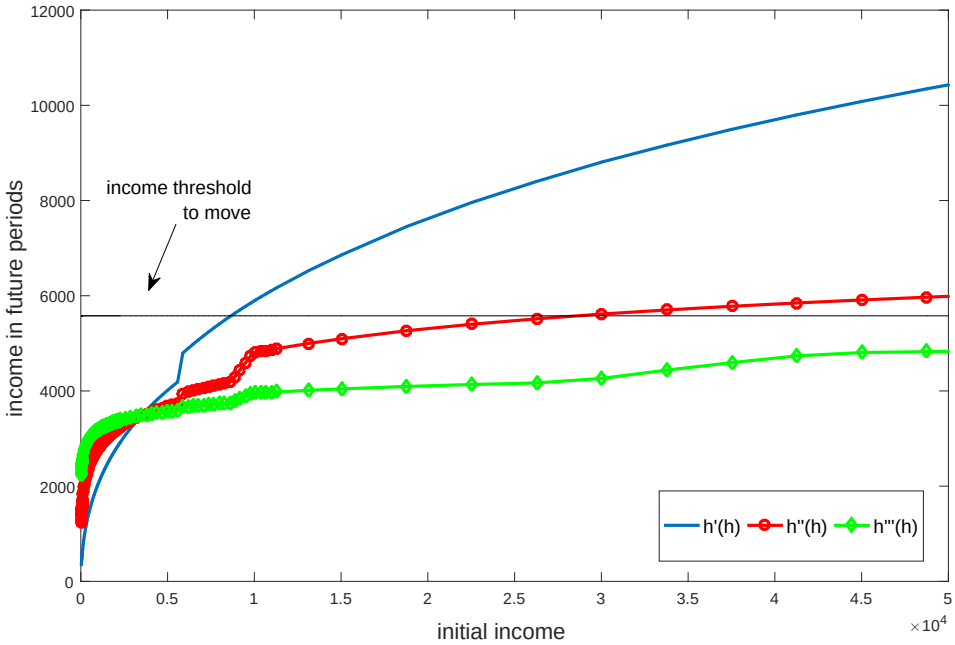


Figure 12: Low a Movements

When production technologies remain unequal (in the Wilson steady state), the income threshold at which households move to N2 is so low that N1 is empty. Furthermore, for h below this threshold, the human capital decision rule is above the 45 degree line for all ability types, meaning N1 is not a poverty trap. Even if a zero-measure group of households were exogenously relocated from N2 to N1, they would eventually accumulate enough income to escape, after which they would never revisit N1.

5.2.1 Multiplicity

Multiple steady state equilibria are almost certain to arise when there are location-specific externalities and households are allowed to sort freely between locations. We search for multiple equilibria as follows: Given fixed neighborhood externalities χ_{N1}, χ_{N2}, we record what the calibrated model implies to be the steady state average human capital in neighborhoods N1 and N2, H_{N1} and H_{N2}. By recording the zeros of the functions $f_1(\chi_{N1}, \chi_{N2}) = H_{N1} - \chi_{N1}$ and $f_2(\chi_{N1}, \chi_{N2}) = H_{N2} - \chi_{N2}$, we can search for equilibria by finding the sets

$$\{(\chi_{N1}, \chi_{N2}) : f_1(\chi_{N1}, \chi_{N2}) = f_2(\chi_{N1}, \chi_{N2}) = 0\}.$$

We find that there are equilibria in addition to those just studied in Section 5.2, but that the only empirically relevant equilibria are those studied in Section 5.2. Figure 13 plots the zeros of f_1 and f_2 in the Wilson counterfactual. The intersections of these zero functions indicate steady state equilbria.

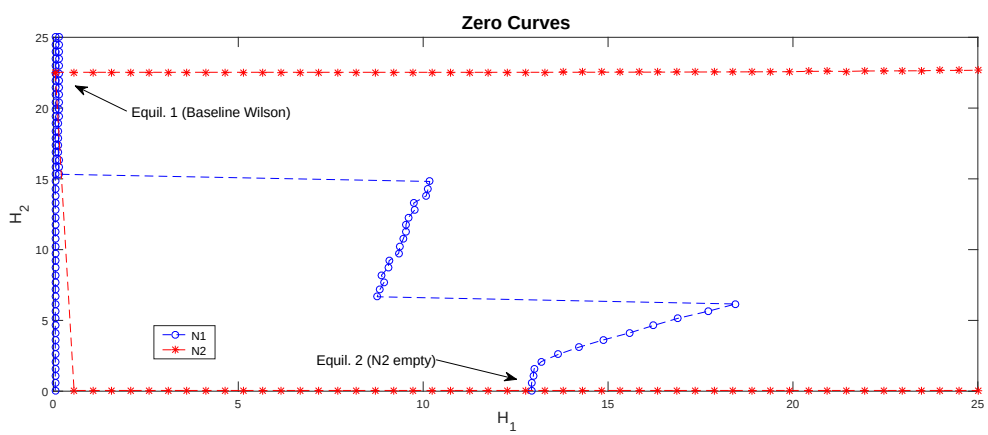


Figure 13: Implied Externalities in the Wilson Counterfactual

The first thing to notice is that there are two equilibria, and both feature one empty neighborhood. Equilibrium 1 is the equilibrium described in Section 5.2: N1 is empty, and there is high congestion in N2. The other equilibrium is a bit perverse. This would result if once moving restrictions were lifted, households in N2 moved en masse into N1. This would cause the externality in N2 to fall to a sufficiently low level that no household would want to move back even if the house price we very low.

We do not consider Equilibrium 2 to be empirically relevant. While Equilibrium 2 exists, it clearly did not occur, with the transition to Equilibrium 1 being far closer to what we observe in the data. Equilibrium 2 is also unambiguously worse than Equilibrium 1. Equilibrium 2 arises from an extreme coordination failure where households all choose to live in a

smaller area and thus pay a much higher housing cost (since there is much less land in N1). All N2 households would have to believe that all of their neighbors would want to move to this unambiguously worse situation. The fact that N1 is initially poorer, and hence has a lower externality, reinforces this logic.

Both zero functions display discontinuities. For f_2 this occurs near $(\chi_{N1}, \chi_{N2}) = (h_{min}, h_{min})$. Fixing $\chi_{N2} = h_{min}$, for sufficiently large χ_{N1} the implied average human capital in N2 will be h_{min}, since N1 is assumed to have a much better externality in these cases. However, as χ_{N1} approaches h_{min}, the difference between χ_{N1} and χ_{N2} becomes small. Since N2 has higher TFP in the Wilson counterfactual, it becomes the unambiguously better location to accumulate human capital, and f_2 jumps to Equilibrium 1.

There are two points of discontinuity for f_1. If χ_{N1} is high and χ_{N2} is low, then implied human capital H_{N1} is near its value in Equilibrium 3. As χ_{N2} gets larger, the TFP difference between the two locations swamps the externality difference. After crossing a threshold χ_{N2} many (but not all) N1 households would choose to live in N2, and H_{N1} is about 50 percent lower after crossing this threshold. A similar discontinuity arises as χ_{N2} increases.

Figure 14 shows the zero-curves for the King counterfactual. When TFP is equalized the number of steady state equilibria increases. First, there are still two equilibria with one empty neighborhood (marked 'Equil. 2' and 'Equil. 3' in the figure). Reaching either equilibrium requires that the initial guess at χ_n be extremely low (so that households must believe that almost everyone will exit). Absent a very low initial guess, the economy goes to the only equilibrium where neighborhoods are asymmetric and non-trivial measures of the population reside in both locations (marked 'Equil. 1'). This corresponds to the steady state King equilibrium just examined in Section 5.2.

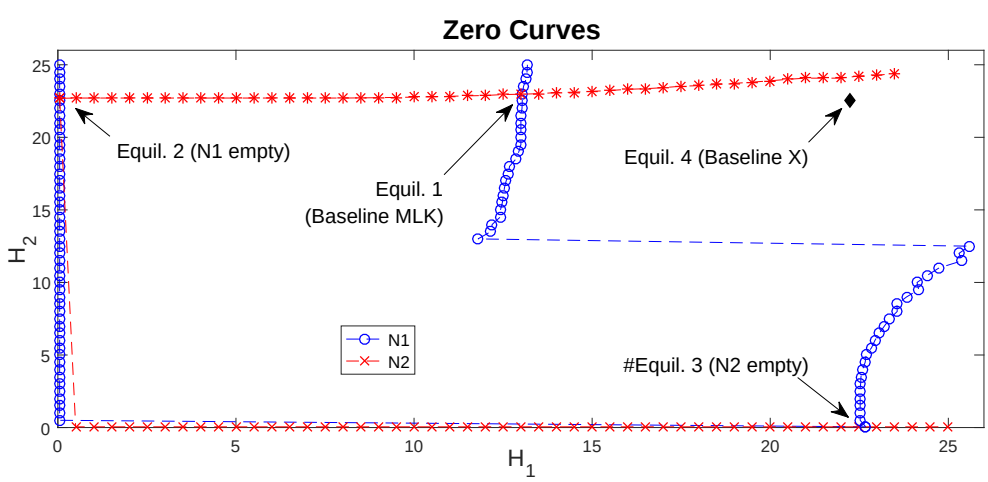


Figure 14: Implied Externalities in the King Counterfactual

We find only one symmetric equilibrium in the King counterfactual. In this case, N1 is a less populated replica of N2, and the population shares are equal to their values in the segregated steady state (so there is no excess congestion in either neighborhood). We do not study this equilibrium because it is very unstable; small deviations in prices or externalities cause the economy to move away from this steady state, and it never returns.

5.3 Dynamics under Myopic Beliefs

It is difficult to study our model with myopic beliefs because it is fundamentally different from a repeated static model. In general, solving the model without rational expectations requires imposing a partial equilibrium condition on the household problem. For example, if households believe that they are in a steady state when they are not, solve their problem, and move according to their decision rules, then the house price and externality they face will not be the ones for which they solved their problem.

Since the price of housing services enters the household's budget constraint, beliefs and reality about this price must be the same for the household to have a well-posed problem. However, because the externality does not enter the household's budget constraint, beliefs about the externality can be mistaken. This allows us to solve a version of the model in which households' beliefs about current-period prices are correct but beliefs about the externality are mistaken.

Define $b_s(H_t)$ as a household's time s beliefs about the time t vector of human capital externalities, and define $b_s(p_t)$ analogously. Under myopia, households believe that the intra-period average human capital is equal to beginning period average human capital, $b_t(H_{t+m}) = \widehat{H}_t$ for all $m \geq 0$. This allows beliefs about the human capital externality to be incorrect both within periods and across periods:

$$b_t(H_{t+m}) \neq H_{t+m} \quad \forall \quad m \geq 0.$$

In contrast, beliefs about housing prices are only mistaken across periods:

$$\begin{aligned} b_t(p_{t+m}) &= p_{t+m} \quad \text{for} \quad m = 0; \\ b_t(p_{t+m}) &\neq p_{t+m} \quad \forall \quad m \geq 1. \end{aligned}$$

Figure 15 shows belief updating in this scenario.

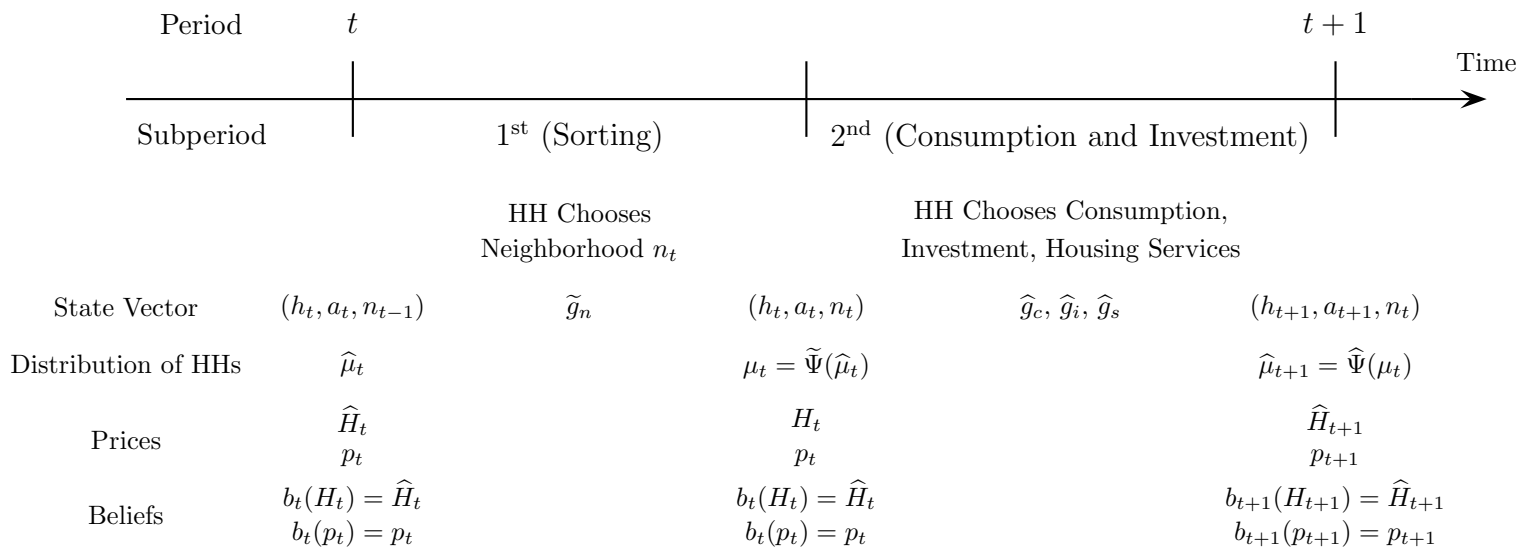


Figure 15: Myopic Beliefs about Human Capital Externalities

Myopic beliefs lead to dynamics in the Wilson counterfactual that could hardly be more different from the dynamics under rational expectations. Myopia essentially solves the households' coordination problem in N1: Since no one forecasts the downfall of the neighborhood, no one leaves. As a result, the neighborhood does not decline.

Figure 16 plots the path of N1's externality in the Wilson counterfactual under rational expectations (blue) along with the paths of N1's externality and beliefs under myopia (red). In the myopic case, the externality remains near its initial value since a significant measure of households remain in N1. Because households do not take into account how moving behavior changes the externality in their neighborhood, they are consistently overly optimistic about the value of the externality. As a result, many households choose to stay that would leave under rational expectations. This, along with the fact that households overinvest in human capital (again because they misjudge the marginal benefit of investing), causes average human capital to actually increase in N1 after a small initial decline.

Contrast this with the neighborhood decline generated in the forward-looking case. In that case, households forecast the decline of N1 and all households eventually leave. High-income households leave first, since they are best-suited to pay for the higher housing price in N2. This leads to a reduction in the human capital in N1, which then leads to the depopulation of the neighborhood.

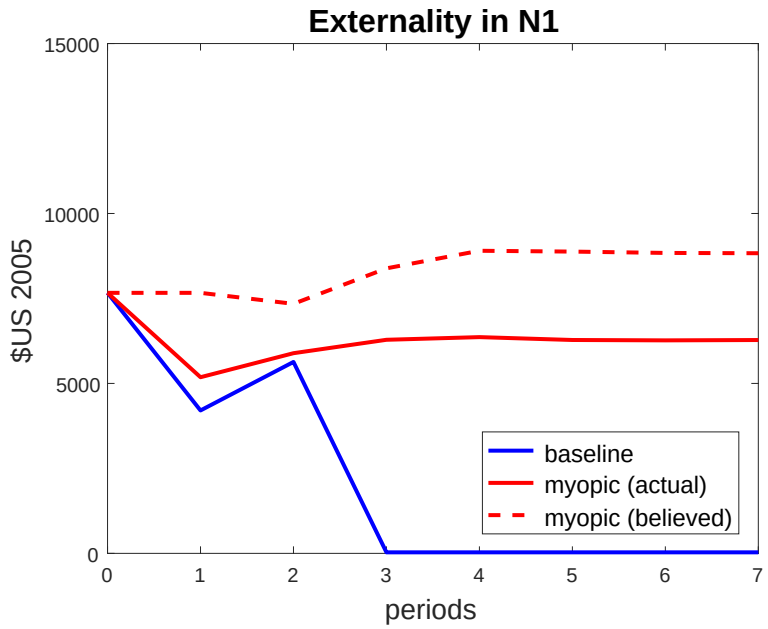


Figure 16: Implied Path of N1 in the Wilson Counterfactual

In the King counterfactual where TFP is equalized, myopia again works in N1's favor as there is slightly higher human capital on average in the long run. The overestimation of future externality values again leads more high income residents to stay in N1. In addition, households invest more in their children than they do under the baseline.

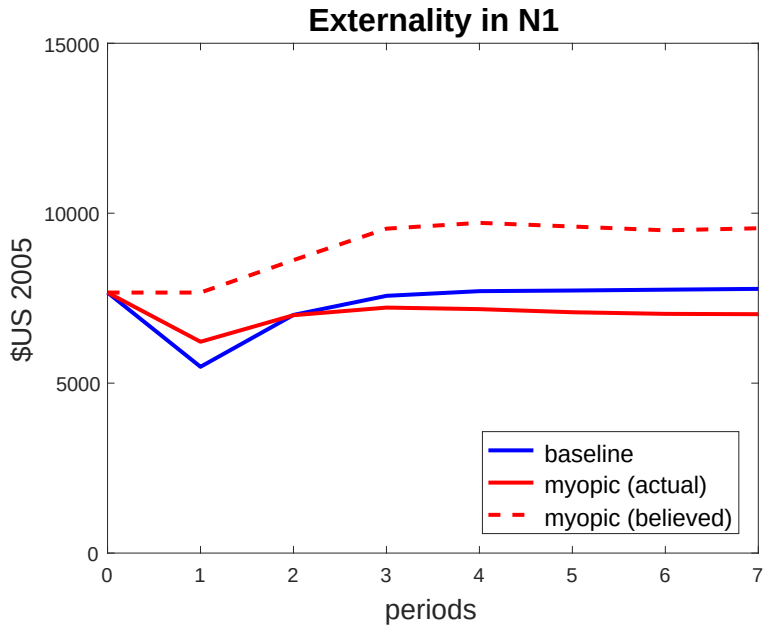
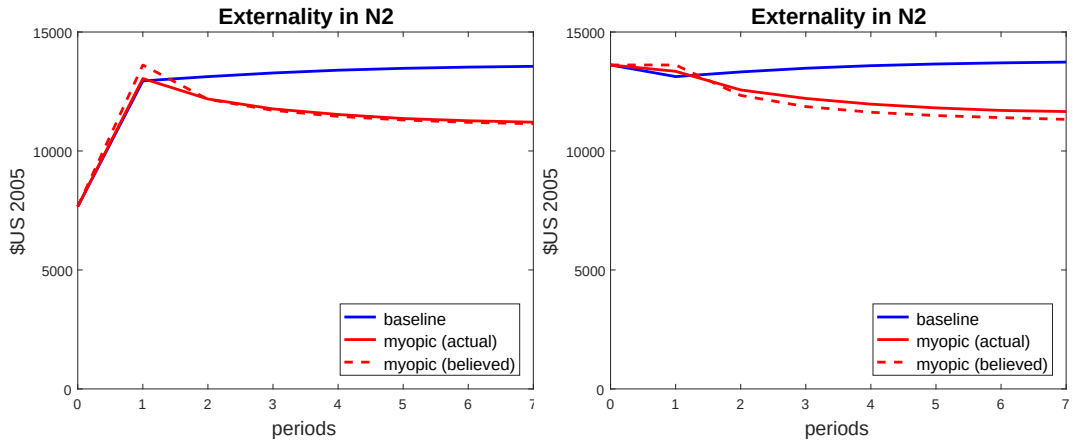


Figure 17: Implied Paths of N1 in the King Counterfactual

The paths for N2 are little changed in either the Wilson or the King counterfactuals

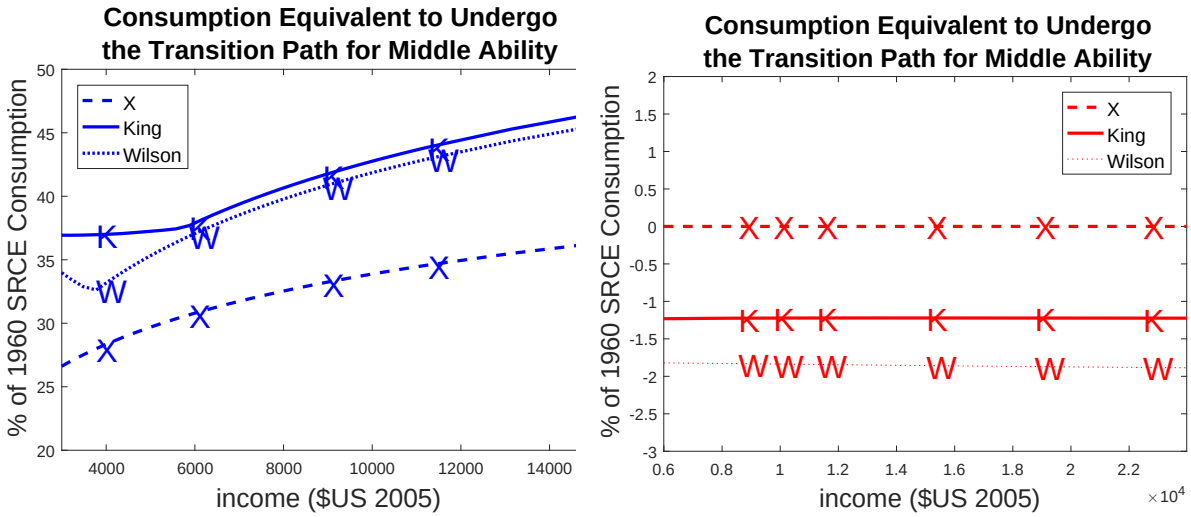
because the differences from mistaken beliefs are not as consequential as they are for N1. While households continually underestimate the externality, the error is small, so that the differences between the baseline and myopic cases are not as significant.



(a) The Wilson Counterfactual (b) The King Counterfactual

Figure 18: Implied Paths of N2 under Myopia

5.4 Welfare



(a) Middle Ability in N1 (b) Middle Ability in N2

Figure 19: Welfare

Figure 19 shows consumption equivalents for the counterfactual policies. This is the percent change in a household's consumption in the original steady state that would be required for them to be indifferent between remaining in the 1960 steady state and undergoing the transition to a steady state under a new policy. A first detail to notice is that every

household in N1 prefers any of the three transitions to remaining in the initial segregated equilibrium.

Starting first with households initially in N1, the King transition is preferred to the other two polices by every household. The intuition for why the King transition is preferred to Wilson is straightforward. Both N1 and N2 are more attractive locations under King than they are under Wilson.

The reasons households initially in N1 prefer the King transition to the X transition are a bit more nuanced, and are different across the income distribution. For high income households in N1, the King policy allows households to move to N2 where they benefit from the higher externality. For poorer households in N1, the King transition is better in two respects. First, the price of housing services in N1 one falls considerably. Because these households have little resources to invest, they do not benefit as much from a high externality. Second, under the King policy a poor household has the option to switch locations if it becomes richer in the future (for instance, through a string of high ability draws) so expected future utility is increased. While it is true that the X policy leads to a similar steady state in N2 as King does, it takes some periods for the transition to approach that level. Under King, a high human capital location can be accessed immediately.

In contrast to their N1 counterparts, households initially located in N2 would prefer that residential sorting remain restricted. There is very little difference in the N2 externality across the three transitions; however the price of housing is different due to congestion. Under continued residential segregation, the X policy and initial SRCE policy are equivalent so there is zero welfare change. At the other extreme, the Wilson transition leads to greatest amount of congestion in N2 and is ranked lowest by initial N2 residents.

Table 4: Welfare Gain across Neighborhoods

Average Welfare (% of 1960 SRCE Consumption)			
	N1	N2	City
X	32.1	0.0	3.7
King	40.2	−1.2	3.5
Wilson	39.1	−1.9	2.8

The welfare consequences differ across neighborhoods not only in their sign but also in their magnitudes. Table 4 reports the welfare gain for each neighborhood as well as for the entire city. For N1, the largest factor for increasing welfare is residential mobility. King produces an additional welfare gain of 8.1 percentage points over X but only an extra 1.1 percentage points over Wilson. Although this finding may seem striking at first, it is not so

surprising given how quickly households exit N1 under Wilson. While the consequences of the Wilson policy are devastating for the N1 location, no households experience it for long. Overall, the X policy produces the largest average welfare gain for the city, but this is driven entirely by the larger relative size of N2.

There are changes to the land rents accruing to the absentee landlord. Table 5 displays the percentage change in the present discounted value rents to the absentee landlord when the landlord discounts the future at the same rate as households. The first column shows the total percentage change. The next two columns calculate the percentage change when either the congestion ratio or the average housing demand is fixed at its initial value. Rents increase in all three cases. Under X, the rise is entirely due to greater housing demand in N1. In the later two cases, however, housing demand would *decrease* rents but increased congestion outweighs this effect.

Table 5: Percentage Change in PDV of Absentee Landlord Rents

Percent Change from Initial Steady State			
	Total	No Congestion Change	No Demand Change
X	2.0	0.0	2.0
King	3.1	7.5	−1.0
Wilson	4.1	12.3	−7.0

6 Conclusion

In this paper, we solve a dynamic heterogeneous agents incomplete market model with location choice and local production externalities. We use a two-neighborhood model without moving to represent Chicago when it was legally segregated, and calibrate the model allowing for unequal technologies. We then consider three policy experiments. First, we remove location restrictions but leave production technologies unequal across neighborhoods. This results in a rapid and complete abandonment of N1 and high housing costs in N2 due to increased congestion. Second, we maintain the restrictions to location choice but equalize technologies. In stark contrast to the first experiment, average income in N1 slowly rises until in the long run N1 is a less populated replica of N2. Finally, we allow households to sort across locations and equalize technologies. This increases average income in both locations, but also leads to a stratified equilibrium where households with sufficiently high income locate in N2, while N1 becomes a haven where low-to-moderate income households move to consume cheap housing. Contrasting the last two experiments highlights the powerful role

location choice plays in determining the effectiveness of policies designed to increase income in impoverished neighborhoods.

To underscore the importance of forward-looking behavior for the predictions of our model, we solve a version of the model in which households are myopic. We find completely different dynamics under myopia. Because they do not take into account the choices of other agents, myopic households are consistently over-optimistic about the future of their neighborhood. As a result, N1 actually improves after allowing for sorting, as opposed to completely declining in the baseline.

This work is meant to be an initial step in quantifying neighborhood transitions with sorting and externalities. Some extensions could be promising. For example, this paper focuses in detail on a two neighborhood model, but extending the framework to more than two neighborhoods would be productive in general and likely necessary for some questions. While more neighborhoods could help to more accurately capture the externalities experienced in N1 (Pattillo (2003)), additional neighborhoods might be more important for characterizing N2. This could allow for "white flight," and could also allow for high-income black enclaves (Bayer et al. (2014)). Both of these topics would be best studied with additional considerations for racial preferences, which could be an important mechanism in accounting for some of the patterns documented in the data (Sharkey (2014), Sharkey (2008)), but we have abstracted from race preferences. We could also imagine using new data to explore the inputs to and functional form of the production function. Finally, an adaptation of our model could be used to study persistent wealth inequality between populations which have experienced segregation and adverse lending policies. We leave these ideas to future work.

Outline of the Appendix

We begin by presenting a proof of the existence of a general equilibrium in a model with uninsurable idiosyncratic ability risk and a production externality, but no housing sector and no mobility. We then generalize this proof to allow for housing, showing the existence of a Segregated Recursive Competitive Equilibrium (SRCE) from the model without moving used in the paper. The broad outline is as follows:

- Simplified Model without Moving

  A.2-A.4 One neighborhood and households get no utility from housing

  A.2 We first state the household problem and use results from Stokey et al. (1989), henceforth SLP, to show that it has a unique solution (i.e., a value function and optimal decision rule) and that the associated value function and decision rule have desirable properties.

  A.3 We then show that a unique stationary distribution of human capital and ability (h, a) exists for each parameterization of the model by appealing to Theorem 2 of Hopenhayn and Prescott (1992).

  A.4 Because all of these results apply to an economy in which the externality is fixed to be some level χ, we conclude by showing that there is an externality χ^* satisfying the required general equilibrium conditions.

- Model without Moving Used in the Text

  A.5 We then generalize this proof to the model under segregation used in the text in which households get utility from both consumption and housing

- Model with Moving Used in the Text

  B.1 We then discuss the conditions that would be required for existence of a general equilibrium of the model with moving used in the text.

A Appendix: Proof of Existence of a SRCE

A.1 Proof: The Household's Problem Has a Unique Solution for Fixed Externality χ

A.1.1 The Household's Problem

We begin by analyzing a model without a moving decision in which the externality χ is not an equilibrium object, but rather is externally set to some fixed value. Because

households optimize facing this value, we often use the subscript $_\chi$ to explicitly indicate that sets, decision rules, etc. pertain to the model with the fixed value χ. The state vector of an infinitely-lived household is (h_t, a_t), where the endogenous state variable is human capital $h_t \in \mathbb{H}_\chi \equiv [\underline{h}_\chi, \overline{h}_\chi] \subset R_+$ with $0 < \underline{h}_\chi < \overline{h}_\chi < \infty$, and the exogenous shock is ability $a_t \in \mathbb{A} \equiv \{\underline{a} = a_1, a_2, \ldots, a_k = \overline{a}\} \subset R_+$, which is a stationary Markov chain with transition probabilities denoted by $\pi(a_i|a_j)$ (We assume $a_i < a_{i+1}$ for $i = 1, \ldots, k-1$.). The correspondence $\Gamma_\chi : \mathbb{H}_\chi \times \mathbb{A} \rightrightarrows \mathbb{H}_\chi$ describes the set of all feasible actions taken by the household at time t, whose image is $\Gamma_\chi(h_t, a_t) \subset \mathbb{H}_\chi$. We denote the graph of Γ_χ by

$$\mathrm{gr}\Gamma_\chi \equiv \{(h_t, a_t, h_{t+1}) \in \mathbb{H}_\chi \times \mathbb{A} \times \mathbb{H}_\chi : h_{t+1} \in \Gamma_\chi(h_t, a_t)\},$$

and at times we will also refer to the feasibility correspondence for a fixed $a_i \in \mathbb{A}$, Γ_{χ,a_i} and its graph $\mathrm{gr}\Gamma_{\chi,a_i} \equiv \{(h, h') \in \mathbb{H}_\chi \times \mathbb{H}_\chi : h' \in \Gamma_\chi(h, a_i)\}$. Letting $\mathcal{H}_\chi$ be an element of the Borel $\sigma-$algebra over $\mathbb{H}_\chi$ (denoted $\mathcal{B}(\mathbb{H}_\chi)$) and $\mathcal{A}$ be an element of the finite $\sigma-$algebra generated by the singletons $\{a_i\}_{i=1}^k$ (denoted $\sigma\big(\{a_i\}\big)$), let $x \in S = \mathbb{H}_\chi \times \mathbb{A}$ and define $\mathcal{S}_\chi$ to be the product $\sigma-$algebra generated by $\mathcal{B}(\mathbb{H}_\chi)$ and $\sigma\big(\{a_i\}\big)$, containing subsets of the form $B = \mathcal{H}_\chi \times \mathcal{A}$.

The household's preferences over streams of consumption $\{c_t\}_{t=0}^\infty$ are described by the discounted expected sum of period utility $u(c_t)$:

$$U = E_0 \sum_{t=0}^{\infty} \beta^t u(c_t) \quad \text{with} \quad \beta \in (0,1).$$

Each period a household chooses consumption c_t and next period's human capital h_{t+1} while respecting the period budget constraint

$$c_t + F(a_t, \chi_t, h_{t+1}) \leq h_t$$

where $F : \mathbb{A} \times \mathbb{H}_\chi \times \mathbb{H}_\chi \to \mathbb{R}$ is the cost of producing h_{t+1} units of human capital given this period's ability a_t and neighborhood externality χ_t. If the budget constraint holds with equality, the feasible set of consumption given χ and a_i is

$$\mathbb{C}_{\chi,a_i} \equiv \{c \in \mathbb{R} | c(h, h') = h - F(a_i, \chi, h') \text{ with } (h, h') \in \mathrm{gr}\Gamma_{\chi,a_i}\}.$$

We assume the cost function $F(a_t, \chi_t, h_{t+1})$ is twice continuously differentiable. In Section A.1.2 we state assumptions about the derivatives of F that will be crucial in our proofs.

We can formulate the household problem recursively as:

$$V_\chi\left(h, a_i\right) = \max_{c,h'} \;\; u\left(c\right) \; + \; \beta \sum_{j=1}^{n} V_\chi\left(h', a_j\right) \pi(a_j|a_i) \tag{7}$$

$$\text{s.t.} \quad c + F\left(a_i, \chi, h'\right) \leq h \tag{8}$$

because the following conditions hold for our model:

Condition 1: $\Gamma_{\chi,a_i} : \mathbb{H}_\chi \rightrightarrows \mathbb{H}_\chi$ is non-empty, compact-valued, and continuous for all $a_i \in \mathbb{A}$;
Condition 2: $u : \mathbb{C}_{\chi,a_i} \to \mathbb{R}$ defined by $u\left[c(h, h')\right]$ is bounded and continuous;[23]
Condition 3: $u : \mathbb{H}_\chi \to \mathbb{R}$ defined by $u\left[c(\cdot, h')\right]$ is strictly increasing;
Condition 4: $\Gamma_{\chi,a_i} : \mathbb{H}_\chi \rightrightarrows \mathbb{H}_\chi$ is increasing for all $a_i \in \mathbb{A}$;
Condition 5: For all $a_i \in \mathbb{A}$, $\theta \in (0, 1)$, and pairs (h_1, h_1'), $(h_2, h_2') \in \text{gr}\Gamma_{\chi,a_i}$, $u : \mathbb{C}_{\chi,a_i} \to \mathbb{R}$ satisfies

$$u\left[\; c\Big(\theta(h_1, h_1') + (1-\theta)(h_2, h_2')\Big) \;\right] \geq \theta u\left[\; c(h_1, h_1') \;\right] + (1-\theta) u\left[\; c(h_2, h_2') \;\right];$$

Condition 6: $\text{gr}\Gamma_{\chi,a_i}$ is convex for all $a_i \in \mathbb{A}$.

More specifically, Conditions 1 and 2 ensure there exists a unique value function $V_\chi : \mathbb{H}_\chi \times \mathbb{A} \to \mathbb{R}$ solving Equation 7 with a non-empty set of feasible plans (Thm 9.6 of SLP). Furthermore, the recursive formulation of the problem has desirable properties. Conditions 3 and 4 ensure that $V_\chi(\cdot, a) : \mathbb{H}_\chi \to \mathbb{R}$ is strictly increasing (Thm 9.7 of SLP), while Conditions 5 and 6 ensure that $V_\chi(\cdot, a)$ is strictly concave and that the optimal policy rule (ie, the decision rule $g_{h,\chi} : \mathbb{H}_\chi \times \mathbb{A} \to \mathbb{H}_\chi$) is a continuous (single-valued) function (SLP Thm 9.8), and therefore also measurable.

A.1.2 Proofs that the Model Satisfies Conditions 1-6

Throughout our proofs we use assumptions about the first derivatives of F:

$F_1 \leq 0$**:** A higher child's ability does not increase the cost of producing h' units for tomorrow;
$F_2 \leq 0$**:** A higher externality does not increase the cost of producing h' units for tomorrow;
$F_3 > 0$**:** The marginal cost of producing h' is positive;

as well as maintained assumptions about the second derivatives, cross-derivatives, and combinations of derivatives:

[23]Note that $\mathbb{C}_{\chi,a_i}$ is closed and bounded since $\mathbb{C}_{\chi,a_i} = [\underline{h}_\chi - F(a_i, \chi, h^{max}(\underline{h}_\chi, a_i)),\ \overline{h}_\chi - F(a_i, \chi, \underline{h}_\chi)] = [\underline{c}_{\chi,a_i}, \overline{c}_{\chi,a_i}]$. The return function $F : A \to \mathbb{R}$ in SLP is in our model $u : \mathbb{C}_{\chi,a_i} \to \mathbb{R}$. We define the function from SLP $F(\cdot, y, z) : A_{yz} \to \mathbb{R}$ as $u[c(\cdot, h'; a_i)] : \mathbb{H}_\chi \to \mathbb{R}$.

$F_{11}, F_{22} \geq 0$**:** There is a (weakly) diminishing marginal return from ability and the externality;

$F_{33} > 0$**:** There is a (strictly) increasing marginal cost of producing h', ensures $\mathbb{H}_\chi$ is bounded;

$F_{12} \geq 0$**:** The cross-effect of marginal benefits is (weakly) diminishing;

F_{13}**,**$F_{23} \leq 0$**:** The marginal cost is (weakly) falling in current child's ability and the externality.

$F_3 > -F_2$**:** Ensures $\mathbb{H}$ is globally bounded.

$F_{22} + F_{33} > -2\,(F_{23})$**:** Ensures $\mathbb{H}$ is globally bounded.

We also make use of the fact that we can define $\Gamma_{\chi,a_i} : \mathbb{H}_\chi \rightrightarrows \mathbb{H}_\chi$, the correspondence describing the feasibility constraints for a fixed a_i, by:

$$\Gamma_{\chi,a_i}(h) = \left\{ \ h' \in \mathbb{H}_\chi \ \middle| \ h' \in [\underline{h}_\chi, h_\chi^{max}(h, a_i)] \ \right\}. \tag{9}$$

Establishing the representation in Equation 9 requires showing that for any χ there exists a maximum h' denoted by $h_\chi^{max}(h, a_i)$ in the choice set of agents with state vector (h, a_i), and that agents can choose any $h' \leq h_\chi^{max}(h, a_i)$ but cannot choose any $h' > h_\chi^{max}(h, a_i)$. This follows from the strict convexity of F in h' ($F_{33} > 0$), as there exists a unique $h_\chi^{max}(h, a_i)$ such that

$$\begin{aligned}
h &> F\Big(a_i, \ \chi, \ h_\chi^{max}(h, a_i)\Big) && \forall \ h' < h_\chi^{max}(h, a_i); \\
h &= F\Big(a_i, \ \chi, \ h_\chi^{max}(h, a_i)\Big) && \text{if } h' = h_\chi^{max}(h, a_i); \text{ and} \\
h &< F\Big(a_i, \ \chi, \ h_\chi^{max}(h, a_i)\Big) && \forall \ h' > h_\chi^{max}(h, a_i).
\end{aligned}$$

Thus the feasibility correspondence is defined as in Equation 9.

We now show that **Conditions 1, 4, and 6 hold in our model** by showing that the feasibility correspondence has the following properties:

Γ_{χ,a_i} **Is Non-empty:** Γ_{χ,a_i} is clearly non-empty as long as we choose $\underline{h}$ such that $0 < \underline{h} \leq h_\chi^{max}(\underline{h}, a_i)$. We can be assured that $0 < \underline{h} \leq h_\chi^{max}(\underline{h}, a_i)$ for all $a_i \in \mathbb{A}$ as long as $F\Big(\underline{a}, \ \chi \ , \underline{h}\Big) \leq 0$.

Γ_{χ,a_i} **Is Compact Valued:** Γ_{χ,a_i} is compact valued because we can see from Equation 9 that $\Gamma_{\chi,a_i}(h) \subset R_+$ is a closed and bounded, and therefore compact, set for each (h, a_i).

Γ_{χ,a_i} **Is lower hemi-continuous:** We follow the proof of Exercise 3.13 (b) in Irigoyen et al. (2003). Looking at the budget constraint in Equation 8, we can see that $h_\chi^{max}(h, a_i)$ is continuous in h by our assumptions on F. So let $y \in \Gamma_\chi(h, a_i)$. Given a sequence $\{h_m\}_{m=1}^{\infty}$ with $\lim_{m\to\infty} h_m = h$, let $\gamma = y/h_\chi^{max}(h, a_i)$ and define $y_m = \gamma h_\chi^{max}(h_m, a_i)$.

Then $y_m \in \Gamma_\chi(h_m, a_i)$ for all m, and since h_χ^{max} is continuous in h, we have that

$$\lim_{m\to\infty} y_m = \gamma \lim_{m\to\infty} h_\chi^{max}(h_m, a_i) = \gamma h_\chi^{max}(h, a_i) = y.$$

Γ_{χ,a_i} **Is upper hemi-continuous:** Consider a sequence $\{(h_m, a_i)\}_{m=1}^\infty$ such that $\lim_{m\to\infty}(h_m, a_i) = (h, a_i)$. If there is a sequence $\{h'_m\}_{m=1}^\infty$ such that $\lim_{m\to\infty} h'_m = h'$ and $h'_m \in \Gamma_\chi(h_m, a_i)$ for all m, then $h'_m \in [\underline{h}, h_\chi^{max}(h_m, a_i)]$ for all m. By the continuity of h_χ^{max}, this implies that

$$\begin{aligned}
&\lim_{m\to\infty} h'_m = h' \geq \underline{h} = \lim_{m\to\infty} \underline{h} \quad \text{and} \\
&\lim_{m\to\infty} h'_m = h' \leq h_\chi^{max}(h, a_i) = \lim_{m\to\infty} h_\chi^{max}(h_m, a_i),
\end{aligned}$$

so

$$h' \in [\underline{h}, h_\chi^{max}(h, a_i)],$$

proving that $h' \in \Gamma_{\chi,a_i}(h)$.

grΓ_{χ,a_i} **Is Convex:** Consider any two points $h_1, h_2 \in \mathbb{H}_\chi$, and suppose that $h'_1 \in \Gamma_{\chi,a_i}(h_1)$ and $h'_2 \in \Gamma_{\chi,a_i}(h_2)$. As long as $F(a, \chi, h')$ is convex in h' as assumed and the budget constraint is defined as in Equation 8, we know that $h_\chi^{max}(h, a_i)$ is concave. That is, as long as $F(a_i, \chi, h')$ is convex in h',

$$\theta h'_1 + (1-\theta)h'_2 \quad \in \quad \left[\ \underline{h}_\chi \ , \ \ h_\chi^{max}\Big(\theta h_1 + (1-\theta)h_2, \ \ a_i \ \Big)\right],$$

so the set grΓ_{χ,a_i} is convex.

Γ_{χ,a_i} **Is Increasing:** Because $h_1 \leq h_2$ implies $h_1 \leq h_2$ and F is increasing in h', we know that $h_1 \leq h_2$ implies $h_\chi^{max}(h_1, a_i) \leq h_\chi^{max}(h_2, a_i)$. Thus $\Gamma_\chi(h_1, a_i) \subseteq \Gamma_\chi(h_2, a_i)$.

Condition 2 holds in our model: Given χ and a_i, a given value of (h, h') determines consumption when the budget constraint holds with equality, $c : \mathbb{H}_\chi \times \mathbb{H}_\chi \to \mathbb{C}_{\chi,a_i}$ defined by $c(h, h') = h - F(a_i, \chi, h')$. Since c is continuous and $u : \mathbb{C}_{\chi,a_i} \to \mathbb{R}$ is continuous, $u[c(\cdot,\cdot)] : \text{gr}\Gamma_{\chi,a_i} \to \mathbb{R}$ is also continuous. $u[c(\cdot,\cdot)]$ is bounded since grΓ_{χ,a_i} is compact valued. **Condition 3 holds in our model** since $u[c(\cdot, h')] : \mathbb{H}_\chi \to \mathbb{R}$ is strictly increasing.

And **Condition 5 holds in our model** since given any $(h_1, h_1'), (h_2, h_2') \in \mathrm{gr}\Gamma_{\chi,a_i}$,

$$
\begin{aligned}
u\Big\{c\Big[\theta(h_1, h_1') + (1-\theta)(h_2, h_2')\Big]\Big\} & \\
&= u\Big\{\Big[\theta h_1 + (1-\theta)h_2\Big] - F\Big(a_i, \chi, \theta h_1' + (1-\theta)h_2'\Big)\Big\} \\
&> u\Big\{\Big[\theta h_1 + (1-\theta)h_2\Big] - \Big[\theta F(a_i, \chi, h_1') + (1-\theta)F(a_i, \chi, h_2')\Big]\Big\} \qquad (10) \\
&= u\Big\{\theta c(h_1, h_1') + (1-\theta)c(h_2, h_2')\Big\} \\
&> \theta u\Big\{c(h_1, h_1')\Big\} + (1-\theta)u\Big\{c(h_2, h_2')\Big\} \qquad (11)
\end{aligned}
$$

where 10 follows from the strict convexity of F in h' and 11 follows from the strict concavity of u.

A.2 Proof: A Unique Stationary (h, a) Distribution Exists for Fixed χ

Next, we must show that a unique stationary equilibrium exists for a fixed value of χ. That is, we now prove that there exists a unique distribution $\mu_\chi^* : \mathcal{S}_\chi \to [0,1]$ perpetually reproducing itself when the optimal decision rule is followed by agents in the economy facing the given ability shock process and an imposed/fixed externality χ.

Given the measurable decision rule $g_{h,\chi} : \mathbb{H}_\chi \times \mathbb{A} \to \mathbb{H}_\chi$ and the Markov chain ability process alternatively written using the function $Q : \mathbb{A} \times \sigma\big(\{a_i\}\big) \to [0,1]$ defined by

$$
Q\left(a_i, \mathcal{A}\right) \equiv \sum_{a_j \in \mathcal{A}} \pi(a_j | a_i),
$$

we know from Theorem 9.13 in SLP that we can write a well-defined transition function $P_\chi : S \times \mathcal{S}_\chi \to [0,1]$ for the Markov process induced by the household problem as:

$$
P_\chi\left(x, B\right) = P_\chi\left((h, a_i), \mathcal{H}_\chi \times \mathcal{A}\right) = \mathbf{1}\left\{g_{h,\chi}\left(h, a_i\right) \in \mathcal{H}_\chi\right\} Q\left(a_i, \mathcal{A}\right), \qquad (12)
$$

where $\mathbf{1}\{\cdot\}$ is the indicator function. The Markov process described by P_χ induces a mapping

$$
\Psi_\chi : \mathcal{P}(\mathcal{S}_\chi) \to \mathcal{P}(\mathcal{S}_\chi)
$$

from the set of probability measures $\mathcal{P}(\mathcal{S}_\chi)$ into itself that updates probability measures as

follows:

$$\mu_{i+1,\chi}(B) = \Psi_\chi(\mu_{i,\chi}(B)) = \int P_\chi\left(x, B\right) \mu_{i,\chi}(dx)$$

Because the following three conditions hold in our model,

Condition I: P_χ is increasing;

Condition II: $\mathbb{H}_\chi \times \mathbb{A}$ has a lower bound $(\underline{h}_\chi, \underline{a})$ and an upper bound $(\overline{h}_\chi, \overline{a})$; and

Condition III: There exists $(h^*_\chi, a^*) \in \mathbb{H}_\chi \times \mathbb{A}$ and a natural number m such that $P^m_\chi\big((\underline{h}_\chi, \underline{a}),\ [\underline{h}_\chi, h^*_\chi] \times \{\underline{a}, \ldots, a^*\}\big) > 0$ and $P^m_\chi\big((\overline{h}_\chi, \overline{a}),\ [h^*_\chi, \overline{h}_\chi] \times \{a^*, \ldots, \overline{a}\}\big) > 0$;

we can appeal to Theorem 2 in Hopenhayn and Prescott (1992) as a proof that for each χ there exists a unique stationary distribution μ^*_χ such that

$$\mu^*_\chi(B) = \Psi_\chi(\mu^*_\chi(B)) = \int P_\chi\left(x, B\right) \mu^*_\chi(dx).$$

A.2.1 Proof of Condition I: P_χ Is Increasing

We begin by proving that $\boldsymbol{g_{h,\chi}}$ **is strictly increasing in** h**:** Given χ and a_i, let h_1 and $h_2 \in \mathbb{H}_\chi \subset \mathbb{R}_+$ with $h_1 > h_2$. Supposing by way of contradiction that $g_{h,\chi}\left(h, a_i\right)$ is not strictly increasing in h, this would imply:

$$g_{h,\chi}\left(h_1, a_i\right) \leq g_{h,\chi}\left(h_2, a_i\right). \tag{13}$$

However, because $h_1 > h_2$ and $F\Big(a_i, \chi, g_{h,\chi}\left(h_1, a_i\right)\Big) \leq F\Big(a_i, \chi, g_{h,\chi}\left(h_2, a_i\right)\Big)$ (by our assumptions that $F_3 > 0$), we know that

$$g_{c,\chi}(h_1, a_i) = h_1 - F\Big(a_i, \chi, g_{h,\chi}\left(h_1, a_i\right)\Big) > h_2 - F\Big(a_i, \chi, g_{h,\chi}\left(h_2, a_i\right)\Big) = g_{c,\chi}(h_2, a_i).$$

This implies $u_c\big[c(h_1)\big] < u_c\big[c(h_2)\big]$, which again by our assumptions on F ($F_{33} > 0$) implies

$$u_c\big[g_{c,\chi}(h_1, a_i)\big] F_3\Big(a_i, \chi, g_{h,\chi}\left(h_1, a_i\right)\Big) < u_c\big[g_{c,\chi}(h_2, a_i)\big] F_3\Big(a_i, \chi, g_{h,\chi}\left(h_2, a_i\right)\Big).$$

By the Euler Equation this tells us that

$$V_h\big(g_{h,\chi}\left(h_1, a_i\right), a'\big) < V_h\big(g_{h,\chi}\left(h_2, a_i\right), a'\big),$$

implying that $g_{h,\chi}\left(h_1, a_i\right) > g_{h,\chi}\left(h_2, a_i\right)$ by the strict concavity of V, contradicting Inequality 13.

We next show that $\boldsymbol{g_{h,\chi}}$ **is strictly concave in** h**:** The optimal investment function

$g_{i,\chi} : \mathbb{H}_\chi \times \mathbb{A} \to \mathbb{R}$ and human capital decision rule $g_{h,\chi} : \mathbb{H}_\chi \times \mathbb{A} \to \mathbb{H}_\chi$ imply an optimal investment cost function that can be defined as

$$i^*_\chi(h, a) = F(a, \chi, g_{h,\chi}(h, a))$$

Fixing both χ and ability a_j, we can write the production cost function as a function of this period's human capital alone, h. We denote this function as

$$F_{\chi;a_j} : \mathbb{H}_\chi \to \mathbb{R},$$

which further allows us to write $i^*_{\chi;a_j} : \mathbb{H}_\chi \to \mathbb{R}$ defined by the rule

$$i^*_{\chi;a_j}(h) = F_{\chi;a_j}(a_j, \chi, g_{h,\chi}(h, a_j)).$$

Thus one can define $g_{h,\chi;a_j} : \mathbb{H}_\chi \to \mathbb{H}_\chi$

$$g_{h,\chi;a_j}(h) = F^{-1}_{\chi;a_j}(i^*_{\chi;a_j}(h))$$

Since F is strictly convex in h', we know that $F^{-1}_{\chi;a_j}$ is strictly concave. To ensure that $g_{h,\chi;a_j}$ is strictly concave in h, therefore, we need only prove that $i^*_{\chi;a_j}$ is strictly increasing.

To see that $i^*_{\chi;a_j}$ is strictly increasing, consider that the Euler Equation can be written as

$$F_3(a_i, \chi, h')u'(c) = \beta E[V_1(h', a')].$$

Recall that the binding budget constraint implies $c + i^*_{\chi;a_j}(h) = h$. Supposing by way of contradiction that $i^*_{\chi;a_j}(h)$ were weakly decreasing in h, this would imply that c were strictly increasing in h. Then $u'(c)$ would decrease in h, requiring that $E[V_1(h', a')]$ would decrease, or that h' were increasing in h. But this provides the contradiction to the supposition that $i^*_{\chi;a_j}(h)$ was weakly decreasing in h, so $i^*_{\chi;a_j}$ must be strictly increasing.

Finally, we show that $\boldsymbol{g_{h,\chi}}$ **is strictly increasing in** a**:** Suppose not. Then there exist a_1, a_2, and h_0 such that $a_1 > a_2$ and $g_{h,\chi}(h_0, a_1) \leq g_{h,\chi}(h_0, a_2)$. Since F_3 is strictly increasing in h' and strictly decreasing in a, this implies that $F(a_1, \chi, g_{h,\chi}(h_0, a_1)) < F(a_2, \chi, g_{h,\chi}(h_0, a_2))$. Since wage income is the same in either case, by the budget constraint it must be that $g_c(h_0, a_1) > g_c(h_0, a_2)$. By the properties of u, $u_c(g_c(h_0, a_1)) < u_c(g_c(h_0, a_2))$. From the Euler equation,

$$EV_h\big(g_{h,\chi}(h_0, a_1), a'\big)F_3(a_2, \chi, g_{h,\chi}(h_0, a_2)) > EV_h\big(g_{h,\chi}(h_2, a_i), a'\big)F_3(a_1, \chi, g_{h,\chi}(h_0, a_1))$$

By the properties of F, $F_3\left(a_2, \chi, g_{h,\chi}\left(h_0, a_2\right)\right) < F_3\left(a_1, \chi, g_{h,\chi}\left(h_0, a_1\right)\right)$ so

$$EV_h\big(g_{h,\chi}\left(h_0, a_1\right), a'\big) < EV_h\big(g_{h,\chi}\left(h_0, a_2\right), a'\big)$$

If the ability process is i.i.d so that $E\left(a'|a_1\right) = E\left(a'|a_2\right)$, then this implies $g_{h,\chi}\left(h_0, a_1\right) > g_{h,\chi}\left(h_0, a_2\right)$ which is a contradiction. If instead, $E\left(a'|a_1\right) > E\left(a'|a_2\right)$, then by the Envelope condition

$$V_h\big(h', a'\big) = u_{c'} w'$$

where $w' = w$ in a steady state. Differentiating $V_h\big(h', a'\big)$ with respect to a' yields

$$\frac{u^2}{c'^2}\left(-F_1 w'\right) > 0$$

The sign comes from the properties of u and F, specifically that the second derivative of u is negative and $F_1 \leq 0$. Thus, we return to the same contradiction that $g_{h,\chi}\left(h_0, a_1\right) > g_{h,\chi}\left(h_0, a_2\right)$.

A.2.2 Proof of Condition II: $\mathbb{H}_\chi$ Is Closed and Bounded

For a given χ, take (h, a_i) as given, and suppose that $\{h_n\}_{n=1}^{\infty} \subseteq \mathbb{H}_\chi$ with $h_n \to h_0$. Then $c + F(a_i, \chi, h_n) \leq h$ for all $n \in \mathbb{N}$, which by the continuity of F implies that $c + F(a_i, \chi, h_0) \leq h$, so that $h_0 \in \mathbb{H}_\chi$. Thus $\mathbb{H}_\chi$ is closed for each χ.

Recall that the cost of endowing h' units of human capital in a child is a function $F : \mathbb{A} \times \mathbb{H}_\chi \times \mathbb{H}_\chi \to \mathbb{R}$ defined by $F\left(a, \chi, h'\right)$, where $F \geq 0$ and equality holds at $F\left(\underline{a}, \chi, \underline{h}_\chi\right)$. Central to our proof that $\mathbb{H}_\chi$ is bounded is the **lowest maintenance cost function** $f : \mathbb{H}_\chi \times \mathbb{R} \to \mathbb{R}$, whose rule tells us the cost of maintaining human capital level h given the best possible ability shock when facing the (non-equilibrium/externally set) externality χ:

$$f(h, \chi) = F(\overline{a}, \chi, h).$$

As long as f is strictly convex in h, then given any χ there is some $\overline{h}_\chi < \infty$ such that

$$\begin{aligned} F(\overline{a}, \chi, h) &< h \quad \text{if} \quad h < \overline{h}_\chi; \\ F(\overline{a}, \chi, h) &= h \quad \text{if} \quad h = \overline{h}_\chi; \\ F(\overline{a}, \chi, h) &> h \quad \text{if} \quad h > \overline{h}_\chi, \end{aligned}$$

so that $[\underline{h}_\chi, \overline{h}_\chi]$ determines the feasible set of h'.

A strictly convex f in h requires that f has a strictly-positive third derivative, or that

$$\frac{\partial f}{\partial h} = F_3 f(h) > 0,$$

or that

$$F_3 > 0. \tag{14}$$

That is, as the level of h to be maintained increases, the increase in the marginal cost of maintaining h tomorrow must dominate the increase in the marginal benefits from today's human capital and today's externality.

For f to have a strictly-positive second derivative in h, it must be the case that

$$\begin{aligned} \frac{d^2 f}{dh^2} &= \left[\frac{\partial}{\partial h}(F_{33})\right] f(h) + (F_3)\frac{\partial f}{\partial h} \\ &= F_{33} f(h) + F_3^2 f(h) \\ &> 0. \end{aligned} \tag{15}$$

If F is strictly convex, then this condition is clearly satisfied.

A.2.3 Proof of Condition III: The Monotone Mixing Condition (MMC)

Recall that $[\underline{h}_\chi, \overline{h}_\chi] \equiv \mathbb{H}_\chi$ is the interval bounded by the minimum and maximum h attainable for a given χ. Following Huggett (1993), define the sequences $\{y_{\chi,n}\}_{n=1}^{\infty}$ and $\{z_{\chi,n}\}_{n=1}^{\infty}$

$$\begin{array}{llll} y_{\chi,1} = \underline{h}_\chi & y_{\chi,2} = g_{h,\chi}(y_{\chi,1}, \overline{a}) & y_{\chi,3} = g_{h,\chi}(y_{\chi,2}, \overline{a}) & \cdots \\ z_{\chi,1} = \overline{h}_\chi & z_{\chi,2} = g_{h,\chi}(z_{\chi,1}, \underline{a}) & z_{\chi,3} = g_{h,\chi}(z_{\chi,2}, \underline{a}) & \cdots \end{array}$$

The fact that $g_{h,\chi}$ is strictly concave in h and strictly increasing in a implies that $y_{\chi,n} \to \overline{h}_\chi$ and $z_{\chi,n} \to \underline{h}_\chi$, as Figure 20 helps to illustrate.

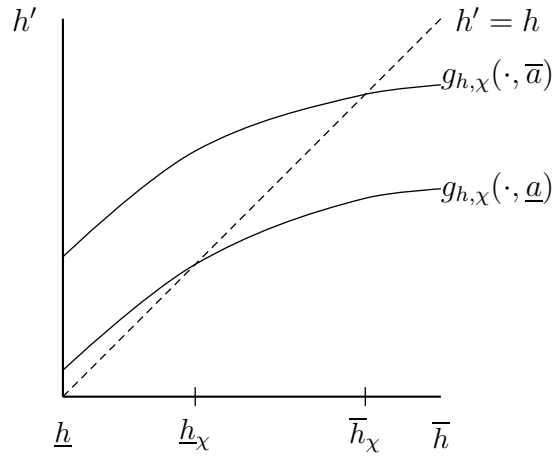


Figure 20: Decision Rules Illustrating Proof from Huggett (1993)

A.3 Proof: A Stationary General Equilibrium Exists

We now show that there exists a general equilibrium, or a human capital level χ^* whose associated steady state equilibrium implied by the model yields an externality of χ^*. Given the global feasible set $\mathbb{H} = [\underline{h}, \overline{h}]$, consider the aggregate supply of human capital that would be implied if χ were externally set to some value. We showed earlier that for any externally fixed χ there is a unique steady state partial equilibrium, which can be characterized by the associated distribution of (h, a), μ^*_χ. Thus we can define the aggregate supply of human capital implied by a given value of χ as:

$$H^S(\chi) = \left\{ H : H = \int g_{h,\chi}(h, a) \; d\mu^*_\chi(h, a) \right\}$$

The equilibrium conditions in the model define a self-map $J : \mathbb{H} \to \mathbb{H}$ by the equation

$$J_1(\chi) = \mathrm{H}(\chi) = \int g_{h,\chi}(h, a) \; d\mu^*_\chi(h, a). \tag{16}$$

The self-map J_1 tells us the externality/level of aggregate human capital implied by the model in which everything is determined in equilibrium except the experienced externality χ, which is set outside the model. A fixed point of J is therefore a general equilibrium of our model.

The Theorem of the Maximum (Thm 4.10.22 in Corbae et al. (2009)) implies $g_\chi(h, a)$ is continuous in χ. Because we also know that

a) $\mathbb{H} \times \mathbb{A}$ is compact;

b) Given a sequence $\{\big(\chi_n, (h_n, a_n)\big)\}$ that converges to $\big(\chi_0, (h_0, a_0)\big)$, $P_{\chi_n}\left((h_n, a_n), \mathcal{H} \times \mathcal{A}\right)$ converges weakly to $P_{\chi_0}\left((h_0, a_0), \mathcal{H} \times a_0\right)$ (due to Lebesgue's Dominated Convergence Theorem);

c) For each χ, there is a unique fixed point $\mu^*_\chi : \mathcal{S} \to [0, 1]$ of $\Psi_\chi : \mathcal{P}(\mathcal{S}) \to \mathcal{P}(\mathcal{S})$;

we can appeal to Theorem 12.13 of SLP as a proof that

$$\int g_{h,\chi}\left(h, a\right)\ d\mu^*_\chi(h, a)$$

is continuous in χ, so $J_1(\chi)$ is continuous as a result.

By the continuity of J_1 and the compactness and convexity of $\mathbb{H}$, we know by Brouwer's Fixed Point Theorem (Aliprantis and Border (2006), Corollary 17.56) that J_1 has a fixed point χ^*, or that our model has a general equilibrium.

To show that J_1 is indeed a self-map on the set $\mathbb{H} = [\underline{h}, \overline{h}] \subset \mathbb{R}_+$, we need to show that the set of feasible household human capital levels, $\mathbb{H}$, is globally closed and bounded. We start by generalizing the proof of $\mathbb{H}_\chi$ being bounded. Consider the specific **lowest maintenance cost function** $f : \mathbb{H} \to \mathbb{R}$ telling us the cost of maintaining human capital level h given the best possible ability shock when facing the externality $\chi = h$,

$$f(h) = F(\overline{a}, \chi = h, h),$$

or assuming the entire population also has the same level of human capital and the highest ability shock (ie, assuming $\mu_\chi\Big((h, \overline{a})\Big) = 1$). As long as $f(h)$ is strictly convex, then we know that there exists a global $\overline{h}$ for which

$$\begin{aligned} F(\overline{a}, \chi = h, h) &< h \quad \text{if} \quad h < \overline{h}; \\ F(\overline{a}, \chi = h, h) &= h \quad \text{if} \quad h = \overline{h}; \\ F(\overline{a}, \chi = h, h) &> h \quad \text{if} \quad h > \overline{h}. \end{aligned}$$

A strictly-positive first derivative of f requires that

$$\frac{df}{dh} = \left[F_2 + F_3\right] f\left(h\right) > 0.$$

This condition is satisfied if and only if

$$F_3 > -F_2. \tag{17}$$

That is, as the level of h to be maintained increases, the increase in the marginal cost must dominate the increase in the marginal benefits from the externality.

Showing f has a strictly-positive second derivative is more tedious. Assuming $\frac{d\chi}{dh} = 1$ and $\frac{d^2\chi}{dh^2} = 0$, we can express the second derivative of f as

$$\begin{aligned} \frac{d^2 f}{dh^2} &= \left[\frac{\partial}{\partial h}\left(F_2 + F_3\right)\right] f\left(h\right) + \left(F_2 + F_3\right)\frac{df}{dh} \\ &= \left[F_{22} + 2F_{23} + F_{33}\right] f\left(h\right) + \left(F_2 + F_3\right)\frac{df}{dh} \\ &> 0. \end{aligned} \tag{18}$$

Now if the term inside of the [] brackets in Equation 18 is positive, then this condition is satisfied. This term will be positive if

$$F_{22} + F_{33} > -2F_{23}. \tag{19}$$

The left hand side terms are disincentives to produce more human capital; the decreased marginal benefit of a higher externality and the increased marginal cost of h'. The right hand side is an incentive to produce more human capital; the decreased marginal cost of producing more h' induced by having higher externality.

A.4 Generalization of Proof to the SRCE Model in the Text (with Housing)

To generalize the proof to a model that includes housing, we first approach the problem with fixed price ρ of housing services $s \in \mathbb{S} \subset \mathbb{R}_+$, thus subscripting by (χ, ρ) rather than χ alone.[24] Some changes to the household problem are that now the feasible set is defined over a different range $\Gamma_{(\chi,\rho),a_i} : \mathbb{H}_{(\chi,\rho)} \rightrightarrows \mathbb{H}_{(\chi,\rho)} \times \mathbb{S}_{(\chi,\rho)}$, the utility function is defined over a different domain $u : \mathbb{C}_{(\chi,\rho),a_i} \times \mathbb{S}_{(\chi,\rho),a_i} \to \mathbb{R}$, there is a decision rule for housing services $g_{s,(\chi,\rho)}(h, a)$, and the transition function $P_{(\chi,\rho)} : S \times \mathcal{S} \to [0, 1]$ for the Markov process induced by the household problem also depends on ρ:

$$P_{(\chi,\rho)}\left((h, a), \mathcal{H} \times \mathcal{A}\right) = \mathbf{1}\left\{g_{h,(\chi,\rho)}\left(h, a_i\right) \in \mathcal{H}\right\} Q\left(a_i, \mathcal{A}\right). \tag{20}$$

As long as we specify an additive utility function $u(c, s) = \phi(c) + \psi(s)$ with properties analogous to those of $u(c)$ used in the earlier proofs, analogues to the proofs in Sections A.1

[24] To be clear here, recall that $s \in \mathbb{S}$ is housing services, $x \in S$ is the state vector, and $B \in \mathcal{S}$ is a Borel set in the associated product σ-algebra.

and A.2 all hold, up to and including the proof that a stationary distribution $\mu^*_{(\chi,\rho)}$ exists and is unique for fixed χ and ρ. The key insight is that there is an intra-temporal equilibrium condition in this model rendering s as a function of c, so that the arguments using the budget constraint $c + \rho s + F(a, \chi, h') \leq h$ all still hold.

To generalize the proof from Section A.3 to accommodate housing, note that $\mathbb{P}$ is bounded because $\mathbb{H}$ is bounded. To see this, first define the point mass distributions by

$$\underline{\mu}\left(\underline{h}, \underline{a}\right) = 1 \quad \text{and}$$
$$\overline{\mu}\left(\overline{h}, \overline{a}\right) = 1,$$

which allow us to define $\mathbb{P} = [\underline{\rho}, \overline{\rho}]$, where given the associated human capital externalities and prices $(\underline{\chi}, \underline{\rho})$, $(\overline{\chi}, \overline{\rho})$, we have:

$$\underline{\rho} = \phi\left(\int g_{s,(\underline{\chi},\underline{\rho})}\left(h, a\right)\ d\underline{\mu}(h, a)\right)$$
$$\overline{\rho} = \phi\left(\int g_{s,(\overline{\chi},\overline{\rho})}\left(h, a\right)\ d\overline{\mu}(h, a)\right)$$

Recall that we have assumed the absentee landlord supplies housing to meet demand. Thus the equilibrium condition in the housing market is that the price of housing services supports this allocation, which can be used to define $J_2 : \mathbb{H} \times \mathbb{P} \to \mathbb{P}$ by

$$J_2(\chi, \rho) = \rho(\chi, \rho) = \phi\left(\int g_{s,(\chi,\rho)}\left(h, a\right)\ d\mu^*_{(\chi,\rho)}(h, a)\right). \tag{21}$$

Equation 21 allows us to define the self-map $J : \mathbb{H} \times \mathbb{P} \to \mathbb{H} \times \mathbb{P}$ using J_2 and the analogue to J_1 in Equation 16. We can again apply Brouwer's Fixed Point Theorem to ensure existence of a fixed point (χ^*, ρ^*).

B Appendix: Discussion of the Existence of a MRCE

B.1 Conditions Required to Prove Existence of a MRCE

Here we sketch an outline of a proof of existence of a MRCE (ie, a general equilibrium of the model with two neighborhoods and mobility). The outline does not constitute a proof because there is one condition that we cannot show analytically. Nevertheless, we believe this condition does hold for a region of the parameter space, and the sketch should give the reader some intuition about the model with moving.

To begin, the state vector becomes (h, a, n), where $n \in \{N1, N2\}$ denotes neighborhood. Since we now must keep track of the externalities χ_{N1}, χ_{N2} and prices of housing services ρ_{N1}, ρ_{N2} in both neighborhoods, for the sake of exposition we refer to this vector as

$$\theta \equiv (\chi_{N1}, \chi_{N2}, \rho_{N1}, \rho_{N2}).$$

We again start by analyzing a model in which the externalities χ_{N1}, χ_{N2} and prices of housing services ρ_{N1}, ρ_{N2} are not equilibrium objects, but rather are externally set to some fixed values.

Again define $\mathcal{S}$ to be the product $\sigma-$algebra generated by $\mathcal{B}(\mathbb{H}_\theta)$, $\sigma\big(\{a_i\}\big)$, and $\sigma\big(\{n\}\big)$ containing subsets of the form $B = \mathcal{H} \times \mathcal{A} \times \mathcal{N}$. Now there is a residential decision rule $g_{n;\theta} : \mathbb{H}_\theta \times \mathbb{A} \times \{N1, N2\} \to \{N1, N2\}$ in addition to the human capital decision rule $g_{h;\theta} : \mathbb{H}_\theta \times \mathbb{A} \times \{N1, N2\} \to \mathbb{H}_\theta$. Note that because there are no moving costs and externalities/prices are set outside the model, the state space and domain of the decision rules could also be restricted to $\mathbb{H}_\theta$ and $\mathbb{A}$. Given these measurable decision rules and the Markov chain ability process alternatively written using the function $Q : \mathbb{A} \times \sigma\big(\{a_i\}\big) \to [0, 1]$ defined by

$$Q\left(a_i, \mathcal{A}\right) \equiv \sum_{a_j \in \mathcal{A}} \pi(a_j | a_i),$$

we know from Theorem 9.13 in SLP that we can write a well-defined transition function $P_\theta : S_\theta \times \mathcal{S}_\theta \to [0, 1]$ for the Markov process induced by the household problem as:

$$P_\theta\Big((h, a_i, n_-),\ \mathcal{H}\times\mathcal{A}\times\mathcal{N}\Big) \ = \ \mathbf{1}\Big\{g_{n;\theta}\left(h, a_i, n_-\right) \in \mathcal{N}\Big\}\ \mathbf{1}\Big\{g_{h;\theta}\left(h, a_i, g_{n;\theta}\left(h, a_i, n_-\right)\right) \in \mathcal{H}\Big\}\ Q\left(a_i, \mathcal{A}\right) \tag{22}$$

where $\mathbf{1}\{\cdot\}$ is the indicator function. The Markov process described by P_θ induces a mapping

$$\Psi_\theta : \mathcal{P}(\mathcal{S}_\theta) \to \mathcal{P}(\mathcal{S}_\theta)$$

from the set of probability measures $\mathcal{P}(\mathcal{S}_\theta)$ into itself that updates probability measures as follows:

$$\widehat{\mu}_{i+1,\theta}(B) = \Psi_\theta(\widehat{\mu}_{i,\theta}(B)) = \int P_\theta\big(x, B\big)\ \widehat{\mu}_{i,\theta}(dx)$$

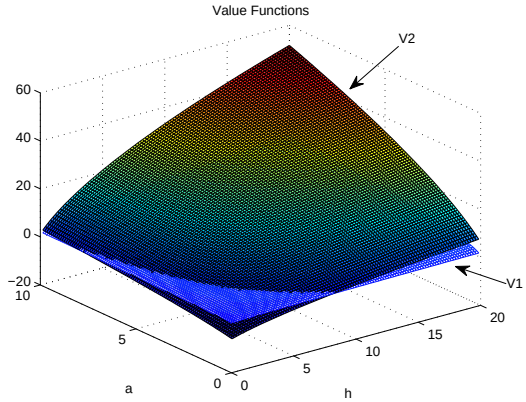
A proof would require the definition of a subset of the parameter space

$$\Theta^\star = \big\{(\chi_{N1}, \chi_{N2}, \rho_{N1}, \rho_{N2}) \in \mathbb{H}_\theta \times \mathbb{H}_\theta \times \mathbb{P} \times \mathbb{P} : \quad \chi_{N1} \leq \chi_{N2}, \quad \rho_{N1} \leq \rho_{N2}, \quad \text{and} \quad \phi(\chi_{N1}, \chi_{N2}, \rho_{N1}, \rho_{N2}) \leq 0\big\}$$

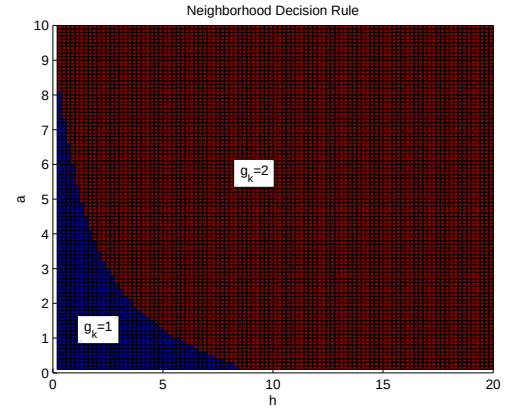
where the last condition would ensure that under a parameterization $\theta_i \in \Theta^\star$, the parameter θ_{i+1} (externalities and prices) implied by the model is itself in $\Theta^\star$.

It is difficult to analytically determine what restrictions ϕ would guarantee that $\Theta^\star$ is a compact, convex set for which the parameters implied by the model under a parameterization in $\Theta^\star$ is a self-map. We generally would need the following to hold: At low levels of h and a, the lower price of housing in N1 is worth more to a household than is the higher externality in N2. This would ensure that the implied externality in N1 would be less than the implied externality in N2. At the same time, this could not be true for too many households, or else enough people would want to live in N1 so that the implied (by the optimizing decision rules and current distributions) price of housing there would be higher than in N2.

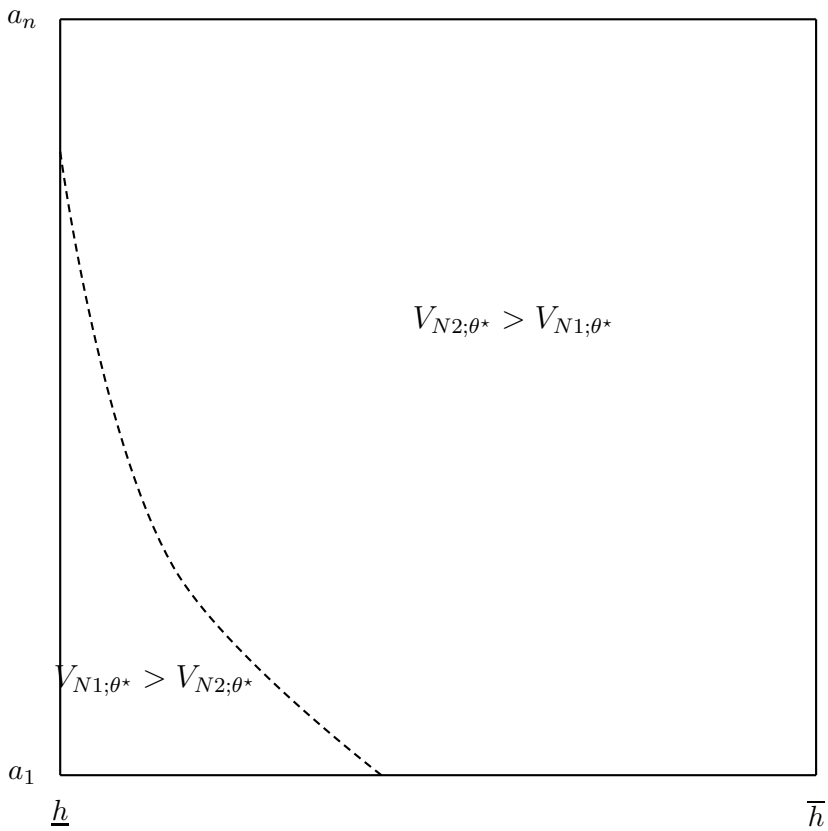
Figure 21 shows two example value functions satisfying these conditions. Here it is worth pointing out a computational difficulty, which is that the neighborhood-specific value functions $V_{N1}, V_{N2} : \mathbb{H}_\theta \times \mathbb{A} \to \mathbb{R}$ are not simply the value functions of each neighborhood in the model under segregation (ie, the value functions $V_{(\chi_{N1},\rho_{N1})}$ and $V_{(\chi_{N2},\rho_{N2})}$ of the two associated SRCEs with prices set outside the models). These value functions are the value of counterfactually residing in either neighborhood this period, with a continuation value that also accounts for future mobility between neighborhoods (ie, the value functions $V_{N1;(\chi_{N1},\chi_{N2},\rho_{N1},\rho_{N2})}$ and $V_{N2;(\chi_{N1},\chi_{N2},\rho_{N1},\rho_{N2})}$ with partial equilibrium prices) depending on where in $\mathbb{H}_\theta \times \mathbb{A}$ the household moves.



(a) Value Function Manifolds for $(\chi_{N1}, \chi_{N2}, \rho_{N1}, \rho_{N2})$



(b) Nbd Decision Rule for $(\chi_{N1}, \chi_{N2}, \rho_{N1}, \rho_{N2})$



(c) The Projection of $V_{N1;\theta^\star} \cap V_{N2;\theta^\star}$ onto $\mathbb{H} \times \mathbb{A}$ for a fixed $\theta^\star \in \Theta^\star$

Figure 21: Value Functions for Parameters $(\chi_{N1}, \chi_{N2}, \rho_{N1}, \rho_{N2})$ Likely to Be in $\Theta^\star$

Again, we have not analytically specified a rule ϕ defining $\Theta^\star$. However, if we could do so, we could show that the model with moving also satisfies the hypotheses of Theorem 2 in Hopenhayn and Prescott (1992), which would serve as a proof that there is a unique stationary distribution associated with each $\theta^\star$.

Given the unique invariant distribution $\widehat{\mu}^*_{\theta^\star}$, we would re-define

$$J : \mathbb{H}^\star \times \mathbb{H}^\star \times \mathbb{P}^\star \times \mathbb{P}^\star \to \mathbb{H}^\star \times \mathbb{H}^\star \times \mathbb{P}^\star \times \mathbb{P}^\star,$$

or

$$J : \Theta^\star \to \Theta^\star,$$

by:

$$\begin{aligned} J_1(\theta^\star) = \chi_{N1}(\theta^\star) &= \frac{\int\ h\ \mathbf{1}\big\{g_{n;\theta}\left(h,a,n_-\right) = N1\big\}\ d\widehat{\mu}^*_{\theta^\star}}{\int\ \mathbf{1}\big\{g_{n;\theta}\left(h,a,n_-\right) = N1\big\}\ d\widehat{\mu}^*_{\theta^\star}} \\ &= \frac{\int\ h\ d\mu^*_{\theta^\star}}{\int\ d\mu^*_{\theta^\star}} \end{aligned}$$

$$J_2(\theta^\star) = \chi_{N2}(\theta^\star) = \frac{\int\ h\ d\mu^*_{\theta^\star}}{\int\ d\mu^*_{\theta^\star}}$$

$$J_3(\theta^\star) = \rho_{N1}(\theta^\star) = f\bigg(\int g_{s;\theta^\star}\left(h,a,g_{n;\theta^\star}(h,a,n_-)\right)\ \mathbf{1}\Big\{g_{n;\theta}\left(h,a,n_-\right) = N1\Big\}\ d\widehat{\mu}^*_{\theta^\star}\bigg)$$

$$J_4(\theta^\star) = \rho_{N2}(\theta^\star) = f\bigg(\int g_{s;\theta^\star}\left(h,a,g_{n;\theta^\star}(h,a,n_-)\right)\ \mathbf{1}\Big\{g_{n;\theta}\left(h,a,n_-\right) = N2\Big\}\ d\widehat{\mu}^*_{\theta^\star}\bigg)$$

By the definition of $\Theta^\star$, we would know that

$$J : \Theta^\star \to \Theta^\star$$

is a self-map. Thus we would again appeal to Theorem 12.13 of SLP as proof that J is continuous, and then again apply Brouwer's Fixed Point Theorem to show existence of a general equilibrium.

C Computational Appendix

C.1 Calibration to SRCE Steady State

Outer loop:

I. Guess parameter vector x^k.

Inner loop: Finding the SRCE associated with the parameter vector x^k.

1. Guess $p_n^k = p_n^0$ and $H_n^k = H_n^0$ for $n = 1, 2$.

2. Over a coarse grid of 150 points in h, iterate on the Bellman equation to solve the household problem, using cubic splines to interpolate the value function between grid points.

3. Linearly interpolate over the decision rules found in Step 2 to convert them from the coarse grid to a fine grid of 5000 h points.

4. Find the invariant distributions in each neighborhood. Γ_1^0 and Γ_2^0 are initial distributions (stored as histograms on the fine human capital grid for each ability type). Starting with Γ_1^0 and Γ_2^0, apply the decision rules produced in Step 3 to generate new distributions Γ_1^1 andΓ_2^1. Continue iterating until $\left\|\Gamma_1^m - \Gamma_1^{m+1}\right\|^\infty$ and $\left\|\Gamma_2^m - \Gamma_2^{m+1}\right\|^\infty$ are below some tolerance. Save $\hat{\Gamma}^k = \Gamma^{m+1}$.

5. Calculate the implied housing demands, $\hat{S}_1^0$, $\hat{S}_2^0$, and human capital levels, $\hat{H}_1^0$, $\hat{H}_2^0$. For $n = 1, 2$, find the implied market clearing house price:

$$\hat{p}_n^0 = \frac{1}{\alpha}\left(\frac{\hat{S}_n^0}{\bar{L}_n^0}\right)^{\frac{1-\alpha}{\alpha}}.$$

6. Update price guesses: $p_n^1 = \zeta_p \hat{p}_n^0 + (1 - \zeta_p)\, p_n^0$, where $\zeta_p \in (0, 1)$. Repeat Steps 2-5 until $\left\|p_1^m - p_1^{m+1}\right\|^\infty$ and $\left\|p_2^m - p_2^{m+1}\right\|^\infty$ are below some tolerance. Save $\hat{p}^k = p^{m+1}$. Then go to 7.

7. Update per capita human capital guesses: $H_n^1 = \zeta_H \hat{H}_n^0 + (1 - \zeta_H)\, H_n^0$ where $\zeta_H \in (0, 1)$. Repeat Steps 2-6 until $\left\|H_1^m - H_1^{m+1}\right\|^\infty$ and $\left\|H_2^m - H_2^{m+1}\right\|^\infty$ are below some tolerance. Save $\hat{H}^k = H^{m+1}$. Then go to 8.

8. Calculate the sum of squared percentage errors between data statistics and those implied by the model under the parameter x^k using $\hat{\Gamma}^k$ and $\hat{p}^k$.

End of Inner Loop

II. Use an optimization routine to update x^{k+1} until the sum of squared percentage errors is minimized.

End of Outer Loop

C.2 Transition to MRCE Steady State from Initial SRCE Steady State

I. Find a new steady state using the method described above for the SRCE calibration. The important difference between the MRCE steady state and the SRCE steady state is the addition of a moving decision rule.

1. Assume that the steady state is reached in $T+1$ periods.

2. Guess a sequence house prices and externalities from period 0 to T, $\{(p_0^k, H_0^k), (p_1^k, H_1^k), \ldots, (p_T^k, H_T^k)\}$. Beginning in period T, solve the household problem backward to $t = 0$, storing the decision rules and value function in each period.

3. Starting with Γ_0 (the wealth distribution in the initial steady state), simulate forward until period T using the decision rules found in the previous step. Calculate the implied prices and per capita human capital levels during the simulation, $\{(\hat{p}_0^k, \hat{H}_0^k), (\hat{p}_1^k, \hat{H}_1^k), \ldots, (\hat{p}_T^k, \hat{H}_T^k)\}$.

4. Update the transition path guess as a linear combination of the initial guess and the implied value. For example, updating $p_t^{k+1} = \theta \hat{p}_t^k + (1-\theta) p_t^k$ for $\theta \in (0,1)$. We have found that it is better to iterate on prices first, holding the externality guesses constant. Once prices converge, update the externalities.

5. Repeat Steps 2-4 until the maximum difference between the transition path guess and the implied value in any period is less than some small tolerance.

C.3 Calculating a Transition when Households are Myopic

1. Begin at time period $t = 0$ with the initial human capital distributions $\Gamma_{1,0}$ and $\Gamma_{2,0}$ having average human capital levels $H_{1,0}$ and $H_{2,0}$.

2. Guess period $t = 1$ prices in each neighborhood, $p_{n,1}$.

3. Iterate on the Bellman equation under prices ($p_{n,1}$) and externalities ($\chi_{1,1} = H_{1,0}$ and $\chi_{2,1} = H_{2,0}$) until the time $t = 1$ value function converges.

4. Using the initial distributions in each neighborhood, sort households according to their moving decision rules found in Step 3.

5. Find $S_{1,1}$ and $S_{1,2}$, housing services demand in each neighborhood and compute the market clearing price. If the difference between the guessed and the implied prices is below a small tolerance, go to Step 6. Otherwise, update prices and repeat steps 2-5.

6. Find average human capital in each neighborhood after sorting using the decision rules found in step 3. These are the actual externality levels. Get h' for each household using investment decisions and actual externalities.

7. Update the $t+1$ distributions $\Gamma_{1,1}$ and $\Gamma_{2,1}$, which imply average human capital levels $H_{1,1}$ and $H_{2,1}$. Repeat steps 2-7 until period $T + 1$ has been reached.

References

Aaronson, D. and B. Mazumder (2008). Intergenerational economic mobility in the United States, 1940 to 2000. *Journal of Human Resources 43*(1), 139–172.

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics 109*(3), 659–684.

Aliprantis, C. D. and K. C. Border (2006). *Infinite Dimensional Analysis* (3rd ed.). Springer.

Aliprantis, D. (2016). Human capital in the inner city. *Empirical Economics*. doi: 10.1007/s00181-016-1160-y.

Aliprantis, D. (2017). Assessing the evidence on neighborhood effects from Moving to Opportunity. *Empirical Economics*. doi:10.1007/s00181-016-1186-1.

Aliprantis, D. and F. G.-C. Richter (2016). Evidence of neighborhood effects from MTO: LATEs of neighborhood quality. *Federal Reserve Bank of Cleveland WP 12-08R*.

Anderson, E. (1999). *Code of the Street: Decency, Violence, and the Moral Life of the Inner City*. W. W. Norton and Company.

Arthur, C. (1999). The 'Pigmentocracy': Racial categories. In C. Arthur and M. Dash (Eds.), *A Haiti Anthology: Libète*, Chapter 1.12. Markus Wiener. Translation and Discussion of Moreau de Saint-Méry's *Description de la partie française de l'Isle de Saint-Domingue, 1797*.

Badel, A. (2010). Understanding permanent black-white earnings inequality. *Federal Reserve Bank of St. Louis Working Paper 2010-047B*.

Baldwin, J. (1989). A television conversation: James Baldwin, Peregrine Worsthorne, Bryan Magee (Encounter). In F. L. Standley and L. H. Pratt (Eds.), *Conversations with James Baldwin*. University Press of Mississippi.

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

Bayer, P., H. Fang, and R. McMillan (2014). Separate when equal? Racial inequality and residential segregation. *Journal of Urban Economics 82*, 32 – 48.

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.

Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2016). A dynamic model of demand for houses and neighborhoods. *Econometrica 84*(3), 893–942.

Bayer, P., S. Ross, and G. Topa (2008). Place of work and place of residence: Informal hiring networks and labor market. *Journal of Political Economy 116*(6), 1150–1196.

Becker, G. S. and N. Tomes (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy 87*(6), pp. 1153–1189.

Becker, G. S. and N. Tomes (1986). Human capital and the rise and fall of families. *Journal of Labor Economics 4*(3), pp. S1–S39.

Behrman, J. and P. Taubman (1985). Intergenerational earnings mobility in the United States: Some estimates and a test of Becker's intergenerational endowments model. *The Review of Economics and Statistics 67*(1), 144–151.

Bénabou, R. (1993). Workings of a city: Location, education, and production. *The Quarterly Journal of Economics 108*(3), pp. 619–652.

Bénabou, R. (1996). Heterogeneity, stratification, and growth: Macroeconomic implications of community structure and school finance. *The American Economic Review 86*(3), pp. 584–609.

Bewley, T. F. (1986). Stationary monetary equilibrium with a continuum of independently fluctuating consumers. In W. Hildenbrand and A. Mas-Collel (Eds.), *Contributions to Mathematical Economics in Honor of Gerard Debreu*, pp. 79–102. North Holland.

Black, S. E., P. J. Devereux, P. Lundborg, and K. Majlesi (2015). Poor little rich kids? the determinants of the intergenerational transmission of wealth. *NBER Working Paper 21409*.

Blackmon, D. A. (2008). *Slavery by Another Name: The Re-enslavement of Black People in America from the Civil War to World War II.* Doubleday.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. Ter Weel (2008). The economics and psychology of personality traits. *Journal of Human Resources 43*(4), 972–1059.

Borjas, G. J. (1995). Ethnicity, neighborhoods, and human-capital externalities. *The American Economic Review 85*(3), pp. 365–390.

Calvó-Armengol, A. and M. O. Jackson (2009). Like father, like son: Social network externalities and parent-child correlation in behavior. *American Economic Journal: Microeconomics 1*(1), 124–150.

Chambers, M., C. Garriga, and D. E. Schlagenhauf (2009). Accounting for changes in the homeownership rate. *International Economic Review 50*(3), 677–726.

Collins, W. and R. Margo (2006). Historical perspectives on racial differences in schooling in the United States. *Handbook of the Economics of Education 1*, 107–154.

Conley, T. G. and G. Topa (2002). Socio-economic distance and spatial patterns in unemployment. *Journal of Applied Econometrics 17*(4), 303–327.

Corbae, D., M. B. Stinchcombe, and J. Zeman (2009). *An Introduction to Mathematical Analysis for Economic Theory and Econometrics*. Princeton University Press.

Cutler, D. M., E. Glaeser, and J. L. Vigdor (1999). The rise and decline of the American ghetto. *Journal of Political Economy 107*(3), 455–506.

Cutler, D. M. and E. L. Glaeser (1997). Are ghettos good or bad? *Quarterly Journal of Economics 112*(3), 827–872.

Durlauf, S. N. (1996). A theory of persistent income inequality. *Journal of Economic Growth 1*(1), 75–94.

Durlauf, S. N. (2004). Neighborhood Effects. In J. V. Henderson and J. E. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Volume 4. Elsevier.

Epple, D., R. Romano, and H. Sieg (2012). The intergenerational conflict over the provision of public education. *Journal of Public Economics 96*(3-4), 255–268.

Fernandez, R. and R. Rogerson (1998). Public education and income distribution: A dynamic quantitative evaluation of education-finance reform. *The American Economic Review 88*(4), pp. 813–833.

Fisher, J. D. M. and M. Gervais (2011). Why has home ownership fallen among the young? *International Economic Review 52*(3), 883–912.

Fuchs-Schuendeln, N. and T. A. Hassan (2015). Natural experiments in macroeconomics. *NBER Working Paper 21228*.

Galiani, S., A. Murphy, and J. Pantano (2015). Estimating neighborhood choice models: Lessons from a housing assistance experiment. *The American Economic Review 105*(11), 3385–3415.

Glomm, G. and B. Ravikumar (1992). Public versus private investment in human capital: Endogenous growth and income inequality. *Journal of Political Economy 100*(4), pp. 818–834.

Glomm, G., B. Ravikumar, and I. C. Schiopu (2011). The political economy of education funding. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of The Economics of Education*, Volume 4, Chapter 9, pp. 615–680. Elsevier.

Graham, B. S. (2016). Identifying and estimating neighborhood effects. *NBER Working Paper 22575*.

Grawe, N. D. (2010). Primary and secondary school quality and intergenerational earnings mobility. *Journal of Human Capital 4*(4), 331–364.

Hickman, C. B. (1997). The devil and the one drop rule: Racial categories, African Americans, and the U.S. Census. *Michigan Law Review 95*(5), 1161–1265.

Hirsch, A. R. (1998). *Making the Second Ghetto: Race and Housing in Chicago 1940-1960*. University Of Chicago Press.

Homel, M. W. (1984). *Down from Equality: Black Chicagoans and the Public Schools, 1920-1941*. University of Illinois Press.

Hopenhayn, H. A. and E. C. Prescott (1992). Stochastic monotonicity and stationary distributions for dynamic economies. *Econometrica 60*(6), 1387–1406.

HUD (1976). *Annual Housing Survey: 1973, Part C: Financial Characteristics of the Housing Inventory* (H-150-73C ed.). Washington, DC: US Department of Housing and Urban Development, US Department of Commerce, and US Bureau of the Census.

HUD (1993). *American Housing Survey: 1991, National Summary Report* (H-150-91 ed.). Washington, DC: US Department of Housing and Urban Development, US Department of Commerce, and US Bureau of the Census.

Huggett, M. (1993). The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control 17*(5), 953–969.

Huggett, M. (1996). Wealth distribution in life-cycle economies. *Journal of Monetary Economics 38*(3), 469–494.

Ichino, A., L. Karabarbounis, and E. Moretti (2010). The political economy of intergenerational income mobility. *Economic Inquiry 49*(1), 47–69.

Ioannides, Y. M. (2010). Neighborhood effects and housing. In J. Benhabib, A. Bisin, and M. O. Jackson (Eds.), *Handbook of Social Economics*, pp. 1281 – 1340. North-Holland.

Irigoyen, C., E. Rossi-Hansberg, and M. L. J. Wright (2003). *Solutions Manual for Recursive Methods in Economic Dynamics*. Harvard University Press.

Johnson, R. C. (2014). Long-run impacts of school desegregation and school quality on adult attainments. *NBER Working Paper 16664*.

Joseph, M. L., R. J. Chaskin, and H. S. Webber (2007). The theoretical basis for addressing poverty through mixed-income development. *Urban Affairs Review 42*(3), 369–409.

King, Jr., M. L. (1998). *The Autobiography of Martin Luther King, Jr.* New York: Warner Books.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental analysis of neighborhood effects. *Econometrica 75*(1), 83–119.

Kopecky, K. A. and R. M. Suen (2010). Finite state Markov-chain approximations to highly persistent processes. *Review of Economic Dynamics 13*(3), 701–714.

Kremer, M. (1997). How much does sorting increase inequality? *The Quarterly Journal of Economics 112*(1), 115–139.

Krusell, P. and A. Smith, Jr (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy 106*(5), 867–896.

Lucas, Jr., R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics 22*(1), 3–42.

Lundberg, S. and R. Startz (1998). On the persistence of racial inequality. *Journal of Labor Economics 16*(2), 292–323.

Margo, R. A. (1986). Educational achievement in segregated school systems: The effects of "separate-but-equal". *The American Economic Review 76*(4), 794–801.

Massey, D. and N. Denton (1993). *American Apartheid: Segregation and the Making of the Underclass*. Cambridge: Harvard University Press.

Mazumder, B. (2005). Fortunate sons: New estimates of intergenerational mobility in the United States using Social Security earnings data. *Review of Economics and Statistics 87*(2), 235–255.

Meyer, S. G. (2000). *As Long as They Don't Move Next Door: Segregation and Racial Conflict in American Neighborhoods*. Rowman & Littlefield.

Minnesota Population Center (2004). *National Historical Geographic Information System* (Pre-release Version 0.1 ed.). Minneapolis, MN: University of Minnesota. http://www.nhgis.org.

Neckerman, K. M. (2007). *Schools Betrayed: Roots of Failure in Inner-City Education*. University of Chicago.

O'Flaherty, B. (2015). *The Economics of Race in the United States*. Cambridge: Harvard University Press.

Orazem, P. F. (1987). Black-white differences in schooling investment and human capital production in segregated schools. *The American Economic Review 77*(4), pp. 714–723.

Pattillo, M. (2003). Extending the boundaries and definition of the ghetto. *Ethnic and Racial Studies 26*(6), 1046–1057.

Pinto, R. (2014). Selection bias in a controlled experiment: The case of Moving to Opportunity. *Mimeo., University of Chicago*.

Polikoff, A. (2006). *Waiting for Gautreaux*. Northwestern University Press.

Polsgrove, C. (2001). *Divided Minds: Intellectuals and the Civil Rights Movement*. W. W. Norton & Company.

President's National Advisory Panel on Insurance in Riot-Affected Areas (1968). *Meeting the Insurance Crisis of Our Cities*, Washington, DC. President's National Advisory Panel on Insurance in Riot-Affected Areas.

Ríos-Rull, J.-V. (1999). Computation of equilibria in heterogeneous-agent models. In R. Marimon and A. Scott (Eds.), *Computational Methods for the Study of Dynamic Economies*, pp. 238–264. Oxford University Press.

Roberts, S. and P. Baker (2010, April 2). Asked to declare his race, Obama checks 'black'. *The New York Times*.

Rubinowitz, L. and I. Perry (2002). Crimes without punishment: White neighbors' resistance to black entry. *Journal of Criminal Law and Criminology 92*(2).

Sampson, R. J. (2012). *Great American City: Chicago and the Enduring Neighborhood Effect*. The University of Chicago Press.

Sampson, R. J., J. D. Morenoff, and F. Earls (1999). Beyond social capital: Spatial dynamics of collective efficacy for children. *American Sociological Review 64*(5), 633–660.

Sharkey, P. (2008). The intergenerational transmission of context. *American Journal of Sociology 113*(4), 931–969.

Sharkey, P. (2014). Spatial segmentation and the black middle class. *American Journal of Sociology 119*(4), pp. 903–954.

Solon, G. (1999). Intergenerational mobility in the labor market (Chapter 29). In O. C. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, pp. 1761–1800. Elsevier.

Squires, G. D. (Ed.) (1997). *Insurance Redlining: Disinvestment, Reinvestment, and the Evolving Role of Financial Institutions*. The Urban Institute Press.

Stokey, N. L., R. E. Lucas, Jr., and E. Prescott (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press.

Todd, P. and K. I. Wolpin (2007). The production of cognitive achievement in children: Home, school and racial test score gaps. *Journal of Human Capital 1*(1), 91–136.

Uzawa, H. (1962). Production functions with constant elasticities of substitution. *The Review of Economic Studies 29*(4), 291–299.

Washington, H. A. (2006). *Medical Apartheid: The Dark History of Medical Experimentation on Black Americans from Colonial Times to the Present*. Harlem Moon Broadway.

Wilson, W. J. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. University of Chicago.

X, M. (1963, January 23). *The Race Problem*. Michigan State University, East Lansing, MI: African Students Association and NAACP Campus Chapter.

X, M. (1990). The Ballot or the Bullet. In G. Breitman (Ed.), *Malcolm X Speaks*, Chapter III. New York: Grove Weidenfeld.