The End of the American Dream? Inequality and Segregation in US Cities *

Alessandra Fogli † Veronica Guerrieri ‡ Mark Ponder § Marta Prato ¶

August 2023

Abstract

Since the 1980s the US has experienced not only a steady increase in income inequality, but also a contemporaneous rise in residential segregation by income. What is the relationship between inequality and residential segregation? How does it affect intergenerational mobility? We first document a positive correlation between inequality and segregation, both over time and across metro areas. We then develop a general equilibrium model where parents choose the neighborhood where they raise their children and invest in their children's education. In the model, segregation and inequality amplify each other because of a local spillover that affects the return to education. We calibrate the model to a representative US metro in 1980 and use the micro estimates of neighborhood exposure effects in Chetty and Hendren (2018b) to discipline the strength of the local spillover. We first use the calibrated version of the model to explore the economy's response to an unexpected skill premium shock. We find that segregation dynamics played a significant role in amplifying the increase in inequality and in dampening intergenerational mobility. We then use the model to explore the effects of policies designed to move poor people to better neighborhoods, like the Moving To Opportunity program. We show that scaling up MTO policies induces general equilibrium effects that limit their efficacy.

*We thank our discussants Elisa Giannone, Ed Glaeser, Richard Rogerson, and Kjetil Storesletten for useful suggestions. We are also grateful to Roland Benabou, Steven Durlauf, Mike Golosov, Luigi Guiso, Erik Hurst, Francesco Lippi, Guido Lorenzoni, Guido Menzio, Alexander Monge-Naranjo, Fabrizio Perri, and Sevi Rodriguez Mora for helpful comments. For outstanding research assistance, we thank Yu-Ting Chiang, Emily Moschini, Luis Simon, and, in particular, Dhananjay Ghei and Francisca Sara-Zaror.

†Federal Reserve Bank of Minneapolis. afogli00@gmail.com

‡University of Chicago, Booth School of Business. vguerrie@chicagobooth.edu

§NERA. ponder.mark@gmail.com

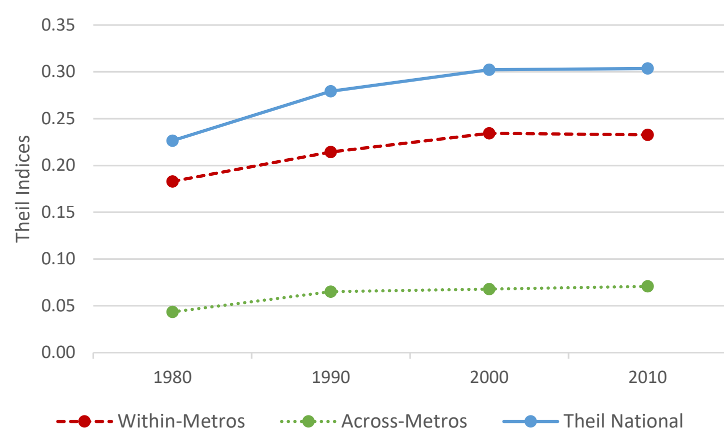
¶Bocconi University. marta.prato@unibocconi.it

1 Introduction

It is a well documented fact that the US has experienced a steady increase in income inequality over the last 40 years. At the same time, there has been a substantial increase in residential segregation by income. What is the link between inequality and residential segregation? In particular, has residential segregation amplified the response of income inequality to underlying shocks, such as skill-biased technical change? How do these patterns in inequality and segregation affect intergenerational mobility? In this paper, we build a model of educational investment and residential choice with local spillovers that can be used to address these questions.

Over the last few decades, inequality across neighborhoods within US metro areas has increased and has been an important driver of the overall income inequality in the US. Figure 1 reports the evolution of the Theil index of family income at the national level (blue solid line) and its decomposition into within-city (red dashed line) and across-cities (green dotted line) inequality.[1]

Figure 1: Inequality Within and Across Metros: Theil Index 1980-2000



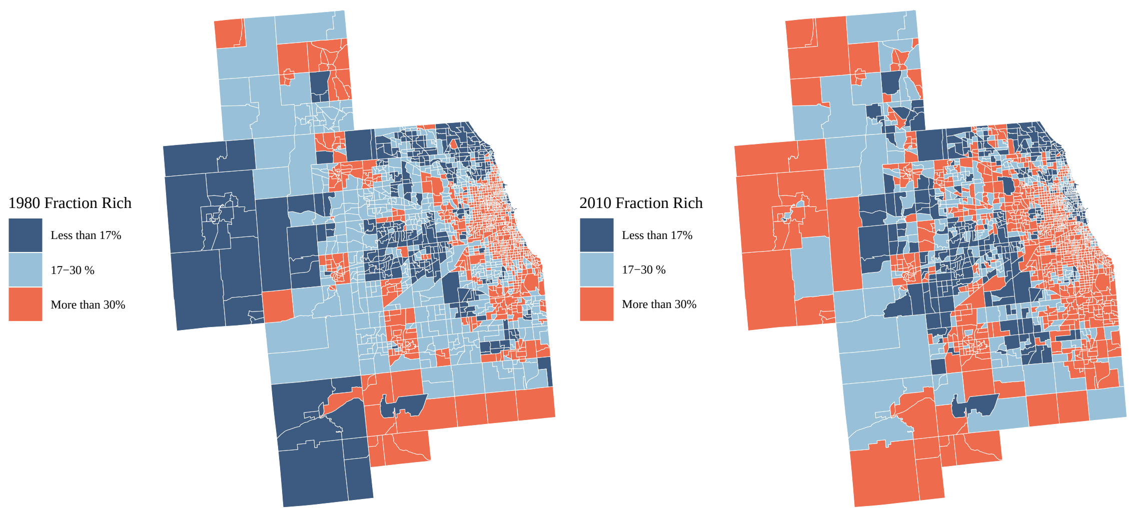
As the figure shows, the within-city component of overall inequality is significantly larger than the across-cities component, and it has increased more over time. A vast literature has focused on the increase in inequality across US cities, but has largely abstracted from the evolution of within-city inequality.[2] At the same time, US cities have experienced an increase in residential

[1]For this figure, we use census tract data on family income described in Section 2.

[2]See, for example, Moretti (2004), Shapiro (2006), Moretti (2012), Eeckhout et al. (2014), Hsieh and Moretti (2015), Diamond (2016), Giannone (2018), and Diamond and Gaubert (2022).

segregation by income. As an example, Figure 2 shows heat maps for the Chicago metro area in 1980 and 2010 to highlight the drastic increase in the number of neighborhoods with a high concentration of either rich (red areas) or poor families (dark blue areas).[3]

Figure 2: Share of Rich Families in Chicago



(a) 1980 (b) 2010

Our paper proposes a theory of within-city inequality that focuses on local spillovers as drivers of residential segregation by income, which, in turn, feeds back into inequality across neighborhoods.

In the 1990s, there was a large theoretical literature focusing on the relation between inequality and local externalities, starting from the seminal work by Benabou (1996a,b), Durlauf (1996a,b), and Fernandez and Rogerson (1996, 1997, 1998). However, only recently has the availability of administrative data allowed for direct estimates of neighborhood spillover effects. In particular, Chetty et al. (2016) and Chetty and Hendren (2018a,b) have shown that children's exposure to different neighborhoods has substantial effects on their future income. We bridge these two strands of literature, by proposing a general equilibrium model calibrated using the micro estimates from Chetty and Hendren (2018b) to understand the contribution of local externalities to

[3] The figure shows the concentration of rich and poor families in all census tracts of the Chicago metro area in 1980 and 2010. We define the families in the top 20th percentile of the metro income distribution as "rich"; all other families as "poor". The increase in concentration of rich and poor families is even more striking when looking at the number of census tracts instead of the areas. The number of census tracts with more than 30% rich families went from 360 to 421, and the number of census tracts with less than 17% rich families went from 894 to 1060.

segregation, inequality, and intergenerational mobility patterns.

In the first part of the paper, we document the positive correlation between income inequality and residential segregation by income at the metro level, both across time and across space. We show that 1) average inequality and residential segregation have increased steadily since 1980; 2) levels of inequality and residential segregation in 1980 are correlated across metros; and 3) changes in inequality and residential segregation between 1980 and 2010 are correlated across metros. We also show that the increase in inequality and segregation is stronger if we restrict the sample to families with children. This points neighborhood exposure effects' role as a key mechanism behind the dynamic relationship between residential segregation and income inequality.

We then build a general equilibrium overlapping generations model with educational and residential choices that features local externalities. The model generates a feedback effect between income inequality and residential segregation that amplifies the response of inequality to underlying shocks. We first use a simple version of the model to explain the mechanism. Agents live for two periods: first they are young and go to school, and then they are old and become parents. There are two neighborhoods, and parents choose both the neighborhood where they raise their children and their children's education level. The key ingredient of the model is a local spillover: investment in education yields higher returns in neighborhoods with higher expected future income of children - that is, neighborhoods with higher parents' income and children's ability. We model the spillover in a general way to capture a variety of mechanisms: differences in the quality of public schools, peer effects, social norms, learning from neighbors' experience, networks, and so forth.[4] We assume that the local spillover is complementary to the children's innate ability and to their level of education. The model generates sorting in equilibrium: richer parents and parents with more talented children choose to pay higher rents to live in the neighborhood with the higher local spillover. It follows that one neighborhood endogenously becomes the "good" one and hence the one where housing is more expensive. This means that in the model, residential choice is a form of investment in children's education, implying that talented children who grow up in poor families may be stuck in worse neighborhoods.[5]

[4]Among the most recent contributions, Agostinelli (2018) shows that peer effects account for more than half of the neighborhood effects in Chetty and Hendren (2018a), while Rothstein (2019) argues that job networks and the structure of local labor and marriage market play a more important role.

[5]We abstract from the fact that educational costs are also endogenous. Cai and Heathcote (2022) show that the rise in inequality explains a large fraction of the increase in college tuition, which would further amplify our mechanism.

We use this simple version of the model to qualitatively understand the dynamic relationship between inequality and segregation and to explore how the model responds to an unexpected permanent skill premium shock. When a skill premium shock hits the economy, inequality increases mechanically, because the wage gap between educated and non-educated workers increases. Moreover, given the complementarity between neighborhood spillover and education, when the skill premium is higher, more parents would like to live in the neighborhood with the larger spillover. However, given the inelastic housing supply, this translates into higher rental rates, and hence into a higher degree of segregation by income. The endogenous change in neighborhood composition, in turn, drives up the spillover differential between the two neighborhoods and translates into even higher future inequality. In particular, poor families with talented children may be pushed into worse neighborhoods, where the incentive to invest in education is lower. This further increases the gap between spillovers and worsens intergenerational mobility over time.[6]

In order to bring the model to the data, we generalize the model in a number of directions. First, we introduce an additional neighborhood to capture richer spatial dynamics. Second, we make the educational choice continuous, so as to not restrict the investment choice set. Third, we introduce two types of preference shocks: one that stands for local amenities and captures an additional force for residential segregation, the other that captures idiosyncratic determinants of the residential choice. We then calibrate the steady state of the model to the average US metro area in 1980. To discipline the calibration, we target a number of features of the US economy in 1980, and to discipline the strength of the local spillover, we use the micro estimates for neighborhood exposure effects obtained in the quasi-experiment of Chetty and Hendren (2018b).

Assuming that the original increase in inequality comes purely from skill-biased technical change, we study the effects of an unexpected, one-time shock to the skill premium on inequality, segregation, and intergenerational mobility over time. Despite the parsimony of the model, the exercise generates patterns for inequality and segregation that resemble the data. We also validate the model with a number of other statistics at the city and neighborhood level, such as house price dynamics in different neighborhoods, neighborhood size dynamics, and intergenerational mobility matrices across family income quartiles. We then use our model to ask our first main

[6]We abstract from redistributive policies. Alesina, Stantcheva and Teso (2018) show that residential segregation by income affects the perception of intergenerational mobility and, in turn, the political support for redistribution.

quantitative question: How much does segregation by income contribute to the rise in inequality? To answer this question, we run two counterfactual exercises where we look at the response of the economy to the same shock, but mute the sorting. In the first exercise, we assume that, after the shock, families are randomly re-located across neighborhoods, which implies that all three neighborhoods have the same distribution of income and ability and there is no residential segregation. In the second one, we assume that after the shock, families cannot re-optimize their residential choice. These exercises show that segregation by income contributes significantly to the total increase in inequality between 1980 and 2010: in the first case, it accounts for 27% of the increase, and in the second, it accounts for 25%. We also show that the increase in inequality in response to the skill premium shock is concurrent with a decrease in intergenerational mobility that is significantly amplified by segregation. We analyze a number of alternative specifications of the model to understand the relevance of some of our simplifying assumptions.

Our finding that residential segregation plays an important role in explaining inequality naturally raises the question whether housing voucher policies aimed at reducing segregation can be effective. In the mid-1990s, the US Department of Housing and Urban Development ran the Moving To Opportunity program (MTO), which offered vouchers to low-income families living in high-poverty neighborhoods to move to better neighborhoods. Chetty et al. (2016) show that the MTO program was quite successful in increasing the adulthood income of the children of families that received vouchers. The MTO experiment offered vouchers to a few hundred families. An important question is then whether a scaled up version of this policy would be effective in reducing overall US inequality and improving economic mobility. The hurdle is that scaling up the program also means generating general equilibrium effects that could potentially reduce the effectiveness of the policy. Our calibrated model is well suited to address this issue. In the paper, we show that once we take into account that housing prices and neighborhood spillovers are endogenous and evolve over time, as we scale up the policy, the income gain for the children of voucher receivers first increases but then starts declining. Moreover, while inequality and residential segregation decrease monotonically with the scale of the policy, its welfare gains also start declining when the scale of the policy is large enough.

Related Literature

Our model builds on a large class of models with multiple communities, local spillovers, and endogenous residential choice that study the effects of stratification ("residential segregation" in our language) on income distribution, going back to the fundamental work by Becker and Tomes (1979) and Loury (1981). Among the seminal papers in this literature, Benabou (1993) explores a steady state model where local complementarities in human capital investment, or peer effects, generate occupational segregation; the paper studies its efficiency properties.[7] Durlauf (1996b) proposes a related dynamic model with multiple communities, where segregation is driven by both locally financed public schools and local social spillovers. The paper shows that economic stratification and strong neighborhood feedback effects generate persistent inequality.[8] Benabou (1996a) embeds growth with complementary skills in production in a similar model, where local spillovers are due both to social externalities (like peer effects) and to locally financed public schools. Fernandez and Rogerson (1996) also study the impact of a number of reforms on public education financing using a related model, with no growth, where residential stratification is purely driven by locally financed public education.[9] Fernandez and Rogerson (1998) calibrate to US data a dynamic version of a similar model to analyze the static and dynamic effects of public school financing reforms. Benabou (1996b) also studies the effects of public-school financing reforms, but he allows for non-fiscal channels of local spillovers, like peers, role models, norms, networks, and so forth. He shows that disentangling financial and social local spillovers is important for evaluating different types of policies.

Relative to this literature, our paper makes several contributions. First, we bridge this theoretical literature and the more recent empirical literature that exploits the advent of large administrative datasets, by using the micro estimates of neighborhood exposure effects in Chetty and Hendren (2018b) to discipline the strength of local spillovers. Second, we use this class of models to analyze the response of an economy with local spillovers to a skill premium shock, one of the main drivers of the rise in income inequality in the mid-1980s.[10] Our model allows us to explore the effects of such a shock on segregation and on the dynamics of the local spillovers, which, in turn, further amplify future inequality and dampen intergenerational mobility. Third, given the

[7]de Bartolome (1990) also studies the efficiency properties of a similar type of model where communities' stratification is driven by peer effects in education. In related papers, the local social externalities take the form of role models (Streufert (2000)), or referrals by neighborhoods (see Montgomery (1991a,b)).

[8]Durlauf (1996a) uses a related model to study how residential stratification can generate permanent relative income inequality (as opposed to absolute low-income or poverty traps) in an economy where everybody's income is growing.

[9]In a similar framework, Fernandez and Rogerson (1997) study the effect of community zoning regulation on allocations and welfare.

[10]See, for example, Katz and Murphy (1992), Autor et al. (1998), Goldin and Katz (2001), Card and Lemieux (2001), Acemoglu (2002), and Autor et al. (2008).

quantitative nature of our analysis, we enrich the theoretical framework in several dimensions to improve its mapping to the data. In particular, the heterogeneity in children's innate ability and its complementarity with the spillover make parents' residential decision a function of the child's ability as well as of parental income, allowing us to obtain a continuous measure of income segregation that can be mapped to the data.[11] Finally, we use the model to run policy experiments. In particular, an important contribution of the paper is to analyze the effects of scaling up the Moving to Opportunity experiment in an environment in which housing prices are endogenous and neighborhood spillovers evolve in response to the policy.

The paper most related to ours is the contemporaneous work of Durlauf and Seshadri (2017). They also build on this class of models to explore the idea that larger income inequality is associated with lower intergenerational mobility, a relationship known as the "Gatsby curve." The model in the paper is close to our model in several dimensions, although the calibration strategy and the main exercise are different and complement each other well.[12] In another contemporaneous paper, Eckert and Kleineberg (2021) estimate a related model of residential and educational choice where local spillovers generate residential sorting, but use it to study the effects of school financing policies. Zheng and Graham (2022) calibrate a similar model also to study the effects of different public school allocation mechanisms. More recently, there has been a growing body of literature focusing on related models with neighborhood effects. Among others, Fogli, Guerrieri, Ponder and Prato (2022) and Chyn and Daruich (2022) analyze the dynamic effects of scaling up Moving to Opportunity and placed-based policies. Agostinelli et al. (n.d.) use a static spatial model to study the effects of vouchers and other policies on residential and school choice.

Another literature that is related to our work focuses on the role of parenting style decisions on children's outcomes.[13] In particular, Doepke and Zilibotti (2017) propose a model where there is a feedback effect between parenting style and socioeconomic conditions. In related work, Agostinelli, Doepke, Sorrenti and Zilibotti (2020) focus on how parenting affects the choice of peer groups. Their parenting style decision is close in spirit to our residential choice decision and is affected by the return to education. They use this model to explore the effects of policy

[11] Previous literature has focused mostly on the extreme cases of complete or zero segregation by income. Our assumption is related in spirit to Hassler and Mora (2000), which proposes a model where innate ability affects technological choices and generates a tight link between growth and intergenerational mobility.

[12] See Durlauf et al. (2022) for a survey on the Gatsby curve.

[13] See Doepke et al. (2019) for a survey.

interventions to move children to better neighborhoods. Using a related framework, Agostinelli, Doepke, Sorrenti and Zilibotti (2022) study the effect on inequality of switching to e-learning during the pandemic.

Our paper is also related to a broad literature on urban development economics.[14] In particular, Ferreira, Monge-Naranjo and Torres de Mello Pereira (2017) is a recent paper that uses a model close to ours to think about the emergence and persistence of urban slums and calibrates it to Brazilian data. A related strand of the literature focuses on spatial sorting generated by local amenities.[15] Among others, Guerrieri, Hartley and Hurst (2013) have focused on the endogenous nature of amenities. In contemporaneous work, Couture, Gaubert, Handbury and Hurst (2019) use a spatial model with endogenous amenities and non-homothetic preferences to study spatial re-sorting within urban areas after the '90s. Another related paper is Bilal and Rossi-Hansberg (2021), which emphasizes that the location choice of individuals is a form of asset investment.

The paper is organized as follows. In Section 2, we document the positive correlation between inequality and segregation across space and time. Section 3 describes the baseline model and shows how the model responds to a skill premium shock. In Section 4 we extend the model, describe our calibration strategy, and show the response of the economy to a skill premium change like the one observed in the data. Section 5 shows our main counterfactual exercises to quantify how much segregation has contributed to the increase in inequality. Section 6 examines the scaling up of MTO policies. Section 7 concludes.

2 Empirical Evidence

Over the last 40 years US cities have experienced a profound transformation in their socio-economic structure: poor and rich families have become increasingly spatially separated over time. As noted by Massey et al. (2009), this is a new phenomenon in US cities, which historically were segregated predominantly on the basis of race.[16] During the last third of the twentieth century, the United States moved toward a new regime of residential segregation characterized

[14] See Bryan, Glaeser and Tsivanidis (2020) for a survey.

[15] Early work by Brueckner, Thisse and Zenou (1999) and Glaeser, Kolko and Saiz (2001) emphasizes the role of urban amenities and spurred a vibrant literature on gentrification.

[16] Massey et al. (2009) documents that from 1900 to the 1970s, what changed over time was the level at which racial segregation occurred, with the locus of racial separation shifting from the macro level (states and counties) to the micro level (municipalities and neighborhoods).

by decreasing racial-ethnic segregation and rising income segregation. Such a shift took place at the same time as a steady increase in income inequality.

In this section, we document the magnitude of these phenomena and explore the correlation between segregation and inequality across time and space.

2.1 Segregation and Inequality over Time

To measure residential segregation by income, we use the dissimilarity index, which is the most common measure of evenness. In our main analysis, we define rich families as those with income above the 80th percentile, within a given metro and all other families as poor. The dissimilarity index for metro j is calculated as follows:

$$D(j) = \frac{1}{2}\sum_i \left| \frac{x_i(j)}{X(j)} - \frac{y_i(j)}{Y(j)} \right|,$$

where $X(j)$ and $Y(j)$ respectively denote the total number of poor and rich families in metro j, while $x_i(j)$ and $y_i(j)$ respectively denote the number of poor and rich families in census tract i in metro j.

We use tract-level family income data from decennial censuses (1980 to 2000) and from the American Community Survey (2008-2012). Our sample includes 380 metropolitan areas using the 2003 OMB definition.[17] We calculate the dissimilarity index for all metro areas in each decade and average at the national level using metro level population weights. Figure 3 plots the resulting measure of segregation (red dashed line) at the national level. The graph shows that the distribution of income has become progressively more uneven across census tracts over time. While in 1980, roughly 32% of families in the average US metro would have had to change residence to achieve an even distribution across census tracts, in 2010, the population that needed to change residence increased to roughly 38%. The increase was especially large between 1980 and 1990 and again between 2000 and 2010.[18]

[17] For summary statistics of our sample, see Appendix A.1.

[18] The increase in residential income segregation over time is a robust finding. Several sociologists have documented this fact using different measures of segregation. In particular, Jargowsky (1996) documents an increase in economic segregation for US metros between 1970 and 1990 using the Neighborhood Sorting Index, Watson (2009) finds an increase in residential segregation by income between 1970 and 2000 using the Centile Gap Index, and most recently, Reardon and Bischoff (2011) and Reardon et al. (2018) document this fact using the information theory index.

Figure 3: Inequality and Segregation over Time

Using the same data on family income at the tract level, we also compute the Gini coefficient at the metro level and similarly average at the national level using metro population weights. Income data at the census tract level are reported in bins and are top coded. Top-coded income data are a significant concern when calculating inequality measures. We follow a recent methodology proposed by von Hippel et al. (2017), who estimate the CDF of the income distribution non-parametrically and then use the empirical mean to fit the top-coded distribution.[19] We plot the resulting measure of the Gini coefficient in Figure 3 (blue solid line), together with the dissimilarity index. Both measures show a significant increase over time, with the Gini coefficient rising from roughly .38 to roughly .44 over the entire period. The figure shows that the increase in spatial segregation by income across neighborhoods happened at the same time as the increase in income inequality. Appendix A.2 documents the robustness of these patterns of increasing segregation and inequality across a host of different measures. [20]

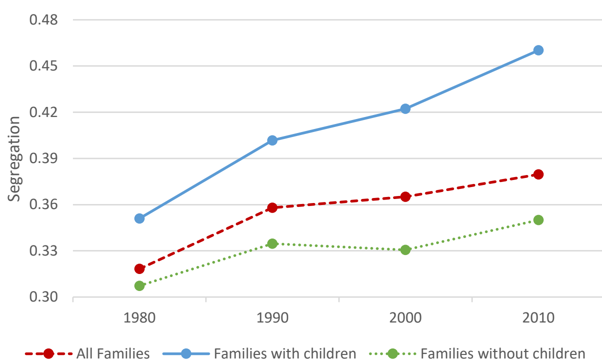
It is interesting to note that if we restrict the sample to families with children, patterns in both

[19] Some papers dealing with individual-level income data, such as Armour et al. (2016), have addressed the issue of top-coded data by estimating a Pareto distribution for the top income bracket. However, this methodology is not feasible when dealing with binned, rather than continuous, income data. The methodology most often used for binned data has been the one proposed by Nielsen and Alderson (1997), who use the Pareto coefficient from the last full income bracket to estimate the conditional mean of the top-coded bracket, as in Reardon and Bischoff (2011), for example. However, such a procedure does not exploit the fact that the census reports the precise empirical average income by census tract. Our method uses this information to improve the estimation of the top-coded distribution. For details, see Appendix A.1.

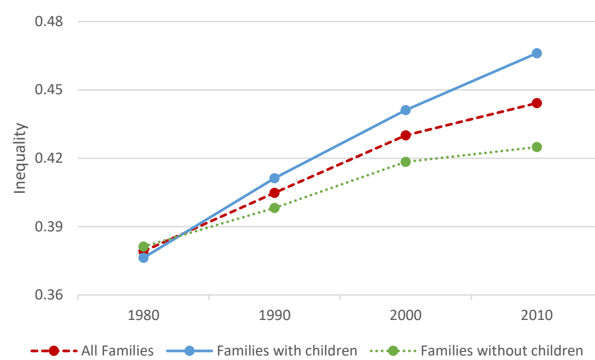
[20] See, for example, Katz and Murphy (1992), Autor et al. (1998), Goldin and Katz (2001), Card and Lemieux (2001), Acemoglu (2002), and Autor et al. (2008).

segregation and inequality are more pronounced. In particular, panel (a) in Figure 4 shows that not only is the level of segregation higher, but the increase over time is larger.[21] The figure shows that in 1980, the dissimilarity index for families with children is 0.35, compared with 0.31 for families without children. By 2010, the dissimilarity index for families with children increases to 0.46, while the dissimilarity index for the other families reaches only 0.35.

Figure 4: Segregation and Inequality: Different Samples



(a) Segregation



(b) Inequality

The increase in inequality is also stronger if we restrict the sample to families with children. Panel (b) in Figure 4 shows that, the level of inequality in 1980 is similar if we look at families with children and families without children; in both cases it is equal roughly to 0.38. However, inequality for families with children rises to 0.47 in 2010, while inequality for families without children reaches only 0.42 in the same year.

These findings point to the presence of children as an important driver of both residential segregation and income inequality. This is one of the reasons why we focus on the role of local spillovers on children's future income as a key mechanism behind the correlation between segregation and inequality.

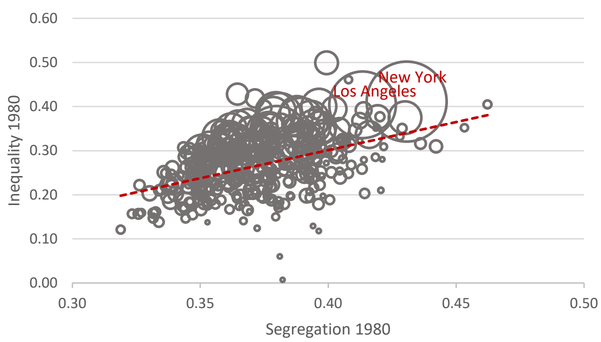
2.2 Segregation and Inequality across US Metros

Next, we document that residential segregation and inequality are also correlated across space. Panel (a) in Figure 5 shows the relationship between the Gini coefficient and the dissimilarity

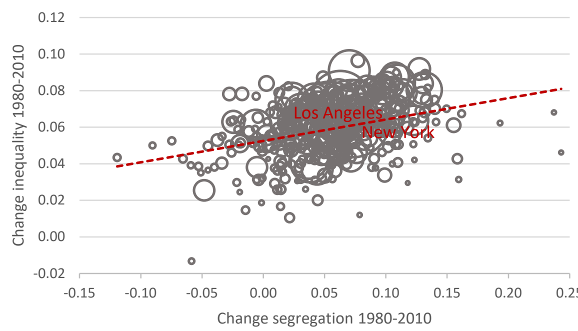
[21]See Appendix A.1 for details on how we construct the sample of families with kids not readily available for 1980 at the census tract level.

index across metro areas in 1980, where the bubbles are proportional to the population of the metro area. The figure shows a positive correlation between segregation and inequality in 1980. The significance of this relationship is robust to the inclusion of controls for demographic and industry composition. It also holds for the other decades in the sample and using the dissimilarity index constructed with other cut-offs to define rich and poor families. If we restrict the sample to families with children, the relationship becomes stronger.[22]

Figure 5: Inequality and Segregation across US Metros



(a) Levels



(b) Changes

The significance of the relationship between inequality and segregation is robust not just in levels but also in differences. Panel (b) in Figure 5 plots the change at the metro level in the Gini coefficient between 1980 and 2010 against the change at the metro level in the dissimilarity index over the same time period. Again, the size of the bubble is proportional to the population of the metro area. The figure shows that between 1980 and 2010, the metro areas that experienced a larger increase in inequality also experienced a larger increase in residential segregation. In Appendix A.3, we show that these results are robust to the inclusion of controls for changes in racial and industrial composition.[23]

Our analysis suggests a positive correlation between inequality and segregation, especially for

[22]The results of the regression of inequality on segregation across US metros in 1980 with and without controls are reported in Table 11, Appendix A.3. The regression coefficient using the full sample is 0.25, with a standard error of 0.015. When we restrict the sample to families with children, it becomes 0.33, with a standard error of 0.017.

[23]The results of the regression of changes in inequality on changes in segregation across US metros between 1980 and 2010 with and without controls are reported in Table 12, Appendix A.3. The regression coefficient for the baseline regression is 0.18, with a standard error of 0.017. If we restrict the sample to families with children, the coefficient becomes 0.24, with a standard error of 0.022.

families with children, both across time and across space. US cities have become increasingly segregated over time, reflecting an increased tendency of families to sort in different neighborhoods according to income.[24]

Prompted by this empirical evidence, we develop a model with local externalities and endogenous residential choice that is able to endogenously generate a feedback effect between inequality and residential segregation. We will use a calibrated version of the model to quantitatively assess the role of residential segregation in the increase in inequality and to explore the effects of scaling up the MTO experiment.

3 Simple Model

We first propose a simple version of a general equilibrium model where parents make residential and educational choices in the presence of endogenous local spillovers that affect children's return to education. We make a number of stark assumptions, as the purpose of the simple model is to explain the mechanism behind the feedback effect between inequality and residential segregation. In section 4, we generalize the model in a number of directions to make it more realistic and useful for the quantitative exercises.

3.1 Setup

The economy is populated by overlapping generations of agents who live for two periods. In the first period, the agent is a child, and in the second period, she is a parent. A parent at time t earns a wage $w_t \in [\underline{w}, \overline{w}]$ and has one child with ability $a_t \in [\underline{a}, \overline{a}]$. The ability of a child is correlated with the ability of the parent. In particular, $log(a_t)$ follows an AR1 process:

$$log(a_t) = x + \rho log(a_{t-1}) + \nu_t,$$

where ν_t is normally distributed with mean zero and variance σ_ν, $\rho \in [0,1]$ is the autocorrelation coefficient, and x is a constant normalized so that the mean of a_t is equal to 1. The joint distribu-

[24] In Appendix A.4, we also analyze the evolution of segregation using school districts instead of census tracts as the geographic subunit of analysis. We find a similar pattern of increase over time that is slightly mitigated in the last part of the period by the rise of private schools. We think that the census tract is a preferable unit of analysis in our context, since it better reflects our flexible notion of neighborhood spillover, is less affected by potential small sample bias, and can be more directly linked to metros.

tion of parents' wages and children's abilities evolves endogenously and is denoted by $F_t(w_t, a_t)$, with $F_0(w_0, a_0)$ taken as given.

There are two neighborhoods, denoted by $n \in \{A, B\}$. All houses are of the same dimension and quality, and the rent in neighborhood n at time t is denoted by R_{nt}. For simplicity, in the simple model we make the extreme assumption that the housing supply is fixed and equal to M in neighborhood A and fully elastic in neighborhood B. We normalize the marginal cost of construction in neighborhood B to 0, so that $R_{Bt} = 0$ for all t. The rental price in neighborhood A, R_{At}, is a key endogenous equilibrium object.[25]

In the simple model, we also assume that there are only two educational levels, $e \in \{e^L, e^H\}$. There is no cost to obtaining the low level of education, while $\tau > 0$ is the cost of investing in the high level of education. We can interpret $e = e^H$ as college education and $e = e^L$ as no college education.

Parents care both about their own consumption and about their children's future wage.[26] In particular, their preferences are given by $u(c_t) + g(w_{t+1})$, where u is a concave and continuously differentiable utility function, and g is increasing and continuously differentiable. A parent with wage w_t and with a child of ability a_t chooses 1) how much to consume, $c_t(w_t, a_t) \in R_+$; 2) where to live, $n_t(w_t, a_t) \in \{A, B\}$; and 3) how much to invest in the child's education, $e_t(w_t, a_t) \in \{e^L, e^H\}$. These choices affect the child's future wage, as explained below.

A key ingredient of the model is the presence of local spillovers that affect the children's return to education, and hence their future income. We denote the size of the local spillover in neighborhood n at time t by S_{nt}. Children's wages are affected by their ability, by their education, by the neighborhood where they grow up, because of the local spillover effect, and also directly by their parents' wage. Formally, the child of an agent (w_t, a_t) who grows up in neighborhood n and gets education level e is going to earn a wage

$$w_{t+1} = \Omega(w_t, a_t, e, S_{nt}, \varepsilon_t), \tag{1}$$

[25] The necessary assumption for our mechanism is that there is at least one neighborhood where housing supply is not fully elastic.

[26] This assumption is common in this class of models. The assumption that agents cannot save (if not by investing in housing or kids'education) is for simplicity. The assumption that agents cannot borrow is for realism, given that typically people cannot borrow against children's future income. An alternative specification could have parents getting utility directly from their children's consumption, with the introduction of a more general borrowing constraint.

where ε_t is an iid noise with cdf Ψ, and Ω is non-decreasing in all its arguments. We assume that Ψ is a normal distribution with mean one and standard deviation σ_ε. Children with higher ability and higher education, who grow up in neighborhoods with larger spillover and have richer parents will have higher future income. Since residential and educational choices are functions of parents' wage and children's ability (w_t, a_t), with a slight abuse of notation, we can write $w_{t+1} = w_{t+1}(w_t, a_t, \varepsilon_t)$. We will show that in equilibrium, for a given child's ability, parents with a higher wage are more likely to choose higher education and the neighborhood with higher spillover. This implies that children's wages will be increasing in parents' wages, both because of the direct effect in (1) and because of indirect effects operating through education and residential choices.

Let us now turn to the spillover. We assume that the size of the spillover effect in neighborhood n at time t is equal to the expected future income of the children growing up in that neighborhood:

$$S_{nt} = \frac{\int\int\int_{n_t(w_t,a_t)=n} w_{t+1}(w_t, a_t, \varepsilon_t) F_t(w_t, a_t) \Psi_t(\varepsilon_t) dw_t da_t d\varepsilon_t}{\int\int_{n_t(w_t,a_t)=n} F_t(w_t, a_t) dw_t da_t}. \tag{2}$$

Given that wages are increasing in ability and in parents' wage, neighborhoods with a higher spillover are neighborhoods with both richer parents and children with higher ability. The presence of this externality implies that the rental rate in neighborhood A, R_{At}, also depends on the relative size of the spillovers in the two neighborhoods, which is endogenous.

We chose this broad specification for S_{nt} because it can capture different sources of pecuniary and social externalities. On the one hand, the fact that neighborhoods with higher spillovers have richer parents allows us to interpret the spillover as linked to better public schools, which are typically locally financed and hence tend to improve with the average taxpayers' income. On the other hand, the fact that neighborhoods with higher spillovers are neighborhoods with more talented children allows us to interpret the spillover as peer effects. Both richer parents and more talented kids may also be the source of stronger networks on the labor market, social norms more conducive to educational investment, and so forth. An alternative specification be to have the spillover equal to the average wage of the parents or to the average level of education of the children in the neighborhood. However, the first would miss the role of innate ability, and the second would underplay the role of parental income. We use the more general specification in equation (2) because it maps better to the empirical estimates from Chetty and Hendren (2018b) that we use in our calibration, which capture the total effect of growing up in a given neighborhood on future income.

In our analysis, we make two assumptions. First, for simplicity, we assume that ability and spillover size affect children's future wages only if they attain the high level of education.

Assumption 1 *The function $\Omega(w,a,e,S,\varepsilon)$ is constant in S and a if $e = e^L$ and is increasing in S and a if $e = e^H$.*

The assumption that the wage of children with low education does not depend on ability stands for the fact that abilities that are relevant in high-skill jobs (which typically require college) may be different from and more heterogenous than abilities that are relevant for low-skill jobs. The assumption that the spillover's size does not affect the wage of children with low education is extreme, but can be interpreted as stating that the quality of K-12 schooling is more important in determining future wages of college graduates than those of non-college-graduates. This second assumption simplifies the analysis because all parents living in the rich neighborhood also pay for their children's college, given that there would be no other reason to pay a higher rent in the first place. We will relax Assumption 1 in the extended model in Section 4, which has a continuous educational choice.

Second, we make the following assumption.

Assumption 2 *The composite function $g(\Omega(w,a,e,S,\varepsilon))$ has increasing differences in a and S, in a and e, in w and S, and in w and e.*

These complementarities assumptions play a crucial role for our mechanism, as we will describe in the next subsections. Two key assumptions are the complementarity between innate ability and education and between innate ability and neighborhood spillover. Although it is hard to get direct estimates of innate ability, these assumptions reflect some of the findings of the recent empirical literature. [27]

[27] Our assumptions of complementarity between innate ability and education and between innate ability and neighborhood spillover are consistent with the latest research on technology of skill formation. Cunha et al. (2010) show that the higher the initial conditions for cognitive and non-cognitive skills of children, the higher the return to parental investment in children at later stages in life. As they highlight, "Family environments and genetic factors may influence these initial conditions." In our model, parental investment in children's future outcomes takes place both through educational investment and through residential choice. Moreover, the recent human capital literature, reviewed in Sacerdote (2011), also highlights the presence of non-linearity in peer effects, which are one of the forces behind our spillover effects. In particular, Sacerdote (2001), Imberman et al. (2012), and Lavy et al. (2012) find that high ability students are the ones who benefit the most from peer effects of other high ability students. Another paper that speaks more specifically to the complementarity between ability and spillover effects is Card and Giuliano

To sum up, a parent with wage w_t who has a child with ability a_t at time t solves the following problem:

$$\begin{aligned} U(w_t, a_t) = & \max_{c_t, e_t, n_t} u(c_t) + E[g(w_{t+1})] \qquad \text{(P1)} \\ & s.t. \;\; c_t + R_{n_t t} + \tau e_t \leq w_t \\ & \qquad w_{t+1} = \Omega(w_t, a_t, e_t, S_{n_t t}, \varepsilon_t), \end{aligned}$$

taking as given spillovers and rental rates in the two neighborhoods, $S_{n_t t}$ and $R_{n_t t}$ for $n_t = A, B$.

3.2 Equilibrium

We are now ready to define an equilibrium.

Definition 1 *For a given initial wage distribution $F_0(w_0, a_0)$, an equilibrium is characterized by a sequence of educational and residential choices, $\{e_t(w_t, a_t)\}_t$ and $\{n_t(w_t, a_t)\}_t$, a sequence of rents and spillover sizes in neighborhoods A and B, $\{R_{nt}\}_t$ and $\{S_{nt}\}_t$ for $n = A, B$, and a sequence of distributions $\{F_t(w_t, a_t)\}_t$ that satisfy the following conditions:*

1. *agents' optimization: for each t, the policy functions e_t and n_t solve problem (P1), for given R_{nt} and S_{nt} for $n = A, B$;*

2. *spillovers' consistency: for each t, equation (2) is satisfied for both $n = A, B$;*

3. *market clearing: for each t, $R_{Bt} = 0$, and R_{At} ensures housing market clearing in neighborhood A:*
$$M = \int \int_{n_t(w_t, a_t) = A} F_t(w_t, a_t) dw_t da_t; \qquad (3)$$

4. *wage dynamics: for each t,*
$$w_{t+1} = \Omega(w_t, a_t, e_t(w_t, a_t), S_{n_t(w_t, a_t) t}, \varepsilon_t). \qquad (4)$$

From now on, we focus on equilibria where the housing market in neighborhood A clears with positive rental rate - that is, $R_{At} > 0$ for all t - which requires also $S_{At} > S_{Bt}$ for all t.[28]

(2016), which shows that high achievers from minority and disadvantaged groups show high returns when included in school tracking programs.

[28] If $S_{At} \leq S_{Bt}$, nobody want to live in A, and the rental rate in A would be zero.

Assumptions 1 and 2 allow us to characterize the equilibrium in a fairly simple way, as shown in the following proposition.

Proposition 1 *Under assumptions 1 and 2, for each time t, there are two non-increasing cut-off functions $\hat{w}_t(a_t)$ and $\hat{\hat{w}}_t(a_t)$, with $\hat{w}_t(a_t) \leq \hat{\hat{w}}_t(a_t)$, such that*

$$e_t(w_t, a_t) = \begin{cases} e^L \text{ if } w_t < \hat{w}_t(a_t) \\ e^H \text{ if } w_t \geq \hat{w}_t(a_t) \end{cases}, \tag{5}$$

and

$$n_t(w_t, a_t) = \begin{cases} B \text{ if } w_t < \hat{\hat{w}}_t(a_t) \\ A \text{ if } w_t \geq \hat{\hat{w}}_t(a_t) \end{cases}. \tag{6}$$

This proposition shows that in equilibrium, the residential and the educational choices can be simply characterized by two monotonic, non-increasing cut-off functions.

Figure 6: Equilibrium Characterization

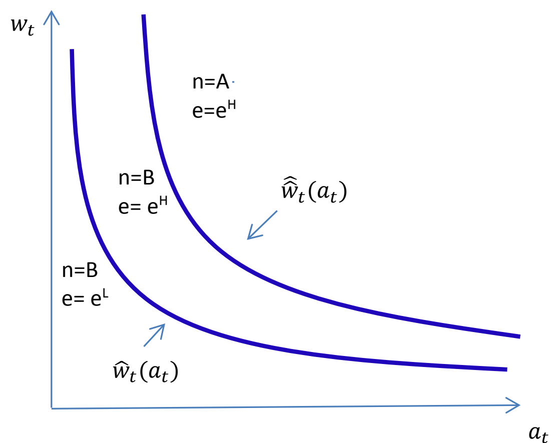


Figure 6 shows a graphical characterization of the equilibrium, for given spillovers and rental rates, with $R_{At} > 0$. The horizontal axis shows the children's ability level a_t and the vertical axis the parents' wage w_t. For any given level of children's ability a_t, there are two thresholds for the parents' wage, $\hat{w}_t(a_t)$ and $\hat{\hat{w}}_t(a_t)$, with $\hat{w}_t(a_t) \leq \hat{\hat{w}}_t(a_t)$, such that parents with wage $w_t < \hat{w}_t(a_t)$ choose to live in B and not to invest in high-level education, parents with wage $\hat{w}_t(a_t) \leq w_t < \hat{\hat{w}}_t(a_t)$ choose to live in B and to invest in high-level education, and parents with wage $w_t \geq \hat{w}_t(a_t)$ choose to live in A and to invest in high-level education. The figure shows that children with richer parents and higher ability tend to be more educated and to live in neighborhood A. On the one hand, for given children's ability, richer parents are more willing to pay the cost of high-level education (τ) and the cost of a higher local externality (higher rental rate). On the other hand, for a given wage, the higher the ability of a child, the more willing the parent is to pay for high-level education and for a higher local externality because of the complementarity between ability and education and between ability and local spillovers, respectively, implied by Assumption 2. For a given ability, a child who grows up in B rather than A has lower probability of getting a high-level education, both because parents living in B are poorer on average and because the size of the local spillover is smaller, reducing the incentive to pay for education even further.

The classic papers in this literature, building on Benabou (1996b) and Durlauf (1996b), typically focus on two extreme cases of segregation by income: either the two neighborhoods are equal to each other and have a representative distribution of income, or they are perfectly segregated, with all the rich agents residing in one neighborhood and the poor in the other. Our model is richer in this dimension, as it allows us to obtain an intensive measure of segregation, which we can match to the data. This is due to the presence of heterogeneity in ability: if all agents had the same ability level, the cut-off function $\hat{\hat{w}}_t(a_t)$ would be horizontal, and the two neighborhoods would feature full segregation by income. However, thanks to the heterogeneity in ability, the two cut-off functions are monotonically non-increasing in ability, and some poorer parents with high ability children choose to live in A to exploit the complementarity with the higher spillover.

Our model also allows us to think about segregation by education. In our simple model, given the binary choice of education, neighborhood A will always be fully segregated, in the sense that all children will get high-level education. However, neighborhood B will generically feature a mix of children with high and low levels of education. In particular, the degree of segregation by education is driven by the distance between the two cut-off functions $\hat{w}_t(a_t)$ and $\hat{\hat{w}}_t(a_t)$. For some parameter configurations, these two functions can coincide, in which case there is perfect segregation by education, as all children living in A will get high-level education and all children in B will not.

3.3 Skill Premium Shock

In this section, we show the model's response to a skill premium shock, which is going to be at the core of the main quantitative exercise in the next section.

To simplify the analysis, we set $e^L = 0$ and $e^H = 1$, and make the following functional form assumptions: $u(c) = g(c) = log(c)$, and

$$\Omega(w, a, e, S_n, \varepsilon) = (b + ae\eta(\beta_0 + \beta_1 S_n)^{\xi})w^{\alpha}\varepsilon. \tag{7}$$

This implies that the wage of a child with a low level of education ($e_t = 0$) is simply equal to $bw^{\alpha}\varepsilon_t$ and depends on neither the child's ability nor the size of the neighborhood spillover, satisfying Assumption 1. Moreover, the wage of a child with high education ($e_t = 1$) is a function of the child's ability as well as of the spillover's size. Notice that β_1 and ξ are the key parameters affecting the strength of the spillover's effect. The specific functional form in (7) also satisfies Assumption 2. In particular, ability is complementary both to education and to the size of the local spillover.

With these assumptions, the cut-off functions that characterize the optimal education and residential choices can be characterized in closed form. Assume that for each ability level a, there is a positive measure of children with high education in neighborhood B.[29] In this case, the two cut-offs are:

$$\hat{w}_t(a) = \tau \left[1 + \frac{b}{a\eta(\beta_0 + \beta_1 S_{Bt})^{\xi}} \right], \tag{8}$$

and

$$\hat{\hat{w}}_t(a) = \tau + R_{At} \left\{ \frac{b + a\eta(\beta_0 + \beta_1 S_{At})^{\xi}}{a\eta[(\beta_0 + \beta_1 S_{At})^{\xi} - (\beta_0 + \beta_1 S_{Bt})^{\xi}]} \right\}. \tag{9}$$

Equation (8) shows that the education cut-off $\hat{w}_t(a)$ is decreasing in ability, as established in Proposition 1, given that the return to education is higher the higher is the ability level. Moreover, for a given ability, the cut-off is decreasing in the local spillover effect in neighborhood B; that is, the higher is the spillover effect in B, the higher are the return to education in that neighborhood and the willingness of parents living there to pay for their children's education. It also shows that, as expected, for a given ability, the willingness of parents living in B to pay for education is higher when the parameters affecting the strength of the return to education and to the spillover, η, β_0, and β_1, are higher and when the cost of education τ and/or the fixed component of the income of children with low levels of education b is lower. Equation (9) shows also that the residential cut-off $\hat{\hat{w}}_t(a)$ is decreasing in a, again in line with Proposition 1, as the return to the

[29] This case arises when the RHS of equation (8) is weakly smaller than the RHS of equation (9) for all ability levels. When this is not the case for some ability a, there is perfect segregation by education - that is, all children with that ability level who grow up in B get the low education level - and the residential and educational cutoff functions coincide and are equal to $\hat{w}_t(a) = \hat{\hat{w}}_t(a) = (\tau + R_{At})[1 + b/a\eta(\beta_0 + \beta_1 S_{At})]$.

larger spillover in neighborhood A is higher the higher is the level of ability. The equation shows that the location decision also depends on the trade-off between the spillover advantage and the cost of living in neighborhood A.

We are now ready to study the response of the economy to an unexpected permanent increase in the skill premium. Through the lens of the model, we can think of an increase in the skill premium as an increase in the parameter η in equation (7), where we interpret high education as college and low education as no college. How is the economy going to respond to such a shock?

First, there is a direct effect of the increase in the skill premium. Keeping the spillovers' size, the house rental price, and the educational and residential choices as given, we note that inequality mechanically increases for two reasons. First, the income gap between college-educated and non-college-educated workers mechanically increases - that is, $\partial^2\Omega/\partial e\partial\eta > 0$ - which is why we interpret a shock to η as a skill premium shock. Second, the return to the local spillover effect, which is complementary to education, is also higher; that is, $\partial^2\Omega/\partial S_n\partial\eta > 0$. This direct effect generates per se an increase in inequality because richer kids have a higher probability both to be college-educated and to live in neighborhood A, where the spillover effect is larger.

The second effect comes from the change in the educational and residential choices, keeping the spillover levels fixed at their pre-shock values. Using equations (8) and (9), we can derive the response of the cut-off functions to an increase in η as follows:

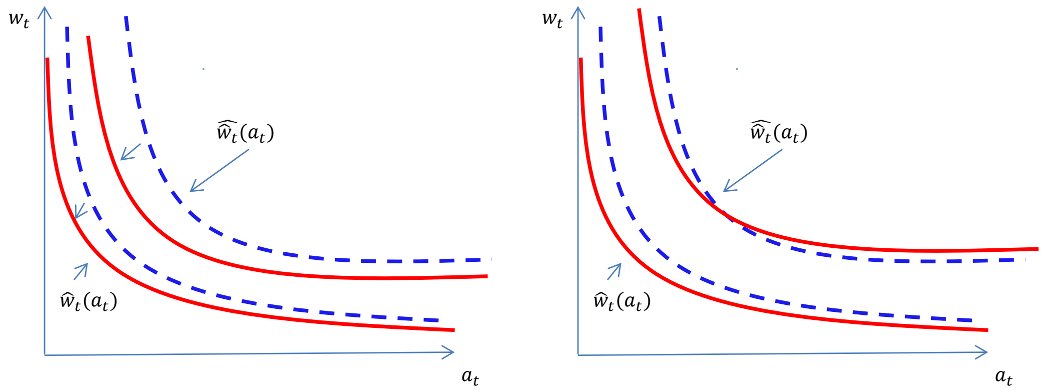
$$\left.\frac{d\hat{w}_t(a_t)}{d\eta}\right|_{S_{At},S_{Bt}} = -\frac{1}{\eta^2}\frac{\tau b}{a_t(\beta_0+\beta_1 S_{Bt})^{\xi}}, \tag{10}$$

and

$$\left.\frac{d\hat{\hat{w}}_t(a_t)}{d\eta}\right|_{S_{At},S_{Bt}} = -\frac{R_{At}b}{\eta^2 a_t[(\beta_0+\beta_1 S_{At})^{\xi}-(\beta_0+\beta_1 S_{Bt})^{\xi}]} + \left\{\frac{b+a\eta(\beta_0+\beta_1 S_{At})^{\xi}}{a\eta[(\beta_0+\beta_1 S_{At})^{\xi}-(\beta_0+\beta_1 S_{Bt})^{\xi}]}\right\}\frac{dR_{At}}{d\eta}. \tag{11}$$

These derivations show that in partial equilibrium - that is, when the rental rate is kept fixed ($dR_{At}/d\eta = 0$) - both cut-off functions shift to the left, so that more children of any ability get higher education and live in neighborhood A. The change in the educational choice is intuitive: the higher the skill premium, the higher the return to college, conditional on any level of ability. Moreover, given that the local spillover is complementary to education, the higher the

Figure 7: Cut-Off Response to Skill Premium Shock



(a) Partial Equilibrium (b) General Equilibrium

skill premium, the higher is the return to the spillover, and hence the higher is the demand to live in neighborhood A, conditional on any level of ability. Panel (a) in Figure 7 qualitatively shows the partial equilibrium response of the educational and residential cut-off functions to the skill premium shock, when spillovers in both A and B and the rental rate in A are kept fixed at the pre-shock levels. The figure shows that both cut-off functions also become flatter after the shock, as it is easy to derive that $d^2\hat{w}_t(a_t)/da_t d\eta\big|_{S_{At},S_{Bt}} > 0$ and $d^2\hat{\hat{w}}_t(a_t)/da_t d\eta\big|_{S_{At},S_{Bt}} > 0$ if $dR_{At}/d\eta = 0$. This means that, with our functional form, the marginal impact of ability on the return to education is smaller when the skill premium is larger.

Next, we analyze the general equilibrium effect, coming from the response of the rental rate in neighborhood A to clear the housing market. Panel (b) in Figure 7 shows that when we consider the general equilibrium, the residential cut-off function shifts back to the right, but in a tilted fashion. As we explained above, taking as given the rental rate and the spillover effects, the demand to live in neighborhood A will increase because of the differential spillover and the complementarity between the spillover and education, shifting the residential cut-off to the left. Given that the housing supply in neighborhood A is fixed, this demand increase pushes up rental rates in that neighborhood, shifting the housing demand back to the right. In particular, the figure shows that the shift back is more pronounced for poorer parents, who won't be able to afford the higher cost of living in the rich neighborhood, irrespective of their children's ability. On net, this generates the tilting that we see in panel (b) in Figure 7, which leads to a higher degree of income segregation: after the shock, some richer families will move to neighborhood A even if their children do not have high ability, at the expense of some talented children from poorer families, who will be induced to move to neighborhood B.[30] This implies that more children from rich families will be exposed to stronger spillover effects and will have even higher future income, while more poor children will grow up in neighborhoods with weaker externalities and will have worse prospects for their future. This, in turn, will amplify the increase in inequality and reduce intergenerational mobility.

The analysis so far has kept the spillover size in the two neighborhoods as given. It has also shown that if a skill premium shock hits a segregated economy, the degree of segregation by income increases, and the response of inequality is amplified as a result. However, in our model, the spillover sizes in the two neighborhoods respond endogenously to the shock. The increase in η increases the future wage of all the educated children, increasing the spillover size in both neighborhoods, S^A and S^B. The shift in the educational cut-off implies that more children get high education in neighborhood B, increasing even further the spillover in that neighborhood. Moreover, the tilting of the residential cut-off implies that neighborhood A will be populated by richer but less talented children. This has two effects. First, it tends to increase the spillover gap between the two neighborhoods: for a given ability, children with richer parents who live in A have higher future income and children of poorer parents who live in B have lower future income (not only because of the direct effect of their parents' wage but also because they will have higher chance to get educated). Second, it tends to decrease the same gap, given that more talented children move from A into B, pushing in the opposite direction. The quantitative exercise in Section 4 will show that the sorting effect by income dominates, so that the spillovers' size in both neighborhoods will increase, but the one in neighborhood A will increase relatively more, generating an additional source of inequality amplification.

4 Quantitative Exercise

As the data show, from 1980 onward, the US experienced a steady increase in labor income inequality. Many factors have contributed to this increase, but in this paper, we focus on skill-biased technical change, which is widely recognized to be a crucial one (see, for example, Katz and Murphy, 1992; Autor, Katz and Krueger, 1998; Goldin and Katz, 2001; Card and Lemieux,

[30]The fact that the new residential cutoff policy crosses the old one is typical in our quantitative exercises, but in general depends on the evolution of the joint distribution of wages and abilities.

2001; Acemoglu, 2002; Autor, Katz and Kearney, 2008).

In this section, we generalize the simple model to explore the quantitative response of the economy to an unexpected, one-time, permanent shock to the skill premium, as described in subsection 3.3. We show that the model is able to replicate well the dynamics of inequality and segregation and the patterns of the intergenerational mobility across income groups. We also validate the model using data on the dynamics of housing prices and neighborhood sizes.

4.1 General Model

For the quantitative analysis, we generalize the model in a number of dimensions to make it more capable of capturing important features of the data.

First, to allow for richer sorting dynamics, we extend the analysis to a city with three neighborhoods instead of two. Neighborhood n is now $\in \{A, B, C\}$. Having an intermediate neighborhood makes the geographic decisions of the agents less extreme, allowing for more realistic sorting dynamics. We focus on equilibria where $R_{At} > R_{Bt} > R_{Ct}$ and $S_{At} > S_{Bt} > S_{Ct}$ for all t, so that kids who grow up in neighborhood A are the ones with highest expected income and those who grow up in C the ones with the lowest.

Second, we generalize the formalization of the housing market, allowing for a general upward-sloping housing supply curve in each neighborhood. In particular, the housing supply curve in neighborhood n at time t is given by

$$H_{nt} = \lambda_n \left(\frac{R_{nt}}{\bar{w}_t} \right)^{\phi_n},$$

where ϕ_n represents the housing elasticity in neighborhood n, λ_n is a shift parameter in the same neighborhood, and $\bar{w}$ is the average parental wage in the city at time t.[31] This implies that neighborhood sizes become endogenous and the model generates dynamic patterns that we can compare with the data.

Third, we introduce two different forms of preference shocks. The first type of preference shock is meant to capture that all agents prefer the richer neighborhoods to the poorer ones for reasons

[31] We normalize the housing rental rates by average wages so that if all prices increase proportionally to wages, there are no real effects on the neighborhood sizes.

other than the education spillover - in particular, fixed amenities, such as parks, water access, restaurants, or even status considerations. We assume that utility from current consumption for an agent who chooses neighborhood n is given by $log[(1+\theta_n)c]$, where $\theta_A > \theta_B = 1 > \theta_C$ with probability π and $\theta_A = \theta_B = \theta_C = 1$ with probability $1-\pi$.[32] The second type of preference shock is an idiosyncratic component that is iid across neighborhoods and agents. Specifically, the utility of an agent is now given by

$$log[(1+\theta_n)c] + log[\Omega(w,a,e,S_n,\varepsilon)] + \sigma_\zeta \zeta_n,$$

where Ω satisfies condition 7 and ζ_n follows a type-I extreme value distribution with scale parameter σ_ζ. This shock introduces some additional randomness that is not systematically related to any particular ranking of the neighborhoods and helps make the model analytically tractable.

Both these types of preference shocks help obtain a more realistic setting, where not all parents who live in the more expensive neighborhood choose high levels of education for their children. In our simple model, the only reason to pay a higher rent to live in neighborhood A is to exploit the higher externality that affects the returns to education. In reality, residential choices are not driven purely by educational considerations. By missing this feature of reality, the simple model might generate a distribution of children growing up in neighborhood A biased towards too high ability and biased toward levels of ability and educational investment that are too high.

Fourth, we make the educational choice continuous to allow for richer investment decisions in education, which we believe are particularly important given the nature of our mechanism. In the simple model with binary educational choice, rich parents are constrained in how much they can invest in their children's education, given that the best they can do is to pay for their college. This means that the binary choice would arbitrarily bound the possible increase in the spillover in response to a skill premium shock. To overcome these limitations, we now assume that the educational choice is continuous, with $e \in R_+$, and that the cost of education is τe^γ.[33] The optimal educational level turns out to be increasing in the parents' wage, innate ability, and, more importantly, the size of the local spillover, because of the complementarity assumption. This

[32]This shock alone would generate residential segregation, even in the absence of local spillovers. In Section E.1, we will explore a model where it is the only driver of residential segregation and spillovers are global.

[33]Calibrating the baseline model would result in too much intergenerational mobility. The continuous educational choice is more appealing also in light of the evidence in Duncan and Murnane (2016) that there has been an increasing polarization between educational investment in rich families and investment in poor ones.

generates an amplification channel for the feedback between inequality and segregation. When the gap in local spillovers increases, parents living in neighborhoods with stronger spillovers invest more in their children's education, and doing so increases the gap in local spillovers in turn.

With these two modifications, the problem of household (w,a) becomes

$$U(w,a) = \max_{e,n} log((1+\theta_n)(w - R_{nt} - \tau e^{\gamma})) + log((b + ae(\beta_0 + \beta_1 S_{nt})^{\xi})w^{\alpha}\varepsilon) + \sigma_{\zeta}\zeta_n. \quad \text{(P3)}$$

The equilibrium definition is a natural extension of Definition 1 in Section 3.2, where the agents' optimization problem is given by P3, there are three neighborhoods, and the housing market clearing conditions for the three neighborhoods are given by equation (4.1).

4.2 Calibration

We now describe our calibration strategy. Because the rise in labor income inequality started in 1980, we assume that in 1980, the economy is in steady state and is hit by an unexpected, one-time, permanent shock to the skill premium. In particular, we change η to match the increase in the skill premium in the data between 1980 and 1990.

In the model, individuals live for two periods: in the first period, they are young and go to school, and in the second period, they are old and work. As noted by Fernandez and Rogerson (1998), in this class of models, individuals spend the same time in period 1 and 2, so we could target the length of a period to the working period or to the schooling period. Given our focus on the educational investment, we choose to interpret one period as 10 years.[34] We interpret period $t = 0$ as 1980, when the economy is in steady state. We assume that at that time, an unexpected, permanent shock hits the economy, and η increases to $\eta' > \eta$, so that the skill premium goes from 0.39 in 1980 ($t = 0$) to 0.55 in 1990 ($t = 1$).

We choose parameters so that the steady state equilibrium of the model matches salient features of the US economy mostly in 1980.[35] Table 1 shows the targets of our baseline calibration, which we are now going to discuss.

[34]The schooling age could be interpreted as 10 or 15, years depending on which level of education one targets. Another factor in our choice of 10 years is that census data are available every 10 years.

[35]Below, we explain that the data available for the rank-rank correlation and the neighborhood exposure effect give us only one data point, which we interpret as an average between 1980 and 2000.

Table 1: *Calibration Targets*

Description	Data	Model	Source
Return to college 1980	0.391	0.391	Goldin and Katz (2009)
Return to college 1990	0.549	0.553	Goldin and Katz (2009)
Gini coefficient	0.376	0.376	Census 1980
Dissimilarity index by income	0.334	0.334	Census 1980
Income 25th/75th p	0.667	0.694	Chetty and Hendren (2018b)
Rank-rank correlation	0.341	0.336	Chetty et al. (2014)
Return to spillover 25th p	0.062	0.062	Chetty and Hendren (2018b)
Return to spillover 75th p	0.046	0.046	Chetty and Hendren (2018b)
Neighborhood A size 1980	0.194	0.193	Census 1980
Neighborhood A size 1990	0.217	0.215	Census 1990
Neighborhood B size 1980	0.301	0.301	Census 1980
Neighborhood B size 1990	0.250	0.278	Census 1990
Average population growth rate	1.089	1.089	Census 1980-2010
Share of rich in A 1980	0.437	0.459	Census 1980
Share of rich in B 1980	0.225	0.212	Census 1980
R_A/R_B	1.253	1.252	Census 1980
R_B/R_C	1.277	1.279	Census 1980
College share A	0.340	0.351	Census 1980
College share B	0.178	0.200	Census 1980
Dissimilarity index by education	0.243	0.266	Census 1980

The first two targets in Table 1 are the US college premia in 1980 and 1990 from Goldin and Katz (2009). In the model, we map the skill premium in 1980 to the steady state difference between the average log wage of college-educated individuals and the average log wage of the others. Given that the educational choice is continuous, we define a cut-off $\hat{e}$ so that individuals with an education level above $\hat{e}$ are college educated and the ones with an education level below are not. We choose $\hat{e}$ so that in 1980, 17.8% of the population is college educated, as in the census data.[36] Finally, we map the skill premium in 1990 to the same difference one period after the shock, keeping the college cut-off $\hat{e}$ constant.

The next target is the value of the Gini coefficient in 1980 with the sample restricted to families with children, which we described in Section 2. Another measure of inequality at the metro area that we target is the ratio of the average income for families in the top 25th percentile of the income distribution to the average income for families in the bottom 25th percentile.

[36]To calculate this number, we look at the number of people older than 25 years of age who completed college at the census tract level.

Next, we target the 1980 value of the dissimilarity index by income. In order to calculate the dissimilarity index, we first map the three neighborhoods in the model - A, B, and C - to the data. For each MSA, we group the census tracts according to the share of rich families that live there, and we define as "rich" the families in the top 20th percentile of the income distribution of that MSA. In particular, for each MSA, we define neighborhood A as the group of census tracts with more than 30% rich families, neighborhood C as the group of tracts with less than 17% rich, and neighborhood B as the group of residual tracts. Given that, as we showed in Section 2, the rise in inequality and segregation has been driven by the top of the distribution, we choose the cut-offs of 17% and 30% so as to have roughly 50% of the population in neighborhood C and the rest split between A and B. Once we have grouped the census tracts in three neighborhoods for each MSA, we calculate the dissimilarity indexes for all the MSAs and then average them, weighting by population. Once again, we restrict the sample to families with children.

Another feature of the US data we target is the level of intergenerational mobility. To this end, we target the rank-rank correlation between log wages of parents and children estimated using administrative records by Chetty et al. (2014).[37] They use children born between 1980 and 1982 and calculate parent income as mean family income between 1996 and 2000 and child income as mean family income between 2011 and 2012, when the children are approximately 30 years old. Given that this correlation is calculated over several decades, we map it in the model to the average rank-rank correlation across 1980, 1990, and 2000, where the 1980 value corresponds to the steady state and the 1990 and 2000 values are calculated after the skill premium shock hits the economy.

A key target for our exercise is what we call the "return to spillover" - that is, the effect of neighborhood exposure on children's income in adulthood. This effect is difficult to measure in the data. Fortunately, there has been a recent growing literature that uses micro data to estimate it. In particular, we use the results from the quasi-experiment in Chetty and Hendren (2018b). Using tax return data for all children born between 1980 and 1986, Chetty and Hendren (2018b) estimate the effect of local spillovers on children's future income, by looking at movers across US

[37]The rank-rank correlation is the relationship between the rank based on income of children relative to others in the same birth cohort and the rank based on parents' income relative to others in the same birth cohort. We choose this statistic instead of the log-log correlation or other measures, because as emphasized by Chetty et al. (2014), it provides a more robust summary of intergenerational mobility.

counties.[38] We focus on their estimates for families moving across counties within the same commuting zone, given that we use the metro area as our geographic unit of analysis. Their estimate implies that for a child with parents at the 25th percentile of the national income distribution, growing up in a 1 standard deviation better county from birth would increase household income in adulthood by approximately 6.2%. This number becomes 4.6% for a child with parents at the 75th percentile of the income distribution.[39] These are the values that we target in our calibration. Let us explain how we map these targets to our model. First, we map the "movers" in Chetty and Hendren (2018b) to the parents who decide to live in a neighborhood different from the one where they grew up - that is, the one chosen by their parents. Then, we calculate the standard deviation of the expected future wage of the children of "movers" at the 25th percentile and at the 75th percentile of the income distribution and divide that by the average wage of the parents.[40] Given that these children are born between 1980 and 1986, in 1984-1998 and 1990-2004 they will be in pre-kindergarten to 12th grade and hence exposed to the local spillover. Thus, as we do for the rank-rank correlation, we map these numbers to the average "spillover effects" in the model across 1980, 1990, and 2000; again, the values of 1990 and 2000 are calculated after the shock.[41]

Our model is able to match a higher spillover return for the lower percentile of the distribution. Given that innate ability is positively correlated with parents' income, one would expect that the assumption of complementarity between ability and spillover would make the neighborhood effect stronger for richer families. This is true for the population, but not for the selected sample

[38]Chetty and Hendren (2018b) control for selection effects by looking at families that move from one county to another with children of varying age, so they were exposed for different fractions of their childhood to the new county. Building on this logic, they effectively use a sample of cross-county movers to regress children's income ranks at age 26 on the interaction of fixed effects for each county and the fraction of childhood spent in that county. The identification assumption is that children's future income is orthogonal to the age they move to a new county.

[39]See table II in Chetty and Hendren (2018b). The table also shows estimates for families moving across counties that are not necessarily within the same commuting zone. In Section 5.1, we explore how these alternative estimates would change our main results.

[40]The exposure effect is equal to $\frac{1}{\bar{w}}\sqrt{\frac{1}{2}\sum_{n\in\{A,B,C\}}(E(w'|n,m)-E(w'|m))^2}$, where $E(w'|n,m)$ is the expected income of children of movers (m) from neighborhood n and $E(w'|m)$ is the average expected income of all movers (m).

[41]The fact that we match the rank-rank correlation and the spillover effects to averages in the model over the period 1980-2000 is why we simultaneously calibrate the parameters of the model and the size of the shock. An alternative calibration strategy would be to calibrate all the parameters of the model so that the steady state matches only the targets for 1980, then use the rank-rank correlation from Chetty et al. (2014) and the spillover effects from Chetty and Hendren (2018b) as if they were numbers for 1980. This alternative calibration would generate a larger increase in the implied spillover effect after the shock, so our choice is conservative.

of movers that we use to map the statistics in Chetty and Hendren (2018b). Let us use the simple two-neighborhood model in Section 3 to explain the mechanism. As we have shown in Figure 6, the cut-off function that represents the residential decision is non-increasing in the space ability/parent's wage. This implies that families that live in A tend to both be richer and have children with higher ability. This also implies that in equilibrium, rich families who decide to live in B instead of A have children with lower ability than a random rich family in the population. Moreover, a poor family that moves from neighborhood B to neighborhood A will tend to have a child with higher ability than a richer one, because she will have a higher ability cut-off for the family to be willing to pay the higher housing cost. So, on average, poor families that move from B to A will tend to have children with higher ability than rich families that move from B to A, and hence higher return from the spillover.

In the model, the size of the neighborhoods is endogenously determined and evolves over time in response to the shock. We use micro data on the evolution of population shares across census tracts to calculate the size of the three neighborhoods in terms of population for the average MSA. We target the population growth in the average metro area and the size of neighborhoods A and B in both 1980 and 1990.[42] We also target the share of rich families in the different neighborhoods in 1980.

Another important object in our model is the relative cost of housing in the different neighborhoods. We use housing values in 1980 at the census tract level from the census data and convert them into rental rates.[43] Using the same methodology to aggregate census tracts described above, we calculate the ratio of rental rate in neighborhood A to that in neighborhood B and the ratio of rental rate in neighborhood B to that in neighborhood C. We then average these ratios across MSAs, weighting by population.

Finally, we use census tract data to calculate the number of college graduates older than 25 residing in neighborhoods A, B, and C for the average MSA. We then target the share of college educated individuals in neighborhood A and the share of college educated individuals in neighborhood B. Moreover, we can use these data to calculate the dissimilarity index by education for the average MSA, where the two exclusive categories are college educated and non-college-educated.

[42] The size of neighborhood C can be calculated as a residual.

[43] We use a standard coefficient of 0.05 for the conversion.

Table 2: *Parameters*

Parameter	Value	Description
ρ	0.44	Autocorrelation of log ability
σ_v	1.06	St. dev. of log ability
σ_ε	0.65	St. dev. of log wage noise shock
α	0.25	Wage function parameter
β_0	0.14	Wage function parameter
β_1	0.07	Wage function parameter
ξ	1.07	Wage function parameter
b	1.90	Wage fixed component for no-college
γ	4.78	Education cost parameter
θ_A	1.20	Preference shock for neighborhood A
θ_C	0.46	Preference shock for neighborhood C
π	0.53	Preference shock probability
σ_ζ	0.15	St. dev. of idiosyncratic preference shock
λ_A	0.13	Shift parameter of housing supply in A
λ_B	0.30	Shift parameter of housing supply in B
ϕ_A	2.37	Elasticity of housing supply in A
ϕ_B	0.11	Elasticity of housing supply in B
ϕ_C	13.62	Elasticity of housing supply in C
η'	2.78	Skill premium shock
n	1.09	Average population growth

Table 2 shows the parameters that we are using to calibrate the model, their calibrated value, and their description. We normalized the values of η, τ, the mean of the ability process a_t, and the mean of the noise shock to the wage process ε to be all equal to 1 in steady state. Moreover, λ_C is pinned down by normalizing the average wage to be equal to 2.44, which is its empirical counterpart in 1980.[44] Notice that we have 20 parameters to match 20 targets.

4.3 Skill Premium Shock

We are now ready to show the response of the economy to a skill premium shock. As we explained above, we assume that in 1980, the economy is in steady state and that at the end of the period, it is hit by an unexpected, one-time, permanent increase in η.

Table 3 shows the response of the economy to such a shock, one, two, and three periods ahead.

The first row shows the dynamics of the return to college. Remember that we choose the shock to

[44]For details about the normalizations see Appendix C.

Table 3: *Response to a Skill Premium Shock*

	t = 0	t = 1	t= 2	t= 3
Return to college	0.391	0.553	0.591	0.610
Gini coefficient	0.376	0.402	0.431	0.447
Dissimilarity index	0.334	0.382	0.418	0.438
A/B spillover ratio	1.112	1.540	1.787	1.920
B/C spillover ratio	1.126	1.181	1.166	1.162
R_A/R_B	1.135	1.495	1.703	1.824
R_B/R_C	1.162	1.234	1.228	1.229
SizeA	0.193	0.215	0.230	0.235
SizeB	0.301	0.278	0.256	0.237
SizeC	0.506	0.507	0.514	0.528

match the increase in the return to college between 1980 and 1990 in the data. What is interesting is that the one-time, unexpected, permanent shock generates persistence in the return to college that keeps increasing after 1990, just as the return to college in the data does. In particular, the return to college in the data is equal to 0.61 and 0.68 in 2000 and 2010 (from Autor et al. (2020)) respectively, which is close to the predicted path in our model.

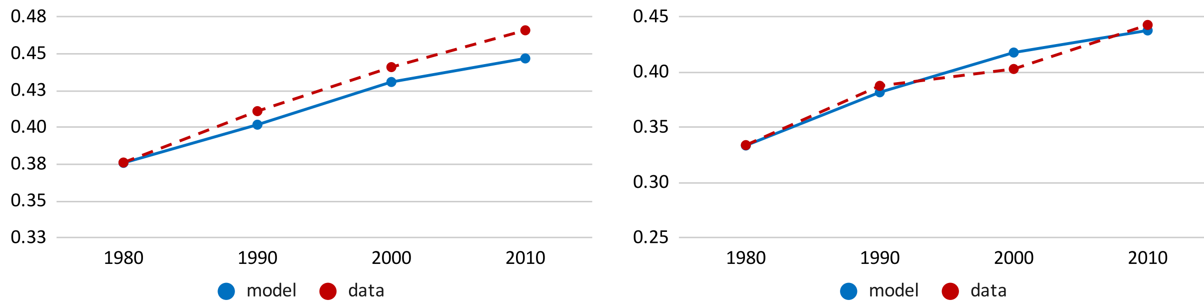
The second and third rows show the response of inequality and segregation, captured by the Gini coefficient and the dissimilarity index. To visualize these results, Figure 8 shows the responses (solid lines) of inequality and segregation to the shock in the model, together with their pattern in the data (dashed lines). While the values of inequality and segregation in 1980 are targets of the calibration, their path over time after 1980 is an outcome of the model and can be compared with the one in the data as a form of validation.

Panel (a) shows that the dynamics of inequality in response to the skill premium shock in the model are close to those in the data. However, the model generates a bit less inequality growth than the data, which is to be expected, given that there are other sources of inequality increase that are outside the model. Panel (b) shows that the model generates a response of segregation to the skill premium shock that also does a good job of replicating the pattern in the data.[45]

Table 3 also shows that in response to the shock, both the spillover in A relative to that in B and the spillover in B relative to that in C increase, and so do the respective rental rate ratios. At

[45]Note that the dissimilarity in Figure 8 is calculated aggregating the census tracts into three neighborhoods, as explained in Section 4.2, and so it is not exactly the same dissimilarity measure reported in Figure 4.

Figure 8: Responses to a Skill Premium Shock



(a) Inequality

(b) Segregation

the same time, the sizes of neighborhood A and C increase, while the size of neighborhood B shrinks. Moreover, these effects are persistent but declining over time.

As we discussed in subsection 3.3, there is a rich feedback effect between inequality and segregation at the heart of our model. First, as the skill premium increases, inequality mechanically increases because college educated workers earn even more than the non-college-educated ones. Given that educated workers are more likely to grow up in neighborhood A than B and in B than C, segregation by income also mechanically increases. Second, as the return to education increases, and given the complementarity between the neighborhood spillover and education, the return to living in neighborhood A relative to that of living in B and the return to living in B relative that of living in C increases. Given that housing supply is somewhat elastic, this shows up in part as a response of housing supply and in part as a response of rental rates. In particular, neighborhood A becomes more attractive relative to the others, so not only its size but also its rental rate increase more than those of the others. This implies that some households move to A because it is more attractive, but at the same time, some poorer households who previously lived in A may be pushed out into neighborhood B or C because of the higher rental rate. On net, the size of A increases. At the same time, neighborhood B becomes more attractive than neighborhood C, implying that the rental rate in B increases relative to rental rate in C. This again implies that more talented kids would be attracted to neighborhood B, but the increase in rental rate may actually force poorer kids, even talented ones, to move from B to C. On net, the table shows that the size of B declines. These sorting patterns imply that segregation by income increases, and this will increase the spillover gaps between A and B and B and C, as shown in the table. Such an increase in spillover differentials feeds back into even higher future inequality.

The dynamics of the rental rate ratios and of the neighborhood sizes can also be compared with the data for further validation. Table 4 shows the dynamics of the rental rate in neighborhood A relative to that in B and the rental rate in B relative to that in C and the sizes of the three neighborhoods. Comparing them with the corresponding model dynamics in Table 3 shows that the model, although stylized, is able to replicate the qualitative properties of these patterns. In particular, the increase in inequality and segregation happened at the same time as an increase in population concentration in the more extreme neighborhoods. The population living in neighborhoods characterized by a percentage of rich families between 17% and 30% went from 30% to 22%, while neighborhoods with an extreme concentration of either rich or poor households expanded.

Table 4: *Neighborhood Sizes and Rental Rates: Data*

	1980	1990	2000	2010
R_A/R_B	1.253	1.282	1.326	(1.265)
R_B/R_C	1.277	1.318	1.343	(1.291)
SizeA	0.193	0.217	0.228	0.251
SizeB	0.301	0.250	0.229	0.215
SizeC	0.506	0.533	0.543	0.534

NOTE: The rental rate ratios in 2010 are in parenthesis because the definition of house prices in the census changed in 2010 from single family housing units to a broader category that also included condos and mobile homes.

4.4 Intergenerational Mobility

An important implication of the model is that the same mechanism behind the feedback between segregation and inequality also generates low intergenerational mobility. As living in neighborhoods with higher spillover is expensive, richer families can afford to expose their children to strong local spillover effects, while poorer families are forced to live in less attractive but more affordable neighborhoods. This inevitably makes it more difficult for poor children to climb up the social ladder and easier for richer children to perpetuate their status. One summary statistic that captures the degree of intergenerational mobility is the rank-rank correlation that we target in our calibration. A richer picture of the degree of intergenerational mobility is given by the intergenerational mobility matrix that reports the probability of a child's being in a given income quintile, conditional on the parents' income quintile. In Table 5, Panel (a), we show the intergenerational mobility matrix generated by the model. The table shows averages across the period 1980-2000 to make it comparable to the data. Panel (b) of Table 5 shows the same matrix calculated using administrative data by Chetty et al. (2014).

Table 5: *Intergenerational Mobility Matrix*

Panel (a): Model

	Children's Quintile				
Parents' Quintile	1	2	3	4	5
1	34.0%	24.7%	19.3%	14.4%	7.6%
2	23.3%	22.9%	21.5%	19.1%	13.3%
3	18.2%	20.8%	21.6%	21.5%	18.0%
4	13.8%	18.2%	20.9%	23.2%	24.0%
5	8.3%	13.4%	17.9%	23.7%	36.6%

Table 6: *Panel (b): Data*

	Children's Quintile				
Parents' Quintile	1	2	3	4	5
1	33.7%	28.0%	18.4%	12.3%	7.5%
2	24.2%	24.2%	21.7%	17.6%	12.3%
3	17.8%	19.8%	22.1%	22.0%	18.3%
4	13.4%	16.0%	20.9%	24.4%	25.4%
5	10.9%	11.9%	17.0%	23.6%	36.5%

Without targeting it, our model is able to replicate quite well the intergenerational mobility matrix, which is a good validation of the model. The main feature of the data, well matched by the model, is that the highest probabilities are in the diagonal and, in particular, in the two extreme quintiles. That is, children tend to stay with higher probability in the same income quintile as their parents, and this is especially true for the highest and, even more, for the lowest quintiles. The data show that the probability that a child whose parents are in the lowest quintile of the distribution stays in the same income quintile (Q5-Q5 probability) is 36.5%, and the model implies 36.6%. This is an important number to take into consideration for policy prescriptions, and in future work, we will think about how it is affected by alternative policies. Another important statistic is the probability that a child whose parents are in the highest quintile of the distribution stays in the same income quintile (Q1-Q1 probability), which is 33.7% in the data and 34% in the model.

Another implication of the model is that intergenerational mobility declines over time in response to the skill premium shock. For example, the rank-rank correlation goes from 0.34 in 1980 to roughly 0.35 in 2010, while the Q5-Q5 probability goes from 29.2 in 1980 to 42.8 in 2010. Unfortunately, given the limited availability of data, it is hard to calculate a reliable time-series for the rank-rank correlation or for the intergenerational mobility matrix. However, Aaronson and Mazumder (2008) show some indirect evidence of a positive relationship between the skill premium and the IGE (intergenerational elasticity) that is consistent with our findings. Moreover, Kulkarni and Malmendier (2022) use cross sectional variation to show that cities with higher homeownership segregation have lower intergenerational mobility.

5 Segregation's Contribution to Inequality

We now use the model to perform a number of exercises in order to answer our main question: How important is segregation in amplifying the effects of a skill premium shock to income inequality? We also explore how segregation affects intergenerational mobility. In the next subsection, we propose two main exercises that help us answer this question. Next, we explore a number of exercises to better understand the quantitative role of different channels in the model.

5.1 Main Counterfactual Exercises

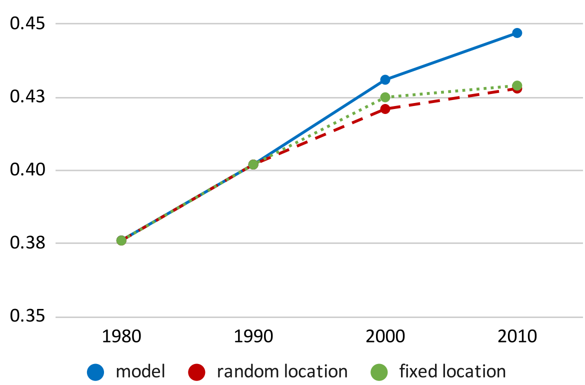
In this section, we show the main counterfactual exercises that quantify the role of segregation in amplifying inequality. In the model, the presence of local spillovers generates sorting of richer parents into the better neighborhoods, generating residential segregation by income. As explained in detail in Section 4.3, segregation amplifies the response of inequality to a skill premium shock because of two main effects. First, even if the strength of the spillover in the two neighborhoods is kept fixed, the increase in income segregation implies that more of the rich children will benefit from exposure to the better neighborhood and will become even richer, while more of the poor children will be forced to grow up in worse neighborhoods, which will worsen their income prospects. Second, the higher degree of income segregation will, in turn, endogenously translate into a larger gap between the spillover effect in the different neighborhoods, further increasing inequality.

One natural way to assess the contribution of segregation to inequality is to shut down families' residential choice in response to the shock, which is going to mute the sorting process. We can do that either by randomly relocating households across neighborhoods or by not allowing agents to move away from the neighborhood where they grew up. We construct two counterfactual exercises corresponding to these two alternatives.

First, we consider a counterfactual exercise where, at the moment of the shock and at any time after that, families are randomly relocated across the three neighborhoods.[46] This implies that the sizes of the local spillover effects in the three neighborhoods are equalized, given that the distribution of families is identical, and so is the expected income of the children growing up there. We impose that the rental rate in the three neighborhoods is the same and is equal to the rental rate that clears a single metro-wide housing market.

In this exercise, children are still exposed to a positive externality that evolves over time, but the strength of the spillover is the same for any location, so the spillover becomes global instead of local. Parents' income still affects children's wages through the direct effect on the wage function Ω and through the educational choice, but the location is not relevant for their future earnings. This mitigates the effect of intergenerational linkages on income and hence mitigates the response of income inequality to a skill premium shock.

Figure 9: Counterfactuals: Inequality and Segregation



(a) Inequality

(b) Segregation

In panel (a) of Figure 9, we compare the response of inequality to the skill premium shock in the baseline model (blue solid line) with the response of the economy when families are randomly relocated across the neighborhoods every period after the shock (red dashed line). The figure shows that segregation contributes to 27% of the increase in inequality over the whole period between 1980 and 2010. The same exercise implies that segregation also amplifies the decrease in intergenerational mobility in response to the shock. In particular, it contributes to 32% of the increase in the rank-rank correlation between 1980 and 2010.

[46]We use the same parameters calibrated in Section 4.2, given that we assume that the economy is in the same steady state in 1980.

Second, we consider a counterfactual exercise where, after the shock, parents are not allowed to move to a neighborhood different from the one where they grew up; that is, locations are fixed. We then keep the rental rates in the three neighborhoods fixed at their steady state levels.[47] The green dotted line in Figure 9 (a) represents the response of inequality in this exercise and shows that according to this counterfactual exercise, segregation contributes to 25% of the inequality increase between 1980 and 2010. We also calculated that segregation contributes to 33% of the increase in the rank-rank correlation in the same period. It is reassuring to notice that the two exercises generate similar results.

Although the two counterfactual exercises deliver a similar pattern for the dynamics of inequality, the patterns of segregation are very different. Figure 9 (b) compares the dynamics of the dissimilarity index in the baseline model and in the two counterfactuals. In the first exercise, where we randomly re-allocate households at every time after the shock, all neighborhoods end up having the same distribution of households, and the dissimilarity index is constant and equal to zero (dashed red line). To be more conservative, we consider the second exercise, where we keep the location of households fixed and the level of the dissimilarity index at the time of the shock does not change. However, in this exercise, even if location is fixed, the dissimilarity index does not stay constant; rather, it declines over time after the shock (dotted green line). This decline reduces the contribution to inequality.

The main reason why the dissimilarity index declines in the counterfactual with fixed location is the evolution of the distribution of ability across neighborhoods. In particular, in our model, given the complementarity between spillover and ability, the neighborhoods with higher spillover tend to attract families with children with higher ability. Once we shut down the sorting process, the average ability in the neighborhoods tends to converge over time, given mean reversion in the ability process.[48] One may think that endogenous education might be another reason behind the decline in segregation, but in Appendix D, we show that quantitatively, this is not the case.

A key moment behind these results is the estimate of the strength of the neighborhood spillover, which we take from Chetty and Hendren (2018b). They find that growing up in a 1 standard

[47] In Appendix D, we also explore the alternative exercise where agents make residential choices, but the housing supply in the three neighborhoods is fixed at the their steady state levels, and rental rates are such that the housing markets clear. Figure 22 shows that the contribution of segregation to inequality calculated with this alternative exercise is very similar.

[48] See Figure 20 in Appendix D.

deviation better county within the same commuting zone would increase adult income by 6.2% for a child in the lowest 25th percentile of the income distribution and by 4.6% for a child in the top 25th percentile. We also explore how these results would change if we recalibrate the model targeting 10.4% and 6.4%, for the lowest and the highest 25th percentile respectively, which are the estimates that Chetty and Hendren (2018b) calculate by looking at families moving across counties, but not necessarily within the same commuting zone. In this case, the contribution of segregation to inequality is, as expected, even higher. It is equal to 54% according to the first counterfactual exercise and to 53% according to the second.[49]

5.2 Understanding the Mechanism

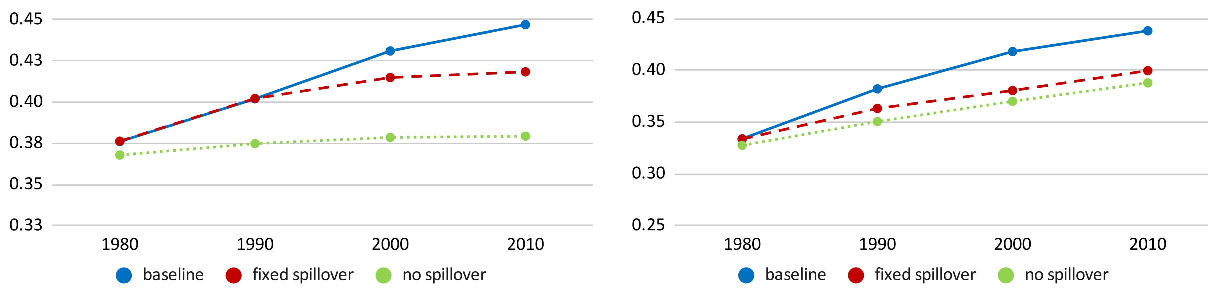
To better understand how local spillover amplifies the effect of inequality on segregation, we now show two alternative counterfactual exercises that complement the previous ones. First, we explore what would happen to inequality if there were no local spillovers at all. Second, we explore what would happen if there were local spillovers, but they were not responsive to the shock. The first exercise is extreme, as it shuts down any form of externality, and the second one keeps the spillovers constant, focusing only on the endogenous response of the spillovers' strength.

The first additional exercise is to consider the case with no spillover effects - that is, where the wage function Ω is not affected by the spillover, or $\beta_1 = 0$. In this case, the only difference between the three neighborhoods is the existence of different "fixed amenities," which in our model are captured by the fact that a random fraction of the households always prefer to live in neighborhood A relative to B and in B relative to C. This generates some degree of segregation in the initial steady state that is driven purely by income: richer families would be the only ones willing to pay to live in the better neighborhoods. The blue solid lines in Figure 10 report the response of the baseline model to the skill premium shock in subsection 4.3, while the dotted green lines show the economy's response to the same shock, with $\beta_1 = 0$ and all the other parameters unchanged. The figure shows that both inequality and segregation increase much less when $\beta_1 = 0$. We can interpret the distance between the blue solid lines and the green dotted lines as the contribution of the spillover effect to the increase in inequality. The figure shows that the

[49] For robustness, we also recalibrate the model using smaller values for the lowest and the highest 25th percentile: 5% and 3%, respectively. In this case, the contribution of segregation to inequality is equal to 15% and 13% according to the two counterfactual exercises.

presence of spillover effects contributes significantly to the increase in inequality in the model. This exercise is extreme, as it rules out not only local spillovers, but any type of externality in education returns. In Appendix E, we consider a version of the model where we eliminate a local component of the spillover, but we keep a global externality.[50]

Figure 10: Counterfactuals with No Spillover and No Spillover Feedback



(a) Inequality

(b) Segregation

The second exercise that we consider aims at assessing the contribution to the rise in inequality coming from the feedback effect due to the endogeneity of the local spillovers. To this end, we explore the response of the economy to the same skill premium shock if local spillovers were present but did not change endogenously - that is, if S^A, S^B, and S^C were fixed at their initial steady state levels. The red dashed lines in panels (a) and (b) of Figure 10 show the responses of inequality and segregation, respectively, in this exercise. When we compare them with the results in the baseline model, we can interpret the differential response as the amplification due to the feedback effect coming from the endogeneity of the local externality. The figure shows that the spillover feedback effect contributes to 40% of the increase in inequality. Moreover, the contribution to the increase in segregation is 37%.[51]

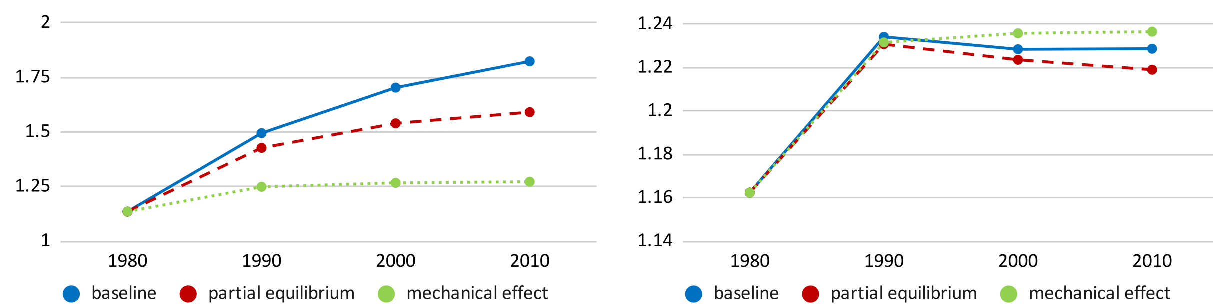
Why does the endogenous change in the spillover effects further amplify inequality? The local spillover effects in the three neighborhoods - S_t^A, S_t^B, and S_t^C - increase in response to the skill premium shock, given that all college educated workers have higher wages, and moreover ev-

[50] Figure 10 is realized without recalibrating the parameters to focus on the decomposition of the response in the model. In the next subsection, we will consider a different exercise. In it, we recalibrate the model so that the difference among neighborhoods is driven purely by amenities and the spillover effects to education returns are global.

[51] Notice that when we compare the response with fixed spillovers with the response in the baseline, we are shutting down not only the endogenous response of the local spillover, but also any form of externality.

erybody invests more in education. However, the strength of the local spillover in neighborhood A increases relatively more than that of neighborhood B, and the strength of the spillover in B increases relatively more than the one in C.[52] Table 3 reports the increase in the spillover ratios in response to the shock, which is illustrated by the solid blue lines in Figure 11.

Figure 11: Decomposing the increase in the spillover ratio



(a) S_A/S_B (b) S_B/S_C

We can decompose the response of S_t^A/S_t^B and S_t^B/S_t^C to the skill-premium shock into three effects. First, there is a mechanical effect: children in richer neighborhoods benefit more from the increase in the skill premium because they are more highly educated and are exposed to stronger spillover effects. This mechanically increases their expected income, and hence the strength of the spillover in the richer neighborhoods. The green dotted lines in the figure show the increase in the spillover ratios due to this mechanical effect - that is, fixing the rental rate and the optimal choices of the parents, but letting the spillover adjust. Second, there is an effect coming from the endogenous response of the optimal educational and residential choice of the parents. As the return to education increases, all parents increase their investment in education, but the richer parents, who are more concentrated in richer neighborhoods, do so even more, increasing the gap in returns. Moreover, parents will tend to move to neighborhoods with higher returns, so families will move from B to A and from C to B. This implies that the ratio S_A/S_B will increase even further, while the ratio S_B/S_C could go in either direction. The dashed red lines in Figure 11 show the increase in the spillover ratios due the sum of the mechanical effect and the effect coming from the endogenous change in educational and residential choices, what we call the "partial equilibrium effect." It shows that the increase in S_A/S_B is amplified, while the ratio

[52] The ratio S_t^B/S_t^C increases on impact in response to the shock and overall between 1980 and 2010, although it is not always monotone.

S_B/S_C starts declining after the first period. Finally, there is a general equilibrium effect coming from the increase in the rental rate in the richer neighborhoods, which increases the degree of sorting by income. Although the more talented children will benefit more from the increase in skill premium, only richer families will be able to pay the higher cost of living in richer neighborhoods, irrespective of their children's ability. This further raises the gap between the spillovers' strengths in the different neighborhoods. We can interpret the difference between the blue solid lines and the red dashed line as the contributions of the general equilibrium effect to the increase in the spillover ratio. The figure shows that the endogenous reallocation of households across neighborhoods due to the general equilibrium effect plays a crucial role in producing a large increase in the spillover ratio, contributing to 34% of the increase in S_t^A/S_t^B and to 15% of the increase in S_t^B/S_t^C between 1980 and 2010.

In Appendix E, we also explore alternative versions of our main model to understand the role of some of our modeling choices. In particular, we explore a version of the model where segregation is driven purely by local amenities and an exercise where we use our main model but we consider a shock to wage dispersion instead of a skill premium shock. For both cases, we show that our main exercise is able to better replicate important features of the data. Next, we investigate a version of the model where we mute the complementarity between ability and local spillovers, and a version of the model where we define the local spillover as the average income of parents living in the neighborhoods instead of the average expected income of the children growing up there. We find that our main results are broadly robust to both these alternative specifications.

6 Scaling Up Moving to Opportunity Policies

In our model, local spillovers generate residential segregation by income that, in turn, reduces intergenerational mobility. Talented children from poor families who grow up in the poorer neighborhoods do not have the same opportunities as children of richer families who can afford a neighborhood with higher spillover and higher amenities. To alleviate these issues, in the mid-90s, the US Department of Housing and Urban Development ran the Moving To Opportunity program (MTO), offering vouchers to low-income people living in high-poverty neighborhoods to move to richer neighborhoods.

Chetty et al. (2016) show that children younger than 13 years old whose families benefited from the vouchers of this program had an income 31% higher than the control group's. Given the success of this program, a natural question is whether we should think about scaling up this experience and run similar programs on a larger scale. Our model is well suited to address this question, given that one natural concern is whether the general equilibrium effects of moving a larger mass of families to better neighborhoods could undermine the benefits of the policy. In particular, as poorer families move to the better neighborhood, the selection of families living in the three neighborhoods endogenously changes, affecting the local spillovers as well as the local rental rates.

6.1 MTO program

The MTO experiment involved 4,604 low-income families living in five cities (Baltimore, Boston, Chicago, Los Angeles, and New York) from 1994 to 1998. The program consisted of offering subsidized housing vouchers to a randomly selected subset of families with children that satified two criteria: they 1) lived in a neighborhood with a poverty rate of 40% or more in 1990 and 2) had an income below 50% of the median income for the metropolitan area. Families were randomly allocated to one of three groups: (i) the experimental group, which could use the voucher only to move to census tracts with 1990 poverty rates below 10%; (ii) the Section 8 group, which could use the voucher without any specific relocation constraint; and (iii) a control group, which received no assistance through MTO. Voucher recipients were required to contribute 30% of their annual household income toward rent and utilities and received housing vouchers that covered the difference between their rent and the family's contribution, up to a maximum amount, defined as the 40th percentile of rental costs in a metro area.

To map the MTO program to our model, we assume that the poorest x families living in the worst neighborhood (neighbohorood C) are offered a voucher to move to a better neighborhood. As in the program, families accepting a voucher are required to pay 30% of their income toward their rent. We focus on the experimental group and require voucher recipients to move to the best neighborhood (neighborhood A). To finance the policy, we assume that the vouchers are covered by a proportional tax on income levied on all families in the city. [53]

[53] In order to characterize the equilibrium, we can naturally extend our equilibrium definition with the presence of the policy. The only difference is that we need to introduce an additional state variable, which is the neighborhood where a parent grew up, or birth neighborhood. Appendix E.5 describes the equilibrium in detail.

6.2 Calibrated MTO Policy

We first introduce in the calibrated model a voucher policy that resembles the actual MTO experiment implemented in 1994-1998. That is, we introduce a voucher policy with $p = 0.001$, to reflect the number of families eligible for a voucher on average across the cities involved in the MTO program. We then compare the predictions of our model with the data on the take-up rates of the eligible families and the income gain for the kids in families taking up the vouchers relative to that of the control group.

First, we look at take-up rates for the MTO experimental group. In our model, about 75% of eligible families in the experimental group choose to accept the voucher. There are two reasons why the model can generate a take-up rate smaller than 100%. First, the rental rate in the worst neighborhood may be lower than the required down-payment to move to the best neighborhood, payment of which is mandatory for the experimental group. In addition, the MTO experiment imposes a cap on the voucher amount, expressed as a fraction of the rental rate in the best neighborhood. Second, our model includes an idiosyncratic preference shock for different neighborhoods, so that a fraction of the eligible families may prefer to remain in their original neighborhood.

In the data, the fraction of families eligible for the voucher in the experimental group taking up the voucher is equal to 48%. Our abstraction from moving costs might explain why the number is a bit smaller than the one in the data. [54]

Second, we examine the income gain for the kids in the families taking up the voucher once they reach adulthood, relative to that of the control group. Chetty et al. (2016) found that relative to the control group, kids in the experimental group who were less than 13 years old when their family received the voucher experienced 31% higher income once they reached their mid-twenties.[55] On the other hand, they find essentially null or slightly negative effects for kids who entered the program when they were past the age of 13.

In our model, we compute the income gain for kids in families that receive the voucher once they become adult, relative to kids in families in the same income percentile in neighborhood C if the

[54] It would be easy to introduce moving costs in the model, but it is hard to discipline those numbers with data.

[55] This corresponds to what Chetty et al. (2016) call TOT (treatment-on-the treated) estimates.

Table 7: Share of Eligible Families with Different Income Cut-Offs

	0.001%	5%	10%	15%	20%	25%
1980	0	0	0	0	0	0
1990	0.0009	0.0385	0.0650	0.1019	0.1242	0.1489
2000	0.0009	0.0309	0.0492	0.0732	0.1062	0.1219
2010	0.0007	0.0341	0.0540	0.0947	0.1104	0.1310
2020	0.0007	0.0258	0.0541	0.0780	0.0932	0.1090

voucher policy were not in place.[56] Given that our model does not distinguish between young and old kids, we compare the income gain from the voucher policy in the model to a simple average income gain of young and old kids in the data, which corresponds to 15.5%. Our model generates an income gain of 13%, which is quite close to the average gain in the data.

6.3 Scaling Up the MTO Program

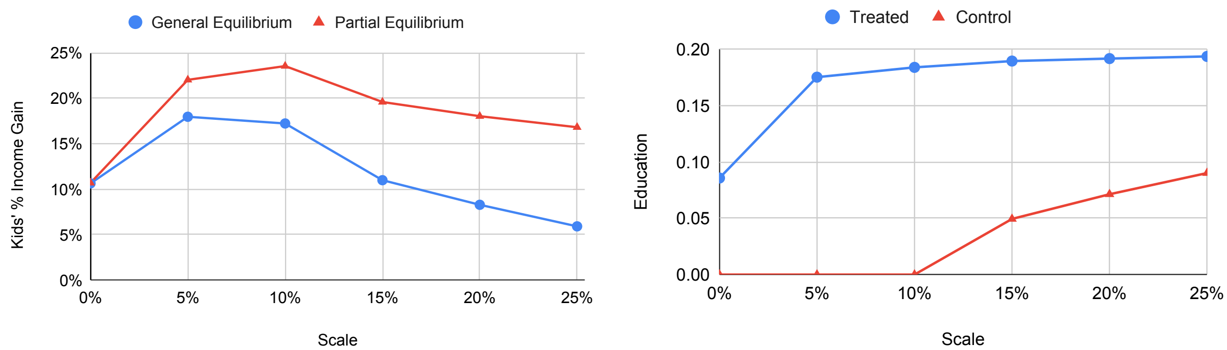
We now use our model to study the effects of scaling up the MTO experiment. To do so, we relax the eligibility requirements for voucher assignment by increasing the income threshold for families living in neighborhood C. In particular, we explore the effects of increasing the percentile of eligible families from the .001th (which is approximated to 0 in the x axis) up to the 25th percentile. Table 7 shows the proportion of families eligible for the voucher over time for different income cut-offs. This also means that as we scale up the program, we change the distribution of ability and income of the families who receive the voucher.

To evaluate the effectiveness of the policy, we first consider the income gain of the children of voucher recipients, who benefit from growing up in a neighborhood with higher spillover. The income gain is calculated as the percentage difference between the expected income of the kids of voucher receivers and the expected income of the same kids if there were no policy in place.

The blue line in panel (a) of Figure 12 shows how the income gain of voucher recipients changes as we scale up the policy. As more families receive the voucher, the average income gain increases at first, but then it starts to decline. To better understand this pattern, we first study the income gain that would arise in partial equilibrium (red line) when both rental rates and spillover

[56]Given that the policy is very small, general equilibrium effects are negligible.

Figure 12: Income Gain and Education for Voucher Recipients' Kids



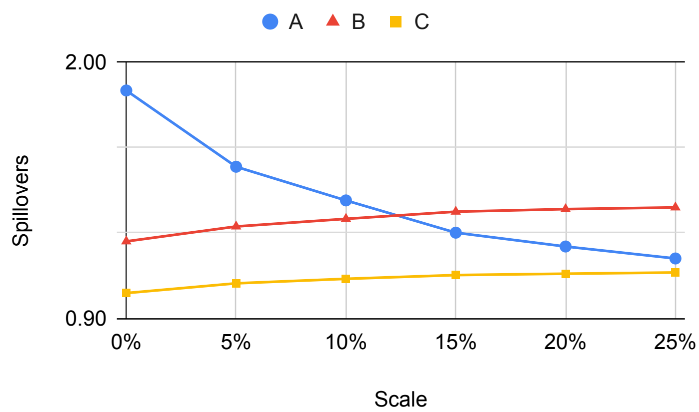
(a) Income gain for voucher recipients' kids (b) Education level for treated vs. control group

sizes are kept constant. The figure shows that the non-monotonicity already arises in partial equilibrium. This is due to different forces working in opposite directions. On the one hand, as the income cut-off of the eligible families increases, richer families will be able to take up the voucher. This also means that families with more talented kids will be able to move to the better neighborhood, because of the correlation between parents' ability and kids' ability. Given that there is complementarity between the local spillover and education, between spillover and parents' wage, and between spillover and ability, the richer families will have a larger income gain from moving to a neighborhood with higher spillover. Moreover, the presence of a cap on the voucher policy contributes to most of the increase in income gain for the initial scale's increase.[57] On the other hand, children of richer families have a smaller gain from the voucher policy because their parents would have invested in education even in the absence of the voucher. This second effect dominates when the scale of the policy becomes large enough, generating the inverted-U shape.

To clarify the education investment channel, panel (b) in Figure 12 compares the education level chosen by the families receiving the voucher when the policy is in place (treated group, in blue) with the education level chosen by the same families if there were no policy (control group, in

[57] When the scale is small, a large fraction of the eligible families are so poor that they end up not taking up the voucher. The reason is that the housing cost is higher than 30% of their income and is beyond the cap imposed by the policy. This means that these families would have to pay additional out-of-pocket expenses to move to A, and a large of fraction of them end up deciding not to take up the voucher, even if they have high ability children. When the scale increases, a larger fraction of families take up the voucher, so more high ability children have a chance to move to A. Given the complementarity between ability and spillover, the average income gain becomes larger.

Figure 13: Spillover in the Three Neighborhoods



red). The figure shows that the families that receive the voucher when the percentile is below or equal to 10% are poor enough that if there were no voucher policy, they would have chosen not to invest in education. This is why when they receive the voucher to move to a neighborhood with high spillover and start investing in education, their children expect a large income gain. When the scale increases above 10%, the families receiving the voucher would have invested in education even with no policy, and so the income gain of their children from the policy gets smaller.

On top of the effects described, as we scale up the policy, there is a general equilibrium effect that further reduces the income gain of the voucher recipients and becomes stronger the larger is the scale. This can be observed in panel (a) of Figure 12 by comparing the blue line, which represents the actual income gain, with the red line, which represents the income gain that would arise in partial equilibrium. The general equilibrium effect dampens the income gain of the children of the voucher recipients because of the endogenous change of spillovers and rental rates in the three neighborhoods.

The first general equilibrium effect is due to the change in spillovers. Figure 13 shows the spillover's response on impact in the three neighborhoods to the introduction of the policy, as a function of its scale. As the scale increases, more poor families move from C to A. This reduces the spillover in neighborhood A, while it increases the spillover of neighborhood C, reducing the advantage of moving from C to A. At the same time, as the spillover in neighborhood A decreases, more families decide to move from neighborhood A to neighborhood B, increasing the spillover in B. As we scale up the policy, such a process gets more pronounced until it leads to a switch in the ranking of the neighborhoods' spillovers. In particular, when the cut-off is equal to or larger than the 15th percentile, the spillover in B becomes larger than the spillover in A, making B the most desirable neighborhood.[58]

The other general equilibrium effect is driven by the change in the rental rates and in the sizes of the three neighborhoods. Panels (a) and (b) in Figure 14 show, respectively, the rental rates and the sizes of the three neighborhoods as a function of the policy's scale. There are two opposite forces that affect the housing demand in neighborhood A. On the on hand, the spillover decline tends to decrease the demand. On the other hand, demand increases because of the voucher recipients moving to A. Overall, the second effect dominates and results in an increase in both the rental rate and the size of neighborhood A. Demand to live in B also increases, because of the increase in families that decide to move from A to B, due to the decline in the spillover gap, while the rental rate in A increases. This increases the rental rate and also slightly increases the size of neighborhood B. The demand to live in neighborhood C decreases because of the voucher recipients who move to A, although this effect is partially offset by the increase in the spillover due to the improvement of the composition of families living there. As a result, both the size and the rental rate in neighborhood C decline. The increase in the rental rate in A and the decrease in rental rate in C further reduce the advantage of moving from C to A, contributing to the dampening of the kids' income gain due to the policy.

In sum, as we scale up the MTO program, the effectiveness of the policy in terms of the adult income gain of the children of voucher recipients at first increases, but eventually it declines as the scale gets larger. This decline is due both to composition effects and to general equilibrium effects.

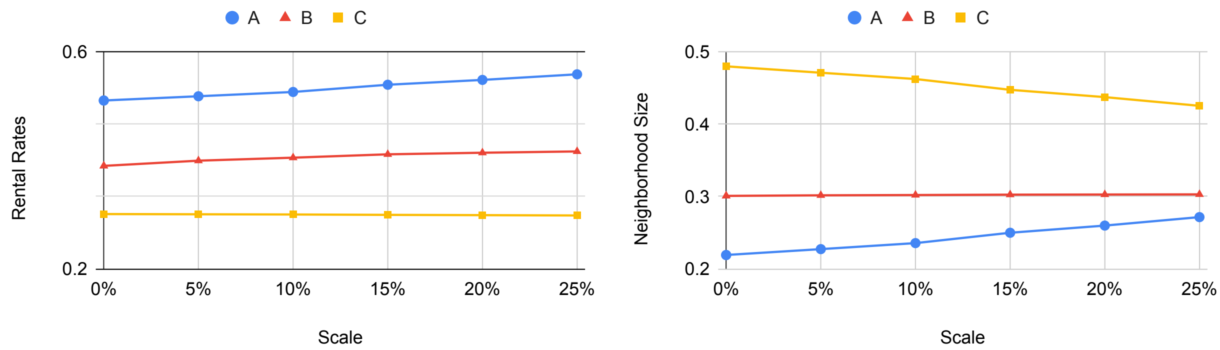
6.4 Aggregate effects

In this section, we look at how the aggregate effects of MTO program change as the scale increases.

Figure 15 shows the level of income inequality, measured by the Gini coefficient (panel a), and of residential segregation, measured by the dissimilarity index (panel b), as a function of the scale

[58] If the scale of the program increases further, the spillover in A further decreases to the point that for some parameters and a large enough scale, it can become even lower than the spillover in C.

Figure 14: Rental Rates in the Three Neighborhoods and Neighborhood Sizes



(a) Rental Rates

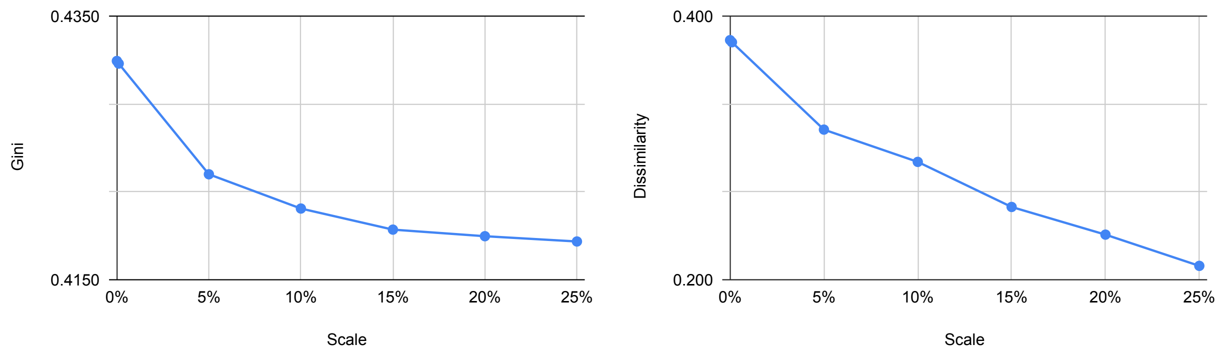
(b) Neighborhood Sizes

of the policy, measured in the period in which the policy is implemented.[59] When the scale of the policy is 0.001, as in the actual MTO experiment, the aggregate effects of the policy are negligible, and both inequality and residential segregation are substantially at the same level they would be at if there were no policy. However, as the scale increases, the policy becomes more successful in reducing both inequality and residential segregation, for two reasons. First, inequality decreases because voucher recipients, who are on the low end of the income distribution, move to the best neighborhood, and thus their kids are exposed to the highest spillover and have higher adult income. Second, at the same time, the kids of the richer families living in A are exposed to a slightly lower spillover because of the general equilibrium effect. This reallocation of individuals across the city also reduces residential segregation, as poorer families move to neighborhood A, making it more mixed. At the same time, neighborhood C becomes less poor because the poorer families move out.

Finally, Figure 16 shows welfare, measured as average expected utility, in the period in which the policy is implemented, as a function of the policy's scale, both in partial equilibrium (red dashed line) and in general equilibrium (blue solid line). In particular, we plot the difference in welfare relative to the baseline case with no policy. It is interesting to note, that while inequality and segregation improve monotonically with the scale, the effect on welfare is non-monotonic. More specifically, the figure shows that the non-monotonicity is due to the general equilibrium effect.

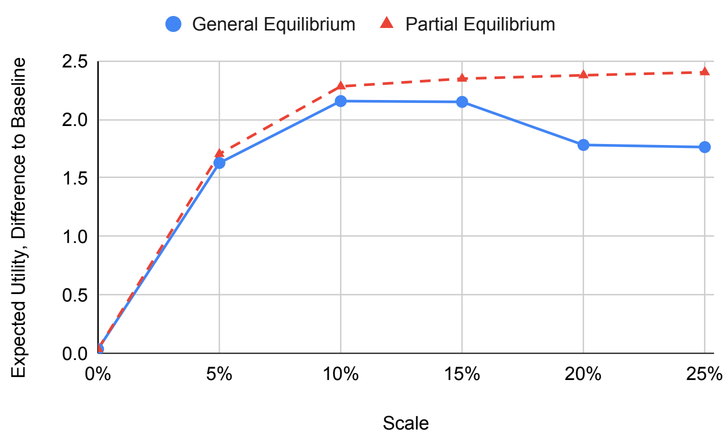
[59]We measure the Gini coefficient on the distribution of expected adult income of kids growing up in the period in which the policy implemented, who are the first generation being affected by the policy.

Figure 15: Income Inequality and Segregation



(a) Inequality (b) Segregation

Figure 16: Welfare



As the scale increases, there are more families who benefit from the voucher, and their expected utility increases. However, in general equilibrium, the benefit of moving to neighborhood A declines with the scale increases as the spillover in A declines. In addition, more families end up paying the highest rent and fewer families end up paying the lowest, so the total amount of resources that ends up in the pocket of house owners (who are outside of the model) increases, reducing average welfare.

7 Concluding Remarks

In this paper, we propose a model where segregation and inequality amplify each other because of a local spillover that affects the returns to education. After calibrating the model using US data and using the micro estimates for neighborhood exposure effects in Chetty and Hendren (2018b), we look at the response of the economy to an unexpected permanent shock to the skill premium. We find that local spillovers and the resulting residential segregation across neighborhoods play a significant role in amplifying the rise in aggregate income inequality and the decline in intergenerational mobility. Differences between neighborhoods in terms of educational opportunities grow over time, and children from poor families do not get a chance to live the American dream. We also use the model to explore the effects of scaling up a voucher policy in the spirit of the MTO program and show that as the scale increases, general equilibrium effects dampen the effectiveness of the policy.

There are a number of interesting directions for future research. One is to explore the cross-sectional implications of our model and the heterogeneity of experiences across US metro areas. Another is to think about the link between residential segregation by income and residential segregation by race.[60] Moreover, the decreasing effectiveness of MTO at larger scales opens the question of whether other types of policies, such as place-based policies, might be more effective. Finally, it would be interesting to endogenize the cost of education or the creation of new neighborhoods, both of which could further magnify the increase in income inequality.

References

Aaronson, Daniel and Bhashkar Mazumder, "Intergenerational Economic Mobility in the United States, 1940 to 2000," *Journal of Human Resources*, 2008, *43* (1), 139–172.

Acemoglu, Daron, "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, 2002, *40* (1), 7–72.

Agostinelli, Francesco, *Investing in Children's Skills: An Equilibrium Analysis of Social Interactions and Parental Investments*, University of Pennsylvania, mimeo, 2018.

__ , Matthias Doepke, Giuseppe Sorrenti, and Fabrizio Zilibotti, "It Takes a Village: the Economics of Parenting with Neighborhood and Peer Effects," 2020.

[60] In Fogli, Garcia-Vasquez, Guerrieri and Prato (2023), we explore this relationship and we endogenize racial discrimination by analyzing its geographical dimension.

—, —, —, and —, "When the Great Equalizer Shuts Down: Schools, Peers, and Parents in Pandemic Times," *Journal of Public economics*, 2022, *206*, 104574.

—, Paolo Martellini, and Margaux Luflade, *On the Spatial Determinants of Educational Access*, University of Chicago.

Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso, "Intergenerational Mobility and Preferences for Redistribution," *American Economic Review*, 2018, *108* (2), 521–554.

Armour, Philip, Richard V. Burkhauser, and Jeff Larrimore, "Using the Pareto Distribution to Improve Estimates of Topcoded Earnings," *Economic Inquiry*, 2016, *54* (2), 1263–1273.

Autor, David, Claudia Goldin, and Lawrence Katz, "Extending the Race between Education and Technology," *American Economic Association: Papers and Proceedings*, 2020, *110*, 347–351.

Autor, David H., Lawrence F. Katz, and Alan B Krueger, "Computing Inequality: Have Computers Changed the Labor Market?," *Quarterly Journal of Economics*, 1998, *113* (4), 1169–1213.

—, —, and Melissa S. Kearney, "Trends in US Wage Inequality: Revising the Revisionists," *Review of Economics and Statistics*, 2008, *90* (2), 300–323.

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

Benabou, Roland, "Workings of a City: Location, Education, and Production," *Quarterly Journal of Economics*, 1993, *108* (3), 619–652.

—, "Equity and Efficiency in Human Capital Investment: The Local Connection," *Review of Economic Studies*, 1996, *63* (2), 237–264.

—, "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *The American Economic Review*, 1996, *86* (3), 584–609.

Bilal, Adrien and Esteban Rossi-Hansberg, "Location as an Asset," *Econometrica*, 2021, *89* (5), 2459–2495.

Brueckner, Jan K, Jacques-Francois Thisse, and Yves Zenou, "Why is Central Paris Rich and Downtown Detroit Poor? An Amenity-Based Theory," *European Economic Review*, 1999, *43* (1), 91–107.

Bryan, Gharad, Edward Glaeser, and Nick Tsivanidis, "Cities in the Developing World," *Annual Review of Economics*, 2020, *12*, 273–297.

Cai, Zhifeng and Jonathan Heathcote, "College Tuition and Income Inequality," *American Economic Review*, 2022, *112* (1), 81–121.

Card, David and Laura Giuliano, "Can Tracking Raise the test Scores of High-Ability Minority Students?," *American Economic Review*, 2016, *106* (10), 2783–2816.

— and Thomas Lemieux, "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, 2001, *116* (2), 705–746.

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

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

— **,** — **, and Lawrence F Katz**, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 2016, *106* (4), 855–902.

— **,** — **, Patrick Kline, and Emmanuel Saez**, "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Chyn, Eric and Diego Daruich, *An Equilibrium Analysis of the Effects of Neighborhood-Based Interventions on Children*, NBER Working Paper 29927, 2022.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst, *Income Growth and the Distributional Effects of Urban Spatial Sorting*, University of Chicago, mimeo, 2019.

Cunha, Flavio, James J Heckman, and Susanne M Schennach, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 2010, *78* (3), 883–931.

de Bartolome, Charles A.M., "Equilibrium and Inefficiency in a Community Model with Peer Group Effects," *Journal of Political Economy*, 1990, *98* (1), 110–133.

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

— **and Cecile Gaubert**, "Spatial Sorting and Inequality," *Annual Review of Economics*, 2022, *14*, 795–819.

Doepke, Matthias and Fabrizio Zilibotti, "Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission," *Econometrica*, 2017, *85* (5), 1331–1371.

— **, Giuseppe Sorrenti, and Fabrizio Zilibotti**, "The Economics of Parenting," *Annual Review of Economics*, 2019, *11*, 55–84.

Duncan, Greg J and Richard J Murnane, "Rising Inequality in Family Incomes and Children's Educational Outcomes," *RSF*, 2016, *2* (2), 142–158.

Durlauf, Steven N, "Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality," in G. Gandolfo W. Barnett and C. Hillinger, eds., *G. Gandolfo W. Barnett and C. Hillinger, eds.*, Dynamic Disequilibrium Modelling: Proceedings of the Ninth International Symposium on Economic Theory and Econometrics, Cambridge University Press, 1996, pp. 505–534.

— , "A Theory of Persistent Income Inequality," *Journal of Economic Growth*, 1996, *1* (1), 75–93.

— **and Ananth Seshadri**, *Understanding the Great Gatsby Curve*, NBER Macroeconomics Annual, 2017.

— , **Andros Kourtellos, and Chih Ming Tan**, "The Great Gatsby Curve," *Annual Review of Economics*, 2022, *14*, 571–605.

Eckert, Fabian and Tatjana Kleineberg, *Saving the American Dream? Education Policies in Spatial General Equilibrium*, Opportunity and Inclusive Growth Working Paper 47, Federal Reserve Bank of Minneapolis, 2021.

Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny, "Spatial Sorting," *Journal of Political Economy*, 2014, *122* (3), 554–620.

Fernandez, Raquel and Richard Rogerson, "Income Distribution, Communities, and the Quality of Public Education," *Quarterly Journal of Economics*, 1996, *111* (1), 135–164.

— **and** — , "Keeping People Out: Income Distribution, Zoning, and the Quality of Public Education," *International Economic Review*, 1997, *38* (1), 23–42.

— **and** — , "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform," *American Economic Review*, 1998, *88* (4), 813–833.

Ferreira, Pedro C, Alexander Monge-Naranjo, and Luciene Torres de Mello Pereira, *Of Cities and Slums*, Working Paper 2016-0224, Federal Reserve Bank of St. Louis, 2017.

Fogli, Alessandra, Martin Garcia-Vasquez, Veronica Guerrieri, and Marta Prato, "Neighborhood Segregation and Endogenous Racial Bias," 2023.

— **, Veronica Guerrieri, Mark Ponder, and Marta Prato**, "Scaling up the American Dream: a Dynamic Analysis of Moving to Opportunity Policies," 2022.

Giannone, Elisa, *Skill-Biased Technical Change and Regional Convergence*, Pennsylvania State University, mimeo, 2018.

Glaeser, Edward L, Jed Kolko, and Albert Saiz, "Consumer City," *Journal of Economic Geography*, 2001, *1* (1), 27–50.

Goldin, Claudia and Lawrence F. Katz, *Decreasing (and then Increasing) Inequality in America: A Tale of Two Half-Centuries*, University of Chicago Press,

Goldin, Claudia Dale and Lawrence F Katz, *Race between Education and Technology*, Harvard University Press, 2009.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst, "Endogenous Gentrification and Housing Price Dynamics," *Journal of Public Economics*, 2013, *100*, 45–60.

Hassler, John and Jose V Rodriguez Mora, "Intelligence, Social Mobility, and Growth," *American Economic Review*, 2000, *90* (4), 888–908.

Hsieh, Chang-Tai and Enrico Moretti, *Why do Cities Matter? Local Growth and Aggregate Growth*, Kreisman Working Papers Series in Housing Law and Policy 30, 2015.

Imberman, Scott A, Adriana D Kugler, and Bruce I Sacerdote, "Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees," *American Economic Review*, 2012, *102* (5), 2048–82.

Jargowsky, Paul A., "Take the Money and Run: Economic Segregation in U.S. Metropolitan Areas," *American Sociological Review*, 1996, *61* (6), 984–998.

Katz, Lawrence F and Kevin M Murphy, "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 1992, *107* (1), 35–78.

Kulkarni, Nirupama and Ulrike Malmendier, "Homeownership Segregation," *Journal of Monetary Economics*, 2022, *129*, 123–149.

Lavy, Victor, M Daniele Paserman, and Analia Schlosser, "Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom," *Economic Journal*, 2012, *122* (559), 208–237.

Loury, Glenn C, "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 1981, *49* (4), 843–867.

Massey, Douglas S, Jonathan Rothwell, and Thurston Domina, "The Changing Bases of Segregation in the United States," *Annals of the American Academy of Political and Social Science*, 2009, *626* (1), 74–90.

Montgomery, James D, "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis," *American Economic Review*, 1991, *81* (5), 1408–1418.

— , *Social Networks and Persistent Inequality in the Labor Market*, Northwestern University, unpublished manuscript, 1991.

Moretti, Enrico, "Human Capital Externalities in Cities," in J. Vernon Henderson and Jacques-FranÃ§ois Thisse, eds., *Cities and Geography*, Vol. 4 of *Handbook of Regional and Urban Economics*, Elsevier, 2004, pp. 2243–2291.

— , *The New Geography of Jobs*, Houghton Mifflin Harcourt, 2012.

Nielsen, Francois and Arthur S. Alderson, "The Kuznets Curve and the Great U-Turn: Income Inequality in U.S. Counties, 1970 to 1990," *American Sociological Review*, 1997, *62* (1), 12–33.

Reardon, Sean F and Kendra Bischoff, "Income Inequality and Income Segregation," *American Journal of Sociology*, 2011, *116* (4), 1092–1153.

— , — , Ann Owens, and Joseph B. Townsend, "Has Income Segregation Really Increased? Bias and bias Correction in Sample-Based Segregation Estimates," *Demography*, 2018, *55* (6), 2129–2160.

Rothstein, Jesse, "Inequality of Educational Opportunity? Schools as Mediators of the Intergenerational Transmission of Income," *Journal of Labor Economics*, 2019, *37* (S1), S85–S123.

Sacerdote, Bruce, "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *The Quarterly Journal of Economics*, 2001, *116* (2), 681–704.

— , "Peer Effects in Education: How Might they Work, How Big are They and How Much do We Know Thus Far?," *Handbook of the Economics of Education*, 2011, *3*, 249–277.

Shapiro, Jesse M, "Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital," *The Review of Economics and Statistics*, 2006, *88* (2), 324–335.

Streufert, Peter, "The Effect of Underclass Social Isolation on Schooling Choice," *Journal of Public Economic Theory*, 2000, *2* (4), 461–482.

von Hippel, Paul T., David J. Hunter, and McKalie Drown, "Better Estimates from Binned Income Data: Interpolated CDFs and Mean-Matching," *Working paper*, 2017.

Watson, Tara, "Inequality and the Measurement of Residential Segregation by Income in American Neighborhoods," *NBER Working Paper 14908*, 2009.

Zheng, Angela and James Graham, "Public Education Inequality and Intergenerational Mobility," *American Economic Journal: Macroeconomics*, 2022, *14* (3), 250–282.

Appendix (For Online Publication)

A Data Methodology

A.1 Segregation and Inequality over Time

Data sources and sample selection. We use tract-level income data from decennial censuses (1980 to 2000) and from the American Community Surveys (ACS) for the 5 year period spanning 2008-2012. Our sample includes metropolitan areas using the 2003 OMB definition. Table 8 reports the sample size, in terms of number of MSAs, census tracts, and all families. Census tracts are small, relatively permanent statistical subdivisions of a county and are designed to have an optimum size of 4,000 people. Census tracts are merged or added over time to keep population size constant. The number of census tracts has increased over time, reflecting the increase in the population.

Table 8: All Families: Summary Statistics

Year	MSAs	Census Tracts	All Families
1980	379	42,406	46,154,644
1990	380	48,412	52,853,972
2000	380	53,033	59,087,771
2010	380	59,842	63,325,283

In our calibration, we restrict the sample to families with children. The data on families with and without children at the census tract level are available for 2000 and 2010, but not for 1980 and 1990. The data in 2000 and 2010 are available for six groups at each income bracket level: 1) married couple family with own children under 18 years; 2) male householder (no wife present) with own children under 18 years; 3) female householder (no husband present) with own children under 18 years; 4) married couple family without own children under 18 years; 5) male householder (no wife present) without own children under 18 years; and 6) female householder (no husband present) without own children under 18 years. We calculate the number of families with children at the census tract-income bracket level as the sum of (1), (2), and (3). Families without children at the census tract-income bracket level are calculated as the sum of (4), (5) and (6).

For 1980 and 1990, these data are not directly available at the census tract level. However, the

Table 9: Families with Children: Summary Statistics

Year	MSAs	Census Tracts	Families with Kids
1980	347	41,246	23,325,537
1990	373	47,184	24,922,747
2000	380	53,033	29,209,867
2010	380	59,842	29,155,384

data are available for those years at a compound geographic level. In particular, from IPUMS NHGIS we use "Census Tract/Block Numbering Area (by State–Standard Metropolitan Statistical Area - County–Place)" for 1980 and "Census Tract/Block Numbering Area (by State–County–Metropolitan Statistical Area/Consolidated Metropolitan Statistical Area/Remainder–Primary Metropolitan Statistical Area/Remainder)" for 1990.

The data are available for the following 9 groups at each income bracket level:1) married couple family with own children under 6 years old; 2) married couple family with own children between 6 and 17 years old; 3) male householder (no wife present) with own children under 6 years old; 4) male householder (no wife present) with own children between 6 and 17 years old; 5) female householder (no husband present) with own children under 6 years old; 6) female householder (no husband present) with own children between 6 and 17 years old; 7) married couple family without own children; 8) male householder (no wife present) without own children; and 9) female householder (no husband present) without own children.

We extract the state, county, and census tract codes from the unique GISJOIN identifier. The unique GISJOIN identifier has information on state, county, census tract, and block code. Since we use compound geographic levels, there are multiple observations for census tracts that lie along multiple county subdivisions. For 1980, we have 47,974 observations with 41,246 unique census tracts. For 1990, we have 47,271 observations with 47,184 unique census tracts. We aggregate the counts of (1)-(9) at the census tract level using the extracted census tract codes. The data for families with children at the census tract-income bracket level is calculated by summing up (1)-(6). Families without children at the census tract-income bracket level are calculated by summing up (7)-(9).

The data on non-family households and all households are available at the census tract level for 1990. This means that we do not have to use the compound geographic level information for the two series, and we can check if we have the correct numbers for families with and without children at the census tract level using the census tract codes extracted from the unique GIS identifier. We find that the sum of families with children, families without children, and non-family households is equal to the number of total households at the census tract level, as shown in Table 10 below.

Table 10: Sample Size in 1990

Year	Metro	Counties	Census Tracts	Families		Non-Family Households	All Households
				with children	without children		
1990	373	870	47,184	24,922,747	26,684,647	22,122,914	73,730,308

Computing the Dissimilarity Index. The dissimilarity index uses the following formula:

$$D(j) = \frac{1}{2}\sum_{i}\left|\frac{x_i(j)}{X(j)} - \frac{y_i(j)}{Y(j)}\right|,$$

where $X(j)$ and $Y(j)$ respectively denote the total number of poor and rich families in metro j, while $x_i(j)$ and $y_i(j)$ respectively denote the number of poor and rich families in census tract i in metro j. To use this formula, we must define poor and rich families within an MSA. To this end, we rank family income buckets from lowest to highest and calculate the cumulative population across buckets. We then find the bucket with a cumulative share closest to our cut-off percentile (we calculated the dissimilarity index using the 50th, 80th, and 90th percentiles). All families with an income greater than the cut-off bucket are labeled "rich," and all families with a lower income are labeled "poor." This definition is then applied to all census tracts within the relevant MSA. The dissimilarity index is then calculated for each MSA, and the results are aggregated to the national level using metro level population weights.

Computing the Gini The Gini coefficients in this paper are calculated following the method of von Hippel et al. (2017). First, a non-parametric estimation of the income CDF is calculated for each metropolitan area. The non-parameteric CDF is calculated using the function binsmooth, provided by von Hippel et al. (2017). This function linearly interpolates between the upper bounds of each income bracket to calculate the CDF, preserving the empirical cumulative distribution for each bin. It then uses the empirical mean income to calculate the implied upper bound for the support of the PDF, choosing the upper bound and the scale parameter so that the mean of the estimated CDF matches the empirical mean. Three methods are proposed to characterize the distribution of the top bracket: linear, Pareto, and exponential. The default method is linear and is what is used here. The binsmooth function returns a non-parameteric CDF function, which can be used to calculate the Gini coefficient (and the conditional mean income of the top-coded bracket). Define

$$\mu = \int x f(x) dx.$$

Then, the Gini coefficient is calculated as

$$G = 1 - \frac{1}{\mu} \int_0^E (1 - F(x))^2 dx.$$

These integrals must be calculated numerically; however, because the CDF is piecewise linear, the approximation error is small. Importantly, the μ from the non-parametric CDF matches the empirical mean. We calculate the Gini coefficient for each MSA and then take the weighted average using metro level population weights to aggregate at the national level.

Example: Segregation in Chicago over Time. The dissimilarity index captures the deviation from an even distribution of rich and poor families. Given that we define as rich the families in the top 20th percent of the metro distribution, the index is equivalent, up to a constant, to a weighted sum of the deviations of the share of rich families in all census tracts from 20%, with weights given by the census tract population relative to the metro population. Figure 2 in the main text plots the share of rich families in each census tract of Chicago in 1980 and 2010.[61] If there were no segregation, each census tract would have the same share of rich families equal to 20%. To visualize how segregation has changed over the period, we use a heat map. We use orange to identify census tracts with a share of rich families higher than 30% (which correspond to neighborhood A in our calibrated model), dark blue for the census tracts with a share of rich families below 17% (which correspond to neighborhood C), and light blue for census tracts with a share of rich between 17% and 30% (which correspond to neighborhood B). We observe that over time, the number of tracts with either high or low concentration of rich families increases at the expense of tracts with an intermediate fraction of rich families. To do these figures, we

[61] To construct this figure, we use the sample of all families.

keep the geography constant over time, and map the 2010 distribution using the 1980 geographic borders. As we explained above, the census tracts are comparable in terms of population but not necessarily in terms of geographic area (larger in suburban and less densely populated areas).[62]

A.2 Inequality and Segregation: Robustness

Figure 3 in the main text shows that the increase in spatial segregation by income across neighborhoods happened at the same time as the increase in income inequality.

We now check the robustness of these patterns, using alternative measures of income segregation and income inequality.

Figure 17: Dissimilarity Index: Different Cut-Offs

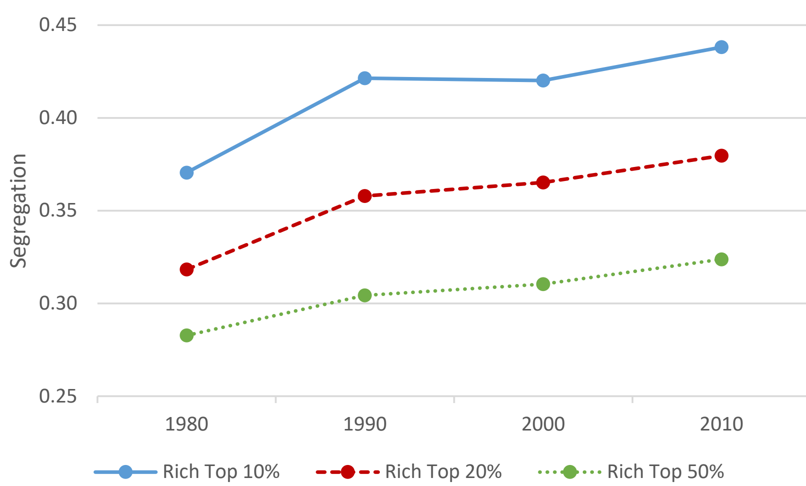


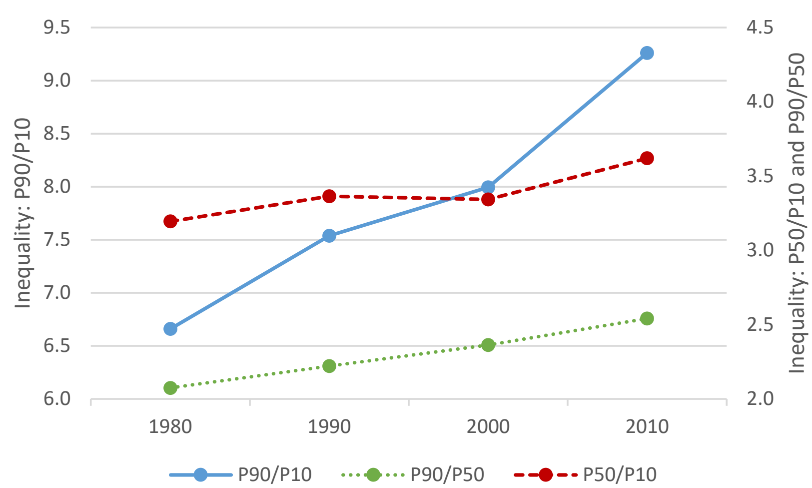
Figure 17 plots the dissimilarity index calculated using different percentiles to define the income groups. The red dashed line shows our benchmark dissimilarity index, while the solid blue line and the dotted green line show the dissimilarity index constructed using the $10th$ and the $50th$ percentiles, respectively. The figure shows that the dissimilarity index shifts up as the cut-off

[62]Note that this measure is different from the one proposed by Reardon and Bischoff (2011), since it is not affected by changes in inequality. Their measure defines the rich in terms of distance from the median income in the metro. With the recent, large increase in inequality, the share of people at the tails of the income distribution has increased even without changes in segregation. The measure we propose is not affected by this issue, as it keeps constant the percentile to define the rich group.

percentile decreases, given that as groups become progressively more homogeneous with respect to income, they are also characterized by higher levels of segregation. However, regardless of the level, all measures show an increasing trend over time.

The increase in inequality is also a robust finding. Figure 18 plots other three measures of income inequality that have been widely used in the literature: the 90/10 ratio that measures the ratio of family income in the top 90th percentile of the population relative to that in the bottom 10th percentile, and, similarly, the 50/10 ratio, and the 90/50 ratio.[63] Figure 18 shows that both the 90/10 and the 90/50 ratios have increased steadily since 1980, while the 50/10 ratio is flat or even slightly decreasing after 1990. This confirms that the rise in income inequality has been driven by the top of the distribution, as already shown by Autor et al. (2008) for individual wage inequality.

Figure 18: Inequality: Different Measures



A.3 Regression Analysis

To explore the relationship between segregation and inequality, we run several regressions. First, we regress the Gini coefficient on the dissimilarity index at the MSA level.[64] Table 11 shows

[63]The procedure implemented to calculate these ratios from binned data at the census tract level is described in Appendix A.1.

[64]For this analysis, we define the cut-off between rich and poor as the 80th percentile. We use population weighting in the regressions, although the results do not change significantly if the observations are unweighted.

the results both without and with controls for racial and industrial composition. Racial shares are reported at the MSA level and are from the decennial census and the American Community Survey. Industry employment shares are from the Quarterly Census of Employment and Wages, provided by the Bureau of Labor Statistics. The results are largely similar to the regressions without controls.

Table 11: *Regression Analysis: Levels*

	Dependent variable: Gini_{1980}			
	(1)	(2)	(3)	(4)
$\text{Dissimilarity}_{1980}$	0.250***	0.107***	0.171***	0.077***
	(0.015)	(0.016)	(0.017)	(0.017)
Constant	0.300***	0.562***	0.151**	0.346***
	(0.005)	(0.023)	(0.063)	(0.059)
Race_{1980}	No	Yes	No	Yes
Industry_{1980}	No	No	Yes	Yes
Observations	379	379	379	379
R^2	0.421	0.632	0.607	0.724
Adjusted R^2	0.420	0.627	0.596	0.713
Residual Std. Error	9.637 (df = 377)	7.725 (df = 373)	8.044 (df = 367)	6.781 (df = 363)
F Statistic	274.303*** (df = 1; 377)	128.136*** (df = 5; 373)	51.629*** (df = 11; 367)	63.505*** (df = 15; 363)

Note: *p<0.1; **p<0.05; ***p<0.01 Regressions use population weights

Table 12 shows the results of regressing changes in Gini coefficient between 1980 and 2010 on changes in dissimilarity index in the same period at the MSA level. We report the results both without and with controls for changes in the racial and industrial composition in the same period. Table 15 shows the crosswalk we used to construct consistent time series for the industrial composition for the MSAs. We also run regressions controlling for the initial level of racial and industrial composition, and the results are robust to this change. Tables 13 and 14 report the summary statistics for the different variables in both regressions.

We also run the same regressions, restricting the sample to families with kids, and find larger coefficients. In particular, the coefficient for the level regression is 0.33 (0.12 with both controls), and the coefficient for the regression in changes is 0.24 (0.20 with both controls).

A.4 School District Analysis

In our model, local spillovers are broadly defined to include many channels. However, school quality is an important one, which makes it interesting to explore the evolution of residential segregation by income at the school district level.

Table 12: *Regression Analysis: Changes*

	Dependent variable: $\Delta\text{Gini}_{2010-1980}$			
	(1)	(2)	(3)	(4)
$\Delta\text{Dissimilarity}_{2010-1980}$	0.176***	0.170***	0.158***	0.154***
	(0.017)	(0.018)	(0.016)	(0.016)
Constant	0.054***	0.054***	0.031***	0.031***
	(0.001)	(0.002)	(0.003)	(0.004)
$\Delta\text{Race}_{2010-1980}$	No	Yes	No	Yes
$\Delta\text{Industry}_{2010-1980}$	No	No	Yes	Yes
Observations	379	379	379	379
R^2	0.212	0.231	0.402	0.420
Adjusted R^2	0.209	0.221	0.384	0.396
Residual Std. Error	6.183 (df = 377)	6.139 (df = 373)	5.458 (df = 367)	5.407 (df = 363)
F Statistic	101.158*** (df = 1; 377)	22.431*** (df = 5; 373)	22.424*** (df = 11; 367)	17.493*** (df = 15; 363)

Note: *p<0.1; **p<0.05; ***p<0.01 Regressions use population weights

Table 13: *Regression in Levels: Summary Statistics*

Variable	Mean	Min	Max	Std
$\text{Dissimilarity}_{1980}$	0.27	0.01	0.50	0.07
Gini_{1980}	0.37	0.32	0.46	0.02
Pct. White_{1980}	0.86	0.33	0.99	0.11
Pct. Black_{1980}	0.10	0.00	0.44	0.10
Pct. Indian_{1980}	0.01	0.00	0.33	0.02
Pct. Asian_{1980}	0.01	0.00	0.60	0.03
Pct. Other_{1980}	0.03	0.00	0.38	0.05
Pct. Agriculture, Forestry, and Fishing_{1980}	0.02	0.00	0.41	0.05
Pct. $\text{Construction}_{1980}$	0.05	0.02	0.17	0.02
Pct. Finance, Insurance, and Real Estate_{1980}	0.05	0.02	0.19	0.02
Pct. $\text{Manufacturing}_{1980}$	0.24	0.00	0.60	0.12
Pct. Mining_{1980}	0.01	0.00	0.25	0.03
Pct. Nonclassifiable $\text{Establishments}_{1980}$	0.00	0.00	0.01	0.00
Pct. Public $\text{Administration}_{1980}$	0.06	0.00	0.49	0.06
Pct. Retail Trade_{1980}	0.19	0.11	0.35	0.03
Pct. Services_{1980}	0.27	0.11	0.53	0.06
Pct. Transportation and Public Utilities_{1980}	0.06	0.03	0.16	0.02
Pct. Wholesale Trade_{1980}	0.05	0.00	0.12	0.02

Dissimilarity Index at the School District Level. The National Center for Education Statistics (NCES) collaborates with the US Census Bureau to provide demographic data for school districts. The data are provided from the 1990 and 2000 decennial census, as well as the 2008-

Table 14: *Regression in Changes: Summary Statistics*

Variable	Mean	Min	Max	Std
ΔDissimilarity$_{2010-1980}$	0.05	-0.12	0.24	0.05
ΔGini$_{2010-1980}$	0.06	-0.01	0.10	0.02
ΔPct. White$_{2010-1980}$	-0.07	-0.27	0.22	0.05
ΔPct. Black$_{2010-1980}$	0.01	-0.10	0.16	0.02
ΔPct. Indian$_{2010-1980}$	0.00	-0.02	0.04	0.00
ΔPct. Asian$_{2010-1980}$	0.02	-0.06	0.24	0.02
ΔPct. Other$_{2010-1980}$	0.03	-0.22	0.19	0.03
ΔPct. Agriculture, Forestry, and Fishing$_{2010-1980}$	-0.01	-0.41	0.07	0.04
ΔPct. Construction$_{2010-1980}$	-0.01	-0.12	0.04	0.02
ΔPct. Finance, Insurance, and Real Estate$_{2010-1980}$	0.01	-0.11	0.08	0.02
ΔPct. Manufacturing$_{2010-1980}$	-0.12	-0.34	0.11	0.07
ΔPct. Mining$_{2010-1980}$	-0.01	-0.11	0.03	0.01
ΔPct. Nonclassifiable Establishments$_{2010-1980}$	0.00	0.00	0.00	0.00
ΔPct. Public Administration$_{2010-1980}$	0.00	-0.31	0.26	0.04
ΔPct. Retail Trade$_{2010-1980}$	-0.05	-0.17	0.06	0.03
ΔPct. Services$_{2010-1980}$	0.22	0.00	0.38	0.06
ΔPct. Transportation and Public Utilities$_{2010-1980}$	-0.02	-0.10	0.05	0.02
ΔPct. Wholesale Trade$_{2010-1980}$	-0.01	-0.07	0.05	0.02

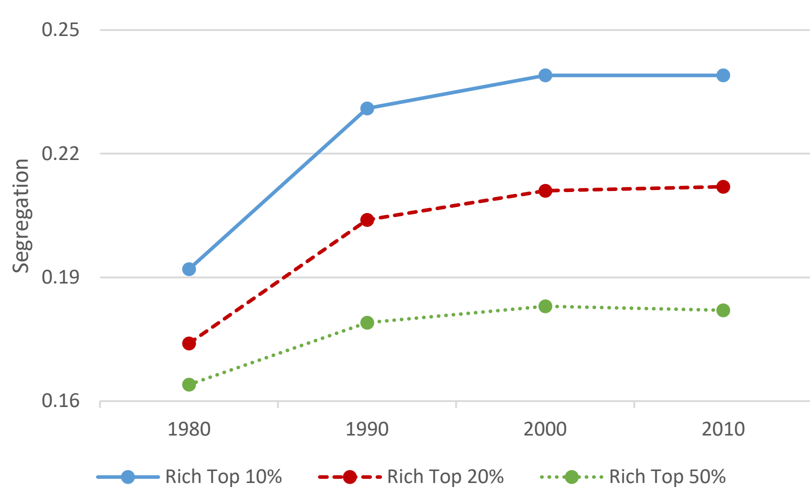
2012 American Community Survey. The data for 1980 are taken from the Census of Population and Housing, Summary Tape 3F and are provided by ICPSR. After combining these files, we calculate the dissimilarity index for all families using school districts as the relevant sub-unit.[65] Figure 19 shows the results of these calculations. The first thing to note is that the overall trend is almost identical to what we get with census tracts. The main difference is that there is a greater increase in dissimilarity from 1990 to 2000 at the school district level and less of an increase from 2000 to 2010. One possible explanation for this trend is the increase in the attendance of private school, which has taken place precisely in the last twenty years and mostly on the East Coast, where there are some of the most populated metros in the US (which have larger weight in our estimates). The increase in the share of children attending private schools weakens the incentive to segregate across school district lines.

[65] Income data for families with children are not available.

Table 15: Industry Crosswalk

NAICS Industry Classification	SIC Industry Classification
Real estate and rental and leasing Finance and insurance	Finance, Insurance, and Real Estate Division
Transportation and warehousing Utilities	Transportation and Public Utilities Division
Educational services Accommodation and food services Administrative and waste services Other services, except public administration Arts, entertainment, and recreation Professional and technical services Management of companies and enterprises Information Health care and social assistance	Services Division
Agriculture, forestry, fishing and hunting	Agriculture, Forestry, and Fishing Division
Construction	Construction Division
Manufacturing	Manufacturing Division
Mining, quarrying, and oil and gas extraction	Mining Division
Unclassified	Nonclassifiable Establishments
Public administration	Public Administration Division
Retail trade	Retail Trade Division
Wholesale trade	Wholesale Trade Division

Figure 19: Inequality and Segregation over Time



Census Tracts vs. School Districts. Census tracts have several advantages over school districts as our unit of analysis. Census tracts are determined by the Census Bureau and are largely fixed over time. When initially determined, the Census aims to include roughly 4,000 people per tract and attempts to define the tract over a homogeneous population. Further, boundaries for census tracts generally follow local government boundaries, such as state, MSA, and county borders, allowing for a clean mapping between sub-units and metros.

In contrast, school districts are locally administered, and their geographic structure can vary by region. Like those of census tracts, the definitions of school districts are relatively stable over time. However, many states have seen a significant consolidation in school districts over time. School districts follow state boundaries but not necessarily MSA lines, complicating our ability to cleanly map sub-units to metros. The degree to which school districts coincide with government boundaries differs across the nation. For instance, on the East Coast, school districts tend to coincide with counties, townships, or city boundaries while in the Midwest, they are almost entirely, independent of municipal boundaries. Finally, the dissimilarity index can be misleading when there are not enough sub-units available. For example, consider an MSA that has a single school district. The dissimilarity index would necessarily be 0 in this case, since the population at the district level necessarily coincides with the population at the metro level. This result may potentially hide significant income segregation within the MSA. The literature has noted that over the past three decades, segregation has increased both between school districts and between schools. Using census tracts will reflect these changes, whereas using school districts would mask the latter trend.

Table 16 reports summary statistics at the district level. The average number of districts in a metro is much smaller that the number of census tracts. This also explains why the sample size is not the same when using different geographic sub-units: districts may span multiple counties, only some of which may belong to a metro area. Several metros have only one school district. The dissimilarity index is necessarily equal to zero in such cases, since the income distribution at the district level coincides with the income distribution of the metro.

Table 16: School Districts: Summary Statistics

Year	**MSAs**	**School Districts**	**All Families**
1980	379	6611	75,233,974
1990	379	6669	63,218,899
2000	379	6849	70,998,529
2010	380	6838	76,071,068

B Proof of Proposition 1

Given that we focus on equilibria with $R_t^A > R_t^B = 0$, we require $S_t^A > S_t^B$ for all t. Also, this requirement together with Assumption 1, implies that agents who choose low education strictly prefer neighborhood B to neighborhood A, so nobody chooses $e = e^L$ and $n = A$. Hence, agents choose among three options: 1) high education and neighborhood A-for short, HA; 2) high education and neighborhood B, HB; and 3) low education and neighborhood B, LB.

Let us consider a given time t and drop the time subscript to simplify notation. Also, to simplify notation, let us drop ε, given that it is iid, so it does not play any role for the optimal policies. Consider an agent with wealth w and ability a who chooses HA. It must be that he prefers that to HB or LB; that is,

$$u(w - R^A - \tau) + g(\Omega(w, a, e^H, S^A)) \geq u(w - \tau) + g(\Omega(w, a, e^H, S^B)) \tag{12}$$

and

$$u(w - R^A - \tau) + g(\Omega(w, a, e^H, S^A)) \geq u(w) + g(\Omega(w, a, e^L, S^B)). \tag{13}$$

Take any $w' > w$. By concavity of u and $R^A > 0$, we have

$$u(w' - R^A - \tau) - u(w' - \tau) \geq u(w - R^A - \tau) - u(w - \tau)$$

and

$$u(w' - R^A - \tau) - u(w') \geq u(w - R^A - \tau) - u(w).$$

Combining these conditions with the assumption that the composite function $g(\Omega)$ has increasing differences in w and S and in w and e (from Assumption 2), we obtain

$$u(w' - R^A - \tau) + g(\Omega(w', a, e^H, S^A)) \geq u(w' - R^B - \tau) + g(\Omega(w', a, e^H, S^B))$$

and

$$u(w' - R^A - \tau) + g(\Omega(w', a, e^H, S^A)) \geq u(w' - R^B) + g(\Omega(w', a, e^L, S^B))$$

for all $w' > w$ and given a. Let us call $w_1(a)$ and $w_2(a)$ the values of w that respectively make conditions (12) and (13) hold with equality for given a. We can then define the cut-off function as

$$\hat{w}(a) = \max\{w_1(a), w_2(a)\}.$$

This proves that all agents with $w \geq \hat{w}(a)$ choose the option HA for a given a. If we use Assumption 1 and 2 and the implicit function theorem, it is straightforward to show that both $w_1(a)$ and $w_2(a)$ are non-increasing functions, and hence that $\hat{w}(a)$ is a non-increasing function as well.

Next, consider an agent with wealth w and ability a who chooses LB. By revealed preferences, he must prefer that to HA or HB; that is,

$$u(w-R^B)+g(\Omega(w,a,e^L,S^B)) \geq u(w-R^A-\tau)+g(\Omega(w,a,e^H,S^A)) \tag{14}$$

and

$$u(w-R^B)+g(\Omega(w,a,e^L,S^B)) \geq u(w-R^B-\tau)+g(\Omega(w,a,e^H,S^B)). \tag{15}$$

Following steps analogous to the ones before, we can show that for a given a, all agents with $w' < w$ prefer LB to both HA and HB. Notice that the value w that makes equation (14) hold with equality is the cut-off value $w_2(a)$ defined above. Moreover, let us call $w_3(a)$ the value of w that makes condition (15) hold with equality for given a. We can then define the cut-off function as

$$\hat{\hat{w}}(a) = \min\{w_2(a), w_3(a)\}.$$

This proves that all agents with $w \leq \hat{\hat{w}}(a)$ choose the option LB for given a. If we use Assumption 2 and the implicit function theorem, it is straightforward to show that $w_3(a)$ is also a non-increasing function, and hence that $\hat{\hat{w}}(a)$ is a non-increasing function as well. Given that both $\hat{w}(a)$ and $\hat{\hat{w}}(a)$ are non-increasing functions, it must be that $\hat{w}(a) \geq \hat{\hat{w}}(a)$ for all a. If there was an a' such that $\hat{w}(a) < \hat{\hat{w}}(a)$, then all agents with $w \in (\hat{w}(a), \hat{\hat{w}}(a))$ would find strictly optimal both HA and LB, which is a contradiction. This proves that an equilibrium is characterized by two non-increasing functions $\hat{w}(a)$ and $\hat{\hat{w}}(a)$, with $\hat{w}(a) \geq \hat{\hat{w}}(a)$ for all a, such that all agents with (w,a) such that $w > \hat{w}(a)$ choose $e = e^H$ and $n = A$ and all agents with (w,a) such that $w < \hat{\hat{w}}(a)$ choose $e = e^L$ and $n = B$.

C Normalizations

For convenience, let us report the optimization problem for a household with wage w and a child with ability a,

$$u(w,a) = max_{e,n} \ln(w - R_n - \tau e^\gamma) + \ln\left(b + ae(\beta_0 + \beta_1 S_n)^\xi\right) + \alpha \ln w + \ln \varepsilon + \sigma_\zeta \zeta_n, \tag{16}$$

and her future wage,

$$w'(w,a) = \left(b + ae(\beta_0 + \beta_1 S_n)^\xi\right) w^\alpha \varepsilon. \tag{17}$$

First note that average ability is not independent of β_0 and β_1, as we can scale a by a constant c_a and scale both β_0 and β_1 by $c_a^{-\frac{1}{\xi}}$ while leaving the optimization problem and the wage expression unchanged. Specifically, we can set $c_a = \frac{1}{\mu_a}$ so that the adjusted average ability is equal to 1.

Moreover, we can scale ε by a constant c_ε and, at the same time, scale b by c_ε^{-1} and both β_0 and β_1 by $c_\varepsilon^{-\frac{1}{\xi}}$, again leaving the problem unchanged. We can normalize the mean of ε to 1 by setting $c_\varepsilon = \frac{1}{\mu_\varepsilon}$.

Next, notice that we can multiply w by a constant c_w. We can scale b by $c_w^{-(1-\alpha)}$ and β_0 and β_1 by $c_w^{\frac{-(1-\alpha)}{\xi}}$. This leaves the wage dynamics unchanged. Moreover, from the housing market condition, R_n is going to be automatically scaled up by the same constant, and we can multiply τ by c_w so that the optimization problem is unchanged as well. This means that we can choose $c_w > 0$ so that the average wage in the economy is equal to 2.44, the average income for a family with children in 1980, in $\$10,000s$.

Finally, we show that we can normalize $\tau = 1$. In particular, we can make the transformation $\tilde{e} = \tau^{\frac{1}{\gamma}} e$ and scale β_0 and β_1 by $\tau^{\frac{1}{\gamma\xi}}$. This leaves the optimization problem (where we now optimize over n and $\tilde{e}$ instead of n and e) and the wage equation unchanged.

D More on Counterfactual Exercise II

In Section 5.1, we show that in the second counterfactual exercise, where we keep location fixed, segregation is not constant but declines over time. This is driven mainly by the endogenous dynamics of the ability distribution in the different neighborhoods. In particular, in our model, given the complementarity between spillover and ability, the neighborhoods with higher spillover tend to attract families with children with higher ability. This is even more true in response to the skill premium shock. Once we shut down the sorting process by fixing the families' locations, the average ability in the neighborhoods tend to converge over time, given mean reversion in the ability process. Panel a in Figure 20 shows that in response to the skill premium shock, average ability in the three neighborhoods in our baseline model tends to diverge. Panel b in the same figure shows that, once we keep fixed family residential choice, average ability in the three neighborhoods tend to converge. The speed of this convergence is affected by the persistence of the ability process that we calibrated in Section 4.2.

Figure 20: Evolution of Average Ability in the Neighborhoods

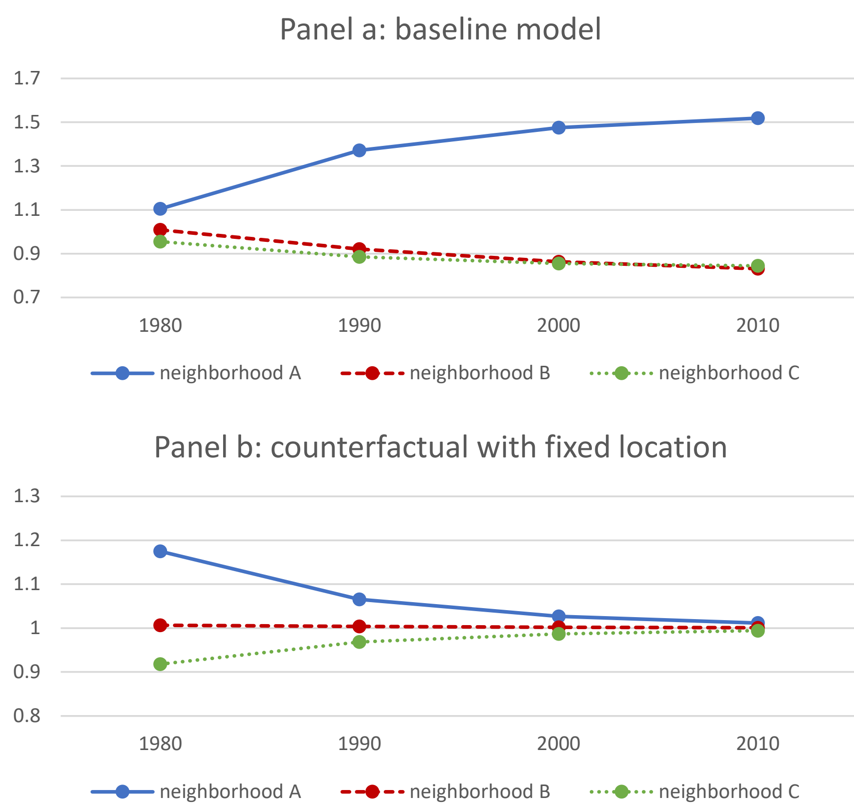


Figure 21 compares the dynamics of segregation in the baseline model and in the counterfactual with fixed location with those in an alternative counterfactual where keep fixed not only the families' location but also their investment in education. This figure shows that the pattern of segregation does not change much when we keep education fixed; if anything, it decreases even further. This implies that the endogenous choice of education cannot be behind the decline in segregation.

In the counterfactual with fixed location that we explore in section 5.1, parents do not have a residential choice, and rental rates are fixed at their steady state values. In Figure 22, we explore an alternative version of the counterfactual. In this version, agents optimally choose their location, but housing supply in the three neighborhoods is fixed so that their sizes are at the steady state levels and rental rates adjust to clear the housing markets. The figure shows that this choice does not quantitatively affect the results.

Figure 21: Segregation: Fixed Location and Education

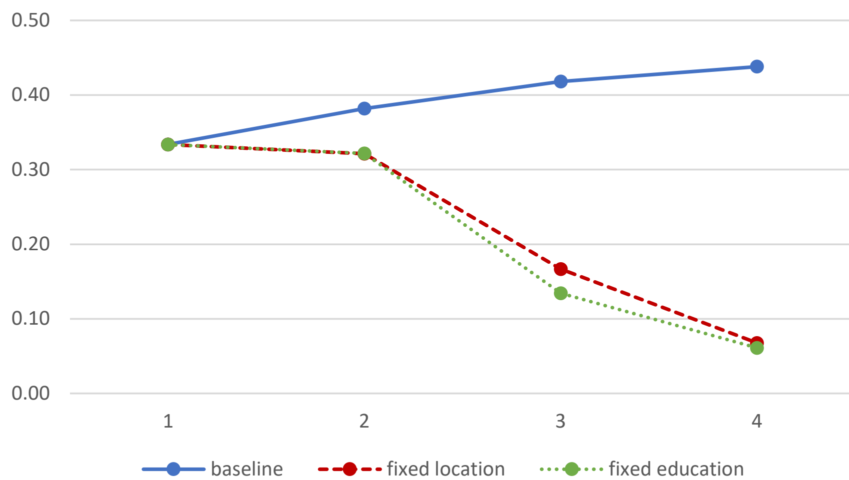
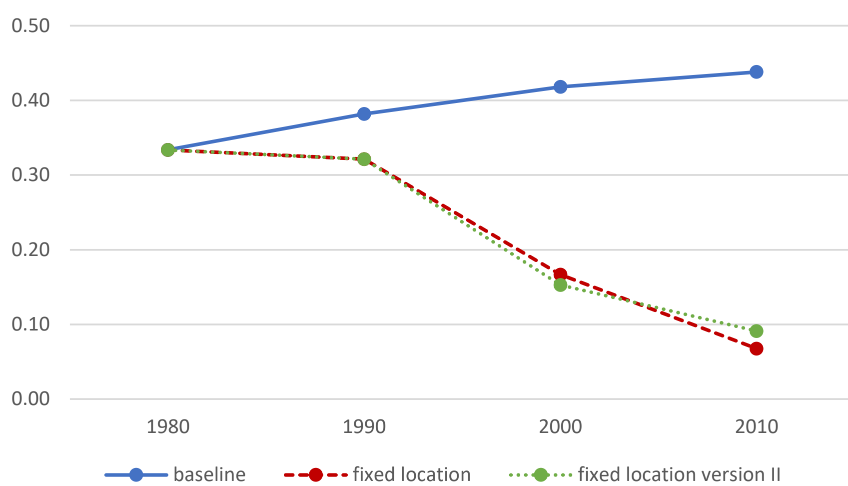


Figure 22: Segregation: Fixed Location and Housing Supply



E Alternative Model Specifications and Robustness

E.1 Model with Global Spillovers and Local Amenities

In our model, spatial segregation is driven not only by the presence of local spillovers but also by the presence of local amenities. In this section, we try to disentangle the two by studying a version of the model where spillovers are global, not local, and local amenities are the only source of segregation.

In particular, we consider a version of the model where the spillover is the same in all three neighborhoods and equal to the expected wage of the children in the city, $S_t = E_t(w_{t+1})$, so that future wages do not depend on the neighborhood where children grow up. Rental rates in different neighborhoods are still different, but just to reflect the presence of different local amenities that make a random fraction π of the agents prefer neighborhood A to B and neighborhood B to C, given that $\theta_A > \theta_B > \theta_C$. This generates segregation by income, given that richer families can afford to pay higher rents to live in neighborhoods with better amenities. However, such a spatial segregation does not translate in differential returns to education. This implies that a skill premium shock does not affect directly the marginal advantage of living in one neighborhood relative that of living in another. When the skill premium shock hits the economy, inequality increases, making richer people even richer and more willing to pay higher rental rates for better amenities. This, in turn, increases residential segregation. However, such an increase in segregation does not feed back into higher inequality, because residential choices do not affect expected wages in this version of the model.

Figure 23 compares the pattern of inequality and segregation after the skill premium shock in the baseline model (blue solid lines) with the same patterns in the version of the model just described (red dashed lines). The figure shows that after recalibrating the model to hit the same moments as in the baseline exercise, inequality and segregation behave quite similarly in the two models. However, such a model would miss important features of the data that our baseline model with local spillover can generate. In particular, Table 17 shows that this model would not be able to replicate two salient patterns in the data: the increase over time in the size of neighborhood A and the increase over time of the ratio of rental rates in neighborhood A relative to those in B, which are documented in Table 4. As we have shown in Table 3 in Section 4.3, our baseline model can match both these dynamics. The reason is that when the skill premium increases, everybody wants to invest more in education. This implies that when educational spillovers are equalized across neighborhoods, relatively poorer households may find it more attractive to live in neighborhoods with worse amenities but cheaper housing, where they can afford to invest more in education, hence reducing the demand to live in neighborhood A. On the contrary, in our baseline model with local spillover, neighborhood A becomes relatively more attractive because

Figure 23: Local Spillover vs. Local Amenities

of the complementarities between spillover and education, thus matching the increase in demand to live in neighborhood A that is a feature of the data. In the next section, we will explore an alternative shock that does not directly affect the return to education.

E.2 Different Shock to Inequality

Another important choice that we make in our baseline exercise is about the nature of the shock. One natural question is if the implied dynamics of the model would be similar if we considered a different type of inequality shock-for example, a simple increase in the volatility of the noise of the wage process. It is interesting to note that given the nature of our mechanism, which is centered on the effect of local spillover on returns to education, such a shock has different implications relative to a skill premium shock, which directly affects the return to education.

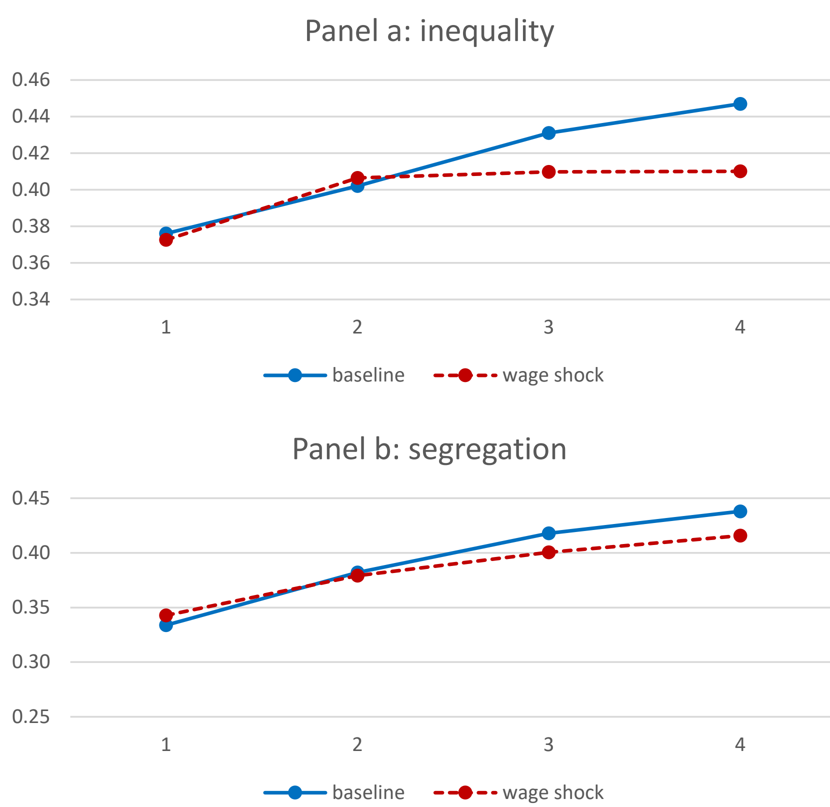
Figure 24 shows the dynamics of inequality and segregation implied by our baseline model in re-

Table 17: *Neighborhood Sizes and Rental Rates*

	1980	1990	2000	2010
R_A/R_B	1.421	1.398	1.384	1.367
R_B/R_C	1.412	1.581	1.805	1.980
$SizeA$	0.193	0.183	0.174	0.163
$SizeB$	0.301	0.277	0.255	0.234
$SizeC$	0.506	0.541	0.573	0.602

sponse to an increase in the volatility of the wage noise σ_ε. The figure shows that both inequality and segregation increase less over time in response to such a shock, relative to a skill premium shock, and this difference is quantitatively larger for inequality. This is because a skill premium shock has a more persistent effect on educational investment, which increases the gap between rich and poor.

Figure 24: Shock to Volatility of Wage Process



Another important difference with our baseline exercise is that, after a wage volatility shock, the model is not able to replicate the relative rental rates dynamics. Table 18 shows that after a wage volatility shock, both the rental rate ratio of neighborhood A versus B and that of B versus C are decreasing over time in contrast with the data. This happens because there is no change in the return to education, so the pool of families that select into the better neighborhoods is determined mainly by income and not by ability. In contrast, in response to a skill premium shock, richer families with more talented kids tend to move to better neighborhoods. They end up investing more in education, thereby increasing the spillover gaps between neighborhoods. This, in turns, feeds back into more future segregation and inequality, generating a persistent effect over time.

Table 18: *Neighborhood Sizes and Rental Rates*

	1980	1990	2000	2010
R_A/R_B	1.286	1.260	1.222	1.186
R_B/R_C	1.306	1.284	1.249	1.226
SizeA	0.193	0.207	0.205	0.210
SizeB	0.301	0.285	0.272	0.259
SizeC	0.506	0.507	0.523	0.531

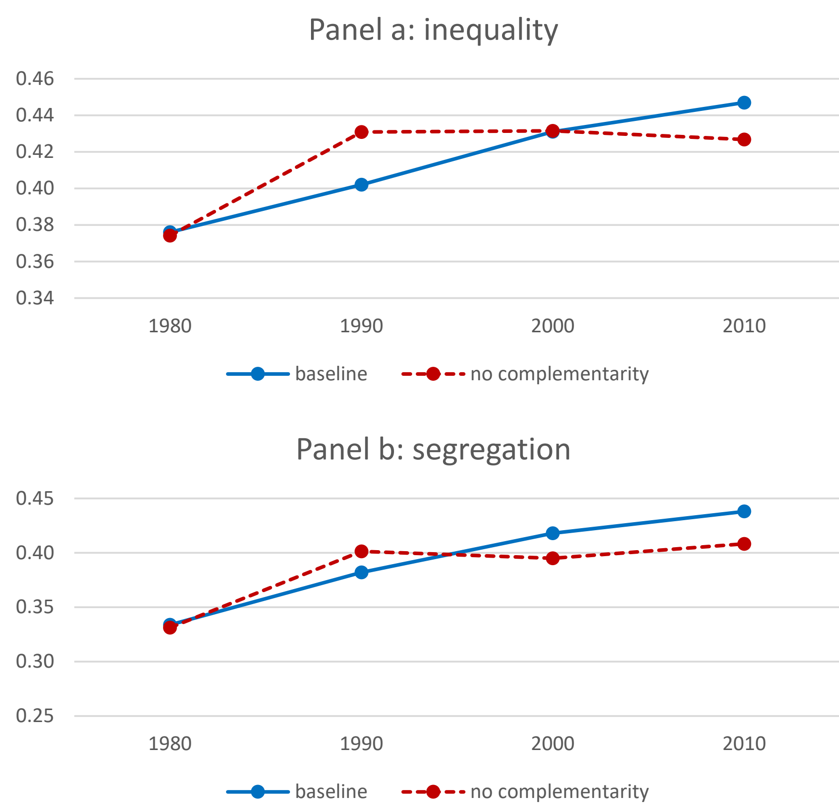
E.3 Role of Complementarities

In the baseline model, we assume complementarity between the spillover and innate ability in the wage process. Although we believe this assumption is consistent with some recent empirical literature on skill formation (see papers cited in Section 3), we also find it interesting to relax it and explore the case where spillover and ability enter linearly in the wage process. In particular, we assume that

$$w_{t+1} = (b + e_t(a_t + (\beta_0 + \beta_1 S_{nt})^{\xi})w^{\alpha}\varepsilon.$$

Figure 25 compares our baseline model with a recalibrated version of the model with the above wage process. The figure shows that this assumption is not particularly crucial for the qualitative properties of the model, because both inequality and segregation increase in response to the shock. However, the persistency of both inequality and segregation decrease substantially relative to our baseline. When ability and spillover are complementary, in response to the skill premium shock, families with more talented kids tend to move to the better neighborhood to exploit the higher spillover where the return to education is higher. This, in turn, increases the spillover gap and makes the effect persistent. However, when spillover and ability enter linearly the wage process, there is no particular incentive for families with more talented kids to move, and so this effect is muted.

Figure 25: Role of Complementarity in Wage Process



E.4 Alternative Spillover Specification

Our choice of spillover definition is motivated by the desire not to take a stand on the specific source of the local spillover, to better replicate the quasi-experiment studied in Chetty and Hendren (2018b). In particular, we define the spillover as average expected wage of the children growing up in a neighborhood. Future wages are affected both by parental income, through a direct effect and an indirect effect on the residential and educational choice, and by children's ability, which affects the returns to education and to spillover. This means that in our formulation both average ability and average parental income affect the size of the spillover, making it possible to capture a variety of mechanisms: peer effects, public school quality, networks, social norms, and so on. However, the past literature on local spillovers (e.g. Benabou (1996a), Benabou (1996b), Fernandez and Rogerson (1996), Fernandez and Rogerson (1997), Fernandez and Rogerson (1998), Eckert and Kleineberg (2021)) has focused on the effect of the quality of public schools. Given that public schools are locally financed, these papers define the spillover as the average parental income in the neighborhood. In this section, we explore this alternative spillover specification and compare the implied dynamics of inequality and segregation in response to a skill premium shock to our main results.

Figure 26: Spillover Equal to Parents' Average Income

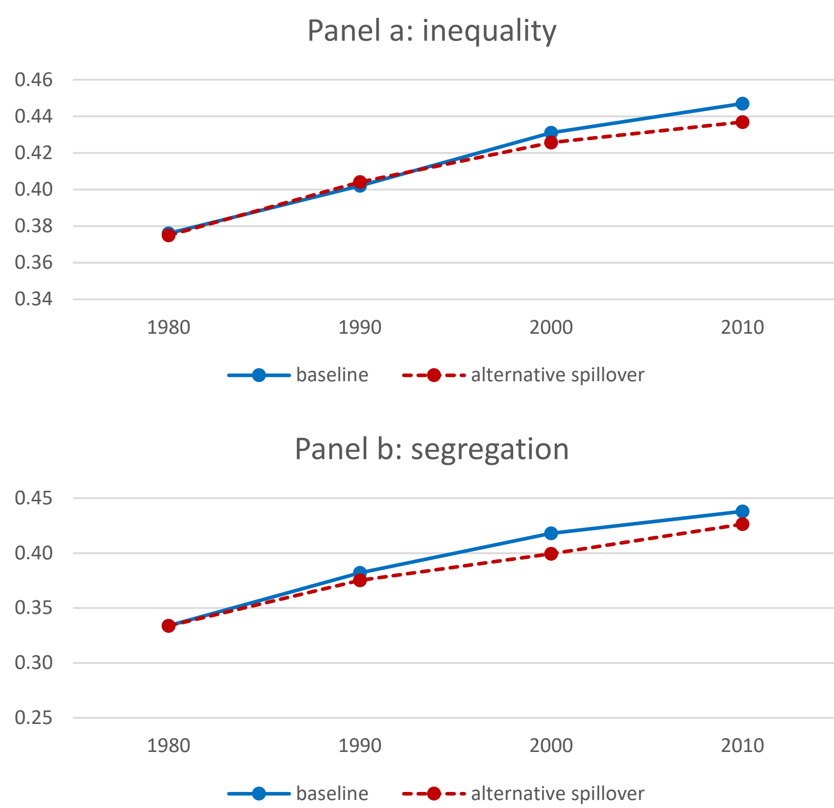


Figure 26 compares the dynamics of inequality and segregation in the baseline model (blue solid lines) with the same dynamics in the model where the neighborhood spillover is defined as the average income of the parents living in the neighborhood (red dashed lines). The figure shows that in the baseline model, there is more propagation of the skill premium shock both for inequality and segregation. This is expected, given that in the baseline model the spillover does not depends only on parents income but also on children's ability, which allows additional mechanisms to affect local spillover, such as, for example, peer effects. However, the figure shows that quantitatively, the difference is not large, which suggests that parental income is the most important factor in local spillover effects-for example, by affecting the quality of public schools. For future research, it would be interesting to use micro data to better disentangle better these different effects.

If we run the same counterfactual exercises that we performed for the baseline model in Subsection 5.1, we obtain that the contribution of segregation to the increase in inequality between 1980 and 2010 with this spillover specification is equal to 22% and to 16%, relative to 27% and 25%, respectively.

Overall, this exercise also shows that our mechanism's capability to explain the increase in inequality due to segregation is sizeable, irrespective of the specific definition of the local spillover.

E.5 Equilibrium Definition with MTO

To resemble the MTO program, we assume that the poorest x families living in the worst neighborhood (neighborhood C) are offered a voucher to move to a better neighborhood. Families accepting a voucher are required to pay 30% of their income toward their rent. The voucher covers the difference between their rent and the family's contribution up to a maximum amount, known as the Fair Market Rent, defined as the 40th percentile of rental costs in the metro area.

We examine the "experimental policy," which requires that a family must move to the best neighborhood (the neighborhood with the highest spillover) if they accept the voucher.

To finance the policy, we assume that the vouchers are covered by a proportional tax ν on income, levied on all families in the city.

In order to characterize the equilibrium with the presence of the policy, we need to introduce an additional state variable, which is the neighborhood where a parent grew up, or birth neighborhood, b_t.

A parent is now characterized by the triplet (w_t, a_t, b_t). Let $\chi(w_t, a_t, b_t)$ denote the eligibility indicator, so that $\chi_t(w_t, a_t, b_t) = 1$ if the parent is eligible and equal to 0 otherwise. The MTO policy prescribes that $\chi(w_t, a_t, b_t) = 1$ if $b_t = \arg\min_k \in A, B, C\{S_A, S_B, S_C\}$ and $w_t \leq \tilde{w}_t$, where

$\tilde{w}_t$ is such that

$$G[\tilde{w}_t|a_t,b_t] = p,$$

where p is a given percentile of the income distribution of the metro area. We define the worst neighborhood as the one with the lowest spillover. Given that the local spillovers are endogenous and evolve over time, the worst neighborhood may change in response to the policy.

Parents who are eligible for the voucher will accept it if they get higher utility from accepting it than from not accepting it. Define $v_t(w_t,a_t,b_t)$ as the voucher acceptance indicator. For all parents such that $\chi_t(w_t,a_t,b_t) = 0$, $v_t(w_t,a_t,b_t) = 0$, while for all parents such that $\chi_t(w_t,a_t,b_t) = 1$, $v_t(w_t,a_t,b_t)$ solves

$$max_{v_t}\{U^V(w_t,a_t,b_t), U^N(w_t,a_t,b_t)\}, \tag{18}$$

where

$$\begin{aligned} U^V(w_t,a_t,b_t) = \max_{e_t} & \log(1+\theta_{\hat{n}})((1-\nu)qw_t - \tau e_t^{\gamma})+ \\ & + \log[(y + a_t e_t \eta(\beta_0 + \beta_1 S_{\hat{n}})^{\xi})w_t^{\alpha}\varepsilon_t] + \sigma_\zeta \zeta_{\hat{n}t} \end{aligned} \tag{19}$$

and

$$\begin{aligned} U^N(w_t,a_t,b_t) = \max_{e_t,n_t} & \log(1+\theta_{n_t})((1-\nu)w_t - R_{n_t} - \tau e_t^{\gamma})+ \\ & + \log[(y + a_t e_t \eta(\beta_0 + \beta_1 S_{n_t})^{\xi})w_t^{\alpha}\varepsilon_t] + \sigma_\zeta \zeta_{n_t t}, \end{aligned} \tag{20}$$

where $U^V(w_t,a_t,b_t)$ is the value of accepting the voucher, $U^N(w_t,a_t,b_t)$ is the value of not accepting it, $\hat{n}$ is the neighborhood with the lowest spillover, and q is the fraction of the income remaining after contributing to the rent (equal to 70%). A parent accepting the voucher pays taxes on the remaining income qw_t and has to move to the neighborhood with the highest spillover. The voucher covers the difference between $R_{\hat{n}t}$ and her contribution. A parent not accepting the voucher has to pay the full rent and pays taxes on the total income, but can choose the optimal neighborhood.

We are now ready to define an equilibrium with voucher policy.

Definition 2 *Equilibrium. For a given initial wage distribution $G_0(w_0,a_0,b_0)$, an equilibrium is characterized by a sequence of educational and residential choices, $\{e_t(w_t,a_t,b_t)\}_t$ and $\{n_t(w_t,a_t,b_t)\}_t$, a voucher eligibility indicator $\{\chi_t(w_t,a_t,b_t)\}_t \in \{0,1\}$, a voucher acceptance choice $\{v_t(w_t,a_t,b_t)\}_t$,* *a sequence of rents and spillover sizes in the three neighborhoods,* $\{R_{kt}\}_t$ *and* $\{S_{kt}\}_t$ *for* $k = A,B,C$*, a sequence of tax rates* $\{\nu_t(w_t,a_t,b_t)\}_t$*, and a sequence of distributions* $\{G_t(w_t,a_t,b_t)\}_t$ *that satisfy:*

1. *agents' optimization:*

   (a) *for each* t *and* (w_t,a_t,s_t) *such that* $\chi(w_t,a_t,b_t) = 0$*, the policy functions* e_t *and* n_t *solve problem (20), for given* R_{kt} *and* S_{kt} *for* $k = A,B,C$*;*

   (b) *for each* t *and* (w_t,a_t,b_t) *such that* $\chi(w_t,a_t,b_t) = 1$*, the policy functions* e_t *and* n_t *solve problem (18), for given* R_{kt} *and* S_{kt} *for* $k = A,B,C$*;*

2. *spillover consistency: for each* t*, equation (2) is satisfied for* $n = A,B,C$*;*

3. *market clearing: for each* t *and* $k \in \{A,B,C\}$*,* R_{kt} *ensures housing market clearing in neighborhood* k*:*

$$\lambda_k \left(\frac{R_{kt}}{\bar{w}_t}\right)^{\phi_k} = \int\int\int_{n_t(w_t,a_t,b_t)=k} G_t(w_t,a_t,b_t)dw_t da_t ds_t; \tag{21}$$

4. *wage dynamics: for each* t*,*

$$w_{t+1} = \Omega(w_t,a_t,e_t(w_t,a_t),S_{n_t(w_t,a_t,b_t)t},\varepsilon_t); \tag{22}$$

5. *budget balance: for each* t*,* ν_t *is such that*

$$\int\int\int_{v_t(w_t,a_t,b_t)=1}(R_A-(1-q)w_t)G_t(w_t,a_t,b_t)dw_t da_t ds_t \leq \nu\left(\int\int\int w_t G_t(w_t,a_t,b_t)dw_t da_t ds_t\right).$$

In equilibrium, eligible parents optimally choose whether to accept a voucher or not.