

**WORKING PAPER** · NO. 2019-99

# The End of the American Dream? Inequality and Segregation in US cities

*Alessandra Fogli and Veronica Guerrieri*

JULY 2019

# The End of the American Dream? Inequality and Segregation in US cities \*

Alessandra Fogli  
Minneapolis Fed

Veronica Guerrieri  
University of Chicago

July 2019

## Abstract

Since the '80s the US has experienced both an increase in income inequality and an increase in residential segregation by income. After documenting this fact, we develop a general equilibrium model where parents choose the neighborhood where to raise their children. Segregation and inequality amplify each other because of local spillovers that affect the education returns. We calibrate the model using 1980 US data and the estimates for neighborhood exposure effects in Chetty and Hendren (2018b). We then show that segregation contributes to 28% of the increase in inequality between 1980 and 2010 after an unexpected permanent skill premium shock.

---

\*Email addresses: afogli00@gmail.com; vguerrie@chicagobooth.edu. For helpful comments, we are grateful to Roland Benabou, Jarda Borovicka, Steven Durlauf, Cecile Gaubert, Mike Golosov, Luigi Guiso, Erik Hurst, Francesco Lippi, Guido Lorenzoni, Guido Menzio, Alexander Monge-Naranjo, Fabrizio Perri, numerous seminar participants, and, in particular, to Elisa Giannone, Ed Glaeser, Richard Rogerson, Kjetil Storesletten, and Nick Tsvanidis for the useful discussions. For outstanding research assistance, we thank Yu-Ting Chiang, Gustavo Gonzalez, Hyunju Lee, Qi Li, Emily Moschini, Luis Simon, and, in particular, Mark Ponder and Francisca Sara-Zaror.

# 1 Introduction

It is a well documented fact that over the last 40 years, the US has experienced a steady increase in income inequality. 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 contributed to amplify the response of income inequality to underlying shocks, such as skill-biased technical change? In this paper, we build a model of human capital accumulation with local spillovers and residential choice that can be used to address these questions.

There has been a large theoretical literature in the '90s 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). More recently, administrative data have been used to propose direct estimates of neighborhood spillover effects. In particular Chetty et al. (2016), and Chetty and Hendren (2018a,b) have shown that there are substantial effects of children's exposure to different neighborhoods 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 segregation and to the recent rise in inequality.

In figure 1, we use the Theil index to decompose the increase in income inequality at the national level (blue solid line) in two parts: the increase in income inequality within metro areas (red dashed line) and the increase in income inequality across metro areas (green dotted line).<sup>1</sup> The Figure shows that both types of income inequality have increased steadily since the '80s and have substantially contributed to the rise in the US income inequality.

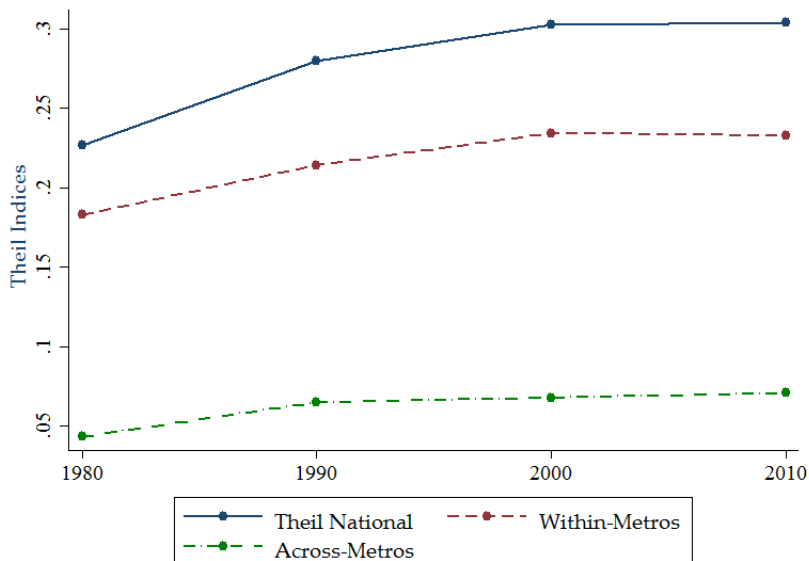
A recent vibrant literature has analyzed the divergence in economic outcomes across metro areas.<sup>2</sup> In our paper, we focus on the divergence in economic outcomes across neighborhoods within metro areas. We first document a positive correlation between income inequality and residential segregation by income at the MSA level, both across time and across space. We use US Census tract data on family income between 1980 and 2010 to construct measures of inequality

---

<sup>1</sup>For this figure, we use the Theil index because it is well suited for these types of decompositions. We use the same Census tract data on family income between 1980 and 2010 that we describe in section 2.

<sup>2</sup>See for example Moretti (2004), Shapiro (2006), Moretti (2012), Eeckhout et al. (2014), Hsieh and Moretti (2015), Diamond (2016), Giannone (2018).

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



and residential segregation at the MSA level. To measure inequality, we use the Gini coefficient as baseline indicator. To measure segregation, we use the dissimilarity index, which is a measure of how uneven is the distribution of two exclusive groups across geographical areas. In particular, we divide the population in two income groups, rich and poor, using the 80th income percentile, and compute the dissimilarity index across census tracts belonging to the same MSA. Using these measures, we show that 1) average inequality and residential segregation have increased steadily since 1980; 2) inequality and residential segregation in 1980 are correlated across MSA; 3) the changes in inequality and residential segregation between 1980 and 2010 are correlated across MSA. We also check the robustness of our main findings with alternative measures of income inequality, such as the 90/10 ratio, and of income segregation, such as the dissimilarity index calculated with different percentiles and an alternative index of segregation,  $H^R$ , that has been used in the literature.<sup>3</sup>

We then build a general equilibrium overlapping generation model with human capital accumulation and residential choice that features local externalities. The model generates a feedback effect

<sup>3</sup>The  $H^R$  index has been proposed by Reardon and Firebaugh (2002) and Reardon and Bischoff (2011), and also used by Chetty et al. (2014).

between income inequality and residential segregation that amplifies the response of inequality to underlying shocks. 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 the level of their children's education. The key ingredient of the model is a local spillover: investment in education yields higher returns in neighborhoods with higher average level of human capital. Such a spillover can capture a variety of mechanisms: differences in the quality of public schools, peer effects, social norms, learning from neighbors' experience, networks, and so forth. It is outside the scope of the paper to determine which spillover channel is most important.<sup>4</sup> The relevant assumption is that the local spillover is complementary to the children's innate ability and to their level of education. This generates sorting in equilibrium: richer parents with more talented children choose to pay higher rents to live in the neighborhood with higher average human capital. It follows that in equilibrium one neighborhood becomes endogenously the "good" one and hence the one where houses are more expensive. This means that in this model, the residential choice is a form of human capital investment.

First, we use a baseline version of the model with binary education choice to understand qualitatively the feedback effect 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 stronger spillover. However, given spatial constraints, this translates into higher housing costs, and hence into 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 inequality.

Next, in order to bring the model to the data, we extend it to embed both a continuous education choice and a local preference shock, and we 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 use the micro estimates for neighborhood exposure effects obtained in the

---

<sup>4</sup>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 and marriage market play a more important role.

quasi-experiment of Chetty and Hendren (2018b).

We then perform our quantitative exercise. We assume that the original increase in inequality comes purely from skill-biased technical change and 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 can then use our model to ask our main quantitative question: how much does segregation by income contribute to the rise in inequality? To answer this question, we run a counterfactual exercise where we look at the response of the economy to the same shock, but assume that, after the shock, families are randomly re-located between the two neighborhoods. The exercise shows that segregation by income contributes to 28% of the total increase in inequality between 1980 and 2010.

We also perform a number of different exercises that assess the importance of local spillovers from different angles. These complementary counterfactuals give results broadly in line with our main exercise.

### **Related Literature.**

Our model builds on a large class of models with multiple communities, local spillovers, and endogenous residential choice, studying 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 and studies its efficiency properties.<sup>5</sup> 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 together with strong neighborhood feedback effects generate persistent inequality.<sup>6</sup> Benabou (1996a) embeds growth with complementary skills in production in a similar model, where local spillovers are due both to social externalities (as peer effects) and locally financed public school. The paper analyzes the trade-off coming from the fact that stratification helps

---

<sup>5</sup>De Bartolome (1990) also studies efficiency properties of a similar type of model where communities stratification is driven by peer effects in education. In similar papers, the local social externalities take the form of role models (Streufert (2000)), or referrals by neighborhoods (see Montgomery (1991a,b)).

<sup>6</sup>Durlauf (1996a) uses a related model to study how it can generate permanent relative income inequality (opposed to absolute low-income or poverty traps) in an economy where everybody's income is growing.

growth in the short run due to the complementarities in skills, while integration helps growth in the longer run, as generates less inequality, and hence heterogeneity in skills, over time. It also studies how alternative systems of education financing affect the economy. 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.<sup>7</sup> 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 in a similar model, but he allows for non-fiscal channels of local spillovers, like peers, role models, norms, networks, and so forth and shows that disentangling between financial and social local spillover is important for assessing different types of policies.

Similarly to this class of paper, our model builds on the idea that stratification, due to a local spillover, generates more inequality over time. We focus on a model that can be calibrated and brought to the data, while, most of the papers discussed, with the notable exception of Fernandez and Rogerson (1998), focus on the qualitative implications of the models. In that spirit, most of them analyze the two extreme scenarios of full stratification and full integration. Given our quantitative direction, we enrich the model to obtain a continuous measure of segregation. In order to discipline the model with data on education, we also introduce an endogenous educational choice, that is absent in the previous papers. Moreover, differently from the literature, we model the local spillover as a black box, that can be interpreted as driven either by a financial or a social channel. While for normative questions that have been explored in the literature the specification of the spillover is clearly important, for positive questions like the ones we address in this paper, it is less so. This is why we prefer to leave the framework more flexible to possibly incorporate different types of local spillover effects.

The most related paper to our work 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 to lower intergenerational mobility, the "Gatsby curve". The model in the paper is close to our model in many dimensions, although the calibration strategy and the main exercise are different and complement well each other.

---

<sup>7</sup>In a similar framework, Fernandez and Rogerson (1997) study the effect of community zoning regulation on allocations and welfare.

In contemporaneous work, Eckert and Kleineberg (2019) study 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. To this end, they structurally estimate the model using regional data of the US geography to match model cross-sectional predictions. Another related paper is Zheng (2017), who calibrates a similar model to study the effects of different public school allocation mechanisms.

Another recent related paper is Ferreira et al. (2017), who use a model close to ours to think about the emergence and persistence of urban slums and calibrate it to Brazilian data. They propose a model with overlapping generation of individuals with different skills, where local spillovers take the form of human capital externalities. They embed growth in the model to think about structural transformation together with urban evolution. They use the model to ask what are the effects of slums on human capital accumulation, structural transformation, urban development and mobility.

Another related strand of the literature focuses on spatial sorting generated by local amenities. The early work by Brueckner et al. (1999) and Glaeser et al. (2001) emphasizes the role of urban amenities and spurred a vibrant literature on gentrification. Among the others, Guerrieri et al. (2013) have focused on the endogenous nature of amenities, by introducing a consumption externality that comes from the average income of the neighbors. In contemporaneous work, Couture et al. (2019) study a spatial model with locations with different endogenous amenities and non-homothetic preferences. The paper focuses on the growth and welfare effects of spatial resorting within urban areas after the '90s. Another related paper is Bilal and Rossi-Hansberg (2019) who emphasize that the location choice of individuals is a form of asset investment.

Our work is also related to the literature investigating the evolution of race-based segregation in US cities and its consequences on individual outcomes. The seminal paper of Cutler and Glaeser (1997) shows that blacks living in more segregated metros have significantly worse outcomes than blacks living in less segregated cities. Given the correlation between income and race, these findings are relevant for our analysis. Interestingly, however, Cutler et al. (1999) show that the American ghetto, rapidly expanding between 1890 and 1970 as blacks migrated to the cities, eventually started declining. Income-based segregation has progressively replaced race-based segregation in US cities.



Besides the vast literature on city segregation, there are also papers that investigate the consequences of high levels of segregation in a cross section of countries. Alesina and Zhuravskaya (2011), using a measure of segregation similar to ours, show that countries where different linguistic and ethnic groups are more segregated across regions are characterized by significantly lower government quality.

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 as in the data. Section 5 shows our main counterfactual exercises to quantify how much segregation has contributed to the increase in inequality. Section 6 concludes.

## 2 Empirical Evidence

Over the last forty 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, historically predominantly segregated on the basis of race.<sup>8</sup> During the last third of the twentieth century, the United States moved toward a new regime of residential segregation characterized by decreasing racial-ethnic segregation and rising income segregation. Such a shift took place at the same time of a steady increase in income inequality.

In this section we document the magnitude of these phenomena and show the correlation between segregation and inequality across time and space. These measures will be used for our calibration exercise in Section 4.

---

<sup>8</sup>Massey et al. (2009) documents that from 1900 to 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).

## 2.1 Segregation and Inequality over Time

The term segregation refers to the spatial distribution of different groups of a population in a geographic unit across geographic subunits. The groups can be defined according to different categories, such as race, education and income, and segregation can be measured at different geographic levels, such as state, county or metro. We are interested in measuring the residential segregation by income within US cities.

In general, segregation is a multidimensional concept, capturing different aspects of the spatial distribution of the population.<sup>9</sup> In this paper, we follow Massey et al. (2009) and focus on the dimension known as evenness, that is, the degree to which different groups are distributed evenly over a set of geographic units.<sup>10</sup> In particular, we use the index of dissimilarity, which is the most common measure of evenness, to measure the segregation of rich and poor families across census tracts within metro areas. In our main analysis, we define rich all the families with income above the 80th percentile of the metro family income distribution, and poor all the others. 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  $Y(j)$  denote the total number of, respectively, poor and rich families in metro  $j$ , while  $x_i(j)$  and  $y_i(j)$  denote the number of, respectively, poor and rich families in census tract  $i$  in metro  $j$ .<sup>11</sup>

We use tract level family income data from Decennial Censuses (1980 to 2000) and from the American Community Surveys (2008-2012). Our sample includes 380 metropolitan areas using the 2003 OMB definition.<sup>12</sup> We calculate the dissimilarity index for all metro areas in each decade and average at the national level using metro level population weights. Figure 2 plots the resulting measure of segregation at the national level. The graph shows that the distribution

---

<sup>9</sup>Massey and Denton (1988) grouped the measures into 5 key dimensions: evenness, exposure, concentration, centralization, and clustering.

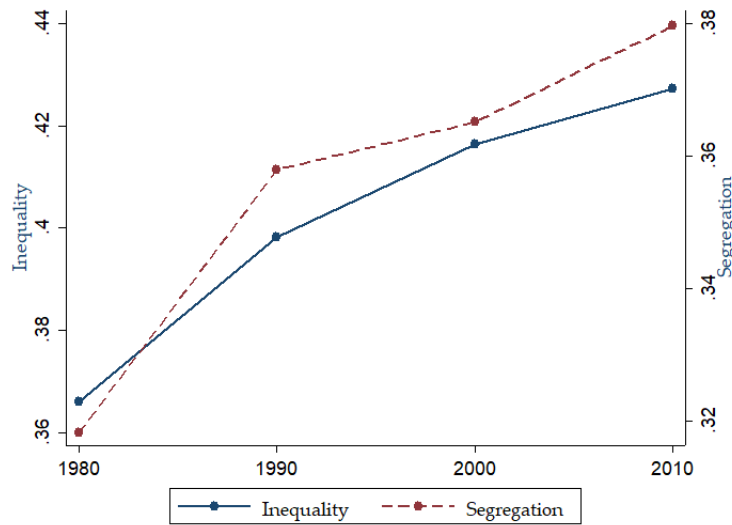
<sup>10</sup>A group is evenly distributed when each geographic subunit has the same percentage of group members as the population in the geographic unit.

<sup>11</sup>The dissimilarity index varies from 0 to 1, with the former value indicating perfect evenness and the latter maximum separation.

<sup>12</sup>For summary statistics of our sample see Appendix B.1.

of income has become progressively more uneven across census tracts over time.<sup>13</sup> If in 1980, roughly 32% of the population of the average US metro 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.<sup>14</sup>

Figure 2: Inequality and Segregation over Time



Using the same data on family income at the tract level that we use to calculate the dissimilarity index, 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

<sup>13</sup>In Appendix B.1 we use the metro area of Chicago as an example to visualize the change in segregation across census tracts.

<sup>14</sup>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.

the top-coded distribution.<sup>15</sup> We plot the resulting estimate of the Gini coefficient in Figure 2 together with the dissimilarity index. Both measures show a significant increase over time, especially between 1980 and 1990, with the Gini coefficient rising from roughly .36 to roughly .42 over the entire period. The figure shows that the increase in spatial segregation by income across neighborhoods happened at the same time of the increase in income inequality.

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

Figure 3: Dissimilarity Index: different cutoffs

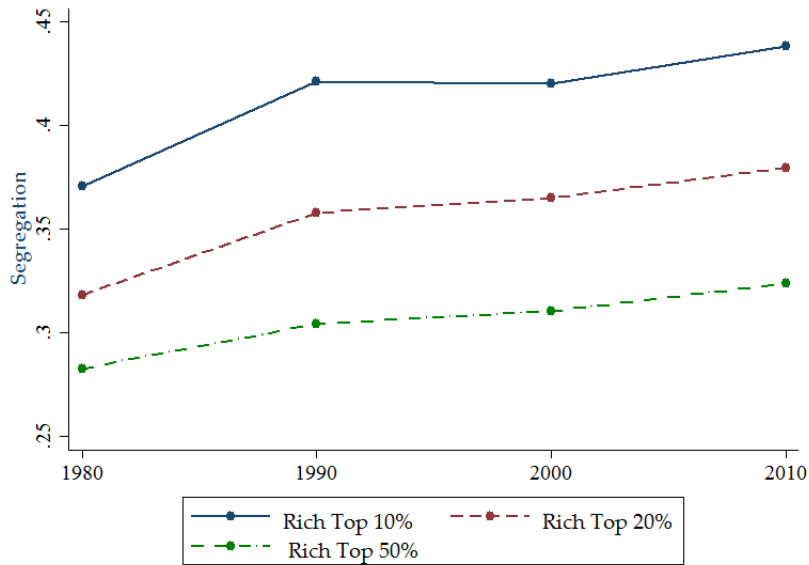


Figure 3 plots the dissimilarity index calculated using different percentiles to define the income groups. The red dashed line shows our benchmark dissimilarity index, while the solid blue line and the dotted green line show the dissimilarity index constructed using the 10<sup>th</sup> and the 50<sup>th</sup>

<sup>15</sup>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 mostly 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, for example, in Reardon and Bischoff (2011). However, such procedure does not exploit the fact that the Census reports the precise empirical average income by census tract. Our method uses this information that can be useful to improve the estimation of the top-coded distribution. For details, see Appendix B.1.

percentiles respectively. The figure shows that the dissimilarity index shifts up as the cut-off percentile decreases, suggesting that groups progressively more homogenous according to income are also characterized by higher levels of segregation. However, regardless of the level, all measures show an increasing trend over time. From now on, when we refer to the dissimilarity index, we refer to the average, population-weighted, of the dissimilarity indexes calculated at the metro level, using the 80th percentile as cut-off to define the rich and the poor.

Figure 4 shows our benchmark dissimilarity index together with the  $H^R$  index, which is another common measure of income segregation proposed by Reardon and Firebaugh (2002) and Reardon and Bischoff (2011). As before, we calculate the  $H^R$  index at the metro level and then aggregate at the national level using population weights.<sup>16</sup> The figure shows that the  $H^R$  index has also increased monotonically over the last four decades, confirming the segregation pattern emerging from the analysis of the dissimilarity index.

Relative to the dissimilarity index, the  $H^R$  index has the advantage of not relying on a single cut-off to define rich and poor, but has the drawback of being more likely to suffer from small sample bias.<sup>17</sup> In Reardon et al. (2018) the authors develop a methodology to correct the potential bias in the  $H^R$  index coming from the small number of observations that tend to overestimate the extent of segregation. Such a bias is also more likely to occur in the last decade of the sample period when Census data are not available and we use ACS data that are characterized by a significantly lower sampling rate than the decennial Census. Figure 4 also plots the "bias-corrected  $H^R$ " and shows that it is systematically lower than its uncorrected counterpart, and that this difference is larger, as expected, in the last part of the sample when we use ACS data. Nevertheless, the figure shows that also according to this index, segregation has increased after 1980.

The increase in inequality is also a robust finding.<sup>18</sup> Figure 5 plots three other measures of income inequality that have been widely used in the literature: the 90/10 ratio that measures the ratio of the family income in the top 90th percentile of the population relative to the income in the bottom 10th percentile, and, similarly, the 50/10 ratio, and the 90/50 ratio.<sup>19</sup> Figure 5 shows that both

---

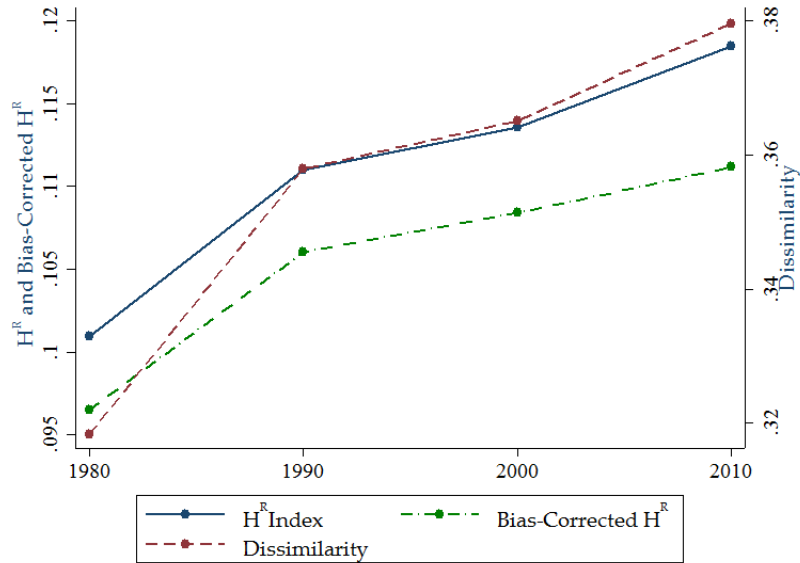
<sup>16</sup>In the appendix, we describe how this index is constructed.

<sup>17</sup>This issue is particularly salient for multigroup indexes cutting the distribution in many groups and simultaneously reducing the number of observations for each.

<sup>18</sup>See, for example, Katz and Murphy (1992); Autor et al. (1998); Goldin and Katz (2001); Card and Lemieux (2001); Acemoglu (2002); Autor et al. (2008).

<sup>19</sup>The procedure implemented to calculate these ratios from binned data at the census tract level is described in Appendix B.1.

Figure 4: Segregation: different measures

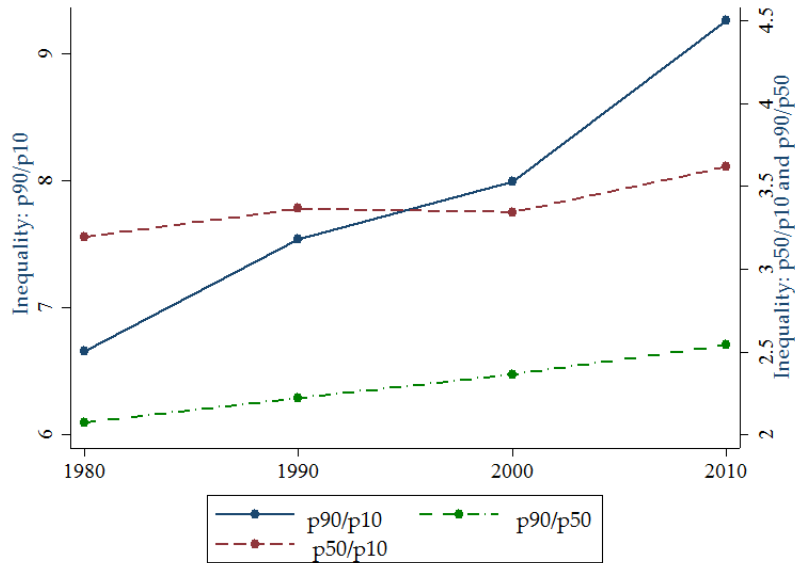


the 90/10 and the 90/50 ratios have increased steadily since 1980, while the 50/10 ratio is flat or even slightly decreasing after 1990. This confirms that the rise in income inequality has been driven by the top of the distribution, as already shown by Autor et al. (2008) for individual wage inequality.

## 2.2 Segregation and Inequality Across US Metros

Next, we document that residential segregation and inequality are also correlated across space. Figure 6 shows the relationship between the Gini coefficient and the dissimilarity 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. We estimate a regression coefficient of 0.25, with a standard error of 0.01. The significance of this relationship is robust to the inclusion of controls for demographics and industry composition. It also holds for the other decades in the sample and using the dissimilarity index constructed with other cut-off

Figure 5: Inequality: different measures



to define rich and poor families.<sup>20</sup>

The significance of the relationship between inequality and segregation is robust not just in levels but also in differences. Figure 7 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 the metro areas that experienced a larger increase in inequality between 1980 and 2010 are also those that experienced a larger increase in residential segregation over the same time period. The regression coefficient is 0.176 with a standard error of 0.017.<sup>21</sup>

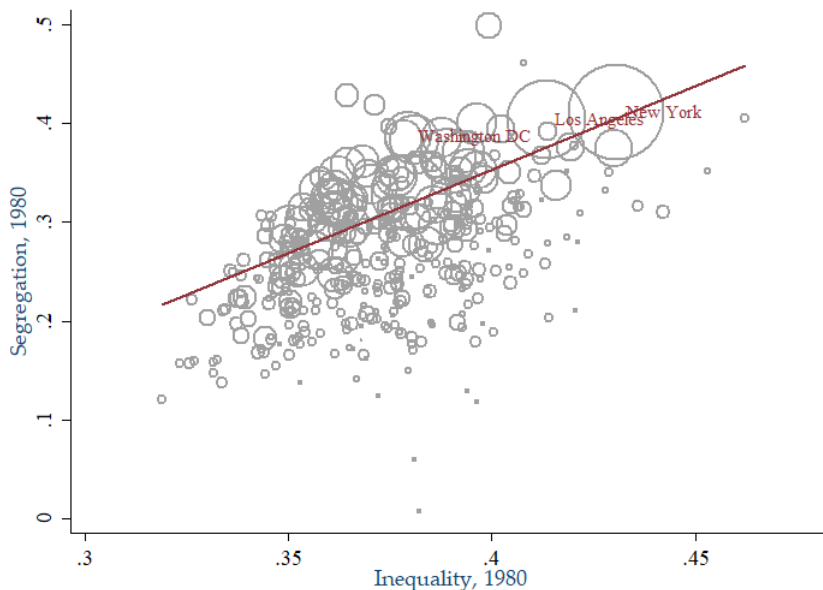
Our analysis suggests a positive correlation between inequality and segregation, both across time and across space. US cities have become increasingly segregated over time reflecting an increased tendency of families to sort in different neighborhoods according to income.<sup>22</sup>

<sup>20</sup>The results of the regression of inequality on segregation across US metros in 1980 with and without controls are reported in Table 5, Appendix B.2.

<sup>21</sup>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 also reported in Table 5, Appendix B.2.

<sup>22</sup>One of the important drivers of sorting is the quality of public school. The relevant subunit of analysis for public school driven type of segregation is the school district. We analyze the evolution of segregation across US school

Figure 6: Inequality and Segregation across US Metros



### 3 Model

We now propose a model of a metro area where families choose the neighborhood where to live taking into consideration that there are local spillovers affecting their children’s future income.

#### 3.1 Set up

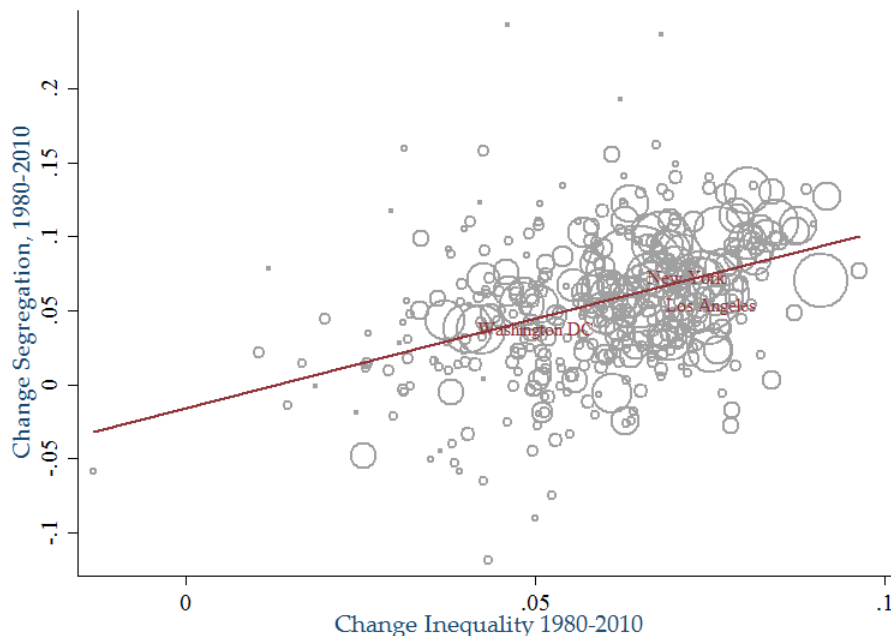
The economy is populated by overlapping generations of agents who live for two periods. In the first period, the agent is a child and accumulates human capital. In the second period, the agent is a parent. A parent at time  $t$  earns a wage  $w_t \in [\underline{w}, \bar{w}]$  and has one child with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

---

districts in Appendix B.3. We find a similar pattern of increase over time slightly mitigated in the last part of the period by the rise of private schools. In the Appendix we also argue that the census tract is a preferable unit of analysis in our context since better reflects our flexible notion of neighborhood spillover, is less affected by potential small sample bias and can be more directly linked to metros.



Figure 7: Change in Inequality and Segregation 1980-2000



AR1 process

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

where  $v_t$  is normally distributed with mean zero and variance  $\sigma_v$ , and  $\rho \in [0, 1]$  is the autocorrelation coefficient. The joint distribution 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, we make the extreme assumption that the housing supply is fixed and equal to  $M$  in neighborhood  $A$  and fully elastic in neighborhood  $B$ .<sup>23</sup> 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.

<sup>23</sup>We make this assumption for simplicity, but one could introduce an intermediate level of housing elasticity in both neighborhoods. The necessary assumption is that there is at least one neighborhood where housing supply is not fully elastic.

In the baseline model we assume that there are two educational levels, that is,  $e \in \{e^L, e^H\}$ .<sup>24</sup> Also, suppose that there is no cost to obtain the low level of education, while  $\tau > 0$  is the cost of investing in high education.

Parents care both about their own consumption and about their children's future wage.<sup>25</sup> 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 a local spillover that affects the children's human capital accumulation, and hence their future income. Children's wages are affected by their ability shock, by their education, by the neighborhood where they grow up because of the local spillover effect, and also directly by their parents' wage.<sup>26</sup> 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

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

where  $\varepsilon_t$  is an iid normally distributed noise with cdf  $\Psi$ ,  $S_{nt}$  is the size of the local spillover in neighborhood  $n$  at time  $t$ , and  $\Omega$  is non-decreasing in all its arguments. Children with higher ability and higher education, who grow up in neighborhoods with larger spillover and have richer parents will accumulate more human capital, and hence earn higher wages. Because the residential and the educational choice are functions of the parents' wage and child'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 parents with a higher wage, for given child's ability, will choose more 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 neighborhood choices.

---

<sup>24</sup>In the quantitative exercise we extend the model to allow for a continuous choice of education.

<sup>25</sup>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, but with the introduction of a borrowing constraint.

<sup>26</sup>Parents' wages also affect children's wages indirectly through the educational 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 average human capital of children growing up in that neighborhood, which in our model translates into the children's expected future average wage:

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

Given that wages are increasing in ability and in parents' wage, neighborhoods with higher spillover tend to be neighborhoods with richer parents and children with higher ability. The idea is that children growing up in these neighborhoods will accumulate more human capital and hence earn higher future income because of the stronger local spillover effect.<sup>27</sup> This formalization of local spillovers can capture different sources of pecuniary and social externalities: neighborhoods with richer families have better public schools that are typically locally financed, children who grow up in such neighborhoods have better peers and establish stronger social networks that will help them on the labor market, parents who live there invest more in education because they learn more successful stories, social norms are more conducive to educational investment, and so forth.<sup>28</sup> The presence of this externality implies that the rental rate in neighborhood  $A$ ,  $R_{At}$ , also depends on the level of the average human capital in that neighborhood  $S_{At}$ , which is endogenous.

In our analysis, we make two assumptions. First, for simplicity, we assume that ability and spillover's size affect children's future wages only if they get 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  $e = e^H$ .*

One could interpret children with high education as college graduates and children with low education as less than college graduates. 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 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

---

<sup>27</sup>Alternative specifications could 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. Also, in the baseline model, the second specification would not be particularly appealing because of the binary nature of the education level.

<sup>28</sup>We label the spillover "average human capital", but the same mathematical expression can stand for any externality that is affected by average parents' income and/or average children's ability. For example, the spillover could represent network connections on the labor market, and so forth.

children with low education is extreme, but can be interpreted as stating that the quality of k-12th schooling is more important in determining future wages of college graduates than of no-college graduates. This second assumption simplifies the analysis because all parents living in the rich neighborhood also pay for their children to get high education, 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.<sup>29</sup>

Second, we assume that there are complementarities between the spillover's size and children's ability, between education and ability, between parents' wage and the spillover's size, and between parents' wage and education. In particular, 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. We discuss it in detail in the next two subsections.

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

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

taking as given spillovers and rental rates in the two neighborhoods,  $S_t^n$  and  $R_t^n$  for  $n = A, B$ .

## 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's 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:*

---

<sup>29</sup>In particular, it will be enough to impose complementarity between the spillover and the education level, and between the ability and the education level.

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*

$$M = \int \int_{n_t(w_t, a_t) = A} F_t(w_t, a_t) dw_t da_t; \quad (3)$$

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

$$w_{t+1} = \Omega(w_t, a_t, e_t(w_t, a_t), S_{n_t(w_t, a_t)}, \epsilon_t). \quad (4)$$

From now on, we focus on equilibria where the housing market in neighborhood A clears with positive rents, that is,  $R_{At} > 0$  for all  $t$ , which requires also  $S_{At} > S_{Bt}$  for all  $t$ .<sup>30</sup>

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*

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

and

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

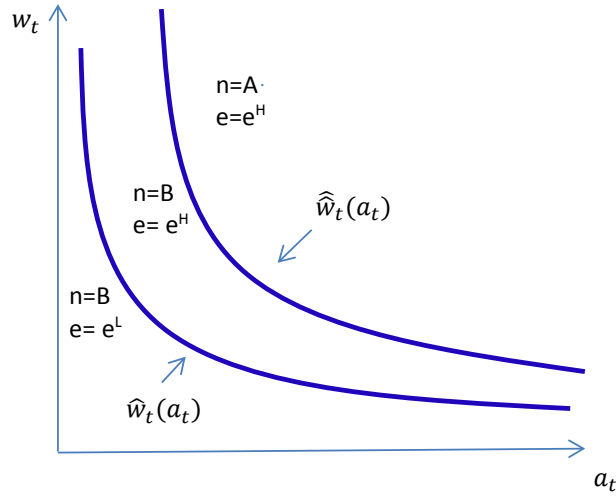
This proposition shows that in equilibrium the residential and the educational choices can be simply characterized by two monotonic cut-off functions.<sup>31</sup>

Figure 8 shows a graphical characterization of the equilibrium, for given spillovers and rental rates, with  $R_{At} > 0$ . The x-axis shows the children's ability level  $a_t$  and the y-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 pay for a high level of education for their children, parents with wage  $\hat{w}_t(a_t) \leq$

<sup>30</sup>If  $S_{At} \leq S_{Bt}$ , nobody would like to live in A and the rental rate in A would be zero.

<sup>31</sup>Assumptions 1 and 2 are needed to obtain the monotonicity result.

Figure 8: Equilibrium Characterization



$w_t < \hat{w}_t(a_t)$  choose to live in B and pay for a high level of education, and parents with wage  $w_t \geq \hat{w}_t(a_t)$  choose to live in A and pay for a level of 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 (cost  $\tau$ ) and the cost of a higher local externality (higher rental rate). On the other hand, for 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 random 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 richest agents residing in one and all the poorest 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  $\hat{w}_t(a_t)$  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 baseline 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 the high- and the low-level 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  $\hat{w}_t(a_t)$ . 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  $e^L = 0$ ,  $e^H = 1$ , and make the following functional form assumptions:  $u(c) = g(c) = \log(c)$ , and

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

On the one hand, this implies that the wage of a child with low education ( $e_t = 0$ ) is simply equal to  $bw^\alpha \varepsilon_t$  and does not depend on either the child's ability or the size of the neighborhood spillover, satisfying assumption 1. On the other hand, 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  $\beta_1$  and  $\xi$  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.<sup>32</sup> In this case, the two cut-offs are:

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

and

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

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 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 is the return to education in that neighborhood, and the higher is the willingness of parents living there to pay for their children's education. It also shows that, as expected, for 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,  $\eta$ ,  $\beta_0$ ,  $\beta_1$ , and  $\xi$  are higher, and when the cost of education  $\tau$  and/or the fixed component of the income of low-educated children  $b$  are lower. Equation (9) shows that also the residential cut-off  $\hat{\hat{w}}_t(a)$  is decreasing in  $a$ , again in line with Proposition 1, as the return to the 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 relative to 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. Under the lens of the model, we can think of an increase in the skill premium as an increase in the parameter  $\eta$  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, inequality mechanically increases for two reasons. First, the income gap between college 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  $\eta$  as a

---

<sup>32</sup>This case arises when the RHS of equation (8) is weakly smaller than the RHS of equation (9) for all ability levels. When instead 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  $\hat{w}_t(a) = \hat{\hat{w}}_t(a) = (\tau + R_{At})[1 + b/a\eta(\beta_0 + \beta_1 S_{At}^\xi)]$ .



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  $\eta$  as follows:

$$\frac{d\hat{w}_t(a_t)}{d\eta} = -\frac{1}{\eta^2} \frac{\tau b}{a_t(\beta_0 + \beta_1 S_{Bt}^\xi)}, \quad (10)$$

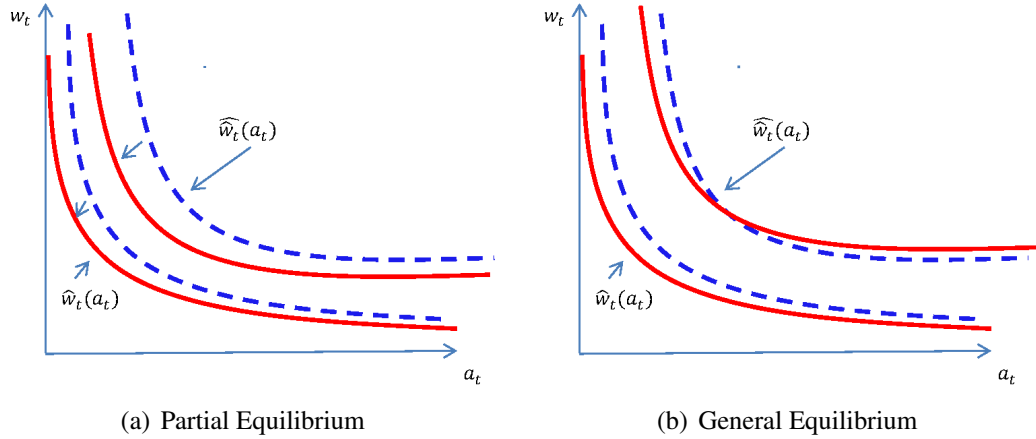
and

$$\frac{d\hat{w}_t(a_t)}{d\eta} = -\frac{R_{At}b}{\eta^2 a_t \beta_1 (S_{At}^\xi - S_{Bt}^\xi)} + \frac{b + a_t \eta (\beta_0 + \beta_1 S_{At}^\xi)}{a_t \eta \beta_1 (S_{At}^\xi - S_{Bt}^\xi)} \frac{dR_{At}}{d\eta}. \quad (11)$$

These derivations show that in partial equilibrium, that is, when the rental rate is kept fixed, 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 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 9 shows qualitatively 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 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  $d^2 \hat{w}_t(a_t) / da_t d\eta > 0$  and  $d^2 \hat{r}_t(a_t) / da_t d\eta > 0$  if  $dR_{At} / d\eta = 0$ . This means that, with our functional form, the marginal impact of ability on the return to education is larger when the skill premium is smaller.

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 9 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 pushes up rental rates in that

Figure 9: Cut-off Response to Skill Premium Shock



neighborhood, shifting the housing demand back to the right. In particular, the figure shows that the shift back is more pronounced for the 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 9, 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. 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 kept the spillover size in the two neighborhoods as given and showed that if a skill premium shock hits a segregated economy, the degree of segregation by income increases and the response of inequality is amplified because of that. However, in our model the spillover sizes in the two neighborhoods respond endogenously to the shock. The increase in  $\eta$  increases human capital of all the educated children, increasing average human capital 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 human capital even more in that neighborhood. Moreover, the tilting of the residential cutoff implies that neighborhood A will be populated by richer but less talented children. This has two effects. First, it tends to increase the gap between

human capital in the two neighborhoods, given that, everything else equal, children of richer parents tend to accumulate higher levels of human capital increasing  $S^A$ , while children of poorer parents tend to accumulate less human capital partially offsetting the increase in  $S^B$  due to the increase in education. Second, it tends to decrease the same gap, given that more talented children move away 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, the US experienced a steady increase in labor income inequality starting in 1980. 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 source of inequality (see, for example, Katz and Murphy, 1992; Autor, Katz and Krueger, 1998; Goldin and Katz, 2001; Card and Lemieux, 2001; Acemoglu, 2002; Autor, Katz and Kearney, 2008).

In this section, we explore the quantitative response of the economy to an unexpected, one-time, permanent shock to the skill premium, as described in subsection 3.3. In the next section, we will use the model to quantify the contribution of residential segregation by income to the increase in income inequality experienced in the US after the '80s.

### 4.1 Extended Model

For the quantitative analysis, we use the specific functional forms for the utility and the wage dynamics function that we have described in subsection 3.3, that is,  $u(c) = g(c) = \log(c)$ , and  $\Omega$  satisfying equation (7). Moreover, we extend the model in two main dimensions: first, we introduce a residential preference shock and, second, we make the educational choice continuous.

The introduction of the preference shock is important to 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 baseline 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 purely driven by educational considerations. Families may prefer more expensive neighborhoods for a number of different reasons, such as better amenities, or higher status. By missing this feature of reality, the baseline model might generate a distribution of children growing up in neighborhood A biased towards too high ability. We then assume that utility from current consumption is now given by  $\log[(1 + \theta I_{n=A})c]$ , where  $\theta \in \{0, \bar{\theta}\}$  is a preference shock with  $\bar{\theta} \geq 0$  and  $\pi = \text{Prob}(\theta = \bar{\theta})$ , so that families with  $\theta = \bar{\theta}$  enjoy their consumption more if they live in neighborhood A.

The other change relative to the baseline model is the assumption that the educational choice is continuous, which we believe is particularly important given the nature of our mechanism. In the baseline 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.<sup>33</sup> This means that the binary choice would arbitrarily bound the possible increase in the spillover in response to a skill premium shock.<sup>34</sup> We then assume that the educational choice is continuous, with  $e \in R_+$  and that the cost of education is quadratic,  $\tau e^2$ .

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

$$U(w, a) = \max_{e, n} \log((1 + \theta I_{n=A})(w - R_{nt} - \tau e^2)) + \log(b + ae(\beta_0 + \beta_1 S_{nt}^\xi))\varepsilon. \quad (\text{P3})$$

In order to understand better the role of the educational choice, let us, for a moment, shut down the preference shock, that is, set  $\hat{\theta} = 0$ . In this case, the first order condition for the educational choice gives

$$e(w, a|n) = \frac{w - R_{nt}}{2\tau} - \frac{b}{2a(\beta_0 + \beta_1 S_{nt}^\xi)},$$

where  $e(w, a|n)$  is the educational choice of a parent with wage  $w$  and a child with innate ability  $a$  conditional on living in neighborhood  $n$  and on  $e(w, a|n)$  being positive. The expression shows that, as expected, education is increasing in income  $w$ , innate ability  $a$ , and in the size of the local spillover  $S_{nt}$ . It is also increasing in  $\beta_0$ ,  $\beta_1$ , and  $\xi$ , which affect the return to education.

<sup>33</sup>If one calibrates the baseline model would obtain too much intergenerational mobility, given that the only way parents can invest in their children's education is to pay for college, but the data discipline the skill premium and hence bound how much rich parents can pay to increase their children's expected income.

<sup>34</sup>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 and poor families.

Moreover, it is decreasing in the cost of education  $\tau$  and in  $b$ , which is the average wage of non-college educated workers.

The equilibrium definition is a natural extension of Definition 1 in Section 3.2, except that in the extended model the policy functions also depend on the preference shock. Moreover, while the policy for the residential choice can still be represented by a cut-off function, the policy for the educational choice is going to be continuous.

## 4.2 Calibration

We now describe our calibration strategy. As 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  $\eta$  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 human capital accumulation, we choose to interpret one period as 10 years.<sup>35</sup> We interpret period  $t = 0$  as 1980, when the economy is in steady state. Then, we assume that at that time an unexpected, permanent shock hits  $\eta$  that becomes  $\eta' > \eta$ , where  $\eta'$  is such that the skill premium goes from 0.30 in 1980 ( $t = 0$ ) to 0.45 in 1990 ( $t = 1$ ), using the estimates, based on CPS data, from Valletta (2018). We describe below how we map the skill premium to the model.

We choose parameters so that the steady state equilibrium of the model matches salient features of the US economy mostly in 1980.<sup>36</sup> Table 1 shows the targets of our baseline calibration, which we are now going to discuss.

The first two targets are the 1980 values of the Gini coefficient and the dissimilarity index that we have described in Section 2, as baseline measures of inequality and income segregation for the

---

<sup>35</sup>The schooling age could be interpreted as 10 or 15 years depending on which level of education one targets. We choose 10 years also considering that Census data are available every 10 years.

<sup>36</sup>Below we explain that the data available for the rank-rank correlation and the neighborhood exposure effect give us only one data point that we interpret as an average between 1980 and 2000.

Table 1: *Calibration Targets*

Description	Data	Model	Source
Gini coefficient	0.366	0.365	Census 1980, family income
Dissimilarity index	0.318	0.318	Census 1980, family income
$H^R$ index	0.100	0.094	Census 1980, family income
B/A average income	0.516	0.459	Census 1980
$R^A-R^B$ normalized	0.073	0.074	Census 1980
Rank-rank correlation	0.341	0.330	Chetty et al. (2014)
Return to spillover 25th p	0.104	0.104	Chetty and Hendren (2018b)
Return to spillover 75th p	0.064	0.070	Chetty and Hendren (2018b)
Return to college 1980	0.304	0.306	Valletta (2018)
Return to college 1990	0.449	0.449	Valletta (2018)

average metro area.<sup>37</sup> As described in Section 2, we use Census data to calculate both the Gini coefficient and the dissimilarity index at the metro level and then we average them across metro areas, weighting by population. We have also discussed that there are alternative measures of income segregation that are used in the literature. In particular, we have shown another measure that has also been widely used in the more recent literature, which is the  $H^R$  index we introduced in Section 2. Given that this index measures segregation using the entire income distribution, we include it as an additional target.

We also want our model to capture the relative average income across neighborhoods. Given that we have two neighborhoods, we divide the census tracts in each metro area in two groups that correspond to neighborhoods A and B in the model. In order to do so, for each MSA, first we rank the census tracts by average income. Then, we look at their population and define neighborhood A as the richest census tracts with population above the 10th percentile (given that this is the percentile closest to the calibrated value of  $M$ , which is the size of neighborhood A), and define neighborhood B as the remaining ones. Finally, we calculate the average income in these two fictitious neighborhoods for each MSA, and then we average them across MSAs weighting by population to obtain the average income in A and in B. The ratio between these two values is the targeted moment.

Another important object in our model is the relative cost of housing in the two neighborhoods. We use housing values at the census tract level from the Census data and convert them to rental

<sup>37</sup>For the calibration we use our baseline dissimilarity index, where we define rich the households in the top 20th percentile of the metro income distribution, and poor the others.

rates.<sup>38</sup> Using the same methodology described above to aggregate census tracts, we calculate the difference between rental rates in A and B at the MSA level and normalize that by the average MSA income. We then average the normalized difference across MSAs weighting by population.

Another feature of the US data we want to 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).<sup>39</sup> In particular, they use children born between 1980 and 1982, 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 the 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 returns 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 counties.<sup>40</sup> Their baseline estimation 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 10%. This number becomes 6.4% for a child with parents at the 75th percentile of the income distribution. These are the values that we target in our calibration. Let us explain how we map these targets to our model. Given that in our model nobody literally moves, we map the “movers” in Chetty and Hendren

---

<sup>38</sup>We use a standard coefficient of 0.05 for the conversion.

<sup>39</sup>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 chose this statistic instead than the log-log correlation or other measures because Chetty et al. (2014) argue that it provides a more robust summary of intergenerational mobility.

<sup>40</sup>Chetty and Hendren (2018b) control for selection effects by looking at families who move from one county to another with kids of different age, so that 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.

(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 own parents. Then, we calculate the difference between the expected future income of the children of “movers” at the 25th percentile and at the 75th percentile of the income distribution if they grew up in neighborhood  $A$  and the expected future income of the same children if they grew up in neighborhood  $B$ . We then divide these numbers by the standard deviation of the spillover’s size  $S^n$  across the two neighborhoods.<sup>41</sup> Given that these children are born between 1980 and 1986, they will be in pre-Kindergarten to 12th grade, and hence exposed to the local spillover, in 1984-1998 and 1990-2004. Hence, 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, where again the values of 1990 and 2000 are calculated after the shock.

Finally, the last targets in table 1 are the US skill premia in 1980 and 1990, which are calculated in Valletta (2018) using CPS data. 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  $\hat{e}$  such that individuals with an education level above  $\hat{e}$  are college educated, and the ones with education below are not. We choose  $\hat{e}$  so that, in 1980, 17.8% of the population is college educated, as in the Census data.<sup>42</sup> Finally, we map the skill premium in 1990 to the same difference between the average log wage of college-educated individuals and the average log wage of the others one period after the shock, keeping the college cut-off  $\hat{e}$  constant. Notice that we target the skill premium also in 1990, after the shock, because we are simultaneously calibrating the parameters of the model and the size of the shock. We need to do that in order to ensure that the average of the rank-rank correlation and of the spillover effects over 1980-2000 match the targets.<sup>43</sup>

Table 2 shows the parameters that we are using to calibrate the model, their calibrated value, and their description. We normalized the value of  $\eta$  in steady state to 1. Notice that the number of

---

<sup>41</sup>This is simply equal to  $\sqrt{M(1-M)}(S^A - S^B)$ , where  $M$  is the housing supply in the rich neighborhood and  $S^A$  and  $S^B$  the steady state level of the spillover’s size in the two neighborhoods.

<sup>42</sup>To calculate this number, we look at the number of people above 25 year old who completed college at the census tract level.

<sup>43</sup>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, and using the rank-rank correlation from Chetty et al. (2014) and the spillover effects from Chetty and Hendren (2018b) as if they were numbers for the 1980. We do not believe this would be reasonable. Moreover, this alternative calibration would generate a larger increase in the implied spillover effect after the shock, so our choice is conservative.



Table 2: *Parameters*

Parameter	Value	Description
$M$	0.08	Size of neighborhood A
$\alpha$	0.24	Wage function parameter
$\beta_0$	2.09	Wage function parameter
$\beta_1$	0.27	Wage function parameter
$\xi$	0.80	Wage function parameter
$\tau$	0.32	Cost of education
$b$	1.61	Wage fixed component for no-college
$\rho$	0.39	Autocorrelation of ability
$\sigma$	0.48	Standard dev. of log innate ability
$\mu_a$	-3.33	Average of log innate ability
$\mu_\varepsilon$	0.41	Average of log wage noise shock
$\sigma_\varepsilon$	0.65	Standard dev. of log wage noise shock
$\bar{\theta}$	0.05	Preference shock value
$\pi$	0.33	Preference shock probability
$\eta'$	3.57	skill premium shock

parameters is higher than the number of targets because the model is highly non-linear.

### 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, is hit by an unexpected, one-time, permanent increase in  $\eta$ .

Table 3 shows the response of the economy to such a shock, one, two, and three periods ahead. In particular, we show the behavior of the return to college, the Gini coefficient, the dissimilarity index, the  $H^R$  index, the relative income and the relative housing rent in the two neighborhoods, the rank-to-rank correlation, and the ratio of the spillover in neighborhood A over the one in neighborhood B.

The first row shows the dynamics of the return to college. Remember that we chose the shock to 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, similarly to the data. In particular, the return to college in the CPS data is equal to 0.52 and 0.57 respectively in 2000 and 2010, which is very close to the predicted

Table 3: *Response to a Skill Premium Shock*

	t = 0	t = 1	t = 2	t = 3
Return to college	0.306	0.449	0.516	0.548
Gini coefficient	0.365	0.395	0.413	0.424
Dissimilarity index	0.318	0.397	0.404	0.405
$H^R$ index	0.094	0.129	0.136	0.140
$B/A$ average income	0.459	0.318	0.271	0.246
$R^A-R^B$ normalized	0.074	0.191	0.300	0.380
Rank-rank correlation	0.252	0.343	0.394	0.417
$A/B$ spillover ratio	1.229	1.816	2.146	2.340

path in our model.

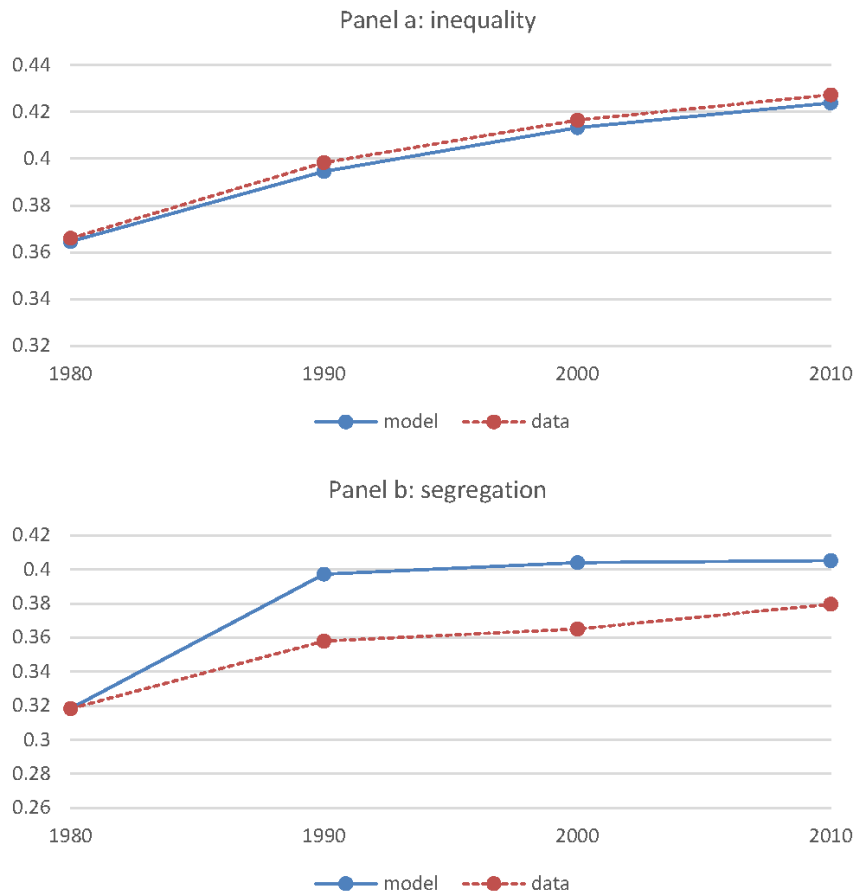
The second and third rows show the response of inequality and segregation, captured by the Gini coefficient and the dissimilarity index. To visualize these results, Figure 10 shows the model response (solid lines) of inequality and segregation to the shock together with their pattern in the data (dashed lines). As explained in subsection 3.3, there are several effects at work behind these dynamics, and the figure shows the quantitative response of inequality and segregation over time in response to the skill premium increase. Although the model is stylized in many dimensions, these responses are in the ballpark of what happened in the data, which is a reassuring validation of the model, given that we do not target these dynamics in the calibration.

Panel a shows the dynamics of inequality in response to the skill premium shock. While the value in 1980 is one of the targets, the path of inequality after 1980 is an outcome of the model. The figure shows that our model generates inequality dynamics very close to the data not only in terms of levels, but also in terms of concavity of the dynamics. Panel b shows that the model generates a response of segregation to the skill premium shock that is in the ballpark of the data, although a bit higher.

Table 3 also shows that the model predicts that the other measure of segregation by income, the  $H^R$  index, increases as well in response to the skill premium shock, similarly to its pattern in the data, where it goes from 0.10 in 1980 to 0.12 in 2010. Another symptom of the increase in segregation is the fact that the relative average income in neighborhood B decreases in response to the shock.<sup>44</sup>

<sup>44</sup>We can see the same qualitative pattern in the data, although it is quantitatively less pronounced.

Figure 10: Responses to a skill premium shock



As we have 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 ones with no college. Given that more educated workers are more likely to grow up in neighborhood A, segregation by income also mechanically increases. On top of that, as the return to education increases and given the complementarity between the neighborhood spillover and education, the return to live in neighborhood A increases, pushing up the housing rental rate differential and hence increasing segregation by income even further. Table 3 shows that in fact the differential rental rate in the two neighborhood increases in response to the shock and keeps increasing as the feedback effect kicks in. On top of that, an important feature of our model is that the sizes of the spillover effects,

that is, the average human capital in the two neighborhoods, are endogenous. As the return to college increases, average human capital increases in both neighborhoods, but, as table 3 shows, it increases more in neighborhood A. This is because of the increase in segregation by income and the fact that children from richer families tend to accumulate higher levels of human capital. It is exactly such an increase in the spillover differential between the two neighborhoods that feeds back into even higher inequality.

Finally, the table shows that intergenerational mobility decreases, as highlighted by the rank-rank correlation going from 0.25 in 1980, to roughly 0.42 in 2010. This is a natural effect of the main mechanism in the model: as living in neighborhood A becomes more expensive, richer families move to A exposing their children to stronger local spillover effects, while poorer families are forced to move out. This, inevitably makes it more difficult for poor children to climb up the social ladder and easier for the richer children to perpetuate their status. Unfortunately, given the limited availability of data, it is hard to calculate a reliable time-series for the rank-rank correlation.<sup>45</sup> However, Chetty et al. (2017) show some evidence that intergenerational mobility has declined in the last half century, looking at the fraction of children earning more than their parents. Moreover, 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.

## 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? In the next subsection, we propose the exercise that we believe is better suited to answer this question. Next, we perform a battery of alternative exercises that take different perspectives and also help us to better understand the quantitative role of different channels in the model.

---

<sup>45</sup>This is the reason why we target the average value between 1980 and 2010.

## 5.1 Main Counterfactual Exercise

We are now ready to show the main counterfactual exercise that quantifies the role of segregation in amplifying inequality. In the model, the presence of local spillovers generates sorting of richer parents into the better neighborhood, 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 keeping fixed the strength of the spillovers effects in the two neighborhoods, the increase in income segregation implies that more rich children will benefit from the exposure to the better neighborhood and will become even richer, while more poor children will be forced to grow up in the worse neighborhood, 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 neighborhood A relative to neighborhood B, further increasing inequality.

One natural way to assess the contribution of segregation to inequality is to shut down the residential choice of the households in response to the shock, which is going to mute the sorting process. In particular, we consider a counterfactual exercise where at the moment of the shock and at any time after that, families are forced to be randomly re-located to the two neighborhoods.<sup>46</sup> This implies that the size of the local spillover effects in the two neighborhoods,  $S^A$  and  $S^B$ , is equalized, given that the distribution of families in the two neighborhoods is identical, and hence the human capital accumulated in the two neighborhoods is identical as well.<sup>47</sup> So, children are still exposed to a human capital accumulation externality that evolves over time, but the strength of the spillover is the same for everybody, so, in a sense, this becomes a global instead of local spillover. The two neighborhoods are the same and parents' income affects children's wages only through the direct effect on the wage function  $\Omega$  and through the educational choice. This mitigates the effect of intergenerational linkages on income and hence mitigates the response of income inequality to a skill premium shock.

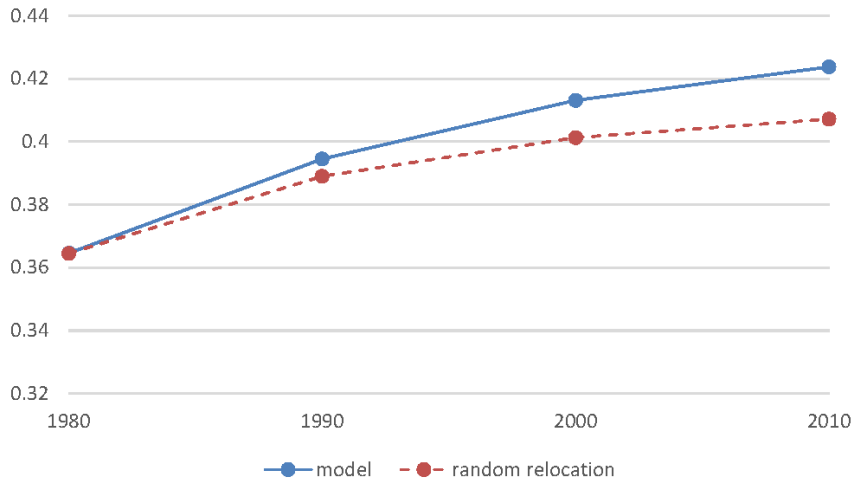
Figure 11 compares the response of inequality to the skill premium shock in the baseline model (blue solid line) to the response of the economy when families are randomly re-located between

---

<sup>46</sup>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.

<sup>47</sup>In the exercise, we impose that the rental rate in neighborhood A is also equal to zero, given that there is no endogenous residential choice and the rental rate in A would be undetermined.

Figure 11: Inequality: counterfactual with random re-location



the two neighborhoods every period after the shock (red dashed line). The figure shows that segregation contributes to 18% of the increase in inequality one period after the shock, that is, between 1980 and 1990 (short run from now on), and to 28% of the increase in inequality over the whole period between 1980 and 2010 (long run from now on). In turn, this exercise shows that segregation also amplifies the decrease in intergenerational mobility in response to the shock. In particular, it contributes to 12% of the increase in the rank-rank correlation between 1980 and 2010.

In this counterfactual exercise, the increase in inequality is dampened because families do not segregate by income and so, there are no differential spillover effects depending on the neighborhood where children grow up. However, in this exercise families cannot even segregate by ability, which could have a positive effect on intergenerational mobility. In future work, it would be interesting to consider more explicit policy experiments that help disentangle the effect of segregation by income with respect to the ability dimension.<sup>48</sup>

In order to check how sensitive our main result is to the exact estimation of the neighborhood exposure effects in Chetty and Hendren (2018b), we also perform some sensitivity analysis. In

<sup>48</sup>One possibility would be to consider house subsidies targeted to low-income families who live in a given location.

particular, we re-calibrate the model to match the same targets as in the benchmark calibration, except that we target a neighborhood exposure effect 30% lower and 30% higher than the values from Chetty and Hendren (2018b). We then perform the same counterfactual exercise for both recalibrations and show that segregation would explain 22% of the increase in inequality between 1980 and 2010 if the neighborhood exposure effect was 30% lower than in the benchmark, and 38% if it was 30% higher. These results show that, although the estimates of the neighborhood exposure effect determines the exact contribution of segregation to inequality, the overall effect remains sizeable even with substantially lower estimates.

## 5.2 Understanding the Mechanism

In the previous counterfactual exercise we explored what would happen to inequality if families were randomly re-located, in order to shut down the endogenous sorting due to the different local spillovers. As we explained above, this exercise can be interpreted as exploring what would happen if spillover effects were present, but their nature was global instead of local. This is our favorite counterfactual to quantify the effect of segregation on inequality. However, there are alternative approaches to assess the contribution of local spillovers to inequality. Next, we show two different and complementary exercises: first, we explore what would happen to inequality if there were no local spillovers at all, and, second, if there were local spillovers, but they were not responsive to the shock. The reason we prefer the previous counterfactual is that we believe that the first exercise is too extreme, as it shuts down any form of human capital externality, and that the second exercise only highlights a piece of the mechanism, as it focuses only on the endogenous response of the spillovers' strength. However, these additional counterfactual are also informative and help decomposing the different channels through which local spillovers affect inequality.

The first additional exercise is to consider an economy where there are no spillover effects at all, that is, where the wage function  $\Omega$  is not affected by the spillover, or  $\beta_1 = 0$ . In this case, the only difference between the two 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. This generates some degree of segregation in the initial steady state that is purely driven by income: richer families would be the only ones willing to pay to live in the

better neighborhood. The blue solid lines in figure 12 report the response of the baseline model to the skill premium shock as in subsection 4.3, while the dotted green lines show the response to the same shock of the economy with  $\beta_0 = 1$  and all the other parameters unchanged. Panel b shows that segregation in steady state is of the same order of magnitude than in the baseline model, even if a bit lower, but, most important, it does not respond to the shock, as the preference shock is random and not correlated with the returns to education. The green dotted line in panel a represents the correspondent response of inequality to the shock and shows that inequality would increase much less in response to the shock if there were no local spillovers. In particular, we interpret the distance between the blue solid lines and the green dotted lines as the contribution of the existence of local spillovers to the increase in inequality. The figure shows that the existence of local spillover effects contribute to 53% of the increase in inequality in the model in the short run and to 65% in the long run.<sup>49</sup> This exercise is quite extreme as it rules out not only local spillovers, but any type of human capital accumulation externality.

The other exercise that we consider aims to assess 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 not changing endogenously, that is, if  $S^A$  and  $S^B$  were fixed at their initial steady state levels. The red dashed lines in panel a and b of figure 12 show the responses of inequality and segregation respectively in this exercise. When we compare them to the baseline model responses, 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 21% and 35% of the increase in inequality respectively in the short and in the long run. Moreover, the contribution to the increase in segregation is 47% in the short run and 23% in the long run. We can also use this exercise to quantify the relevance of exogenous differences in local spillovers. We can interpret the difference between the red dashed lines and the green dotted lines as the amplification that would arise if the spillover effects were exogenous. The figure shows that roughly 40% of the amplification of inequality due to spillover effects can be attributed to the exogenous component and the rest (roughly 60%) to the endogenous feedback.

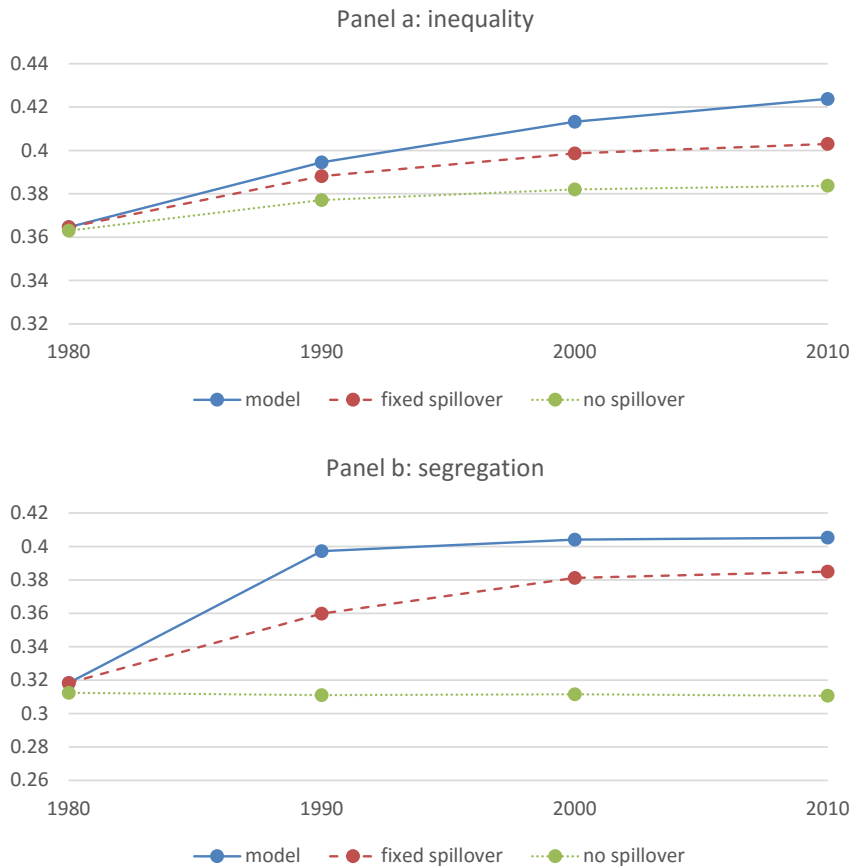
Why does the endogenous change in the spillover effects further amplify inequality? The local

---

<sup>49</sup>Figure 12 is realized without re-calibrating the parameters to focus on the decomposition of the response in the model. In the next subsection we will consider a different exercise where we re-calibrate the model.



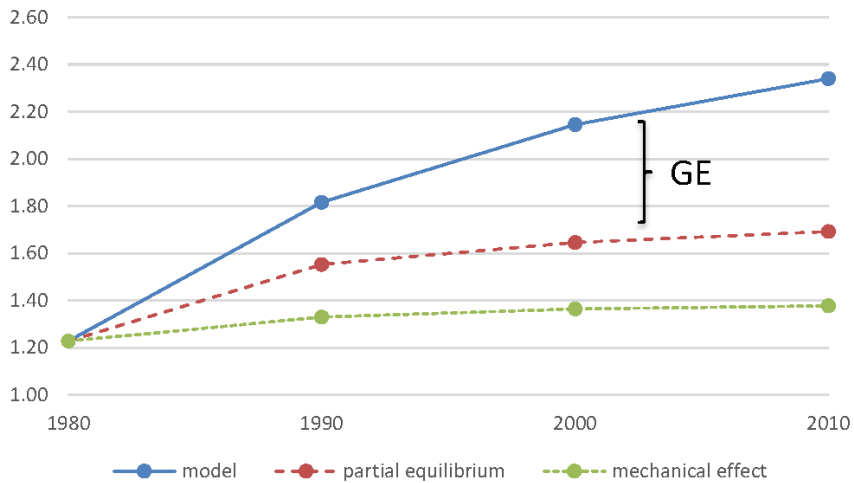
Figure 12: Counterfactuals with no spillover and no spillover feedback



spillover effects in both neighborhoods,  $S_t^A$  and  $S_t^B$ , increase in response to the skill premium shock, given that all college educated workers have higher wages and, moreover, everybody invests more in education. However, as we have already emphasized, the strength of the local spillover in neighborhood  $A$  increases relatively more than in neighborhood  $B$ . The last line in table 3 reports the model increase in the spillover ratio  $S_t^A/S_t^B$  in response to the shock, which is illustrated by the solid blue line in Figure 13.

We can decompose the response of  $S_t^A/S_t^B$  to the skill-premium shock in three effects. First, there is a mechanical effect: children in the rich neighborhood benefit more from the increase in the skill premium because they are more highly educated and are exposed to the stronger spillover effect in neighborhood  $A$ . This mechanically increases the level of their human capital, and hence

Figure 13: Decomposing the increase in the spillover ratio



the strength of the spillover in neighborhood *A* more than in neighborhood *B* where children have lower level of education and are exposed to a weaker spillover effect. The green dotted line in the figure shows the increase in the spillover ratio 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 neighborhood *A*, do it even more, increasing the gap in returns. The dashed red line in Figure 13 shows the increase in the spillover ratio due the sum of the mechanical effect and this effect coming from the endogenous change in educational and residential choice, which represents the partial equilibrium effect. Finally, there is a general equilibrium effect coming from the increase in the rental rate in the rich neighborhood, 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 the rich neighborhood, irrespective of their children’s ability. This further raises the gap between the spillovers’ strengths in the two neighborhoods. We can interpret the difference between the blue solid line and the red dashed line as the contribution 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 58% of the increase in  $S_t^A / S_t^B$  between 1980 and 2010.

### 5.3 Model Comparison Exercise

Finally, an alternative perspective in evaluating the relevance of neighborhood externalities for the dynamics of US inequality is to consider an alternative model that is the same of our model with the only exception that does not feature the local externality, that is, the same model with  $\beta_1 = 0$ . This exercise is similar to one of the counterfactual exercises we considered in the previous subsection, with the difference that now we re-calibrate the model. So, the question that we ask is slightly different: how much more inequality increase our model predicts relative to the same model with no spillover effects, in response to a comparable skill premium shock? The difference with the green dotted lines in Figure 12 is that now we re-calibrate the model using exactly the same targets as in Table 1, except for the ones that are not relevant in the new model, that is, except for the two values for the neighborhood exposure effects from Chetty and Hendren (2018b). We then consider a comparable skill premium shock, that is, a one-time unexpected permanent increase in  $\eta$  that matches the increase in the skill premium in the new model to the increase in the data between 1980 and 1990, as we did for the baseline calibration.

Figure 14: Inequality: model with no spillover effects

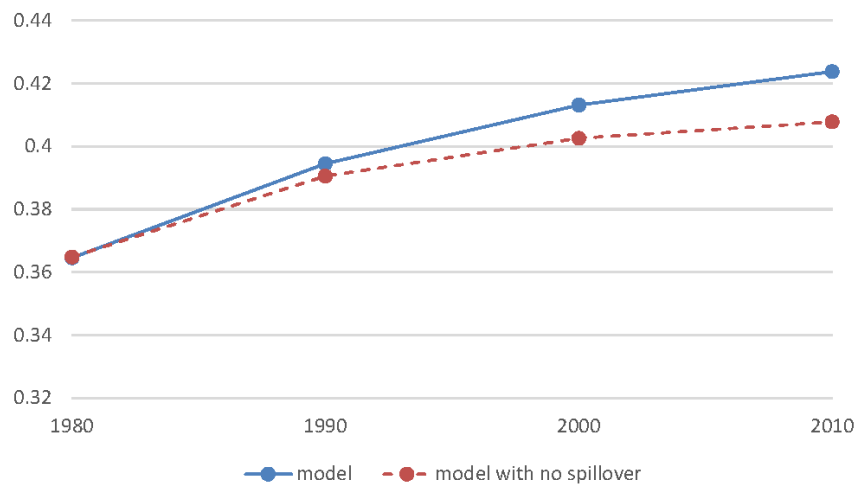


Figure 14 compares the response of inequality to the skill premium shock in the two models, where the blue solid line represents our model and the red dashed line represents the alternative model with no spillover externality. The figure shows that in the model with no spillover effects on the returns to education, inequality increases 14% less in the short run and 28% less in the long run relative to our benchmark model.

Overall, we have performed a number of different exercises that try to assess from different angles the importance of local spillover effects and the resulting income segregation on amplifying the response of inequality to a skill premium shock. Our preferred exercise is the main counterfactual in section 5.1 that implies that segregation contributed to 28% of the increase in inequality between 1980 and 2010. The decomposition exercises in section 5.2 show that 65% of the increase in inequality generated by our model is due to the presence of spillover effects and 35% to the feedback effect due to the endogeneity of the spillovers' strengths. Finally, our last exercise shows that a model with no spillover effects would generate an increase in inequality 28% smaller than our benchmark model.

## 6 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. We calibrate the model using US data in 1980, and using the micro estimates of neighborhood externalities that Chetty and Hendren (2018b) proposed using administrative data. We then hit the economy with an unexpected permanent shock to the skill premium and look at the responses over time of inequality, residential segregation, and intergenerational mobility. We use a number of counterfactual exercises to show the role of local spillovers and of the resulting segregation in amplifying inequality. Our main exercise shows that segregation contributes to 28% of the increase in inequality between 1980 and 2010, in response to a skill-biased technical change shock.

In work in progress, we are considering a model with multiple cities to exploit the cross-sectional richness of the data. This allows us to use the model to think about the correlation of inequality, segregation, and intergenerational mobility across metro areas and also to understand the differential responses to a common skill premium shock.

There are multiple interesting directions for future research. One interesting dimension to extend the model would be to endogenize the cost of education, that has also increased over time, and think about possible feedback effects. Another interesting avenue for future work would be to push the normative implications of the model and think about alternative policy experiments that exploit the spatial nature of the model.

## 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,” *UPenn, mimeo*, 2018.
- Alesina, Alberto and Ekaterina Zhuravskaya**, “Segregation and the Quality of Government in a Cross Section of Countries,” *American Economic Review*, 2011, 101 (5), 1872–1911.
- 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 H, Lawrence F Katz, and Alan B Krueger**, “Computing inequality: have computers changed the labor market?,” *The Quarterly Journal of Economics*, 1998, 113 (4), 1169–1213.
- Autor, David H., Lawrence F. Katz, and Melissa S. Kearney**, “Trends in US Wage Inequality: Revising the Revisionists,” *The Review of Economics and Statistics*, 2008, 90 (2), 300–323.
- Bartolome, Charles AM De**, “Equilibrium and inefficiency in a community model with peer group effects,” *Journal of Political Economy*, 1990, 98 (1), 110–133.
- 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,” *The Quarterly Journal of Economics*, 1993, 108 (3), 619–652.
- , “Equity and efficiency in human capital investment: the local connection,” *The 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, pp. 584–609.
- Bilal, Adrien and Esteban Rossi-Hansberg**, “Location as an asset,” *Princeton University, mimeo*, 2019.

- 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.
- Card, David and Thomas Lemieux**, “Can falling supply explain the rising return to college for younger men? A cohort-based analysis,” *The 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,” *The Quarterly Journal of Economics*, 2018, 133 (3), 1107–1162.
- **and** — , “The impacts of neighborhoods on intergenerational mobility II: County-level estimates,” *The Quarterly Journal of Economics*, 2018, 133 (3), 1163–1228.
- , **David Grusky, Maximilian Hell, Nathaniel Hendren, Robert Manduca, and Jimmy Narang**, “The fading American dream: Trends in absolute income mobility since 1940,” *Science*, 2017, 356 (6336), 398–406.
- , **Nathaniel Hendren, 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,” *The Quarterly Journal of Economics*, 2014, 129 (4), 1553–1623.
- Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst**, “Income Growth and the Distributional Effects of Urban Spatial Sorting,” *University of Chicago, mimeo*, 2019.
- Cutler, David M. and Edward L. Glaeser**, “Are Ghettos Good or Bad?,” *Quarterly Journal of Economics*, 1997, 112 (3), 827–872.
- , — , **and Jacob L. Vigdor**, “The Rise and Decline of the American Ghetto,” *Journal of Political Economy*, 1999, 107 (3), 455–506.
- 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.
- Duncan, Greg J and Richard J Murnane**, “Rising inequality in family incomes and children’s educational outcomes,” *RSF*, 2016.
- Durlauf, Steven N**, “Neighborhood feedbacks, endogenous stratification, and income inequality,” in *Dynamic Disequilibrium Modelling: Proceedings of the Ninth International Symposium on Economic Theory and Econometrics*, W. Barnett, G. Gandolfo, and C. Hillinger, eds., 1996.
- , “A theory of persistent income inequality,” *Journal of Economic growth*, 1996, 1 (1), 75–93.
- **and Ananth Seshadri**, “Understanding the Great Gatsby Curve,” in “NBER Macroeconomics Annual 2017, volume 32,” University of Chicago Press, 2017.

- Eckert, Fabian and Tatjana Kleineberg**, “Can We Save the American Dream? A Dynamic General Equilibrium Analysis of the Effects of School Financing on Local Opportunities,” *Yale University, mimeo*, 2019.
- 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,” *The 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, pp. 813–833.
- Ferreira, Pedro C, Alexander Monge-Naranjo, and Luciene Torres de Mello Pereira**, “Of Cities and Slums,” *St Louis Fed, mimeo*, 2017.
- Giannone, Elisa**, “Skill-Biased Technical Change and Regional Convergence,” *Penn 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,” *The causes and consequences of increasing inequality*, 2001, pp. 37–82.
- Guerrieri, Veronica, Daniel Hartley, and Erik Hurst**, “Endogenous gentrification and housing price dynamics,” *Journal of Public Economics*, 2013, 100, 45–60.
- Hsieh, Chang-Tai and Enrico Moretti**, “Why do cities matter? Local growth and aggregate growth,” 2015.
- 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.
- Loury, Glenn C**, “Intergenerational transfers and the distribution of earnings,” *Econometrica: Journal of the Econometric Society*, 1981, pp. 843–867.
- Massey, Douglas S and Nancy A Denton**, “The dimensions of residential segregation,” *Social forces*, 1988, 67 (2), 281–315.
- , **Jonathan Rothwell, and Thurston Domina**, “The changing bases of segregation in the United States,” *The 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,” *The American economic review*, 1991, 81 (5), 1408–1418.
- , *Social networks and persistent inequality in the labor market*, Center for Urban Affairs and Policy Research Evanston, 1991.
- Moretti, Enrico**, “Human capital externalities in cities,” in “Handbook of regional and urban economics,” Vol. 4, 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 Glenn Firebaugh**, “Measures of multigroup segregation,” *Sociological methodology*, 2002, 32 (1), 33–67.
- **and Kendra Bischoff**, “Income inequality and income segregation,” *American Journal of Sociology*, 2011, 116 (4), 1092–1153.
- Reardon, Sean F, Kendra Bischoff, Ann Owens, and Joseph B. Townsend**, “Has Income Segregation Really Increased? Bias and Bias Correction in Sample-Based Segregation Estimates,” *Demography*, 2018, *forthcoming*.
- Rothstein, Jesse**, “Inequality of educational opportunity? Schools as mediators of the intergenerational transmission of income,” *Journal of Labor Economics*, 2019, 37 (S1), S85–S123.
- 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.
- Valletta, Robert G**, “Recent flattening in the higher education wage premium: Polarization, skill downgrading, or both?,” in “Education, skills, and technical change: Implications for future US GDP growth,” University of Chicago Press, 2018.
- 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 No. 14908*, 2009.
- Zheng, Angela**, “Public Education Inequality and Intergenerational Mobility in an Overlapping Generations Model,” *mimeo*, 2017.



# Appendix (For Online Publication)

## A 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 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,  $HB$ ; 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  $\varepsilon$ , given that it is iid, so 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,

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

and

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

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 make respectively conditions (12) and (13) hold with equality for given  $a$ . We can then define the cutoff function

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

This proves that all agents with  $w \geq \hat{w}(a)$  choose the option  $HA$  for given  $a$ . Using 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,

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

and

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

Following analogous steps to before, we can show that, for 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 cutoff function

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

This proves that all agents with  $w \leq \hat{\hat{w}}(a)$  choose the option  $LB$  for given  $a$ . Using 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  $n = B$ .

## B Data Methodology

### B.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 4 reports our summary statistics at the metro and census tract level. 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 4: *Summary Statistics*

year	no_metros	no_CTs	ave_no_CTs
1980	379	42406	111.9
1990	380	48412	127.4
2000	380	53033	139.6
2010	380	59842	157.5

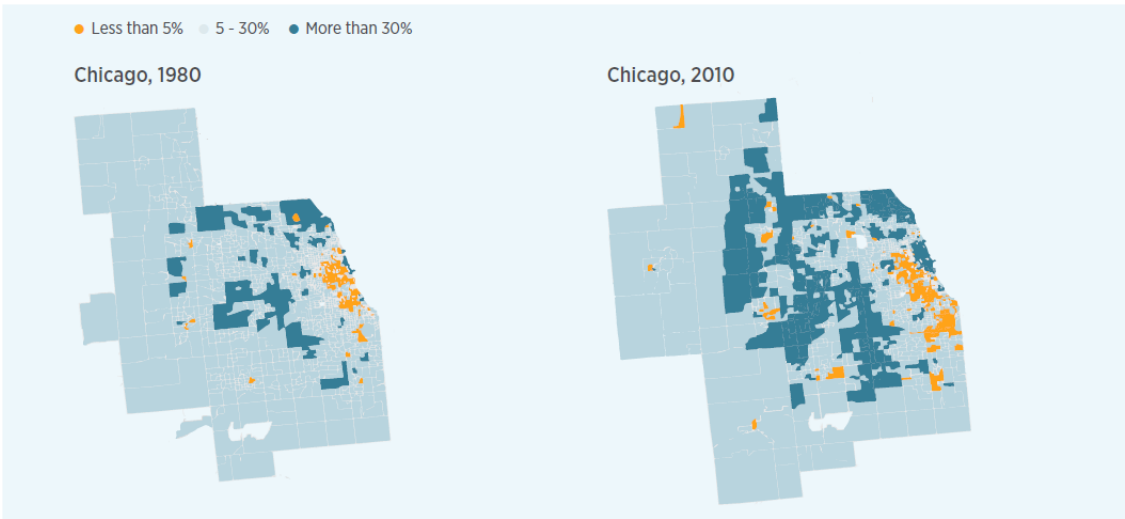
**Computing the Dissimilarity Index** The dissimilarity index uses the following formula

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

where  $X(j)$  and  $Y(j)$  denote the total number of, respectively, poor and rich families in metro  $j$ , while  $x_i(j)$  and  $y_i(j)$  denote the number of, respectively, 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 deemed "rich" and all families with a lower income are deemed "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.

**Example: Segregation in Chicago over time** The dissimilarity index captures the deviation from an even distribution of rich and poor households. Given that we define rich the households in the top 20 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 in all census tracts from the 20%, with weights given by the census tract population relative to the metro population. Figure 15 plots the share of rich households in each census tract of Chicago in 1980 and 2010. If there was no segregation, each census tract would have the same share of rich households equal to 20%. We associate to census tracts with a share of rich between 5 and 30 percent a light blue color. To visualize how segregation has changed over the period we use a heat map. In particular, we use dark blue to identify census tracts with a share of rich households higher than 30%, yellow for the census tracts with a share of rich households below 5%, and light blue for census tracts with a share of rich between 5% and 30%. We observe that over time the light color is replaced by darker colors, as several census tracts in Chicago display larger deviations from an even distribution of rich households in 2010 compared to 1980.<sup>50</sup>

Figure 15: Share of Rich Households in Chicago



<sup>50</sup>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.

**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-parametric CDF is calculated using the function `binsmooth`, provided by von Hippel et. al. in R. 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 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-parametric CDF function which can be used to calculate the Gini coefficient (and the conditional mean income of the top-coded bracket). Define:

$$\mu = \int xf(x)dx$$

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, there is little approximation error. Importantly, the  $\mu$  from the non-parametric CDF matches the empirical mean. After a Gini coefficient is calculated for each MSA, the weighted average of these coefficients is taken, using the count of family units in the MSA as weights.

**Computing the  $H^R$  Index** To construct the  $H^R$  index in a given metro, first we define the information theory index (or Theil index)  $H(p)$  that measures the segregation of rich and poor individuals across census tracts, where the rich are the individuals above the  $p$ -th percentile of the family income in the metro area and the poor are the others:

$$H(p) = 1 - \frac{1}{E(p)} \sum_{j=1}^J \frac{t_j}{T} E_j(p),$$

where  $E(p) = -[p \log(p) + (1 - p) \log(1 - p)]$  is the entropy at the metro level using  $p$  as the cutoff percentile,  $E_j(p)$  is the entropy at the census tract  $j$  level, and  $t_j/T$  is the share of the metro population in census tract  $j$ . Since income is available only in coarsened form (income data are reported into 16 income categories in U.S. census and ACS data), we follow Reardon and Bischoff (2011) and estimate  $H(p)$  by first computing  $H$  at the set of finite values that correspond to the percentiles of the thresholds used to coarsen the data; we then fit a polynomial function through the resulting points and then use the fitted polynomial as an estimate of  $H(p)$ . The resulting income segregation index is

$$H^R = \frac{1}{\int_0^1 E(p) dp} \int_0^1 E(p) H(p) dp = 2 \int_0^1 E(p) H(p) dp.$$

**Computing the Income Ratios** To calculate the various income ratios, the income brackets at the tract level are collapsed to the MSA level. The income brackets are then sorted by income level (by year and MSA) and the cumulative distribution of persons within each income bracket is calculated. Next, we find the income bracket associated with the bottom 10% of the population and the top 10% of the population and calculate the ratio of these two incomes. Because the income distribution is discrete, the exact cut-offs cannot be calculated. To deal with this, the cut-offs are defined to be the income level which had the minimum distance to the relevant threshold.

## B.2 Regression Analysis

To explore the relationship between segregation and inequality, we run several regressions. First, we regress the MSA level Gini coefficients on the MSA dissimilarity indices.<sup>51</sup> The results in Table 5 show a strong correlation between the two variables: metros with higher level of inequality in 1980 are also those that display higher level of segregation in the same year. This finding is robust to different definitions of the two variables and it holds for all decades in our sample. Additionally, we calculate the change in dissimilarity and inequality between 1980 and 2010 for each MSA, and regress the change in inequality on the change in dissimilarity. The results are presented in the last two columns of Table 5.

---

<sup>51</sup>For this analysis, define the cut-off between rich and poor as the 80th percentile. We used population weighting in the regressions, although the results do not change significantly if the observations are unweighted.

The coefficient of the change in inequality on the change in segregation is positive and statistically significant. The average change in the dissimilarity index was 0.061 and the average change in the Gini coefficient was 0.064. The results indicate that a 10% change in dissimilarity leads to a 1.7% change in the Gini. Similar results hold if we restrict to the 1980-2000 sample period.

To check the robustness of these results, we ran the same regressions including controls for the racial composition of each MSA, as well as employment shares by industry. 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. Most significantly, the magnitude of the coefficient using levels decreases by a factor of four, although it is still statistically significant. The coefficient on the differences also decreases, but much more modestly, going from 0.176 to 0.149.

Table 5: Regression Analysis

	<i>Dependent variable:</i>			
	Gini 1980		Change Gini: 1980 to 2010	
	(1)	(2)	(3)	(4)
Segregation	0.250*** (0.0151)	0.0724*** (0.0164)		
Change Dissimilarity: 1980 to 2010			0.176*** (0.0175)	0.149*** (0.0163)
Constant	0.300*** (0.00489)	0.677 (0.800)	0.0542*** (0.00125)	1.361* (0.649)
Race and Industry Controls	No	Yes	No	Yes
Observations	379	378	379	378
$R^2$	0.421	0.724	0.212	0.424
Adjusted $R^2$	0.420	0.713	0.209	0.401

*Note: Regressions use population weights.*

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

### **B.3 School District Analysis**

An important driver of segregation in US cities is the quality of public schools which mostly varies at the school district level and not necessarily at the census tract level. While our model abstracts from any specific channel that gives rise to neighborhood income segregation and proposes a more general and flexible approach, it is interesting to explore the evolution of income segregation at the school district level.

**Computing the Dissimilarity Index and H Index** The National Center for Education Statistic (NCES) collaborates with the U.S. Census Bureau to provide demographic data for school districts. Data is provided from the 1990 and 2000 Decennial Census, as well as the 2008-2012 American Community Survey. Data for 1980 was taken from the Census of Population and Housing, Summary Tape 3F and is provided by ICPSR. After combining these files, we calculate the dissimilarity index, H index, and adjusted H index using school districts as the relevant sub-unit. Figures 16 and 17 show 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. A similar pattern holds for the H indexes. 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 US (which have larger weight in our estimates). The increase in the share of children attending private school weakens the incentive to segregate across school districts lines.

**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 census tracts, the definitions are relatively stable over time. However, many states have seen a significant consolidation in school districts over time.<sup>52</sup> 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 instance, 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 as well as between schools. Using census tracts will reflect these changes, whereas using school districts would mask the latter trend. Table 6 reports summary statistics at the district level. The average number of districts in a metro is much smaller than the number of census tracts. Several metros only have one school district and 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 6: *Summary Statistics*

year	no_metros	no_SDs	ave_no_SDs
1980	379	6611	17.5
1990	379	6669	17.6
2000	379	6849	18.1
2010	380	6838	18.0

---

<sup>52</sup>For instance, Massachusetts saw a significant consolidation in 2009.

Figure 16: Inequality and Segregation over Time

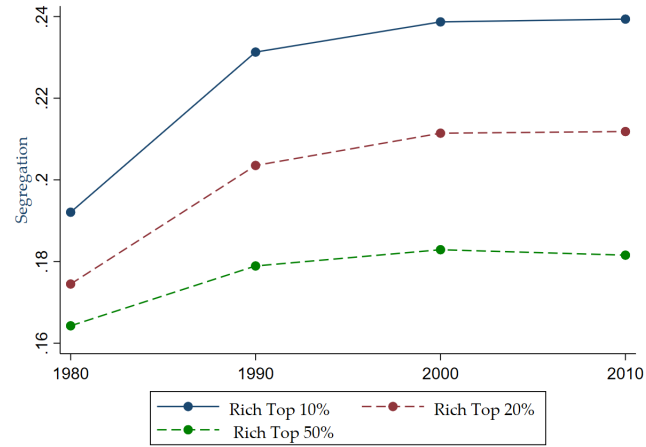


Figure 17: Inequality and Segregation over Time

