Journal of Urban Economics 82 (2014) 32–48

Contents lists available at ScienceDirect

Journal of Urban Economics

www.elsevier.com/locate/jue

Separate when equal? Racial inequality and residential segregation

Patrick Bayer [a,d,*], Hanming Fang [b,d], Robert McMillan [c,d]

[a] Department of Economics, Duke University, 213 Social Sciences Building, Box 90097, Durham, NC 27708, United States
[b] Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, United States
[c] Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada
[d] NBER, United States

ARTICLE INFO

Article history:
Received 8 May 2012
Revised 9 May 2014
Available online 24 May 2014

Keywords:
Segregation
Racial inequality
Neighborhood formation
Negative feedback loop
General equilibrium

ABSTRACT

This paper sets out a new mechanism involving the emergence of middle-class black neighborhoods that can lead segregation in American cities to increase as racial inequality narrows. The formation of such neighborhoods requires a critical mass of highly educated blacks in the population, and leads to an increase in segregation when those communities are attractive for blacks who would otherwise reside in middle-class white neighborhoods. To assess the empirical importance of this "neighborhood formation" mechanism, we propose a two-part research design. First, inequality and segregation should be negatively related in cross section for older blacks if our mechanism operates strongly, as we find using both the 1990 and 2000 Censuses. Second, a negative relationship should also be apparent over time, particularly for older blacks. Here, we show that increased educational attainment of blacks relative to whites in a city between 1990 and 2000 leads to a significant rise in segregation, especially for older blacks, and to a marked increase in the number of middle-class black communities. These findings draw attention to a negative feedback loop between racial inequality and segregation that has implications for the dynamics of both phenomena.

1. Introduction

At first glance, the relationship between racial inequality and residential segregation in American cities may seem obvious. Not only does segregation exacerbate racial inequality – Cutler and Glaeser (1997) show, for example, that young blacks have significantly worse education and labor market outcomes than young whites in more segregated cities.[1] Increased racial inequality also plausibly leads to greater levels of residential segregation as households sort across communities on the basis of education and income.[2] In this way, both directions of causality appear to give rise to a positive connection between residential segregation and racial inequality.

This standard intuition misses an important aspect of the sorting equilibrium in a city that can instead create a significant *negative* relationship between racial inequality and residential segregation. In particular, when racial inequality is substantial and black households make up a relatively small fraction of a city's population – conditions that hold in many American cities – it is impossible for highly educated (or high-income) black neighborhoods to form in equilibrium. As a result, highly educated blacks must generally choose between largely white middle-class communities or predominantly black poor communities.[3] In such circumstances, a decline in racial inequality has the potential to relax this binding neighborhood choice constraint. Specifically, as the number of highly educated blacks in the population increases beyond a critical mass, the formation of middle-class black neighborhoods becomes feasible.[4] If these neighborhoods prove to be an attractive alternative

* Corresponding author at: Department of Economics, Duke University, 213 Social Sciences Building, Box 90097, Durham, NC 27708, United States. Fax: +1 919 684 8974.

E-mail addresses: patrick.bayer@duke.edu (P. Bayer), hanming.fang@econ.upenn.edu (H. Fang), mcmillan@chass.utoronto.ca (R. McMillan).

[1] See also Ananat (2011).

[2] Given the strong correlation between race and other sociodemographic characteristics, Schelling (1969, 1971) noted that racial segregation would arise in the housing market even without explicit sorting on the basis of race. Sorting based on income is the focus of a number of papers, notably LeRoy and Sonstelie (1983), Glaeser et al. (2008), and Brueckner and Rosenthal (2009); Bayer and McMillan (2012) explore factors that lead to departures from Tiebout income stratification. The contributions of socioeconomic characteristics more generally in explaining cross-sectional variation in racial segregation are examined by Miller and Quigley (1990), Harsman and Quigley (1995), and Bayer et al. (2004), among others.

[3] For expositional simplicity, we use the terms "middle-class neighborhoods" throughout the paper to refer to communities with a significant fraction of highly educated or high-income households and "poor neighborhoods" to refer to those with few highly educated or high-income households.

[4] By "formation" of highly educated black neighborhoods, we have in mind *either* an increased concentration of highly educated blacks within existing neighborhoods *or* the development of new neighborhoods via housing construction.

http://dx.doi.org/10.1016/j.jue.2014.05.002
0094-1190/

for those blacks who would have chosen middle-class white neighborhoods, residential segregation may increase markedly.

In this paper, we introduce the neighborhood formation mechanism and explore its empirical importance in American cities. To formalize the mechanism, we set out a simple equilibrium model of decentralized residential choice, which serves to link a city's overall sociodemographic composition with its level of neighborhood racial segregation.[5] If households only value vertically-differentiated neighborhood amenities and do not care about the race of their neighbors when deciding where to live, we show that socioeconomic inequality and racial segregation indeed exhibit a monotonic positive relationship, as suggested by the conventional intuition. Such monotonicity breaks down, however, when racial considerations also affect location choices. In this case, if the proportion of highly educated blacks is sufficiently low, the choice set is restricted in that middle-class black neighborhoods are scarce. As the proportion of highly educated blacks in a city increases, the set of available neighborhood options expands through the formation of new middle-class black neighborhoods, providing an avenue for segregation to rise as highly educated blacks leave predominantly white high-amenity neighborhoods.

Our focus on neighborhood formation as a possible channel linking inequality with racial segregation is motivated by three stylized observations about the current state of American cities; we begin our empirical analysis by documenting these facts in detail. First, the vast majority of metropolitan areas contain very few, if any, middle-class black neighborhoods. Second, given the limited availability of such neighborhoods, a substantial fraction of highly educated blacks (education proxying for socioeconomic status (SES) more generally) reside in both middle-class white neighborhoods and poor black neighborhoods. This suggests that many highly educated blacks might well prefer to locate in middle-class black neighborhoods, *were* they available.[6] Third, middle-class black neighborhoods indeed emerge in those metropolitan areas with a sufficiently high proportion of highly educated blacks. Taken together, these facts suggest that our proposed neighborhood formation mechanism may be empirically relevant, given the makeup of many American cities.

The main empirical goal of this paper is to shed light on the importance of our hypothesized neighborhood formation channel in the presence of other possible mechanisms, particularly the neighborhood effects mechanism of Cutler and Glaeser (1997) – hereafter "CG" – which yield opposing predictions for the inequality-segregation relationship.[7,8] According to CG's mechanism, growing up in a more segregated neighborhood leads to greater racial inequality. Other mechanisms implying a positive segregation-inequality relationship include sorting on the basis of sociodemographic factors correlated with race (along the lines envisaged by Schelling), and discrimination in the housing market.[9]

In the absence of a suitable natural experiment or compelling instruments for racial inequality, we propose a two-part research design, taking advantage of differential relations between black–white education inequality and neighborhood segregation across an individual's life cycle. The key idea is that CG's neighborhood effects mechanism, which leads to a positive relationship, is strongest for young blacks, namely those either of school age or who recently completed their education – the other alternative mechanisms we previously discussed would lead to a similarly positive relationship for all ages. In contrast, our neighborhood formation mechanism generates a negative relationship between racial inequality and segregation for blacks of all ages, and should be especially strong for older blacks, whose education has been long pre-determined.

Building on this idea, we argue first that if our neighborhood formation mechanism operates strongly in the data, one would expect to see a negative cross-sectional correlation between inequality and segregation for *older* blacks. This is indeed what we find using Census data: controlling for white educational attainment, the proportion of highly educated blacks aged 40 and above in a metropolitan area increases in the level of neighborhood segregation, implying a strong negative cross-sectional relationship between racial inequality and segregation for this older group. This new finding is surprising because it implies that our neighborhood formation mechanism not only overcomes the force of CG's neighborhood effects channel, but also the various other mechanisms that work in the opposite direction.

The second part of our research design focuses on evidence using first differences over time. Given that CG's mechanism operates only for younger blacks and our neighborhood formation mechanism operates throughout the life cycle, the latter should dominate upon differencing. Further, the strength of our mechanism should be identified by the first-difference effect of changes in segregation regressed on changes in neighborhood educational attainment for *older* blacks, after controlling for changes in the education of whites. Implementing this first-differencing approach using Census data, we show that increases in the proportion of highly educated blacks in a metropolitan area between 1990 and 2000 are associated with significant increases in overall racial segregation, after controlling for the educational attainment of whites and fixed city-level factors. Alongside the cross-sectional evidence, this is also surprising, given the operation of competing mechanisms that would tend to produce reductions in segregation under the same conditions. When we look specifically at older blacks, we find that increases in the proportion of highly educated blacks (again controlling for white education) are associated with strongly positive increases in city-wide segregation.[10] We also find that such changes are associated with significant increases in the number of middle-class black neighborhoods, as hypothesized under our neighborhood formation mechanism. In combination, these findings have implications for the inter-related dynamics of segregation and racial inequality, which we elaborate on after presenting the results.

The remainder of the paper is organized as follows: In Section 2, we set out a simple equilibrium sorting model intended to capture the role of neighborhood formation; Section 3 provides empirical motivation for our neighborhood formation hypothesis; in Section 4, we describe our two-part research design and present our main empirical evidence, with complementary findings in Section 5; in Section 6, we discuss the implications of our results; and Section 7 concludes.

[5] In practice, segregation may result from discrimination in the housing market as well as from household choices. This limits the extent to which the model can be used to carry out welfare analyses.

[6] This is consistent with Vigdor's (2003) finding that "the nationwide proportion of Black households with *few or no* Black neighbors exceeds the proportion stating a preference for such neighborhoods" (p. 589).

[7] However, Footnote 45 and Table 9 address the possibility that the strength of CG's neighborhood effects mechanism may depend on the proportion of highly educated blacks in the metropolitan area.

[8] Given our focus, the current paper will have relatively little to say about the contrasting legal and welfare implications of alternative mechanisms linking segregation and inequality, important though these are.

[9] Given that higher-income households typically sort into neighborhoods with bigger houses and better amenities, so greater racial inequality should lead to more racial segregation because of sorting in dimensions correlated with race. And if discrimination contributes to residential segregation, as is likely, then to the extent that less-educated or lower-income blacks are more discriminated against (see Arrow (1973), so one would expect increases in racial inequality to lead to greater segregation.

[10] In Section 5, we present evidence indicating that the positive relation is due primarily to within- rather than across-metropolitan area sorting. Both sources can be viewed as variants of the same general sorting mechanism.

2. The neighborhood formation mechanism in theory

In this section, we formalize our neighborhood formation mechanism, presenting a simple equilibrium sorting model to clarify the relationship between the sociodemographic composition of a metropolitan area and neighborhood segregation.[11,12]

Neighborhoods. Consider a metropolitan area with a total mass of households equal to 1. Suppose that a fraction $\lambda \in (0,1)$ of these households is black, with the remainder $1-\lambda$ being white. The total number of neighborhoods is fixed at J. Let the measure of available houses in neighborhood $j \in \{1,\ldots,J\}$ be n_j, and assume that all houses are physically identical, with the total units of available houses across all neighborhoods being equal to the total number of households, i.e. $\sum_{j=1}^{J} n_j = 1$.

From household i's perspective, neighborhood $j \in \{1,\ldots,J\}$ is characterized by three attributes. The first is the exogenous amenity level of neighborhood j, denoted by q_j. Without loss of generality, we assume that $q_1 \geqslant q_2 \geqslant \cdots \geqslant q_J$.[13] The second attribute is the fraction of neighbors of the *same* race as household i in neighborhood j, denoted by r_{ij}. It is endogenous and will be determined in equilibrium. The third attribute is the price of houses in neighborhood j, denoted by p_j, which will also be determined in equilibrium.

Households. Households are heterogeneous in their tastes for the amenity, denoted by α_i for household i, and also their preferences for the race of their neighbors, which denoted by β_i.[14] The utility that household i with preferences (α_i, β_i) receives from living in neighborhood j with attributes (q_j, r_{ij}, p_j) is given by

$$U_{ij} = \alpha_i q_j + \beta_i r_{ij} - p_j. \tag{1}$$

We assume that a household's taste for the amenity, α_i, varies with its education level, which is either *high* or *low*. If a household is highly educated, then its amenity taste parameter α is drawn from a continuous CDF $F_h(\cdot)$, while if a household is less-educated, then its α is drawn from a continuous CDF $F_l(\cdot)$, where $F_h(\cdot)$ first-order stochastically dominates $F_l(\cdot)$. This captures the idea that highly educated households are more willing to pay for amenities than less-educated households. We denote the fraction highly educated among all black households in the city by $\rho_B \in (0,1)$ and the fraction highly educated among whites, $\rho_W \in (0,1)$. For simplicity, assume that the taste parameter for same-race neighbors, β_i, is identical for all households, i.e. $\beta_i = \beta \geqslant 0$ for all i. Given their preferences, households simply choose to reside in one of the J neighborhoods in order to maximize utility.[15]

Equilibrium. An equilibrium in this model is characterized by a rule assigning households to neighborhoods and a vector of housing prices $(p_1,\ldots,p_J)$, where p_J is normalized to zero, such that the housing markets in all neighborhoods clear, and all households are in their most preferred location given the amenity levels, racial compositions, and housing prices in all neighborhoods.

Given this simple structure, we now describe how to solve the model, first in the simpler case where tastes over the race of one's neighbors are switched off, i.e., when $\beta = 0$. For a given equilibrium, we calculate a standard segregation measure; then we examine how the segregation index changes as we increase the proportion of highly educated blacks in the metropolitan area population, given the education of whites. The results from this exercise provide a benchmark against which we compare the more general case where households have tastes over the race of their neighbors in addition to preferences over exogenous amenity levels.[16]

2.1. No same-race preferences ($\beta = 0$)

In the case where households do not care about the race of their neighbors, neighborhoods differ in two relevant dimensions only: their amenity levels q_j and their housing prices p_j. The (essentially) unique equilibrium of the one-dimensional model is a positive assortative matching equilibrium, where households with a high preference for amenities sort into high-amenity neighborhoods, with housing prices in neighborhood j set at a level that makes the marginal household indifferent between living in neighborhood j and neighborhood $j-1$, the next level down in terms of amenity quality.

The equilibrium in this case is straightforward to characterize, and can be solved for analytically. The first step involves finding the threshold values of α recursively that will equate demand and supply of houses in each neighborhood; the second step is then to find the housing prices in each neighborhood to ensure that the marginal households are indifferent between the neighborhoods with adjacent values of amenities. Under the assumption that the race of residents in a particular community is randomly drawn from blacks and whites given their educational attainment – reasonable given that there are no same-race preferences – we can infer the racial compositions of each neighborhood, which we can then use to compute segregation indices.

The segregation measure that we use in this simple model is the *exposure rate*.[17] Our primary interest lies in the consequences for racial segregation, measured by the exposure rate of black households to white neighbors, when there is a reduction in racial inequality. Specifically, we increase the fraction of highly educated blacks (ρ_B), holding fixed the educational attainment of whites (ρ_W), i.e., the proportion of highly educated blacks is increased at the expense of less-educated blacks, starting from $\rho_B = 0$. When same-race preferences are absent, i.e., when $\beta = 0$, the average exposure of blacks

---

[11] Our stylized model abstracts from several considerations likely to be relevant in practice: commuting cost is ignored in locational decisions, and neighborhood composition does not affect the production of individual attributes, such as educational attainment. In our empirical analysis in Section 4, we do allow for the operation of this latter "neighborhood effects" channel.

[12] Sethi and Somanathan (2004) develop a model to explain the persistence of high levels of racial segregation in many U.S. cities. Their treatment focuses on the stability of equilibria in the context of a transparent two-community model, while our analysis demonstrates explicitly how inequality and segregation can be negatively linked (via neighborhood formation) in a model with potentially many communities.

[13] As in Sethi and Somanathan (2004), it is possible to endogenize the amenity level, for instance, by making it equal to the fraction of highly educated in the neighborhood. For our purposes, this generalization is not essential.

[14] The preference for same-race neighbors can either represent a pure taste for living in neighborhoods with others of the same race or arise through indirect channels. For example, individuals of the same race might cluster together in residential neighborhoods because they have correlated preferences for local amenities including retail outlets, restaurants, newspapers, and churches (see Berry and Waldfogel, 2003; Waldfogel, 2007). For various theoretical arguments why individuals might care about the racial composition of their neighborhoods, see, e.g. Cornell and Hartmann (1997), Farley et al. (1994), O'Flaherty (1999) and Lundberg and Startz (1998); for empirical evidence, see, e.g. Ihlanfeldt and Scafidi (2002), Vigdor (2003), Charles (2000, 2001), King and Mieszkowski (1973), Yinger (1978) and Galster (1982).

[15] The assumption that blacks are free to choose from the whole set of neighborhoods is made to simplify our argument. To the extent that blacks may be excluded from living in some neighborhoods due to discrimination, the phrase that blacks make "choices" should be viewed as shorthand, capturing both locational preferences and discrimination.

[16] In the Appendix, we solve an illustrative six-community example, which underlies the figures later in this section.

[17] At the individual level, the exposure rate of a household i in group g to another group g' is the percentage of household i's neighbors that belong to group g'. Our arguments also go through if we use alternative segregation measures, such as the dissimilarity index (adjusting for the fact that it is inversely related to the exposure rate). Dissimilarity indices are used in our main empirical analysis in Section 4.

to white neighbors will be monotonically increasing in ρ_B over the empirically relevant range, $\rho_B < \rho_W$. Intuitively, as blacks shift up the education distribution conditional on the education of whites, their tastes for higher amenity neighborhoods strengthen, leading to greater residential integration as blacks and whites become more similar in this dimension.

For illustration, we plot the relationship using the parameterization given in the six-community example, developed in the Appendix, in Fig. 1(a), which has racial equality in education on the horizontal axis and segregation on the vertical axis. Note that when sorting occurs solely on the basis of education and the associated taste for the amenity, some racial segregation arises initially simply because race is correlated with education and thus taste for amenity. This corresponds to the logic in Schelling's argument (see Footnote 2) that some degree of racial segregation would be expected even in the absence of any direct preference over the race of one's neighbors.

2.2. Strictly positive same-race preferences ($\beta > 0$)

We now provide an intuitive characterization of the equilibria for the case where households care about the race of their neighbors in addition to amenity levels and housing prices.[18] When households care about the race of their neighbors, the allocation rule described above for the case without same-race preferences needs to be modified. Since the highest amenity neighborhoods are predominantly white, whites with any given taste for amenity will now be willing to pay more than (and thus outbid) blacks with the same taste for the amenity, due to same-race preferences. This will drive the proportion of whites even higher, leading other whites to find these neighborhoods even more attractive.

To fix the ideas related to our neighborhood formation mechanism, suppose that the proportion of whites who are highly educated, ρ_W, is fairly close to one, and contrast two extremes. First, consider a situation where the proportion of highly educated blacks among all blacks, ρ_B, is very low. In such a case, it is impossible to have a large fraction of blacks in the highest amenity neighborhoods. Given that, the threshold taste level above which highly-educated blacks will be willing to pay to live in such high-amenity neighborhoods, denoted by α_B^*, must be higher than the threshold for highly-educated whites α_W^*, i.e., $\alpha_B^* > \alpha_W^*$. Nonetheless, highly-educated blacks with very high amenity taste draws will find it optimal to live in predominantly white neighborhoods with high amenity levels. As ρ_B increases in a range of small values starting from 0, we would thus expect there to be more highly-educated blacks with exceptionally high values of α who choose to live in predominantly white high-amenity neighborhoods rather than lower-amenity neighborhoods that have greater proportions of blacks. Thus initially, we expect black households' exposure to white neighbors to be increasing in ρ_B.

Now consider the other extreme case, where ρ_B is high and close to ρ_W. Here, it becomes possible for the highly-educated blacks with a high taste for the neighborhood amenity to bid for houses in one of the high-amenity neighborhoods and achieve a racial majority there. Once blacks become a majority in a high-amenity neighborhood, the same-race preference will lead more blacks (with somewhat lower α's) to move into that neighborhood, and this process could lead to the emergence of a predominantly black high-amenity neighborhood. In this case, in contrast, the exposure rate of black households to white neighbors tends to be low.[19]

Combining these pieces of reasoning, we would expect the relation between black exposure to whites – our measure of racial *integration* – and the fraction of highly-educated blacks ρ_B to exhibit an inverted-U relationship, with a range of values for ρ_B over which the exposure rate of black households to white neighbors declines in ρ_B. In this range, segregation and racial inequality are negatively related. We verify that this is indeed the case in the context of our stylized residential choice model.

Fig. 1(b), drawn from the computational sorting equilibrium of the simple model, illustrates the above argument.[20] As shown, when $\rho_B < \rho_B^*$, there is no possibility of a majority-black high-amenity neighborhood; thus, as ρ_B increases, more and more highly-educated black households with high-α preferences live in white-majority high-amenity neighborhoods, and so blacks' average exposure to whites increases in ρ_B. But at $\rho_B = \rho_B^*$, a black majority high-amenity neighborhood becomes sustainable; and as a result, when ρ_B gets larger than ρ_B^*, blacks' exposure to white neighbors starts to decline with ρ_B as more and more highly-educated blacks move into high-amenity black-majority neighborhoods.[21]

A complementary way to depict the effects of an exogenous increase in the proportion of highly educated blacks ρ_B, while holding ρ_W fixed, is to directly examine the evolution of available neighborhoods that emerge in equilibrium. Using the simulated equilibrium outcomes for the model outlined above by varying ρ_B, for a given $\beta > 0$, Fig. 2 plots the available equilibrium neighborhood configurations in "% Black" (horizontal axis) and "Amenity" (vertical axis) space for two different values of ρ_B. The left panel 2(a) shows that when ρ_B is small, the sorting equilibrium is unable to support majority-black high-amenity neighborhoods (i.e., neighborhoods in the northeast quadrant) due to an insufficient number of highly educated blacks with strong tastes for amenities; instead, the small measure of highly-educated blacks with strong tastes for amenities live in white-majority high-amenity neighborhood. However, the right panel 2(b) shows that, as ρ_B becomes sufficiently big, so high-amenity, black-majority neighborhoods start to emerge in the north-east portion of the figure. The presence of such neighborhoods provides an opportunity for racial segregation to increase, as we hypothesize.

The stylized depiction in Fig. 2 has a useful analog in terms of scatterplots describing actual cities. As we will see, Fig. 3 in Section 3 presents scatterplots analogous to those in Fig. 2, showing how the range of available communities can expand when the underlying demographic structure of the MSA changes. Specifically, Fig. 3 is constructed using actual cross-sectional Census data from U.S. cities, where Boston and St. Louis represent MSAs with low proportions of highly educated blacks (low ρ_B) and Atlanta

[18] Because analytical solutions are difficult to obtain in this more general case, we confirm the main intuition by solving for the illustrative model's equilibria numerically.

[19] Potential multiple equilibria complicate our discussion. Here we are just referring to the possibility of such a predominantly black high-amenity equilibrium. It should be intuitively clear that with same-race preferences, the equilibrium with the highest degree of racial segregation actually maximizes landowner profits from house sales, i.e. it is the equilibrium that maximizes the total housing prices of the neighborhoods. We assume that such an equilibrium is likely to be selected. This allows us to assume away the coordination problem, and instead focus on the *small numbers problem*, according to which middle-class black neighborhoods may not arise because of an insufficient mass of highly educated blacks. Coordination problems are likely to be a short-term phenomenon, as developers and other entrepreneurs have an incentive to solve them.

[20] We apply a variant of the algorithm that solves numerically for sorting equilibria presented in Bayer et al. (2011) (see Appendix for more details).

[21] The empirics we present in Section 4 support the view that in the current configuration of U.S. cities, the relationship between blacks' educational attainment (relative to whites) and residential segregation is likely to be on the decreasing portion of the curve, as shown in Fig. 1(b). There, we restrict attention to cities with more than 10,000 blacks, which might be viewed as a proxy for the critical-mass threshold.

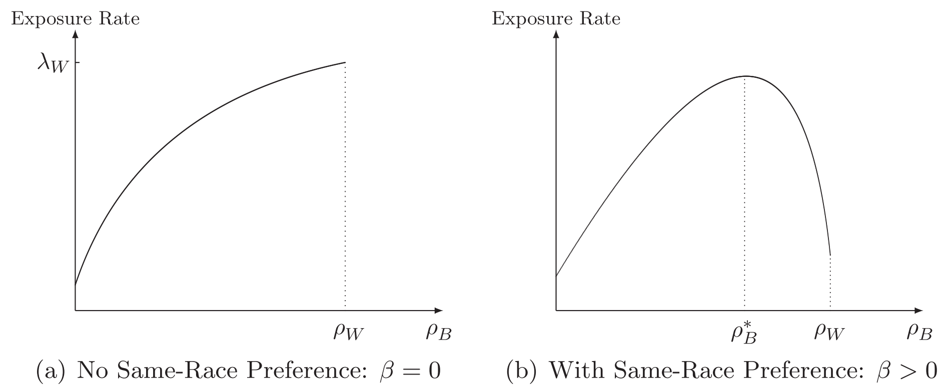

Fig. 1. Black households' average exposure rate to white neighbors as a function of ρ_B.Notes: ρ_B and ρ_W denote the fraction of highly-educated blacks and whites respectively; λ_W denotes the fraction of whites in the MSA population. The figures are drawn from the calculated equilibrium of the model described in the text as ρ_B varies from 0 to 3/5, and $\rho_W = 3/5$ (see Appendix for the full parameterizations). At $\rho_B = \rho_B^*$, a majority-black high-amenity neighborhood becomes sustainable.

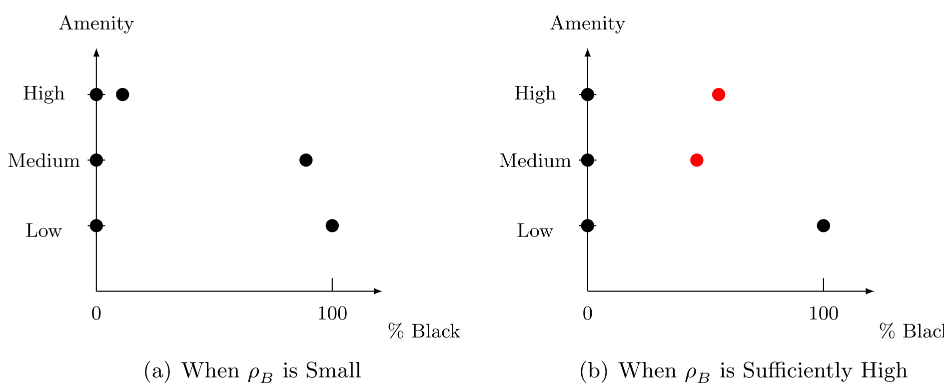

Fig. 2. Neighborhoods in "% Black-Amenity" Space as ρ_B increases, when households have same-race preferences.

and Baltimore-Washington DC represent MSAs with high proportions (high ρ_B). We discuss the relevant patterns in some detail below.

3. Neighborhood availability in U.S. metropolitan areas

In this section, we describe three stylized empirical facts about the availability of neighborhoods in U.S. metropolitan areas. These help motivate our focus on the neighborhood formation mechanism.

The 2000 U.S. Census provides the primary data source for the descriptive analysis of this section. Our sample consists of 276 Metropolitan Statistical Areas (MSAs).[22] Within each MSA, we examine the characteristics of its *neighborhoods*. In our analysis, a neighborhood corresponds to a Census tract, which typically contains between 3000 and 5000 individuals. Using publicly-available Census Tract Summary Files (SF3) from the 2000 Census, we characterize each neighborhood on the basis of two dimensions: *the fraction of residents who are black* and *the fraction of residents with four-year college degrees*.[23,24]

Fact 1. *In almost every MSA, there are very few neighborhoods combining high fractions of both college-educated and black individuals.*

Table 1 provides very clear evidence relating to Fact 1. Panel A lists the overall number of tracts in which more than 0%, 20%, 40% and 60%, respectively by column headings, of individuals 25 years and older are at least college-educated. Panel B then shows the number of tracts in the U.S. by both education and race (specifically, the percentage of individuals with a college degree and the percentage of individuals who are black). As the corresponding numbers show, a much smaller fraction of the tracts with a high percentage black also have a high proportion of college-educated individuals. For example, while 22.6% of all tracts are at least 40% college-educated, only 2.5% of tracts that are at least 40% black are at least 40% college-educated, and only 1.1% of tracts that are at least 60% black are at least 40% college-educated. In marked contrast, Panel C presents analogous numbers for whites, showing a far greater fraction of neighborhoods with at least 40, 60, and 80% white meeting the education criteria listed in the column headings.[25]

Fact 2. *College-educated blacks live in a very diverse set of neighborhoods in each MSA. Substantial fractions live in predominantly white high-SES neighborhoods and substantial fractions also live in predominantly black low-SES neighborhoods.*

[22] These include free-standing Metropolitan Statistical Areas (MSAs) and Consolidated Metropolitan Statistical Areas (CMSAs), which consist of two or more economically and socially linked metropolitan areas.

[23] Our focus in this section is on non-Hispanic black and non-Hispanic white individuals 25 years and older residing in U.S. metropolitan areas.

[24] The Census Summary Files necessitate the use of a single dimension to characterize socioeconomic status as they only provide the joint distribution of race-by-income or race-by-education for a given neighborhood. In light of this constraint, we use educational attainment to proxy socioeconomic status more generally, on the basis that it is a better predictor of permanent income than current income in the Census year.

[25] While Table 1 reveals a scarcity of high-SES black neighborhoods in the U.S. as a whole, these tracts are concentrated in only a handful of MSAs, and most notably Baltimore-Washington, DC. (see Table A.1). This implies that in most MSAs, the availability of high-SES black neighborhoods is even more limited. See Pattillo (2005) for a thorough discussion of such neighborhoods.

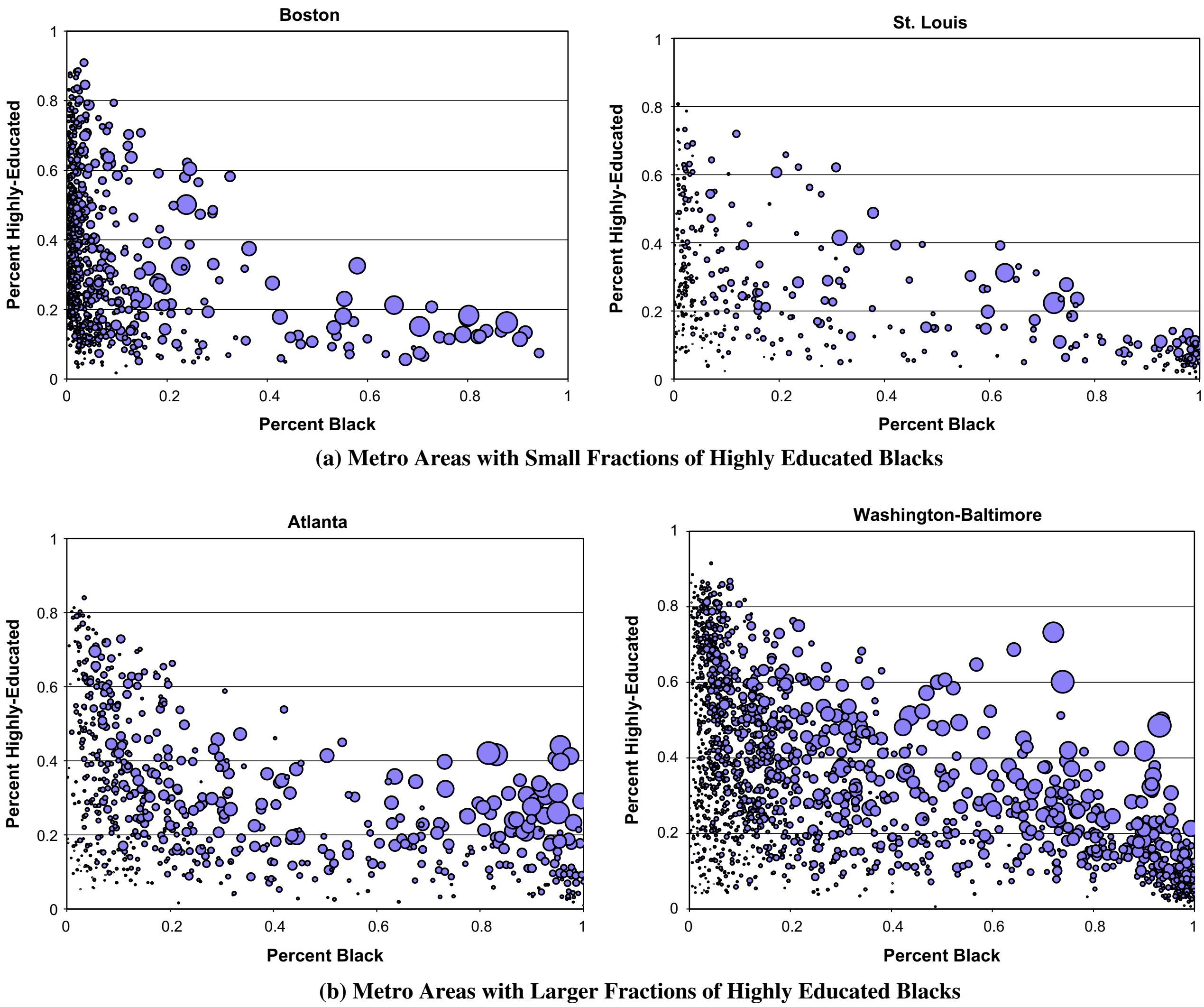

Fig. 3. Neighborhood characteristics in illustrative metropolitan areas: Boston and St. Louis (Panel (a)); Atlanta and Baltimore-Washington DC (Panel (b)).

Table 2 provides evidence relevant to Fact 2, summarizing the characteristics of neighborhoods in MSAs throughout the United States in which college-educated blacks reside. Given the absence of mixed- or high-SES black neighborhoods, highly educated blacks live in a diverse set of neighborhoods, ranging from those that are predominantly white and highly educated to neighborhoods that are predominantly black with much lower levels of education on average. The numbers point to a clear trade-off for college-educated blacks between the fraction of their neighbors who are black and the fraction who are highly educated: the average fraction of highly educated neighbors falls from 38.0% for those college-educated blacks living with the smallest fraction of black neighbors to 13.8% for those living with the largest fraction.[26]

Two aspects of the pattern in the table are pertinent to our neighborhood formation mechanism. First, the fact that such a high fraction of college-educated blacks live in segregated neighborhoods with relatively low average educational attainment suggests that – whether due to preferences or discrimination – race remains an important factor in the location decisions of a large number of college-educated blacks. This helps to rule out an obvious potential explanation for the absence of mixed- or high-SES black neighborhoods, namely that college-educated households simply demand college-educated neighborhoods without regard to racial composition. Second, the fact that a significant number of college-educated blacks reside in predominantly white neighborhoods makes it possible for an increase in the availability of mixed- or high-SES black neighborhoods to lead to greater segregation.

Fact 3. *While predominantly black high-SES neighborhoods are concentrated in only a handful of MSAs, the availability of these neighborhoods is increasing in the proportion of college-educated blacks in the MSA population.*

In support of our third stylized fact, Table 3 reports four regressions that relate the log of the number of tracts in an MSA that meet the race and education criteria specified in the column heading to metropolitan socioeconomic characteristics (proportion highly educated black, highly educated white, less-educated black and less-educated white) and the log of metropolitan area population. Holding the size of the MSA constant, a one percentage-point increase – just under a standard deviation – in the proportion of college-educated blacks in an MSA (at the expense of the omitted category, Asians and Hispanics) increases the number of tracts that

[26] Comparison of Panels A and B in Table 2 reveals that college-educated blacks in each metropolitan area who reside with the smallest fraction of other blacks have roughly the same fraction of college-educated neighbors as college-educated whites do on average. However, college-educated blacks living in tracts with the highest fraction of black neighbors have only about *one-third* of the fraction of highly educated neighbors as whites do on average.

Table 1
Number of tracts in United States in 2000 by race and education.

	(1)	(2)	(3)	(4)
	Percent college degree or more			
	All	>20%	>40%	>60%
Panel A: All tracts				
(1) All	49,021	26,351	11,094	3005
	100.0%	53.8%	22.6%	6.1%
Panel B: Tracts by Percent Black				
(2) >20% Black	9149	2567	641	59
	100.0%	28.1%	7.0%	0.6%
(3) >40% Black	5657	1164	142	14
	100.0%	20.6%	2.5%	0.2%
(4) >60% Black	3921	623	44	5
	100.0%	15.9%	1.1%	0.1%
(5) >80% Black	2559	271	21	1
	100.0%	10.6%	0.8%	0.0%
Panel C: Tracts by Percent White				
(6) >20% White	43,179	25,178	11,041	2999
	100.0%	58.3%	25.6%	6.9%
(7) >40% White	39,602	24,566	10,839	2967
	100.0%	62.0%	27.4%	7.5%
(8) >60% White	35,154	22,543	10,214	2870
	100.0%	64.1%	29.1%	8.2%
(9) >80% White	26,910	17,539	8102	2339
	100.0%	65.2%	30.1%	8.7%

Note: The top number in each cell gives the number of tracts meeting both the education criterion described in the column heading (e.g., greater than 40% college-educated) and the race criterion in the row heading (e.g., greater than 40% black); the bottom number in each cell gives the number of tracts meeting each race and education criterion as a fraction of the number of tracts meeting each race criterion. Tract compositions are calculated using individuals 25 years and older in U.S. metropolitan areas. Tracts considered in this table have a minimum of 800 such individuals (the average tract in the full sample has slightly over 3000).

are least 60% black and 40% college-educated by 42%, and the number that are at least 60% black and 20% college-educated by 56%. The embolded values in the table are there to emphasize the main estimates of interest. The sizes of these effects are substantially in excess of the mechanical increase that would occur were the additional blacks distributed evenly across all the typical MSA's tracts – unsurprising given the small fraction of the typical MSA population accounted for by college-educated blacks.

Neighborhood scatterplots using census data. Related to the regression evidence in Table 3, Fig. 3 shows scatterplots of available neighborhoods in four metropolitan areas: Boston and St. Louis in Panel A, and Atlanta and Baltimore-Washington DC in Panel B. Note that in Boston and St. Louis, around 11% of the blacks have college degrees, while the fractions of blacks in Atlanta and Baltimore-Washington DC with college degrees are approximately twice as high.[27]

In each scatterplot, a circle represents a Census tract and its coordinates describe the fraction of blacks (horizontal axis) and the fraction of college-educated individuals (vertical axis) in the tract. The diameter of each circle is proportional to the number of college-educated blacks in the tract; thus the largest circles correspond to the tracts where highly educated blacks are most likely to live. Panel A reveals a short supply of neighborhoods in Boston and St. Louis that combine high fractions of both highly educated and black individuals – few neighborhoods appear in the north-east corner of the plot. Panel B shows that a substantially greater number of neighborhoods combining relatively high fractions of both black and highly educated individuals – those populating the north-east corner of each figure – are found in the Atlanta and Baltimore-Washington DC metropolitan areas.[28] These scatterplots resemble stylized Fig. 2, which illustrates neighborhood formation derived from our model when residents have same-race preferences.

It is this third stylized fact along with the documented small number of middle-class black neighborhoods in the vast majority of U.S. metropolitan areas (Fact 1) that motivates the idea that an increase in the proportion of highly educated blacks within a metropolitan area should allow middle-class black neighborhoods to form more readily. As these neighborhoods are likely to be attractive to highly educated blacks, and indirectly through same-race preference to less-educated blacks as well, their emergence may lead to an empirically sizable increase in residential segregation on the basis of race once households re-sort, along the lines of the model presented in Section 2. The potential for such re-sorting is apparent from Fact 2 which documented that a nontrivial fraction of highly educated blacks currently reside in predominantly white neighborhoods.

4. Research design and main results

The theoretical and descriptive analyses of the previous two sections motivate our main empirical hypothesis– that residential segregation and racial inequality will be *negatively* related, given the racial and socioeconomic compositions of most U.S. metropolitan areas. Further, this negative relationship arises, so we argue, through a process of neighborhood formation.

One possible approach to shedding light on this hypothesis is to mimic the stylized exercise in Section 2 by specifying household tastes over locational attributes, then estimating an equilibrium residential choice model using data drawn from a single metropolitan area.[29] In this paper, we take a different tack, making use of across-MSA data in order to assess whether the neighborhood formation mechanism is important in practice. The observational data we use for our analysis make it extremely difficult to isolate exogenous variation in the sociodemographic variables of interest; yet even in the absence of compelling instruments, we argue that the pattern of observed correlations between MSA-wide segregation and inequality, both cross-sectionally and over time, can be informative as to which of the potential mechanisms are operating strongly in the data.

To explain the logic of our approach, consider as a starting point estimates of the cross-sectional relationship between an MSA's level of residential segregation and the fraction of highly educated blacks there, controlling for the educational attainment of whites. Such estimates will clearly reflect the overall impact of several alternative mechanisms, discussed in the Introduction. In order to distinguish the impact of our hypothesized neighborhood formation mechanism from the alternative mechanisms, including CG's neighborhood effects mechanism, we take advantage of the *differential timing* of these mechanisms over the life cycle. In particular, the neighborhood effects mechanism implies a negative relationship between concurrent measures of segregation and the educational attainment of *young* blacks; as the metropolitan area evolves and individuals move within and across metropolitan areas, this negative relationship should generally weaken with age. In contrast, our neighborhood formation mechanism gives rise to a cross-sectional relationship between concurrent measures of segregation and the educational attainment of blacks that should

[27] For reference, blacks and whites constitute 11.1% and 69.5%, respectively, of the U.S. population 25 years and older residing in MSAs. Among blacks, 15.4% have at least a four-year college degree, while the comparable number for whites is over twice as high, at 32.5%. Blacks with four-year college degrees constitute a mere 1.7% of the U.S. population residing in MSAs.

[28] See Gabriel and Painter (2012) for a recent discussion of segregation in Washington DC.

[29] See Bayer et al. (2011) for such an approach.

Table 2
Neighborhood patterns for college-educated individuals in the United States in 2000.

Quintile	1	2	3	4	5	Overall
Panel A: College-Educated Blacks						
College-educated blacks first ranked within each MSA by percent black in Census tract						
Average tract composition reported by corresponding quintile, averaging across all MSAs						
Percent Black	5.7	14.4	28.3	54.6	78.9	32.0
Percent College-Educated	38.0	31.6	26.2	18.4	13.8	27.2
Percent Black and College-Educated	1.3	3.3	6.2	8.0	10.0	5.2
Panel B: College-Educated Whites						
College-educated whites first ranked within each MSA by percent white in Census tract						
Average tract composition reported by corresponding quintile, averaging across all MSAs						
Percent White	55.0	77.9	86.6	90.4	94.5	77.4
Percent College-Educated	27.0	36.2	40.7	39.3	39.2	35.3
Percent White and College-Educated	20.1	30.4	36.2	36.1	37.4	30.4

Note: The panels of the table summarize the average distribution of neighborhoods in which college-educated blacks and whites in U.S. metropolitan areas reside, respectively, using data from the 2000 Census. To construct the numbers in Panel A, college-educated blacks in each MSA are ranked by the fraction of blacks in their tract and assigned to one of five quintiles for that MSA. Average neighborhood sociodemographic characteristics are then reported for each quintile, averaging across all MSAs. Overall averages are given in the last column. Panel B reports analogous figures for college-educated whites, first ranking by their tract-level exposure to whites within each MSA. Tract compositions are calculated using individuals 25 years and older in U.S. metropolitan areas.

Table 3
The availability of middle-class black neighborhoods in 2000.

Dependent variable	log(number of tracts in MSA >60% black and >40% college-educated) (1)	log(number of tracts in MSA >60% black and >20% college-educated) (2)	log(number of tracts in MSA >40% black and >40% college-educated) (3)	log(number of tracts in MSA >40% black and >20% college-educated) (4)
Metropolitan composition				
% Black with College Degree	**42.16*****	**55.70*****	**36.14*****	**38.51*****
	(10.28)	**(11.33)**	**(13.35)**	**(11.37)**
% Black with less than College Degree	−4.51***	0.49	−1.52	5.22***
	(1.83)	(2.34)	(2.42)	(2.21)
% White with College Degree	−1.64	1.06	0.49	3.52***
	(1.20)	(1.49)	(1.55)	(1.62)
% White with less than College Degree	0.06	1.77***	−0.21	1.84**
	(0.57)	(0.73)	(0.65)	(0.81)

Notes: The four regressions reported in this table relate various measures of the availability of middle-class black neighborhoods to the sociodemographic composition of the metropolitan area using 2000 Census data. Metropolitan-level observations ($N = 276$) are weighted by population, and the log of the population is included as an additional control. Standard errors are reported in parentheses. *** Denotes significance at the 1% level; ** denotes significance at the 5% level.

be positive for households of all ages, and potentially be *even stronger* for older households, who are more likely to have made multiple residential location decisions during their lives. Consideration of this life-cycle pattern suggests two complementary ways to distinguish the neighborhood formation mechanism from the neighborhood effects mechanism empirically, which we describe next.

4.1. Cross-sectional evidence

The first approach is cross-sectional. If we extend the analysis of Cutler and Glaeser (1997) to *older* individuals (their paper focuses on ages 20–30), we should see a significant weakening of the effects that they find. To that end, we follow CG and estimate regressions of the form:

$$y_i = \boldsymbol{X}_i'\boldsymbol{\beta} + \beta_1 \mathrm{Seg}_i + \beta_2 \mathrm{Seg}_i \times \mathrm{Black}_i + \epsilon_i, \quad (2)$$

where y_i represents an individual outcome variable, $\textsc{Seg}_i$ is an MSA-level measure of segregation of blacks and whites, $\textsc{Black}_i$ is a dummy variable taking value 1 if individual i is black, and $\boldsymbol{X}_i$ includes individual demographic and MSA-level characteristics.[30] We do so separately for individuals aged 20–24 and 25–30, as in CG, but also for older age groups, between the ages of 30 and 70, focusing on the effect of living in a more segregated metropolitan area for blacks relative to whites, summarized by the coefficient (β_2) on the segregation-black dummy interaction term.

To get a sense of possible combined age patterns, the neighborhood formation mechanism is hypothesized to lead to a positive relationship between concurrent measures of segregation and the educational attainment of blacks relative to whites across the age range; the neighborhood effects mechanism is hypothesized to have a negative impact, strongest among young blacks and declining with age. If both mechanisms are operating strongly in the data, we might expect the neighborhood effects mechanism to

[30] While CG's framework was developed to explore the importance of neighborhood effects – i.e., the impact of MSA-wide segregation on the educational and labor-market outcomes of blacks relative to whites, controlling for other factors – it also provides us with a useful means of estimating conditional correlations between residential segregation and inequality.

dominate at younger ages, and the neighborhood formation mechanism to dominate at older ages. Thus the *net* relationship between concurrent segregation in a metropolitan area and the educational attainment of blacks relative to whites (captured by the relevant age group-specific coefficient, β_2) should be negative for younger blacks relative to whites, but become positive for older blacks. Furthermore, if we do not distinguish blacks by age, we might observe that the average effects across all ages cancel out; thus conducting the analysis disaggregated by age allows us to separately identify these two effects.

We implement the above cross-sectional research design using data that combine variables from the 5% sample of the 1990 Census with the same set of MSA characteristics used in CG.[31] Descriptive statistics for the MSA variables are shown in Table A.2, the sample being drawn from the 209 metropolitan areas that have populations of at least a hundred thousand and at least ten thousand blacks in 1990. Following CG, we measure residential segregation using dissimilarity indices constructed for each MSA from racial compositions – the proportions of blacks and non-blacks – at the tract level.[32] The mean value for the dissimilarity index is 56%, with a standard deviation of 12.9%. The most segregated MSA in 1990 in the sample is Detroit (87.3%), and the least is Jacksonville, NC (20.6%).

We capture racial inequality in an MSA using the educational attainment of blacks relative to whites. Accordingly, racial inequality will be taken to have narrowed in a cross-sectional context when the proportion of highly educated blacks (the proportion with at least a college degree) increases across MSAs, given white educational attainment. In our 1990 Census sample, 22.7% of the adult population have a college degree or more. For whites, the mean proportion is 24.6%, while for blacks, it is under half of that at 11.4%. Around these average differences, there is considerable variation in educational attainment by race across MSAs.[33]

Given our interest in the age profile of educational attainment by race, we further disaggregate by age in Table A.3. The pattern is similar for blacks and whites, with educational attainment rising then falling across the age distribution. Educational inequality is apparent throughout the age distribution, with black educational attainment being markedly lower than for whites, this across-age variation for blacks versus whites being relevant for the first part of our research design. The table also shows descriptive statistics for labor market outcome variables – log wages and whether idle (both not working and not in school) – for the same age categories, and subdivided by race, along with a set of individual demographic control variables included in the main regressions.

Table 4 reports the coefficient estimates for β_2 on the interaction term $\text{Seg}_i \times \text{Black}_i$ in the specification described by Eq. (2). The individual outcomes we examine include college graduation (Column 1) and log earnings (Column 2), relevant to our notion of high-SES individuals, along with high school graduation (Column 3) and whether idle (i.e., neither unemployed nor in school, Column 4). All the specifications include a rich set of controls. The specifications in CG that relate most directly to our analysis are those using educational attainment as the dependent variable. Table 4 replicates CG's results for age groups 20–24 and 25–30, but extends the analysis for individuals between the ages of 31–50, 51–70, respectively, the latter two groups further broken down into 10-year age spans.

[31] These latter data were kindly made available to us by Jacob Vigdor.

[32] Dissimilarity indices range from zero to one, and can be interpreted as measuring the proportion of blacks who would need to change tracts in order for races to be evenly distributed throughout the metropolitan area (see Cutler et al. (1999) for more discussion).

[33] Stamford, CT has the highest gap between the proportions of whites and blacks with a college degree, at 38.6%, while Houma-Thibodoux, LA, Danville, VA, and Fayetteville, NC all have gaps between 7% and 8%.

Table 4
Segregation and metropolitan composition – age profile in 1990.

Dependent variable	College graduation (1)	Ln(Earnings) (2)	High school graduation (3)	Idle (4)
Coefficient on interaction between black and metropolitan segregation (***dissimilarity index***)				
Age category				
20–24	−0.104***	−0.549***	−0.266***	0.339***
	(0.031)	(0.108)	(0.037)	(0.033)
25–30	−0.075	−0.190***	−0.214***	0.302***
	(0.056)	(0.074)	(0.040)	(0.038)
31–50	0.004	0.196**	−0.117***	0.196***
	(0.058)	(0.083)	(0.041)	(0.033)
31–40	−0.034	0.146*	−0.125***	0.208***
	(0.062)	(0.087)	(0.037)	(0.037)
41–50	0.058	0.268***	−0.100**	0.176***
	(0.055)	(0.090)	(0.050)	(0.032)
51–70	0.073**	0.430***	−0.072	0.101***
	(0.034)	(0.124)	(0.061)	(0.031)
51–60	0.061	0.417***	−0.118*	0.109***
	(0.040)	(0.114)	(0.064)	(0.035)
61–70	0.087***	0.480***	−0.004	0.104***
	(0.031)	(0.185)	(0.065)	(0.036)

Notes: This table reports coefficients from a series of regressions based on the specification used in Cutler and Glaeser (1997) to generate their Table 4. The specification includes individual characteristics [Black, Asian, Other nonwhite, Hispanic, Female], metropolitan characteristics [Segregation, ln (population), Percent black, ln (median household income), Manufacturing share] and interactions of these metropolitan characteristics with whether the individual is black. The coefficient on Black*Segregation is reported here for four individual outcomes and for six age ranges. Cutler and Glaeser report results for individuals between the ages of 20–24 and 25–30, respectively. The coefficients in the table for these ages are not identical to those reported in Cutler and Glaeser but are very close, most likely attributable to the fact that we use the 5% sample of the 1990 Census while the 1% sample is used by Cutler and Glaeser. All other measures are identical, as we used the same metropolitan characteristics used by Cutler and Glaeser, generously made available by Jacob Vigdor on his website. *** Denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

The estimated β_2 coefficients in Table 4 for age groups 20–24 and 25–30 are very similar to those reported in Cutler and Glaeser (1997, Table IV), with any minor discrepancies being attributable to differences in Census sample (we use the 5% while CG used the 1% Census sample). For each outcome, the estimate for the β_2 coefficient on the "$\text{Seg} \times \text{Black}$" interaction implies a significantly worse outcome for younger blacks relative to whites that is also economically significant, as noted in CG.[34] Taking the age group 20–24 for illustration, a one standard deviation increase in segregation (12.6%) would lower the probability of graduating from high school for blacks relative to whites by around 3.3% points, and lower the earnings of blacks relative to whites by around 6.8%.

The point estimates for the 25–30 age group are all of the same sign but, for each outcome, are *lower in absolute value* for the 25–30 age group than those for the 20–24 age group. More strikingly, as we examine even older age groups (see Row 3 and below in Table 4), the effects of MSA segregation continue to dampen and, for key outcomes, change sign relative to the young age groups, consistent with the predicted age profile of net effects we explained in the previous section. For college graduation, the negative effect of increased segregation on blacks relative to whites becomes indistinguishable from zero for ages 31–50 – the point estimate is negative for ages 31–40, and positive for the 41–50 age group. It then becomes positive and significant for ages 51–70. In this case, a one standard deviation increase in segregation is associated with a 9% point increase in the probability that blacks

[34] The sole exception is college graduation for the age group 25–30, though the point estimate still indicates that blacks perform worse than whites.

graduate relative to whites, which is a large effect. The effect for the 61–70 subcategory is even larger. Similar pattern also holds for the other outcomes. For racial inequality in earnings, coefficients on "Seg × Black" also switch from being negative to being positive and statistically significant, now even for the 31–50 age group, and the sizes of the effect are monotonically increasing with age; for high school graduation (Column 3) and idleness (Column 4), the effects of MSA segregation continue to dampen for older age groups.

We take these results as strong baseline evidence in support of both the existence of the neighborhood formation mechanism that is the focus of this paper and the neighborhood effects channel identified by CG. At face value, the results suggest that in the very same highly segregated metropolitan areas, older blacks have significantly higher levels of educational attainment and earnings relative to whites (compared to their counterparts in less segregated cities), while younger blacks have significantly worse outcomes relative to whites. This pattern is exactly what one would expect given the combined operation of the neighborhood formation and neighborhood effects mechanisms, working differentially across black households of different ages.

4.2. Evidence in first differences

In the second part of our research design, we use the fact that the life cycle patterns exploited in the first part also give rise to a strong prediction concerning the relationship between segregation and the socioeconomic status of blacks relative to whites in *first differences*. Consider the relationship between the *change* in segregation in an MSA and the *change* in black socioeconomic status over time – for instance, comparing across decennial censuses, as we will do below – while controlling for changes in the socioeconomic status of whites. In this case, the operation of the neighborhood effects mechanism in CG implies that the change in segregation can only directly affect the educational attainment of younger blacks relative to younger whites and should have no effect on that of older age groups (because their educational (and human capital investment) decisions were largely complete around age 25). The neighborhood formation mechanism, in contrast, should continue to generate a positive relationship between the relative educational attainment of blacks (versus whites) and segregation at *all* ages. Relative to the cross-sectional relationship in levels, therefore, the neighborhood formation mechanism should more fully dominate the neighborhood effects mechanism in first differences. Moreover, the relationship between these variables attributable to the neighborhood formation mechanism should be identified by the correlation between changes in segregation and changes in black educational attainment relative to whites observed for older individuals.

To implement our second approach using changes over time, we estimate equations at the metropolitan area level of the form:

$$\Delta \mathrm{Seg}_j = \gamma_1 \Delta\%\mathrm{Highly\_Edu\_Black}_j + \boldsymbol{\Delta X_j'}\boldsymbol{\gamma} + v_j, \quad (3)$$

where $\Delta\mathrm{Seg}_j$ represents the change between the 1990 and 2000 Censuses in MSA j's segregation (captured by a relevant segregation index), $\Delta\%\mathrm{Highly\_Edu\_Black}_j$ the change in percent highly educated black in MSA j, and $\boldsymbol{\Delta X_j}$ includes MSA-level changes in other sociodemographics, including changes in the percentage of highly educated whites. Our interest focuses on the coefficient γ_1, which we hypothesize to be positive if the neighborhood formation mechanism dominates. Note that the first-differences research design also allows us to deal with identification issues associated with time-invariant omitted MSA-level characteristics that may influence neighborhood availability and are also correlated with the MSA's demographic structure.

Eq. (3) is estimated with the same individual and MSA variables used in the cross-sectional analysis for 1990 (see Table 4), but for the 2000 Census as well. We average the variables up to the MSA level, and construct first differences for each MSA based on 1990 and 2000 MSA averages. Descriptive statistics are given in Table A.4 for the sample of 214 MSAs that appear in both waves.[35]

The first feature to note from Table A.4 is that segregation, measured at the MSA level using dissimilarity indices, fell quite sharply over the decade: on average, dissimilarity indices were 5.4% lower, with a standard deviation of 4.1%. This accords with a fact that has been well-documented: as shown for example in Iceland et al. (2002), residential segregation in U.S. cities has been following a downward trend over the three decades since the 1980 Census, a conclusion that is invariant to the way segregation is measured.[36] Table A.4 also provides suggestive aggregate evidence that racial inequality has increased over the same decade. While the proportion of blacks with a college degree increased only very slightly between 1990 and 2000, the proportion of whites with a college degree rose by around 2.2% points; and in the same broad direction, while the proportion of less educated blacks remained virtually unchanged, the proportion of less educated whites fell sharply. The table also reports first-difference changes in MSA characteristics of the same variables we controlled for in levels in Table 4.

Table 5 reports the estimation results for four specifications given by Eq. (3). We regress the change in the MSA-level dissimilarity index between 1990 and 2000 on a variety of measures of the change in the sociodemographic composition of the metropolitan area over the same period, along with other metropolitan controls.

Column 1 reveals a strong positive relationship between the change in the fraction of blacks with a college degree in the MSA population and the change in segregation, controlling for other changes in the education composition of the MSA and in log population. Specifically, our estimate indicates that a one-standard deviation increase in the fraction of highly educated blacks, holding fixed the education composition of whites, would lead to about a one-percent increase in the dissimilarity index. This is a large positive effect of the order of a quarter of a standard deviation in the change in the dissimilarity index over the decade. The finding is robust to the inclusion of additional MSA-level covariates, measuring changes in median household income and manufacturing share, as shown in Column 2.

To further investigate the role of age structure along the lines hypothesized above, we break the effect of changes in the proportion of highly educated blacks in an MSA relative to whites down by age in Columns 3 and 4. Specifically, we measure the effects of changing the proportion of highly educated blacks in two separate age categories, 25–44, and 45 or older, respectively, on the change in the MSA segregation; we also break down the other education controls (for less-educated blacks, and highly and less-educated whites) in the same way. In Column 3, we control only for changes in log population, while in Column 4 we also control for changes in the MSA median household income and manufacturing share between 1990 and 2000. The results from this age disaggregation reported in Columns 3 and 4 make very clear that it is changes in the proportion of *older* highly educated blacks, aged 45 and above rather than 25–44, that affect residential segregation in first differences. Indeed, the estimates indicate that effectively all the positive impact of changes in the proportion of highly

[35] For this analysis, unlike in Section 4.1, we no longer restrict attention to MSAs that have at least 100,000 individuals and 10,000 blacks, which increases the sample of MSAs slightly. Our results are not sensitive to this.

[36] In a careful study using data from the U.S. Postal Service, Boustan and Margo (2009) find evidence that the relationship between black postal employment and segregation has declined in recent decades.

Table 5
Segregation and metropolitan composition – first differences (1990–2000).

Dependent variable		Δ dissimilarity index (1990–2000) (1)	Δ dissimilarity index (1990–2000) (2)	Δ dissimilarity index (1990–2000) (3)	Δ dissimilarity index (1990–2000) (4)
Δ Metropolitan Characteristics					
% Black with College Degree	**All adults**	**2.573***** **(0.728)**	**2.924***** **(0.788)**		
	Age 25–44			−1.042 (1.311)	−0.273 (1.346)
	Age 45+			**4.489***** **(1.319)**	**3.504**** **(1.461)**
% Black with less than College Degree	All adults	−0.031 (0.226)	−0.159 (0.167)		
	Age 25–44			−0.153 (0.337)	−0.614 (0.356)
	Age 45+			0.528 (0.487)	0.867 (0.506)
% White with College Degree	All adults	−0.009 (0.127)	−0.009 (0.123)		
	Age 25–44			−0.053 (0.198)	0.015 (0.201)
	Age 45+			0.088 (0.190)	0.021 (0.188)
% White with less than College Degree	All adults	0.396*** (0.095)	0.275*** (0.092)		
	Age 25–44			0.233 (0.148)	0.157 (0.145)
	Age 45+			0.427*** (0.131)	0.295** (0.142)
Controls for Change in log (Population)?		Yes	Yes	Yes	Yes
Controls for Changes in Other Metropolitan Variables?		No	Yes	No	Yes

Notes: The table reports coefficients and standard errors from two regressions of the change in the metropolitan dissimilarity index between 1990 and 2000 on the changes in metropolitan sociodemographic composition over this same period. All four columns include the change in log (population) as a control. Columns (2) and (4) also include changes in log (median income) and in manufacturing share. Regressions are based on the sample of metropolitan areas that appear in both 1990 and 2000 ($N = 214$).
*** Denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

educated blacks comes through the older age category, with estimated effect sizes similar to those for highly educated blacks in Columns 1 and 2; in contrast, the effects for the younger group are actually slightly negative, if indistinguishable from zero. This striking age pattern is again consistent with the prior discussion, to the effect that the neighborhood formation mechanism and the neighborhood effects mechanism of CG seem to cancel out for younger adults, leaving no significant net relationship.

5. Complementary analysis and robustness

5.1. Neighborhood formation

We now provide evidence of neighborhood formation using the same organization of the first-differenced MSA data as in Section 4.2. Specifically, we examine the relationship between the changes in the proportion of highly educated blacks and the changes in the number of middle-class black neighborhoods within an MSA, conditioning on changes in other MSA sociodemographics.

Table 6 provides evidence relating to the formation of middle-class black neighborhoods. Columns 1–4 show results based on different definitions of "middle-class black neighborhood," the dependent variable being the change in the log number of tracts satisfying the given definition in the column heading. Column 1 shows that a one percentage point increase in the proportion of highly educated blacks, controlling for the education of whites, is associated with a 22% increase in the number of middle-class black communities, defined as tracts that are both at least 60% black and 40% college educated. The estimated effects are even larger when considering broader definitions of "middle-class black neighborhood."[37]

While Table 6 shows that increases in the proportion of college-educated blacks are associated with sharp increases in the number of middle-class neighborhoods in the MSA, the life-cycle logic we emphasized in Section 4 suggests that, to the extent that residential choices are made mostly by relatively older individuals, we should expect to see stronger associations between the changes in the number of middle-class neighborhoods and the changes in the proportion of *older* college-educated blacks. This is confirmed in Table 7, where we report the effects of changing the proportion of younger versus older highly educated blacks in an MSA on middle-class black neighborhood formation, again conditioning on other sociodemographics. It shows the consistently positive impact of increasing the proportion of older college-educated blacks (aged 45 and above), while the effects of changes in the proportions of younger college-educated blacks (aged 25–44) tend to be smaller, and are insignificant for the narrowest definition of middle-class black neighborhoods (at least 60% black and at least 40% college-educated).[38]

[37] In each column, we control for changes in log population of the MSA. The results are robust to the inclusion of changes in log (median income) and manufacturing share.

[38] Using the Neighborhood Change Database from Geolytics, we find suggestive evidence that as a neighborhood transitions into a middle-class black neighborhood, blacks (especially college educated blacks) move in the neighborhood, while college-educated whites move out and less-than-college-educated whites move in.

Table 6
Middle-class black neighborhoods and metropolitan composition – first differences.

Dependent variable	Δ log (number of tracts in MSA >60% black and >40% college-educated) (1)	Δ log (number of tracts in MSA >60% black and >20% college-educated) (2)	Δ log (number of tracts in MSA >40% black and >40% college-educated) (3)	Δ log (number of tracts in MSA >40% black and >20% college-educated) (4)
Δ Metropolitan characteristics				
% Black with College Degree	**22.03*****	**39.54*****	**35.98*****	**53.36*****
	(4.91)	**(5.96)**	**(5.17)**	**(5.88)**
% Black with less than College Degree	−6.76***	−9.67***	−6.76**	−7.33***
	(1.52)	(1.86)	(1.60)	(1.83)
% White with College Degree	−1.85**	0.21	−1.73*	0.73
	(0.85)	(1.03)	(0.89)	(1.02)
% White with less than College Degree	−1.45**	−0.29	−0.78	0.15
	(0.63)	(0.76)	(0.66)	(0.75)

Notes: The table reports coefficients and standard errors from four regressions that relate changes in the number of middle-class black neighborhoods between 1990 and 2000 to the change in metropolitan educational composition over this same period, along with the change in the log of the metropolitan area population. Regressions are based on the sample of metropolitan areas that appear in both 1990 and 2000. *** Denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

Table 7
Segregation, middle-class black neighborhoods and metropolitan composition – first differences.

Dependent variable	Δ log (number of tracts in MSA >60% black and >40% college-educated) (2)	Δ log (number of tracts in MSA >60% black and >20% college-educated) (3)	Δ log (number of tracts in MSA >40% black and >40% college-educated) (4)	Δ log (number of tracts in MSA >40% black and >20% college-educated) (5)
Δ Metropolitan characteristics				
% Black with College-Degree				
Adults – age 25–44	10.44	31.01***	31.97***	38.85***
	(9.12)	(11.13)	(9.61)	(10.85)
Adults – age 45+	33.72***	47.49***	47.31***	61.92***
	(9.11)	(11.12)	(9.61)	(10.84)

Notes: The table reports coefficients and standard errors from four regressions that relate changes in the number of middle-class black neighborhoods between 1990 and 2000 to the change in metropolitan educational composition over this same period, focusing on changes in the percentage of younger versus older blacks with a college degree. Regressions are based on the sample of metropolitan areas that appear in both 1990 and 2000. *** Denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

5.2. Across-MSA sorting

The results presented in Table 4–7 provided evidence consistent with our neighborhood formation mechanism. The development of that mechanism focused implicitly on within-MSA sorting, yet one can envisage a more general version of the same sorting story that involves migration *across* MSAs. In this subsection, we consider the extent to which a positive relationship between segregation and racial inequality might be due to *across*-MSA sorting, where highly educated blacks differentially migrate to MSAs with more middle-class black neighborhoods, rather than within-MSA sorting.

To address the likely strength of the across-MSA sorting channel, we make use of rich Census microdata providing information on the metropolitan area in which each individual lived *five years prior to the Census*. These data allow us to examine the extent to which highly educated blacks are drawn disproportionately to metropolitan areas that have a larger number of middle-class black neighborhoods. Such a migration pattern could generate the kinds of cross-sectional results shown for older adults in Table 4 if black in-migrants were significantly more educated than those who already lived in segregated metropolitan areas.

Table 8 reports the results of a series of regressions that relate the neighborhood structure in an individual's current metropolitan area to a set of individual education-race categories for a sample of individuals aged 20–30.[39] The dependent variable in the set of regressions shown in Columns 1–3 is the number of tracts in the individual's current MSA that are at least 60% black and 40% college-educated.

The regression shown in Column 1 is estimated on a sample of individuals who moved to a new MSA between 1995 and 2000 and includes fixed effects for the MSA the individual resided in 5 years prior to the Census year. In essence, this specification compares the characteristics of newly-chosen metropolitan areas for two individuals who resided in the same metropolitan area five years ago. The results demonstrate clearly that college-educated blacks are indeed more likely to choose MSAs with a greater number of neighborhoods that are at least 60% black and 40% college-educated than all other types of individuals. For example, relative to college-educated whites leaving the same MSA, college-educated blacks choose MSAs that have an average of 0.9 more tracts meeting these criteria (the average number of such tracts for all U.S. metropolitan areas is only 0.3). Such across-MSA sorting is clearly consistent with the notion that metropolitan areas with a higher fraction of middle-class black neighborhoods are particularly attractive to college-educated blacks. This finding accords both

[39] We focus on these younger adults on the basis that they are more likely than others to move to a new metropolitan area during a given five-year period.

Table 8
Assessing across-metropolitan area sorting.

Dependent variable	Number of tracts in MSA >60% Black and >40% College-Educated			Number of tracts in MSA >40% Black and >40% College-Educated		
Sample	Movers (1)	Movers (2)	Stayers (3)	Movers (4)	Movers (5)	Stayers (6)
Individual characteristic						
Black with College Degree	1.075*** (0.107)	1.165*** (0.147)	0.903 (0.812)	2.702*** (0.254)	3.104*** (0.326)	2.798*** (1.198)
Black with less than College Degree	0.197*** (0.054)	0.253*** (0.087)	0.380 (0.681)	0.079 (0.129)	0.372** (0.186)	1.463 (1.293)
White with College Degree	0.157*** (0.053)	0.17* (0.094)	−0.248 (0.577)	0.833 (0.110)	1.144*** (0.160)	0.126 (0.950)
White with less than College Degree	−0.499*** (0.052)	−0.561*** (0.075)	−0.704 (0.562)	−1.38*** (0.139)	−1.446*** (0.141)	−1.609 (0.969)
Includes fixed effects for MSA of residence 5 years prior to Census?	Yes	No	No	Yes	No	No

Notes: The six regressions reported in this table each relate a measure of the availability of middle-class black neighborhoods to an individual's race-education category. All regressions use a sample of individuals aged 20–30 in 2000. Separate regressions are reported for individuals who moved between metropolitan areas and those who did not in the five years prior to the 2000 Census. For movers, a specification that includes fixed effects for the MSA of residence in 1995 is also reported. Standard errors adjusted for clustering at the metropolitan level are reported in parentheses. *** Denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

with individuals' same-race preference as specified in our model and the fact that most U.S. MSAs contain a very limited number of middle-class black neighborhoods.[40]

This kind of across-MSA sorting is *unlikely* to be responsible for the negative relationship between segregation and racial inequality we documented earlier. To that end, Columns 2 and 3 in Table 8 report the results of corresponding specifications for individuals who, respectively, do and do not migrate across MSAs during this five-year period, dropping the fixed effects for the lagged MSA.[41] The resulting coefficients reveal a remarkably similar pattern to those reported in Column 1. That an almost identical pattern obtains for stayers as movers implies that the proportion of college-educated blacks in the sample of migrants into MSAs with a greater number of middle-class black neighborhoods is *roughly the same* as the proportion of college-educated blacks already residing in these MSAs. Thus, while college-educated blacks do systematically migrate to MSAs with a high number of middle-class black neighborhoods, this migration *does not* systematically change the socioeconomic structure of these MSAs. In turn, this pattern of migration does not contribute to cross-sectional differences in MSA educational composition of the blacks in a systematic way, allowing us to rule out this type of sorting as an explanation for the positive relationship between segregation and black educational attainment relative to whites.[42]

5.3. Cohort effects? Cross-sectional results for 2000 Census

Another potential concern is that the age pattern we document in Table 4 is not a life-cycle effect as we argue, but rather represents cohort effects instead. To distinguish age effects from cohort effects, we report in Table A.5 the same analysis we carried out in Table 4 but using the 2000 instead of the 1990 Census.[43] Comparing interaction coefficients in each column of this table against the corresponding entries in Table 4 reveals a similar pattern and similar point estimates. In the case of college education, shown in Column 1, there is evidence of a mild steepening of the profile in 2000 relative to 1990 – slightly more negative to slightly more positive – and the estimates are somewhat more precise. For log earnings, the profile is flatter, becoming positive for the 41–50 age group rather than the 31–40 age group in 1990 (though in this latter case, the point estimate is imprecise). These estimates make clear that a very similar age profile to that reported in Table 4 for the 1990 Census emerges using 2000 Census data.[44]

6. Implications

The combined presence of the *neighborhood formation* mechanism (which predicts a negative segregation-inequality relationship) and the *neighborhood effects* mechanism of CG (which predicts a positive segregation-inequality relationship) points to the operation of a negative feedback loop that affects the joint evolution of residential segregation and racial socioeconomic inequality. Suppose that government policies aimed at improving inner city schools are able to reduce racial educational inequality in an MSA. Our neighborhood formation mechanism predicts that this will lead to an increase in segregation among blacks of all education levels; the increase in segregation will then, via CG's neighborhood effects mechanism, lead to lower educational attainment among young blacks relative to whites, undoing some of the initial reduction in racial inequality over time.[45] The operation of this negative feedback loop implies that the movement towards racial convergence will tend to be inhibited.[46]

We also note that the effects of the negative feedback may be mitigated when the proportion of highly educated blacks in an

[40] Frey (2004) provides interesting descriptive evidence relating to the "New Great Migration" since the 1990s, with blacks and especially college educated blacks moving to the South in increasing numbers.

[41] Additional fixed effects for the lagged MSA cannot be included for stayers since they did not move.

[42] As a further robustness check, Columns 4–6 repeat the analysis using the number of tracts in the individual's current MSA that are at least 40% black and 40% college-educated. These results are similar to those presented in Columns 1–3 in that there is little discernible difference when comparing movers and stayers.

[43] The summary statistics for the 2000 micro Census data are provided in Table A.6.

[44] Related, Collins and Margo (2000) report the interaction coefficient from a series of CG-style regressions for the log (earnings) of individuals aged 20–30 as far back as the 1940 Census. They estimate effects of roughly the same magnitude (though statistically insignificant) as that reported by CG for 1990, and interpret this as evidence supporting the notion that "ghettos did not *turn* bad" in more recent decades.

[45] The strength of CG's neighborhood effects mechanism relating racial segregation to black outcomes crucially depends on where the middle-class blacks forming their own neighborhoods are moving from. Indeed Table 9 below shows that the negative feedback may be mitigated when the proportion of highly educated blacks in an MSA is sufficiently high. In such cases, the increase in racial segregation is a result of middle-class blacks trading places with middle-class whites as middle-class neighborhoods emerge; as such, the increase in racial segregation does not coincide with an increase in segregation in SES. As a result, CG's neighborhood effects mechanism is weaker.

[46] Loury (1977) draws attention to a negative externality in the accumulation of human capital, which gives rise to persistent differences in income across race.

Table 9
Enhanced Cutler–Glaeser regressions: The effect of metropolitan segregation on individual outcomes.

	Age 20–24				Age 25–30			
	HS Graduate (1)	College Graduate (2)	Ln (Earnings) (3)	Idle (4)	HS Graduate (5)	College Graduate (6)	Ln (Earnings) (7)	Idle (8)
Coefficients on interactions between black and metropolitan segregation (dissimilarity index) and proportion college-educated blacks in metro area reported								
Cutler–Glaeser Regressions								
Black * Metro Dissimilarity Index (Segregation)	−0.269*** (0.041)	−0.094*** (0.032)	−0.788*** (0.140)	0.340*** (0.031)	−0.201*** (0.039)	−0.064 (0.062)	−0.433*** (0.094)	0.310*** (0.038)
Adding Interactions with (% Metro Black and College-Educated)								
Black * Segregation	−0.412*** (0.080)	−0.101*** (0.039)	−1.123*** (0.260)	0.387*** (0.070)	−0.241*** (0.072)	−0.016 (0.065)	−0.505*** (0.164)	0.394*** (0.083)
Black * Segregation * (%Metro Black and College Educated)	**13.60*** (4.32)**	**2.40 (3.30)**	**20.06* (12.25)**	**−5.26* (3.09)**	**6.21* (3.67)**	**−0.88 (5.94)**	**6.07 (7.93)**	**−7.45** (3.06)**
Black* (%Metro Black and College Educated)	−8.89*** (2.95)	−2.13 (2.36)	−9.36 (8.81)	4.00* (2.49)	−3.02 (2.59)	0.54 (4.22)	−0.82 (5.32)	4.51** (2.03)

Notes: This table reports the results of a series of OLS regressions based on the specifications in Table 4 of Cutler and Glaeser (1997). Each specification includes individual characteristics [Black, Asian, Other nonwhite, Hispanic, Female], metropolitan characteristics [segregation, ln (population), % black, ln (median household income), manufacturing share] and interactions of these metropolitan characteristics with whether the individual is black. The upper panel replicates their results, reporting the coefficient on the interaction between whether the individual is black and metropolitan segregation. The lower panel reports the results of regressions that add interactions with the proportion of the metropolitan population that is college-educated and black. This measure is included directly and interacted with the level of segregation in the metropolitan area. Both of these variables are in turn interacted with a dummy indicating whether the individual is black. Coefficients are reported for only interactions with whether the individual is black. All regressions use the metropolitan variables used by Cutler and Glaeser, which Jacob Vigdor has generously made available on his website. Standard errors are reported in parentheses. *** Denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

MSA is sufficiently high. To see this, Table 9 reports a series of OLS regressions using specifications similar to that of Eq. (2) in Table 4, with the exception that we now add the triple interaction $\text{Seg}_i \times \text{Black}_i \times$ (%Metro Black and College Educated).[47] Columns 1–4 focus on the sample of 20–24 age group and Columns 5–8, the 25–30 age group. As in Table 4, we examine the same four outcomes: high school graduation, college graduation, log earnings and whether idle. The coefficient estimates indicate that even though segregation is negatively correlated with black outcomes relative to whites for these two young age groups – a result confirming Cutler and Glaeser (1997, Table IV) – a significant exposure to highly educated blacks actually has a positive effect on individual outcomes. For example, for high school graduation, the coefficient estimate of the term $\text{Seg}_i \times \text{Black}_i \times$ (%Metro Black and College Educated) suggests that being exposed to the negative influence of segregation on younger blacks' high school graduation rate will be reduced by about 15% due to the exposure of more highly educated blacks. This result thus suggests the possibility that, when there is a sufficiently high proportion of highly educated blacks in an MSA, we may break out of the negative feedback loop and achieve a simultaneous reduction in residential segregation *and* racial inequality.

7. Conclusion

In this paper, we have argued that residential segregation may rise, somewhat counter-intuitively, when racial differences in education and other sociodemographics narrow. Motivated by the scarcity of middle-class black neighborhoods in many U.S. cities, we proposed a mechanism that could generate such a negative inequality-segregation relationship, involving a process of neighborhood formation. Increases in black socioeconomic status relative to whites would lead to the formation of new middle-class black neighborhoods, likely to be attractive to blacks (particularly those who are highly educated), permitting increases in residential segregation as inequalities across race narrow.

In order to examine the importance of this neighborhood formation mechanism in practice, we set out a two-part research design based on the distinctive cross-sectional and time-series predictions of the neighborhood formation mechanism vis-a-vis competing mechanisms that are also likely to influence the inequality-segregation relationship. Implementing this two-part design using Census data, we show that there is a negative cross-sectional relationship between inequality and segregation for older blacks, based on both the 1990 and 2000 Censuses. Across time, we show that increases in the proportion of highly educated blacks in a metropolitan area between 1990 and 2000, controlling for the education of whites, are associated with greater racial segregation; and we find even stronger effects among older blacks, likely reflecting the strength of neighborhood formation mechanism alone. Further, increases in average educational attainment of blacks relative to whites are associated with sharp rises in the number of middle-class black neighborhoods, consistent with the neighborhood formation mechanism.

We noted that racial inequality and residential segregation are linked in an intergenerational feedback loop, with segregation influencing racial inequality in the present, which then affects household residential sorting and so future segregation, in turn influencing socioeconomic inequality. Our results, in combination with the findings of Cutler and Glaeser (1997), indicate that the feedback is negative in character, likely to inhibit reductions in segregation and racial inequality over time. We also identified conditions under which the effects of this negative feedback loop are mitigated, when the proportion of highly educated blacks in an MSA is sufficiently high.

It is worth drawing attention to another channel, related to the formation of preferences, that might also serve to weaken the negative feedback loop. The results in our paper are based entirely on the recent range of data across U.S. cities, which never get very close to racial equality. It is possible that if racial inequalities in education and income were reduced significantly, so the general strength of racial preferences might also weaken, with racial segregation declining as a result. Endogenizing preferences in this fashion presents an important, as well as challenging, area for future work. To explore the inter-related dynamics of segregation and inequality more explicitly calls for a dynamic model incorporating both neighborhood sorting and education production. The task of developing and estimating such a model awaits further research.

Acknowledgments

We are grateful to Editor Stuart Rosenthal and two anonymous referees, whose numerous suggestions helped us improve the

[47] Of course, we also add the interaction $\text{Black} \times$ (%MetroBlack and College Educated) in Table 9.

paper considerably. We would also like to thank Joe Altonji, Victor Couture, Christoph Esslinger, Richard Freeman, Roland Fryer, Mike Gilraine, Ed Glaeser, Caroline Hoxby, Matt Kahn, Larry Katz, Robert Moffitt, Derek Neal, Steve Pischke, Richard Rogerson, Kim Rueben, Olmo Silva, Will Strange, Matt Turner, Chris Udry, Jacob Vigdor, Bruce Weinberg, and seminar/conference participants at Harvard, LSE, Minnesota, Penn State, Toronto, UBC, USC, UVA, Washington University at St. Louis, Yale and the NBER for useful comments and suggestions. Branko Boskovic, Jon James, and Hugh Macartney provided excellent research assistance. The U.S. Department of Education, the NSF, and SSHRC provided financial support for this research. All remaining errors are our own.

Appendix A

In this appendix, we illustrate our mechanism using a six-community example, which serves as the basis for Figs. 1 and 2.

We assume that the six neighborhoods are equal-sized, i.e. $J = 6$ and $n_j = 1/6$ for all j. The neighborhoods differ in amenity levels, with $q_1 = q_2 = 2, q_3 = q_4 = 1$ and $q_5 = q_6 = 0$. Also suppose that $\lambda = 3/8,\ \rho_B = 1/3$ initially, and $\rho_W = 3/5$, so the total fraction of highly educated is 1/2, i.e. $\sum_{r\in\{b,w\}}\rho_r\lambda_r = 0.5$. Finally, we assume that α among the highly educated is distributed uniformly on $[400, 1000]$, while it is distributed uniformly on $[0, 600]$ among the less-educated, thereby allowing highly educated households

Table A.1
Metropolitan Areas in 2000 with Tracts Combining High Fractions of Black and College-Educated Individuals.

	Number of tracts meeting both race and education criteria			Population 25 years and older	Fraction black	Fraction of blacks with college degree
Percentage black	>80%	>60%	>40%	(in millions)		
Percentage with college degree	>40%	>40%	>40%			
Baltimore-Washington	5	14	33	5.06	0.24	0.21
Detroit	5	8	19	3.51	0.19	0.13
Chicago		3	16	6.11	0.16	0.15
New York		4	15	14.88	0.15	0.17
Los Angeles	4	6	10	11.50	0.06	0.18
Atlanta	5	5	8	2.65	0.26	0.22
Cleveland		1	6	1.96	0.15	0.11
Philadelphia		1	5	4.12	0.17	0.13
San Francisco-Oakland			5	4.95	0.06	0.19
Raleigh–Durham		1	3	0.65	0.12	0.22
Indianapolis			3	1.05	0.12	0.14
Jackson, MS	1	1	2	0.44	0.25	0.17
Houston	1	1	2	3.10	0.15	0.18
Columbia, SC			2	0.59	0.17	0.17
New Orleans			2	0.85	0.33	0.13
All U.S. Metropolitan Areas	21	44	142	154.84	0.11	0.15

Notes: Tract compositions are calculated using individuals 25 years and older in U.S. metropolitan areas. Tracts considered in this table have a minimum of 800 such individuals.

Table A.2
Metropolitan area characteristics for 1990.

Variable	Residential Segregation (Dissim. Index)	ln (MSA population)	Percent Black	ln (median income)	Manufacturing Share
MSA-Level Descriptives for 1990					
Number of MSAs	209	209	209	209	209
Mean	0.586	13.1	0.138	10.3	0.172
Standard Deviation	0.126	1.0	0.092	0.2	0.069
Minimum	0.206	11.6	0.009	9.9	0.036
Maximum	0.873	16.0	0.457	11.0	0.456

Notes: The sociodemographic variables are obtained using the 5% sample of the 1990 Census.

Table A.3
Summary statistics for 1990 census micro data.

Variable	Age 20–24		Age 25–30		Age 31–50		Age 51–70	
	White	Black	White	Black	White	Black	White	Black
Education								
High school graduate	87.7%	75.9%	89.9%	79.2%	90.9%	77.4%	76.4%	48.8%
College graduate	14.8%	5.0%	28.8%	12.3%	31.5%	15.3%	19.7%	9.2%
Work and income								
Idle	12.6%	30.6%	14.7%	27.6%	15.5%	23.7%	48.8%	52.8%
ln (earnings)	9.3	9.0	9.8	9.5	10.1	9.8	10.0	9.7
Demographic variables								
Black	15.1%		13.3%		12.3%		11.2%	
Asian	1.4%		1.1%		0.9%		0.9%	
Other nonwhite	4.3%		3.4%		2.3%		1.3%	
Hispanic	8.0%		6.2%		4.0%		2.8%	
Female	51.6%		51.6%		51.6%		53.5%	
N	417,838		627,503		1,766,671		1,051,655	

Notes: This table is analogous to Cutler and Glaeser (1997) Table 2, though using the 1990 Census 5% rather than the 1% sample.
It adds columns for ages 31–50 and 51–70. The education categories are not exclusive. 'Idle' corresponds to not working and not being in school.

Table A.4
Segregation and metropolitan characteristics – first-differenced (1990–2000) at the MSA level.

Variable	Mean	Standard Deviation	Minimum	Maximum
Dissimilarity Index	−0.054	0.041	−0.197	0.115
% Black with College Degree	0.002	0.004	−0.008	0.018
Age 25–44	0.000	0.002	−0.011	0.010
Age 45+	0.002	0.003	−0.002	0.011
% Black with less than College Degree	0.001	0.020	−0.060	0.142
Age 25–44	−0.002	0.012	−0.039	0.057
Age 45+	0.003	0.010	−0.021	0.085
% White with College Degree	0.023	0.023	−0.035	0.125
Age 25–44	−0.011	0.014	−0.061	0.049
Age 45+	0.035	0.016	−0.003	0.103
% White with less than College Degree	−0.066	0.057	−0.179	0.593
Age 25–44	−0.056	0.031	−0.107	0.257
Age 45+	−0.009	0.033	−0.106	0.336
Log (Population)	0.177	0.175	−0.161	1.009
Log (Median Household Income)	0.344	0.055	0.188	0.542
Manufacturing Share	−0.046	0.039	−0.196	0.037
Percent Black	0.005	0.013	−0.038	0.051

Notes: This table is based on a sample of 214 MSAs.

Table A.5
Segregation and metropolitan composition – age profile in 2000.

Dependent variable	College Graduation (1)	Ln (Earnings) (2)	High School Graduation (3)	Idle (4)
Coefficient on interaction between black and metropolitan segregation (dissimilarity index)				
Age category				
20–24	−0.137***	−0.478***	−0.271***	0.321***
	(0.026)	(0.103)	(0.037)	(0.053)
25–30	−0.101**	−0.187***	−0.151***	0.291***
	(0.044)	(0.072)	(0.034)	(0.042)
20–30	−0.118***	−0.299***	−0. 207***	0.304***
	(0.033)	(0.074)	(0.032)	(0.040)
31–50	0.005	0.026	−0.104***	0.237***
	(0.035)	(0.058)	(0.039)	(0.027)
31–40	−0.038	−0.087	−0.138***	0.247***
	(0.037)	(0.074)	(0.040)	(0.031)
41–50	0.047	0.148**	−0.063	0.223***
	(0.036)	(0.062)	(0.044)	(0.028)
51–70	0.084***	0.306v	−0.037	0.076***
	(0.033)	(0.088)	(0.058)	(0.027)
51–60	0.095**	0.305***	−0.034	0. 126***
	(0.041)	(0.091)	(0.053)	(0.028)
61–70	0.063**	0.307*	−0.030	0.013
	(0.031)	(0.162)	(0.076)	(0.040)

Notes: This table reports interaction coefficients using the same specifications underlying Table 4, though using 2000 rather than 1990 Census data. *** Denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

Table A.6
Summary statistics for 2000 micro data.

Variable	Age 20–24		Age 25–30		Age 31–50		Age 51–70	
	White	Black	White	Black	White	Black	White	Black
Education								
High school graduate	85.8%	76.1%	87.7%	82.0%	89.5%	80.9%	83.8%	66.1%
College graduate	13.7%	6.2%	33.1%	16.4%	32.2%	16.7%	27.3%	13.8%
Work and income								
Idle	14.3%	29.1%	17.6%	28.1%	18.8%	29.3%	43.6%	52.1%
ln (earnings)	9.6	9.3	10.1	9.9	10.4	10.1	10.3	10.1
Demographic variables								
Black	13.9%		12.9%		12.3%		10.6%	
Asian	5.4%		6.1%		5.0%		4.1%	
Other nonwhite	15.1%		13.6%		9.0%		5.1%	
Hispanic	21.2%		19.9%		13.3%		8.5%	
Female	49.5%		50.3%		50.9%		52.4%	
N	640,546		823,613		3,019,416		1,724,116	

to have higher willingness to pay for amenities, though with some overlap. (We assume uniform distributions for analytic convenience.)

A.1. No same-race preferences ($\beta = 0$)

Given the illustrative parameterization, in the essentially unique sorting equilibrium, the high-amenity neighborhoods 1 and 2 will be occupied only by highly educated households with α in the interval $[600, 1000]$; the medium-amenity neighborhoods 3 and 4 will be occupied by a 50/50 mixture of highly educated and less-educated residents with their α lying in the interval $[400, 600]$; and the low-amenity neighborhoods 5 and 6 will be occupied only by the less-educated, with their α's in the interval $[0, 400]$. The equilibrium housing prices are $p_1 = p_2 = 1000$, $p_3 = p_4 = 400$ and $p_5 = p_6 = 0$.

Under the assumption that the race of residents in a particular community is randomly drawn from blacks and whites given their educational attainment, the racial compositions of each neighborhood in the equilibrium described above are as follows: the fraction of residents in neighborhoods 1 and 2 who are black is 25%, the fraction in neighborhoods 3 and 4 is 37.5%, and in neighborhoods 5 and 6, it is 50%. Thus in neighborhoods 1 and 2, a black household's exposure rate to whites is 3/4, given that 75% of the residents are white; similarly, in neighborhoods 3 and 4, black households' exposure rate to whites is 5/8; and black households' exposure to whites in neighborhoods 5 and 6 is 1/2. Since the fraction of blacks living in neighborhoods 1 and 2, 3 and 4, and 5 and 6, respectively, are 1/9, 1/3 and 4/9, the average exposure rate of blacks to whites in this initial equilibrium is given by $\frac{2}{9} \times \frac{3}{4} + \frac{1}{3} \times \frac{5}{8} + \frac{4}{9} \times \frac{1}{2} = 43/72$.

We noted in the text that when sorting occurs solely on the basis of education and the associated taste for the amenity, some racial segregation arises initially simply because race is correlated with education. The exposure rate of 43/72 in the sorting equilibrium when $\rho_B = 1/3$ and $\rho_W = 3/5$ is lower than the overall proportion of whites in the population, $1 - \lambda = 5/8$, which is the exposure rate that would arise under "random spreading." This conforms to the logic in Schelling's argument rehearsed in the Introduction.

A.2. Strictly positive same-race preferences ($\beta > 0$)

As noted in the main text, analytical solutions are difficult to obtain in the more general case where households care about the race of their neighbors in addition to amenity levels. We thus solve for the model's equilibria using numerical methods. Here, we apply a variant of the algorithm that solves numerically for sorting equilibria presented in Bayer et al. (2011). Given some starting allocation of households to communities and a vector of initial house prices, the first step of the algorithm involves calculating household demands over the available communities, allowing for same-race preferences over neighborhood racial composition. From these demands, we compute a set of prices to clear the housing market. Next, households are re-allocated to their preferred communities at these market-clearing prices. Then we re-calculate household demands over communities, given the new neighborhood compositions, compute a new set of market-clearing prices, and continue iteratively until the process converges. The results are shown in Figs. 1(b) and 2.

References

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

Arrow, Kenneth J., 1973. The theory of discrimination. In: Ashenfelter, O., Rees, A. (Eds.), Discrimination in Labor Markets. Princeton University Press, Princeton.

Bayer, Patrick, McMillan, Robert, 2012. Tiebout sorting and neighborhood segregation. Journal of Public Economics 96, 1129–1143.

Bayer, Patrick, McMillan, Robert, Rueben, Kim, 2004. What drives racial segregation? New evidence using census microdata. Journal of Urban Economics 56 (3), 514–535.

Bayer, Patrick, McMillan, Robert, Rueben, Kim, 2011. An Equilibrium Model of Sorting in an Urban Housing Market. Mimeo, Duke University.

Berry, Steven., Waldfogel, Joel., 2003. Product Quality and Market Size, NBER Working Paper 9675.

Boustan, Leah Platt, Margo, Robert A., 2009. Race, segregation, and postal employment: new evidence on spatial mismatch. Journal of Urban Economics 65, 1–10.

Brueckner, Jan K., Rosenthal, Stuart S., 2009. Gentrification and neighborhood housing cycles: will America's future downtowns be rich? Review of Economics and Statistics 91 (4), 725–743.

Charles, Camille Z., 2000. Neighborhood racial-composition preferences: evidence from a multi-ethnic metropolis. Social Problems 47, 379–407.

Charles, Camille Z., 2001. Processes of racial residential segregation. In: O'Connor, Alice, Tilly, Chris, Bobo, Lawrence D. (Eds.), Urban Inequality: Evidence from Four Cities. Russell Sage Foundation, New York.

Collins, William J., Margo, Robert A., 2000. Residential segregation and socioeconomic outcomes: when did ghettos go bad? Economics Letters 69, 239–243.

Cornell, S., Hartmann, D., 1997. Ethnicity and Race: Making Identities in a Changing World. Pine Forge Press, Thousand Oaks.

Cutler, David, Glaeser, Edward, 1997. Are ghettos good or bad? Quarterly Journal of Economics 112, 826–872.

Cutler, David, Glaeser, Edward, Vigdor, Jacob, 1999. The rise and decline of the American ghetto. Journal of Political Economy 107 (3), 455–506.

Farley, R., Steeh, C., Krysan, M., Jackson, T., Reeves, K., 1994. Stereotypes and segregation: neighborhoods in the Detroit area. American Journal of Sociology 100, 750–780.

Frey, William H., 2004. The New Great Migration: Black Americans' Return to the South, 1965-2000. The Living Cities Census Series. Brookings Institution.

Gabriel, Stuart A., Painter, Gary D., 2012. Household location and race: a 20-year retrospective. Journal of Regional Science 52 (5), 809–818.

Galster, G., 1982. Black and white preferences for neighborhood racial composition. AREUEA Journal 10, 39–66.

Glaeser, Edward L., Kahn, Matthew E., Rappaport, Jordan, 2008. Why do the poor live in cities? The role of public transportation. Journal of Urban Economics 63, 1–24.

Harsman, Bjorn, Quigley, John M., 1995. The spatial segregation of ethnic and demographic groups: comparative evidence from Stockholm and San Francisco. Journal of Urban Economics 37, 1–16.

Iceland, John, Weinberg, Daniel H., Steinmetz, Erika, 2002, Racial and Ethnic Residential Segregation in the United States: 1980–2000. U.S. Census Bureau, Census Special Report, CENSR-3, Washington, DC. U.S. Government Printing Office.

Ihlanfeldt, Keith, Scafidi, Ben, 2002. Black self-segregation as a cause of housing segregation: evidence from the multi-city study of urban inequality. Journal of Urban Economics 51, 366–390.

LeRoy, Stephen F., Sonstelie, Jon, 1983. Paradise lost and regained: transportation innovation. Income, and Residential Location Journal of Urban Economics 13, 67–89.

Loury, Glenn C., 1977. A Dynamic Theory of Racial Income Differences. Women, Minorities and Employment Discrimination, Lexington Books, Lexington, MA. pp. 153–188 (Chapter 8).

Lundberg, Shelly, Startz, Richard, 1998. Race, Information and Segregation. Mimeo, University of Washington at Seattle.

Miller, V., Quigley, John M., 1990. Segregation by racial and demographic group: evidence from the San Francisco bay area. Urban Studies 27, 3–21.

O'Flaherty, Brendan, 1999. Troubled Transactions and their Consequences: Race in the United States. Mimeo, Columbia University.

Pattillo, Mary, 2005. Black middle-class neighborhoods. Annual Review of Sociology 31, 305–329.

Schelling, Thomas C., 1969. Models of segregation. American Economic Review 59 (2), 488–493.

Schelling, Thomas C., 1971. Dynamic models of segregation. Journal of Mathematical Sociology 1, 143–186.

Sethi, Rajiv, Somanathan, Rohini, 2004. Inequality and segregation. Journal of Political Economy 112 (6), 1296–1321.

Thomas King, A., Mieszkowski, Peter, 1973. Racial discrimination, segregation, and the price of housing. Journal of Political Economy 81, 590–606.

Vigdor, Jacob L., 2003. Residential segregation and preference misalignment. Journal of Urban Economics 54 (3), 587–609.

Waldfogel, Joel, 2007. The Tyranny of the Market: Why You Can't Always Get What You Want. Harvard University Press.

Yinger, John, 1978. The black–white price differential in housing: some further evidence. Land Economics 54, 187–206.