Separate When Equal? Racial Inequality and Residential Segregation*

Patrick Bayer Hanming Fang Robert McMillan

July 6, 2005

Abstract

In contrast to conventional wisdom, this paper identifies a powerful mechanism which can lead to persistent and even *increasing* residential segregation when racial differences in education and other sociodemographics narrow. We document that middle-class black neighborhoods are in short supply in many U.S. metropolitan areas, forcing highly educated blacks either to live in white neighborhoods with high amenity levels or in more black neighborhoods with lower amenity levels. A simple model then shows that increases in the proportion of highly educated blacks in a metropolitan area may lead to the *emergence* of new middle-class black neighborhoods, relieving the prior neighborhood supply constraint and causing increases in residential segregation. Cross-MSA evidence from the 2000 Census indicates that this mechanism does in fact operate: as the proportion of highly educated blacks in an MSA increases, so the segregation of educated blacks and blacks more generally goes up. Our empirical findings are robust and have important implications for the evolution of residential segregation.

Keywords: Segregation, Racial Sorting, Racial Inequality, Neighborhood Formation.
JEL Classification Numbers: H0, J7, R0, R2.

---

*We are grateful to Joseph Altonji, Richard Freeman, Roland Fryer, Edward Glaeser, Caroline Hoxby, Larry Katz, Richard Rogerson, Kim Rueben, Matthew Turner, Chris Udry, Jacob Vigdor, Bruce Weinberg and seminar/conference participants at Harvard, Penn State, USC, Virginia, Yale and the NBER for helpful comments and suggestions. We are responsible for all remaining errors. Contact addresses: Bayer and Fang, Department of Economics, Yale University, P.O. Box 208264, New Haven, CT 06520-8264; McMillan, Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada.

1 Introduction

Racial segregation is a pervasive phenomenon in cities throughout the United States. In the substantial literature that studies its causes and consequences, a number of researchers have attempted to evaluate the contributions of socioeconomic characteristics other than race in explaining segregation.[1] Such studies typically find that a significant proportion of observed segregation can be explained by across-race differences in socioeconomic variables such as education and income.[2] These findings accord with Schelling [28]'s intuition that, because race is correlated with socioeconomic characteristics and these characteristics affect residential choices, some residential segregation would emerge even in the absence of explicit sorting on the basis of race; racial segregation would be a by-product, for example, of the selection of higher income individuals into bigger houses and nicer neighborhoods. A seemingly natural corollary is that a reduction in racial differences in socioeconomic variables would lead to a reduction in racial segregation.

This conventional wisdom is based on partial equilibrium perspective that takes the neighborhood structure in a metropolitan area to be fixed as the sociodemographics of the black population change. In a general equilibrium analysis, where neighborhood structure adjusts to changes in the distribution of socioeconomic characteristics, we hypothesize that the opposite may actually occur: reductions in black-white differences in socioeconomic characteristics may lead to *increases* in the segregation of educated blacks and blacks more generally.

Our hypothesis is motivated by a careful examination of the supply of middle-class black neighborhoods in US cities.[3] Three facts are particularly relevant. First, in almost every metropolitan area, few if any neighborhoods combine high fractions of both black and highly educated individuals. This shortage of middle-class black neighborhoods forces highly educated blacks to choose between predominantly black neighborhoods with low average education and predominantly white neighborhoods with high levels of education. Second, faced with a limited choice set in terms of

---

[1] Important contributions to this literature include Cutler and Glaeser [10], Cutler, Glaeser and Vigdor [11], and Massey and Denton [22].

[2] See Miller and Quigley [24], for example. Following a similar approach, Bayer, McMillan and Rueben [2] used restricted-access 1990 Census microdata to show that a set of sociodemographic variables that include education, income and language can explain 30 percent of Black segregation and 93 percent of Hispanic segregation in the Bay Area housing market. Sethi and Somanathan [26] propose a different method for decomposing segregation measures into one component that can be attributed to the effect of racial income disparities alone, and another component that combines the effects of neighborhood preferences and discrimination, and reached similar conclusions.

[3] See Section 2 for details.

neighborhood alternatives, highly educated blacks do in fact live in a very diverse set of communities: while a fraction live in neighborhoods with very few other black and many college-educated residents, many live in neighborhoods that have a high fraction of blacks and very few other college-educated residents. This raises the possibility that highly educated blacks might have preferred to live in highly educated majority-black neighborhoods, *were they available* – the lack of availability of middle-class black neighborhoods may be sharply binding.[4]

The third fact is that the supply of middle-class black neighborhoods is an increasing function of the number of highly educated blacks in the population: an increase in the average educational attainment of the black population of a typical metropolitan area can lead to a marked increase in the supply of middle-class black neighborhoods. Our hypothesis follows naturally, as an increase in supply may then lead to an increase in segregation, given that such neighborhoods provide an attractive alternative to highly-educated blacks in many metropolitan areas.

In Section 3, we present a stylized equilibrium model of decentralized residential choice, within-metropolitan area sorting and neighborhood formation which formalizes this mechanism.[5] We show that as more middle-class majority-black neighborhoods become available as a result of increased black educational attainment, highly educated blacks who currently live in middle-class white neighborhoods may move to more preferable middle-class black neighborhoods, leading to higher levels of residential segregation. Our model also makes the less obvious prediction that the exposure of highly educated blacks to other highly educated blacks, and blacks in general, may also increase with the proportion of highly educated blacks in the metropolitan area.

The empirical analysis at the heart of the paper takes seriously the neighborhood formation mechanism. Our primary analysis, presented in Section 4, uses Census Tract Summary Files from the 2000 Census to examine how changes in the composition of the population within a metropolitan area affect the way that individuals sort on the basis of race and education. The results show that, relative to others in the MSA, highly educated blacks are increasingly exposed to other blacks as the education level of blacks in the MSA increases. This change is driven primarily by a large increase in relative exposure to other highly educated blacks, but is more than completely offset by a decrease in relative exposure to highly educated whites. These changes lead to a slight decrease in the average educational attainment in the neighborhoods in which highly educated blacks reside.

---

[4]This is entirely consistent with Vigdor's [29] 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).

[5]A recent paper by Sethi and Somanathan [27] presents a different model in which they show that racial segregation and income inequality do not exhibit a monotonic relationship. See Section 3 for more discussion of this paper.

At the same time, highly educated blacks are also increasingly exposed to less educated blacks and vice versa. These empirical regularities are robust to controls for metropolitan area size and region; and they are all consistent with the comparative statics prediction of our model when the proportion of highly educated blacks in an MSA increases from a low to a moderate level.

Using this cross-sectional analysis as a baseline, we then explore in Section 5 the possibility that this cross-sectional positive correlation between segregation and black educational attainment may not be related to within-metro sorting as we propose, but may instead arise due to another mechanism. Before turning to any specific analysis, we emphasize that most of the leading alternative explanations for a correlation between these measures would imply a negative rather than positive correlation. Explanations that can be ruled out on this ground include statistical discrimination in either the housing or mortgage market,[6] or standard explanations related to within-metro sorting on the basis of socioeconomic characteristics (the conventional wisdom). We explore the following potential explanations in greater detail: (i) the impact of segregation on socioeconomic outcomes (reverse causation); (ii) across-metro sorting on the basis of observables; and (iii) across-metro sorting on unobservables.

Previous research, most notably Cutler and Glaeser [10, CG thereafter], suggests that the channel of reverse causation would result in a negative correlation. Specifically, using the 1990 Census, CG found that segregation at the metropolitan level substantially reduces relative educational and labor market outcomes for blacks *aged 20-30*. In light of this finding, it is quite surprising that we find a clear positive correlation between black educational attainment and segregation at the metropolitan level. We present a detailed analysis in Section 5.1 that reconciles our findings with CG's results: applying CG's analysis to older populations in the same dataset yields a large statistically significant positive effect. Given this age profile, we conclude that both mechanisms operate in the data, with each working to obscure the other.

The second alternative explanation that we explore in greater detail relates to across-metro sorting on observables (Section 5.2). In particular, using Census PUMS microdata, which characterizes where an individual resided five years prior to the survey, we examine whether highly educated blacks are drawn disproportionately to metropolitan areas that have a larger number of middle-class black neighborhoods. We find that this is indeed the case. This type of migration is clearly consistent with the broad narrative developed in this paper – that in many metropolitan areas,

[6] Taste-based discrimination is captured by our model, as it would be a reason for why agents prefer to live with neighbors of their own race, as specified by the utility function (1).

highly educated blacks are constrained by the short supply of middle-class black neighborhoods; as a result, they are more likely to migrate to metropolitan areas with their preferred middle-class neighborhoods. Equally importantly, however, the proportion of highly educated blacks among those migrating into metropolitan areas with a large number of middle-class black neighborhoods is comparable to the proportion in the population already residing there. Thus, this pattern of migration does not systematically contribute to cross-sectional differences in metropolitan composition, allowing us to rule out this type of sorting as an explanation for our baseline positive cross-sectional relationship between segregation and black educational attainment.

We then examine the possibility of across-metropolitan sorting on the basis of unobservable taste for segregation (Section 5.3). Such sorting would give rise to a classic form of selection bias if those highly educated blacks living in metro areas with a more educated black population have stronger unobserved tastes for segregation. To study this issue, we run a regression that essentially compares the neighborhood composition of individuals migrating into metropolitan areas that have a higher fraction of highly educated blacks to those who already reside there. This analysis reveals that highly educated blacks migrating into these metro areas choose less segregated neighborhoods, suggesting that, if anything, selection bias of this kinds works to slightly attenuate our main finding.

Taken together, these analyses support the notion that the positive correlation between metropolitan segregation and black educational attainment is in fact related to *within-metro area* sorting, confirming our main hypothesis. We conclude in Section 5.4 by presenting time-series evidence on the relationship between metropolitan population structure and segregation. Specifically, we regress the change in a measure of segregation (a dissimilarity index) in a metropolitan area between 1990 and 2000 on the changes in the sociodemographic composition of its population. We find that an increase in black educational attainment in a metropolitan area over time significantly increases its segregation, thus providing additional time-series support for our main hypothesis.

Our results have several important implications. First, in contrast with the conventional wisdom, they suggest that racial segregation is unlikely to disappear as racial differences in socioeconomic characteristics narrow.[7] The mechanism uncovered in our analysis indicates that an overall increase in the educational attainment of blacks may lead to a decrease in residential segregation *only if* highly educated blacks are dispersed in many, instead of concentrated in few, MSAs. This echoes Glaeser and Vigdor's [16] finding that segregation is lowest among the rapidly growing cities

[7] The conventional view has been embraced by many scholars in this literature. See, for example, Durlauf [13], Wilson [31] and Mayer [23].

in the West, where there is as yet no high concentration of highly educated blacks. Our findings also relate to Wilson [31], who argues that reductions in institutional discrimination in the housing market in the middle of the 20th century led to large-scale reductions in the exposure of less-educated to more-educated blacks, as more-educated blacks left the inner city neighborhoods to which they were formerly restricted. Based on our findings, this trend may not have been so severe in cities in which the black population was more educated initially; and it may partially reverse itself as the black population becomes relatively more educated over time.

The remainder of the paper is organized as follows. Section 2 documents empirically the types of neighborhood available across different metropolitan areas in the United States; Section 3 presents a simple model of neighborhood formation that highlights the key features of the mechanism underlying our empirical results; Section 4 presents our main empirical finding that the exposure of highly educated blacks, and blacks more generally, to other blacks increases as the proportion of highly educated blacks in the metropolitan area increases. Section 5 evaluates leading alternative hypotheses; and Section 6 concludes.

2 The Supply of Neighborhoods in U.S. Metropolitan Areas

In this section, we present some empirical facts regarding the supply of neighborhoods in U.S. metropolitan areas. These facts motivate our central hypothesis.

Throughout our analysis, we define metropolitan areas as either (i) free-standing Metropolitan Statistical Areas (MSAs) or (ii) Consolidated Metropolitan Statistical Areas (CMSAs) consisting of two or more economically and socially linked metropolitan areas – Primary Metropolitan Statistical Areas (PMSAs). (Henceforth, for expositional convenience, we will just use the term MSA.) For the most part, a "neighborhood" in our analysis corresponds to a Census tract, which typically contains 3,000 to 5,000 individuals; and we use publicly available Census Tract Summary Files (SF3) from the 2000 Census, which provide information on the distribution of education by race for each Census tract. Our focus is on non-Hispanic black and non-Hispanic white individuals 25 years and older residing in U.S. metropolitan areas. We characterize each neighborhood in a metropolitan area on the basis of two dimensions: *the fraction of residents that are black* and *the fraction of residents that are college-educated.*[8]

[8] Educational attainment is used to proxy socioeconomic status more generally: it is a better predictor for one's permanent income than current income in the Census year.

We establish four stylized empirical facts about neighborhood choice sets in the United States:

FACT 1. College-educated blacks constitute a small fraction of the population living in the typical metropolitan area;

FACT 2. Neighborhoods that combine high fractions of both college-educated and black individuals are in extremely short supply in almost every metropolitan area;

FACT 3. College-educated blacks choose to live in a very diverse set of neighborhoods in each metropolitan area;

FACT 4. Middle-class black neighborhoods are concentrated in only a few metropolitan areas that have sizeable numbers of college-educated blacks.

[Table 1 About Here]

Table 1 describes the joint distribution of education and race for blacks and whites. Based on our race definitions, blacks and whites respectively constitute 11.1 and 69.5 percent of the U.S. population 25 years and older residing in metropolitan areas. Among blacks, 15.4 percent have at least a four-year college degree, while the comparable number for whites is 32.5 percent. For the U.S. population as a whole 27.7 percent have at least a four-year college degree (not shown in Table 1).

[Table 2 About Here]

Table 2 documents the number of tracts in the U.S. by the percentage of individuals with a college degree and the percentage of individuals that are black and white, respectively. Panel A describes the number of tracts in which more than 0, 20, 40 and 60 percent of individuals 25 years and older are at least college-educated, respectively. Panel B reports the number of tracts in each of the categories listed in the column headings that contain a minimum fraction of blacks equal to 20, 40, 60, and 80 percent, respectively. As the corresponding numbers show, a much smaller fraction of the tracts with a high fraction black also have a high fraction of individuals with a college degree. For example, while 22.6 percent (row 1, column 3) of all tracts are at least 40 percent college-educated, only 2.5 percent (row 3, column 3) of tracts that are at least 40 percent black are at least 40 percent college-educated, and only 1.1 percent (row 4, column 3) of tracts that are at least 60 percent black are at least 40 percent college-educated. Panel C of Table 2 presents analogous numbers for whites. They show a markedly different pattern of neighborhood choices for whites, with a greater fraction of neighborhoods with at least 40, 60, and 80 percent whites meeting the education criteria listed in the column headings.

[Table 3 About Here]

While Table 2 reveals a scarcity of neighborhoods with high fractions of both black and college-educated individuals in the U.S. as a whole, Table 3 further shows that such tracts, to the extent that they exist, are concentrated in only a handful of metropolitan areas, most notably Washington, DC. This implies that the supply of such neighborhoods in most metropolitan areas is even more limited. Table 3 illustrates, for example, that of the 44 tracts (see row 4, column 3 of Table 2) that are at least 60 percent black and 40 percent college-educated, 14 are in Baltimore-Washington DC, 8 in Detroit, 6 in Los Angeles, and 5 in Atlanta. Almost 75 percent of these tracts can thus be found in these four MSAs only. Of the 142 tracts (see row 3, column 3 of Table 2) that are at least 40 percent black and 40 percent college-educated, almost two-thirds are in the MSAs listed above along with Chicago and New York.

Tables 2 and 3 taken together show clearly that while neighborhoods that combine high fractions of both college-educated and white individuals are abundant in all metropolitan areas, neighborhoods that combine high fractions of both college-educated and black individuals are in extremely short supply. This suggests that college-educated blacks in most metropolitan areas may face a trade-off between living with other black versus other college-educated neighbors.

To graphically illustrate this potential trade-off faced by highly educated blacks, Figure 1 shows the scatterplots of available neighborhoods in four metropolitan areas: Boston, Dallas, Philadelphia, and St. Louis. In each scatterplot, a circle represents a Census tract and its coordinates represent the fraction of college-educated individuals (vertical axis) and the fraction of blacks (horizontal axis) in the tract. The diameter of the 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.[9] For these four metropolitan areas, the scatterplots demonstrate the short supply of neighborhoods that combine high fractions of both highly educated and black individuals, neighborhoods that would have appeared in the north-east corner of the plot. They are strongly suggestive of the notion that highly educated blacks face a trade-off when making their residential choices.

[9] Note that tracts that do not contain any highly educated blacks do not appear in these scatterplots.

[Figure 1 About Here]

Figure 1 also demonstrates that, facing the constrained choice set, highly educated blacks choose to live in a diverse set of neighborhoods: while a sizeable fraction of college-educated blacks in each of the four MSAs choose neighborhoods with few black and many college-educated neighbors (neighborhoods in the north-western corner of the plots), another sizeable fraction choose neighborhoods with many black and few college-educated neighbors (neighborhoods in the south-eastern corner of the plots).

[Table 4 About Here]

Panel A of Table 4 summarizes the characteristics of neighborhoods chosen by college-educated blacks in metropolitan areas throughout the U.S. We first rank highly educated blacks in each metropolitan area by the fraction of blacks in their Census tract and assign individuals to their corresponding quintile of this distribution. This corresponds to drawing four vertical lines in the scatterplot for each metropolitan area such that an equal number of college-educated blacks fall into each of the resulting five regions. Panel A of Table 4 then summarizes the average fractions of black and college-educated individuals in the tract corresponding to the quintiles of this distribution, averaged over all U.S. metropolitan areas.

The numbers corresponding to different quintiles show 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 percent for those college-educated blacks living with the smallest fraction of black neighbors to 13.8 percent for those living with the largest fraction.

Panel B of Table 4 reports analogous numbers for college-educated whites. Comparison of Panels A and B reveals that college-educated blacks in each metropolitan area who live in the bottom quintile of tracts (in terms of 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 the top quintile of tracts (those with the greatest fraction of other blacks) have only about *one-third* of the fraction of highly educated neighbors. That such a high fraction of college-educated blacks in U.S. metropolitan areas choose segregated neighborhoods with relatively low average education attainment suggests that race remains an important factor in the location decisions of a large number college-educated blacks.[10]

[10] In Section 3, we present more evidence regarding race preferences.

[Figure 2 About Here]

Figure 2 depicts the scatterplots of neighborhoods in Atlanta, Chicago, Detroit and Washington DC – metropolitan areas that contain a more sizeable number of college-educated blacks, as shown in Table 3. Figure 2 illustrates that these metropolitan areas supply a substantially greater number of neighborhoods combining relatively high fractions of both black and highly educated individuals, and thus the constraint on the neighborhood choice set for highly educated blacks is relaxed there. As the neighborhood supply constraint is relieved for highly educated blacks, highly educated blacks may be increasingly exposed to other blacks.

Figures 1 and 2 together suggest that the constraint on the neighborhood choice sets for highly educated blacks will be systematically relaxed as the number of highly educated blacks in a metropolitan areas increases.[11,12] It is this neighborhood formation mechanism along with the documented short supply of middle-class black neighborhoods in the vast majority of US metro areas that motivates our central hypothesis. In particular, we hypothesize that an increase in the average educational attainment of blacks within a metro area allows middle-class black neighborhoods to form more readily and, consequently, leads to an increase in the residential segregation of highly educated blacks.

3 A Model

Having motivated our central hypothesis by examining the supply of neighborhoods in US metro areas, we now present a simple model of residential choice with endogenous neighborhood emergence. The simple model formalizes our idea that the supply of middle-class black neighborhoods is an increasing function of the average education of the black population in the metropolitan area.

[11] Indeed, regressions of the number or fraction of tracts in an MSA that are at least 40 percent college-educated and 40 percent black on metropolitan socioeconomic characteristics reveal a strong positive relationship with the fraction of college-educated blacks in the MSA. The number of such tracts is also, not surprisingly, increasing in the population of the MSA and a similar pattern holds for any combination of education and race criterion that count the number of tracts in the upper-right portion of the scatterplots.

[12] We also examined a series of quantile regressions designed to fit the 90th percentile of the relationship between neighborhood education and race shown in the scatterplots for college-educated blacks – that is, to approximate the implicit neighborhood availability constraint defined by the absence of neighborhoods in the upper-right portion of these scatterplots. These regressions demonstrate that the neighborhood availability constraint shifts significantly outward as the fraction of college-educated blacks in the MSA population is increased. This result holds no matter whether the fraction black or fraction of college-educated households in the MSA is held constant.

With this increased supply of middle-class black neighborhoods, the model predicts a clear increase in the segregation of highly educated blacks and, more subtly, a possible increase in the segregation of less educated blacks provided the number of highly educated blacks is reasonably small. Within the context of the broader paper, this stylized model serves to clarify the potential role of endogenous neighborhood formation in affecting segregation levels; the empirical analysis that follows does not rely on the specific assumptions of the model.

Sethi and Somanathan [27] present an alternative model in which they show that low levels of racial inequality are consistent with extreme and even rising levels of segregation in cities where the minority population is large. Their model does not explicitly emphasize the idea of neighborhood emergence since they treat the total number of neighborhoods as being exogenously fixed. In contrast, our model emphasizes the emergence of new middle-class neighborhoods, consistent with the empirical facts documented in Section 2.

Basic Ingredients. Before describing the detailed features of the model, we highlight three key ingredients that drive our results. The first key ingredient is an assumption that population preferences are such that, *taking housing prices into account*, individuals prefer to live near others of the same race and education level. This assumption is a statement about the indirect rather than the direct utility function. In terms of education-related sorting, it allows for the possibility that all individuals prefer to live with highly educated neighbors due to, say, positive externalities in human capital production (see Benabou [5] and Cutler and Glaeser [10], for example). Given the capitalization of these externalities into housing prices, our assumption essentially implies that highly educated individuals are able to outbid less educated individuals to live in more educated neighborhoods; it is a convenient reduced-form simplification that allows us to place the role of housing prices in the background of the analysis.

In support of this assumption, there is ample evidence that, taking housing prices into account, individuals prefer to live in neighborhoods with others of the same race.[13, 14] Again, our assumption

[13] Cornell and Hartmann [9], Farley *et al.* [14], O'Flaherty [21] and Lundberg and Startz [20] provide various theoretical arguments as to why individuals might care about the racial composition of their neighborhoods.

[14] For example, in the Multi-City Survey of Urban Inequality (MCSUI), respondents were shown a card representing a neighborhood with fifteen houses (in three parallel rows of five houses each), and then asked to illustrate the racial composition of their "ideal" neighborhoods, where they were presumed to live in the house located at the center of the middle row. Using data from the MCSUI conducted between 1993-1994 in the Atlanta, Detroit, and Los Angeles metropolitan areas, Ihlanfeldt and Scafidi [17] found that, 35-43 percent of blacks designated an all-black

does not require that individuals absolutely prefer to live near others of the same race, but simply that individuals are more willing to live with others of their own race than others in the metro area. Importantly, preferences for the neighborhoods with higher fractions of individuals of the same race need not arise through direct preferences for the race of one's neighbors, but might also come about through a number of indirect channels. In particular, individuals of the same race may cluster together in residential neighborhoods because they have correlated preferences for local public and private goods including retail, restaurants, newspapers, and churches (see Waldfogel (2004)) or because they have preferences to live near family members. What is important from the point of view of our model is that race proxies for an important dimension of the residential choice process that has a considerable impact on location decisions.

The second basic ingredient in our model is the notion of a critical neighborhood size. To capture this notion, we define a neighborhood as a collection of individuals residing at a particular point in space and assume that each resident incurs a cost that is decreasing in the total number of residents in the neighborhood. Rather than simply assuming an exogenous neighborhood size, we assume a decreasing average cost function because it more readily captures the idea that a larger population of individuals can sustain a larger number (and higher quality) local goods in line with the preferences of those in the neighborhood (see, e.g., Berry and Waldfogel [4], and references cited therein).

The final important component of the model is the specification of idiosyncratic location preferences that are unrelated to sorting on the basis of education or race. We capture heterogeneity in preferences for locations throughout the metropolitan area by assuming that individuals have employment locations distributed in space and would prefer to commute shorter distances. Such an assumption is standard in the "spatial mismatch" literature (see Kain [18], Ross [25] and Weinberg [30]). The introduction of preferences for location unrelated to neighborhood race and education brings the physical geography of a metropolitan area into the model in a natural way and renders the density as well as the size of a given race-education category important to the neighborhood

neighborhood or mostly black neighborhood (eleven blacks and four whites) as their top choice; and 81-92 percent of the blacks chose all black or mostly black neighborhoods as one of their top two choices. See also Vigdor [29] and Charles [6][7] for related evidence.

It is important to emphasize that such evidence has to be at best considered as suggestive, as the MCSUI survey questions make no mention of neighborhood amenities, housing prices, or other factors that might influence residential choices. Thus such evidence does not necessarily reveal fundamental racial preferences. King and Mieszkowski [19], Yinger [32] Galster [15] report evidence of segregating preferences based on housing prices and rents.

formation process.

Model. Consider a metropolitan area located on a straight line with length 2, represented by the interval $[-1, 1]$. The population density in the metropolitan area is given by $N > 0$, so its total population is $2N$. There are two racial groups $r \in \{b, w\}$, a proportion $\lambda_w \in (0, 1)$ being white, with the remaining proportion $\lambda_b = 1 - \lambda_w$ being black. Individuals within each racial group differ in their educational attainment: a fraction $\rho_r \in (0, 1)$ of race-r individuals are highly educated (denoted by type-h) and the remaining fraction $1 - \rho_r$ are less educated (denoted by type-l). Cross-race inequality in socioeconomic characteristics is reflected by the difference $\rho_w - \rho_b$. For all metropolitan areas in the U.S., the relevant case is $\rho_w > \rho_b$. Thus a narrowing in the racial gap in educational attainment can be represented by an increase in ρ_b while keeping ρ_w fixed.

For simplicity, we assume that whites' residential locations are fixed: at each endpoint of the line, there are two communities, one for highly educated whites (called communities WH and WH') and one for less educated whites (called communities WL and WL').

We focus our analysis on the residential location choices of blacks and the *emergence* of black neighborhoods. Accordingly, we model idiosyncratic locational preferences of blacks by assuming that their job locations are uniformly distributed along the straight line. Commuters experience a cost of $\theta > 0$ per unit distance between their work and place of residence.

There is a cost of maintaining a community, and the average per-resident cost is given by $c(n)$ where n is the number of residents in the community.[15] We assume that $c(\cdot)$ decreases in n.

We now describe blacks' preferences. Consider a black individual with education $e \in \{l, h\}$, whose job location is at point $z \in [-1, 1]$ on the straight-line. Her utility from living in a community $j \in J$, where J is the set of available communities to be determined in equilibrium, is given by:

$$u(j; z, e) = \alpha \left[p_b(j) + \gamma_1 p_w(j)\right] + \beta \left[p_e(j) + \gamma_2 p_{e'}(j)\right] - \theta D(j, z) - c\left(n(j)\right), \qquad (1)$$

where $e' \neq e$ is the other education category; $p_r(j)$ is the proportion of residents in community j of race r; $p_e(j)$ is proportion of residents in j with education attainment e; $D(j, z)$ is the commuting distance between community j and z's job location; $n(j)$ is the number of residents in community j; and $\alpha > 0, \beta > 0, \gamma_1 \in (0, 1)$, and $\gamma_2 \in (0, 1)$ are constants.

In utility function (1), the first term $\alpha[p_b(j)+\gamma_1 p_w(j)]$ captures the utility from interacting with people of different races in the same community, where $1/\gamma_1 > 1$ measures the same-race preference

[15] Technically, this rules out tiny enclaves of individuals claiming to form a neighborhood of their own.

discussed earlier. The interpretation of the second term $\beta[p_e(j) + \gamma_2 p_{e'}(j)]$ is more subtle. As we explained previously, it is meant to capture, in a reduced-form way, the idea that highly educated individuals will *on net* (taking into account both human capital externalities and housing prices) prefer to live in more expensive neighborhoods with many other highly educated residents, while less educated individuals will prefer on net to live in cheaper neighborhoods with other less educated residents.

We define an equilibrium of this simple model to be a set of neighborhoods J^* (including the existing neighborhoods WH, WH', WL, WL') and the residential choices of all blacks such that: (1) given J^*, all black individuals' residential choices are utility-maximizing; (2) no coalitions of blacks in $j \in J^*$ can be better off by forming their own neighborhood; and (3) there is a positive measure of residents in all neighborhoods $j \in J^*$.[16] It is important to remark that our equilibrium condition (2) assumes away the coordination problem among highly educated blacks in their decision to form their own neighborhood. Indeed a *coordination problem*, if it exists, is likely a short-term phenomenon, as developers and other entrepreneurs are likely to solve it. In our model, the lack of middle-class black neighborhoods in a metropolitan area is a result of a *small numbers problem* - that is, an insufficient density of highly educated blacks given the distribution of idiosyncratic preferences - instead of a coordination problem.

As in any model of residential sorting, there are multiple equilibria. Thus, we focus on a particular equilibrium in which the sizes of black neighborhoods, if formed, are maximized. Given the uniform distribution of the population on the city 'line,' this implies that black neighborhoods are formed in the center of the city.

In the equilibrium selected above the set of neighborhoods J^* depends on the parameters of the model. We are particularly interested in the way the set J^* is affected by an increase in ρ_b – the fraction of highly educated blacks in the metropolitan area. Consider an equilibrium in which a single black community, community B, emerges at point 0. Clearly community B, were it to emerge, would consist of blacks whose job locations were close to point 0. Thus given $J^* = \{$WH, WH', WL, WL', B$\}$, blacks' optimal residential choices can be characterized by a pair $\{x_h^*, x_l^*\}$ such that all highly educated (less educated, respectively) blacks will choose to live in community B if and only if their job location z satisfies $|z| \leq x_h^*$ ($|z| \leq x_l^*$, respectively). The marginal types

[16] Note that we do not need to directly impose a threshold neighborhood size in our model. The existence of the four white neighborhoods, together with the assumption that $c(n)$ is decreasing in n, endogenously ensures that small enclaves of blacks will not form their own neighborhoods.

$\{x_h^*, x_l^*\}$ can be determined from the indifference conditions (see Appendix A for details). Figure 3 depicts this type of equilibrium when ρ_b is small.[17]

[Figure 3 About Here]

Imagine that we now increase the fraction of highly educated blacks ρ_b from a low level initially. First, note that as ρ_b increases, the proportion of highly educated blacks in community B, $p_h(\mathrm{B})$, will increase even if the thresholds $\{x_h^*, x_l^*\}$ were hypothetically unchanged. As $p_h(\mathrm{B})$ increases, community B becomes more attractive vis-à-vis community WH and WH' for highly educated blacks. As a result, the marginal highly educated black who commutes to community B, x_h^*, will increase, raising the probability that a highly educated black chooses community B This shift has the effect of increasing the exposure of highly educated blacks to both highly and less educated blacks at the expense of their exposure to highly educated whites.

The results for less educated blacks are more ambiguous. On the one hand, community B becomes more educated, which makes it less desirable for less educated blacks according to their preference specified in (1); on the other hand, the increase in the total population in community B drives down the per-resident community cost c. Thus, whether or not community B becomes more desirable for less educated blacks is indeterminate. It is thus possible that exposure of highly and less educated blacks to one another may increase. (See Figure 4b for a graphical illustration).

[Figure 4 About Here]

When ρ_b is sufficiently high, however, a point may be reached where it becomes profitable for highly educated blacks in community B to form their own community at point 0, called BH, leaving behind a less educated black community BL (see Figure 4c). The exact point at which the highly educated black neighborhood BH emerges is determined by the balancing of the following two forces. First, by separating from the less educated blacks living in community BL, highly educated blacks have to incur a higher per-resident community cost c as a result of the smaller population size; second, because community BH consists only of highly educated blacks, the utility component

[17] If such an equilibrium exists with a sufficiently small ρ_b, one can show that $x_l{}^* > x_h{}^*$. The reason is simple: when ρ_b is small, community B is necessarily a predominantly less educated all-black community. Because $\gamma_2 < 1$, the utility for a less educated black from community B is always higher than that for a highly educated black at any job location. Thus less educated blacks are more willing to commute to community B. This is not important for the analysis but explains the ranking of x_l^* and x_h^* in Figure 3.

$p_h\,(\mathrm{BH}) = 1 > p_h\,(\mathrm{B}) + \gamma_2 p_l\,(\mathrm{B})$ because $\gamma_2 < 1$. We capture the above discussion in the following proposition:

Proposition 1 ***(Comparative Statics)***

1. *An increase in ρ_b from small to moderate values will lead to a higher exposure of highly educated blacks to both highly and less educated blacks, and decrease their exposure to highly educated whites.*

2. *When ρ_b is sufficiently high, all-black highly-educated neighborhoods may emerge; and the exposure of highly educated blacks to whites, as well as to less educated blacks, will decrease.*

To summarize, the key insight of our simple model is that the nature of available neighborhoods for highly educated blacks is likely to change as the average education level of blacks in a metropolitan area increases. The change in the available neighborhoods for highly educated blacks occurs both when ρ_b is moderate and when it is high: when ρ_b is moderate, community B will contain more highly educated blacks even though it is not yet stratified on the basis of education; when the proportion of highly educated blacks ρ_b is sufficiently high, a highly educated black community BH emerges and results in a more dramatic change in neighborhood structure. It is worth pointing out that the emergence of community BH is likely to induce an accelerated migration of highly educated blacks from community WH and WH' to community BH, resulting in greater racial segregation in residential locations.

Given the empirical facts presented in Section 2, which demonstrate the relatively small number of highly educated blacks and the short supply of middle-class black neighborhoods in the vast majority of US MSAs, we would generally expect the first comparative static described above (an increase in ρ_b from small to moderate level) to apply to the vast majority of US MSAs. This comparative static thus forms the basis of our main hypothesis: that the segregation of highly educated blacks (and blacks more generally) is an increasing function of the average educational attainment of blacks in a metropolitan area. We also note that the emergence of middle-class black neighborhoods also depends positively on the population density N and the overall proportion of blacks in the metropolitan area λ_b.[18] As we show below, an increase in average black educational attainment in large metropolitan areas leads to especially strong increases in the exposure of highly educated blacks to one another, a result in line with the second comparative static prediction above.

[18] It also depends indirectly on the commuting cost θ and the community cost function $c\,(n)$ via their effects on ${x_h}^*$.

4 Empirical Analysis

We now present a series of empirical analyses designed to test our main hypothesis: that the segregation of highly educated blacks (and blacks more generally) is an increasing function of the average educational attainment of blacks in a metropolitan area. We present results using a variety of organizations of the 1990 and 2000 US Censuses of Population and a number of distinct empirical specifications, following the sequence described in the Introduction.

4.1 Cross-Sectional Analysis: Segregation Patterns in U.S. Metropolitan Areas

We begin our cross-sectional analysis by considering the general pattern of segregation in the U.S. as a whole. Panel A of Table 5 reports the average cross-exposure of individuals by race-education categories *relative to the fraction in an individual's MSA*.[19] The first row of Panel A states that, relative to an average individual in the same metropolitan area, blacks without a college degree are exposed to 19.6 percentage points more blacks without a college degree and 2.1 percentage points more college educated blacks, etc.

[Table 5 About Here]

Panels B and C of Table 5 report segregation patterns in a manner analogous to Panel A, but separately for metropolitan areas with above and below the median fraction (1.23 percent) of college-educated blacks. Comparison of Panels B and C provide some initial evidence as to how segregation patterns vary with the sociodemographic composition of the metropolitan area. It shows that the relative exposure of blacks in each education category to both highly and less educated blacks is significantly greater in metropolitan areas with above-median fractions of college educated blacks. For both highly and less educated blacks, the average tract-level exposure to blacks relative to the fraction of blacks in MSAs above the median is more than double that for MSAs below the median.

[19] The exposure rates are constructed as follows (see Bayer, McMillan and Rueben [2]). Let r_j^i be a set of indicator variables that take the value 1 if individual i is of race j and 0 otherwise, and let R_k^i be the fraction of individuals of race k in individual i's neighborhood (the Census tract, for example). The average exposure of individuals of race j to households of race k is $E_{jk} = \sum_i r_j^i R_k^i / \sum_i r_j^i$.

4.2 Cross-Sectional Analysis: Regression Results

To control more formally for the sociodemographic structure of the metropolitan area, Table 6 reports the results of a series of regressions of various tract composition measures on individual and MSA characteristics. Econometrically, the regressions reported in Table 6 are of the following form:

$$Y_{i,m} = \alpha_m + \boldsymbol{\beta}\mathbf{X}_i + \boldsymbol{\gamma}\mathbf{X}_i \cdot \mathbf{X}_m + \varepsilon_{i,m}, \tag{2}$$

where $Y_{i,m}$ denotes the Census-tract level exposure rate for an individual i living in MSA m; $\mathbf{X}_i$ is $i's$ individual characteristics; $\mathbf{X}_m$ is the characteristics of MSA m; and the term α_m represents the MSA fixed effect.[20] The dependent variable $Y_{i,m}$ varies by the heading listed in each column. For example, the dependent variable for the regression in Column 1 is the fraction of college educated blacks in i's Census tract, and the dependent variable for Column 2 is the fraction of less educated blacks in i's Census tract, etc. The inclusion of the MSA fixed effects ensures that all of the other parameters characterize the effect on average tract composition *relative to the MSA average* for each set of individuals, i.e., that the regressions account for the mechanical increase in exposure that would follow from a change in the metropolitan area's composition.

The key coefficients in the regressions are the coefficients $\boldsymbol{\gamma}$ on the interactions of individual and MSA characteristics, which characterize how the average race-education composition of an individual's tract relative to the MSA as a whole, for individuals of different characteristics, varies with MSA characteristics. For example, Column 1 shows that a one percentage-point increase in the proportion of highly educated blacks (at the expense of the omitted race category, Hispanics) in an MSA will increase the exposure *relative to the MSA average* of highly educated blacks to other highly educated blacks by 0.968 percentage points, and it also increases the relative exposure of less educated blacks to other highly educated blacks by 1.022 percentage points. It is also useful to look across a particular row. For example, the first row tells us that a one percentage-point increase in the proportion of highly educated blacks in an MSA will increase the relative exposure of highly educated blacks to other highly educated blacks by 0.968 percentage points (Column

[20] Because we use the Census Tract Summary Files in our empirical analysis, in practice, $\mathbf{X}_i$ are simply $i's$ race-education categories. In the basic regressions reported in Table 6, $\mathbf{X}_m$ include the MSA's population compositions, namely, the proportion of the MSA population that is white/highly educated, black/highly educated, white/less educated and black/less educated. It is also worth mentioning that the Census Summary Tract Files provide the *number* of individuals in each race/education category by Census tract, thus our regression in practice is equivalent to running weighted OLS where the weight is given by the number of individuals in each race/education cell.

1), to less educated blacks by 3.04 percentage points (Column 2), and to blacks overall by 4.008 percentage points (Column 3); but it decreases the relative exposure of highly educated blacks to highly educated overall by 1.261 percentage points (Column 4).

[Table 6 About Here]

Table 7 summarizes the coefficient estimates in Table 6 (and other similar regressions using alternative definition of "highly educated" and/or controlling for region and population of the metropolitan area) by reporting results from a statistical experiment that examines how average tract compositions for highly and less educated blacks change as the proportion of college-educated blacks in the MSA is increased by one percentage point, holding constant the fraction of blacks. This corresponds to examining the impact of an increase in the average education level of the black population holding the characteristics of the rest of the population constant.[21]

[Table 7 About Here]

Panel 1 of Table 7 shows that when the fraction of college-educated blacks in a metropolitan area increases by 1 percent at the expense of less educated blacks, the relative exposure of college educated blacks to other college educated blacks increases by 1 percentage point and is statistically significant; the relative exposure of less educated blacks to college educated blacks also increases by 1.1 percentage points. Overall, the relative exposures of blacks with and without a college degree to other blacks increase by 4 and 6.1 percentage points respectively, highlighting the increased segregation of blacks of all education levels following an increase in the average education of the black population. This empirical finding is consistent with our model's prediction when ρ_b lies in an intermediate range (Figure 4b), which we think is the plausible scenario for most U.S. metropolitan areas.[22]

[21] We also conducted another experiment that examines how average tract compositions for highly and less educated blacks change as the fraction of college-educated blacks in the MSA is increased by one percentage point at the expense of college-educated whites. This corresponds to examining the impact of an increase in the fraction of the educated population that is black. As the coefficients reported in Table 6 imply, the results from this experiment are qualitatively and quantitatively very similar to those reported in Table 7.

[22] The results of Table 6 are not driven by the specific form of the dependent variable that we employ. We conducted a series of regressions analogous to those reported in Table 6 except that the dependent variable is defined as the fraction of individuals in a given category in an individual's tract divided by the fraction in the metropolitan area as a whole. In this way, an increase in tract-level exposure to individuals in a given category from 6 to 12 percent

Panel 2 of Table 7 reports results analogous to those reported in Panel 1 with the exception that the underlying measure of "highly educated" is changed to include those individuals having at least attended college. With this broader definition, the fraction of individuals 25 years and older in U.S. metropolitan areas who are highly educated is 54 percent, the fraction who are both highly educated and black is 5 percent, and the fraction of blacks who are highly educated is 45 percent. Our primary objective in examining an alternative is to consider a definition of "highly educated" that includes a larger fraction of individuals and especially black individuals. A comparison of Panel 2 and Panel 1 reveals a qualitatively similar pattern. With the expanded definition of highly educated, the relative increase in exposure of both highly and less educated blacks to other blacks is more evenly split between highly and less educated blacks.

In sum, both definitions of "highly educated" reveal a pattern of increased relative exposure of both highly and less educated blacks to blacks in each education category when the fraction of highly educated blacks in the metropolitan area increases. This pattern is consistent with the predictions of our model corresponding to an increase in ρ_b (the fraction of blacks who are highly educated) from low to moderate levels. With an increase in the average education level of the black population, highly educated blacks move on net into more segregated neighborhoods, increasing the average education level in some of the most segregated neighborhoods. In terms of the scatterplots, this pattern is consistent with the formation of segregated, black neighborhoods with mixed education levels along with a corresponding shift of highly-educated blacks away from highly-educated predominantly white neighborhoods to these newly formed neighborhoods.

4.3 Cross-Sectional Analysis: Robustness and Heterogeneity

One potential concern with the results presented in Panels 1 and 2 of Table 7 is that they may be driven by unobserved factors related to historic patterns of black settlement, migration, and segregation in the U.S. To address such concerns, Panel 3 of Table 7 reports the results of a set of regressions analogous to those reported in Panels 1 and 2 with the addition of a complete set of interactions between each individual's race-education category and a measure of metropolitan size

following an increase in the proportion of these individuals in the metropolitan area from 3 to 6 percent would not result in an increase in the dependent variable in this case, while it would have resulted in a 3 percentage point increase in the dependent variable used in the regressions reported in Table 6. The resulting parameter estimates led to a very similar set of conclusions, ensuring that our initial results are not driven by the functional form of the dependent variable. Throughout the remainder of the paper, we present the results of regressions analogous to those reported in Table 6.

and four dummies for Census region (Northeast, Midwest, South, West).[23]

A comparison of the results in Panel 3 and Panel 2 reveals a qualitatively similar pattern both in magnitude and in statistical significance. In particular, with the additional controls, the increase in the relative exposure of both highly and less educated blacks to other blacks declines by 15-20 percent in magnitude, but remain highly significant. Changes in relative exposure to highly educated neighbors also decline and remain insignificant in each case. Taken together, these results give us confidence that the main conclusions of the paper are not driven by obvious omitted variable biases.

While we added metropolitan size and interactions in the results reported in Panel 3 of Table 7, we still assumed that the effects of the fraction of highly educated blacks on segregation do not depend on metropolitan area size. The critical mass story implicit in our model implies that not only the fraction but also the number of highly educated blacks in the metropolitan area may be important for the formation of more-educated and segregated black neighborhoods. Given the same fraction of highly educated blacks, highly educated black neighborhoods might more easily form in large (population-wise) rather than small metropolitan areas.

Panels 4-6 estimate separate regressions, including the additional 16 control variables added in Panel 3, for small (0-200k), medium (200-600k), and large (600k+) metropolitan areas. A clear pattern emerges in the table: following an increase in the average education level of the black population, the increased relative exposure of both highly and less educated blacks to other blacks is much greater in large versus small metropolitan areas. For highly educated blacks, the magnitude of the effect rises from 0.002 in small, to 0.025 in medium-sized, and 0.040 in large metro areas. The results tend to have higher statistical significance in larger metropolitan areas. These results are consistent with the notion of critical neighborhood size – as a percentage point increase in the fraction of highly educated blacks in a large versus small metropolitan area obviously represents a larger increase in the number of highly educated blacks in the MSA.

In line with the predictions of our theoretical model a qualitatively different pattern begins to emerge in large metropolitan areas. In particular, the increased exposure of highly educated blacks to other blacks is dominated by an increased exposure to other highly educated blacks. Thus, for this subsample, an increase in the average education of the black population might be associated with the formation of predominantly highly educated, segregated black neighborhoods. Not surprisingly, the relative exposure of less educated blacks to educated neighbors declines most markedly in this

[23] A total of 16 interaction terms are added to the regression.

sub-sample, marking the only specification where this effect even borders on statistical significance. Thus, the results for large metropolitan areas correspond well to the predictions of the model related to an increase in the fraction of highly educated blacks from a moderate to a large number.

5 Robustness to Alternative Explanations

The cross-sectional finding of a positive relationship between black educational attainment and residential segregation at the metropolitan level is in many ways a surprising result. As we mentioned in the Introduction, most first-order explanations for such a relationship examined in the literature would imply a negative relationship. In this paper, motivated by the short supply of middle class black neighborhoods in most MSAs, we argue that the impact of increased black educational attainment on the formation of such neighborhoods provides a potential explanation for this result. Here, we consider other potential explanations for a positive correlation between black segregation and educational attainment.

Note that for the empirical analysis reported in Section 4, we used Census Tract summary files (SF3) data. The advantage of the tract level data is that we are able to observe an individual's neighborhood at the geographically disaggregated level of a Census tract. The disadvantage, however, is the data is summarized at the tract level and we are unable to observe the whole vector of individual characteristics that Census actually collects. In this section, we will instead rely on another organization of the Census, the Public Use Microsample (PUMS data). Relative to the data from the Census summary files, the PUMS data specify the geographic location of an individual's residence to a much larger region that is about 20-30 times larger than a tract – a Census PUMA; but PUMS data have the advantage that they are at the individual level.[24]

5.1 Reverse Causality: Reconciling with Cutler and Glaeser (1997)

As discussed above, the primary reason that the finding of a positive correlation between black segregation and educational attainment measured at the metropolitan level is surprising is that Cutler and Glaeser [10] report a negative correlation when essentially running the reverse regression

[24] Using PUMAs rather than tracts as the geographic unit to define the left-hand side variables in the regressions specified in regression equation (2) results in coefficient estimates that are qualitatively similar to those reported in Table 7. As would be expected, the magnitudes are slightly smaller in this case due to the use of a larger geographic unit. Details are available from the authors upon request.

for individuals *aged 20-30*. In this subsection, we present a detailed analysis that reconciles the findings of these two studies.

CG ran a series of regressions that relate individual education, fertility, and labor market outcomes to individual and metropolitan characteristics. Their primary focus is on isolating the effect of living in a more segregated metropolitan area on these outcomes for blacks relative to whites. This effect is summarized as the coefficient on the interaction of a measure of metropolitan segregation and a dummy variable that indicates whether the individual is black. Their regressions that relate most directly to our primary findings involve college education as the dependent variable. Here, the coefficient estimates on the interaction term describe the correlation between metropolitan segregation and the relative educational attainment of blacks, holding the attainment of whites constant. They report results both from OLS regressions and IV regressions, where they instrument for segregation with a number of alternative variables designed to isolate the causal effect of residential segregation on outcomes.

[Table 8 About Here]

To reconcile our results with CG's, we begin with their OLS results. Table 8 reports the coefficient on the interaction between their metropolitan dissimilarity measure and whether an individual is black, first replicating their results for age groups 20-24 and 25-30 and then reporting analogous coefficients for individuals between the ages of 31-40, 41-50, 51-60, and 61-70, respectively. The coefficients for older individuals reveal a markedly different pattern from those for younger individuals. Focusing specifically on college education and earnings, which most closely correspond to our definition of highly educated (or high SES), the coefficients reverse sign from negative to positive starting for individuals slightly older than those studied by CG. From a purely mechanical perspective, this age profile reconciles the results presented in our paper with those in CG, thereby implying that the overall positive correlation that we report in Section 4 is driven primarily by older individuals.

The primary results presented in CG are not the OLS results replicated here, but a series of IV estimates that instrument for metropolitan segregation with three alternative instruments designed to isolate the causal impact of segregation on outcomes. CG motivate this IV approach by suggesting that their negative coefficient estimates from OLS regressions might be attributable to within-metropolitan sorting, namely that segregation might be higher in metropolitan areas where blacks had poor socioeconomic characteristics relative to whites as a result of sorting on the basis of socioeconomic characteristics. Importantly, however, when they instrument for segregation, the point estimate on the interaction between black and segregation in the college degree and log earnings regressions becomes *more negative* in every case (for both age groups and with each alternative instrument, a total of 12 regressions). This suggests that the reverse channel of causality (within metropolitan-area sorting) is actually working against their result, causing the correlation between black socioeconomic status and metropolitan segregation to move in a positive rather than a negative direction.[25]

In this way, the full set of results reported in CG (OLS and IV) along with our results in this paper can be fully reconciled as the operation of the mechanisms that form the focus of the two papers, with each mechanism working to obscure the other in the data. Because many individuals migrate across metropolitan areas in early adulthood and metropolitan segregation evolves (slowly) over time, one would generally expect the negative relationship between segregation in an individual's current MSA and educational outcomes related to the CG mechanism to be strongest for the youngest cohort of adults. (This is why they study young adults in the first place.) Conversely, the positive correlation between average black educational attainment and metropolitan segregation related to the within-metro sorting mechanism that we identify should be strongest among older cohorts. These individuals collectively have had the greatest amount of time to influence the metropolitan neighborhood structure and, consequently, also the segregation levels in the MSA in which they reside.

It is important to point out that potential alternative explanations for the age profile revealed in Table 8 do not appear plausible. First, a similar age profile to that reported in Table 8 for the 1990 Census (not shown in the tables) emerges in the 1980 and 2000 Censuses for both college degree and earnings regression. This suggests that alternative interpretations of the profile as a cohort- rather than age- profile are unlikely to hold.[26] For 1980, 1990, and 2000, we have also conducted

[25] Another superficial difference between Cutler and Glaeser [10] and our work, which is not important in explaining the differences in the cross-sectional correlations reported in each paper, relates to the measure of segregation. In particular, we use race/education specific exposure rates as our measure of segregation while they use MSA-level dissimilarity indices. The dissimilarity index, proposed by Duncan and Duncan [12], is an aggregate-level measure capturing the fraction of blacks that would have to switch areas to achieve an even racial distribution citywide (see Cutler, Glaeser and Vigdor [11] for more discussion). See footnote 19 for the construction of exposure rates.

[26] Also of note, Collins and Margo (2000) report the key coefficient from a series of CG style regressions for ln(earnings) of individuals aged 20 to 30 as far back as the 1940 and 1950 Censuses. They report a statistically insignificant effect of roughly the same magnitude as that reported by CG for 1990.

analyses of across-metro sorting analogous to the next subsection of our paper and the across-metro sorting analysis in CG. These analyses suggest little change in the nature of across-metro sorting as it relates to the correlation of metropolitan segregation and black educational attainment over the past three decades, making alternative explanations for the age profile related to across-metro sorting unlikely as well.

5.2 Across-Metro Sorting: Do Middle-Class Black Neighborhoods Attract Highly Educated Blacks?

Another potential explanation for the existence of a correlation between metropolitan population characteristics and segregation relates to across-metropolitan area sorting, instead of the within-metro sorting we highlighted in our endogenous neighborhood formation story. Fortunately, the Census PUMS microdata contain information on the metropolitan area in which each individual lived *five years prior to the Census*. Using these data, we consider two aspects of across-metropolitan sorting. We begin by examining whether highly educated blacks are drawn disproportionately to metropolitan areas that have a larger number of middle-class black neighborhoods. Such migration pattern would be a prediction from an extension of our model when we allow for migration across MSAs. However, such migration patterns might generate another form of a reverse causation problem for our primary cross-sectional results. That is, while our main hypothesis relates to the impact of population characteristics on the equilibrium structure of neighborhoods in a city (e.g., segregation levels), sorting of this kind implies that neighborhood structure affects population characteristics through its impact on migration.

[Table 9 About Here]

To explore this issue further, Table 9 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. These younger adults are much more likely than others to move to a new metropolitan area during a given five-year period. 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 that moved to a new MSA between 1995 and 2000 and includes fixed effects for the MSA in which the individual resided in 5 years ago. In essence, this specification compares the characteristics of newly chosen metropolitan areas for two individuals that resided in the same metropolitan area five years ago. The results clearly demonstrate that college-educated blacks are more likely to choose metropolitan areas 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 US metropolitan areas is only 0.3). Such 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, a finding consistent with our specification of blacks' preferences in our model and the fact that most U.S. metropolitan areas have very limited number of middle-class black neighborhoods.

To explore whether this kind of across-metropolitan sorting is likely to lead to the aforementioned reverse causation problem, columns 2-3 in Table 9 report the results of corresponding specifications for individuals that do and do not migrate across MSAs during this five-year period, respectively, dropping the fixed effects for the lagged MSA.[27] The resulting coefficients reveal a remarkably similar pattern to those reported in column 1. That an almost identical pattern obtains for stayers 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, in fact, 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. Thus, this pattern of migration does not systematically contribute to cross-sectional differences in metropolitan area composition. This allows us to rule out this type of sorting as an explanation for our baseline positive cross-sectional relationship between segregation and black educational attainment.

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 again demonstrate the same patterns.

5.3 Across-Metropolitan Area Sorting: Selection Bias

A second aspect of across-metro sorting that might pose problems for our analysis is related to unobservable characteristics. In particular, if those highly educated blacks that live in metropolitan

[27] We could not include additional fixed effects for the lagged MSA for stayers since they did not move.

areas with a more educated black population have stronger unobserved tastes for segregation, this would lead to a positive cross-sectional correlation between segregation and educational attainment unrelated to within-metro sorting.

To explore the possibility of selection bias, we again make use of the information on the metropolitan area in which each individual resided in 1995. In particular, we decompose the sociodemographic composition of each individual's current metropolitan area $\mathbf{X}_{m(2000)}$ into two components: the first component, the "lagged measure," $\mathbf{X}_{m(1995)}$, gives the composition of the metropolitan area in which that person lived five years ago; the second component, called the "differenced measure," $\boldsymbol{\Delta}\mathbf{X}_m \equiv \mathbf{X}_{m(2000)} - \mathbf{X}_{m(1995)}$, is the difference between the composition of the current metropolitan area and the lagged measure.[28] Note that $\mathbf{X}_{m(2000)} \equiv \mathbf{X}_{m(1995)} + \boldsymbol{\Delta}\mathbf{X}_m$. For the 90 percent of the population who did not move, the differenced measure is zero; while for movers, this difference reflects the change in metropolitan area sociodemographics associated with their move. We then include distinct interaction terms with both measures in a specification analogous to regression (2) that we used to generate our baseline cross-sectional results in Tables 7:

$$Y_{i,m(2000)} = \alpha_m + \boldsymbol{\beta}\mathbf{X}_i + \boldsymbol{\gamma}_1\mathbf{X}_i \cdot \mathbf{X}_{m(1995)} + \boldsymbol{\gamma}_2\mathbf{X}_i \cdot \boldsymbol{\Delta}\mathbf{X}_m + u_{im}. \tag{3}$$

The estimated coefficients on the lagged measure $\mathbf{X}_{m(1995)}$ versus differenced measures $\boldsymbol{\Delta}\mathbf{X}_m$ indicate the direction of the selection bias: $\boldsymbol{\gamma}_1 > \boldsymbol{\gamma}_2$ would indicate a negative selection bias, while $\boldsymbol{\gamma}_1 < \boldsymbol{\gamma}_2$ indicates a positive selection bias. To see this concretely, suppose for illustration that we run a regression in which $Y_{i,m(2000)}$ measures exposure to blacks and we are interested in the coefficient on interactions with the fraction of highly educated blacks in the metropolitan area. If sorting is purely random with respect to the metropolitan population characteristics, we would expect that the coefficient estimates of $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ be the same (and be equal to the estimate of $\boldsymbol{\gamma}$ in specification (2)) because current metropolitan characteristics affect all individuals equally. If $\boldsymbol{\gamma}_1$ exceeds $\boldsymbol{\gamma}_2$, this implies that, relative to the existing residents of an MSA, individuals migrating from a metropolitan area with a smaller fraction of highly educated black households choose neighborhoods with a smaller fraction of blacks. This implies that, on average, the unobserved taste for segregation (u_{im}) among in-migrants from metropolitan areas with a lower fraction of highly educated blacks must be lower than that of existing residents, thus implying a negative selection bias.[29]

[28] The subscripts $m\,(2000)$ and $m\,(1995)$ denote an individual's metropolitan area in 2000 and five years prior, respectively.

[29] More formally, one needs to make an assumption about the initial distribution of the unobserved taste for

[Table 10 About Here]

Table 10 shows results for analysis comparable to that of Table 7. Because the analysis presented here is based on the 2000 Census PUMS, the definition of neighborhood is the Census PUMA. In each of panels, therefore, the first specification repeats our baseline analysis making this change. As in Table 7, Table 10 summarizes the effect of a one percent increase in the fraction of college-educated blacks in the MSA, holding the fraction of blacks constant. The two panels summarize results for blacks with some college or more and a high school degree or less, respectively, and should be compared with the results presented in column 2 of Table 7. As the initial specification reported in each panel reveals, the results remain statistically significant and only slightly smaller in magnitude than those presented in column 2 of Table 7.

Examining the results when metropolitan measures are decomposed into lagged and differenced measures reveals that $\gamma_1 > \gamma_2$ in all cases. This implies that new migrants to MSAs with high average black educational attainment coming from MSA with lower levels of black educational attainment tend to locate in less segregated neighborhoods upon arriving than otherwise identical longer-term residents of these MSAs. Taken together, these results suggest that across-metro sorting on the basis of unobservables introduces a slight negative selection bias in our main results.

5.4 Time Series Evidence

Having explored the robustness of our cross-sectional results in great detail, we now conclude our empirical analysis by presenting time-series evidence on the relationship between metropolitan structure and segregation. In particular, we regress a measure of the change in metropolitan level dissimilarity index between 1990 and 2000 on measures of the changes in the sociodemographic composition of the metropolitan area. The results of this regression are summarized in Table 11, which show a strong positive relationship between the change in the fraction of blacks with a college degree in the population and segregation.

[Table 11 About Here]

---

segregation u_{im} across metropolitan areas. The comparison of γ_1 and γ_2 serves as a test for selection bias if one assumes that u_{im} is fixed and distributed independently across metropolitan areas at birth. Given some form of migration costs, a positive selection bias would imply that the average u_{im} among in-migrants from metropolitan areas with a smaller fraction of educated blacks would be higher than that of existing residents even allowing for the possibility of migration in previous periods.

Again, it is straightforward to use the coefficients of Table 11 to compute the effect of a one percent increase in the fraction of highly-educated blacks holding the fraction of blacks constant, by subtracting the second coefficient from the first. Regression results are presented with and without weighting by the population of the metropolitan area. The results clearly show that metropolitan areas that experienced an increase in average black educational attainment between 1990 and 2000 saw an marked increase in segregation over this period.

6 Implications and Conclusion

This paper has explored the relationship between metropolitan level sociodemographic composition, particularly racial inequality in education, and residential segregation. We have presented a theoretical argument and empirical evidence indicating that the conventional wisdom, which suggests that residential segregation will fall when racial differences in education and other sociodemographics decline, may not hold.

Our analysis began by showing that middle-class black neighborhoods are in short supply given the current black sociodemographics in many U.S. metropolitan areas, forcing high SES blacks either to live in white neighborhoods with high levels of neighborhood amenities or in more black neighborhoods with lower amenity levels. We presented a model showing that, under certain conditions, increases in black sociodemographics in metropolitan areas will lead to the emergence of new middle-class black neighborhoods, relieving the prior neighborhood supply constraint and leading to increases in residential segregation. We then presented across-MSA evidence from the 2000 Census indicating that this mechanism does in fact operate: as the proportion of highly educated blacks in an MSA increases, so the segregation of educated blacks and blacks more generally goes up.

This change is driven primarily by a large relative increase in exposure to other highly educated blacks and is more than completely offset by a decrease in exposure to highly educated whites. At the same time, highly educated blacks are also increasingly exposed to less educated blacks and vice-versa. This effect is consistent with the predictions of our theoretical model when moving from a low to a moderate proportion of highly educated blacks. We have also shown, as far as possible, that our results are robust to concerns related to omitted variable and selection biases.

Our results have a number of important implications. First, in contrast to the conventional wisdom, they imply that racial segregation is unlikely to disappear with convergence in racial differences in socioeconomic characteristics.[30] The mechanism uncovered in our analysis indicates that an overall increase in the educational attainment of blacks may lead to a decrease in residential segregation *only if* highly educated blacks are dispersed in many, instead of concentrated in few, MSAs. This echoes Glaeser and Vigdor's [16] finding that segregation is lowest among the rapidly growing cities in the West, where there is as yet no high concentration of highly educated blacks. Second, our results also have implications concerning the impact of racial sorting in the housing market on the long-run convergence of educational attainment across race. In particular, the results indicate that *given the current sociodemographic structure of U.S. metropolitan areas*, increases in the average education level of blacks may result in a slight decrease in the relative exposure of both highly and less educated blacks to educated neighbors. A third implication relates to the literature following Wilson (1987), which demonstrates that reductions in institutional discrimination in the housing market in the middle of the 20th century led to large-scale reductions in the exposure of less educated to more educated blacks as more educated blacks left the inner city neighborhoods to which they were formerly restricted. The evidence we present here suggests that this trend may not have been severe in cities in which the black population was more educated initially and may partially reverse itself as the black population becomes relatively more educated over time.

References

[1] Arrow, Kenneth J. (1973), "The Theory of Discrimination," in O. Ashenfelter and A. Rees, Eds., *Discrimination in Labor Markets*. Princeton: Princeton University Press.

[2] Bayer, Patrick, Robert McMillan and Kim Rueben (2004), "What Drives Racial Segregation? New Evidence Using Census Microdata." Forthcoming, *Journal of Urban Economics*.

[3] Bayer, Patrick, Robert McMillan, and Kim Rueben (2004), "Residential Segregation in General Equilibrium," mimeo, Yale University.

[4] Berry, Steven and Joel Waldfogel (2003), "Product Quality and Market Size." NBER Working Paper 9675.

[5] Benabou, Roland (1992), "Working of a City: Location, Education and Production," *Quarterly Journal of Economics*, Vol. 108: 619-652.

[30] The conventional view has been embraced by many scholars in this literature. See, for example, Durlauf [13], Wilson [31] and Mayer [23].

[6] Charles, Zubrinsky C. (2000), "Neighborhood Racial-Composition Preferences: Evidence from a Multi-ethnic Metropolis," *Social Problems*, Vol. 47, 379-407.

[7] Charles, Zubrinsky C. (2001), "Processes of Racial Residential Segregation," in *Urban Inequality: Evidence from Four Cities*, edited by Alice O'Connor, Chris Tilly and Lawrence D. Bobo. New York: Russell Sage Foundation.

[8] Coate, Stephen and Glenn C. Loury (1993), "Will Affirmative Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, 83, 1220-40.

[9] Cornell, S. and Hartmann D. (1997), *Ethnicity and Race: Making Identities in a Changing World*. Thousand Oaks: Pine Forge Press.

[10] Cutler, David and Edward Glaeser (1997), "Are Ghettos Good or Bad?" *Quarterly Journal of Economics*, 112, 826-72.

[11] Cutler, David, Edward Glaeser and Jacob Vigdor (1999), "The Rise and Decline of the American Ghetto," *Journal of Political Economy*, 107(3): 455-506.

[12] Duncan, O. and B. Duncan (1955), "A Methodological Analysis of Segregation Indices." *American Sociological Review*, 20: 210-217.

[13] Durlauf, Steven N. (1996), "A Theory of Persistent Income Inequality," *Journal of Economic Growth*, Vol. 1, 75-93.

[14] Farley, R., C. Steeh, M. Krysan, T. Jackson and K. Reeves (1994), "Stereotypes and Segregation: Neighborhoods in the Detroit Area," *American Journal of Sociology*, 100: 750-780.

[15] Galster, G. (1982), "Black and White Preferences for Neighborhood Racial Composition." *AREUEA Journal*, 10, 39-66.

[16] Glaeser, Edward L. and Jacob L. Vigdor (2001), "Racial Segregation in the 2000 Census: Promising News," Brookings Center on Urban & Metropolitan Policy.

[17] Ihlanfeldt, Keith and Ben Scafidi (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.

[18] Kain, John (1968), "Housing Segregation, Negro Employment, and Metropolitan Decentralization," *Quarterly Journal of Economics*, 82, 175-197.

[19] King, A. Thomas and Peter Mieszkowski (1973), "Racial Discrimination, Segregation, and the Price of Housing," *Journal of Political Economy*, 81, 590-606.

[20] Lundberg, Shelly and Richard Startz (1998), "Race, Information and Segregation," mimeo, University of Washington at Seattle.

[21] O'Flaherty, Bredan (1999), "Troubled Transactions and their Consequences: Race in the United States," mimeo, Columbia University.

[22] Massey, Douglas S. and Nancy A. Denton (1993), *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, Massachusetts: Harvard University Press.

[23] Mayer, Susan E. (2004), "Income Inequality, Economic Segregation and Children's Educational Attainment," Working Paper, Harris Graduate School of Public Policy, University of Chicago.

[24] Miller, V. and John M. Quigley (1990), "Segregation by Racial and demographic Group: Evidence from the San Francisco Bay Area," *Urban Studies*, 27: 3-21.

[25] Ross, Stephen L. (1998), "Racial Differences in Residential and Job Mobility: Evidence concerning the Spatial Mismatch Hypothesis," *Journal of Urban Economics*, 43(1): 112-35.

[26] Sethi, Rajiv and Rohini Somanathan (2001), "Racial Income Disparities and the Measurement of Segregation," mimeo.

[27] Sethi, Rajiv and Rohini Somanathan (2004), "Inequality and Segregation," forthcoming, *Journal of Political Economy.*

[28] Schelling, Thomas C. (1971), "Dynamic Models of Segregation," Journal of Mathematical Sociology, 1: 143-186.

[29] Vigdor, Jacob L. (2003), "Residential Segregation and Preference Misalignment," *Journal of Urban Economics,* Vol. 54 No. 3, 587-609.

[30] Weinberg, Bruce (2000), "Black Residential Centralization and the Spatial Mismatch Hypothesis," *Journal of Urban Economics*, 48(1): 110-34.

[31] Wilson, William Julius (1987), *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* Chicago: University of Chicago Press.

[32] Yinger, John (1978), "The Black-White Price Differential in Housing: Some Further Evidence," *Land Economics*, 54, 187-206.

A Model Appendix

In this appendix, we explain precisely how the black marginal types $\{x_h^*, x_l^*\}$ are determined. We first restrict attention to equilibria in which, if a highly educated (less educated, respectively) black is to choose not to live in community B, she will choose community WH or WH' (community WL or WL' respectively) depending on proximity. Given a pair of thresholds $\{x_l, x_h\}$, the total measure of less- and highly educated blacks in community B are, respectively, $2N\lambda_b(1-\rho_b)x_l$ and $2N\lambda_b\rho_b x_h$. Thus the total population in community B is $2N\lambda_b[\rho_b x_h + (1-\rho_b)x_l]$. Moreover, the relevant proportions for community B are

$$p_b(\mathrm{B}) = 1, p_w(\mathrm{B}) = 0, p_h(\mathrm{B}) = \frac{\rho_b x_h}{\rho_b x_h + (1-\rho_b)x_l}, p_l(\mathrm{B}) = \frac{(1-\rho_b)x_l}{\rho_b x_h + (1-\rho_b)x_l}.$$

The utilities for a highly and less educated black individuals with job location $z \in [0,1]$ from living in community B are then, respectively:

$$\begin{aligned} V_{\mathrm{B}}^h(z; x_h, x_l) &= \alpha + \beta\left[\frac{\rho_b x_h}{\rho_b x_h + (1-\rho_b)x_l} + \gamma_2 \frac{(1-\rho_b)x_l}{\rho_b x_h + (1-\rho_b)x_l}\right] \\ &\quad -\theta z - c\left(2N\lambda_b\left[\rho_b x_h + (1-\rho_b)x_l\right]\right); \\ V_{\mathrm{B}}^l(z; x_h, x_l) &= \alpha + \beta\left(\frac{(1-\rho_b)x_l}{\rho_b x_h + (1-\rho_b)x_l} + \gamma_2 \frac{\rho_b x_h}{\rho_b x_h + (1-\rho_b)x_l}\right) \\ &\quad -\theta z - c\left(2N\lambda_b\left[\rho_b x_h + (1-\rho_b)x_l\right]\right). \end{aligned}$$

We can also calculate the utilities from living in communities WH for a highly educated black with job location $z \in [0,1]$. Given the postulated threshold x_h, the measure of highly educated blacks in community WH is $N\lambda_b\rho_b(1-x_h)$. Taking account of the measure of highly educated whites in community WH, we have the following proportions:

$$p_b(\mathrm{WH}) = \frac{\lambda_b\rho_b(1-x_h)}{\lambda_b\rho_b(1-x_h) + \lambda_w\rho_w}, p_w(\mathrm{WH}) = \frac{\lambda_w\rho_w}{\lambda_b\rho_b(1-x_h) + \lambda_w\rho_w}, p_h(\mathrm{WH}) = 1, p_l(\mathrm{WH}) = 0.$$

Thus the utility for a highly educated black from living in community WH is:

$$\begin{aligned} V_{\mathrm{WH}}^h(z; x_h) &= \alpha\left[\frac{\lambda_b\rho_b(1-x_h)}{\lambda_b\rho_b(1-x_h) + \lambda_w\rho_w} + \gamma_1 \frac{\lambda_w\rho_w}{\lambda_b\rho_b(1-x_h) + \lambda_w\rho_w}\right] + \beta \\ &\quad -\theta(1-z) - c\left(N\left[\lambda_b\rho_b(1-x_h) + \lambda_w\rho_w\right]\right). \end{aligned}$$

Similarly, the utility for a less educated black with job location $z \in [0,1]$ from living in community WL is:

$$\begin{aligned} V_{\mathrm{WL}}^l(z; x_l) &= \alpha\left[\frac{\lambda_b(1-\rho_b)(1-x_l)}{\lambda_b(1-\rho_b)(1-x_l) + \lambda_w(1-\rho_w)} + \gamma_1 \frac{\lambda_w\rho_w}{\lambda_b(1-\rho_b)(1-x_l) + \lambda_w(1-\rho_w)}\right] \\ &\quad +\beta - \theta(1-z) - c\left(N\left[\lambda_b(1-\rho_b)(1-x_l) + \lambda_w(1-\rho_w)\right]\right). \end{aligned}$$

The equilibrium pair of thresholds (x_l^*, x_h^*) must satisfy

$$V_{\mathrm{B}}^h(x_h^*; x_h^*, x_l^*) = V_{\mathrm{WH}}^h(x_h^*; x_h^*), \tag{4}$$

$$V_{\mathrm{B}}^l(x_l^*; x_h^*, x_l^*) = V_{\mathrm{WL}}^l(x_l^*; x_l^*). \tag{5}$$

Equation (4) requires that the marginal type for highly educated blacks, ${x_h}^*$, is indifferent between living in community B (an all-black mixed-education community) and community WH (a highly educated community with a white majority). Equation (5) requires that the marginal type for less educated blacks ${x_l}^*$ is indifferent between living in community B and community WL (a less educated community with white majority). We assume that the parameters of the model are such that equation system (4) and (5) has solutions.

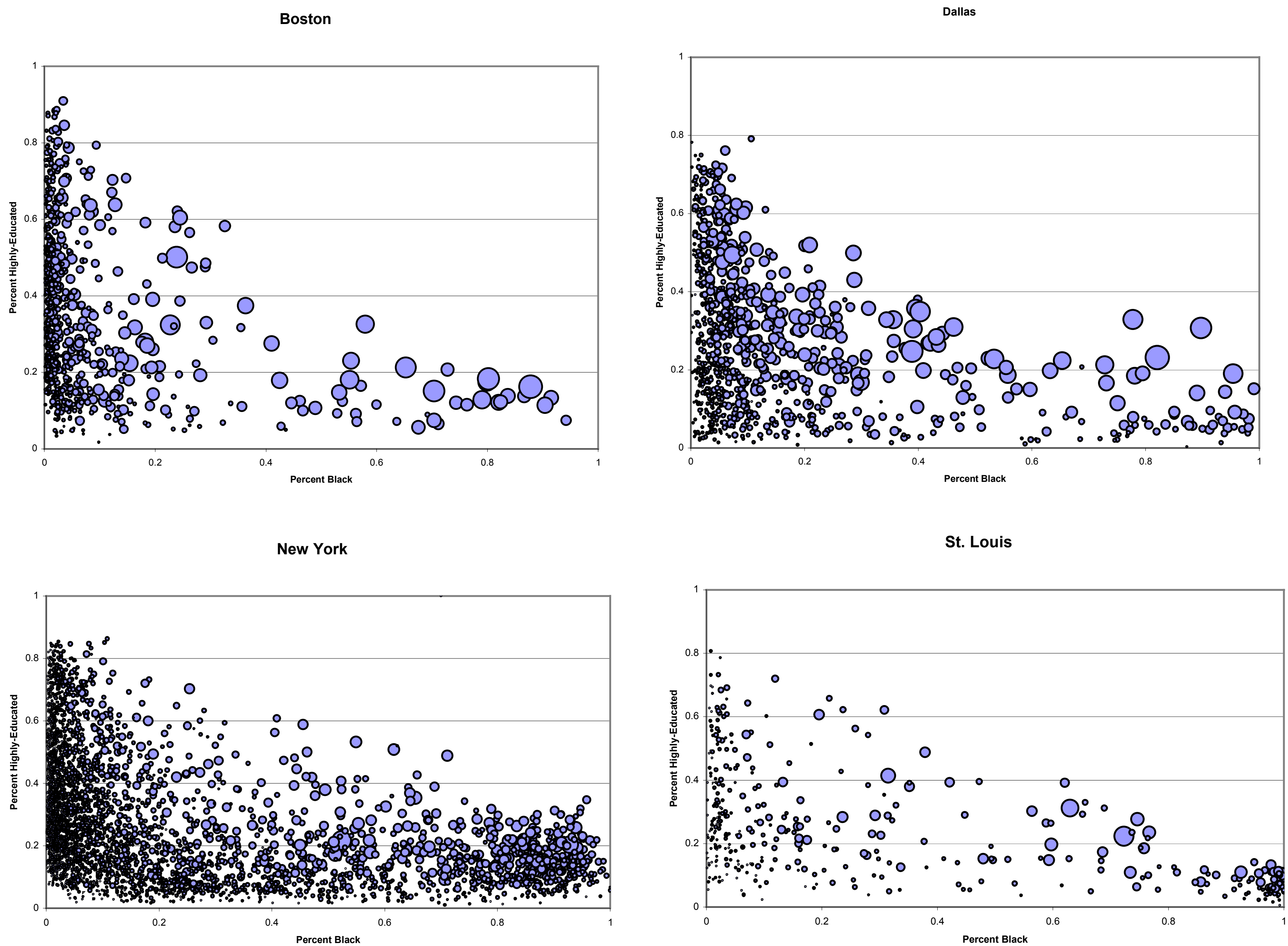

Figure 1: Neighborhood Choice Sets in Boston, Dallas, New York and St. Louis.

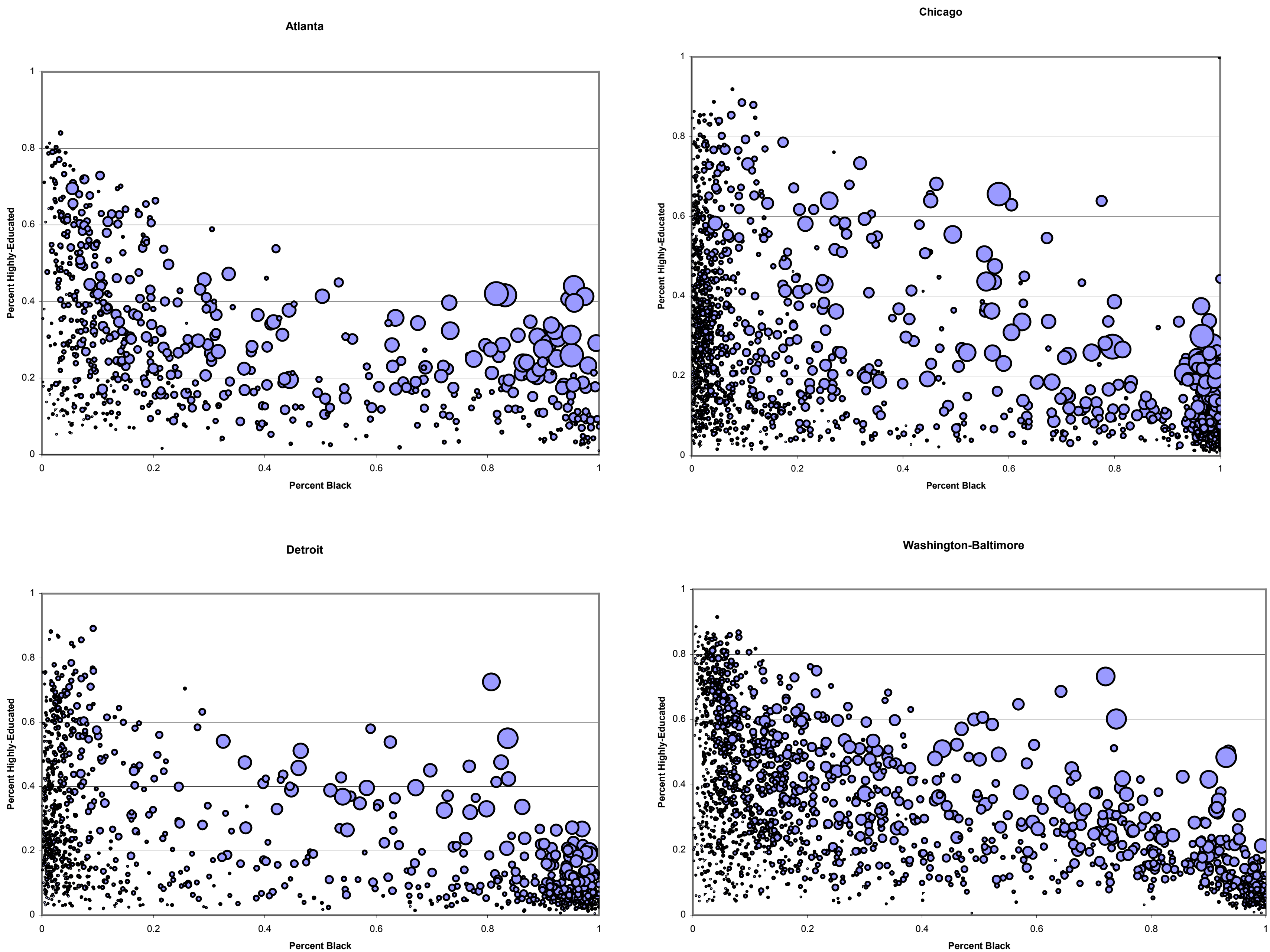

Figure 2: Neighborhood Choice Sets in Atlanta, Chicago, Detroit and Washington DC-Baltimore.

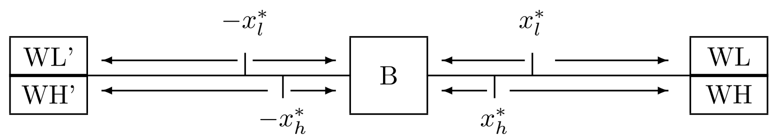

Figure 3: A Graphical Illustration of an Equilibrium.

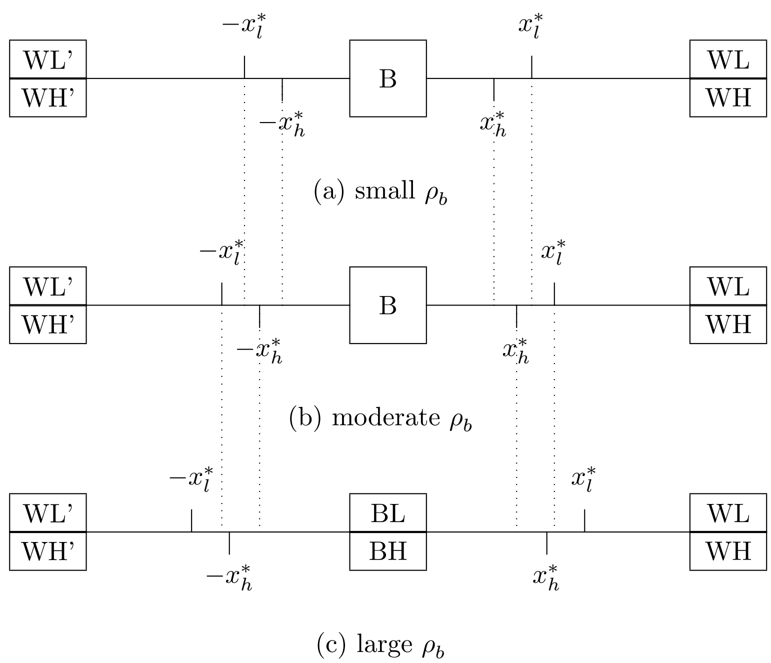

Figure 4: Comparative Statics with Respect to ρ_b.

Table 1: Make-Up of Population Living in US Metropolitan Areas

Race	Education	(1) **Percentage of Overall Population**	(2) **Percentage by Race**
Black	Less than HS	0.029	0.258
Non-Hispanic	HS	0.032	0.291
	Some College	0.033	0.297
	College Degree	0.011	0.102
	Advanced Degree	0.006	0.052
White	Less than HS	0.091	0.132
Non-Hispanic	HS	0.185	0.266
	Some College	0.192	0.277
	College Degree	0.124	0.178
	Advanced Degree	0.102	0.147

Note: The universe for this table are individuals 25 years and older in US metropolitan areas (approximately 155 million individuals in total). The columns report the fraction of individuals in each race-education category as a percentage of the total population and the population of the same race, respectively.

Table 2: 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	3,005
	100.0%	53.8%	22.6%	6.1%
Panel B: Tracts by Percent Black				
(2) > 20% Black	9,149	2,567	641	59
	100.0%	28.1%	7.0%	0.6%
(3) > 40% Black	5,657	1,164	142	14
	100.0%	20.6%	2.5%	0.2%
(4) > 60% Black	3,921	623	44	5
	100.0%	15.9%	1.1%	0.1%
(5) > 80% Black	2,559	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	2,999
	100.0%	58.3%	25.6%	6.9%
(7) > 40% White	39,602	24,566	10,839	2,967
	100.0%	62.0%	27.4%	7.5%
(8) > 60% White	35,154	22,543	10,214	2,870
	100.0%	64.1%	29.1%	8.2%
(9) > 80% White	26,910	17,539	8,102	2,339
	100.0%	65.2%	30.1%	8.7%

Note: The top number in each cell reports the number of tracts meeting both the education criterion described in the column heading (e.g., greater than 40 percent college-educated) and the race criterion in the row heading (e.g., greater than 40 percent black); the bottom number in each cell reports 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 3,000).

Table 3: Metropolitan Areas 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 US Metro Areas	21	44	142	154.84	0.11	0.15

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

Table 4: Neighborhood Patterns for College-Educated Individuals in the United States

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

Quintile	1	2	3	4	5	Total
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

Quintile	1	2	3	4	5	Total
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. metro areas reside, respectively. To construct the numbers in Panel A, college-educated blacks in each metro area are ranked by the fraction of blacks in their tract and assigned to one of five quintiles. Average neighborhood sociodemographic characteristics are then reported for each quintile, averaging across all metro areas. Panle 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 5: Neighborhood-Level Composition by Race and Education

Panel A: All Metropolitan Areas

	Neighborhood Composition relative to Metro Area (Z_n - Z_m)			
Individual	% Black College Degree	% Black < College Deg	% Black	% College Degree
Black with less than college degree	0.021	0.196	0.217	-0.063
Black with college degree	0.026	0.133	0.159	-0.021
White with less than college degree	-0.003	-0.014	-0.017	-0.009
White with college degree	-0.001	-0.010	-0.011	0.044

Panel B: Metropolitan Areas Below Median Fraction of College-Educated Blacks (<1.23 percent)

	Neighborhood Composition relative to Metro Area (Z_n - Z_m)			
Individual	% Black College Degree	% Black < College Deg	% Black	% College Degree
Black with less than college degree	0.013	0.111	0.124	-0.038
Black with college degree	0.016	0.078	0.094	0.005
White with less than college degree	-0.001	-0.006	-0.007	-0.004
White with college degree	0.000	-0.007	-0.007	0.045

Panel C: Metropolitan Areas Above Median Fraction of College-Educated Blacks (>1.23 percent)

	Neighborhood Composition relative to Metro Area (Z_n - Z_m)			
Individual	% Black College Degree	% Black < College Deg	% Black	% College Degree
Black with less than college degree	0.024	0.220	0.244	-0.070
Black with college degree	0.028	0.147	0.175	-0.027
White with less than college degree	-0.005	-0.024	-0.029	-0.014
White with college degree	-0.002	-0.012	-0.014	0.043

Note: Table reports average neighborhood (tract) characteristics Z_n for individuals in the race-education category shown in row heading relative to average composition of the individual's metropolitan area Z_m. Average compositions are reported for all metropolitan areas (Panel A) and for metro areas in which less than (Panel B) and more than (Panel C) 1.23 percent of the population is college educated and black, respectively. Tract and metropolitan compositions are calculated using individuals 25 years and older in U.S. metropolitan areas.

Table 6: Fixed Effects Regression of Neighborhood Composition on Interactions of Individual and Metro Characteristics

Dependent Variable:	**Neighborhood (Tract) Composition**			
	% Black Col Deg	**% Black < Col Deg**	**% Black**	**% College Degree**
	(1)	**(2)**	**(3)**	**(4)**
Individual_BlackHighEd* Metro_%BlackHighEd	0.968	3.040	4.008	-1.261
	(0.201)	(1.575)	(1.720)	(0.861)
Individual_BlackHighEd* Metro_%BlackLowEd	-0.065	0.069	0.004	0.128
	(0.062)	(0.360)	(0.413)	(0.124)
Individual_BlackHighEd* Metro_%WhiteHighEd	-0.058	-0.257	-0.315	0.070
	(0.018)	(0.091)	(0.104)	(0.042)
Individual_BlackHighEd* Metro_%WhiteLowEd	0.000	0.148	0.149	0.066
	(0.024)	(0.091)	(0.113)	(0.033)
Individual_BlackLowEd* Metro_%BlackHighEd	1.022	4.911	5.933	-2.340
	(0.119)	(1.983)	(2.030)	(1.295)
Individual_BlackLowEd* Metro_%BlackLowEd	-0.086	-0.062	-0.148	0.319
	(0.040)	(0.399)	(0.427)	(0.185)
Individual_BlackLowEd* Metro_%WhiteHighEd	-0.039	-0.325	-0.364	0.191
	(0.014)	(0.094)	(0.105)	(0.045)
Individual_BlackLowEd* Metro_%WhiteLowEd	0.014	0.218	0.232	0.079
	(0.017)	(0.124)	(0.140)	(0.044)
Individual_BlackHighEd	0.023	0.030	0.053	-0.007
	(0.015)	(0.060)	(0.073)	(0.024)
Individual_BlackLowEd	0.011	0.049	0.059	-0.086
	(0.011)	(0.078)	(0.088)	(0.029)
Individual_WhiteHighEd	-0.004	-0.038	-0.041	0.094
	(0.002)	(0.011)	(0.011)	(0.014)
Individual_WhiteLowEd	-0.005	-0.037	-0.041	0.035
	(0.002)	(0.010)	(0.012)	(0.012)
Includes MSA Fixed Effects	Yes	Yes	Yes	Yes

Note: All regressions include metropolitan area fixed effects. 'High Ed' refers to individuals with a college degree and 'Low Ed' refers to those with less than a college degree. Each regression is estimated on the sample of individuals 25 years and older in US metropolitan area (about 155 million observations when Census weights are applied). Tract and metropolitan compositions

Table 7: Predicted Change in Neighborhood Composition relative to Metropolitan Average (Z_n - Z_m)

Estimated effect of a one percent increase in fraction of college-educated blacks in MSA holding the fraction of blacks constant.

	Panel 1		Panel 2		Panel 3		Panel 4		Panel 5		Panel 6	
Sample:	Full Sample		Full Sample		Full Sample		MSA Pop < 200K		MSA Pop 200-600K		MSA Pop > 600K	
Defintion of 'High Ed':	Col. Deg. or More		Some Col. or More		Some Col. or More		Some Col. or More		Some Col. or More		Some Col. or More	
Individual:	*Black High Ed*	*Black Low Ed*	*Black High Ed*	*Black Low Ed*	*Black High Ed*	*Black Low Ed*	*Black High Ed*	*Black Low Ed*	*Black High Ed*	*Black Low Ed*	*Black High Ed*	*Black Low Ed*
Change in Rel. Neighborhood Exposure (Z_n - Z_m)												
% Black & High Ed	0.010	0.011	0.027	0.027	0.023	0.021	0.001	0.003	0.012	0.012	0.027	0.024
	0.000	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.644*	*0.311*	*0.015*	*0.001*	*0.000*	*0.001*
% Black & Low Ed	0.030	0.050	0.017	0.031	0.014	0.025	0.001	0.011	0.013	0.016	0.012	0.025
	0.119	*0.038*	*0.159*	*0.017*	*0.176*	*0.035*	*0.892*	*0.104*	*0.063*	*0.046*	*0.431*	*0.204*
% Black	0.040	0.061	0.044	0.058	0.038	0.046	0.002	0.014	0.025	0.028	0.040	0.049
	0.057	*0.012*	*0.014*	*0.003*	*0.006*	*0.003*	*0.810*	*0.119*	*0.022*	*0.012*	*0.067*	*0.055*
% Highly Educated	-0.014	-0.027	-0.003	-0.013	-0.002	-0.011	-0.005	0.002	-0.001	-0.002	-0.003	-0.017
	0.156	*0.126*	*0.680*	*0.094*	*0.869*	*0.351*	*0.471*	*0.653*	*0.849*	*0.758*	*0.763*	*0.195*
Includes interactions with region and population?	No		No		Yes		Yes		Yes		Yes	

Note: This table summarizes the predicted change in the average composition of tracts in which blacks with and without a college degree, respectively, reside given a one percent increase in the fraction of college-educated blacks in the metropolitan area holding the fraction fo black households constant. The coefficients report the change in tract composition relative to the metropolitan average, i.e., over and above the mechanical effect of changing the metropolitan composition. P-values adjusted for clustering at the metropolitan level are reported in italics.

Table 8. Cutler-Glaeser Education and Earnings Regressions by Age

Coefficient on interaction between black and metropolitan segregation (dissimilarity index)

	Age of Sample					
Dependent Variable	20-24	25-30	31-40	41-50	51-60	61-70
College Graduation	-0.094	-0.064	0.002	0.074	0.070	0.034
	(0.032)	(0.062)	(0.069)	(0.059)	(0.046)	(0.054)
Ln(Earnings)	-0.786	-0.433	-0.026	0.239	0.411	0.081
	(0.140)	(0.094)	(0.084)	(0.092)	(0.131)	(0.280)

Notes: This table reports the results of a series of regressions based on the specification used in Cutler and Glaeser (1997) to generate Table IV. 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 reported for these ages are not identical to those reported in Cutler and Glaeser but are very close. This is most likely attributable to the fact that we use the 5 percent sample of the 1990 Census while the 1 percent sample is used in Cutler and Glaeser. All other measures should be identical as we used the metropolitan characteristics used by Cutler and Glaeser, which Jacob Vigdor has graciously made available on his website.

Table 9: Assessing Across-Metropolitan Sorting on Observable Characteristics

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	Movers	Stayers	Movers	Movers	Stayers
	(1)	(2)	(3)	(4)	(5)	(6)
Individual Characteristic:						
Black with college degree	1.075	1.165	0.903	2.702	3.104	2.798
	(0.107)	(0.147)	(0.812)	(0.254)	(0.326)	(1.198)
Black with less than college degree	0.197	0.253	0.380	0.079	0.372	1.463
	(0.054)	(0.087)	(0.681)	(0.129)	(0.186)	(1.293)
White with college degree	0.157	0.170	-0.248	0.833	1.144	0.126
	(0.053)	(0.094)	(0.577)	(0.110)	(0.160)	(0.950)
White with less than college degree	-0.499	-0.561	-0.704	-1.380	-1.446	-1.609
	(0.052)	(0.075)	(0.562)	(0.139)	(0.141)	(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 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 1990. Separate regressions are reported for individuals that moved between metro areas and those that did not in the five years prior to the 2000 Census. For movers, a specification that includes fixed effects for the metro area of residence in 1995 is also reported. Standard errors adjusted for clustering at the metropolitan level are reported in parentheses.

Table 10: Assessing Sorting on Unobservables: Including Lagged and Differenced Metro Area Composition

	Panel 1			Panel 2			Panel 3		
Dependent Variable:	**PUMA Composition % Black - HS Degree or less**			**PUMA Composition % Black - Some College or more**			**PUMA Composition % Black**		
		Single Regression			***Single Regression***			***Single Regression***	
	Actual	***Lagged***	***Differenced***	***Actual***	***Lagged***	***Differenced***	***Actual***	***Lagged***	***Differenced***
Black with some college or more	0.010	0.011	0.004	0.023	0.024	0.017	0.033	0.035	0.020
	0.092	*0.085*	*0.722*	*0.000*	*0.000*	*0.000*	*0.001*	*0.001*	*0.076*
Black with HS degree or less	0.022	0.023	0.011	0.024	0.024	0.020	0.046	0.047	0.031
	0.008	*0.010*	*0.226*	*0.000*	*0.000*	*0.000*	*0.000*	*0.001*	*0.010*

Note: The first column of each panel corresponds to results presented in column 3 of table 7 using Census PUMAs rather than tracts as the definition of neighborhood instead of tracts. The second and third columns report the corresponding coefficients when 2000 MSA sociodemographics are decomposed into a lagged measure based on where the householded lived in 1995 and the difference between the 2000 and 1995 measure. For these results, a complete set of interactions correpsonding to those shown in Table 6 are included in the underlying regressions for both the lagged and differenced terms. P-values adjusted for clustering at the metropolitan level are reported in italics.

Table 11: Relating Changes in Segregation to Changes in Metropolitan Composition

Dependent Variable:	**Change in Dissimilarity Index (1990-2000)**	
Weights:	**MSA Population**	**None**
Change in Metropolitan Characteristics (1990-2000)		
% Black w/ College-Degree	3.013	3.472
	(0.822)	(0.632)
% Black w/ Less Than College-Degree	-0.214	-0.040
	(0.293)	(0.238)
% White w/ College-Degree	0.051	-0.004
	(0.191)	(0.174)
% White w/ Less Than College-Degree	0.052	0.310
	(0.170)	(0.148)
Population (in millions)	-0.0017	-0.0031
	(0.0009)	(0.0012)
Constant	-0.056	-0.042
	(0.009)	(0.007)
N	220	220

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 change in metropolitan composition over this same period. Rsegressions are based on the sample of metropolitan areas that appear in both 1990 and 2000 and are reported with and without weighting by metropolitan population.