

NBER WORKING PAPER SERIES

RACIAL SORTING AND NEIGHBORHOOD QUALITY

Patrick Bayer  
Robert McMillan

Working Paper 11813  
<http://www.nber.org/papers/w11813>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
December 2005

We would like to thank Fernando Ferreira for outstanding research assistance. We are grateful to Pat Bajari, Steve Berry, Matt Kahn, Tom Nechyba, Kim Rueben and Chris Timmins for many valuable discussions about this research, and to Maureen Cropper, David Cutler, Dennis Epple, Ed Glaeser, James Heckman, Vernon Henderson, Costas Meghir, Robert Moffitt, David Neumark, Steve Ross, Holger Sieg, Kerry Smith, and Jacob Vigdor for additional helpful comments. We also thank conference participants at the AEA, ERC, IRP, NBER, PET, SIEPR, and SITE, and seminar participants at Brown, Chicago, Colorado, Columbia, Duke, Johns Hopkins, Northwestern, NYU, PPIC, Simon Fraser, Stanford, Toronto, UBC, UC Berkeley, UC Irvine, UCLA, and Yale for helpful suggestions. This research was conducted at the California Census Research Data Center; our thanks to the CCRDC, and to Ritch Milby in particular. We gratefully acknowledge financial support for this project provided by the National Science Foundation under grant SES-0137289, the Public Policy Institute of California, and SSHRC. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2005 by Patrick Bayer and Robert McMillan. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Racial Sorting and Neighborhood Quality  
Patrick Bayer and Robert McMillan  
NBER Working Paper No. 11813  
December 2005  
JEL No. H0, J7, R0, R2

### **ABSTRACT**

In cities throughout the United States, blacks tend to live in significantly poorer and lower-amenity neighborhoods than whites. An obvious first-order explanation for this is that an individual's race is strongly correlated with socioeconomic status (SES), and poorer households can only afford lower quality neighborhoods. This paper conjectures that another explanation may be as important. The limited supply of high-SES black neighborhoods in most U.S. metropolitan areas means that neighborhood race and neighborhood quality are explicitly bundled together. In the presence of any form of segregating preferences, this bundling raises the implicit price of neighborhood amenities for blacks relative to whites, prompting our conjecture — that racial differences in the consumption of neighborhood amenities are significantly exacerbated by sorting on the basis of race, given the small numbers of blacks and especially high-SES blacks in many cities. To provide evidence on this conjecture, we estimate an equilibrium sorting model with detailed restricted Census microdata and use it to carry out informative counterfactual simulations. Results from these indicate that racial sorting explains a substantial portion of the gap between whites and blacks in the consumption of a wide range of neighborhood amenities — in fact, as much as underlying socioeconomic differences across race. We also show that the adverse effects of racial sorting for blacks are fundamentally related to the small proportion of blacks in the U.S. metropolitan population. These results emphasize the significant role of racial sorting in the inter-generational persistence of racial differences in education, income, and wealth.

Patrick Bayer  
Department of Economics  
Yale University  
Box 208264  
New Haven, CT 06520-8264  
and NBER  
patrick.bayer@yale.edu

Robert McMillan  
University of Toronto  
mcmillan@chass.utoronto.edu

# 1 INTRODUCTION

In cities throughout the United States, blacks tend to live in significantly poorer and lower-amenity neighborhoods than whites. Many researchers posit that such neighborhood differences play a central role in the perpetuation of racial inequality (see Wilson (1987) and Massey and Denton (1993), for example).<sup>1,2</sup> In trying to understand how these substantial gaps in neighborhood quality across race arise, an obvious first-order explanation is that an individual's race is strongly correlated with socioeconomic status (SES),<sup>3</sup> and poorer households can only afford lower quality neighborhoods. This paper conjectures that another explanation may be important, due to a combination of racial sorting and the short supply of predominantly black, high-amenity neighborhoods in almost all U.S. metropolitan areas.

The short supply is stark: while over 11,000 Census tracts in U.S. metropolitan areas are at least 40 percent college-educated, for example, a mere 44 of these tracts are also at least 60 percent black.<sup>4</sup> As a consequence of this short supply, neighborhood race and many other neighborhood characteristics are explicitly linked in the set of residential options available to most households: in order to choose high-amenity neighborhoods, households must typically live with a higher fraction of white neighbors, given that the full range of possible neighborhood options is not spanned.

The bundling of neighborhood race and other neighborhood amenities would be of little import if households had identical preferences for neighborhood racial composition or if race played no role in household location decisions. But given any form of segregating preferences in the population,<sup>5</sup> this bundling drives a wedge between the *implicit price* that whites versus blacks must pay in order to consume higher levels of a given neighborhood amenity. Imagine, for example, that black and white households preferred to live with neighbors of the same race, other things equal. In this situation, while all households would pay the direct cost of residing in high-

---

<sup>1</sup> In the words of Massey and Denton (1993), “a significant share of black America is condemned to experience a social environment where poverty and joblessness are the norm,... where educational failure prevails, and where social and physical deterioration abound. Through prolonged exposure to such an environment, black chances for social and economic success are drastically reduced” (page 2).

<sup>2</sup> The intergenerational persistence of income inequality across race is also an important theme in the work of Glenn Loury (see, for example, Loury (1977)), drawing attention to the role of negative externalities in the human capital accumulation process for blacks, some of which operate at the neighborhood level.

<sup>3</sup> For example, 15 percent of black adults compared to 33 percent of white adults had attained a 4-year college degree in U.S. metropolitan areas in 2000.

<sup>4</sup> Of the 44 tracts, 33 are located in one of four metropolitan areas: Baltimore-Washington, Detroit, Los Angeles, and Atlanta. Only 142 tracts are at least 40 percent black and at least 40 percent college-educated.

<sup>5</sup> Note that the implicit price wedge described here would arise under any of the following forms of segregating preferences: (i) if whites and blacks each preferred neighbors of the same race (ii) if whites and blacks both preferred black neighbors but blacks were willing to pay for them or (iii) if whites and blacks both preferred white neighbors but whites were willing to pay more for them.

amenity neighborhoods via higher housing prices, the increased fraction of whites in these neighborhoods would be welfare-enhancing for whites and welfare-decreasing for blacks. In turn, as each household responded in a decentralized way to its own implicit price when deciding where to live, the resulting consumption of neighborhood amenities by blacks would be lower than that of otherwise-identical whites.<sup>6</sup>

This discussion prompts our main conjecture: given the short supply of high-amenity, predominantly black neighborhoods in most metropolitan areas, racial sorting – whether driven by decentralized preferences or discrimination – is likely to exacerbate the gap between blacks and whites in the consumption of many neighborhood amenities.<sup>7,8</sup> Further, the resulting adverse effects of racial sorting for blacks may diminish as the proportion of (especially high-SES) blacks in the metropolitan population grows larger and high-amenity black neighborhoods form, given evidence (see Bayer, Fang, and McMillan (2005)) that the availability of mixed- and high-SES black neighborhoods is increasing in the proportion of highly educated blacks in the metropolitan population.

The central task of the current paper is to shed light on our conjecture empirically. To do so, one would ideally like to compare observed racial differences in the consumption of neighborhood amenities with those arising in a world in which households chose neighborhoods without regard for race. To investigate the role played by the proportion of highly educated blacks, in turn one would ideally like to examine the impact of increasing the proportion of blacks in the population on the consumption effects of racial sorting, while holding the strength of ‘preferences’ for neighborhood race constant.

The primary analysis presented in this paper is designed to implement these ideal thought experiments in an intuitive way. Specifically, using restricted-access Census data that precisely match nearly a quarter of a million households to their neighborhoods in the San Francisco Bay Area, we estimate a flexible equilibrium model of residential sorting and use it to conduct counterfactual simulations that correspond directly to these thought experiments.<sup>9</sup> At its heart,

---

<sup>6</sup> As we discuss in more detail below, discriminatory practices that increase the implicit price that blacks versus whites pay for housing in predominantly white neighborhoods clearly have the same implications regarding the consumption of neighborhood amenities by blacks versus whites. The goal of this paper is the measure the total impact of racial sorting not to distinguish between these alternative explanations.

<sup>7</sup> A gap between blacks and whites in the consumption of neighborhood amenities would arise even in the absence of racial sorting because of the substantial underlying differences between blacks and whites in socioeconomic characteristics (e.g., income, education) that contribute to residential sorting.

<sup>8</sup> While not generally explicitly linked to the size of the black population or the availability of high-SES black neighborhoods, the literature examining the spatial mismatch hypothesis, first proposed by Kain (1968), has spawned innumerable studies that suggest that residential segregation can exacerbate existing inequalities in employment access and outcomes. See Ihlanfeldt and Sjoquist (2000) for a recent survey.

<sup>9</sup> This model is presented fully in Bayer, McMillan, and Rueben (2005).

our primary estimation approach relies on the standard economic notion of revealed preference: by examining how location decisions vary on average with household characteristics (such as income, education and race) given the set of neighborhoods available in the market, we infer how the demand for housing and neighborhood attributes varies with these household characteristics. In estimating the model, we are also careful to address an important endogeneity problem that arises due to the correlation of neighborhood sociodemographic characteristics with unobserved aspects of housing and neighborhood quality (a correlation that is induced by residential sorting), implementing a boundary fixed effects strategy closely related to that of Black (1999). The resulting preference estimates are reasonable in magnitude across a wide set of housing and neighborhood attributes, providing evidence of strong interactions in the utility function between individual and neighborhood sociodemographic characteristics, consistent with stratification along both racial and socioeconomic dimensions in equilibrium.

Using the estimated household preferences along with our equilibrium model, we then carry out two complementary simulations that provide evidence on our conjecture. In the first, we compare the actual state of the world to one in which race is eliminated as a factor in each household's residential location decision.<sup>10,11</sup> Eliminating racial sorting – whether driven by preferences or discrimination – leads to a substantial decrease in racial segregation across all socioeconomic levels in the new equilibrium, although some segregation persists due to the sizeable average differences remaining between blacks and whites in education, income, and wealth. At the same time, average black-white differences in the consumption of many neighborhood amenities are also substantially reduced. Specifically, the results imply that eliminating neighborhood race from the location decision cuts the black-white gap in neighborhood school quality and crime by more than half and the gap in average neighborhood income and education by nearly this amount. This indicates that racial sorting is as important as the underlying socioeconomic differences between whites and blacks in driving differences in the consumption of neighborhood amenities.

---

<sup>10</sup> In light of the potential for racial discrimination in the housing market, it is important to point out that the 'preference' parameters estimated in our analysis combine the impact of preferences and discrimination. That is, discriminatory practices that increase the implicit price that blacks versus whites pay for housing in predominantly white neighborhoods will generally result in the estimation of stronger 'preferences' among blacks for predominantly black neighborhoods. In our first simulation, setting the 'preference' parameters in the model associated with neighborhood race to zero corresponds to eliminating both decentralized racial preferences and any centralized discriminatory practices that exist in the current state of the world.

<sup>11</sup> Mechanically, this involves calculating a new housing market equilibrium that allows households to re-sort and housing prices to adjust, having first set all of the utility function parameters associated with neighborhood racial composition equal to zero.

In the second simulation, we equalize (counterfactually) the proportions of blacks and whites in the metropolitan area, thereby relieving the existing bundling constraint while leaving all preference parameters at their estimated values. Though this simulation also leaves the socioeconomic distribution for both whites and blacks unchanged, the new equilibrium is characterized by a sharp *increase* in the availability of highly-amenity, predominantly-black neighborhoods. As a result, the observed white-black gaps in the consumption of neighborhood amenities decline by nearly the same amount as in our first simulation: the adverse effects of racial sorting for blacks in the Bay Area would be virtually eliminated were blacks to constitute a larger proportion of the metropolitan area's population. This result makes clear that the small proportion of black households found in the typical U.S. metropolitan area is central to racial sorting having adverse effects on the consumption of neighborhood amenities by blacks relative to whites.

That the effect of racial sorting declines with the proportion of (highly educated) blacks in the population gives rise to an implication that is testable across metropolitan markets. In particular, it implies that the adverse effects of segregation on black outcomes (relative to whites) should decline as the proportion of highly educated blacks increase. To provide complementary evidence on this aspect of our main hypothesis, we augment the primary regressions reported in Cutler and Glaeser (1997, henceforth 'CG') in order to examine how the effect of segregation varies with the proportion of highly educated blacks in the metropolitan population. This analysis demonstrates that the substantial negative effects of segregation for blacks relative to whites shown in CG decline to zero as the fraction of college-educated blacks in metropolitan population reaches 3 to 6 percent (the average fraction in U.S. metro areas in 2000 was 1.7 percent), again implying that the adverse effects of racial sorting are strongly linked to the small number of highly educated blacks in the population.

These empirical results draw attention to an important consequence of racial sorting – the distortion affecting the neighborhood *quality* choice<sup>12</sup> – and they provide clear evidence of the sheer size of this distortion. The substantial reductions in neighborhood quality for blacks that we find are likely to have a significant impact on the inter-generational persistence of racial differences in education, income, and wealth. Further, our analysis provides a new explanation as to why the distortion to neighborhood quality choice arises in the first place, due to a combination of a 'small numbers' problem and racial sorting.

---

<sup>12</sup> See Massey and Denton (1993) for a discussion in the sociology literature of the losses experienced by both middle-class and poor blacks as a result of "the limitation of black residential options through segregation" (page 9).

Our analysis informs a much larger body of research that examines the impact of segregation on individual outcomes. Most importantly, it implies that segregation *measured at either the neighborhood or metropolitan level* does not provide a sufficient statistic for estimating the impact of racial sorting. In particular, because racial sorting is directly responsible for a sizeable portion of the difference in neighborhood amenity consumption by blacks versus whites, analyses of the impact of neighborhood measures of segregation on individual outcomes that *condition* on these other neighborhood characteristics may inadvertently attribute too little of the black-white differences in outcomes to racial sorting.<sup>13</sup> Moreover, our second simulation makes clear that the adverse effect of racial sorting for blacks is not fully captured by an estimate of the impact of metropolitan segregation on individual outcomes. In particular, it shows plainly that an increase in the proportion of blacks in the population of a metropolitan area can reduce the adverse effects of racial sorting for blacks while simultaneously *increasing* the level of segregation.

The rest of the paper is organized as follows: Section 2 presents descriptive evidence from all metropolitan areas in the United States on neighborhood availability and segregation using 2000 Census data, helping motivate our conjecture. Section 3 outlines the key feature of our detailed San Francisco Bay Area dataset, and Section 4 provides descriptive evidence related to neighborhood supply and sorting just using Bay Area data. Sections 5, 6 and 7 describe the main analytical tool used in this paper - an equilibrium model of residential sorting - describing the model, its estimation, and the estimated parameters in turn. Section 8 then uses the estimated model to conduct two counterfactual simulations that shed light on our conjecture. Section 9 provides complementary evidence based on augmented Cutler-Glaeser regressions and Section 10 concludes.

## **2 PATTERNS OF RACIAL SORTING IN U.S. METROPOLITAN AREAS**

Our primary analysis will be based on detailed microdata drawn from the San Francisco Bay Area. In this section, we provide motivation for that analysis by characterizing the supply of neighborhoods in all U.S. cities. The primary goal of this initial analysis is to demonstrate that the fundamental conditions motivating our main conjecture exist in most U.S. metropolitan areas. In particular, we show that (i) neighborhoods combining high-fractions of both college-educated and black individuals are in short supply in almost every metropolitan area in the United States and (ii) that faced with the resulting trade-off between black versus other college-educated neighbors, college-educated blacks choose a very diverse set of neighborhoods in each metropolitan area.

---

<sup>13</sup> See, for example, Kain (1968), Ihlanfeldt (1992), Ihlanfeldt and Sjoquist (1990), O'Regan and Quigley (1998), Ross (1998), Weinberg (2000, 2004), and Ross and Zenou (2004), among many others.

This analysis also serves to highlight the fact that the racial composition and set of available neighborhoods in the San Francisco Bay Area are in fact comparable to the average patterns for U.S. metropolitan areas as a whole.

Our approach is straightforward. Using publicly available Census Tract Summary Files (SF3) from the 2000 Census, we characterize the distribution of race and neighborhood quality for all neighborhoods in U.S. metropolitan areas (MSAs).<sup>14</sup> We also plot the distribution of available neighborhoods in individual MSAs, drawing attention to important patterns in terms of neighborhood availability.

In this portion of the analysis, a ‘neighborhood’ corresponds to a Census tract, which typically contains 3,000 to 5,000 individuals and we summarize neighborhood quality in a single dimension - the fraction of residents who are college-educated. (Using far richer data for the San Francisco Bay Area, we consider a much wider set of neighborhood attributes below.) In terms of racial composition, we focus on non-Hispanic black and non-Hispanic white individuals 25 years and older residing in U.S. metropolitan areas. Non-Hispanic 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 a four-year college degree, while the comparable number for whites is 32.5 percent.

We begin by showing that, while neighborhoods combining 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. To that end, Table 1 documents the number of tracts in the U.S. by the percentage of individuals with a college degree and the percentage of individuals who are black or 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 then 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 1 presents analogous numbers for whites. They show a



markedly different pattern of neighborhood choices for whites, with a far greater fraction of neighborhoods with at least 40, 60, and 80 percent whites meeting the education criteria listed in the column headings.

In addition to being in short supply overall, neighborhoods combining high fractions of both college-educated and black individuals are concentrated in only a handful of metropolitan areas, most notably Baltimore-Washington DC, indicating that the supply of such neighborhoods in most metropolitan areas is even more limited.<sup>15</sup> The absence of these neighborhoods means that neighborhood race and many other neighborhood characteristics are explicitly linked in the set of residential options available to most households: in order to choose neighborhoods with more college-educated neighbors, households must typically live with a greater fraction of whites. It is the explicit bundling of neighborhood race and other neighborhood amenities in the absence of high-amenity black neighborhoods that motivates our main hypothesis.

To illustrate this potential trade-off, Figure 1 shows scatterplots of available neighborhoods in three metropolitan areas: Boston, Dallas, San Francisco and St. Louis. In the scatterplots, 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.

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 each plot, *had they existed*. Figure 1 also demonstrates that, facing this constrained choice set, college-educated blacks choose to live in a diverse set of neighborhoods: while a sizeable fraction of college-educated blacks in each of these 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).

To show that the patterns shown in Figure 1 are representative of those for U.S. metropolitan areas as a whole, Panel A of Table 2 summarizes the characteristics of neighborhoods in which

---

<sup>14</sup> In this section, 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).

<sup>15</sup> Of the 44 tracts that are at least 60 percent black and 40 percent college-educated, for example 14 are in Baltimore-Washington DC, 8 in Detroit, 6 in Los Angeles, and 5 in Atlanta. 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. See Bayer, Fang, and McMillan (2005) for more details.

college-educated blacks reside in all U.S. MSAs. We first rank college-educated blacks in each MSA 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 2 then summarizes the neighborhood characteristics corresponding to these quintiles, averaged over all U.S. metropolitan areas.

The table shows 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 2 reports analogous numbers for college-educated whites. Comparison of Panels A and B reveals that the 40 percent of college-educated blacks in each metropolitan area who live in the tracts with the smallest fraction of other blacks have roughly the same fraction of college-educated neighbors as college-educated whites do on average; however, college-educated blacks living 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 in Table 2 choose segregated neighborhoods with relatively low average education attainment is consistent with two aspects our main conjecture. First, it suggests that, whether due to preferences or discrimination, race remains an important factor in the location decisions of a large number of college-educated blacks. This helps to cast doubt on an obvious potential explanation for the absence of neighborhoods combining high fractions of both college-educated and black individuals – namely, that college-educated blacks simply chose college-educated neighborhoods without regard for the racial composition. Second, the especially low levels of average neighborhood educational attainment for those college-educated blacks living in the most segregated neighborhoods suggests that racial sorting may indeed have a substantial impact on the consumption of neighborhood amenities by blacks. It is important to recognize, however, that the evidence presented in Table 2 is far from conclusive in this regard. Other explanations for this pattern related to heterogeneity in other important individual attributes (e.g., income and wealth) or in neighborhood amenities other than average educational attainment might explain these observed patterns. It is precisely these alternative potential explanation that our primary analysis, which conditions on a wide set of individual and neighborhood attributes, is designed to address.

Finally, we motivate the idea that the short supply of high-amenity black neighborhoods may be systematically relaxed as the number of college-educated blacks in a metropolitan areas

increases with a combination of regression analysis and scatterplots for a different set of U.S. cities.<sup>16,17</sup> In particular, Figure 2 depicts analogous scatterplots to those presented in Figure 1 for Atlanta, Chicago, Detroit and Washington DC – MSAs that contain a much more sizeable number of college-educated blacks than those MSAs shown in Figure 1. As Figure 2 clearly shows, the supply of neighborhoods combining relatively high fractions of both black and highly educated individuals is substantially greater in these MSAs. As a result, we would expect the adverse effects of racial sorting for the consumption of neighborhood amenities by blacks versus whites to be less severe in these MSAs. We provide evidence on this aspect of our main conjecture in the second simulation and with additional across-metropolitan analysis, which reveal that the adverse effects of racial sorting for black neighborhood amenity consumption are indeed directly related to the fact that blacks (college-educated blacks in particular) represent such a small fraction of the population of the San Francisco Bay Area and, more generally, the U.S. metropolitan population.

### 3 DATA FOR PRIMARY ANALYSIS

Given this broad characterization of racial sorting in MSAs throughout the U.S., we now turn to the much more detailed dataset for our primary analysis. The particular dataset that we construct is based primarily on restricted Census microdata for the San Francisco Bay Area for 1990. These restricted Census data provide the same detailed individual, household, and housing variables found in the public-use version of the Census, but also include information on the location of individual residences and workplaces at a very disaggregate level. In particular, while public-use data specify the PUMA (a Census region with at least 100,000 individuals) in which a household lives, the restricted data specify the Census block (a Census region with an average of 100 individuals), thereby identifying an individual’s neighborhood far more precisely than has been previously possible with such a large data set. These data allow us to estimate a model of residential sorting on the part of individual households, while controlling carefully for a wide set

---

<sup>16</sup> 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.

<sup>17</sup> 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.

of household characteristics and making use of reasonable variation in the data to identify the impact of a wide variety of factors (including neighborhood racial composition) on each individual's location decision.

In assembling our Bay Area dataset, we use data from six contiguous counties: Alameda, Contra Costa, Marin, San Mateo, San Francisco, and Santa Clara. The resulting study area is reasonably self-contained and sizeable along a number of dimensions, including over 1,100 Census tracts, and almost 39,500 Census blocks, the smallest unit of aggregation in the data. Our final sample consists of just over 242,000 households. We also note that, among the largest metropolitan areas in the US, the fraction of black and white households in the San Francisco Bay Area (68.6 percent white, 7.6 percent black) most closely matches that of the country as a whole (69.5 percent white, 11.1 percent black).

The Census provides a wealth of data on the individuals in the sample – race, age, educational attainment, income from various sources, household size and structure, occupation, and employment location.<sup>18</sup> In addition, it provides a variety of housing characteristics: whether the unit is owned or rented, the corresponding rent or owner-reported value,<sup>19</sup> number of rooms, number of bedrooms, type of structure, and the age of the building. We use these housing characteristics directly and in constructing neighborhood characteristics, characterizing stock of housing in the neighborhood surrounding each house, as well as neighborhood racial, education and income distributions based on the households within the same Census block group, a Census region containing approximately 500 housing units. We merge additional data describing local conditions with each house record, constructing variables related to crime rates, land use, local schools, topography, and urban density.<sup>20</sup> The list of the principal housing and neighborhood variables used in the analysis, along with means and standard deviations, is given in the first two columns of Table 3.

---

<sup>18</sup> Throughout our analysis, we treat the household as the decision-making agent and characterize each household's race as the race of the 'householder', assigning households to one of four mutually exclusive categories of race/ethnicity: Hispanic, non-Hispanic Asian, non-Hispanic black, and non-Hispanic white. To maintain a streamlined exposition of results, we focus on black and white households, although it is important to point out that our primary analysis also controls separately for Asian and Hispanic households.

<sup>19</sup> As described in the Data Appendix, we construct a single price vector for all houses, whether rented or owned. Because the implied relationship between house values and current rents depends on expectations about the growth rate of future rents in the market, we estimate a series of hedonic price regressions for each of over 40 sub-regions of the Bay Area housing market. These regressions return an estimate of the ratio of house values to rents for each of these sub-regions.

<sup>20</sup> For each of these measures, a detailed description of the process by which the original data were assigned to each house is provided in a Data Appendix.

#### 4 RACIAL SORTING AND NEIGHBORHOOD AMENITIES IN THE BAY AREA

Before turning to our model of residential sorting, we first describe the patterns of racial sorting and consumption of neighborhood amenities in the San Francisco Bay Area. The main goal of this section is to illustrate that the pattern demonstrated for average neighborhood education for all U.S. metropolitan areas extends to a wider set of neighborhood amenities. This analysis also illustrates how housing prices vary with neighborhood amenities and racial composition.

**Segregation Patterns.** We begin by describing the pattern of racial segregation in the Bay Area. We do so by examining the average compositions of the neighborhoods (Census block groups in this case) in which households in a particular category of race and income reside.<sup>21</sup> These measures are reported in Panel A of Table 4. The measures in the first row show neighborhood compositions averaged over all of the households in the Bay Area. The remaining rows report neighborhood compositions averaged over the set of households described in the row heading. The second row, for example, indicates that black households live in neighborhoods in which an average of 20.9 percent of the households are both black and in the lowest income quartile, 9.0 percent are black and in the second income quartile, etc. Average neighborhood compositions are reported for blacks and whites as a whole and for blacks and whites in the lowest and highest income quartile, respectively.

Panel B of Table 4 re-summarizes these average neighborhood composition measures reported in Panel A in a way that is more meaningful for discussing segregation, reporting the average neighborhood composition for a particular type of household relative to the average for the Bay Area as a whole. For example, the first entry of Panel B is calculated as the average exposure of black households to blacks in the lowest income quartile (20.9 percent), divided by the average exposure of all households in the Bay Area to blacks in the lowest income quartile (3.4 percent). This implies that a black household in the Bay Area is exposed on average to 6.1 times the fraction of blacks in the lowest income quartile that the average household in the Bay Area is exposed to.

Panel B reveals a clear pattern of racial segregation for Bay Area blacks that cuts across all income categories. While blacks in the lowest income quartile are exposed to 6.2 times the fraction of blacks relative to the average household in the metropolitan area, the comparable figure for blacks in the highest income quartile is 3.2, which indicates a substantial amount of

---

<sup>21</sup> We use income throughout the remainder of the paper as a proxy for socioeconomic status in describing the results or our analysis. Results based on education or income conditional on education are completely analogous.

segregation even for high-income blacks. Moreover, while high-income blacks are especially highly exposed to blacks in the highest income quartile (4.4 times the Bay Area average), their exposure to blacks in the lowest income quartile also remains high at 2.9 times the Bay Area average. Thus race continues to play a large role in the residential choice process even for high-income blacks.

**Racial Sorting and the Consumption of Housing and Neighborhood Amenities.** To explore this possibility directly, Table 5 describes the distribution of neighborhoods in which blacks and whites in the highest quartile of the income distribution reside, respectively.<sup>22</sup> In each case, as in Table 2, neighborhoods are first ranked by the fraction of a household's neighbors that are of the same race, and quintiles of the distribution are then reported.

Panel A shows the distribution of neighborhoods in which blacks in the highest income quartile reside. We order these households into five quintiles based on percentage black in the neighborhood. Thus the first column provides average housing and neighborhood characteristics for the 20 percent of high-income blacks who live in neighborhoods with the lowest fraction of black households, neighborhoods in which less than 4 percent of the population is black. As one reads across the columns, the neighborhoods have a higher fraction of black households by construction; the final column indicates that fully 20 percent of blacks in the highest income quartile reside in neighborhoods in which over 54 percent of the population is black.

As in Table 2, what emerges from Panel A of Table 5 is a clear picture of the wide range of neighborhoods in which high-income blacks reside. Comparing the neighborhoods at either end of the spectrum, the levels of school quality, public safety, average neighborhood income, and fraction college-educated are each 1.5 to 2 standard deviations greater in the neighborhoods with the least versus the greatest fraction of black households. Panel B of Table 5 shows the same distribution for whites in the highest income quartile, first ranking neighborhoods by percent white and again reporting the quintiles of this distribution. For whites, increases in the fraction of white neighbors are accompanied by increases rather than decreases in the consumption of housing and neighborhood amenities. Thus, while increased consumption of neighborhood amenities such as school quality, public safety, and neighborhood education and income comes at the expense of increased housing prices for households of each race, these increases are accompanied by sharp decreases in the fraction of households of the same race for blacks and increases in the fraction of households of the same race for whites.

---

<sup>22</sup> An analogous table for blacks and whites in the lowest income quartile is included as Appendix Table 1. This table generally shows a similar, although muted, pattern as Table 5.

While highly suggestive, consumption patterns like those presented in Table 5 do not provide conclusive evidence concerning the impact of that racial sorting on the consumption of neighborhood amenities. First, the wide variation in neighborhood amenity consumption among high-income blacks might reflect heterogeneity in other important individual attributes (e.g., income and wealth) or in neighborhood amenities other than those shown in the table. Second, as Panel A reveals, those high-income blacks that live in the lowest amenity neighborhoods spend substantially less on housing. Thus, another potential explanation for the lower average consumption of neighborhood amenities by blacks relative to whites is that blacks have a lower average willingness-to-pay for these amenities relative to whites.

The primary empirical analysis that we now present isolates the impact of racial sorting from these alternative explanations by allowing each of these channels to contribute to observed residential location decisions. Specifically, by conditioning on a wide set of individual and neighborhood attributes, this analysis explicitly conditions out any of the variation in neighborhood amenity consumption that can be explained by heterogeneity in these attributes. Second, by allowing willingness-to-pay for each housing and neighborhood amenity to vary completely flexibly by race, our analysis explicitly allows for the possibility that racial differences in demand for amenities might explain the observed pattern of neighborhood amenity consumption. Our subsequent empirical analysis does imply that each of these alternative channels do play a role in explaining consumption patterns. Even after controlling for these alternatives in a very flexible way, however, our analysis leads to the clear conclusion that sorting on the basis of neighborhood racial composition – whether due to preferences or discrimination – drives a substantial fraction of observed racial differences in neighborhood amenity consumption.

## **5 A MODEL OF RESIDENTIAL SORTING**

To measure the impact of racial sorting on the consumption of neighborhood amenities, we now turn to a model of the residential location decision of households in the Bay Area. In developing such a model, our goal is to provide the simplest analytical tool that can account for (i) heterogeneity in both household attributes and the attributes of houses/neighborhoods and (ii) the endogenous determination of housing prices and neighborhood demographic compositions.

To this end, we adopt the equilibrium model of an urban housing market developed in Bayer, McMillan, and Rueben (2004b). This equilibrium model consists of two key elements: the household residential location decision problem and a market-clearing condition. While maintaining a simple structure, the model allows households to have heterogeneous preferences defined over housing and neighborhood attributes in a very flexible way; it also allows for

housing prices and neighborhood demographic compositions to be determined in equilibrium. In estimating the model, we are careful to account for the correlation that naturally arises between unobserved housing and neighborhood attributes and both housing prices and neighborhood composition. Having estimated the model, we then use it to conduct two equilibrium counterfactual simulations that provide direct evidence on our main conjecture.<sup>23</sup>

**The Residential Location Decision.** We model the residential location decision of each household as a discrete choice of a single residence from a set of houses available in the market. The utility function specification is based on the random utility model developed in McFadden (1973, 1978) and the specification of Berry, Levinsohn, and Pakes (1995), which includes choice-specific unobservable characteristics.<sup>24</sup> Let  $X_h$  represent the observable characteristics of housing choice  $h$  including characteristics of the house itself (e.g., size, age, and type), its tenure status (rented vs. owned), and the characteristics of its neighborhood (e.g., school, crime, and topography). We use  $\bar{Z}$  to represent the average sociodemographic characteristics of the corresponding neighborhood, writing it separately from the other housing and neighborhood attributes to make explicit the fact that these characteristics are determined in equilibrium.<sup>25</sup> Let  $p_h$  denote the price of housing choice  $h$  and let  $d_h^i$  denote the distance from residence  $h$  to the primary work location of household  $i$ . Each household chooses its residence  $h$  to maximize its indirect utility function  $V_h^i$ :

$$(1) \underset{(h)}{\text{Max}} \quad V_h^i = \alpha_X^i X_h + \alpha_Z^i \bar{Z}_h - \alpha_p^i p_h - \alpha_d^i d_h^i + \xi_h + \varepsilon_h^i.$$

The error structure of household indirect utility is divided into a correlated component associated with each house valued the same by all households,  $\xi_h$ , and an individual-specific term,  $\varepsilon_h^i$ . A

---

<sup>23</sup> A long line of theoretical studies, including important papers by Epple, Filimon and Romer (EFR) (1984, 1993), Benabou (1993, 1996), Anas and Kim (1995), Anas (2002), Fernandez and Rogerson (1996, 1998), and Nechyba (1999, 2000) have developed and used models of sorting to analyze the way that interdependent individual decisions in the housing market aggregate up to determine the equilibrium structure of a metropolitan area. In recent years, a new line of empirical research has sought to take these models to the data. Epple and Sieg (1999) develop an estimator for the equilibrium sorting model of EFR, providing the first unified treatment of theory and empirics in the literature. In the same vein, Sieg *et al.* (2004) use this approach to explore the general equilibrium impacts of air quality improvements in the Los Angeles Basin.

<sup>24</sup> Discrete choice applications in the urban economics literature include Anas (1982), Quigley (1985), Gabriel and Rosenthal (1989), Nechyba and Strauss (1998), and Bajari and Kahn (2001). Only the latter paper includes choice-specific unobservables. Brock and Durlauf (2001) study a general class of discrete choice models with social interactions that do not include choice-specific unobservables.

<sup>25</sup> This component of the utility function allows for endogenous sorting on the basis of race, as in Schelling (1969, 1971), as well as on the basis of other characteristics such as income and education.



useful interpretation of  $\xi_h$  is that it captures unobserved housing quality, including any unobserved quality associated with the surrounding neighborhood.<sup>26</sup>

Each household's valuation of choice characteristics is allowed to vary with its own characteristics,  $Z^i$ , including education, income, race, employment status, and household composition. Specifically, each parameter associated with housing and neighborhood characteristics and price,  $\alpha_j^i$ , for  $j \in \{X, \bar{Z}, d, p\}$ , varies with a household's own characteristics according to:

$$(2) \alpha_j^i = \alpha_{0j} + \sum_{k=1}^K \alpha_{kj} Z_k^i,$$

with equation (2) describing the parameters of household  $i$ 's preference for choice characteristic  $j$ .

**Characterizing the Housing Market.** As with all models in this literature, the existence of a sorting equilibrium is much easier to establish if the individual residential location decision problem is smoothed in some way. To this end, we assume that the housing market can be fully characterized by a set of housing types that is a subset of the full set of available houses, letting the supply of housing of type  $h$  be given by  $S_h$ .<sup>27</sup>

Given the household's problem described in equations (1)-(2), household  $i$  chooses housing type  $h$  if the utility that it receives from this choice exceeds the utility that it receives from all other possible house choices - that is, when

$$(3) V_h^i > V_k^i \Rightarrow W_h^i + \varepsilon_h^i > W_k^i + \varepsilon_k^i \Rightarrow \varepsilon_h^i - \varepsilon_k^i > W_k^i - W_h^i \quad \forall k \neq h$$

where  $W_h^i$  includes all of the non-idiosyncratic components of the utility function  $V_h^i$ . As the inequalities in (3) imply, the probability that a household chooses any particular choice depends in general on the characteristics of the full set of possible house types. Thus the probability  $P_h^i$  that household  $i$  chooses housing type  $h$  can be written as a function of the full vectors of house/neighborhood characteristics (both observed and unobserved) and prices  $\{X, p, \xi\}$ :

$$(4) P_h^i = f_h(Z^i, \mathbf{X}, \mathbf{p}, \xi)$$

---

<sup>26</sup> We employ an indirect utility function that is linear in housing prices. Alternative specifications of the indirect utility function could certainly be estimated, as the linear form is not essential to the model.

<sup>27</sup> We also assume that each household observed in the sample represents a continuum of households with the same observable characteristics, with the distribution of idiosyncratic tastes  $\varepsilon_h^i$  mapping into a set of choice probabilities that characterize the distribution of housing choices that would result for the continuum of households with a given set of observed characteristics. For expositional ease and without loss of generality, we assume that the measure of this continuum is one.

as well as the household's own characteristics  $Z^i$ .

Aggregating the probabilities in equation (4) over all observed households yields the predicted demand for each housing type  $h$ ,  $D_h$ :

$$(5) \quad D_h = \sum_i P_h^i.$$

In order for the housing market to clear, the demand for houses of type  $h$  must equal the supply of such houses and so:

$$(6) \quad D_h = S_h, \quad \forall h \quad \Rightarrow \quad \sum_i P_h^i = S_h \quad \forall h.$$

Given the decentralized nature of the housing market, prices are assumed to adjust in order to clear the market. The implications of the market clearing condition defined in equation (6) for prices are very standard, with excess demand for a housing type causing price to be bid up and excess supply leading prices to fall. Given the indirect utility function defined in (1) and a fixed set of housing and neighborhood attributes, Bayer, McMillan, and Rueben (2004b) show that a unique set of prices (up to scale) clears the market.

Given that some neighborhood attributes are endogenously determined by the sorting process itself, we define a sorting equilibrium to be a set of residential location decisions and a vector of housing prices such that the housing market clears and each household makes its optimal location decision given the location decisions of all other households. In equilibrium, the vector of neighborhood sociodemographic characteristics along with the corresponding vector of market clearing prices must give rise to choice probabilities that aggregate back up to the same vector of neighborhood sociodemographics.<sup>28</sup>

Whether this model gives rise to multiple equilibria depends on the distributions of preferences and available housing choices, as well as the utility parameters. In general, it is impossible to establish that the equilibrium is unique a priori. Fortunately, estimation of the model does not require the computation of an equilibrium nor uniqueness more generally, as we describe in the next section.

---

<sup>28</sup> Bayer, McMillan, and Rueben (2004b) establish the existence of a sorting equilibrium as long as (i) the indirect utility function shown in equation (1) is decreasing in housing prices for all households; (ii) indirect utility is a continuous function of neighborhood sociodemographic characteristics; and (iii)  $\epsilon$  is drawn from a continuous density function.

## 6 ESTIMATION

Estimation of the model follows a two-step procedure related to that developed in Berry, Levinsohn, and Pakes (1995). A rigorous presentation of the estimation procedure is included in a technical appendix, including a discussion of methods for simplifying the computation and a description of the asymptotic properties of the estimator. In this section, we outline the estimation procedure, focusing on the identification of the model.

It is helpful to first introduce some notation. In particular, we rewrite the indirect utility function as:

$$(7) \quad V_h^i = \delta_h + \lambda_h^i + \varepsilon_h^i$$

where

$$(8) \quad \delta_h = \alpha_{0X} X_h + \alpha_{0\bar{Z}} \bar{Z}_h - \alpha_{0p} p_h + \xi_h$$

and

$$(9) \quad \lambda_h^i = \left( \sum_{k=1}^K \alpha_{kX} Z_k^i \right) X_h + \left( \sum_{k=1}^K \alpha_{k\bar{Z}} Z_k^i \right) \bar{Z}_h - \left( \sum_{k=1}^K \alpha_{kp} Z_k^i \right) p_h.$$

In equation (8),  $\delta_h$  captures the portion of utility provided by housing type  $h$  that is common to all households, and in (9),  $k$  indexes household characteristics. When the household characteristics included in the model are constructed to have mean zero,  $\delta_h$  is the mean indirect utility provided by housing choice  $h$ . The unobservable component of  $\delta_h$ ,  $\xi_h$ , captures the portion of unobserved preferences for housing choice  $h$  that is correlated across households, while  $\varepsilon_h^i$  represents unobserved preferences over and above this shared component.

The first step of the estimation procedure is equivalent to a Maximum Likelihood estimator applied to the individual location decisions taking prices and neighborhood sociodemographic compositions as given,<sup>29</sup> returning estimates of the heterogeneous parameters in  $\lambda$  and mean indirect utilities,  $\delta_h$ . This estimator is based simply on maximizing the probability that the model correctly matches each household observed in the sample with its chosen house type. In particular, for any combination of the heterogeneous parameters in  $\lambda$  and mean indirect utilities,

---

<sup>29</sup> Formally, the validity of this first stage procedure requires the assumption that the observed location decisions are individually optimal, given the collective choices made by other households and the vector of market-clearing prices *and* that households are sufficiently small such that they do not interact strategically with respect to particular draws on  $\varepsilon$ . This ensures that no household's particular idiosyncratic preferences affect the equilibrium and the vector of idiosyncratic preferences  $\varepsilon$  is uncorrelated with the prices and neighborhood sociodemographic characteristics that arise in any equilibrium. For more discussion, see the Technical Appendix.

$\delta_h$ , the model predicts the probability that each household  $i$  chooses house type  $h$ . We assume that  $\varepsilon_h^i$  is drawn from the extreme value distribution, in which case this probability can be written:

$$(10) \quad P_h^i = \frac{\exp(\delta_h + \hat{\lambda}_h^i)}{\sum_k \exp(\delta_k + \hat{\lambda}_k^i)}$$

Maximizing the probability that each household makes its correct housing choice gives rise to the following log-likelihood function:

$$(11) \quad \ell = \sum_i \sum_h I_h^i \ln(P_h^i)$$

where  $I_h^i$  is an indicator variable that equals 1 if household  $i$  chooses house type  $h$  in the data and 0 otherwise. The first step of the estimation procedure consists of searching over the parameters in  $\lambda$  and the vector of mean indirect utilities to maximize  $\ell$ .

**The Endogeneity of Neighborhood Sociodemographic Composition.** Having estimated the vector of mean indirect utilities in the first stage of the estimation, the second stage of the estimation involves decomposing  $\delta$  into observable and unobservable components according to the regression equation (8).<sup>30</sup> In estimating equation (8), important endogeneity problems need to be confronted. To the extent that house prices partly capture house and neighborhood quality unobserved to the econometrician, so the price variable will be endogenous. Estimation via least squares will thus lead to price coefficients being biased towards zero, producing misleading willingness-to-pay estimates for a whole range of choice characteristics. This issue arises in the context of any differentiated products demand estimation and we describe the construction of an instrument for price in the Technical Appendix.

A second identification issue of particular concern for our application relates to the correlation of neighborhood sociodemographic characteristics in  $\bar{Z}$  (which includes neighborhood race, income and education, as well as school quality) with unobserved housing and neighborhood quality,  $\xi_h$  - a correlation that arises given the sorting of households across locations. To properly estimate preferences in the face of this endogeneity problem, we adapt a technique previously developed by Black (1999) when estimating preferences for school quality. Black's strategy makes use of a sample of houses near school attendance zone boundaries,

---

<sup>30</sup> Notice that the set of observed residential choices provides no information that distinguishes the components of  $\delta$ . That is, however  $\delta$  is broken into components, the effect on the probabilities shown in equation (10) is identical.

estimating a hedonic price regression that includes boundary fixed effects. Intuitively, the idea is to compare houses in the same local neighborhood but on opposite sides of the boundary, exploiting the discontinuity in the right to attend a given school. For our purposes, boundary fixed effects are likely to absorb out differences in many fixed housing and neighborhood attributes, including ones that are unobservable.<sup>31</sup> To the extent that sorting with respect to the school district boundaries that we use is driven by differences in school quality and neighborhood sociodemographics themselves, the use of boundary fixed effects isolates variation in neighborhood sociodemographics that is uncorrelated with variation in unobserved housing and neighborhood quality. Thus, it provides an appealing way to account for the correlation of neighborhood sociodemographics with unobservable neighborhood quality in addition to the correlation of school quality with unobservable neighborhood quality as well as.

Table 3 displays descriptive statistics for various sub-samples related to the boundaries as well as the full sample. The first two columns report means and standard deviations for the full sample while the third column reports means for the sample of houses within 0.25 miles of a school district boundary.<sup>32</sup> Comparing the first column to the third column of the table, it is immediately obvious that the houses near school district boundaries are not fully representative of those in the Bay Area as a whole. To address this problem, we create sample weights for the houses near the boundary.<sup>33</sup> Column 7 of Table 3 shows the resulting weighted means, showing that using these weights makes the sample near the boundary much more representative of the full sample, column 7 typically being much closer to column 1 than column 3 is.

Comparing differences across school district boundaries, displayed in columns 4 and 5, the average characteristics of houses with 0.25 miles of the boundary on the high school quality versus low school quality side of each boundary reveals that houses on the high side cost \$53

---

<sup>31</sup> A number of empirical issues arise in incorporating boundary fixed effects into our analysis. Concerning the choice of jurisdiction for which the boundaries are defined, we use boundaries between school districts in the Bay Area. A central feature of local governance in California helps to eliminate some of the problems that naturally arise with the use of school district boundaries, as Proposition 13 ensures that the vast majority of school districts within California are subject to a uniform effective property tax rate of one percent. Concerning the width of the boundaries, we experimented with a variety of distances and report the results for 0.25 miles, as these were more precise due to the larger sample size.

<sup>32</sup> In addition, the fourth and fifth columns report means on the high versus low average test score side of the school district boundary; the sixth column reports t-tests for difference in means of fourth and fifth columns; and the seventh column reports weighted means for the sample of houses within 0.25 miles of a school district boundary - the weight is described below.

<sup>33</sup> The following procedure is used: we first regress a dummy variable indicating whether a house is in a boundary region on the vector of housing and neighborhood attributes using a logistic regression. Fitted values from this regression provide an estimate of the likelihood that a house is in the boundary region given its attributes. We use the inverse of this fitted value as a sample weight in subsequent regression analysis conducted on the sample of houses near the boundary.

more per month and are assigned to schools with a 43-point average test score increase.<sup>34</sup> Houses on the high quality side of the boundary are more likely to be inhabited by white households and households with more education and income – this pattern is evident when looking at the difference in means test. These types of across-boundary differences in sociodemographic composition are what one would expect if households sort on the basis of preferences for school quality, thereby leading those with stronger tastes or increased ability to pay for school quality to choose the higher school quality side of the boundary. In the results that we present below, the inclusion of boundary fixed effects has the expected effect, reducing the estimated willingness of households to pay for average neighborhood educational attainment and income.

**Racial Preferences and Discrimination.** The strategy of using boundary fixed effects is designed to deal with the correlation of neighborhood sociodemographic characteristics with any unobserved component of neighborhood quality valued the same by households of all races. It is important to point out, however, that this strategy does not help US distinguish the extent to which these estimated racial interactions result from (i) discrimination in the housing market (e.g., centralized discrimination against recent immigrants from China), (ii) direct preferences for the race of one’s neighbors (e.g., preferences on the part of a recent immigrant from China to live with other Chinese immigrants), and (iii) preferences for race-specific portions of unobserved neighborhood quality (e.g., preferences for Chinese groceries which are located in neighborhoods with a high fraction of Chinese residents). That is, these underlying explanations are indistinguishable from one another because they give rise to predicted residential location decisions that are observationally equivalent in the data.

To see this more formally, consider a model of housing market discrimination in which discrimination drives an wedge between the effective price that blacks versus whites pay for housing in neighborhoods with an high fraction of white households. Such a price wedge might capture increased search costs associated with finding a house in the face of discrimination in real estate brokerage, increased borrowing costs associated with finding a mortgage in the face of discrimination in the mortgage lending market, or the increased sales price required to convince a white owner to sell to a prospective black buyer in the face of discrimination by individual home sellers. Because increases in the effective price that blacks face to purchase houses in white neighborhoods are unobserved in the data, their effect on behavior (presumably causing blacks to be less likely to purchase homes in white neighborhoods) will be interpreted by our model as

---

<sup>34</sup> As described in the Data Appendix, we construct a single price vector for all houses, whether rented or owned.

stronger preferences among blacks to reside with black versus white neighbors. In this way, discrimination of any of these kinds gives rise to behavioral predictions that can be matched identically by a model that only allows for preferences over racial composition.

Our inability to distinguish preferences from discrimination implies that the primary counterfactual simulation that we consider below – where we set the estimated ‘preference’ parameters associated with neighborhood race to zero – essentially sets all factors that affect racial sorting, whether as a result of preferences or discrimination, to zero. In this way, while distinguishing preferences from discrimination is important for interpreting the welfare implications of our results, our inability to distinguish these potential explanations for racial sorting does not seriously affect the key aim of our analysis, which is to gauge the impact of racial sorting as a whole in driving racial differences in the consumption of neighborhood amenities.

## 7 PARAMETER ESTIMATES

Estimation of the full model proceeds in two stages, as noted, the first stage recovering interaction parameters and vector of mean indirect utilities, the second stage returning the components of mean indirect utility. We report the estimates of the interaction parameters in Appendix Table 2. As the table demonstrates, the first stage of the estimation procedure returns 165 parameters on terms that interact individual and household characteristics, permitting great flexibility in preferences across different types of households.<sup>35</sup> In particular, the model includes the following household characteristics: total household income, household income from capital sources (a proxy for wealth), race, education, work status, age, the presence of children, and, importantly, interactions of household income and race. These household characteristics are interacted with many housing and neighborhood attributes including house price, owner-occupancy status,<sup>36</sup> number of rooms, the age of the structure, average test score, elevation, population density, crime and eight variables characterizing the neighborhood sociodemographic composition: the fraction of households of each race, the fraction of households college educated, average neighborhood income, and neighborhood income interacted with race. The model also

---

<sup>35</sup> An extended discussion of this broad set of parameter estimates can be found in Bayer, McMillan, and Rueben (2004b), which formally sets out the model and estimation procedure used here. In this paper, we focus primarily on the parameters associated with race.

<sup>36</sup> We treat ownership status as a fixed feature of a housing unit in the analysis. Thus, whether a household rents or owns is endogenously determined within the model by its house choice. In the model, we allow households to have heterogeneous preferences for home-ownership (a positive interaction between household wealth and ownership, for example, will imply that wealthier households are more likely to own their housing unit, as we find below). A single price index is used for owner- and renter-occupied units - see the Data Appendix for details.

captures the spatial aspect of the housing market by allowing households to have preferences over commuting distance.<sup>37</sup>

This specification is especially flexible from the point of view of the main research question addressed in the paper, in two key ways. First, it includes a full set of race interactions permitting, for example, black households to have different preferences for Asian versus white neighbors. Second, it includes interactions of race and income both as household and neighborhood characteristics, thereby permitting high-income Asian households, for example, to have different preferences than low-income Asian households for neighborhoods and for these preferences to depend on whether a neighborhood has high- versus low-income Asian neighbors.

The numbers in Appendix Table 2 are not directly interpretable in dollar values and so we discuss the results in terms of marginal willingness-to-pay measures (MWTP); the results for the mean household are shown in Table 6 and results related to heterogeneity in MWTP are shown in Table 7. The first three columns of Table 6 report the implied measures of the mean MWTP for housing and neighborhood attributes that result for three specifications of the mean indirect utility regressions. These measures are calculated by dividing the coefficient associated with each choice characteristic in these regressions by the coefficient on price.

Results are reported for the full sample and for a sample of houses within 0.25 miles of school district boundaries, with and without including fixed effects. No clear changes emerge when the sample is reduced to only those houses near a school district boundary. Comparing the coefficients on the neighborhood sociodemographic characteristics with and without the inclusion of boundary fixed effects (columns 2 and 3) yields the pattern of results one would expect if the boundary fixed effects control for unobserved components neighborhood quality unrelated to the sorting of households across the boundary. In particular, controlling for fixed effects increases the coefficient on percent black (reported at the mean average neighborhood income) from -\$285 to -\$234; on percent Hispanic from -\$37 to \$104; and on percent Asian from -\$70 to \$150. Doing so also reduces the coefficient on the percent of households with a college degree from \$186 to \$165 and the coefficient on average neighborhood income (/ \$10,000) from \$89 to \$85 per month. Thus boundary fixed effects seem to be effective in controlling for fixed aspects of unobserved neighborhood quality that are correlated with neighborhood sociodemographics, and thus provide an attractive way of estimating preferences for neighborhood sociodemographic characteristics in the presence of this important endogeneity problem.

---

<sup>37</sup> We treat a household's primary work location as exogenous, calculating the distance from this location to the location of the neighborhood in question. Estimates based on a specification without commuting distance are qualitatively similar.



Table 7 reports the implied estimates of the heterogeneity in MWTP for selected housing and neighborhood characteristics for the specification associated with column (3) in Table 7, which includes boundary fixed effects. This is our preferred specification. The first row of Table 8 repeats the MWTP of the mean household and then reports the MWTP for households with the characteristic listed in the row heading, holding all other characteristics at the mean. The table reveals strong segregating racial interactions, with households of each race preferring to live near others of the same race. Interpreted literally as preferences, black households with income equal to the mean (\$55,000), for example, are willing to pay \$67 per month on average to live in a neighborhood with 10 percent more black versus white households. White households with mean income, on the other hand, are willing to pay \$38 per month on average to live in a neighborhood that is 10 percent more white versus black.<sup>38</sup>

Importantly, the equilibrium predictions of the model concerning racial sorting are driven by the differences in preferences across households of different races (as discussed above, this is in essence what makes it impossible to distinguish preferences from discrimination in observational data). At the mean, the *difference* between what whites and blacks are willing to pay for a 10 percent increase in the fraction of white versus black neighbors is about \$100 per month. Table 8 also shows similar figures calculated for households at a higher income level (income=\$120,000). The difference in the WTP on blacks and whites remain near \$90 per month even at this much higher income level. Thus, strong segregating forces in the housing market are relevant at all income levels for all races.<sup>39</sup>

## 8 EQUILIBRIUM COUNTERFACTUAL SIMULATIONS

We now use the estimated parameters to conduct two equilibrium counterfactual simulations that shed direct evidence on our main conjecture. The first simulation is designed to characterize the housing market equilibrium in a world in which neighborhood racial composition has no effect on residential location decisions. In particular, we consider a counterfactual simulation that eliminates all racial interactions in the location decision – that is, one that sets all of the utility parameters that govern preferences for neighborhood racial characteristics (including interactions of neighborhood race and neighborhood income) to zero. As discussed above, this corresponds to

---

<sup>38</sup> While not the focus of our primary analysis, the estimates of Table 8 also reveal strong segregating preferences for Hispanic and Asian households; those with mean incomes are willing to pay \$98 and \$72 per month respectively to live with others of the same race versus whites.

<sup>39</sup> The strong racial interactions that we estimate are in no way implicitly assumed in writing down the model. As is clear from Table 8, households of every income level prefer to live with higher income neighbors. This makes clear that the model does not in any way force the parameters to yield preferences

calculating a new equilibrium in a world in which no one has direct preferences for the race of one's neighbors and there is no racial discrimination in the housing market.<sup>40</sup>

Having characterized the importance of racial sorting in driving racial differences in the consumption of neighborhood amenities in the San Francisco Bay Area, we then conduct a second simulation designed to study the importance of the size of the black population in the metropolitan area in driving these results. In this second simulation, we leave composition and preference parameters the same and simply change overall size of black and white populations. As the analysis of available neighborhoods in Atlanta, Chicago, and Washington DC presented in Section 2 suggested, a significant number of mixed- and high-SES black neighborhoods are likely to form when proportion of blacks in the metropolitan population increases. In turn, with the expanded availability of these neighborhoods, racial sorting might not have nearly the same negative consequences for the consumption of neighborhood amenities by black households.

**Simulation Details.**<sup>41</sup> The basic structure of solving for a new equilibrium consists of a loop within a loop. The outer loop calculates the sociodemographic composition of each neighborhood, given a set of prices and an initial sociodemographic composition of each neighborhood. The inner loop calculates the unique set of prices that clears the housing market, given an initial sociodemographic composition for each neighborhood. Thus having set the parameters associated with neighborhood race to zero, we first calculate a new set of prices that clears the market. Using these new prices and the initial sociodemographic composition of each neighborhood, we calculate the probability that each household chooses each housing type, and aggregating these choices to the neighborhood level, compute the predicted composition of each neighborhood. We then replace the initial neighborhood sociodemographic measures with these new measures and start the loop again – i.e., calculate a new set of market clearing prices with

---

for live with others like oneself, as both high- and low-income households are willing to pay substantially more for higher income neighbors.

<sup>40</sup> It is important to point out that this simulation does not set the coefficients multiplying the interaction of ownership status and household race to zero. To the extent that discrimination in the mortgage market is responsible for the significant negative coefficient on the interaction of black and ownership status, this form of discrimination remains in the new counterfactual equilibrium.

<sup>41</sup> As discussed in Section 5, uniqueness is not a generic property of our sorting model due to the presence of social interactions in the sorting process. Because our first simulation removes the central component of these interactions, sorting along racial lines, the potential of multiple equilibria in this simulation is less of an issue. For our second simulation, which leaves social interactions in the sorting process unchanged, the potential for multiple equilibria certainly exists. The counterfactual equilibrium reported in the paper is calculated by starting from the observed equilibrium and iterating to a new equilibrium following the procedure outlined in this sub-section. While we do not have a proof that this is the only sorting equilibrium in this counterfactual environment, we expect that the general pattern of results would obtain in any such counterfactual equilibrium – given the increased size of the black population.

these updated neighborhood sociodemographic measures. We continue this process until the neighborhood sociodemographic measures converge. The set of household location decisions corresponding to these new measures along with the vector of market clearing housing prices describe the new equilibrium.

Because some neighborhood amenities, such as crime rates and school quality, depend in part on the sociodemographic composition of the neighborhood, it is natural to expect these neighborhood characteristics to adjust as part of the movement to a new sorting equilibrium.<sup>42</sup> Accounting for the impact of neighborhood sociodemographic characteristics on crime rates and test scores is a challenging exercise, as selection problems abound. For example, an OLS regression of crime rates on neighborhood sociodemographic characteristics almost certainly overstates the role of these characteristics in producing crime as it ignores the fact that households sort non-randomly across neighborhoods.

In the light of these difficulties, we adopt an approach that seeks to provide simple bounds for the characteristics of the new equilibrium that results for each of our simulations. For one bound, we calculate a new equilibrium without allowing crime rates and average test scores in each neighborhood to adjust. For the other bound, we calculate a new equilibrium, adjusting crime rates and average test scores in each neighborhood according to the adjustments implied by an OLS regression of the crime rate and average test score on neighborhood sociodemographic composition. The first bound will tend to understate the impact of sociodemographic shifts on the implied crime rate and average test score in each neighborhood, while the second bound will tend to overstate the impact of these sociodemographic shifts. As the results below indicate, these bounds provide a tight range for the predictions from our simulations.

**Eliminating ‘Preferences’ for Neighborhood Racial Composition.** Table 8 reports the results of our simulation, characterizing the way average housing and neighborhood consumption measures change for households in various race and race-income categories. In each case, we report three measures: a pre-simulation measure as well as post-simulation measures from simulations that do and do not adjust school quality and crime as neighborhood sociodemographic characteristics change, respectively.

The first two columns indicate the racial composition exposure rate measures that arise with the elimination of racial interactions. Not surprisingly, the elimination of racial interactions has

---

<sup>42</sup> Such adjustments may arise due to effects that operate through the political system, as in Tiebout (1956), or as the result of productive externalities. The former effects are likely to be limited in our analysis due to nature of the provision of public goods in California, which gives local governments almost no control over taxes or the level of spending.

an enormous effect in reducing segregation. Black exposure to other blacks falls from over 38 percent to around 14.5 percent, and black exposure to whites jumps from 37.7 percent to around 64.5 percent. (Recall that around 67 percent of the households in our study area are white and just under 9 percent are black.) Similarly, the average exposure of white households to blacks increases from just over 4 percent to 7 percent. Thus removing race as a factor in the location decision almost completely eliminates segregation except for a small portion that arises because black and white households still differ markedly along socioeconomic dimensions including education, income, and wealth.

The elimination of race as a factor in location decisions also has important consequences for the consumption of neighborhood amenities by households of each race. The remaining columns of Table 8 report a number of such measures including the home-ownership rate, average monthly house price, average commuting distance, and the average consumption of house size, school quality, crime, neighborhood income and education.<sup>43</sup> The first two main rows of the table show the overall impact of eliminating race-based sorting on black-white consumption differences. The largest reductions in black-white consumption gaps occur for public school quality and crime for which the gap is reduced by 55-65 percent. The overall black-white consumption gap for neighborhood income and education also declines substantially, on the order of 45-50 percent. Again, the ranges for these estimates reflect the results of two simulations that differ in the manner school quality and crime are adjusted with the changing neighborhood sociodemographic composition. Importantly, these substantial reductions in racial differences in consumption come about simply by eliminating racial interactions in the housing market - that is, without changing household income, wealth, education or other household characteristics.

The remaining four main rows of the table show results separately for the highest and lowest income quartile for both races. Focusing on the results for households in the highest income quartile, the numbers reveal that black households in the top income quartile experience increased consumption of every type of neighborhood and housing amenity, including house size and home ownership as a result of eliminating the role of race in the housing market. Moreover, especially large reductions in black-white consumption gaps occur for public safety, 70 percent, and school quality, 60 percent. Blacks in the lowest income quartile also experience increased consumption of each local public good, but actually experience a decline in housing consumption.

These results indicate that race plays a profound role in shaping the equilibrium matching of households to neighborhoods in the urban housing market. As the consumption patterns of Tables

5 have already suggested, because blacks make up only about 8 percent of the population of the Bay Area, consumption decisions regarding neighborhood race and other neighborhood characteristics are not separable; increases in the consumption of local public goods typically mean a decline in the fraction of blacks in a neighborhood. This affects the implicit price that blacks versus whites pay for local public goods, thereby accentuating racial differences in consumption.

An important aspect of the simulation results presented in Table 8 is that blacks at all income levels also spend a considerable amount more on housing in the new equilibrium in which sorting for race-related reasons has been eliminated. This suggests that the observed lower levels of neighborhood amenity consumption by blacks in the actual Bay Area equilibrium are at least partially compensated with lower housing prices. However, given the difficulties in obtaining alternatives for many of these amenities (e.g., highly educated neighbors) in the private market, racial sorting in the housing market almost certainly works to strengthen the persistence of intergenerational racial differences in educational attainment, income, and wealth by accentuating differences in the consumption of important neighborhood amenities that contribute to these life outcomes through the formation of human capital and social networks.<sup>44</sup>

**Equalizing the Number of Blacks and Whites.** Table 9 reports the results of our second simulation, which equalizes the size of the white and black populations in the Bay Area. In so doing, we leave the distributions of socioeconomic characteristics for households of each race and the estimated ‘preference’ estimates unchanged.

In interpreting the results of these simulations, two aspects of this analysis are important to point out. First, by expanding the number of blacks in the population relative to whites while maintaining the socioeconomic distributions for each race, the counterfactual environment is characterized by a sizeable decrease in average education, income, and wealth of the population as a whole. This implies that the average levels of neighborhood income and education reported in Table 9 will necessarily be smaller for the metropolitan area as a whole. Moreover, rather than trying to predict the productive impact of these changes on overall average levels of school

---

<sup>43</sup> We also note that the elimination of racial interactions leads to an overall reduction in commuting distances for both blacks and whites; without needing to adjust their location decisions for race-related reasons, households are able to more easily find suitable locations in other dimensions.

<sup>44</sup> The possibility remains that black households may use the implicit compensation that comes in the form of lower house prices to offset the lower levels of public good consumption with increased private consumption that improves the human capital formation of their children.

quality and crime, we focus attention on only the version of our counterfactual simulations that leave school quality and crime in each neighborhood unchanged.<sup>45</sup>

A second implication of the counterfactual change is that, using the same cut-offs for the income quartiles as in the observed equilibrium, the fraction of households in the original upper income quartile range now is substantially smaller than 25 percent (conversely, the fraction in the lower quartile is now greater than 25 percent) as a result of the overall decline in average socioeconomic characteristics in the metropolitan area. In Table 9, we hold the income quartile cut-offs the same as in the original equilibrium; this implies that households *in each original income quartile range for each race* are better off relative to the average household in the metropolitan area than they were in the original equilibrium. Because most housing and neighborhood amenities are assumed to be in fixed supply, it should be expected that when broken out by these original income quartile ranges, the consumption of most housing and neighborhood amenities increases in each quartile. Moreover, aggregating up, an increase in average housing and neighborhood amenities (other than average neighborhood income and education) should also be expected for both whites and blacks as a whole. While this may seem counterintuitive, it comes about because of the changing relative sizes of the two racial groups in the population.

Turning now to the results themselves, while certainly not identical to the results shown in Table 8, the most striking aspect of the results is their similarity to the previous simulation. Focusing on the implied racial differences in crime, school quality, average education and income, the declines in the white-black gap are roughly the same magnitude – on the order of 50 percent reduction in the difference – as those reported in Table 8.<sup>46</sup>

The most notable difference in the results shown in Table 8 versus Table 9, relates to racial segregation. In Table 8, the reductions in racial differences in housing and neighborhood consumption were directly attributable to the elimination of racial sorting, the resulting limited segregation giving rise to better neighborhood amenities in other dimensions for blacks relative to whites. In Table 9, the increasing number of blacks at all points in the income distribution gives rise to *increases* in own-race exposure. In this case, the reductions in racial differences in the consumption of housing and neighborhood amenities are attributed to formation of mixed- and high-SES black neighborhoods. The presence of these neighborhoods increases segregation in the

---

<sup>45</sup> As in Table 8, versions of the simulations that allow neighborhood school quality and crime rates to vary with changing neighborhood sociodemographic composition led to very similar results.

<sup>46</sup> Again for neighborhood income and education, the magnitudes drop for all individuals because counterfactual increases number of lower SES households – the comparison described in the text relates to the difference in average consumption between blacks and whites.

metropolitan area (measured as the difference between own-race exposure and the proportion in population) but also breaks the implicit bundling of neighborhood race and neighborhood amenities in other dimensions. As can be seen in Table 9, this allows high-income blacks to consume relatively high levels of neighborhood and housing amenities while living in neighborhoods with a relatively high fraction of black household on average.

## 9 ACROSS-METRO EVIDENCE

The results of our second simulation show that the adverse effects of racial sorting for blacks are closely tied to the fraction of blacks in the metropolitan population. As this fraction increases, racial sorting continues to give rise to highly segregated neighborhoods, yet the consumption of neighborhood amenities by blacks relative to whites increases significantly.

This simulation also demonstrates that the level of racial segregation in a metropolitan area is not a sufficient statistic for characterizing the impact of racial sorting on the relative education and labor market outcomes of blacks versus whites. In particular, according to some conventional measures, segregation increases in our second counterfactual simulation while the relative consumption of neighborhood amenities improves dramatically for blacks relative to whites. This suggests that the adverse effects of racial sorting (and the resulting segregation) for blacks are strongest precisely when there are very few highly-educated blacks in population.

To explore this possibility directly, we examine a set of alternative specifications of regressions appearing in Cutler and Glaeser (1997, CG). Specifically, using the 1990 Census, CG run 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. The 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. 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.

To explore the possibility that segregation is most harmful to blacks relative to whites when the fraction of highly educated blacks in the population is small, we augment CG's main specification to include a measure of the fraction of college-educated blacks in the metropolitan population in the level. This measure is included in the level, interacted with the black indicator variable, interacted with metropolitan segregation, and interacted with metro segregation and the black indicator variable. Thus we allow black-white differences (captured by the coefficients on

the interactions with whether the individual is black) to vary directly with the fraction of college-educated blacks in the population and indirectly through the effect of segregation.

The upper-panel of Table 10 replicates the OLS coefficient on the interaction between CG's metropolitan dissimilarity measure and whether an individual is black for age groups 20-24 and 25-30. While not identical, our analysis replicates CG's very closely. The lower panel of Table 10 reports a wider set of coefficients with the inclusion of additional interaction terms associated with the proportion of college-educated blacks in the population. Our coefficient of interest is on the interaction term (black)\*(segregation)\*(proportion metro area - black and college-educated). This coefficient reveals how the adverse effects of segregation for blacks relative to whites changes with the increased share of college-educated blacks in the population.

As can be seen in the table, the negative consequences of segregation for blacks reported in CG's analysis fall significantly with the fraction of college-educated blacks in the metropolitan area. For the high school graduation rates of 20-24 year olds, for example, the adverse effects of segregation for blacks are eliminated when the fraction of college-educated blacks in the population reaches 3 percent. This fraction is about twice the average fraction of college-educated blacks in U.S. metropolitan areas (1.7 percent). More generally, the effects of segregation for blacks aged 20-24 relative to whites for the other labor market, education, and fertility outcomes falls to zero when the fraction of college-educated blacks in the population reaches 4-6 percent.

Collectively, these across-metro regressions provide evidence that is complementary to our primary analysis. These regressions not only emphasizing the role of small numbers in driving the adverse effects of racial sorting for blacks relative to whites in U.S. metropolitan areas, but also provide a link between the study of their impact on the consumption of neighborhood amenities and a series of educational, labor market, and fertility outcomes.

## **10 CONCLUSION**

The fact that race is strongly correlated with socioeconomic characteristics in the U.S. population provides an obvious first-order explanation for the fact that blacks live in poorer and lower-amenity neighborhoods than whites. In this paper, we conjecture that another explanation may be as important. In particular, we have argued that sorting on the basis of race itself significantly exacerbates racial differences in the consumption of neighborhood amenities.

We motivated this conjecture by first providing direct evidence that predominantly black, high-SES neighborhoods are in short supply in almost all U.S. metropolitan areas. As a consequence, neighborhood race and neighborhood are explicitly linked in the set of residential options available to most households: in order to choose high-amenity neighborhoods, households



must typically live with a higher fraction of white neighbors. Given any form of segregating preferences, this bundling drives a wedge between the *implicit price* that whites versus blacks must pay in order to consume higher levels of a given neighborhood amenity. As households respond to these implicit prices, the resulting consumption of neighborhood amenities by blacks would be lower than that of otherwise-identical whites.

Using a flexible equilibrium model of residential sorting estimated with detailed data on the location decisions of nearly a quarter of a million households, we conduct two complementary simulations designed to provide evidence on our conjecture. In the first, we compare the actual state of the world to one in which race is eliminated as a factor in each household's residential location decision. The resulting halving of the racial differences in neighborhood amenity consumption (school quality, crime, average neighborhood education and income) implies that racial sorting is just as important as the underlying socioeconomic differences between whites and blacks in driving differences in the consumption of neighborhood amenities. In the second simulation, we equalize (counterfactually) the proportions of blacks and whites in the metropolitan area, thereby relieving the existing bundling constraint, while leaving all preference parameters at their estimated values. In this case, the observed white-black gaps in the consumption of neighborhood amenities decline by nearly the same amount as in our original simulation - the adverse effects of racial sorting for blacks in the Bay Area in terms of neighborhood amenities would be virtually eliminated were blacks to constitute a larger proportion of the metropolitan area's population. This result is also supported by additional evidence based on across-MSA comparisons that implies that the adverse effect of segregation for blacks versus whites falls to zero as the fraction of college-educated blacks in the population reaches 3 to 6 percent.

Overall, our empirical results draw attention to an important consequence of racial sorting - the significant distortion affecting the neighborhood *quality* choice - and provide a clear explanation as to why this distortion arises in the first place, due to a small numbers problem. The substantial reductions in neighborhood quality for blacks that we find are likely to have a significant impact on the inter-generational persistence of racial differences in education, income, and wealth. Our analysis also informs a much larger body of research that examines the impact of segregation on individual outcomes, drawing attention to the fact that segregation *measured at either the neighborhood or metropolitan level* does not provide a sufficient statistic for estimating the impact of racial sorting.

## References

- Anas, Alex (1982), *Residential Location Markets and Urban Transportation: Economic Theory, Econometrics and Public Policy Analysis*, Academic Press, New York.
- Anas, Alex (2002), "Prejudice, Exclusion, and Compensating Transfers: The Economics of Ethnic Segregation," *Journal of Urban Economics*, 52: 409-32.
- Anas, Alex and Ikki Kim (1996), "General Equilibrium Models of Polycentric Urban Land Use with Endogenous Congestion and Job Agglomeration," *Journal of Urban Economics*, 40: 232-56.
- Bajari, Patrick, and Matthew Kahn (2001), "Why Do Blacks Live in Cities and Whites Live in Suburbs?" unpublished manuscript, Stanford University.
- Bayer, Patrick, Hanming Fang, and Robert McMillan (2005), "Separate When Equal? Racial Inequality and Segregation," NBER Working Paper 11507, August 2005.
- Bayer, Patrick, Fernando Ferreira, and Robert McMillan (2003), "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," Economic Growth Center, Yale University Working Paper No. 872.
- Bayer, Patrick, Robert McMillan, and Kim Rueben (2004), "What Drives Racial Segregation? New Evidence Using Census Microdata," *Journal of Urban Economics*, 56(3), 2004: 514-535.
- Bayer, Patrick, Robert McMillan, and Kim Rueben (2004b), "An Equilibrium Model of Sorting in an Urban Housing Market," NBER Working Paper 10865.
- Benabou, Roland (1993), "The Workings of a City: Location, Education, and Production," *Quarterly Journal of Economics*, 108(3), pp.619-652.
- Benabou, Roland (1996), "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review*, Vol. 86, No. 3., pp. 584-609.
- Berry, Steven (1994), "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, Vol. 25, pp. 242-262.
- Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, Vol 63, pp. 841-890.
- Berry, Steven, Oliver Linton, and Ariel Pakes (forthcoming), "Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems," *Review of Economics Studies*.
- Black, Sandra (1999) "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics*, May 1999.
- Brock, William A., and Durlauf, Steven N. (2001) "Discrete Choice with Social Interactions." *Rev. Econ. Stud.* 68: 235-60.
- Chambers, D.N. (1992), "The Racial Housing Price Differential and Racially Transitional Neighborhood," *Journal of Urban Economics*, 32: 214-232.
- Cutler, David and Edward Glaeser (1997), "Are Ghettos Good or Bad?" *Quarterly Journal of Economics*, August: 826-72.

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

Epple, Dennis (1987), "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products," *Journal of Political Economy*, 107: 645-81.

Epple, D., R. Filimon, and T. Romer (1984), "Equilibrium Among Local Jurisdictions: Towards an Integrated Approach of Voting and Residential Choice," *Journal of Public Economics*, Vol. 24, pp. 281-304.

Epple, D., R. Filimon, and T. Romer (1993), "Existence of Voting and Housing Equilibrium in a System of Communities with Property Taxes," *Regional Science and Urban Economics*, Vol. 23, pp. 585-610.

Epple, Dennis and Holger Sieg (1999), "Estimating Equilibrium Models of Local Jurisdictions," *Journal of Political Economy*, Vol. 107, No. 4., pp. 645-681.

Epple, Dennis, Thomas Romer and Holger Sieg (2001), "Interjurisdictional Sorting and Majority Rule: An Empirical Analysis," *Econometrica*, Vol. 69, No. 6., pp. 1437-1455.

Evans, W.N., W.E. Oates, R.M. Schwab (1992), "Measuring Peer Group Effects: A Study of Teenage Behavior," *Journal of Political Economy*, 966-991.

Fernandez, Raquel and Richard Rogerson (1996), "Income Distribution, Communities, and the Quality of Public Education." *Quarterly Journal of Economics*, Vol. 111, No. 1., pp. 135-164.

Fernandez, Raquel and Richard Rogerson (1998), "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education Finance Reform," *American Economic Review*, Vol. 88, No. 4., pp. 813-33.

Gabriel, S. and S. Rosenthal (1989), "Household Location and Race: Estimates of a Multinomial Logit Model," *Review of Economics and Statistics*, 71: 240-9.

Harsman, Bjorn and John Quigley (1995), "The Spatial Segregation of Ethnic and Demographic Groups: Comparative Evidence from Stockholm and San Francisco," *Journal of Urban Economics*, 37: 1-16.

Ihlanfeldt, Keith (1992), *Job Accessibility and the Employment and School Enrollment of Teenagers*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Ihlanfeldt, K. and David Sjoquist (1990), "Job accessibility and racial differences in youth employment rates", *American Economic Review*, 80, 267-276.

Ihlanfeldt, Keith and David Sjoquist (2000). "The spatial mismatch hypothesis: A review of recent studies and their implications for welfare reform," *Housing Policy Debate*, 9, 4.

Ihlanfeldt, Keith and Ben Scafidi (forthcoming) "Whites' Neighborhood Racial Preferences and Neighborhood Racial Composition: Evidence from the Multi-City Study of Urban Inequality," *Housing Studies*.

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(2002): 366-390.

Kain, John 1968. Housing segregation, negro employment, and metropolitan decentralization. *Quarterly Journal of Economics*, 82, 175-197

Loury, Glenn C. (1977), "A Dynamic Theory of Racial Income Differences," Chapter 8 in *Women, Minorities and Employment Discrimination*, Lexington Books, Lexington, MA: 153-188.

- Massey, Douglas, and Nancy Denton (1993), *American Apartheid: Segregation and the Making of the Underclass*, Harvard University Press, Cambridge MA.
- McFadden, Daniel (1973), "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, eds., *Frontiers of Econometrics*, Academic Press, New York.
- McFadden, Daniel (1978), "Modeling the Choice of Residential Location," in eds. Karlquist, A., et al., *Spatial Interaction Theory and Planning Models*, Elsevier North-Holland, New York.
- 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.
- Nechyba, Thomas J. (1999), "School Finance Induced Migration and Stratification Patterns: the Impact of Private School Vouchers," *Journal of Public Economic Theory*, Vol. 1.
- Nechyba, Thomas J. (2000), "Mobility, Targeting, and Private School Vouchers," *American Economic Review*, Vol. 90(1): 130-46.
- Nechyba, Thomas J., and Robert P. Strauss (1998), "Community Choice and Local Public Services: A Discrete Choice Approach," *Regional Science and Urban Economics*, Vol. 28, 51-73.
- O'Regan, Kathy and John Quigley (1998) "Teenage employment and the spatial isolation of minority and poverty households," *Urban Studies*, 35, 1187-1205.
- Quigley, John M. (1985), "Consumer Choice of Dwelling, Neighborhood, and Public Services," *Regional Science and Urban Economics*, Vol. 15(1).
- 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.
- Ross, Stephen L. and Yves Zenou. (2004) "Effort, Location, and Urban Unemployment," unpublished manuscript, University of Connecticut.
- Schelling, Thomas C. (1969), "Models of Segregation." *American Economic Review*, 59(2): 488-93.
- Schelling, Thomas C. (1971), "Dynamic Models of Segregation," *Journal of Mathematical Sociology*, 1: 143-186 .
- Sieg, Holger, V. Kerry Smith, H. Spencer Banzaf and Randall Walsh (forthcoming) "Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods," *International Economic Review*.
- Tiebout, Charles M. (1956), "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 64: 416-424.
- Weinberg, Bruce (2000) "Black Residential Centralization and the Spatial Mismatch Hypothesis," *Journal of Urban Economics*, 48(1): 110-34.
- Weinberg, Bruce (2004) "Testing the Spatial Mismatch Hypothesis Using Inter-city Variations in Industrial Composition," *Regional Science and Urban Economics*, 34(5): 505-32.
- Wilson, William Julius (1987), *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*, University of Chicago Press: Chicago.

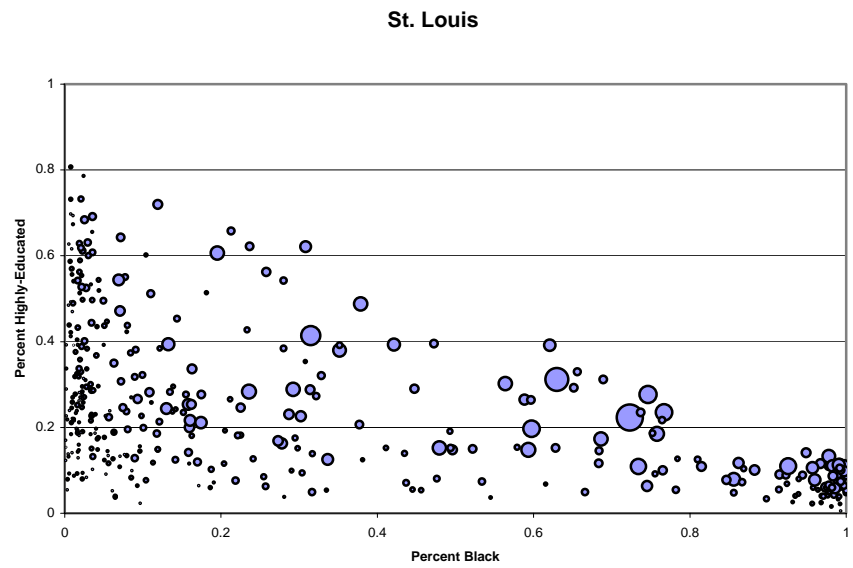
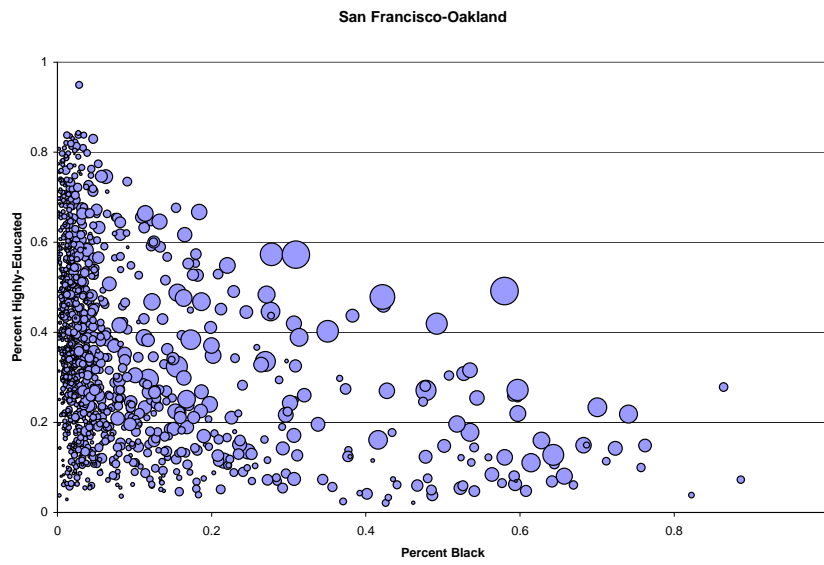
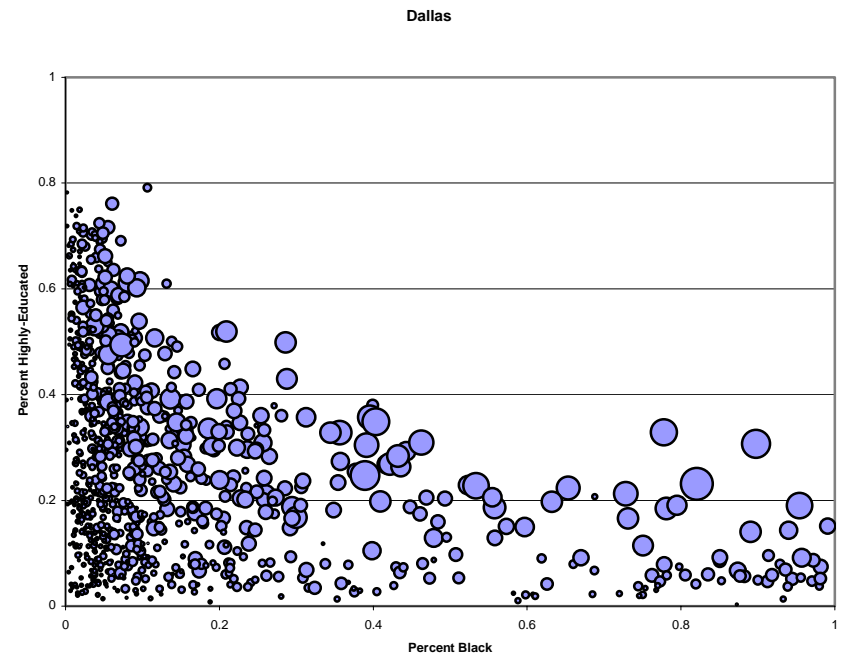
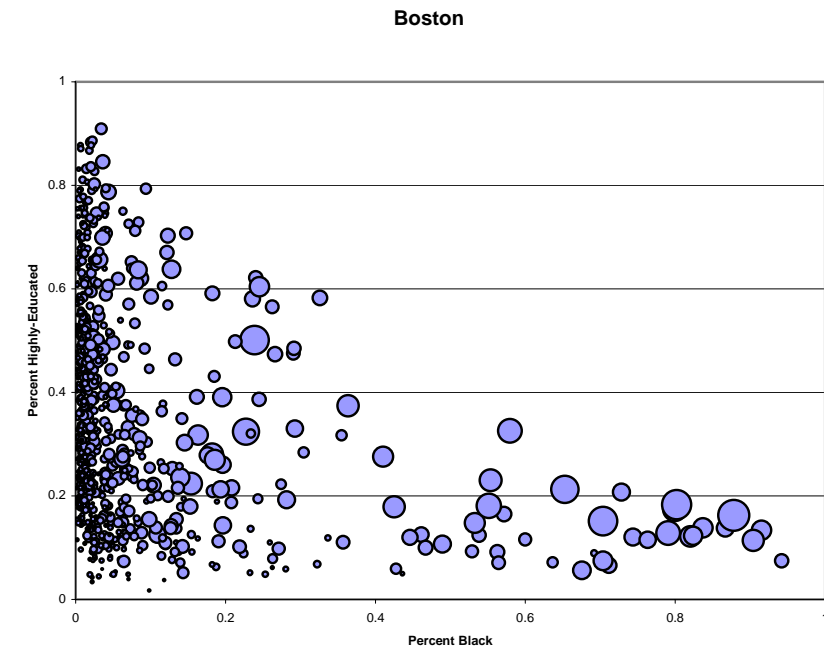


Figure 1: Neighborhood Choice Sets in Boston, Dallas, San Francisco, and St. Louis

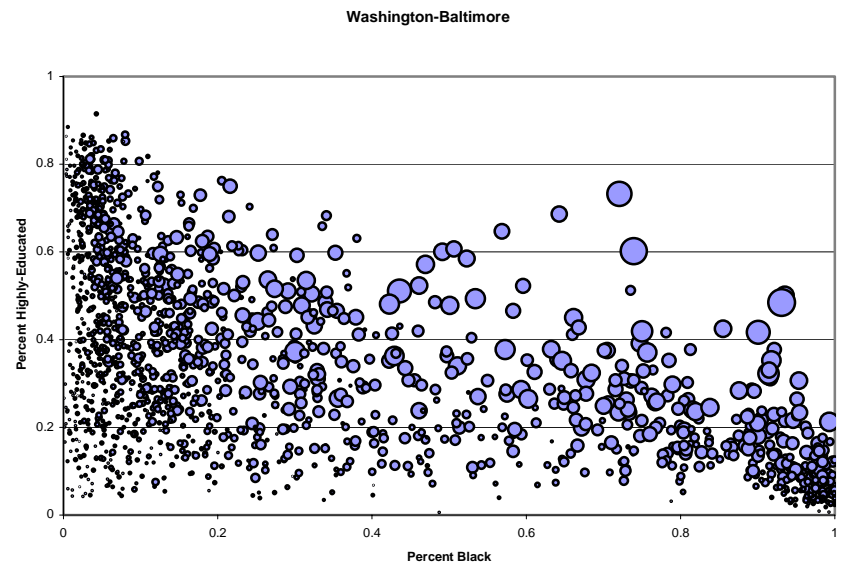
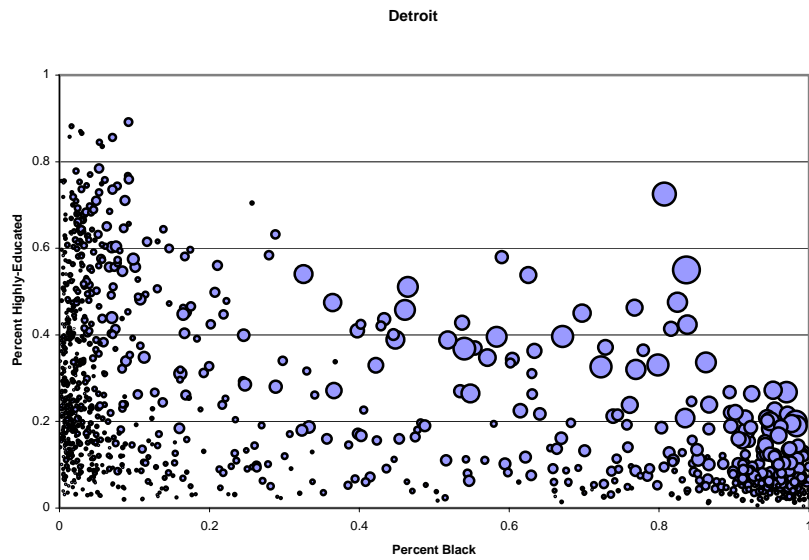
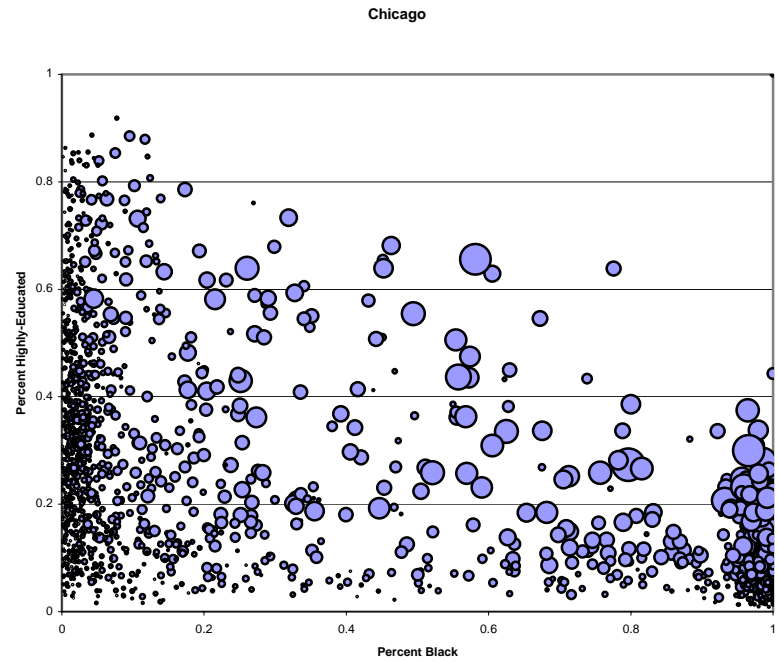
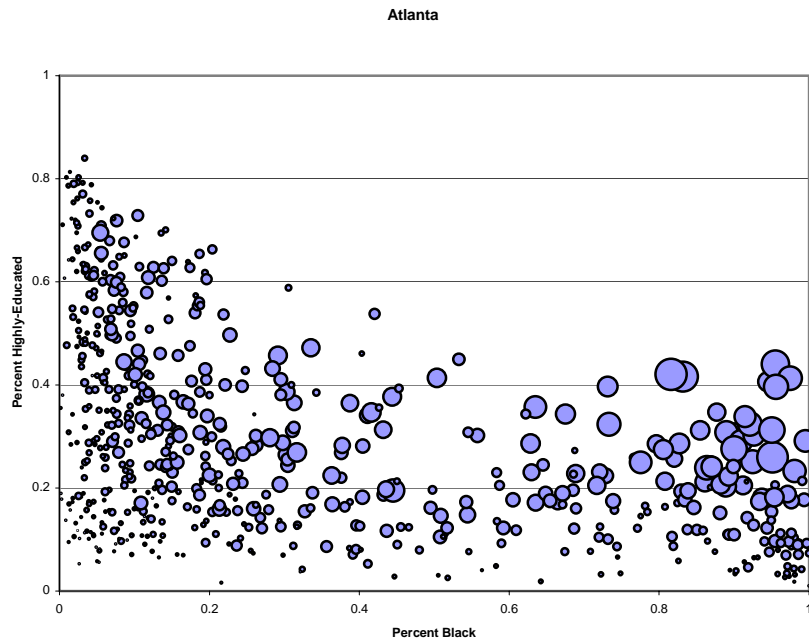


Figure 2: Neighborhood Choice Sets in Atlanta, Chicago, Detroit and Washington DC-Baltimore.

**Table 1: Number of Tracts in United States in 2000 by Race and Education**

<b>All Tracts</b>	<b>Percent College Degree or More</b>			
	<i>at least</i>			
	<i>0%</i>	<i>20%</i>	<i>40%</i>	<i>60%</i>
Number	49,021	26,351	11,094	3,005
Fraction of tracts at least 0% black	100.0%	53.8%	22.6%	6.1%
<b>Percent Black</b>	<i>0%</i>	<i>20%</i>	<i>40%</i>	<i>60%</i>
<i>at least 20%</i>				
Number	9,149	2,567	641	59
Fraction of tracts at least 20% black	100.0%	28.1%	7.0%	0.6%
<i>at least 40%</i>				
Number	5,657	1,164	142	14
Fraction of tracts at least 40% black	100.0%	20.6%	2.5%	0.2%
<i>at least 60%</i>				
Number	3,921	623	44	5
Fraction of tracts at least 60% black	100.0%	15.9%	1.1%	0.1%
<i>at least 80%</i>				
Number	2,559	271	21	1
Fraction of tracts at least 80% black	100.0%	10.6%	0.8%	0.0%
<b>Percent White</b>	<i>0%</i>	<i>20%</i>	<i>40%</i>	<i>60%</i>
<i>at least 20%</i>				
Number	43,179	25,178	11,041	2,999
Fraction of tracts at least 20% black	100.0%	58.3%	25.6%	6.9%
<i>at least 40%</i>				
Number	39,602	24,566	10,839	2,967
Fraction of tracts at least 40% black	100.0%	62.0%	27.4%	7.5%
<i>at least 60%</i>				
Number	35,154	22,543	10,214	2,870
Fraction of tracts at least 60% black	100.0%	64.1%	29.1%	8.2%
<i>at least 80%</i>				
Number	26,910	17,539	8,102	2,339
Fraction of tracts at least 80% black	100.0%	65.2%	30.1%	8.7%

*Note:* Tracts considered have a minimum of 800 households (the average tract in the US has almost 3,000 households)

**Table 2: Neighborhood Patterns for College-Educated Households in the United States**

**Panel A: Neighborhood Patterns for College-Educated Black Households**

*Households first ranked by percent black in Census tract within its MSA*

*Measures reported by household's corresponding quintile within its MSA*

<b>Quintile</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>Total</b>
Percent Black	5.7	14.4	28.3	54.6	78.9	32.0
Percent Highly Educated	38.0	31.6	26.2	18.4	13.8	27.2

**Panel B: Neighborhood Patterns for College-Educated White Households**

*Households first ranked by percent white in Census tract within its MSA*

*Measures reported by household's corresponding quintile within its MSA*

<b>Quintile</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>Total</b>
Percent White	55.0	77.9	86.6	90.4	94.5	77.4
Percent Highly Educated	27.0	36.2	40.7	39.3	39.2	35.3

*Note:* The panels of the table summarize the average distribution of neighborhoods in which college-educated blacks and whites in US metro areas reside, respectively. To construct the numbers in the upper panel, college-educated blacks in each metro area are ranked by the fraction of black households 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. The lower panel reports analogous figures for college-educated whites, first ranking by their tract-level exposure to whites within each MSA.



**Table 3. Overall Sample and Sub-Sample Near School District Boundaries**

Sample Boundary/Weights Observations	full sample		within 0.25 miles of boundaries				weighted sample
	242,100		actual sample	high test score side*	low test score side*	t-test for difference in means (4) versus (5)	
	(1) Mean	(2) S.D.	(3) Mean	(4) Mean	(5) Mean		(6) Mean
<b><u>Housing/Neighborhood Characteristics</u></b>							
monthly house price	1,087	755	1,130	1,158	1,105	5.71	1,098
average test score	527	74	536	558	515	50.96	529
1 if unit owned	0.597	0.491	0.629	0.632	0.626	1.04	0.616
number of rooms	5.114	1.992	5.170	5.207	5.134	3.13	5.180
1 if built in 1980s	0.143	0.350	0.108	0.118	0.099	5.09	0.148
1 if built in 1960s or 1970s	0.391	0.488	0.424	0.412	0.437	4.22	0.406
elevation	210	179	193	194	192	1.14	212
population density	0.434	0.497	0.352	0.349	0.355	2.08	0.374
crime index	8.184	10.777	6.100	6.000	6.192	2.36	7.000
% Census block group white	0.681	0.232	0.704	0.712	0.686	9.62	0.676
% Census block group black	0.081	0.159	0.071	0.065	0.076	6.21	0.080
% Census block group Hispanic	0.110	0.114	0.113	0.107	0.119	8.62	0.117
% Census block group Asian	0.122	0.120	0.112	0.110	0.113	2.50	0.121
% block group college degree or more	0.438	0.196	0.457	0.463	0.451	5.14	0.433
average block group income	54,744	26,075	57,039	58,771	55,457	10.23	55,262
<b><u>Household Characteristics</u></b>							
household income	54,103	50,719	56,663	58,041	55,405	4.20	55,498
1 if children under 18 in household	0.333	0.471	0.324	0.322	0.325	0.54	0.336
1 if black	0.076	0.264	0.066	0.062	0.070	2.69	0.076
1 if Hispanic	0.109	0.312	0.111	0.102	0.119	4.54	0.115
1 if Asian	0.124	0.329	0.112	0.114	0.110	1.06	0.121
1 if white	0.686	0.464	0.706	0.717	0.696	3.86	0.682
1 if college degree or more	0.438	0.497	0.460	0.467	0.454	2.64	0.441
age (years)	47.607	16.619	47.890	48.104	47.699	1.99	47.660
1 if working	0.698	0.459	0.705	0.702	0.709	1.28	0.701
distance to work (miles)	8.843	8.597	8.450	8.412	8.492	0.82	8.490

*Notes:* Columns 1 and 2 report the mean and standard deviation for key variables for the full sample. Column 3 reports means for the sample of houses within 0.25 miles of a school district boundary. Columns 4 and 5 report means on the high versus low test score side of boundaries. Column 6 provides a t-statistic for a test of whether the means reported in columns 4 and 5 are equal. Column 7 reports weighted means for the sample of houses within 0.25 miles of a school district boundary. Weights are constructed so as to make the boundary sample more representative of the full sample and are described in the main text. In constructing columns 4 and 5, we assign each house in the full sample to the nearest school district boundary, noting whether its local school has a higher test score than the school associated with the closest Census block on the other side of the boundary.

**Table 4: Black-White Segregation Patterns for the San Francisco Bay Area**

**Panel A: Average Neighborhood Compositions**

Race Income Quartile	Black					White				
	q1	q2	q3	q4	Total	q1	q2	q3	q4	Total
<i>Household Type</i>										
All Households	0.034	0.020	0.013	0.008	0.076	0.150	0.167	0.176	0.194	0.686
Black	0.209	0.090	0.055	0.025	0.379	0.113	0.113	0.107	0.090	0.424
White	0.019	0.014	0.009	0.006	0.048	0.158	0.181	0.195	0.228	0.761
Black - Lowest Income Quartile	0.281	0.104	0.059	0.023	0.467	0.115	0.084	0.062	0.040	0.302
Black - Highest Income Quartile	0.098	0.060	0.046	0.035	0.240	0.101	0.121	0.141	0.153	0.516
White - Lowest Income Quartile	0.026	0.017	0.010	0.005	0.058	0.203	0.191	0.170	0.154	0.718
White - Highest Income Quartile	0.007	0.007	0.006	0.006	0.026	0.120	0.158	0.202	0.328	0.807

**Panel B: Rate of Exposure Relative to Average Household in Bay Area**

Race Income Quartile	Black					White				
	q1	q2	q3	q4	Total	q1	q2	q3	q4	Total
<i>Household Type</i>										
Black	6.1	4.5	4.1	3.1	5.0	0.8	0.7	0.6	0.5	0.6
White	0.6	0.7	0.7	0.7	0.6	1.0	1.1	1.1	1.2	1.1
Black - Lowest Income Quartile	8.2	5.3	4.4	2.9	6.2	0.8	0.5	0.4	0.2	0.4
Black - Highest Income Quartile	2.9	3.0	3.4	4.4	3.2	0.7	0.7	0.8	0.8	0.8
White - Lowest Income Quartile	0.8	0.8	0.8	0.7	0.8	1.3	1.1	1.0	0.8	1.0
White - Highest Income Quartile	0.2	0.3	0.5	0.8	0.3	0.8	0.9	1.2	1.7	1.2

*Note:* Each entry in Panel A shows the average neighborhood exposure of households of the race or race-income category shown in the row heading to the race or race-income category shown in the column heading. The first entry in the table implies, for example, that for the Bay Area population as a whole an average of 3.4 percent of a household's neighbors are black and in in the lowest income quartile. Each entry in Panel B shows the rate of 'over-exposure' of households of the race or race-income category shown in the row heading to the race or race-income category shown in the column heading.

**Table 5: Housing and Neighborhood Consumption Patterns for Households in Highest Income Quartile****Panel A: Housing and Neighborhood Consumption of Black Households in Highest Income Quartile***Households ranked by % Black in neighborhood - consumption measures summarized by quintile*

<b>Quintile</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>Total</b>
Percent Neighborhood - Black	0-4	4-8	8-20	20-54	54-100	24
Average Test Score	559	528	508	459	418	494
Crime Rate	4	6	8	17	22	11
Average Neighborhood Income	71,150	57,450	55,200	47,300	36,650	53,550
Percent Neighborhood - College Degree of more	54	46	43	41	25	42
House Price	1,584	1,320	1,245	1,163	814	1,225
Number of Rooms	6.8	6.2	6.3	6.1	6.2	6.3
Home Ownership	80	71	76	77	87	78
<i>Average Income of Black Households in Top Income Quartile</i>	107,750	98,150	95,300	96,350	91,250	97,760

**Panel B: Housing and Neighborhood Consumption of White Households in Highest Income Quartile***Households ranked by % White in neighborhood - consumption measures summarized by quintile*

<b>Quintile</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>Total</b>
Percent Neighborhood - White	0-71	71-82	82-87	87-92	92-100	81
Average Test Score	518	548	577	594	602	568
Crime Rate	8	5	4	3	3	4
Average Neighborhood Income	56,250	64,750	76,300	81,150	95,000	74,690
Percent Neighborhood - College Degree of more	45	51	57	62	65	56
House Price	1,301	1,496	1,684	1,786	2,037	1,661
Number of Rooms	6.2	6.5	6.7	6.8	7.0	6.6
Home Ownership	79	82	86	87	91	85
<i>Average Income of White Households in Top Income Quartile</i>	103,250	109,050	117,050	124,850	136,050	118,050

*Notes:* The two panels of this table report statistics for the neighborhoods in which black and white households in the highest income quartile reside. In each case, households are first ranked according to the fraction of households of the same race in their neighborhood and the quintiles of that distribution are shown. In all cases, 'neighborhood' refers to the corresponding Census block group. The first income measure in each case is the average income of the neighborhood. The second income measure reported corresponds to households of the given race in the highest income quartile.

**Table 6: Implied Mean MWTP Measures**

Sample	Residential Sorting Model			Hedonic Price Regressions		
	full sample	within .25 mile of boundaries		full sample	within .25 mile of boundaries	
	No	No	Yes	No	No	Yes
<b>Boundary Fized Effects</b>						
<b>Observations</b>	242,100	27,958	27,958	242,100	27,958	27,958
	(1)	(2)	(3)	(1)	(2)	(3)
% Black*	-316.00 (9.80)	-285.46 (32.06)	-233.94 (38.87)	-101.92 (10.78)	-94.96 (35.28)	-40.46 (42.74)
% Hispanic*	-9.72 (13.59)	-37.19 (46.83)	104.11 (59.01)	142.64 (14.95)	106.60 (51.54)	254.31 (64.88)
% Asian*	-48.97 (11.40)	-69.84 (45.68)	149.77 (55.21)	67.84 (12.54)	-1.69 (50.27)	241.13 (60.71)
% College Degree or More	249.63 (9.19)	185.74 (25.96)	164.78 (39.42)	303.16 (10.11)	235.04 (28.57)	177.11 (43.34)
Average Income*	83.15 (0.74)	89.48 (2.18)	85.44 (2.64)	107.98 (0.82)	113.26 (2.40)	109.22 (2.90)
Average Test Score (in s.d.'s)	18.40 (1.53)	16.69 (4.23)	21.46 (5.29)	24.19 (1.68)	19.01 (4.66)	23.67 (5.81)
Owner-Occupied	154.93 (2.66)	141.08 (7.40)	148.15 (7.38)	133.56 (2.93)	117.59 (8.14)	125.63 (8.12)
Number of Rooms	111.71 (0.69)	111.67 (1.95)	109.28 (1.96)	122.29 (0.76)	123.91 (2.15)	121.72 (2.16)
Built in 1980s	99.60 (3.36)	71.36 (9.29)	87.40 (10.00)	101.88 (3.69)	80.58 (10.23)	108.57 (10.99)
Built in 1960s or 1970s	20.52 (2.41)	1.32 (6.86)	2.48 (7.47)	15.98 (2.65)	-4.40 (7.55)	4.87 (8.21)
Elevation (/100)	-1.70 (0.68)	-14.57 (2.18)	6.15 (3.99)	1.75 (0.75)	-14.08 (2.40)	6.51 (4.39)
Population Density	20.16 (3.23)	40.88 (13.87)	23.12 (17.31)	35.86 (3.56)	49.13 (15.27)	14.65 (19.04)
Crime Index	-0.39 (0.16)	-0.49 (0.65)	-1.02 (1.60)	0.38 (0.18)	0.94 (0.71)	0.36 (1.76)
F-statistic for boundary fixed effects			4.162			8.754

Notes: All neighborhood attributes are measured using the corresponding Census block group. Specifications shown in the table also include controls for interactions between neighborhood racial composition variables and average income as well as land use (% industrial, % residential, % commercial, % open space, % other) in 1, 2, 3, 4, and 5 mile rings around location and six variables that characterize the housing stock in each of these rings. \*Coefficients for % Asian, % Black, % Hispanic, Average Income reported at mean.

**Table 7. Heterogeneity in Marginal Willingness to Pay for Selected Neighborhood and Housing Attributes**

	Neighborhood Sociodemographics					House Characteristics		
	+10% Asian vs. White (at mean)	+10% Black vs. White (at mean)	+10% Hisp vs. White (at mean)	+10% College Educated	Blk Group Avg Income + \$10,000	Own vs. Rent	+1 Room	Built in 1980s vs. pre-1960
<b>Mean MWTP</b>	10.4 (5.9)	-23.4 (3.9)	15.0 (5.5)	16.5 (3.9)	85.4 (2.6)	148.2 (7.4)	109.3 (2.0)	87.4 (10.0)
<b><i>Race (at mean income=\$54,755)</i></b>								
Asian	97.9	-10.5	25.0	10.3	86.9	253.3	78.8	118.3
Black	38.5	66.6	44.1	35.9	65.8	80.3	117.4	96.9
Hispanic	8.7	-9.3	71.1	17.5	91.0	130.7	96.8	73.9
White	-8.1	-37.8	1.1	14.8	86.5	139.5	115.8	82.9
<b><i>Race (at income=\$120,000)</i></b>								
Asian	83.2	-31.7	-0.1	18.4	100.2	394.1	120.2	182.1
Black	37.5	28.6	19.0	41.3	79.2	221.4	159.0	160.8
Hispanic	7.7	-30.6	64.4	22.9	107.1	276.4	141.9	140.5
White	-9.1	-58.9	-23.9	20.1	98.0	277.2	155.0	144.9
<b><i>Education</i></b>								
less than college degree	15.6	-26.0	17.4	-9.1	87.5	134.5	107.0	69.8
college degree	3.8	-20.1	11.9	49.4	82.7	165.6	112.2	110.0

*Notes:* All figures are estimates of marginal willingness to pay for the change shown in the column heading. Figures are reported in terms of a monthly rent - see Data Appendix for a discussion of corresponding price is created for owner-occupied housing units in the sample. The first row of the table reports the mean MWTP measures that correspond to column (3) of Table 5. The remaining rows report MWTP measures for a household with the characteristics shown in the row heading and mean attributes for all other characteristics. All estimates are based on specification that includes boundary fixed effects and all neighborhood variables are measured using the corresponding Census block group.

**Table 8: Counterfactual - Eliminating 'Preferences' for Neighborhood Racial Composition**

<b>Panel A: Overall Results</b>	<i>Neighborhood and Housing Consumption Measures</i>									
	<i>Percent Black</i>	<i>Percent White</i>	<i>Home Ownership</i>	<i>House Size</i>	<i>Average Test Score</i>	<i>Crime Rate</i>	<i>House Value</i>	<i>Commute</i>	<i>Avg. N'hood Income</i>	<i>% N'hood Colleg Educ.</i>
<b>Black</b>										
Pre-Simulation	0.379	0.424	0.40	4.50	458	18.73	740	9.9	37,400	0.31
Post-Simulation - Unadjusted	0.147	0.646	0.35	4.53	502	12.11	868	9.1	44,700	0.38
Post-Simulation - Adjusted	0.146	0.645	0.35	4.53	496	13.01	846	9.0	44,400	0.37
<b>White</b>										
Pre-Simulation	0.048	0.761	0.63	5.36	541	6.14	1160	10.3	57,600	0.47
Post-Simulation - Unadjusted	0.071	0.725	0.63	5.31	531	7.75	1126	9.9	55,700	0.45
Post-Simulation - Adjusted	0.071	0.725	0.63	5.31	531	7.71	1127	9.9	55,700	0.45
<b>Panel B: Results by Income Quartile</b>	<i>Neighborhood and Housing Consumption Measures</i>									
	<i>Percent Black</i>	<i>Percent White</i>	<i>Home Ownership</i>	<i>House Size</i>	<i>Average Test Score</i>	<i>Crime Rate</i>	<i>House Value</i>	<i>Commute</i>	<i>Avg. N'hood Income</i>	<i>% N'hood Colleg Educ.</i>
<b>Black - Lowest Income Quartile</b>										
Pre-Simulation	0.467	0.302	0.31	4.08	445	21.29	612	9.8	33,100	0.28
Post-Simulation - Unadjusted	0.172	0.610	0.22	3.94	490	14.06	689	8.9	39,900	0.34
Post-Simulation - Adjusted	0.172	0.609	0.22	3.94	482	15.30	661	8.8	39,300	0.32
<b>Black - Highest Income Quartile</b>										
Pre-Simulation	0.240	0.516	0.72	5.96	502	11.53	1,261	10.0	53,100	0.43
Post-Simulation - Unadjusted	0.108	0.671	0.76	6.27	535	6.98	1,490	10.1	60,300	0.50
Post-Simulation - Adjusted	0.107	0.671	0.76	6.26	533	7.29	1,477	10.0	60,400	0.50
<b>White - Lowest Income Quartile</b>										
Pre-Simulation	0.058	0.718	0.46	4.36	521	8.42	838	9.9	46,700	0.40
Post-Simulation - Unadjusted	0.068	0.695	0.45	4.31	508	10.78	788	9.5	44,900	0.37
Post-Simulation - Adjusted	0.068	0.696	0.45	4.30	506	10.95	786	9.5	44,700	0.36
<b>White - Highest Income Quartile</b>										
Pre-Simulation	0.026	0.807	0.87	6.56	565	4.18	1,607	11.2	71,800	0.56
Post-Simulation - Unadjusted	0.062	0.716	0.87	6.50	559	5.03	1,585	10.7	69,300	0.55
Post-Simulation - Adjusted	0.063	0.716	0.87	6.51	559	4.93	1,586	10.7	69,400	0.55

*Note:* This table reports the consumption of housing and neighborhood amenities by households of each race. Numbers are reported for a counterfactual simulation that sets all 'preference' parameters associated with neighborhood racial composition to zero.

**Table 9: Counterfactual - Equalizing Proportion of Blacks and Whites in Metropolitan Population**

<b>Panel A: Overall Results</b>		<i>Neighborhood and Housing Consumption Measures</i>								
	<i>Percent Black</i>	<i>Percent White</i>	<i>Home Ownership</i>	<i>House Size</i>	<i>Average Test Score</i>	<i>Crime Rate</i>	<i>House Value</i>	<i>Commute</i>	<i>Avg. N'hood Income</i>	<i>% N'hood Colleg Educ.</i>
<b>Black</b>										
Pre-Simulation	0.379	0.424	0.40	4.50	458	18.73	740	9.9	37,400	0.31
Post-Simulation - Unadjusted	0.720	0.131	0.45	4.71	499	11.89	925	9.1	31,700	0.30
<b>White</b>										
Pre-Simulation	0.048	0.761	0.63	5.36	541	6.14	1160	10.3	57,600	0.47
Post-Simulation - Unadjusted	0.125	0.705	0.65	5.39	537	7.94	1178	10.4	42,900	0.39
<b>Panel B: Results by Income Quartile</b>		<i>Neighborhood and Housing Consumption Measures</i>								
	<i>Percent Black</i>	<i>Percent White</i>	<i>Home Ownership</i>	<i>House Size</i>	<i>Average Test Score</i>	<i>Crime Rate</i>	<i>House Value</i>	<i>Commute</i>	<i>Avg. N'hood Income</i>	<i>% N'hood Colleg Educ.</i>
<b>Black - Lowest Income Quartile</b>										
Pre-Simulation	0.467	0.302	0.31	4.08	445	21.29	612	9.8	33,100	0.28
Post-Simulation - Unadjusted	0.763	0.102	0.31	4.21	489	13.56	741	8.6	25,800	0.25
<b>Black - Highest Income Quartile</b>										
Pre-Simulation	0.240	0.516	0.72	5.96	502	11.53	1,261	10.0	53,100	0.43
Post-Simulation - Unadjusted	0.683	0.194	0.82	6.10	526	7.43	1,428	9.6	48,400	0.43
<b>White - Lowest Income Quartile</b>										
Pre-Simulation	0.058	0.718	0.46	4.36	521	8.42	838	9.9	46,700	0.40
Post-Simulation - Unadjusted	0.178	0.695	0.40	4.35	522	10.08	842	9.8	30,100	0.29
<b>White - Highest Income Quartile</b>										
Pre-Simulation	0.026	0.807	0.87	6.56	565	4.18	1,607	11.2	71,800	0.56
Post-Simulation - Unadjusted	0.089	0.716	0.89	6.42	551	5.79	1,529	11.4	55,000	0.49

*Note:* This table reports the consumption of housing and neighborhood amenities by households of each race. Numbers are reported for a counterfactual simulation that equalizes the fraction of white and black households in the population (38.1%), leaving estimated preference parameters and distribution of socioeconomic characteristics for each race unchanged.

**Table 10. Enhanced Cutler-Glaeser Regressions: The Effect of Metropolitan Segregation on Individual Outcomes**

*Coefficients on interactions between black and metropolitan segregation (dissimilarity index) and proportion college-educated blacks in metro area reported*

	Age 20-24					Age 25-30				
	<i>HS Graduate</i>	<i>College Graduate</i>	<i>Ln(Earnings)</i>	<i>Idle</i>	<i>Any Child Born</i>	<i>HS Graduate</i>	<i>College Graduate</i>	<i>Ln(Earnings)</i>	<i>Idle</i>	<i>Any Child Born</i>
<i>Cutler-Glaeser Regressions</i>										
Black*Metro Dissimilarity Index (Segregation)	-0.269*** (0.041)	-0.094*** (0.032)	-0.788*** (0.140)	0.340*** (0.031)	0.189*** (0.032)	-0.201*** (0.039)	-0.064 (0.062)	-0.433*** (0.094)	0.310*** (0.038)	0.131*** (0.035)
<i>Adding Interactions with (% Metro Black and College-Educated)</i>										
Black*Segregation	-0.412*** (0.080)	-0.101*** (0.039)	-1.123*** (0.260)	0.387*** (0.070)	0.218*** (0.071)	-0.241*** (0.072)	-0.016 (0.065)	-0.505*** (0.164)	0.394*** (0.083)	0.217*** (0.046)
<b>Black* Segregation* (% Metro Black and College Educated)</b>	<b>13.60*** (4.32)</b>	<b>2.40 (3.30)</b>	<b>20.06* (12.25)</b>	<b>-5.26* (3.09)</b>	<b>-4.61 (3.79)</b>	<b>6.21* (3.67)</b>	<b>-0.88 (5.94)</b>	<b>6.07 (7.93)</b>	<b>-7.45** (3.06)</b>	<b>-9.83*** (2.87)</b>
Black* (% Metro Black and College Educated)	-8.89*** (2.95)	-2.13 (2.36)	-9.36 (8.81)	4.00* (2.49)	1.90 (2.75)	-3.02 (2.59)	0.54 (4.22)	-0.82 (5.32)	4.51** (2.03)	6.00*** (2.11)

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



---

**Appendix Table 1: Housing and Neighborhood Consumption Patterns for Households in Lowest Income Quartile**

---

**Panel A: Housing and Neighborhood Consumption of Black Households in Lowest Income Quartile**

*Households ranked by % Black in neighborhood - consumption measures summarized by quintile*

<b>Quintile</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>Total</b>
Percent Neighborhood - Black	0-14	14-37	37-65	65-82	82-100	47
Average Test Score	502	462	440	417	398	444
Crime Rate	13	20	23	25	29	22
Average Neighborhood Income	42,350	33,950	29,550	26,200	23,100	31,030
Percent Neighborhood - College Degree of more	39	34	24	18	12	25
House Price	695	563	501	473	439	534
Number of Rooms	4	4	4	4	4	3.9
Home Ownership	20	20	25	33	32	26
<i>Average Income of Black Households in Bottom Income Quartile</i>	12,550	11,750	11,600	11,000	10,400	11,460

**Panel B: Housing and Neighborhood Consumption of White Households in Lowest Income Quartile**

*Households ranked by % White in neighborhood - consumption measures summarized by quintile*

<b>Quintile</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>Total</b>
Percent Neighborhood - White	0-57	57-71	71-80	80-88	88-100	72
Average Test Score	473	501	518	541	574	521
Crime Rate	18	11	7	4	3	9
Average Neighborhood Income	35,100	40,600	46,050	51,200	63,100	47,210
Percent Neighborhood - College Degree of more	29	39	42	44	52	41
House Price	629	694	789	861	1,011	796
Number of Rooms	4	4	4	4	5	4.2
Home Ownership	37	38	42	48	60	45
<i>Average Income of White Households in Bottom Income Quartile</i>	12,650	13,400	13,550	13,700	13,600	13,380

---

Notes: The two panels of this table report statistics for the neighborhoods in which black and white households in the lowest income quartile reside. In each case, households are first ranked according to the fraction of households of the same race in their neighborhood and the quintiles of that distribution are shown. In all cases, 'neighborhood' refers to the corresponding Census block group. The first income measure reported in each case is the average income of the neighborhood. The second income measure corresponds to households of the given race in the lowest income quartile.

**Appendix Table 2: Interaction Parameter Estimates**

Housing/Neighborhood Attribute	Household Characteristic												
	Hhld Income	Children Under 18	Black	Hispanic	Asian	Some College	College Degree or More	Working	Age	Hhld Capital Income	Black* Hhld Income	Hispanic* Hhld Income	Asian* Hhld Income
Monthly House Price	0.071 (0.003)	0.071 (0.023)	0.087 (0.051)	-0.244 (0.048)	0.208 (0.056)	0.285 (0.068)	0.400 (0.042)	0.197 (0.062)	0.007 (0.001)	0.013 (0.002)	0.030 (0.015)	0.082 (0.025)	0.034 (0.013)
Owner-Occupied	0.142 (0.005)	-0.050 (0.025)	-0.427 (0.058)	-0.046 (0.036)	0.851 (0.058)	0.027 (0.051)	0.191 (0.038)	0.303 (0.065)	0.046 (0.004)	0.094 (0.007)			
Number of Rooms	0.151 (0.005)	0.522 (0.027)	0.010 (0.034)	-0.521 (0.052)	-1.223 (0.071)	0.085 (0.047)	0.036 (0.031)	0.011 (0.043)	0.007 (0.001)	-0.060 (0.005)			
Built in 1980s	0.045 (0.004)	-0.064 (0.021)	0.065 (0.046)	-0.040 (0.039)	0.184 (0.045)	0.192 (0.062)	0.196 (0.037)	0.337 (0.068)	-0.011 (0.001)	0.019 (0.003)			
Built in 1960-79	0.013 (0.003)	0.023 (0.018)	0.315 (0.054)	-0.139 (0.045)	0.221 (0.057)	0.163 (0.065)	0.044 (0.029)	0.209 (0.060)	-0.005 (0.001)	0.002 (0.001)			
Average Test Score	0.001 (0.002)	0.056 (0.023)	-0.229 (0.056)	-0.077 (0.039)	0.086 (0.043)	0.186 (0.066)	0.206 (0.040)	0.141 (0.057)	0.011 (0.002)	0.056 (0.002)			
Elevation	0.019 (0.002)	0.038 (0.013)	-0.097 (0.038)	-0.134 (0.044)	0.006 (0.036)	0.141 (0.064)	0.090 (0.036)	-0.018 (0.042)	0.006 (0.001)	-0.039 (0.006)			
Population Density	0.017 (0.004)	-0.216 (0.024)	-0.561 (0.062)	-0.030 (0.031)	0.004 (0.032)	-0.006 (0.036)	0.159 (0.037)	-0.253 (0.063)	-0.006 (0.001)	0.042 (0.005)			
Crime Index	-0.016 (0.003)	0.010 (0.021)	0.491 (0.065)	0.045 (0.038)	0.017 (0.038)	-0.044 (0.044)	0.235 (0.041)	-0.164 (0.059)	0.014 (0.002)	0.066 (0.010)			
% Black	-0.073 (0.005)	0.114 (0.023)	1.700 (0.069)	0.697 (0.052)	0.680 (0.060)	-0.089 (0.054)	0.145 (0.037)	-0.114 (0.056)	-0.003 (0.001)	-0.110 (0.015)	-0.032 (0.019)		
% Hispanic	-0.063 (0.006)	0.125 (0.019)	0.700 (0.063)	0.891 (0.055)	0.425 (0.058)	-0.192 (0.062)	-0.094 (0.032)	-0.015 (0.043)	-0.010 (0.002)	-0.076 (0.012)		0.126 (0.030)	
% Asian	-0.003 (0.005)	0.088 (0.024)	0.799 (0.061)	0.311 (0.053)	1.708 (0.063)	-0.064 (0.059)	-0.215 (0.040)	-0.056 (0.047)	0.001 (0.001)	-0.079 (0.013)			-0.036 (0.021)
% College Degree or More	0.022 (0.006)	-0.200 (0.025)	0.574 (0.054)	0.080 (0.047)	-0.052 (0.043)	0.375 (0.070)	1.681 (0.058)	-0.338 (0.065)	-0.006 (0.002)	0.093 (0.014)			
Average Income	0.045 (0.006)	0.048 (0.018)	-0.808 (0.057)	0.245 (0.045)	-0.053 (0.041)	-0.028 (0.049)	-0.313 (0.053)	0.100 (0.056)	0.003 (0.001)	0.020 (0.006)			
% Black*Average Income	0.054 (0.009)		0.674 (0.078)								-0.026 (0.032)		
% Hispanic*Average Income	0.067 (0.010)			0.340 (0.067)								-0.081 (0.039)	
% Asian*Average Income	0.010 (0.009)				0.282 (0.071)								-0.003 (0.027)
Distance to Work	-0.022 (0.002)	0.156 (0.023)	-0.272 (0.043)	0.189 (0.035)	0.221 (0.039)	-0.093 (0.058)	0.160 (0.032)	-13.765 (0.056)	-0.010 (0.001)	-0.468 (0.019)			

Note: Parameter estimates reported with all variables normalized to have mean zero, standard deviation one. Standard errors are in parentheses.