

NBER WORKING PAPER SERIES

RESIDENTIAL SEGREGATION IN GENERAL EQUILIBRIUM

Patrick Bayer
Robert McMillan
Kim Rueben

Working Paper 11095
<http://www.nber.org/papers/w11095>

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

We would like to thank Fernando Ferreira for outstanding research assistance. Thanks also to Pedro Cerdan and Jackie Chou for help in assembling the data. We are grateful to Pat Bajari, Steve Berry, Dennis Epple, Tom Nechyba, Holger Sieg and Chris Timmins for many valuable discussions, and to Joe Altonji, Gregory Besharov, Maureen Cropper, David Cutler, Ed Glaeser, James Heckman, Vernon Henderson, Phil Leslie, Costas Meghir, Robert Moffitt, Michael Riordan, Steve Ross, Kerry Smith, Jon Sonstelie, Chris Taber, Chris Udry, and Jacob Vigdor for additional valuable comments. We also thank conference participants at the AEA, ERC, IRP, NBER, PET, SITE, and SIEPR, and seminar participants at Brown, Chicago, Colorado, Columbia, Duke, Johns Hopkins, Northwestern, NYU, PPIC, Stanford, Toronto, UC Berkeley, UC Irvine, UCLA, and Yale for useful 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 Canada. 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, Robert McMillan, and Kim Rueben. 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.

Residential Segregation in General Equilibrium
Patrick Bayer, Robert McMillan, and Kim Rueben
NBER Working Paper No. 11095
January 2005
JEL No. H0, J7, R0, R2

ABSTRACT

Black households in the United States with high levels of income and education (SES) typically face a stark tradeoff when deciding where to live. They can choose neighborhoods with high levels of public goods or a high proportion of blacks, but very few neighborhoods combine both, a fact we document clearly. In the face of this constraint, we conjecture that racial sorting may dramatically lower the consumption of local public goods by high-SES blacks. To shed light on this, we estimate a model of residential sorting using unusually detailed restricted Census microdata, then use the estimated preferences to simulate a counterfactual world in which racial factors play no role in household residential location decisions. Results from this exercise provide the first evidence that sorting on the basis of race gives rise to significant reductions in the consumption of local public goods by black and high-SES black households in particular. These consumption effects lead to significant losses of welfare and are likely to have important intergenerational implications.

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.ca

Kim Rueben
Public Policy Institute of California
krueben@ui.urban.org

1 INTRODUCTION

That race plays a fundamental role in shaping the way that households sort into neighborhoods in US cities has been widely documented in the academic literature.¹ While residential segregation is the most obvious result, perhaps less apparent is the way that the role that race plays in the market shapes the quality of neighborhoods chosen by black versus white households in other dimensions. Consider, for example, the location decision of a college-educated black household. In the vast majority of US metropolitan areas, few if any neighborhoods exist that combine high fractions of both black and college-educated households – this we demonstrate very clearly below. This neighborhood supply constraint presents college-educated blacks with a clear trade-off between living with an increased fraction of black neighbors or living with an increased fraction of college-educated neighbors. Faced with this choice, college-educated blacks in fact choose a very diverse set of neighborhoods: those living with the smallest fraction of other blacks typically live with the same fraction of college-educated neighbors as whites do on average; in marked contrast, those college-educated blacks who live with the highest fraction of other blacks live with a substantially smaller fraction of college-educated neighbors.

This much lower consumption of average neighborhood education (which in this example proxies for neighborhood quality more generally) by those college-educated blacks that choose neighborhoods with the highest fraction of black neighbors prompts the central hypothesis of this paper. Specifically, we conjecture that racial sorting in US metropolitan areas leads to substantial reductions in the consumption of local public goods by blacks, and particularly blacks

¹ See Cutler, Glaeser, and Vigdor (1999) for a detailed analysis of the evolution of segregation patterns in the United States over the twentieth century. Miller and Quigley (1990) and Bayer, McMillan, and Rueben (2004) among others conclude that a substantial portion of black-white segregation remains unexplained by sorting due to racial differences in other socioeconomic variables. Analysis of survey responses concerning racial preferences reaches a

with high levels of income and education. That is, if race were not a factor in the sorting process, many black households would consume far higher levels of local public goods, with important consequences for household welfare in the short term and individual outcomes in the longer term.

To quantify the impact of racial sorting on the consumption of local public goods by households of different races, we would ideally like to compare current consumption levels with those in a world where households chose neighborhoods without regard for race. While it would be fanciful to expect to see this directly in observational data, we adopt an approach that implements this ideal thought experiment in an intuitive and feasible way. Using unusually rich restricted Census microdata on quarter of a million households linked to precise geographic locations in the San Francisco Bay Area in 1990, we estimate a flexible model of residential sorting that allows household preferences for a wide variety of housing and neighborhood attributes to be recovered from their location decisions. At its heart, this 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, one can infer how preferences for housing and neighborhood attributes vary with these sociodemographic characteristics.

With estimates of the relative weight that households with different characteristics place on a wide set of housing and neighborhood attributes in hand,² we then conduct an intuitive counterfactual simulation that quantifies the effect of racial sorting on the consumption of local public goods. Specifically, we compare the current state of the world to one in which race is

similar conclusion. See Ihlanfeldt and Scafidi (2002, forthcoming) for recent evidence from the Multi-City Study of Urban Inequality (MCSUI) data.

² It is worth emphasizing here that in developing estimation strategy below we explicitly account for the correlation of neighborhood sociodemographic characteristics with unobserved aspects of housing and neighborhood quality

eliminated as a factor in each household's residential location decision, calculating a new housing market equilibrium once all preferences associated with neighborhood racial composition have been eliminated. The results provide clear evidence that racial sorting increases black-white differences in the consumption of public safety and school quality by over 100 percent; in the consumption of neighborhood income and education, black-white differences increase by nearly this amount. This decreased consumption by blacks relative to whites is partially compensated by lower housing prices. However, given that close substitutes for many aspects of local public good consumption (such as everyday exposure to highly educated neighbors) are not easily replicated by the private market, such substantial decreases in local public good consumption are still very likely to generate immediate welfare losses in addition to having a significant impact on the inter-generational persistence of racial differences in education, income, and wealth.

In light of the potential for discrimination in the housing market, we emphasize that our empirical analysis captures the total effect of racial sorting - whether driven by preferences or discrimination - on the consumption of local public goods. Because these rival explanations for racial sorting are indistinguishable in observational data,³ notice that the adverse effects of racial sorting on the consumption of local public goods by blacks that we identify might arise even in the absence of discrimination - i.e., they could arise entirely as the result of the collective decentralized location decisions made by the households in a metropolitan area. The fundamental explanation for these adverse effects is the extremely short supply of neighborhoods that combine both a high fraction of black households with high levels of local public goods in

using a boundary fixed effects strategy related to that of Black (1999). Failure to do so would lead to a substantial overstatement of the impact of neighborhood race on neighborhood sorting.

³ As we show in Section 5 below, given the most natural model of discrimination (i.e., that it drives an increasing wedge between the implicit price that blacks versus whites pay for housing in neighborhoods with an increasing

US metropolitan areas, a fact that itself is almost certainly attributable to the even more basic fact that black and especially highly-educated black households constitute such a small fraction of the population (11 percent and 2 percent, respectively). Thus as long as these fundamental conditions continue to hold, the adverse effects of racial sorting on local public good consumption by blacks would likely continue into the future even if all forms of centralized discrimination were to be eliminated.

This paper fits into a larger body of research that examines the impact of segregation on individual outcomes.⁴ In many of these studies, concerns about the non-random selection of individuals into neighborhoods give rise to serious measurement issues and, consequently, a recent literature beginning with Cutler and Glaeser (1997) attempts to draw causal inferences by examining the impact of segregation measured at the metropolitan level on outcomes for young adults.⁵ The current paper relates to this literature in two ways. First, our results point to a specific mechanism through which metropolitan segregation levels might be strongly associated with individual outcomes – the effect of racial sorting on racial differences in the consumption of local public goods and exposure to high-SES neighbors. Second, while the vast majority of the prior literature examines the impact of segregation on individual outcomes, we essentially push the analysis back a step, examining the impact of the role that race plays in the individual location decision problem on the set of neighborhoods that arise in equilibrium. In our analysis, then, segregation is a feature of the collective set of neighborhoods that form – along with the

fraction of whites), discrimination gives rise to behavioral and pricing predictions that can be matched identically by a model that only allows for preferences over racial composition.

⁴ This literature dates to Kain (1968) and includes Ihlanfeldt (1992), Ihlanfeldt and Sjoquist (1990), O'Regan and Quigley (1998), Ross (1998), Weinberg (2000, 2004), and Ross and Zenou (2004) among many others.

⁵ They find little effect for whites but significant negative effects on educational attainment, labor market activity, and the avoidance of teenage fertility for blacks. In a final portion of their analysis, Cutler and Glaeser attempt to distinguish whether three potential mechanisms explain their findings, one of which is the decreased exposure of blacks to educated neighbors in segregated metropolitan areas. They find evidence that this mechanism, which is related to our analysis, explains some of the negative impact of metropolitan segregation on outcomes.

level of local public goods consumed by households of each race – rather than the fundamental explanatory variable in the analysis.

The rest of the paper is organized as follows: Section 2 presents descriptive evidence regarding neighborhood availability and segregation from all the metropolitan areas in the United States from the year 2000. Here, we provide the first evidence of a significant shortage of neighborhoods combining a high fraction of black households with a high fraction of college-educated households; in the face of this constraint, we then show that highly educated blacks reside in a wide variety of neighborhoods. Section 3 outlines the key feature of our detailed San Francisco Bay Area dataset, and Section 4 provides evidence using these data relating to the distribution of neighborhoods available in the Bay Area housing market, as well as the actual neighborhood choices made by households in different race and income categories. Again, we will find remarkably similar patterns for the Bay Area as for the US as a whole. 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 a counterfactual simulation that sheds light on our main conjecture, making clear the effects of eliminating race as a factor in the household location decision on the consumption of local public goods by different types of household. Section 9 concludes.

2 PATTERNS OF RACIAL SORTING IN US METROPOLITAN AREAS

To motivate the central hypothesis of this paper, we begin our analysis by characterizing the two broad patterns for the US as a whole already referred to: (i) that neighborhoods that combine high fractions of both college-educated and black households are in extremely short

supply in almost every metropolitan area throughout the United States and (ii) that faced with the resulting trade-off between black versus other college-educated neighbors, a significant fraction of college-educated blacks in every metropolitan area choose neighborhoods with a high fraction of other black households and very low levels of average neighborhood education. This pattern of choices suggests that race-related factors in the housing market, whether driven by preferences or discrimination, may lead to a substantial decrease in the consumption of local public goods by many college-educated blacks.

The analysis in this section is based on data compiled from Summary Files of the 2000 Census, which give information on the distribution of education by race for each Census tract in the United States. We characterize the race and educational attainment of households as that of the head of household and focus specifically on non-Hispanic black and non-Hispanic white households throughout our analysis.^{6,7} Based on this definition, black and white households constitute 11.1 and 69.5 percent of US households that reside in metropolitan areas, respectively. Among black households, 15.4 percent have a college degree, while the comparable figure for white households is 32.5 percent, and for all US households, 27.7 percent. For this portion of our analysis, we take household education to proxy for socioeconomic status more generally and average neighborhood education proxies for the level of local public goods. Using far richer data from a single metropolitan area, we consider a wider set of both household and neighborhood characteristics below.

⁶ Given the analysis below, it would have been preferable to use information on the joint distribution of race and income rather than education to make the results directly comparable. Unfortunately, the corresponding Census data, which we downloaded at the time of writing, had clear and serious errors. We use Census tracts rather than block groups in this portion of the analysis because tracts are the lowest level at which the joint distribution of education and race is available in the Census Summary files.

⁷ The vast majority of households that checked two races can be characterized as either Hispanic or non-Hispanic Asian or Pacific Islander. Other households that checked two or more races - a very small fraction overall - were dropped from this analysis.

Table 1 documents the number of tracts in the United States by the percentage of households with a college degree and the percentage of households that are black and white. The first row describes the number of tracts in which more than 0, 20, 40, 60, and 80 percent of head of households are college-educated, respectively. The next four rows report the number of tracts in each of these categories that contain a minimum fraction of black households 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 of black households have a high fraction of households with a college degree. For example, while 23 percent of all tracts are at least 40 percent college educated (a number comparable to the fraction of US households with a college degree), only 2.5 percent of tracts that are at least 40 percent black are at least 40 percent college educated, and only 1.1 percent of tracts that are at least 60 percent black are at least 40 percent college educated. Moreover, tracts that combine high fractions of both black and college-educated households are in fact concentrated in just a handful of metropolitan areas, most notably Washington, DC, implying that the supply of such neighborhoods in most metropolitan areas is extremely limited.⁸

The final four rows of Table 1 show analogous numbers for white households, reporting the number of tracts in the US that meet the education criterion described in each column heading subject to a minimum fraction of white households equal to 20, 40, 60, and 80 percent, respectively. As the corresponding figures show in this case, a markedly different pattern emerges for white households, with a slightly greater fraction of neighborhoods with at least 40, 60, and 80 percent white households meeting each education criterion. Figure 1, which plots the

⁸ Of the 44 tracts (less than 0.1 percent of all tracts) that are at least 60 percent black and 40 percent college-educated, 13 are in the Washington DC PMSA, 8 in Detroit, 6 in Los Angeles, and 5 in Atlanta. Almost 75 percent of these tracts can thus be found in one of only four PMSAs. Of the 142 tracts that are at least 40 percent black and 40 percent college-educated, almost two-thirds are in the PMSAs listed above along with Chicago and New York.

percentages shown in Table 1 for blacks and whites respectively, clearly illustrates the difference in patterns. Taken together, this characterization of the composition of US neighborhoods makes clear that while neighborhoods that combine high fractions of both college-educated and white households are amply supplied in metropolitan areas throughout the US, neighborhoods that combine high fractions of both college-educated and black households are in extremely short supply. This implies that college-educated black households in most metropolitan areas face a clear trade-off between living with other black versus other college-educated neighbors

To show this trade-off graphically, Figures 1a-1f show a series of scatterplots for six metropolitan areas: Boston, Dallas, New York, Philadelphia, San Francisco, and Washington, D.C.. In these figures, each small circle represents a Census tract and is located according to the fraction of college-educated households (vertical axis) and the fraction of black households (horizontal axis) in the tract. The size of the circle is based on the number of college-educated black households in the neighborhood – so that the largest circles depict the tracts in which college-educated blacks are most likely to live.⁹ These scatterplots clearly demonstrate a short supply of neighborhoods that combine high fractions of both college-educated and black households and, consequently, the implicit trade-off that college-educated black households must make when choosing a neighborhood.¹⁰

The figures also show the diverse set of neighborhoods in which college-educated black households reside given this constraint. In particular, while a sizeable fraction of college-educated blacks in each of the depicted MSAs chooses neighborhoods with few black and many college-educated neighbors, a sizeable fraction also chooses neighborhoods with many black and

⁹ Note that tracts that do not contain any college-educated blacks do not show up in these scatterplots.

¹⁰ The relatively higher availability of such tracts in Washington D.C. relates to the relatively high fraction of college-educated blacks in the metropolitan area. We return to a discussion of this point when describing the results of our simulations below.

few college-educated neighbors. To examine the range of neighborhoods chosen by college-educated blacks in metro areas throughout the country as a whole, Table 2 summarizes the information contained in the scatterplots in a succinct way. In particular, we first rank college-educated black households in each metropolitan area by the fraction of blacks in the household's Census tract and assign households to their corresponding quintile of this distribution. This corresponds to drawing four vertical lines in each scatterplot so that an equal number of college-educated black households are in each of the resulting regions. The upper panel of Table 2 then summarizes the average fractions of black and college-educated households in the corresponding tract for the quintiles of this distribution, averaging over all US metropolitan areas.

Thus the first column of Table 2, for example, characterizes the average neighborhood composition for the 20 percent of college-educated black households that reside with the smallest fraction of other black households in their metropolitan area. A clear trade-off is apparent between the fraction of a household's neighbors that are black and the fraction that are college-educated; the average fraction of college-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. The lower panel of Table 2 reports analogous results for white households. While not perfectly monotonic, the resulting pattern for whites is almost exactly opposite that for blacks: those whites residing with the greatest fraction of neighbors of the same race within each metropolitan area generally reside with a greater rather than smaller fraction of college-educated neighbors.

The patterns of neighborhood sorting shown in Table 2, while not conclusive on their own, provide suggestive evidence concerning the impact of racial sorting on the consumption of local public goods, indicating a clear trade-off between living with households of the same race

and households of the same education level for college-educated blacks. There is no such tradeoff for whites. Moreover, while those college-educated blacks in each metropolitan area who live with the smallest fraction of other blacks in fact live in tracts with roughly the same fraction of college-educated neighbors as college-educated whites do on average, those college-educated blacks living with the greatest fraction of other blacks live with only about *one-third* that fraction. This suggests that racial sorting in metropolitan areas throughout the United States, whether driven by preferences or discrimination, may lead to a substantial decrease in the consumption of local public goods by a large fraction of college-educated blacks.

3 DATA

Given this broad characterization of racial sorting in metropolitan areas throughout the United States, we now turn to the much more detailed dataset that we have assembled for our primary analysis. Relative to the data used in the previous section to study patterns of racial sorting for the US as a whole, the key feature of this dataset is that it links microdata that provide a wide set of household characteristics to precise geographic indicators that allow us to characterize the neighborhood in which each household lives. As we describe below, the resulting dataset allows us to estimate a model of residential sorting that controls carefully for a wide set of household characteristics and makes 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. Having identified the relative roles of these factors in each type of household's location decision, we then identify the impact of racial sorting on the consumption of local public goods by simulating a new housing market equilibrium in which race has been eliminated as a one of these factors.

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 the public-use data specify the PUMA (a Census region with approximately 100,000 individuals) in which a household lives, the restricted data specify the Census block (a Census region with approximately 100 individuals), thereby identifying the local neighborhood that each individual inhabits and the characteristics of each neighborhood far more accurately than has been previously possible with such a large data set.

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.¹¹ In addition, it provides a variety of housing characteristics: whether

¹¹ Throughout our analysis, we treat the household as the decision-making agent and characterize each household's race as the race of the 'householder' – typically the household's primary earner. We assign 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 throughout, we limit our focus to black and white households, although it is important to point out that our primary analysis below also controls separately for Asian and Hispanic households.

the unit is owned or rented, the corresponding rent or owner-reported value,¹² 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.¹³ 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.

4 RACIAL SORTING AND LOCAL PUBLIC GOODS CONSUMPTION IN THE SFBA

Before turning to our model of residential sorting, we first describe the patterns of racial sorting and local public goods consumption in the Bay Area, thereby providing the reader with a clear sense of the pattern of racial sorting in our primary data set and also motivating our main conjecture – that racial sorting in the housing market has a substantial negative effect on the consumption of local public goods by black households, especially those at high socioeconomic levels.

¹² 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.

¹³ 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.

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.¹⁴ 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

¹⁴ 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.

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 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, suggesting that racial sorting may indeed have a substantial impact on the characteristics of the neighborhoods chosen by high-income blacks along other dimensions.

Racial Sorting and the Consumption of Local Public Goods. To explore this possibility directly, Tables 5 and 6 describe the distribution of neighborhoods in which blacks and whites in the highest and lowest quartile of the income distribution reside, respectively. 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.

Focusing first on Panel A of Table 5, this shows the distribution of neighborhoods in which blacks in the highest income quartile reside. We order high-income 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. 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.

As Table 5 makes clear, while increased consumption of local public goods 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 black households but increases in the fraction of households of the same race for whites. Given segregating racial preferences (as we find below), this implies that high-income blacks face an implicit price of consuming these local public goods that is higher than the price faced by comparable whites. Thus if race were removed as a consideration in the location decision along the lines of our central hypothesis, the implicit price that blacks would face in choosing neighborhoods with higher levels of local public goods would fall relative to whites, likely leading those high-income blacks that currently live in the most segregated neighborhoods to choose neighborhoods in line with those chosen by the high-income blacks living with the smallest fraction of black neighbors. A simple comparison of the average consumption of local public goods by high-income blacks living with the smallest fraction of black neighbors to that of high-income whites as a whole (shown in the final column of Table 5) reveals that a sizeable movement of this kind could lead to a substantial

decline in the gap in the consumption of local public goods between high-income black and white households.

In a similar fashion, Table 6 reveals that blacks in the lowest income quartile also face an implicit price of local public goods that exceeds the direct costs. While not as marked as for households in the highest income quartile, increased consumption of local public goods is again accompanied by sharp decreases in the fraction of black households. Thus we anticipate that racial sorting in the housing market lowers the consumption of local public goods for blacks at all income levels, although perhaps not as markedly for those with lower levels of income.

5 A MODEL OF RESIDENTIAL SORTING

The patterns of racial sorting and local public good consumption just described for the Bay Area and in Section 2 for the US as a whole suggest a strong link between the importance of race in the residential sorting process and the vast differences between whites and blacks in the consumption of local public goods. In particular, those black households that live with the smallest fraction of black neighbors in their respective metropolitan area consume levels of local public goods that are in line with the average levels consumed by comparable whites, while those who live with the highest fraction of black neighbors consume much lower levels.

To measure the impact of racial sorting on the consumption of local public goods more formally, we now turn to the principal analytical tool we use to explore this link - an equilibrium model of a self-contained urban housing market in which households sort themselves among the set of available housing types and locations. Following the model developed in Bayer, McMillan, and Rueben (2004b), the model that we set out here consists of two key elements: the household residential location decision problem and a market-clearing condition. While it has 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 sociodemographic compositions to be determined in housing market equilibrium. This will be valuable for the counterfactual simulation analysis below.¹⁵

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.¹⁶ 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.¹⁷ Let p_h denote the price of housing choice h and let d_h^i denote the distance from residence h to the

¹⁵ A long line of theoretical studies, including important papers by Anas and Kim (1995), Anas (2002), Epple, Filimon and Romer (EFR) (1984, 1993), Benabou (1993, 1996), Fernandez and Rogerson (1996, 1998), Nechyba (1999, 2000), Anas and Kim (1995), and Anas (2002) 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.

¹⁶ 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 but again do not include choice-specific unobservables.

¹⁷ 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.

primary work location of household i . Each household chooses its residence h to maximize its indirect utility function V_h^i :

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

The error structure of household indirect utility is divided into a correlated component associated with each house valued the same by all households, ξ_h , and an individual-specific term, ε_h^i . A useful interpretation of ξ_h is that it captures unobserved housing quality, including any unobserved quality associated with the surrounding neighborhood.¹⁸

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, α_j^i , for $j \in \{X, \bar{Z}, d, p\}$, varies with a household's own characteristics according to:

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

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

¹⁸ 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.

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 .¹⁹

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

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

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 $\{X, p, \xi\}$:

$$(4) \quad P_h^i = f_h(Z^i, X, p, \xi)$$

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:

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

¹⁹ 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 ε_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.

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.²⁰

Whether this model gives rise to multiple equilibria depends on the distributions of preferences and available housing choices, as well as the utility parameters.²¹ 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. Thus, the primary place where the issue of whether the equilibrium is unique arises is in conducting counterfactual simulations; we discuss this issue in Section 8.

²⁰ 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) ϵ is drawn from a continuous density function.

²¹ On the one hand, as described above, when neighborhood sociodemographic characteristics do not enter the utility function, the equilibrium is unique. On the other hand, if households have strong preferences to live with others of the same race and do not value any other housing or neighborhood attributes, multiple equilibria arise, each characterized by complete racial segregation, but with the attachment of a given race to a given neighborhood completely indeterminate. The real world, of course, lies somewhere in between these extreme cases.

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, including a discussion of methods for simplifying the computation and a description of the asymptotic properties of the estimator, is included in a technical appendix. In this section, we outline the estimation procedure, focusing on identification of the model.

It is helpful in describing the estimation procedure 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), δ_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, δ_h is the mean indirect utility provided by housing choice h . The unobservable component of δ_h , ξ_h , captures the portion of unobserved preferences for housing choice h that is correlated across households, while ε_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,²² returning estimates of the heterogeneous parameters in λ and mean indirect utilities, δ_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 λ and mean indirect utilities, δ_h , the model predicts the probability that each household i chooses house type h . We assume that ε_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 λ and the vector of mean indirect utilities to maximize ℓ .

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 δ into observable and unobservable components according to

²² 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 ε . This ensures that no household's particular idiosyncratic preferences affect the equilibrium and the vector of idiosyncratic preferences ε is uncorrelated with the prices and neighborhood sociodemographic characteristics that arise in any equilibrium. For more discussion, see the Technical Appendix.

the regression equation (8).²³ 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, ξ_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, 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.²⁴ To the extent that sorting with respect to the

²³ Notice that the set of observed residential choices provides no information that distinguishes the components of δ . That is, however δ is broken into components, the effect on the probabilities shown in equation (10) is identical.

²⁴ 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.

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.²⁵ 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.²⁶ 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

²⁵ 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.

²⁶ 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.²⁷ 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.

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.

Regardless of whether the sizes of the parameters that multiply the interactions of household race and neighborhood racial composition result from preferences or discrimination, these parameters do inform us about the importance of sorting on the basis of race in the housing

²⁷ As described in the Data Appendix, we construct a single price vector for all houses, whether rented or owned.

market. If one thinks of discrimination as an expression of the preferences of the discriminating group concerning the group discriminated against, then our model essentially misassigns these preferences to the group discriminated against. Thus, while our estimate of the preferences of black households to live with other black households may be overstated, the difference between the preferences of white versus black households to live with black households remains informative. Because it is the differences in estimated preferences that drive the equilibrium predictions of the model, our inability to distinguish centralized discrimination from decentralized preferences does not seriously affect a key aim of our simulation analysis, namely to gauge the impact of racial factors as a whole on consumption of local public goods by households of each race.

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 1. 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.²⁸ 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,²⁹ 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 captures the spatial aspect of the housing market by allowing households to have preferences over commuting distance.³⁰

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 1 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 7 and results related to heterogeneity in MWTP are shown in Table 8. The first three columns of Table 7 report the implied measures of the mean MWTP for housing and neighborhood attributes that result for three specifications of the mean

²⁸ 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.

²⁹ 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.

³⁰ 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.

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.³¹ 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.³²

Table 8 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

³¹ 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.

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.³³ 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.

Importantly, the equilibrium predictions of the model concerning segregation patterns 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). Looking at the difference between what whites versus households in the other race categories are willing to pay for these changes, Asian-White and Black-White differences come to over \$100 per month for a 10 percent change, while Hispanic-White differences amount to \$70 per month. Table 8 also shows similar figures calculated for households at a higher income level (income=\$120,000) in this case Asian-White, Black-White and Hispanic-White differences each remain near \$90 per month. Thus, strong segregating forces in the housing market are relevant at all income levels for all races.³⁴

³² Comparison of our parameter estimates with analogous hedonic price regressions provides further support for their plausibility. We carry out this comparison in a brief Hedonics Appendix.

³³ We discuss the implications of centralized discrimination in the housing market for the interpretation of these estimates in the Hedonics Appendix below.

³⁴ The strong segregating 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 segregating preferences (i.e., preferences for others like oneself), as both high- and low-income households are willing to pay for higher income neighbors.

8 GENERAL EQUILIBRIUM SIMULATION

We now use the estimated parameters to conduct a general equilibrium simulation designed to characterize the housing market equilibrium in a world in which the race of one's neighbors does not directly affect the residential choice problem. 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 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.

Simulation Details.^{35,36} 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

³⁵ It is also important to point out that because the model itself does not perfectly predict the housing choices that individuals make, the neighborhood sociodemographic measures initially predicted by model, $Z_n^{PREDICT}$, do not match the actual sociodemographic characteristics of each neighborhood, Z_n^{ACTUAL} . Consequently, before calculating the new equilibrium for any simulation, we first solve for the initial prediction error associated with each neighborhood n : $\omega_n = Z_n^{ACTUAL} - Z_n^{PREDICT}$. We add this initial prediction error ω_n to the sociodemographic measures calculated in each iteration before substituting these measures back into the utility function.

³⁶ 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. This simulation, however, removes the central component of these interactions,

aggregating these choices to the neighborhood level, compute the predicted sociodemographic 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 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.³⁷ 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 the adjustments implied by an OLS regression of the crime rate and average test score on neighborhood sociodemographic

sorting along racial lines, and, consequently, the presence of multiple equilibria for this particular simulation does not concern us.

³⁷ 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

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 reasonably tight range for the predictions from our simulations.

Simulation Results. Table 9 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 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 local public goods by households of each race. The remaining columns

provision of public goods in California, which gives local governments almost no control over taxes or the level of spending.

of Table 9 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.³⁸ 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 school quality and public consumption 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.

³⁸ We also note that the elimination of racial interactions leads to an overall reduction in commuting distances for all households except Asians; without needing to adjust their location decisions for race-related reasons, households are

These results indicate that race plays a profound role in shaping the equilibrium matching of households to neighborhoods in an urban housing market. As the consumption patterns of Tables 5 and 6 have already suggested, because blacks make up only about 9 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 9 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 public good consumption by blacks in the actual Bay Area equilibrium are at least partially compensated with lower housing prices. However, given the difficulties in obtaining substitute sources of supply of these amenities, racial sorting in the housing market generally works to strengthen the persistence of intergenerational racial differences in educational attainment, income, and wealth by accentuating differences in the consumption of important local public goods that contribute to these life outcomes through the formation of human capital and social networks.³⁹

able to more easily find suitable locations in other dimensions.

³⁹ 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.

9 CONCLUSION

This paper has studied the consequences of residential segregation from a new perspective, recognizing that the types of neighborhoods available in a metropolitan housing market are endogenously determined, governed by the characteristics and preferences of households who reside in the area.

In the paper, we advance and test a new hypothesis: that racial sorting in the housing market serves to lower the consumption of neighborhood amenities by black households, especially those with moderate to high incomes. This hypothesis is motivated by the empirical regularity that in many US cities, neighborhoods combining a high fraction of black households and even moderate levels of average income and education are in very short supply. This constraint has the effect of raising the implicit price that black households pay for school quality, public safety, neighborhood education and income, given that in order to consume more of these other important neighborhood attributes, households are typically required to live in a neighborhood with fewer blacks. In turn, many high-income blacks live in neighborhoods with high fractions of other black households, giving up substantial amounts of consumption of local public goods and average US neighborhood socioeconomic characteristics to do so. Our conjecture is that, were race eliminated as a factor driving location decisions, then consumption of these amenities by many black households would increase.

We have presented the first evidence in the literature relating to this conjecture, based on a simulation of an equilibrium model of residential sorting that we estimate using data on almost a quarter of million households in the San Francisco Bay Area. In estimating the model, we are careful to use reasonable variation in the data that addresses the correlation of neighborhood sociodemographic characteristics with unobserved housing and neighborhood quality. The

model in combination with the estimated preference structure then provides a powerful analytical tool for carrying out a general equilibrium counterfactual experiment that eliminates race as a factor in driving household residential location decisions.

Our results show that racial sorting in the housing market (whether driven by preferences directly or discrimination) leads to significant reductions in the consumption of neighborhood amenities by all blacks, and large reductions in the consumption of housing amenities by high-income blacks. As a consequence, racial sorting in the housing market accentuates differences in the consumption of neighborhood amenities that arise as the result of racial differences in income and wealth. Not only do these consumption differences have immediate welfare implications; they are also likely to be a powerful force that works to slow racial convergence in education, income, and wealth over time.

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, Fernando Ferreira, and Robert McMillan (2003), "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," unpublished manuscript, Yale University.

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

- 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 D. Sjoquist, (1990), Job accessibility and racial differences in youth employment rates, *American Economic Review*, 80, 267-276.
- 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
- 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, (2004) "Testing the Spatial Mismatch Hypothesis Using Inter-city Variations in Industrial Composition," *Regional Science and Urban Economics*, 34(5): 505-32.
- Weinberg, Bruce, (2000) "Black Residential Centralization and the Spatial Mismatch Hypothesis," *Journal of Urban Economics*, 48(1): 110-34.

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

All Tracts	Percent College Degree or More			
	<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%
Percent Black	<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%
Percent White	<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

Quintile	1	2	3	4	5	Total
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

Quintile	1	2	3	4	5	Total
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				
	242,100		actual sample	high test score side*	low test score side*	t-test for	weighted sample
	(1) Mean	(2) S.D.	(3) Mean	(4) Mean	(5) Mean	difference in means ((4) versus (5))	(6) Mean
<u>Housing/Neighborhood Characteristics</u>							
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
<u>Household Characteristics</u>							
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*

Quintile	1	2	3	4	5	Total
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*

Quintile	1	2	3	4	5	Total
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: 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*

Quintile	1	2	3	4	5	Total
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*

Quintile	1	2	3	4	5	Total
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.

Table 7 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
Boundary Fized Effects						
Observations	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 8. 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
Mean MWTP	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)
<i>Race (at mean income=\$54,755)</i>								
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
<i>Race (at income=\$120,000)</i>								
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
<i>Education</i>								
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 9: Counterfactual - Eliminating Racial Interactions in Location Decision

Panel A: Overall Results	<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>
Black										
Pre-Simulation	0.379	0.424	0.40	4.50	458	18.73	740	9.89	37,377	0.31
Post-Simulation - Unadjusted	0.147	0.646	0.35	4.53	502	12.11	868	9.15	44,711	0.38
Post-Simulation - Adjusted	0.146	0.645	0.35	4.53	496	13.01	846	9.05	44,356	0.37
White										
Pre-Simulation	0.048	0.761	0.63	5.36	541	6.14	1160	10.28	57,624	0.47
Post-Simulation - Unadjusted	0.071	0.725	0.63	5.31	531	7.75	1126	9.86	55,665	0.45
Post-Simulation - Adjusted	0.071	0.725	0.63	5.31	531	7.71	1127	9.88	55,696	0.45
Panel B: Results by Income Quartile	<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>
Black - Lowest Income Quartile										
Pre-Simulation	0.467	0.302	0.31	4.08	445	21.29	612	9.76	33,093	0.28
Post-Simulation - Unadjusted	0.172	0.610	0.22	3.94	490	14.06	689	8.94	39,858	0.34
Post-Simulation - Adjusted	0.172	0.609	0.22	3.94	482	15.30	661	8.82	39,258	0.32
Black - Highest Income Quartile										
Pre-Simulation	0.24	0.516	0.72	5.96	502	11.53	1,261	10.00	53,053	0.43
Post-Simulation - Unadjusted	0.108	0.671	0.76	6.27	535	6.98	1,490	10.09	60,295	0.50
Post-Simulation - Adjusted	0.107	0.671	0.76	6.26	533	7.29	1,477	10.02	60,365	0.50
White - Lowest Income Quartile										
Pre-Simulation	0.058	0.718	0.46	4.36	521	8.42	838	9.88	46,710	0.40
Post-Simulation - Unadjusted	0.068	0.695	0.45	4.31	508	10.78	788	9.46	44,893	0.37
Post-Simulation - Adjusted	0.068	0.696	0.45	4.30	506	10.95	786	9.47	44,732	0.36
White - Highest Income Quartile										
Pre-Simulation	0.026	0.807	0.87	6.56	565	4.18	1,607	11.16	71,805	0.56
Post-Simulation - Unadjusted	0.062	0.716	0.87	6.50	559	5.03	1,585	10.71	69,319	0.55
Post-Simulation - Adjusted	0.063	0.716	0.87	6.51	559	4.93	1,586	10.73	69,387	0.55

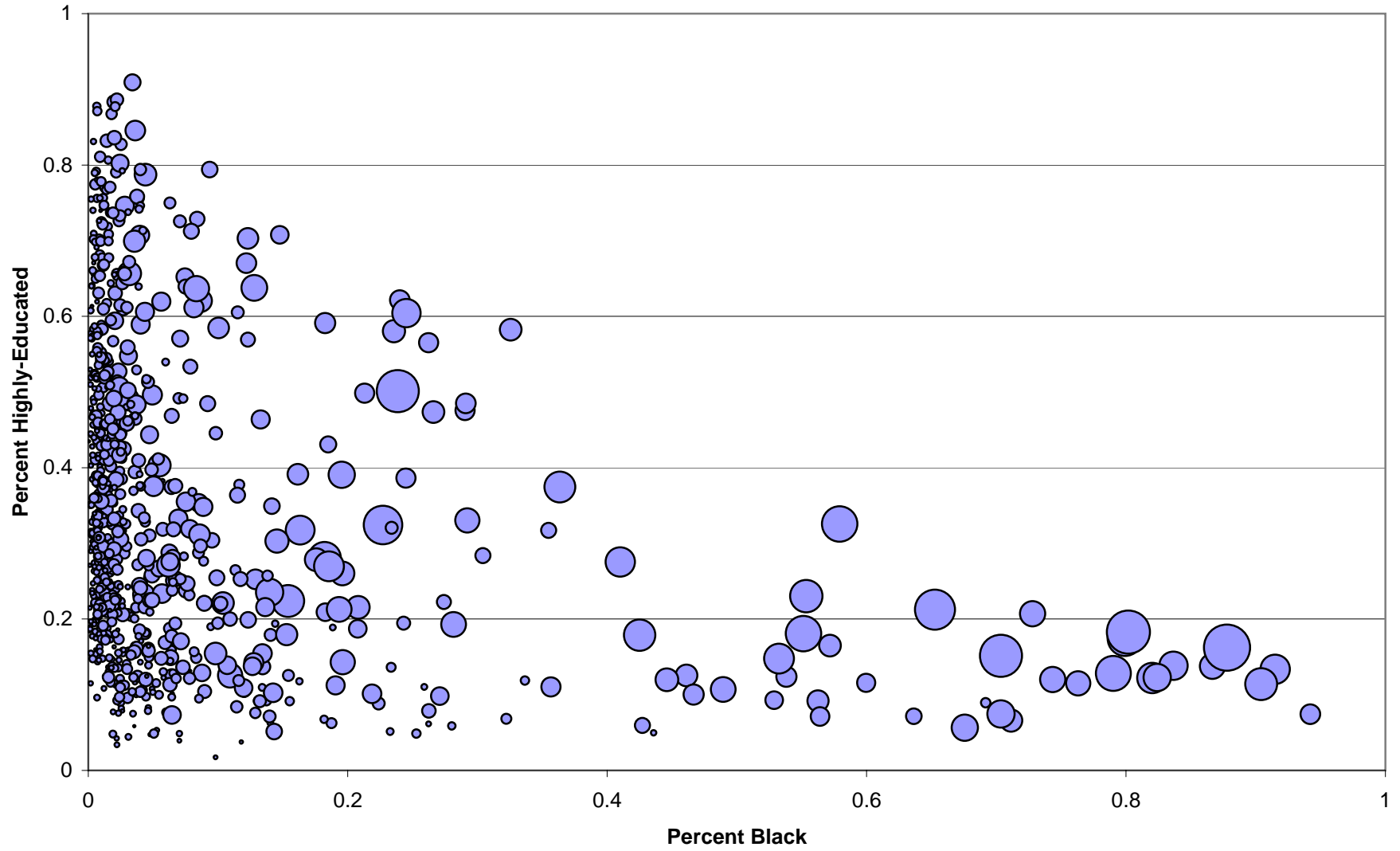
Note: This table reports the consumption of housing and local public goods by households of each race. Numbers are reported for a counterfactual simulation that sets all preference parameters associated with neighborhood racial composition to zero.

Appendix Table 1: Interaction Parameter Estimates

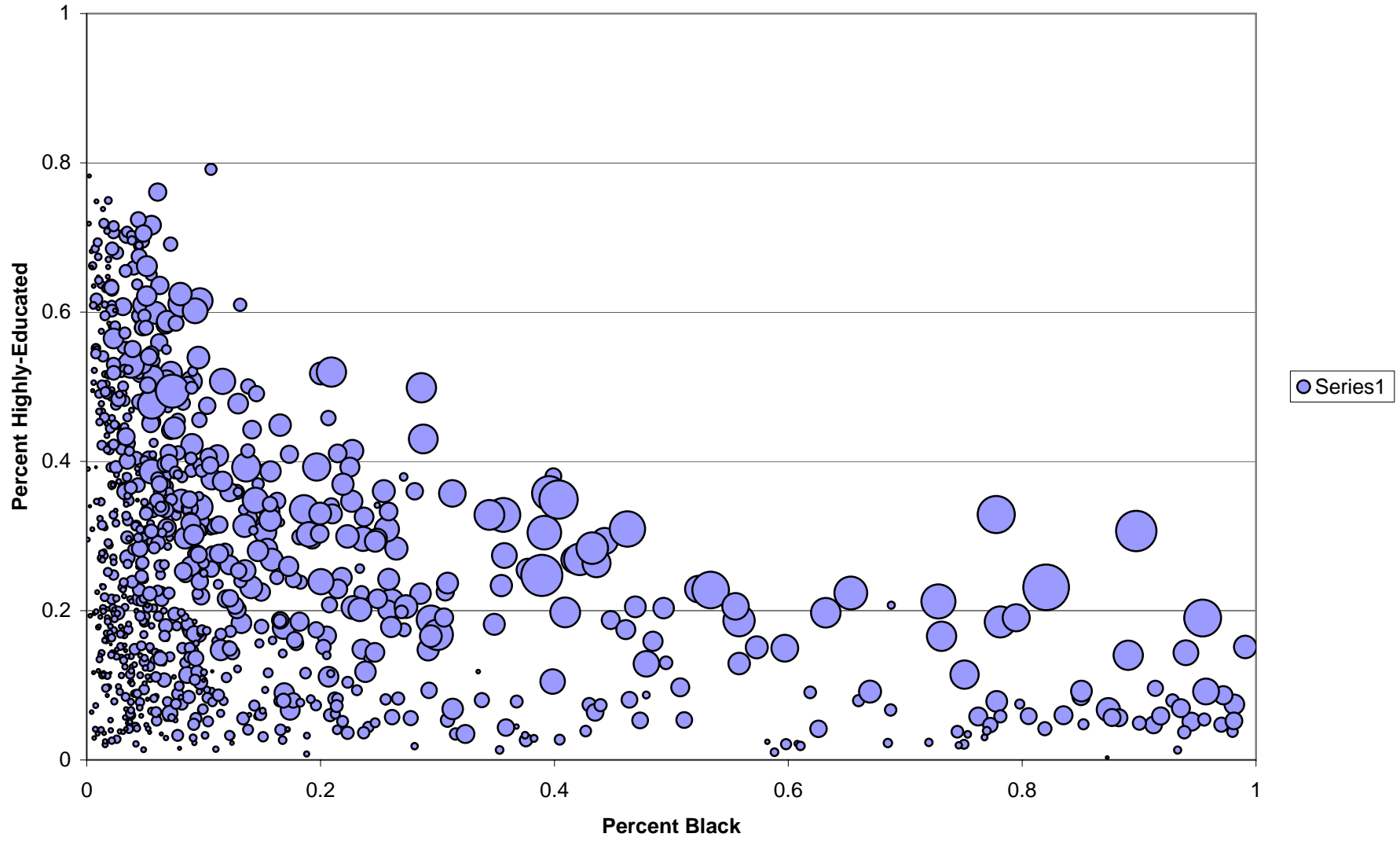
Housing/Neighborhood Attribute	Household Characteristic												
	Hhld Income	Children Under 18	Black	Hispanic	Asian	Some College	College Degree or More	Working	Age	Hhld Capital Income	Black* Hhld Income	Hispanic* Hhld Income	Asian* Hhld Income
Monthly House Price	0.071 (0.003)	0.071 (0.023)	0.087 (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 (0.030)	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.

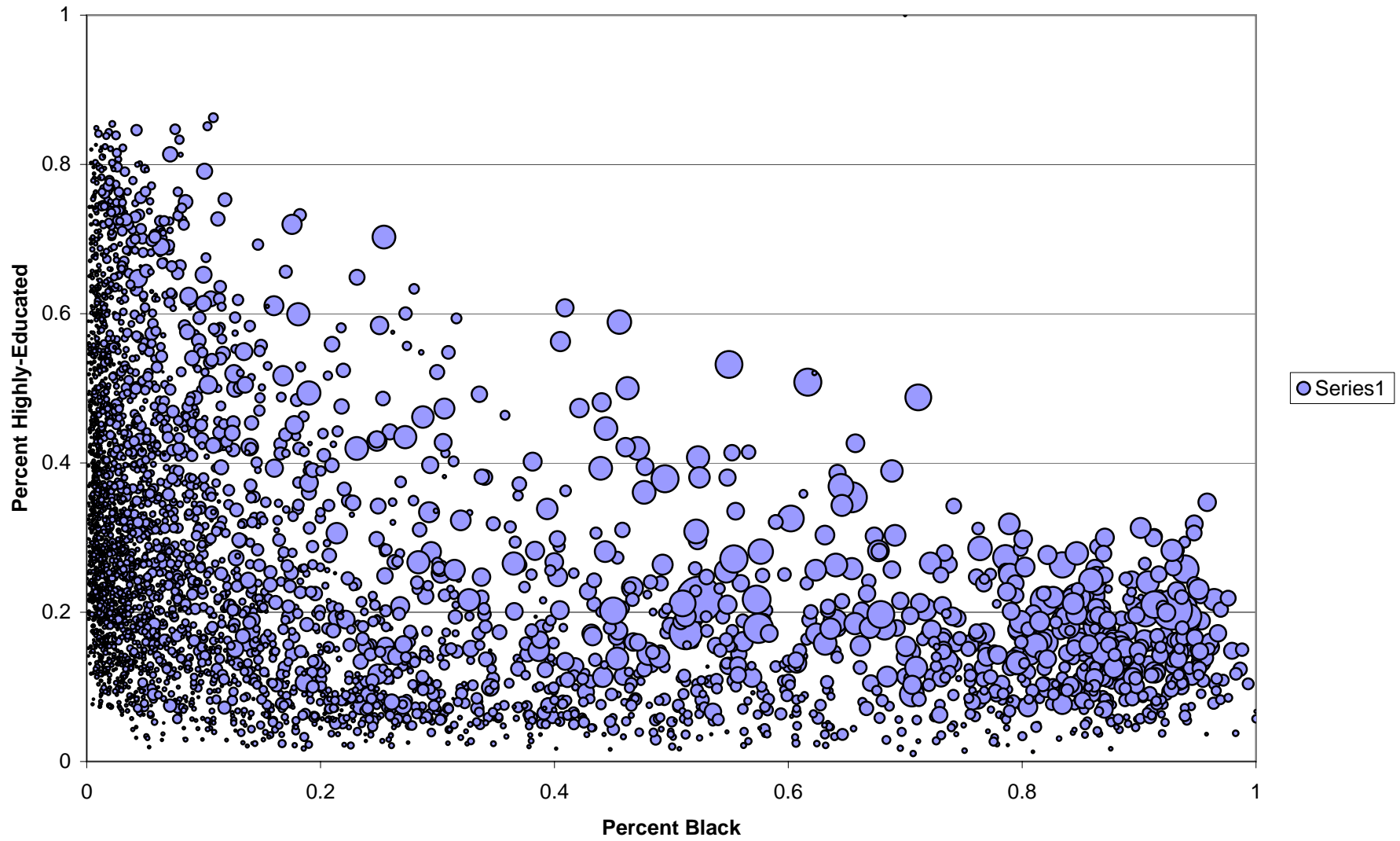
Boston



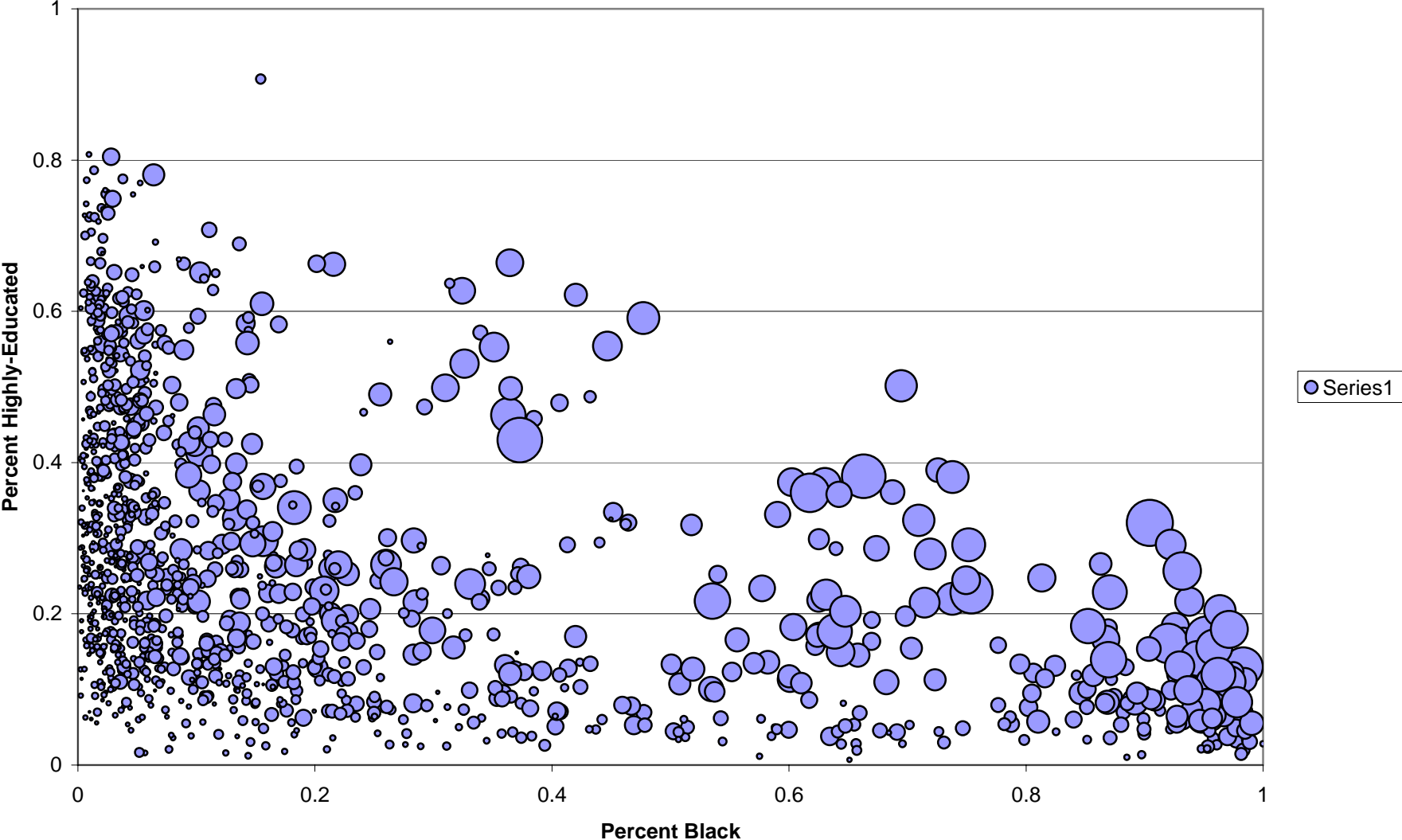
Dallas



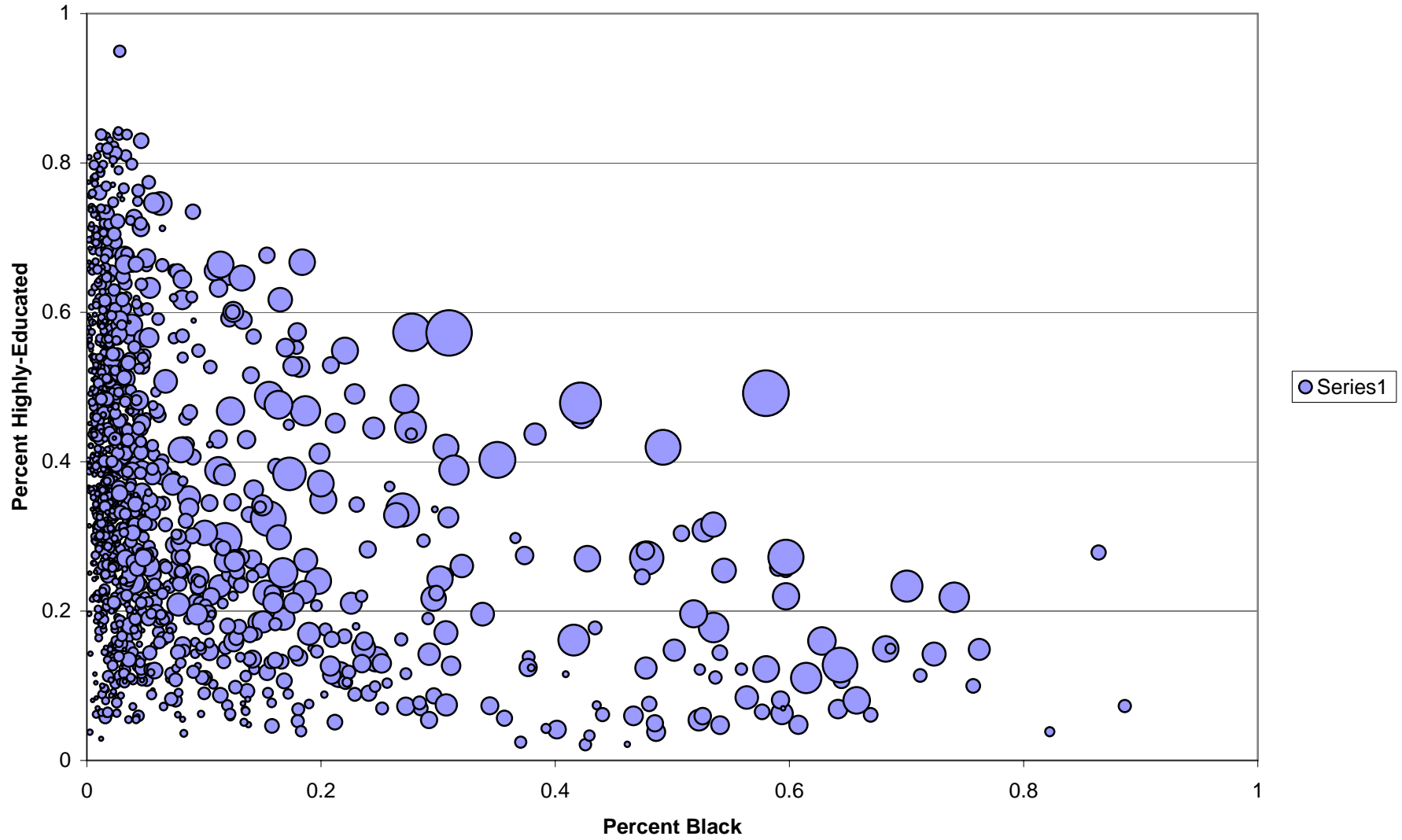
New York



Philadelphia



San Francisco-Oakland



Washington-Baltimore

