Racial Segregation and the Black-White Test Score Gap

David Card
Department of Economics
University of California Berkeley
and NBER

Jesse Rothstein
Department of Economics
Princeton University
and NBER

September 2004

We are grateful to the Andrew Mellon Foundation and the College Board for assistance in obtaining the SAT data used in this study, and to Jacob Vigdor, Jon Guryan, and Sarah Reber for supplying other data. We thank Florence Neymotin and Ashley Miller for outstanding research assistance, and Nicole Fortin, Thomas Lemieux, and Justin McCrary for helpful comments and suggestions.

Racial Segregation and the Black-White Test Score Gap

David Card & Jesse Rothstein

Abstract

The impact of school segregation on student outcomes is a central concern of education policy. We compare within-city black-white test score gaps across metropolitan areas that differ in the extent of school segregation. More segregated cities have robustly larger black-white gaps in average SAT scores and in other measures of educational achievement. Using an estimator derived from "control function" approaches to instrumental variables models, we decompose the effects of school segregation into components attributable to differences in residential sorting patterns, to court-ordered school desegregation, and a residual. The residential component has a large effect on test score gaps, but the remaining components do not. These results do not appear attributable either to endogeneity of residential segregation or to differences in school resources. Our results therefore suggest that the composition of neighborhoods matters, but not that of schools. We uncover evidence, however, that non-residential school segregation is negatively associated with black-white differences in assignment to advanced-track courses, suggesting that within-school segregation may offset much of the variation in non-residential, across-school segregation.

The racial segregation of schools has been a central concern of education policy since the Supreme Court's 1954 *Brown v. Board* decision. Despite the *Brown* decision and several decades of desegregation efforts, U.S. schools remain relatively segregated. Currently, over half the schoolmates of a typical black student are black, compared to less than 10 percent for a typical white student (Orfield and Lee, 2004). Concerns over the impact of segregated schooling are reinforced by the correlation between average test scores and the fraction of black schoolmates, a relationship first documented in the influential Coleman Report (Coleman 1966). As illustrated in Figure 1, the same pattern holds today: average SAT scores are 250 points lower at all-black than at all-white high schools.[1] About half of this difference is a pure composition effect, reflecting black-white test score gaps among students at the same schools. Nevertheless, there is also a strong negative correlation between the average scores of black students and the share of black schoolmates.

Whether this correlation represents the *causal* effect of a school's racial composition is unclear, however, because students who attend schools with different black enrollment shares are different.[2] One approach to the inference problem is to assume that even though students of different abilities may attend different schools within a city, the overall distribution of student ability (conditional on race) is similar across cities. This idea underlies the analysis of Evans, Oates, and Schwab (1992), who use inter-city variation in racial and ethnic composition to identify peer group effects.[3] In this paper we propose a simple extension of this research design, based on the achievement gap between black and white students within each city. Comparisons of black students *relative* to whites control for unobserved variables that may be correlated with the fraction of black

[1] This figure is based on SAT takers who planned to attend college starting in 1998-2002 in states with relatively high SAT participation rates.
[2] On the general problem of inferring peer group effects from observational data, see Manski (1993) and Brock and Durlauf (2001).
[3] Hoxby (2000) similarly assumes that differences across metropolitan areas in unobserved student background characteristics can be ignored. Hsieh and Urquiola (2003) and Rothstein (2004) focus on the within-area sorting process.

students in a city, provided that these variables affect blacks and whites equally. Our primary explanatory variable is the difference between the average fraction of black schoolmates at schools attended by blacks and whites – a standard measure of racial segregation in a city's schools. We also control for many observed city characteristics, including black-white differences in family background, the fraction of black students, and region of the country. We apply this methodology to a new micro data set comprised of over one quarter of students who wrote the SAT between 1998 and 2001 from states with high SAT participation rates.

A majority of inter-city variation in the relative exposure of black and white students to black schoolmates is driven by differences in residential sorting (Cutler and Glaeser 1997; Cutler, Glaeser et al. 1999). Nevertheless, cities also differ in the extent to which local policies (e.g., district boundaries and magnet school programs) enhance or reduce the impact of residential segregation. We develop a simple decomposition of observed school-level segregation into a component attributable to residential segregation and a non-residential component. We further decompose the non-residential component into a part attributable to the pressure of legal desegregation plans adopted in the 1970s and early 1980s, and a remaining "residual" component. Comparisons of the impacts of these three components provide insights into the plausibility of alternative explanations for the link between school segregation and student achievement.

Our analysis leads to three main conclusions. First, controlling for a rich set of individual and peer group characteristics, the black-white test score gap in a city is negatively correlated with the black-white difference in exposure to black schoolmates. Our estimates imply that moving from a metropolitan area with a low degree of school segregation (like Charlottesville, Virginia) to one with a high degree of segregation (like Cleveland) widens the black-white test score gap by roughly 40 points. Second, virtually all of this effect arises through the residential component of school segregation. Deviations in school segregation from the level predicted by the degree of residential segregation have small and insignificant effects. Third, when we decompose the non-residential component of school segregation into a part attributable to court-ordered desegregation programs, and a residual component, we find that neither component has a large or systematic effect on the relative achievement of black students.[4]

A potential concern with our use of SAT scores to measure student achievement is selective test participation. We address this issue by excluding data from states where most college-bound students take the ACT test (Clark 2003), and by including a selection correction based on observed SAT participation rates (Dynarski 1987; Card and Payne 2002). We also estimate a series of models using "trimmed" samples that, under certain assumptions, provide a bound on the effects of school segregation (Manski 1990). Most importantly, we estimate a series of parallel models using 2000 Census data on the enrollment, employment, and high school graduation rates of 16-24 year olds. Like the test score results, these models show that the relative outcomes of black youths are negatively affected by the residential component of school segregation, but are insignificantly related to the non-residential component. The similarity between these findings and those from our SAT-based analysis suggests that selective test participation is not driving our main results.

An examination of alternative explanations for the contrasting impacts of residential and non-residential school segregation leads us to consider the link between desegregation efforts and **within-school** segregation.[5] We analyze this connection using data on the number of honors

[4] Guryan (forthcoming) shows that the implementation of a major desegregation program during the 1970s was associated with a modest but statistically significant 3 percentage point reduction in black dropout rates relative to whites, suggesting that policy-induced changes in school-level racial segregation had an impact on black students' relative performance. Our approach, unfortunately, does not have the power to distinguish effects of the size that Guryan finds, although our point estimates are not even of the right sign to be consistent with his results.
[5] There is growing recognition of the issue of within-school segregation. Cohen (2004) provides a recent account of the phenomenon at Evanston Township High School. A 1993 suit filed in Rockford Illinois led to a court order to integrate the **courses** offered at middle and high schools in the district (Weiler, 2004). See also the discussions in

classes taken by black and white SAT-takers, and teacher reports of the incidence of ability tracking (from the Schools and Staffing Survey). Inter-city differences in the degree of relative school segregation attributable to differences in residential segregation have no significant effect on the black-white gap in honors course-taking. In contrast, there is strong positive relationship between the non-residential component of school segregation and the gap in honors course-taking, driven by higher honors participation by white students in cities with higher non-residential desegregation. This pattern suggests that even though busing, magnet schools, and other policies may reduce racial segregation across schools, they are associated with a countervailing rise in within-school segregation that limits the exposure of black students to white schoolmates and may ultimately weaken the intended effects of the programs.

II. Empirical Framework

a. Basic Model

Our empirical analysis is based on a simple causal model of the form:

(1) $$Y_{ijsc} = X_{ijsc}\alpha_j + X_{jsc}\beta_j + B_{sc}\gamma_j + u_{jsc} + \varepsilon_{ijsc} ,$$

where Y_{ijsc} represents an achievement measure (e.g. SAT score) for student i of race group j in school s and city c, X_{ijsc} is a vector of observed characteristics of the student, X_{jsc} is a vector of average characteristics of students in race group j in school s in city c, B_{sc} represents the fraction of black students in school s, u_{jsc} is a common error component for students of group j in school s and city c, and ε_{ijsc} is an individual-level error (with mean 0 for each race group in each school). The key parameters of interest in this model are the coefficients γ_j, which measure the effects of the exposure to black schoolmates on student achievement of the two race groups.

Clotfelter, Ladd, and Vigdor (2003) and in Clotfelter (2004).

There are several channels through which the racial composition of schoolmates could affect student achievement, leading to a model like (1). Perhaps the simplest is a peer group quality effect. Since black test scores are generally lower than white scores (Jencks and Phillips 1998), a rise in the fraction of black students lowers the mean "test potential" of the student body. If individual achievement is affected by the average test potential of classmates, an increase in the fraction of black students will lower the expected performance of students (holding constant their own characteristics).[6] A closely related channel is the establishment of norms or standards of achievement (Ogbu and Fordham 1986; Jencks and Phillips 1998). If black students value academic achievement less than whites, as claimed by Ogbu (2003), and if individual student performance is affected by peer norms, then a rise in the fraction of black classmates will lower achievement. A third channel is school quality. Boyd et al. (2003), for example, argue that teachers prefer to teach in schools with lower minority enrollment, and that schools compete for more effective teachers, which might lead to an equilibrium in which better teachers are assigned to schools with fewer black students.

Regardless of the specific channel, the key problem in estimating a model like (1) is that non-randomness in the assignment of students to schools may lead to a correlation between the average unobserved characteristics of students in a given school, u_{jsc}, and the black share of students at the school, B_{sc}. By aggregating to the city level and differencing between race groups we can eliminate the effects of sorting, provided that differences in school segregation across cities are uncorrelated with differences in the gap in unobserved characteristics between black (j=1) and white (j=2) students in a city. Specifically, equation (1) implies:

[6] Simplifying notation slightly , this model could be represented as: $Y_{is} = x_{is}\alpha + (x_s\alpha)\lambda + \varepsilon_{is}$, where x_s is a vector of the mean characteristics of the student body at school s. Here, student i's achievement depends on her own characteristics though the index of test potential $x_{is}\alpha$, and on the average of the same index of characteristics of her schoolmates.

(2) $Y_{1c} - Y_{2c} = X_{1c}(\alpha_1 + \beta_1) - X_{2c}(\alpha_2 + \beta_2) + (B_{1c}\gamma_1 - B_{2c}\gamma_2) + u_{1c} - u_{2c}$,

where Y_{jc} is the average outcome for students of race j in city c; X_{jc} is the corresponding average of X_{ijsc}; u_{jc} is the average of the race-group specific school-level residuals u_{jsc} in city c; and B_{jc} is a **group-specific** measure of the average exposure to black schoolmates in city c,[7]

(3) $B_{jc} = N_{jc}^{-1} \Sigma_s N_{jsc} B_{sc}$.

Estimation of equation (2) will provide consistent estimates of the γ coefficients if (and only if) B_{1c} and B_{2c} are uncorrelated with $(u_{1c} - u_{2c})$, conditional on X_{1c} and X_{2c}. If exposure to black schoolmates has the same effect on blacks and whites (i.e., $\gamma_1 = \gamma_2$) then the necessary assumption is somewhat weaker, that $(B_{1c} - B_{2c})$ is uncorrelated with $(u_{1c} - u_{2c})$ conditional on X_{1c} and X_{2c}.

Although equation (2) can be estimated directly on city-level data, such an approach ignores the substantial micro-level variation in the individual X's. Moreover, with only 200-300 cities, it is difficult to estimate flexible city-level models for the effects of family background characteristics on student outcomes. To use the data more efficiently, we proceed in two steps, first estimating the α_j coefficients from an individual-level model that includes a flexible function of the observed family background variables (fully interacted with race) and unrestricted school fixed effects (to absorb the contributions of X_{jsc}, B_{sc}, and u_{jsc}).[8] Using these first stage coefficients, we then construct an adjusted test score for each student:

$$r_{ijsc} = Y_{ijsc} - X_{ijsc}\,\hat{\alpha}_j.$$

We then form the average adjusted scores for students in race group j in a city (r_{jc}), and take the difference between blacks and whites in the same city, leading to

---

[7] In the segregation literature (e.g. Massey and Denton 1988; Iceland, Weinberg et al. 2002) B_{1c} and B_{2c} are known as the indices of exposure of black and white students to blacks, respectively.

[8] The micro samples are very large so the $\hat{\alpha}_j$'s are estimated quire precisely. We include only a single fixed effect for each high school, however, implicitly restricting the effects of the school level peer group variables to be constant across races. We also make no adjustment for the selectivity of test score participation in the first stage model, presumably leading to some attenuation in the α coefficients.

(4) $r_{1c} - r_{2c} = X_{1c}(\beta_1 + \alpha_1 - \hat{\alpha}_1) - X_{2c}(\beta_2 + \alpha_2 - \hat{\alpha}_2) + B_{1c}\gamma_1 - B_{2c}\gamma_2 + u_{1c} - u_{2c}$.

This model has the same functional form as equation (2), but now (assuming that sampling errors in $\hat{\alpha}$ can be ignored) the coefficients on the city-level averages of the X's represent the group-level effects, rather than the sum of the individual and group effects.

Although differencing the adjusted test scores of black and white students eliminates any variables that have equal effects on the potential outcomes of the two race groups, it is still possible that the degree of racial segregation is correlated with the black-white difference in unobserved characteristics of the families or schools in a city. We posit that this difference can be decomposed as:

$$u_{1c} - u_{2c} = Z_c \theta + v_c ,$$

where Z_c is a set of city characteristics (including Census division effects and the overall fraction of black students in the city) and v_c represents all remaining unobserved differences between black and white students in city c. This leads to a model of the form:

(5) $r_{1c} - r_{2c} = X_{1c}\beta_1 - X_{2c}\beta_2 + B_{1c}\gamma_1 - B_{2c}\gamma_2 + Z_c \theta + v_c$.

This specification yields consistent estimates of the γ coefficients if (and only if) v_c, the unexplained difference in black-white test potential, is uncorrelated with B_{1c} and B_{2c}, conditional on X_{1c}, X_{2c}, and the variables included in Z_c.

b. Adjusting For Selective Participation in the SAT

While average SAT scores are widely-cited measures of high school performance, not all students write the test, leading to potential selectivity biases in observed mean test scores. All of our analyses of the SAT data use only observations from the states with the highest SAT-taking rates, where nearly all college-bound seniors take the SAT (rather than a competing test, the ACT, which is dominant in other states). In cities in SAT states, we assume that the true model for the adjusted test score of race group j in city c is

(6) $r_{jc} = X_{jc}\beta_j + B_{jc}\gamma_j + u_{jc}$

and that the mean observed score when a fraction p_{jc} of the group write the test is given by:

(6') $r_{jc} = X_{jc}\beta_j + B_{jc}\gamma_j + u_{jc} + \varrho_j \lambda(\Phi^{-1}(p_{jc}))$,

where λ( ) is the inverse Mills ratio function, Φ is the standard normal c.d.f., and ϱ_j is an unknown coefficient. This specification is appropriate if ε_{ijsc} (the unobserved ability of student i in group j at school s in city c) is normally distributed, and if the probability that i writes the test depends on whether a latent normal variable that is correlated with ε_{ijsc} passes a threshold that varies by group and city (Heckman 1979; Card and Payne 2002).[9] Our estimating model becomes

(7) $r_{1c} - r_{2c} = X_{1c}\beta_1 - X_{2c}\beta_2 + B_{1c}\gamma_1 - B_{2c}\gamma_2 + \varrho_1 \lambda(\Phi^{-1}(p_{1c})) - \varrho_2 \lambda(\Phi^{-1}(p_{2c})) + Z_c\theta + v_c$

or, in the simplified case in which the coefficients are the same for blacks and whites:

(7') $r_{1c} - r_{2c} = (X_{1c} - X_{2c})\beta + (B_{1c} - B_{2c})\gamma + \varrho(\lambda(\Phi^{-1}(p_{1c})) - \lambda(\Phi^{-1}(p_{2c}))) + Z_c\theta + v_c$.

A key assumption underlying our correction procedure is that the correlation between a student's potential test score and the latent variable determining whether or not she writes the test is constant across cities. This assumption may fail if differences in the way students are sorted to schools lead to differences in the way test participation rates vary with ability in different cities.[10]

c. Decomposing School Segregation into Residential and Other Components

---

[9] Simplifying notation slightly, assume that the adjusted outcome of student i in city c is given by $r_{ic} = X_c\beta + u_c + \varepsilon_{ic}$, where ε_{ic} is normally distributed, and assume that i writes the test if $\delta r_{ic}+\zeta_{ic} \geq k_c$, where ζ_{ic} is another normally distributed error and k_c is some city-specific threshold. In this model the test participation rate in city c is $p_c = \Phi[h(k_c, X_c, u_c; \beta, \delta, \sigma)]$, where h( ) is a known function and σ includes the variances and covariance of ε_{ic} and ζ_{ic}. Moreover, $E[\varepsilon_{ic} |$ i writes the test$] = \varrho \lambda[h(k_c, X_c, u_c; \beta, \delta, \sigma)]$, where λ is the inverse Mills ratio function and ϱ is a function of the elements of σ. Since Φ is invertible, $E[\varepsilon_{ic} |$ i writes the test$] = \varrho \lambda[\Phi^{-1}(p_c)]$. The expression in the text follows.

[10] An alternative correction can be developed using school-level participation rates if the correlation between potential test scores and the latent variable determining test participation is constant across schools. We believe this assumption is unappealing, especially for black students.

Intercity differences in school segregation are driven by two main factors: the extent of residential segregation in the city; and institutional features of the school system, such as the number of school districts in the city (Urquiola 1999; Rothstein 2004), the private enrollment share, the design of school catchment areas, and the use of busing and magnet school programs. To understand the links between the black-white achievement gap and school segregation, it is useful to decompose the level of school segregation into a residential component and a remainder, and to compare the relative impacts of these two sources of school-level segregation. Building on equation (7'), we use a simple decomposition method that is formally equivalent to a conventional instrumental variables (IV) approach.

The first step (equivalent to a first stage model in an IV approach) is to project relative school segregation in city c on relative residential segregation in the city and all the other covariates in the main estimating model:

(8) $\Delta B_c = B_{1c} - B_{2c} = W_c \pi_w + (R_{1c} - R_{2c})\pi_R + \zeta_c$,

where W_c includes the difference in the observed characteristics of black and white schoolmates in city c, the difference in the inverse Mills ratio terms for blacks and whites, and the vector of city-level control variables, and R_{1c} and R_{2c} are measures of the average exposure of blacks and whites to black neighbors in the Census tracts in city c. From (8) we have an orthogonal decomposition:

$$\Delta B_c = \Delta \hat{B}_c + (\Delta B_c - \Delta \hat{B}_c) ,$$

where $\Delta \hat{B}_c = W_c \hat{\pi}_w + \Delta R_c \hat{\pi}_R$ is the predicted level of school segregation given $\Delta R_c = (R_{1c} - R_{2c})$.

In the second stage of our analysis we include the two terms $\Delta \hat{B}_c$ and $(\Delta B_c - \Delta \hat{B}_c)$ separately in the model for the black-white test score gap:

(9) $r_{1c} - r_{2c} = (X_{1c} - X_{2c})\beta + \Delta \hat{B}_c \gamma_1 + (\Delta B_c - \Delta \hat{B}_c)\gamma_2 +$

$$+ \varrho(\lambda(\Phi^{-1}(p_{1c})) - \lambda(\Phi^{-1}(p_{2c}))) + Z_c \theta + v_c .$$

It is well known that the estimate of γ_1 from equation (9) is numerically equal to the IV estimate that would be obtained using ΔR_c as an instrument for ΔB_c in (7').[11] The estimate of γ_2 (the coefficient on unexplained or "non-residential" segregation) may differ from the estimate of γ_1, depending on whether relative test scores are more strongly correlated with the residential or non-residential components of school segregation. In the standard IV framework, ΔR_c is assumed to be a valid instrument and there is a constant coefficient γ in the second stage model. In this setting, any discrepancy between $\hat{\gamma}_1$ and $\hat{\gamma}_2$ is interpreted as evidence of endogeneity bias arising from the correlation between the non-residential component of school segregation and the unobserved determinants of black relative test scores in a city. More generally, however, $\hat{\gamma}_1$ can differ from $\hat{\gamma}_2$ because the unobserved determinants of relative test scores are differentially correlated with the residential and non-residential parts of relative school segregation, or because the two components of relative segregation have different causal effects.

If measures of the institutional structure of the school system in each city are available, the non-residential component of school segregation can be further decomposed into a part attributable to the specific measured features, and a "residual" component. In this paper, we focus on the presence and strength of a court-ordered desegregation plan, as documented in Welch and Light (1987). To implement this extended decomposition, we first compute two "first stage" regressions, akin to equation (8):

(10a) $\Delta B_c = W_c\pi_w^1 + \Delta R_c\,\pi_R^1 + \zeta_c^1$,

(10b) $\Delta B_c = W_c\pi_w^2 + \Delta R_c\,\pi_R^2 + D_c\pi_C^2 + \zeta_c^2$,

---

[11] (9) is a simple re-arrangement of the usual control function estimator of the IV model (Garen 1984). OLS standard errors for (9) do not account for the sampling variability in the coefficient estimates $\hat{\pi}_w$ and $\hat{\pi}_R$ that are used to construct $\Delta\hat{B}_c$. We use a variance estimator that is consistent in this situation (Murphy and Topel 1985). The orthogonality of $\Delta\hat{B}_c$ and (ΔB_c - $\Delta\hat{B}_c$) in our version of the control function estimator greatly simplifies the computation. See the Technical Appendix for details.

where, as before, ΔR_c is the relative neighborhood segregation measure, W_c is a vector of control variables included in second stage model, and D_c is a measure of the strength of the court-ordered desegregation plan in city c. We then compute the predicted values $\Delta \hat{B}_c$ and $\Delta \tilde{B}_c$ from regressions (10a) and (10b), respectively, and estimate a second stage model of the form

$$(11) \quad r_{1c} - r_{2c} = (X_{1c} - X_{2c})\beta + \Delta \hat{B}_c\gamma_1 + (\Delta \tilde{B}_c - \Delta \hat{B}_c)\gamma_2 + (\Delta B_c - \Delta \tilde{B}_c)\gamma_3$$
$$+ \varrho(\lambda(\Phi^{-1}(p_{1c})) - \lambda(\Phi^{-1}(p_{2c}))) + Z_c \theta + v_c .$$

Here, γ_1 is the effect of the portion of school segregation that is attributable to residential segregation; γ_2 is the effect of the additional portion that can be predicted given both residential segregation and the legal desegregation measure; and γ_3 is the effect of the residual component. By construction $\Delta \hat{B}_c$, $(\Delta \tilde{B}_c - \Delta \hat{B}_c)$, and $(\Delta B_c - \Delta \tilde{B}_c)$ are mutually orthogonal. As a result, the estimate of γ_1 from the three part decomposition in equation (11) is the same as the estimate obtained from the two part decomposition in equation (9). The estimate of γ_2 from (11) is numerically equal to the IV estimate that would be obtained by instrumenting ΔB_c with D_c, while controlling for W_c and ΔR_c. Any divergence between $\hat{\gamma}_1$ and $\hat{\gamma}_2$ reveals the same information as a conventional over-identification test, and depends on whether the part of relative school segregation explained by residential segregation has the same effect on the black-white achievement gap as the part explained by the pressure of court-ordered desegregation.

III. Data Sources and Overview

Our primary data source consists of test scores and family background information for 100% of all black SAT-takers and 25% of all non-black SAT-takers in the 1998-2001 test cohorts.[12]

[12] We also have and use observations on 100% of white test takers in California and Texas. We use sampling weights in all analyses to account for this stratified sampling scheme. Demographic and family background information is

These files also contain an identifier for the high school attended by each test-taker. We match public school test-takers to school-level information from the Common Core of Data (CCD) for the appropriate school years, and private school test-takers to data from the 1997-8 Private School Survey (PSS). We also use geographic information in the CCD and PSS files to assign Metropolitan Statistical Area (MSA) identifiers, based on 2000 MSA definitions.[13] 92% of SAT-takers in SAT states (defined below) are successfully matched to schools; 87% of these attend school inside an MSA.

We use the CCD and PSS data to construct measures of school segregation for each MSA, and the 2000 Census Summary File 1 to construct residential segregation measures at the Census tract level. We also use the 2000 Census Public Use Microdata Samples to construct various family background measures for high-school age children in each MSA, and for our supplementary analyses of non-test outcomes. Finally, in some of our analyses, we use information on 1950-1990 residential segregation patterns from Cutler, Glaeser, and Vigdor (1999), and information from Welch and Light (1987) on major desegregration programs. More information on our data sources and merging methods is presented in the Data Appendix.

Table 1 gives an overview of the degree of racial segregation in housing and schooling in selected U.S. cities. In particular, we show data for the 10 cities with the lowest levels of neighborhood segregation (Panel A); the 10 cities with the lowest levels of neighborhood segregation among those with at least 10% blacks in the local population (Panel B); and the 10 most

---

collected in the Student Descriptive Questionnaire (SDQ) filled out by SAT-takers just before the test. We exclude observations for students who did not report their race/ethnicity in the SDQ or who reported an ethnicity other than white or black.

[13] Where the Office of Management and Budget designates a larger metropolitan area a Consolidated Metropolitan Statistical Area (CMSA) with several defined sub-areas (Primary Metropolitan Statistical Areas, or PMSAs), we treat the PMSA as the relevant city definition, but estimate standard errors that are robust to arbitrary correlation among PMSAs in the same CMSA.

residentially segregated cities (Panel C).[14] The first three columns of the table present residential segregation data, while the next three columns show school segregation measures based on data for the public and private high school students in each city. Finally, the three right-hand columns shows school segregation indices constructed for SAT-takers. These differ from the corresponding indices for all high school students to the extent that SAT takers are non-randomly drawn from different high schools. Since SAT participation rates are very low in some traditional "ACT-writing" states, we only report the SAT-taker averages for MSAs in states with higher participation rates.[15]

The 10 least-segregated MSAs are smaller cities (mainly in the West) with very low black population shares. In these cities high schools tend to be a little more segregated than neighborhoods, though the black-weighted and white-weighted fractions of black students are uniformly small. By comparison, the least segregated MSAs with substantial black population shares are all in the South. Even in this highly selective set of cities blacks are unevenly distributed across Census tracts, with at least a 10 percentage point gap between the fraction of blacks in the Census Tract of a typical black resident relative to that of a typical white resident. School segregation in these cities is uniformly lower than neighborhood segregation (compare column 3 and column 6). Finally, MSAs with the highest levels of residential segregation are larger cities in the Midwest, Mid-Atlantic, and South. Relative segregation of high schools in most of these cities is similar to that of neighborhoods, though there are some interesting exceptions, including Miami and New York, where high schools show substantially less relative segregation than neighborhoods. Comparisons of the school segregation measures for all high students and SAT takers show that the fraction of

[14] Let fb_t represent the fraction of black residents in Census tract t, let N_{bt} and N_{wt} represent the numbers of black and white residents in the tract, and N_b and N_w represent the numbers of black and white residents in the city. The black-weighted fraction black is $N_b^{-1} \Sigma_t N_{bt} fb_t$, and the white-weighted fraction black is $N_w^{-1} \Sigma_t N_{wt} fb_t$. We use the difference between these two to rank cities by degree of residential segregation.

[15] We define SAT states as those with SAT participation rates of 25% or more. The distribution is clearly bimodal, and there is little question that this cutoff accurately captures the states where the SAT is dominant.

black schoolmates at schools attended by black SAT-takers is a little lower than corresponding fraction for all black high school students, suggesting some selectivity in test participation. For white students the gap is much smaller.

The patterns in Table 1 suggest that there is wide variation across U.S. cities in the relative exposure of black and white students to black schoolmates. At the low end are MSAs like Bellingham, Washington (with about a 1.5% black population share), in which blacks are roughly evenly distributed across neighborhoods and schools. At the high end are MSAs like Gary, Indiana in which the black residents (19% of the population) and black students are concentrated in a few neighborhoods and schools, leading to a very wide gap in the relative exposure of blacks and whites to black schoolmates.[16] In between are some cities with a high degree of residential segregation but relatively more integrated schools (like Miami), as well as many cities with modest degrees of residential and school segregation.

As noted earlier, in this paper we confine our analysis of SAT outcomes to cities in states with relatively high participation rates. Table 2 presents some comparisons between the cities and students in all 331 MSAs (columns 1-2), the 189 MSAs in high SAT participation states (columns 3-4), and the 142 MSAs in low SAT participation states (columns 5-6). Cities in "SAT states" are larger, have slightly higher average income, a lower fraction of blacks, and more Hispanics than cities in non-SAT states.[17] They also have less racial segregation at both the neighborhood and high school levels. On average we estimate that 38 percent of white high school students and 27 percent of black high school students in cities in SAT states write the SAT.

---

[16] It is worth noting that Gary is one of the few large cities with a sizable black population share where there was never a successful lawsuit or consent decree mandating active desegregation of the schools (Welch and Light 1987).

[17] California, Texas, and Florida – three states with the highest fractions of Hispanics – are all SAT states. The averages in Table 2 and in the remainder of our analysis are weighted by $(1/N_b + 1/N_w)^{-1}$ where N_b and N_w are the numbers of blacks and whites in the city population. These weights are inversely proportional to the variance of city-level differences between black and white averages, computed over the entire population.

The bottom four rows in Table 2 show average SAT scores for the different city groups, the mean SAT gap between whites and blacks, the mean gap in the estimated family background contribution to SAT scores, and the mean gap in scores controlling for family background. Average SAT scores are lower in high-participation states (Dynarski 1987; Rothstein 2004), but the black-white difference in raw scores is very similar for cities in SAT and non-SAT states, suggesting that use of within-city differences may moderate problems associated with selective test participation.

As described earlier, we construct background-adjusted SAT scores for the students in our samples using the coefficients from a micro level model that includes unrestricted high school dummies and a flexible set of controls for parental education and income, fully interacted with race.[18] Specifically, we sum the predicted contributions of the family background variables for each test taker and subtract this total from his or her raw SAT score to derive our background-adjusted score. Because of the inclusion of the school dummies in the regression model, the mean adjusted score over all test-takers in a city is equal to zero. The means for whites and blacks in a city will in general differ from zero, however, because the two groups attend different schools, and because of any unexplained gaps between white and black scores within each school. Ignoring the latter effect, the black-white gap in adjusted scores is essentially the difference in a weighted average of the fixed effects for the high schools in a city, with separate weights for each race based on their relative enrollment shares at each school. The fact that the mean gap in adjusted scores is negative implies that black students tend to attend high schools with lower average test scores, conditional on the demographic composition of their students. Interestingly, the black-white gaps in the background indices and adjusted test scores are very similar for cities in SAT states and non-SAT states .

[18] The SDQ asks students about each parents' education (in 10 categories) and family income (in 14 categories). We use a full cross product of the parental education categories plus a full set of dummies for the income categories to yield 114 parental background variables, each interacted with race. Not all students complete the questionnaire. We restrict our analysis to students who report their race/ethnicity, though some of the 114 variables are missing data categories for students who report answers to some but not all of the questions.

As a final descriptive exercise, Figures 2a and 2b show the correlations across cities between the black-white test score gap and the relative segregation of black students. The horizontal axis in each figure is our relative segregation index, calculated for SAT takers using high school racial composition data from the Common Core of Data and Private School Survey. Consistent with Table 1, B_{1c} - B_{2c} calculated in this manner ranges from 0 to 0.7.[19] The vertical axis in Figure 2a shows the difference in mean test scores between whites and blacks, while Figure 2b shows the difference in background-adjusted scores. In both graphs the scatter of points suggests a systematic negative relationship between the relative segregation of black students and their relative test performance. The models in the next section probe the robustness of this simple correlation.

IV. Models for the Black-White Test Score Gap

a. Basic Models

We turn to the task of estimating equation (7'), using cross-city data for MSAs in high SAT participation states. A preliminary issue is how to estimate the average characteristics of schoolmates (i.e., X_{1c} and X_{2c}). The Common Core only provides data on the racial composition of schools, so we consider two alternatives. One choice is to use micro data from the 2000 Census to estimate the average characteristics of black and white teenagers in each MSA.[20] The other is to use the average characteristics of black and white test-takers. The advantage of the Census data is that the characteristics are measured for all students, not just those who write the SAT. The disadvantage is that city-wide averages from the Census ignore any non-randomness in the

---

[19] The MSA with the most segregated schools is Gary, Indiana. Newark, New Jersey is second.

[20] See the Data Appendix for details. In brief, we use the 5 percent public use sample of the 2000 Census. We select all children ages 13-17 and match them to their mother and father (if present). We then calculate mean characteristics of the parents of these children by MSA and race.

distribution of test takers across high schools.[21] Using the mean characteristics of SAT-takers automatically adjusts for the distribution of test-takers across schools, at the cost of some bias from the exclusion of data on non-SAT-taking schoolmates. Bias in either measure may not be such a serious problem if it is constant across cities, or if it is about the same for blacks and whites in a city, since in most of our specifications we control for the *difference* in the average characteristics of white and black schoolmates. Nevertheless, we present results using both the Census-based and SAT-taker averages as controls.

A second and closely related issue concerns which characteristics of schoolmates to include in X_{1c} and X_{2c}. A concern here is parsimony, since we only have 189 cities in our sample. As a one-dimensional summary of family background differences across cities we use the mean predicted test scores for black and white students from our first-step models (i.e., the mean values of $X_{ijsc}\hat{\alpha}_1$ and $X_{ijsc}\hat{\alpha}_2$). We refer these as "average student background indices." We also include a few selected main effects from the background index (for example, mother's years of schooling) to test whether some characteristics are relatively more important at the group level than at the individual level.

A third issue is whether a "first differenced" specification (i.e., equation 7') is appropriate, or if the β, γ, and ϱ coefficients for whites and blacks are different. Again, given the limited number of cities in our data set, we are concerned about including a large number of group-level X's with separate coefficients for whites and blacks. We therefore proceed by imposing the assumption of equal and opposite effects for most of the coefficients, though we explore specifications in which the restriction is removed for some coefficients. A final issue is what city-level variables to include in Z_c. After some experimentation, we narrowed the set to 15 variables: the fraction of black students at the average (white or black) SAT-taker's school, the fraction of Hispanics in the city, the

[21] We noted in the discussion of Table 1, for example, that the exposure of black SAT-takers to black schoolmates is a little lower than that of black students overall.

log of MSA population, the log of MSA land area, the fractions of the city population with 13-15 years of schooling and 16 or more years of schooling, the log of average household income, the Gini coefficient for household income, and 5 dummies for the Census division.[22]

Table 3 presents an initial set of estimates of equation (7') in which we impose a first differenced specification and experiment with alternative choices for the schoolmate characteristics. Column A presents a baseline specification that includes the city-level controls plus 3 other variables: the difference between the fraction of black students at black and white test takers' schools (row 1); the average fraction of black students in test takers' schools (row 2); and the difference in the inverse Mills ratio terms for black and white students (row 3). Consistent with the simple scatterplot in Figure 2b, there is a precisely estimated negative effect of relative segregation on the black-white gap in adjusted test scores. The inverse Mills ratio term also has a negative and significant impact, as would be expected if there was positive selection bias in the mean scores of test takers relative to the underlying population. The effect of the average racial composition of test takers' schools is small and insignificantly different from zero, indicating that black-white performance gaps do not vary substantially with the "blackness" of the city, once the **relative** exposure of black and white students to black schoolmates is held constant.

Columns B-E add additional controls for the relative characteristics of black and white schoolmates, all estimated from 2000 Census data. The model in column B adds the difference in the background indices for blacks and whites. (We use the coefficients from our SAT micro models together with the X's for high-school age children in the Census). This variable has a significant positive effect, and leads to a 15% reduction in the size of the estimated segregation effect. It also halves the effect of the inverse Mills ratio term, so the selection correction term is no longer

[22] Apart from the division dummies and the fraction black "main effect," these are all derived from the 2000 Census Summary File 3. Note that only 6 of the 9 divisions are represented among SAT states.

statistically significant. The coefficient on the differenced background index is surprisingly large in magnitude, considering that that the dependent variable is already adjusted for the background characteristics of the test-takers themselves. The index is measured in SAT points, so the 0.80 coefficient in column B implies that a 10 point widening in the expected test score gap between the black and white high school students in a city widens the gap in *adjusted* test scores between black and white test takers by 8 points. Taken literally, this suggests that schoolmate characteristics have nearly as big an effect as a student's own family background, though another explanation might be that the α coefficients estimated from the microdata are attenuated, with the city averages capturing some of the remaining individual-level background effect. Column C shows that augmenting the model with black-white differences in father's years of schooling, mother's years of schooling, and family income has a relatively small effect. As shown in column D, however, adding finer level information on differences in the distributions of mothers' and fathers' education seems to knock out the family background index, with little effect on the magnitude or precision of the relative segregation coefficient. Finally, adding information on the differences in the fractions of mothers and fathers who were born abroad and the fraction of single-parent families (column E) has little additional impact.

The models in columns F, G, and H repeat the specifications in columns B, C, and D, using estimates of the mean characteristics of black and white students based on SAT-takers in each city. The results in column F are not too different from the analogous results in column B, though the background index coefficient is even larger when the index is computed from SAT takers. In columns G and H we add several of the underlying components of the background index, including the difference in family incomes between black and white students.[23] The latter variable has a very

[23] Since family income is reported in intervals in the SAT data, we scale the income categories using the coefficients

large coefficient and displaces the explanatory power of the background index. Even more importantly, the magnitude of the relative segregation variable falls 30-40% when the background components are added to the model, and the estimate becomes only marginally significant.

We suspect that these models are over-fit, since they include over 20 highly collinear explanatory variables in a model with only 186 observations. Moreover, the effects of some of the background variables seem too large to be taken literally. (For example, in column H, the relative fraction of mothers with some college has a very large negative effect on relative test scores, while the corresponding variable for fathers has a large positive effect). We conclude that in simple first-differenced specifications the estimated effect of school segregation on the black-white test score gap is robustly negative, but that the precise magnitude is somewhat sensitive to the choice of controls for other characteristics of black and white students.

Table 4 addresses the question of whether exposure to black schoolmates has equal effects on black and white students, as is assumed in the models in Table 3. The specifications in Columns A-E of this table are identical to the ones reported in the corresponding columns of Table 3, except that we allow separate coefficients for the average fractions of black students in black and white test taker's schools (B_{1c} and B_{2c}). In each column, greater exposure of black test-takers to black schoolmates lowers black relative scores, while greater exposure of white test-takers to black schoolmates has an opposite effect. The estimated effects of the black exposure variable are larger than the corresponding effects of the white exposure variable, though formal tests of the hypothesis that the coefficients are equal and opposite (reported in the bottom row of the table) are uniformly

---

from our first stage background adjustment models. Using this scaling, the income variable is measured in SAT points. The average parental education variables in the SAT data, which do not correspond exactly to years of completed education, are measured in the same units.

insignificant.[24]

Our differencing framework allows a natural specification test for the hypothesis that the measured effects of relative segregation on the black-white achievement gap reflect the causal effect of exposure, rather than a spurious effect of unobserved determinants of black-white gaps. In particular, if our model is correctly specified, white exposure to black classmates should have no effect on average black scores, once black exposure is controlled, and similarly black exposure to blacks should have no effect on average white scores.[25] Moreover, if the "equal-and-opposite" peer effects restriction of equation (7') holds, the effect of black exposure to blacks on black average scores should be equal to that of white exposure to blacks on white average scores. Columns F and G of Table 4 consider the effects of B_{1c} and B_{2c} on black and white mean residual scores separately.[26] The estimates are remarkably consistent with earlier results. Exposure to black schoolmates has a large negative effect on black scores, and a negative, though imprecisely estimated effect on white scores. Even more supportive of our causal interpretation, the "off-diagonal" terms, measuring the effects of white (black) exposure on black (white) scores are both small and insignificant.

b. Decomposing School Segregation into Residential and Other Components

Our reading of the results in Tables 3 and 4 is that black relative test scores are lower in cities with more racially segregated schools, and that the evidence points to a causal interpretation of

---

[24] The models in this table exclude the control for the overall fraction black in the city. If the within-city black-white test score gap varied with the overall fraction black—Table 3 indicates that it does not—both the B_{1c} and B_{2c} coefficients in Table 4 would be shifted by this effect relative to the causal effects of exposure.

[25] Again, this might be violated if the city fixed effect that is implicit in our differenced models is correlated with the city fraction black. The Appendix develops these models more formally.

[26] The family background variables—our background index, family income, and parents' education, as well as the inverse Mill's ratio in the SAT-taking rate—are entered in Columns F and G as the black and white means, respectively, rather than as the difference between them. As a result, the difference between coefficients in Columns F and G is not identical to those of the otherwise analogous Column C.

this phenomenon. To gain further insights it is useful to decompose the level of school segregation in a city into a component due to geographic segregation of white and black families in the MSA and a remainder. As a starting point, Figures 3A and 3B plot city-level data on the black-white gap in actual and adjusted SAT scores against residential segregation, measured as the difference in fraction of black neighbors in black versus white residents' census tracts. The figures show a clear negative correlation between increased neighborhood segregation and the black-white test gap. By comparison, the plots in Figures 3C and 3D show no correlation between the test score gaps and the component of relative school segregation that is orthogonal to relative neighborhood segregation.[27]

To analyze these patterns more formally, we use the decomposition method described in Section II. We begin by estimating the "first stage" relationship between relative school segregation and relative residential segregation (i.e., equation (8)). The first three columns of Table 5 show alternative specifications for this model estimated on the sample of all major cities. The remaining columns show a parallel set of estimates for the subsample of cities in SAT states. Regardless of sample or the choice of control variables there is a strong relationship between the relative neighborhood segregation of the black population and the relative school segregation of black students. Over the full set of cities a 10 percentage point increase in the relative exposure of blacks to black neighbors is associated with an 8 percentage point increase in a school segregation measure. When we limit the sample to cities in SAT states and measure school segregation by SAT-takers' relative exposure to black schoolmates, the effect ranges from 5 to 7 percentage points.

Table 6 presents estimates of our second-stage model in which the residential and non-residential components of school segregation are entered separately. Estimates of the effects of these two components are shown in the first two rows of the table. Recall that the effect of the

[27] We calculate this component as the residual from a univariate regression of relative school segregation on relative neighborhood segregation.

residential component is equivalent to an instrumental variables (IV) estimate of the effect of school segregation on relative test scores, using neighborhood segregation as an instrument for school segregation. Compared to the OLS estimates from parallel specifications (presented in Table 3) the IV estimates are roughly twice as large in magnitude, and more stable across specifications. The estimates imply that moving from a city with a high degree of residential segregation (e.g. Cleveland) to one with low residential segregation (Charlotte) would be expected to reduce the black-white test score gap by 45 points. By contrast, the estimates in row 2 of the effect of the non-residential component of school of segregation are quite small and statistically insignificant. The results of the decomposition are highly robust to changes in the control variables.

c. Selection into SAT-taking

There are a number of possible explanations for the contrasting effects of the residential and non-residential components of relative school segregation on the black-white achievement gap. Before examining these in more detail we first consider an alternative explanation: that inferences from the SAT data are biased by selective test participation. At the outset, we note that the estimates in Table 6 are based on comparisons across metropolitan areas in high-SAT-participation states. Restricting attention to these cities eliminates the correlation between overall SAT participation rates and average SAT scores (see Rothstein 2004). Moreover, all the models in Table 6 include inverse Mill's ratio control functions based on the relative test participation rates of whites and blacks in different cities. As expected, the estimated coefficients on the selection correction functions are uniformly negative, indicating a positive correlation between latent test scores and the probability of writing the test. Nevertheless, the coefficients on the selection correction terms are somewhat imprecise, and often insignificant, raising the concern that our results may be biased by selective participation.

Our first approach to this concern is to re-estimate the models using artificially trimmed data. The impact of selection bias is maximized when the correlation between latent test scores and the probability of test participation is 1. In this case, the observed distribution of scores in a city with test-taking rate p_c is a truncated version of the true distribution, with a fraction (1- p_c) of the lowest scores removed. If the lowest test taking rate across all cities in the sample is p_{min}, selection bias can be eliminated by discarding the lowest test scores in all other cities until the corrected participation rate is p_{min} in all cities. If the correlation between test participation and latent scores is less than 1, this procedure "over-corrects" the data from high-participation cities, producing to a positive correlation between observed mean test scores (from the trimmed sample) and the original participation rate – the opposite sign of the bias in the untrimmed data.

In practice we implement this approach by choosing separate participation thresholds for blacks and whites (p^b=0.2 and p^w=0.3), excluding cities with participation rates below these thresholds, and trimming the observed distributions for all other cities. We re-adjust the observed SAT scores for the trimmed samples using estimates of $\hat{\alpha}_j$ from this trimmed sample, and then re-estimate our basic models using the black-white difference in average adjusted trimmed scores as the dependent variable. The results are shown in Table 7.

Columns A and B reproduce the OLS models from columns B and C of Table 3 on the trimmed data. In each case, the estimated effect of relative segregation is slightly smaller when the trimmed sample is used, and is estimated with less precision, so that neither coefficient is significant.[28] Columns C and D reproduce the specifications from columns B and C of Table 6. In

[28] Note, however, that in this specification the "main effect" of the fraction black at the average sample student's school may be picking up some of the effect of segregation as well, as if the students in the trimmed sample were completely segregated from other students in the MSA at schools that were perfectly racially balanced, there would be no variation

the trimmed sample the estimated effects of residentially-induced segregation are more negative -- though less precisely estimated -- than in the full sample, while the non-residential component of school segregation has a *positive* (and nearly significant) effect on the black-white test score gap. There is certainly no indication from these estimates that the pattern of findings in the overall sample is driven by selective test participation. If anything, the results from the trimmed sample suggest an even bigger contrast between the effects of the residential and non-residential components of relative school segregation.

A second and arguably more persuasive to evaluate the impact of selective test participation is to examine models for black-white relative achievement based on outcomes that are available for a random sample of youths. We use the 2000 Census microdata samples to construct four outcome measures for 16-24 year olds in each city: the employment rate, the enrollment rate, the fraction either employed or in school, and the fraction who either are currently enrolled or have completed high school. A limitation of the Census data is that there is no family background information for children who are no longer living with their parents. Consequently, we make no individual-level adjustments for family background. Instead, we regress the black-white difference for each outcome measure on relative school segregation, black-white differences in the same Census-based family background measures used in our SAT analysis (measured for high-school-age children in the same city), and the other citywide control variables used in our previous models.

Table 8 presents estimates of our two-stage decomposition models for the Census outcomes. (Note that these models are estimated over the full set of 319 identifiable MSAs). For all four outcome measures we obtain a negative estimate of the effect of the residentially-induced component of relative school segregation. The estimates imply that a 10 percentage point increase

---

in this variable. Consistent with this, these main effects are quite large in Table 7.

in residentially-based school segregation leads to a 1.7 percentage point reduction in the black employment rate, relative to that of whites; a 0.9 percentage point relative reduction in the share who are in school; a 1.2 percentage point relative reduction in the share either employed or in school; and a 1.0 percentage point relative reduction in the fraction who have finished high school. All but the enrollment effect are significantly different from zero. As we found in Table 6, the non-residential component of school segregation has no significant relationship with the black-white gap in the Census-based outcomes. Indeed, the estimate of γ_2 has the "wrong" sign in two of the four models.

It seems clear that the basic pattern of results in Table 6 cannot be explained as the result of endogenous selection into SAT-taking. The same pattern is apparent in trimmed samples designed to mitigate the impact of selection bias, and in Census data that are free of endogenous selection. We conclude that the contrasting relationships between the black-white relative achievement gap and the residential and non-residential components of school segregation are a genuine feature of the data, and not simply an artifact of problems with the SAT sample.

d. Further Decomposing the Non-residential Component of School Segregation

The decomposition underlying the results in Tables 6-8 divides the observed level of relative school segregation in a city into a component attributable to the degree of neighborhood segregation in the city and a residual component. Interpreting the effect of the residual component is difficult, however, because it includes the effects of features like busing and magnet school programs as well as potentially endogenous behavioral reactions such as private school enrollment. To isolate the component of school segregation attributable to direct policy intervention, we use Welch and Light's (1987) summary of court-ordered desegregation programs in the 1970s and early 1980s.

Specifically, for each city we use Welch and Light's estimate of the change in the dissimilarity index of racial segregation for the city's schools from the year just before the city's major desegregation plan to the last year of implementation. This change is set to zero for cities without a major desegregation plan. Cities not included in the Welch and Light sample (or for which they were unable to obtain data) are excluded from the sample. We then use this variable along with relative residential segregation in the two-part decomposition described by equations (10a), (10b), and (11).

Columns A and B of Table 9 report the two first stage models (equations (10a) and (10b)) relating residential segregation and the Welch and Light desegregation measure to our school-level segregation index. These models are estimated using data for all public high school students in the 60 metropolitan areas that are in high SAT participation states and were included in Welch and Light's sample. The estimates in column A confirm that, as in Table 5, there is a strong effect of residential segregation on relative school segregation in this smaller sample of cities. The effect is reduced slightly in column B, when the desegregation measure is added. The coefficient on the latter variable indicates that cities with more effective desegregation programs in the 1970s and early 1980s continued to have substantially less segregated high schools in 2000 than would be expected given their observable characteristics. Columns C and D repeat these estimates with the dependent variable measured over the schools attended by SAT-takers, with very similar results.

Finally, columns E and F report estimates of the decomposition models (9) and (11). As in Table 6, we see a strong effect of residentially-based school segregation on black-white adjusted test score gaps. The remaining portion of school segregation, however, continues to have no effect, regardless of whether it is entered as a single variable (column E) or decomposed into a desegregation-related component and a residual (column F). Indeed, the estimated effect of the component of school segregation attributable to explicit desegregation programs is actually positive, though imprecisely estimated, suggesting that cities with stronger desegregation remedies have even larger black-white test score gaps than do cities with weaker (or no) court orders.

V. Interpreting the Results

One interpretation of the results in Tables 6-10 is that exposure to black schoolmates has no effect on the relative achievement of black students, holding constant exposure to black neighbors. Several other interpretations are also possible, however. First, it may be that the degree of residential segregation is endogenous with respect to unobserved differences in the relative abilities of black and white students in a city. Reber (2003), for example, finds that court-ordered desegregation plans led to significant white flight from the court-supervised districts. If white flight was more pronounced in cities with a bigger gap in the unobserved abilities of black and white students, our estimates of the residential segregation effect will be biased. A second candidate explanation is that higher residential segregation is correlated with bigger resource gaps between the schools attended by black and white students. If this is true, and if resources matter for test performance, our estimates may be reflecting a resource effect, rather than a pure racial segregation effect.[29] Third, it may be that classroom peer groups are what matter for student performance, and that cities with more aggressive between-school integration programs also have more within-school segregation. This mechanism is consistent with anecdotal evidence suggesting that districts with stronger desegregation programs often offer magnet programs, honors courses, or "schools within schools" to try to attract white families to high-minority schools (Clotfelter, Ladd et al. 2003;

[29] There is an important sense in which resource effects may be part of broadly-defined peer effects: If voters are willing to devote greater resources to schools with more white students, or if these schools are able to obtain higher-quality resources with the same spending (as would be the case if, for example, good teachers prefer to teach in white schools), this might be thought of as a consequence of the school enrollment. It is nevertheless informative to understand whether this is the channel through which our estimated effects operate.

Clotfelter 2004).[30]

a. Potential endogeneity of residential segregation: A historical perspective

We investigate the first alternative hypothesis by using historical residential segregation data, rather than current data, to decompose current school segregation. Historical residential segregation will be a valid instrument for current school segregation if the degree of neighborhood segregation in a city in, e.g., 1950 is unrelated to the difference in unobserved family background and ability between black and white students in a city in 2000. One way this could fail is through selective mobility. Higher residential segregation in the past could have led to the net out-migration of blacks with higher unobserved ability (or in-migration of whites with higher unobserved ability). Although we cannot fully discount this link, an examination of mobility patterns over the 1979-2000 period for respondents in the National Longitudinal Survey of Youth shows no evidence of such patterns. The other way that historical neighborhood segregation could be related to the current racial gap in test-taking ability is through an intergenerational link. If neighborhood segregation was correlated with differences in the black and white populations in a city in 1950, some of that difference would persist today through the usual intergenerational channels. We expect that most of these differences would be expressed today in our observed family background measures, including family structure, education, and income. At a minimum, we believe that any correlation between the black-white gap in unobserved ability and the current level of neighborhood segregation in a city will be substantially smaller with historical measures of neighborhood segregation.

[30] See also, for example, the opinion in *People Who Care v. Rockford Board of Education*, 851 F. Supp. 905 (1993), which reads in part: "The court finds that the ability grouping and tracking practices of the Rockford School District (hereinafter 'RSD') did not represent a trustworthy enactment of any academically acceptable theory or practice. The RSD tracking practices skewed enrollment in favor of whites and to the disadvantage of minority students. The court finds that it was the policy of the RSD to use tracking to intentionally segregate white students from minority students…." (p. 940)

We use the indices of residential segregation constructed by Cutler, Glaeser, and Vigdor (1999; hereafter "CGV") from 1940-1990 Decennial Census data as our measures of historical segregation.[31] Table 10 presents estimates of our two-part decomposition model (equation 9) using CGV segregation measures from 1950, 1960, 1970, 1980, and 1990 as instruments for current segregation. Column A reproduces the model from Table 6, column C. Columns B and C restrict the sample to the 170 cities in our SAT sample for which CGV provide 1990 data. The same model is estimated, using current segregation as the instrument in column B and CGV's 1990 isolation and dissimilarity indices as instruments in column C. The remaining columns repeat the exercise using earlier census years, with each pair of columns using the sample of SAT MSAs for which CGV have data for the relevant year. In each pair, the first column uses current segregation as the instrument in the restricted sample, and the second uses the historical segregation measures.

Considering first the models that use the current segregation measure, in columns A, B, D, F, H, and J, the coefficient on the residential component of school segregation falls somewhat as the sample is gradually restricted, and the estimate in column J is only about half of the original estimate. (Note, however, that in the sample of cities with 1950 data the model is overfit, as there are only 37 observations.) Interestingly, the same is not true when the CGV historical measures are used as instruments. In these models—columns C, E, G, I, and K—the coefficient on the residential component of school segregation is remarkably stable, nearly identical in the 2000 and 1970 models and falling by less than a third even in the 1950 model. Consistent with previous results, none of the models in Table 10 indicate an effect of the residual component of school segregation on adjusted test score gaps. We conclude from this table that our earlier results are not substantially biased by

[31] Current levels of residential segregation are fairly highly correlated with lagged levels. Among the 200 MSAs for which CGV provide 1970 segregation measures, for example, the weighted correlation between CGV's isolation index and a similar index computed from 2000 data is 0.78. Even going back to 1950, when only 62 MSAs can be matched to current data, the correlation is 0.59.

the potential endogeneity of recent changes in residential segregation.

b. Segregation and relative access to resources

The second candidate explanation for the contrasting effects of the residential and non-residential components of school segregation is that residential segregation is highly related to the gap in school resources at school attended by black and white schools, leading to a biased estimate of the effect of residentially-based school segregation. Unfortunately, there are few sources of information on resource availability. The only school-level measure that is available for most schools (from the Common Core of Data Elementary/Secondary School Universe data file) is the number of full-time-equivalent (FTE) teachers. We compute the number of teachers per student at (public) schools attended by white and black students in each MSA.[32] Measures of spending are available only at the district level. Using the CCD Local Education Agency Finance Survey (also known as the F-33 portion of the Census of Governments) we compute expenditures per pupil, instructional expenditures per pupil, and per pupil revenue in districts attended by white and black students in each MSA. Unfortunately, if resource allocations are not equal across schools in each district, these variables provide imperfect measures of the actual resources available to black and white students. For information on the *quality* of resources, we turn to data from the Schools and Staffing Survey (SASS) on the qualifications, experience, and characteristics of a sample of teachers, matched to the racial composition of the schools in which they teach.

Table 11 presents estimates of our basic decomposition model, fit to measures of the black-

[32] This calculation includes all grades, even though class sizes are typically smaller in elementary schools, because FTEs are only available at the school level and the separation of grades into elementary, middle, and high schools varies somewhat across metropolitan areas. For the same reason, we use teachers per student rather than the inverse (which has a more natural interpretation as the average class size), as the enrollment-weighted average of the former ratio for each race is insensitive to the way that heterogeneous grades are distributed among schools.

white gap in the various resource measures.[33] Interestingly, the coefficients on the residential component of school segregation are uniformly positive for the CCD measures, and significant in three of the four columns, suggesting that the relative funding of blacks' schools and districts is *higher* in more residentially segregated cities. Turning to the SASS measures, both residential and non-residential components of school segregation are associated with closer racial matches between students and their teachers (i.e. with relatively more black teachers at schools serving black students). Beyond this, the SASS data are quite noisy, though there is no indication that either component of school segregation is associated with worse resources for black students. We conclude from the set of results in Table 11 that differential access to resources is an unlikely explanation for our findings on the black-white achievement gap.

c. Across-school segregation and within-school exposure

Our final candidate explanation is that student achievement is primarily affected by classroom level peers, rather than school-level peers, and that variation across cities in the relative exposure of black and white students to black *schoolmates* is only weakly correlated with the relative exposure to black *classmates*. This hypothesis is difficult to test directly, since to the best of our knowledge there are no national data on the racial composition of high school classrooms. We therefore focus on indirect tests, investigating whether the prevalence of ability tracking of various forms covaries with the residential and non-residential components of school segregation.[34]

[33] In these models, the fraction black is not defined over SAT-takers' schools but over all public schools, weighted by each race's public high school enrollment (columns A-D and H) or enrollment in all grades (E-G, I).

[34] Note that causation need not operate from tracking to relative performance: Greater inequality of test scores presumably creates incentives to institute tracking, and larger black-white test score gaps would naturally lead to racial disparities in course assignments. Similarly, cities with more tracking may be able to achieve greater desegregation, if tracking is a precondition for some white families' enrollment in mixed-race schools. Given the lack of relationship between non-residential school segregation and black-white test score gaps, any relationship of this component of segregation with black-white gaps in course-taking seems unlikely to reflect causality flowing from test score gaps.

Our first effort to examine this issue uses the fraction of secondary teachers in the SASS sample in each MSA who reported that students were assigned to their courses on the basis of their ability. Column A of Table 12 presents estimates of the effects of residential and school segregation on this measure of the prevalence of tracking. There is no evidence here that tracking is more prevalent in cities with less school segregation, in either component, although standard errors are large.

The SAT data provide another source of information about tracking. Students are asked, as part of the Student Descriptive Questionnaire, whether they had taken honors courses in each of several academic subjects, and whether they intended to claim advanced placement credit or course exemptions in college on the basis of high school work. Columns B and C of Table 12 present models for dependent variables measuring whether students indicated that they had taken honors courses in math and English, respectively, while Columns D and E take as outcomes the fraction of students who planned to claim college-level credit in any subject (column D) and specifically in math or English (column E).

We present estimates of the relationships between the two components of school segregation and the white and black means of these variables in Panels B and C. The estimates in Panel B show little relation between school segregation and black honors-taking. Those in Panel C suggest a negligible or slightly positive effect of the residential component of school segregation on white students' participation in honors courses, contrasting with a nearly significant negative effect of the non-residential component. Panel D reports estimates for the black-white difference in honors participation at the city level, while Panel E reports models for the within-high-school version of the same difference (computed as MSA-specific black coefficients in linear probability models for each measure with high school fixed effects). Non-residential school segregation is associated with large positive effects on the black-white gap in honors course taking (though these are only marginally significant in Panel E), while the residential segregation effect is much smaller, less stable, and never significant. These results offer clear support for the within-school segregation hypothesis, suggesting that school desegregation has a positive relationship with the prevalence of ability sorting. This is consistent with there being a weak relationship between the non-residential component of school segregation and a measure of relative classroom exposure, were one available.[35]

VII. Conclusion

One of the most important and controversial issues in U.S. education policy is the racial segregation of schools. In this paper we propose a new methodology, based on the differences in black and white student achievement across cities, to evaluate the effects of relative exposure of black students to black schoolmates. Using this method we find what we believe to be robust and credible evidence that black students' relative achievement is lower in cities with more racially segregated schools. We go on to ask whether the components of school segregation attributable to residential patterns and other factors have similar effects on outcome gaps. To our surprise, we find that they have quite different effects: Residential segregation appears to hurt the relative performance of black students, while variation in school segregation around that predicted from

[35] We have also estimated models for the tracking measures that separate out the component of non-residential school segregation attributable to court-ordered desegregation. Standard errors are large, but the results indicate that, if anything, court-ordered desegregation has larger effects on tracking than does the residual component. Also, though we focus in Table 12 on tracking in secondary grades, we have found two indications that non-residential school segregation may be associated with greater within-school stratification at the *primary* level. First, when we do not restrict our SASS sample to secondary teachers we estimate a positive effect of the non-residential segregation component on the prevalence of tracking. Second, our analysis of data from the Early Childhood Longitudinal Survey on the racial composition of kindergarten classrooms suggests that cities with more non-residential school segregation have schools that are, at the kindergarten level, significantly more internally segregated than are schools in cities with less school segregation.

residential patterns—even when instrumented by measures of the strength of court-ordered school desegregation regimes—is essentially uncorrelated with black-white gaps.

At first glance our results on the effects of the residential component of school segregation appear to be inconsistent with recent results on the effects of moving children from high-poverty to lower-poverty neighborhoods. Using data from the Moving to Opportunity experiment, Kling and Liebman (2004) and Katz, Kling, and Liebman (2001) find that moving from a high-poverty public housing development to a lower poverty neighborhood has negligible impacts on children's average outcomes. Similarly, Jacob (2003) finds little impact on student outcomes when public housing residents are forced to move because of building demolitions. It should be noted, however, that both of these interventions affected children for only a few years, rather than over their entire lives. Our results are more clearly consistent with those of Cutler and Glaeser (1997), who show substantial effects of long-run variation in residential segregation on adult economic outcomes.

Our finding that variation in the non-residential component of school segregation has little or no effect on the black achievement gap also appears to conflict with Guryan's (forthcoming) recent analysis of the effect of desegregation plans on high school graduation rates, though our cross-sectional approach does not provide enough power to distinguish effects of the size estimated by Guryan. In any case, it may be the case that school desegregation programs have a bigger effect on in their first few years of operation than 15-20 years later. In particular, we find suggestive evidence that cities with lower rates of non-residential school segregation have higher rates of within-school tracking, possibly reducing the exposure of black students to white classmates and counteracting desegregation-induced rises in exposure to white schoolmates.

Given the divergence between our results and those in previous work, and the limitations of our research design, we believe that our analysis raises as many questions as it answers. More research is needed to understand the long-run effects of residential segregation on student outcomes, and the apparently small effects of other sources of school segregation.

REFERENCES

Boyd, Donald, Hamilton Lankford, Susanna Loeb and James Wyckoff (2003). "Analyzing the Determinants of the Matching Public School Teachers to Jobs: Estimating Compensating Differentials in Imperfect Labor Markets." National Bureau of Economic Research: Working Paper #9878.

Brock, William A. and Steven N. Durlauf (2001). "Interactions-Based Models," in Handbook of Econometrics, Volume 5, J. J. Heckman and E. Leamer, eds. Amsterdam; London and New York, Elsevier Science, North-Holland**:** 3297-3380.

Card, David and A. Abigail Payne (2002). "School Finance Reform, the Distribution of School Spending, and the Distribution of SAT Scores." Journal of Public Economics **83**(1): 49-82.

Clark, Melissa (2003). "Selection Bias in College Admissions Tests," Mimeo, Princeton University

Clotfelter, Charles T. (2004). After Brown : The Rise and Retreat of School Desegregation. Princeton, N.J., Princeton University Press.

Clotfelter, Charles T., Helen F. Ladd and Jacob L. Vigdor (2003). "Segregation and Resegregation in North Carolina's Public School Classrooms." North Carolina Law Review **81**: 1463-1511.

Coleman, James S. (1966). Equality of Educational Opportunity. Washington, D.C., U.S. Office of Education.

Cutler, David M. and Edward L. Glaeser (1997). "Are Ghettos Good or Bad?" Quarterly Journal of Economics **112**: 827-72.

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

Dynarski, Mark (1987). "The Scholastic Aptitude Test: Participation and Performance." Economics of Education Review **6**(3): 263-273.

Evans, William N., Wallace E. Oates and Robert M. Schwab (1992). "Measuring Peer Group Effects: A Study of Teenage Behavior." Journal of Political Economy **100**(5): 966-91.

Garen, John (1984). "The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable." Econometrica **52**(5): 1199-1218.

Guryan, Jonathan (forthcoming). "Desegregation and Black Dropout Rates." American Economic Review.

Heckman, James J. (1979). "Sample Selection Bias as a Specification Error." Econometrica **47**: 153-161.

Hoxby, Caroline M. (2000). "Does Competition among Public Schools Benefit Students and Taxpayers?" American Economic Review **90**(5): 1209-1238.

Hsieh, Chang-Tai and Miguel Urquiola (2003). "When Schools Compete, How Do They Compete? An Assessment of Chile's Nationwide School Voucher Program," Mimeo, Columbia University, August.

Iceland, John, Daniel H. Weinberg and Erika Steinmetz (2002). "Racial and Ethnic Residential Segregation in the United States: 1980-2000." U.S. Census BureauCENSR-3, Washington, D.C.

Jencks, Christopher and Meredith Phillips, Eds. (1998). The Black-White Test Score Gap. Washington, D.C., The Brookings Institution.

Katz, Lawrence, Jeffrey Kling and Jeffrey Liebman (2001). "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment." Quarterly Journal of Economics **116**(2): 607-654.

Kling, Jeffrey and Jeffrey Liebman (2004). "Experimental Analysis of Neighborhood Effects on

Youth." Princeton University, Industrial Relations Section Working Paper 483.
Manski, Charles F. (1990). "Nonparametric Bounds on Treatment Effects." American Economic Review Papers and Proceedings **80**: 319-323.
--- (1993). "Identification of Endogenous Social Effects: The Reflection Problem." Review of Economic Studies **60**(3): 531-42.
Massey, Douglas S. and Nancy A. Denton (1988). "The Dimensions of Residential Segregation." Social Forces **67**(2): 281-314.
Murphy, Kevin M. and Robert H. Topel (1985). "Estimation and Inference in Two-Step Econometric Models." Journal of Business and Economic Statistics **3**(4): 370-379.
Ogbu, John (2003). Black American Students in an Affluent Suburb: A Study of Academic Disengagement. Mahwah, NJ, Lawrence Erlbaum Associates, Inc.
Ogbu, John and Signithia Fordham (1986). "Black Students' School Success: Coping with the Burden of 'Acting White'." The Urban Review **18**(3): 176-206.
Reber, Sarah (2003). "Court-Ordered Desegregation: Successes and Failures Integrating American Schools since Brown," Unpublished manuscript, November.
Rothstein, Jesse M. (2004). "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions." National Bureau of Economic Research: Working Paper #10666.
Urquiola, Miguel (1999). "Demand Matters: School District Concentration, Composition, and Educational Expenditure." University of California, Berkeley, Center for Labor Economics: Working Paper #14.
Welch, Finis and Audrey Light (1987). New Evidence on School Desegregation. Washington, D.C., United States Commission on Civil Rights.

TECHNICAL APPENDIX

A. Inference in the decomposition models

Our basic two-level decomposition model, (8) and (9), may be rewritten as

(A1) $E[Y \mid X, Z, W] = X\beta + (Z\pi)^*\gamma_1 + (W - Z\pi)^*\gamma_2$, with

(A2) $E[W \mid Z] = Z\pi$.

Let $\varepsilon = Y - E[Y \mid X, Z, W]$ and $u = W - E[W \mid Z]$; by definition, $\text{cov}(\varepsilon, u)=0$. Let $\hat{\pi} = (Z'Z)^{-1}Z'W$ be the sample estimate of the first-stage coefficient π, and let $D = (X \quad Z\pi \quad (W\text{-}Z\pi))$, $\hat{D} = (X \quad Z\hat{\pi} \quad (W\text{-}Z\hat{\pi}))$, and $\Omega = (\beta' \quad \gamma_1' \quad \gamma_2')'$.

If we could obtain the "true" OLS estimate of Ω from (A1), $(D'D)^{-1}D'Y$, inference would be straightforward. Because we observe only $\hat{\pi}$ and not π (and therefore $\hat{D}$ rather than D), we instead estimate

(A3) $\hat{\Omega} = (\hat{D}'\hat{D})^{-1}\hat{D}'Y$.

Inference proceeds by noting that

(A4) $Y = D\Omega + \varepsilon = \hat{D}\Omega + (D - \hat{D})\Omega + \varepsilon = \hat{D}\Omega + Z(\pi - \hat{\pi})(\gamma_1 - \gamma_2) + \varepsilon$, so

(A5) $\hat{\Omega} = \Omega + (\hat{D}'\hat{D})^{-1}\hat{D}'[Z(\pi - \hat{\pi})(\gamma_1 - \gamma_2) + \varepsilon]$.

As a result,

(A6) $$\begin{aligned}\text{var}(\hat{\Omega}) &= \text{var}\{(\hat{D}'\hat{D})^{-1}\hat{D}'[Z(\pi - \hat{\pi})(\gamma_1 - \gamma_2) + \varepsilon]\}\\ &\approx \text{var}\{(D'D)^{-1}D'[Z(\pi - \hat{\pi})(\gamma_1 - \gamma_2) + \varepsilon]\}\\ &= (D'D)^{-1}D'\text{var}\{Z(\pi - \hat{\pi})(\gamma_1 - \gamma_2) + \varepsilon\}D(D'D)^{-1}\\ &= (D'D)^{-1}D'[\text{var}\{Z(\pi - \hat{\pi})(\gamma_1 - \gamma_2)\} + \text{var}(\varepsilon)]\,D(D'D)^{-1},\end{aligned}$$

with the approximation due to Murphy and Topel (1985) and the final equality a result of the fact that

(A7) $Z(\pi - \hat{\pi})(\gamma_1 - \gamma_2)\} = Z((Z'Z)^{-1}Z'u)(\gamma_1 - \gamma_2)$

is uncorrelated with ε. Note that

(A8) $\text{var}\{Z(\pi - \hat{\pi})(\gamma_1 - \gamma_2)\} = (\gamma_1 - \gamma_2)^2 Z\text{var}(\hat{\pi})Z'$, so

(A9) $\text{var}(\hat{\Omega}) = (D'D)^{-1}D'[(\gamma_1 - \gamma_2)^2 Z\text{var}(\hat{\pi})Z' + \text{var}(\varepsilon)]\,D(D'D)^{-1}$.

Under classical assumptions of homoskedasticity and independence, $\text{var}(\varepsilon)=\sigma_\varepsilon^2 I$ and $\text{var}(u)= \sigma_u^2 I$, so

(A9') $\text{var}(\hat{\Omega}) = (D'D)^{-1}D'[(\gamma_1 - \gamma_2)^2 Z'(Z'Z)^{-1}Z\,\sigma_u^2 + \sigma_\varepsilon^2]\,D(D'D)^{-1}$.

In our empirical implementation, we allow for heteroskedasticity and clustering at the CMSA level, so we use the general form (A9). Let Σ be the estimated variance matrix of ε in this case, and let Γ be the heteroskedastic, clustered estimate of $\text{var}(\hat{\pi})$, which can be readily obtained from standard software packages. Canned routines for estimating the clustered variance matrix for the second-stage model, (A3), do not take account of sampling error in $\hat{\pi}$, so omit the first term in (A9) and provide

(A10) $\Psi = (\hat{D}'\hat{D})^{-1}\hat{D}'\Sigma\hat{D}(\hat{D}'\hat{D})^{-1}$.

We adjust this matrix to obtain a consistent variance-covariance matrix for $\hat{\Omega}$:

(A11) $\widehat{\text{var}}(\hat{\Omega}) = \Psi + (\hat{\gamma}_1 - \hat{\gamma}_2)^2(\hat{D}'\hat{D})^{-1}\hat{D}'Z\Gamma Z'\hat{D}(\hat{D}'\hat{D})^{-1}$.

The three-level model—(10a), (10b), and (11)—is only slightly more complex:

(A12) $Y = X\beta + (Z_1\pi_{11})^*\gamma_1 + (Z_1\pi_{21} + Z_2\pi_{22} - Z_1\pi_{11})^*\gamma_2 + (W - Z_1\pi_{21} + Z_2\pi_{22})^*\gamma_3 + \varepsilon$, with

(A13) $W = Z_1\,\pi_{11} + u_1$

$= Z_1\,\pi_{21} + Z_2\,\pi_{22} + u_2$.

Again, let $D=(X\ \ Z_1\pi_{11}\ \ (Z_1\pi_{21} + Z_2\pi_{22} - Z_1\pi_{11})\ \ (W - Z_1\pi_{21} + Z_2\pi_{22}))$, $\Omega=(\beta'\ \ \gamma_1'\ \ \gamma_2'\ \ \gamma_3)'$, and $\hat{D}=(X\ \ Z_1\hat{\pi}_{11}\ \ (Z_1\hat{\pi}_{21} + Z_2\hat{\pi}_{22} - Z_1\hat{\pi}_{11})\ \ (W - Z_1\hat{\pi}_{21} + Z_2\hat{\pi}_{22}))$. Then

(A14) $Y = D\,\Omega + \varepsilon = \hat{D}\,\Omega + Z_1(\pi_{11} - \hat{\pi}_{11})(\gamma_1-\gamma_2) + (Z_1(\pi_{21} - \hat{\pi}_{21}) + Z_2(\pi_{22} - \hat{\pi}_{22}))(\gamma_2-\gamma_3) + \varepsilon$.

By the same logic used above,

(A15)
$$\begin{aligned}\mathrm{var}(\hat{\Omega}) &= \mathrm{var}\{(D'D)^{-1}D'\,[Z_1(\pi_{11} - \hat{\pi}_{11})(\gamma_1-\gamma_2) + (Z_1(\pi_{21} - \hat{\pi}_{21}) \\ &\qquad + Z_2(\pi_{22} - \hat{\pi}_{22}))(\gamma_2-\gamma_3) + \varepsilon]\} \\ &= (D'D)^{-1}D'[Z_1{*}\mathrm{var}(\hat{\pi}_{11}){*}Z_1'(\gamma_1-\gamma_2)^2 \\ &\quad + (Z_1{*}\mathrm{var}(\hat{\pi}_{21}){*}Z_1' + Z_2{*}\mathrm{var}(\hat{\pi}_{21}){*}Z_1' \\ &\qquad + 2{*}Z_1{*}\mathrm{cov}(\hat{\pi}_{21}, \hat{\pi}_{22}){*}Z_2')(\gamma_2-\gamma_3)^2 \\ &\quad + 2{*}Z_1{*}(\mathrm{cov}(\hat{\pi}_{11}, \hat{\pi}_{21}){*}Z_1' + \mathrm{cov}(\hat{\pi}_{11}, \hat{\pi}_{22}){*}Z_2')(\gamma_1-\gamma_2)(\gamma_2-\gamma_3) \\ &\quad + \mathrm{var}(\varepsilon)]D(D'D)^{-1}.\end{aligned}$$

A consistent estimate of $\mathrm{var}(\hat{\Omega})$ is obtained by replacing each term in (A15) with an estimate of it. The variance terms are directly obtainable from the first-stage models. The only complex terms are those involving the covariance between $\hat{\pi}_{11}$ and the coefficients from the second first-stage model. By the standard omitted variables formula, $\hat{\pi}_{11} = \hat{\pi}_{21} + (Z_1'Z_1)^{-1}Z_1'Z_2\hat{\pi}_{22}$; as a result, $\mathrm{cov}(\hat{\pi}_{11}, \hat{\pi}_{21}) = \mathrm{var}(\hat{\pi}_{21}) + (Z_1'Z_1)^{-1}Z_1'Z_2\mathrm{cov}(\hat{\pi}_{22}, \hat{\pi}_{21})$ and $\mathrm{cov}(\hat{\pi}_{11}, \hat{\pi}_{22}) = \mathrm{cov}(\hat{\pi}_{21}, \hat{\pi}_{22}) + (Z_1'Z_1)^{-1}Z_1'Z_2\mathrm{var}(\hat{\pi}_{22})$.

B: Equal and opposite effects?

The specifications in Table 4 allow us to test the assumption that changes in the fraction of black schoolmates have equal effects on black and white students, in which case the first-differenced specifications reported in Table 3 are warranted. To aid in the interpretation of the models in this table, recall that equation (4) relates the adjusted SAT score of black and white students in city c to the mean characteristics of their schoolmates (X_{1c} and X_{2c}), the average fraction of black students in their schools (B_{1c} and B_{2c}), and an average residual (u_{1c} and u_{2c}):

(B1a) $r_{1c} = X_{1c}\beta_1 + B_{1c}\,\gamma_1 + u_{1c}$,

(B1b) $r_{2c} = X_{2c}\beta_2 + B_{2c}\,\gamma_2 + u_{2c}$.

Differencing (B1a) and (B1b) is appropriate if any correlations between (u_{1c}, u_{2c}) and (B_{1c}, B_{2c}) arises through a common component:

$u_{1c} = u_c + \eta_{1c}$, $u_{2c} = u_c + \eta_{2c}$, with $E[\eta_{jc} B_{kc}] = 0$ for all j and k,

since in this case $u_{1c} - u_{2c}$ is uncorrelated with either B_{1c} or B_{2c}. Since the common component can always be decomposed as

$u_c = \pi_1 B_{1c} + \pi_2 B_{2c} + \xi_c$, with $E[\xi_c B_{kc}] = 0$

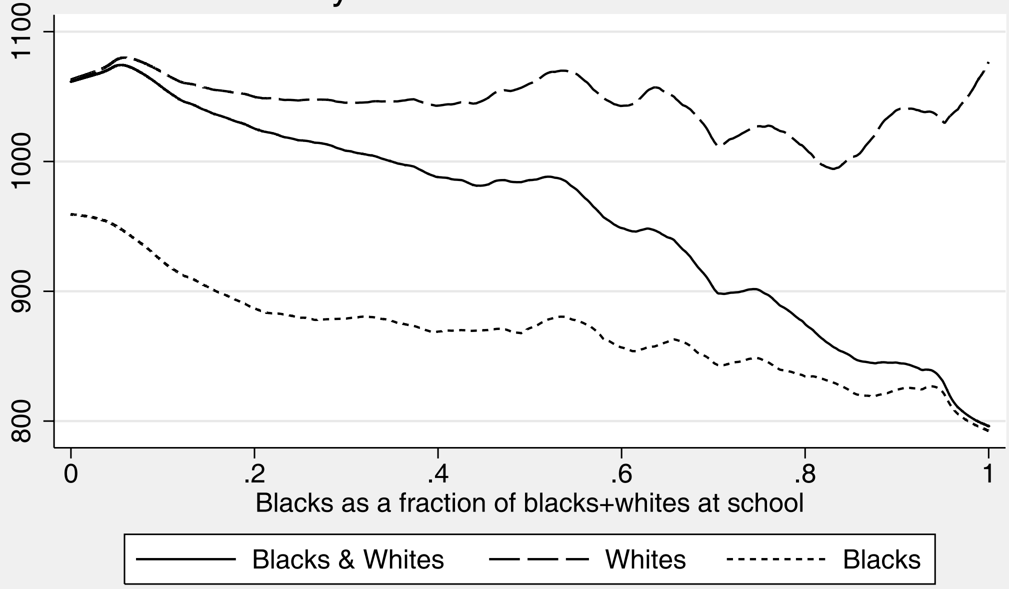
we could also estimate separate models for blacks and whites that include both segregation terms:

(B2a) $r_{1c} = X_{1c}\,\beta_1 + B_{1c}\,(\gamma_1 + \pi_1) + B_{2c}\,\pi_2 + \xi_c + \eta_{1c}$,

(B2b) $r_{2c} = X_{2c}\,\beta_2 + B_{1c}\,\pi_1 + B_{2c}\,(\gamma_2 + \pi_2) + \xi_c + \eta_{2c}$.

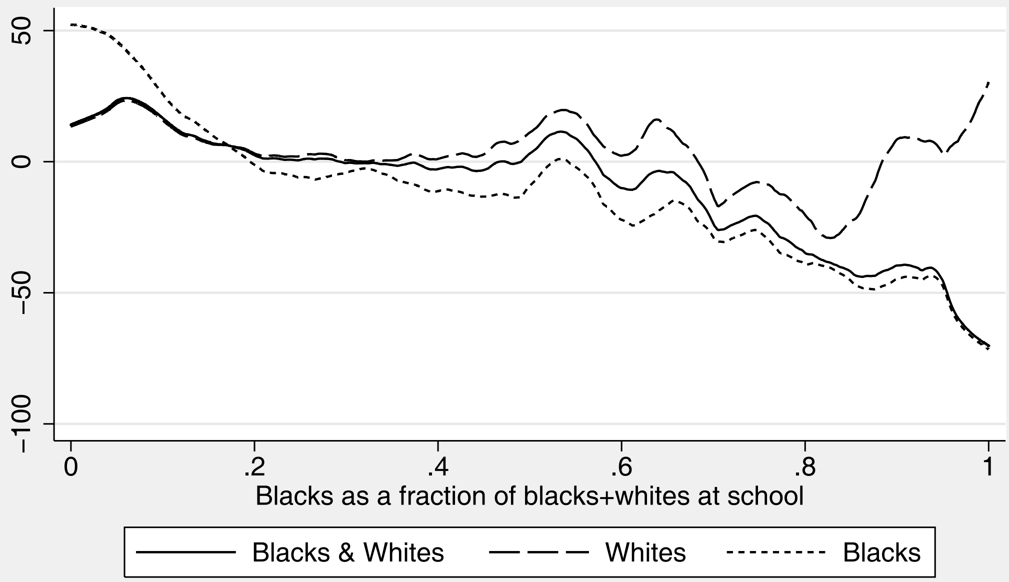
Columns A-E of Table 4 present a differenced model that relates ($r_{1c} - r_{2c}$) to B_{1c} and B_{2c}. If (B2a) and (B2b) are correctly specified, the coefficients on these variables should estimate γ_1 and $-\gamma_2$, respectively. Columns F and G of Table 4 present race-specific models like (B2a) and (B2b). The implied values of π_1 and π_2 are both close to zero.

Figure 1A. Average SAT scores among white and black students, by school fraction black



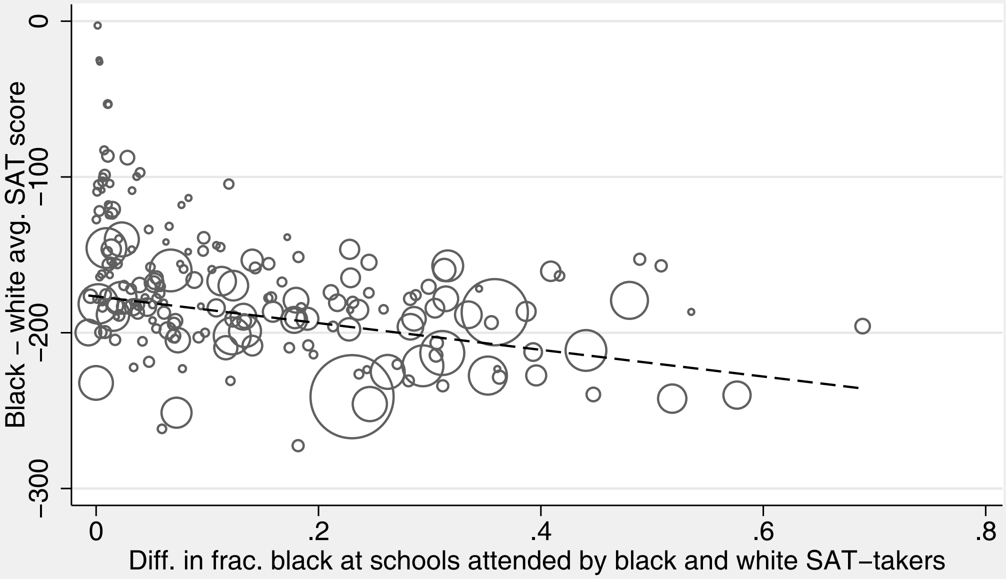
Notes: Sample is schools in metropolitan areas in SAT states. Lines are kernel means, using an Epanechnikov kernel and a bandwidth of 0.02.

Figure 1B. Average residual SATs among white and black students, by school fraction black

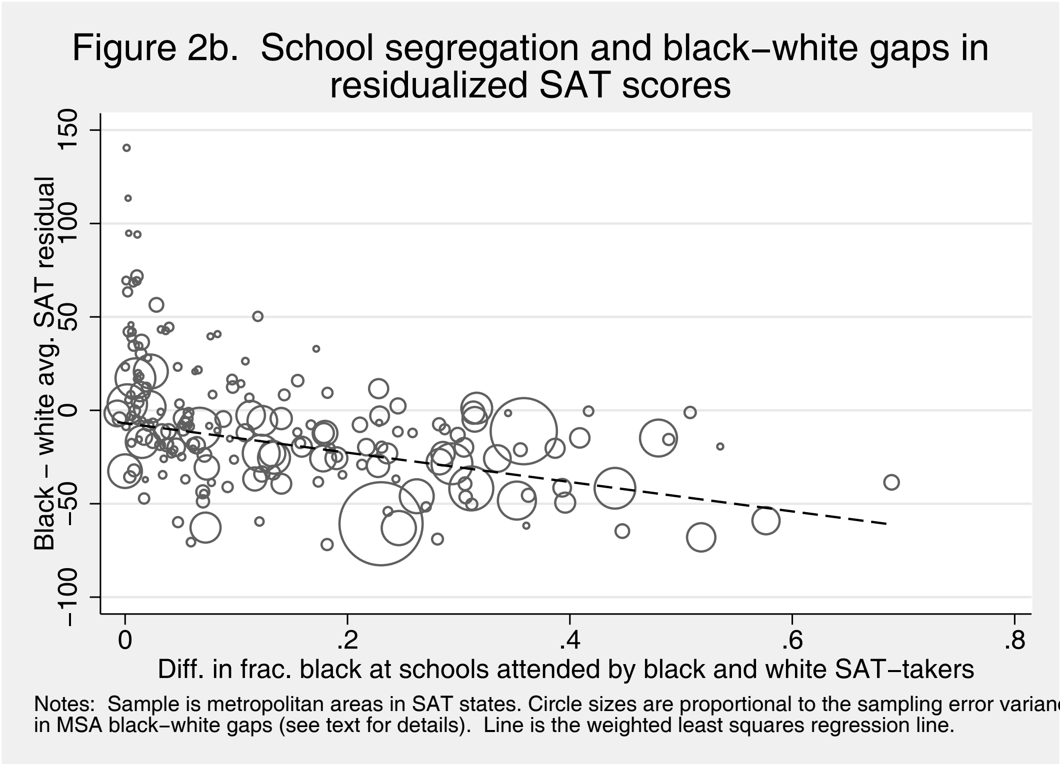


Notes: Sample is schools in metropolitan areas in SAT states. Lines are kernel means, using an Epanechnikov kernel and a bandwidth of 0.02.

Figure 2a. School segregation and black−white gaps in unadjusted SAT scores

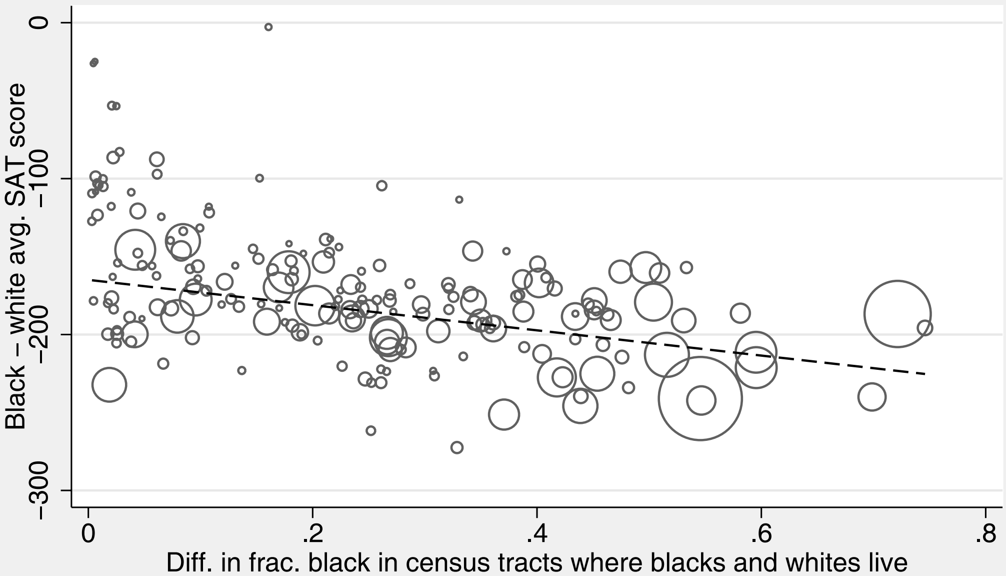


Notes: Sample is metropolitan areas in SAT states. Circle sizes are proportional to the sampling error variance in MSA black−white gaps (see text for details). Line is the weighted least squares regression line.



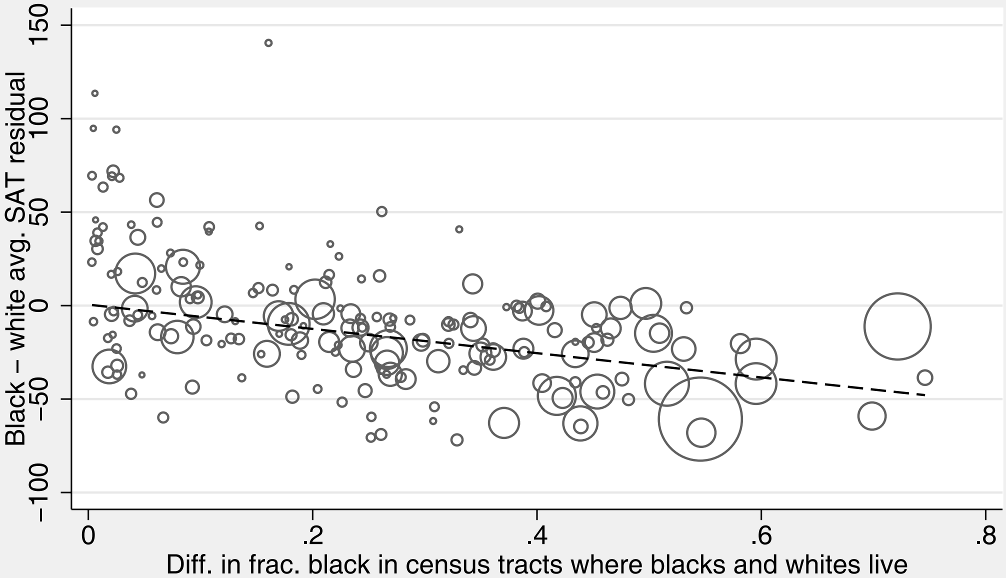
Notes: Sample is metropolitan areas in SAT states. Circle sizes are proportional to the sampling error variance in MSA black−white gaps (see text for details). Line is the weighted least squares regression line.

Figure 3a. Residential segregation and black−white gaps in unadjusted SAT scores



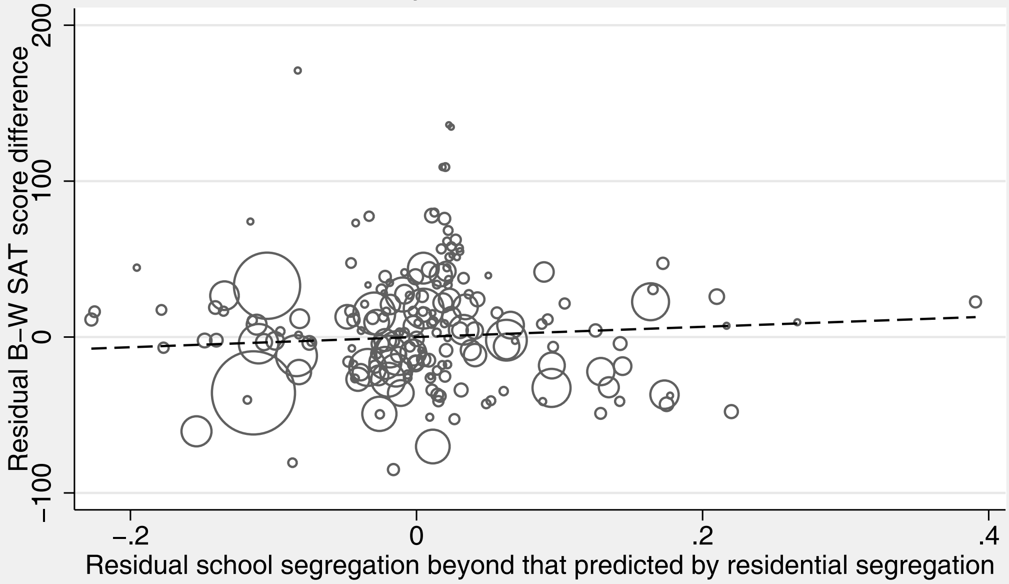
Notes: Sample is metropolitan areas in SAT states. Circle sizes are proportional to the sampling error variance in MSA black−white gaps (see text for details). Line is the weighted least squares regression line.

Figure 3b. Residential segregation and black−white gaps in residual SAT scores



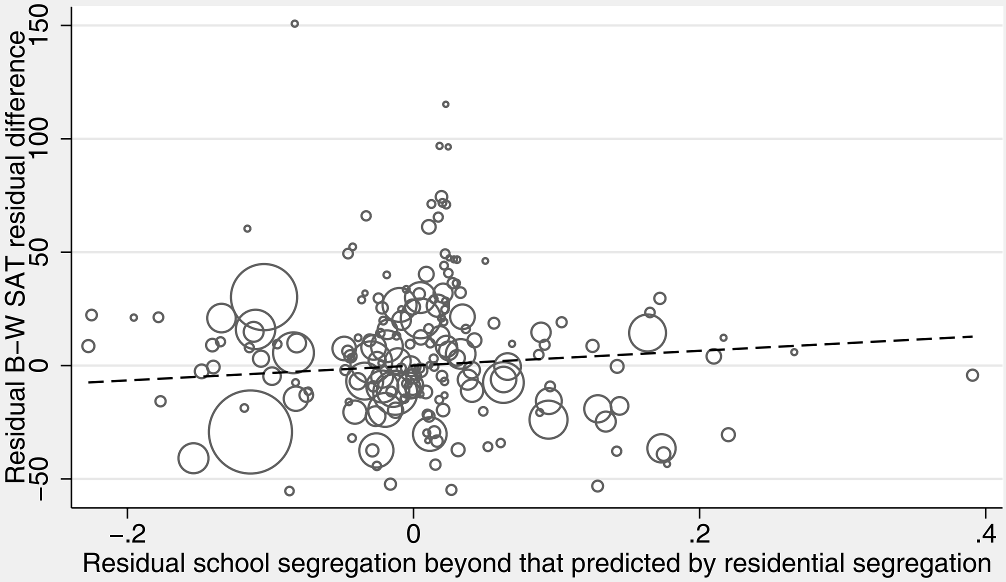
Notes: Sample is metropolitan areas in SAT states. Circle sizes are proportional to the sampling error variance in MSA black−white gaps (see text for details). Line is the weighted least squares regression line.

Figure 3c. Non−residential school segregation and black−white gaps in unadjusted SAT scores



Notes: Sample is metropolitan areas in SAT states. Circle sizes are proportional to the sampling error variance in MSA black−white gaps (see text for details). Line is the weighted least squares regression line.

Figure 3d. Non−residential school segregation and black−white gaps in SAT residuals



Notes: Sample is metropolitan areas in SAT states. Circle sizes are proportional to the sampling error variance in MSA black−white gaps (see text for details). Line is the weighted least squares regression line.

Table 1: Residential and school segregation in representative metropolitan areas

Name	Residential Fraction black in: Blacks' tracts	Whites' tracts	Diff.	High school students Fraction black in: Blacks' schools	Whites' schools	Diff.	SAT-takers Fraction black in: Blacks' schools	Whites' schools	Diff.	SAT State?
Least Segregated Cities										
Missoula, MT MSA	0.6%	0.5%	0.1%	0.4%	0.3%	0.1%				N
Bismarck, ND MSA	0.5%	0.4%	0.1%	0.7%	0.2%	0.5%				N
Provo-Orem, UT MSA	0.7%	0.5%	0.2%	3.5%	0.3%	3.3%				N
Boise City, ID MSA	1.1%	0.8%	0.3%	1.2%	0.8%	0.4%				N
Bellingham, WA MSA	1.5%	1.2%	0.3%	1.6%	1.1%	0.5%	0.9%	0.9%	0.0%	Y
Medford-Ashland, OR MSA	1.0%	0.7%	0.3%	0.7%	0.5%	0.2%	0.6%	0.5%	0.1%	Y
Grand Junction, CO MSA	1.0%	0.7%	0.3%	6.2%	1.0%	5.2%				N
Boulder-Longmont, CO PMSA	1.7%	1.3%	0.4%	1.8%	1.2%	0.6%				N
Bangor, ME MSA	1.4%	1.0%	0.4%	0.9%	0.5%	0.4%	0.9%	0.6%	0.3%	Y
Redding, CA MSA	1.6%	1.2%	0.5%	3.1%	1.2%	2.0%	0.9%	0.8%	0.1%	Y
Least segregated cities with at least 10% blacks										
Jacksonville, NC MSA	29.4%	19.4%	9.9%	32.1%	23.9%	8.2%	31.8%	25.2%	6.6%	Y
Dover, DE MSA	32.0%	20.2%	11.9%	29.1%	22.5%	6.6%	28.2%	22.1%	6.1%	Y
Lawton, OK MSA	33.4%	21.5%	11.9%	29.5%	21.5%	8.0%				N
Charlottesville, VA MSA	27.0%	13.3%	13.7%	27.2%	15.5%	11.6%	23.7%	15.9%	7.8%	Y
Fayetteville, NC MSA	49.4%	34.2%	15.2%	51.5%	39.6%	11.9%	54.4%	36.2%	18.2%	Y
Clarksville-Hopkinsville, TN-KY MSA	35.5%	19.8%	15.7%	33.1%	27.3%	5.8%				N
Killeen-Temple, TX MSA	38.9%	22.5%	16.4%	35.8%	22.1%	13.7%	35.0%	20.7%	14.3%	Y
Danville, VA MSA	44.8%	27.8%	17.0%	46.8%	35.2%	11.6%	43.7%	34.2%	9.4%	Y
Greenville, NC MSA	47.0%	29.5%	17.5%	48.8%	41.2%	7.5%	45.3%	40.2%	5.1%	Y
Myrtle Beach, SC MSA	31.7%	13.4%	18.3%	32.8%	23.2%	9.5%	31.7%	23.8%	7.9%	Y
Most segregated cities										
Philadelphia, PA-NJ PMSA	68.6%	9.1%	59.5%	61.7%	10.1%	51.5%	52.8%	8.7%	44.1%	Y
Miami, FL PMSA	79.4%	19.8%	59.5%	58.5%	22.8%	35.7%	51.8%	22.4%	29.4%	Y
Flint, MI PMSA	72.0%	8.0%	64.1%	70.7%	7.6%	63.1%				N
Cleveland-Lorain-Elyria, OH PMSA	72.4%	6.9%	65.4%	68.9%	7.4%	61.5%				N
Milwaukee-Waukesha, WI PMSA	72.2%	6.0%	66.2%	63.3%	7.7%	55.6%				N
Newark, NJ PMSA	78.2%	8.3%	69.9%	66.7%	7.6%	59.1%	64.5%	6.8%	57.6%	Y
Chicago, IL PMSA	78.0%	7.2%	70.8%	66.5%	7.3%	59.2%				N
New York, NY PMSA	82.6%	10.5%	72.2%	52.2%	11.8%	40.5%	47.5%	11.7%	35.9%	Y
Gary, IN PMSA	80.3%	5.7%	74.6%	76.2%	4.8%	71.4%	73.2%	4.2%	68.9%	Y
Detroit, MI PMSA	81.4%	6.2%	75.1%	80.8%	5.1%	75.7%				N

Notes: Segregation rankings are by difference in fraction black between blacks' and whites' census tracts, as in the fourth column, among 331 MSAs and PMSAs. Fraction black in whites' and blacks' census tracts and schools is blacks divided by blacks plus whites. Cities "with at least 10% blacks," however, require blacks to be 10% of the total population.

Table 2. Summary statistics for cities in the SAT sample

	All Cities		In SAT states		Not in SAT states	
	Mean	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
N	331		189		142	
Population (millions)	2.856	3.010	3.042	3.168	2.412	2.552
Fraction black	0.12	0.09	0.11	0.08	0.14	0.10
Fraction hispanic	0.21	0.21	0.25	0.23	0.09	0.10
log(Mean HH income)	10.98	0.19	10.99	0.20	10.96	0.16
Segregation (Black fraction black - white fraction black)						
Residential (Tract)	0.37	0.21	0.36	0.21	0.39	0.23
High schools	0.26	0.19	0.22	0.16	0.34	0.23
SAT-takers' schools	0.21	0.17	0.20	0.15	0.24	0.20
SAT-taking rate						
All students	0.31	0.15	0.37	0.11	0.16	0.12
White students	0.32	0.14	0.38	0.09	0.16	0.11
Black students	0.21	0.11	0.27	0.07	0.09	0.09
SAT-takers						
Avg. SAT	1033.5	71.2	999.5	45.7	1114.8	53.0
Black-white avg. SAT	-193.3	36.5	-194.0	34.3	-191.6	41.3
Black-white avg. background	-170.8	12.2	-171.2	11.8	-170.0	13.3
Black-white avg. residual SAT	-22.4	30.1	-22.8	26.9	-21.6	36.8

Notes: All summary statistics are weighted by the geometric average of the city's white and black populations.

Table 3. OLS models for school segregation's effect on black-white differences in residual SATs

	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)
Black-white difference: Fr. black in SAT-	-107.7	-89.9	-85.3	-81.4	-74.9	-80.9	-50.3	-44.7
takers' schools	(28.8)	(25.8)	(25.6)	(27.1)	(25.6)	(25.0)	(22.5)	(24.3)
Fr. black "main effect:" Fr. black in white &	14.3	21.2	18.1	30.4	38.2	26.8	31.0	33.2
black SAT-takers' schools	(32.1)	(27.7)	(26.4)	(28.4)	(27.9)	(29.0)	(26.3)	(25.0)
B-W inverse Mills ratio	-61.7	-33.7	-38.2	-28.7	-27.6	-36.2	-21.6	-20.2
	(16.6)	(18.2)	(17.7)	(15.8)	(13.8)	(18.7)	(15.1)	(16.6)
B-W background index (census)		0.80	0.95	-0.12	-0.11			
		(0.26)	(0.42)	(0.45)	(0.46)			
B-W father's education (census)			0.6					
			(4.8)					
B-W mother's education (census)			-8.2					
			(5.7)					
B-W Mothers HS (Census)				-26.4	-22.7			
				(28.2)	(29.4)			
B-W Mothers college (census)				71.4	75.8			
				(52.5)	(52.9)			
B-W Fathers HS (Census)				74.2	69.3			
				(27.1)	(28.0)			
B-W Fathers college (census)				128.9	108.1			
				(47.6)	(46.5)			
B-W family income (Census; $1,000s)			0.22	0.09	0.08			
			(0.27)	(0.23)	(0.23)			
B-W two parents indic. (census)					9.4			
					(24.3)			
B-W father immigrant (census)					21.5			
					(23.2)			
B-W mother immigrant (census)					17.6			
					(29.1)			
B-W background index (SAT-takers)						1.37	-0.07	0.65
						(0.23)	(1.33)	(1.29)
B-W mother's education (SAT-takers)							-1.0	
							(1.3)	
B-W father's education (SAT-takers)							2.2	
							(1.7)	
B-W: Fathers some college (SAT-takers)								125.6
								(74.2)
B-W: Fathers BA+ (SAT-takers)								-20.1
								(96.8)
B-W: Mothers some college (SAT-takers)								-190.8
								(79.2)
B-W: Mothers BA+ (SAT-takers)								106.5
								(59.6)
B-W family income (SAT-takers; in SAT							6.3	4.4
points)							(2.1)	(2.5)
N	186	185	184	185	184	186	186	186
R-squared	0.56	0.60	0.61	0.65	0.68	0.65	0.68	0.70

Notes: All models are weighted by the geometric mean of the city's white and black populations, and include controls for the log of the city population and for the city land area, fraction Hispanic, fraction with some college and BAs, log mean HH income, gini coefficient, and census division effects. All standard errors are clustered on the CMSA.

Table 4. OLS models of the separate effects of the fraction black at white and black students' schools on adjusted SAT scores

Dependent variable	B-W adjusted SAT					Black adj. SAT	White adj. SAT
	(A)	(B)	(C)	(D)	(E)	(F)	(G)
Fr. black in black SAT-takers' schools	-98.9	-81.3	-77.6	-71.3	-63.4	-95.1	0.6
	(24.4)	(23.0)	(22.7)	(23.9)	(22.4)	(23.8)	(10.4)
Fr. black in white SAT-takers' schools	25.1	31.7	31.0	37.4	47.0	12.9	-32.7
	(59.9)	(52.4)	(51.6)	(53.1)	(52.9)	(49.3)	(29.8)
B-W inverse Mills ratio	-62.6	-36.3	-40.6	-31.0	-30.2	11.0	27.7
	(16.1)	(17.6)	(16.9)	(15.6)	(13.8)	(14.8)	(11.8)
Background index (census)		0.76	1.00	-0.04	-0.02	1.02	-0.81
		(0.25)	(0.42)	(0.44)	(0.46)	(0.42)	(0.47)
Father's education (census)			-0.1			-8.3	35.1
			(4.7)			(5.8)	(7.6)
Mother's education (census)			-8.5			-0.3	-0.3
			(5.7)			(5.5)	(12.4)
Mothers HS (Census)				-24.6	-20.9		
				(28.2)	(29.8)		
Mothers college (census)				62.0	67.3		
				(52.8)	(53.2)		
Fathers HS (Census)				78.9	76.0		
				(26.5)	(27.9)		
Fathers college (census)				125.5	106.5		
				(49.1)	(47.2)		
Family income (Census; $1,000s)			0.16	0.03	0.03	0.09	0.36
			(0.27)	(0.23)	(0.23)	(0.39)	(0.35)
Two parents indic. (census)					7.99		
					(25.53)		
Father immigrant (census)					24.89		
					(23.47)		
Mother immigrant (census)					11.21		
					(29.10)		
N	186	185	184	185	184	184	188
R-squared	0.57	0.60	0.62	0.65	0.67	0.62	0.84
P-value: (Black fr. black)=-1*(White fr. black)	0.12	0.22	0.24	0.40	0.68	n/a	n/a

Notes: All models are weighted by the geometric mean of the city's white and black populations, and include controls for the log of the city population and for the city land area, fraction Hispanic, fraction with some college and BAs, log mean HH income, gini coefficient, and census division effects. Listed family background variables are entered as black-white differences in Columns A-E, as black means in F, and as white means in G. All standard errors are clustered on the CMSA.

Table 5. First-stage models for school segregation

	BFB-WFB, all HS students			BFB-WFB, SAT-takers' schools			
	(A)	**(B)**	**(C)**	**(D)**	**(E)**	**(F)**	**(G)**
Black-white difference: Fr. Black in census tract	0.78	0.77	0.77	0.51	0.69	0.67	0.58
	(0.08)	(0.08)	(0.08)	(0.08)	(0.10)	(0.10)	(0.08)
Fr. black "main effect:" Fr. black in white & black residents' tracts	-0.05	0.30	0.30	0.14	0.37	0.36	0.36
	(0.12)	(0.07)	(0.08)	(0.15)	(0.09)	(0.09)	(0.08)
B-W background index (census)			0.0001				
			(0.0007)				
B-W inverse Mills ratio				0.12	-0.08	-0.10	-0.13
				(0.07)	(0.06)	(0.07)	(0.05)
B-W background index (SAT-takers)						-0.0011	0.0069
						(0.0009)	(0.0032)
B-W mother's education (SAT-takers)							-0.003
							(0.004)
B-W father's education (SAT-takers)							-0.006
							(0.003)
B-W family income (SAT-takers; in SAT points)							-0.029
							(0.006)
Controls for 12 MSA-level variables?	n	y	y	n	y	y	y
N	331	331	324	186	186	186	186
R-squared	0.74	0.88	0.88	0.70	0.86	0.86	0.87

Notes: All models are weighted by the geometric mean of the city's white and black populations. MSA-level control variables are the log of the city population and for the city land area, fraction Hispanic, fraction with some college and BAs, log mean HH income, gini coefficient, and census division effects. All standard errors are clustered on the CMSA. In Columns A-C, the dependent variable is the difference in fraction black between schools attended by white and black students; in Columns D-G it is the difference between schools attended by white and black SAT-takers.

Table 6. Decomposition of segregation effects on black-white differences in residual SATs into residential and non-residential components

	(A)	(B)	(C)	(D)	(E)	(F)	(G)
B-W fr. Black: Fitted value using residential	-201.2	-182.8	-177.6	-158.7	-144.0	-170.4	-155.6
segregation	(39.8)	(39.6)	(38.3)	(34.9)	(34.1)	(36.5)	(38.0)
B-W fr. Black: Residual portion	-17.6	-10.5	-3.8	-13.5	-15.2	-3.6	21.7
	(36.1)	(33.2)	(33.1)	(33.5)	(32.5)	(30.3)	(28.6)
Fr. black "main effect:" Fr. black in white &	106.0	104.9	103.4	100.6	100.4	106.6	106.2
black SAT-takers' schools	(44.4)	(41.4)	(40.6)	(39.8)	(38.5)	(41.1)	(38.9)
B-W inverse Mills ratio	-49.8	-30.6	-33.7	-27.1	-26.0	-30.7	-25.1
	(20.8)	(21.5)	(20.1)	(18.8)	(15.4)	(23.3)	(17.6)
B-W background index (census)		0.56	0.47	-0.31	-0.30		
		(0.25)	(0.49)	(0.47)	(0.50)		
B-W father's education (census)			2.49				
			(5.24)				
B-W mother's education (census)			-6.00				
			(6.02)				
B-W Mothers HS (Census)				-28.9	-26.1		
				(27.5)	(28.5)		
B-W Mothers college (census)				69.9	73.8		
				(43.3)	(44.7)		
B-W Fathers HS (Census)				61.8	56.4		
				(29.9)	(30.1)		
B-W Fathers college (census)				111.7	96.8		
				(50.1)	(50.1)		
B-W family income (Census; $1,000s)			0.31	0.16	0.14		
			(0.30)	(0.27)	(0.26)		
B-W two parents indic. (census)					8.8		
					(28.8)		
B-W father immigrant (census)					12.2		
					(24.4)		
B-W mother immigrant (census)					24.2		
					(31.8)		
B-W background index (SAT-takers)						1.09	0.51
						(0.25)	(1.43)
B-W mother's education (SAT-takers)							-0.45
							(1.44)
B-W father's education (SAT-takers)							0.96
							(1.74)
B-W family income (SAT-takers; in SAT							2.47
points)							(2.69)
N	186	185	184	185	184	186	186
R-squared	0.63	0.66	0.67	0.69	0.71	0.69	0.72

Notes: All models are weighted by the geometric mean of the city's white and black populations, and include controls for the log of the city population and for the city land area, fraction Hispanic, fraction with some college and BAs, log mean HH income, gini coefficient, and census division effects. Fitted values of school segregation index are from first-stage regressions on residential segregation index plus all control variables included in the model here. All standard errors are clustered on the CMSA and computed using a version of the Murphy-Topel (1985) estimator.

Table 7. Estimates from trimmed SAT-takers data

	(A)	(B)	(C)	(D)
Black-white difference: Fr. black in SAT-takers' schools	-57.4	-75.1		
	(49.8)	(42.9)		
B-W fr. black: Fitted value using residential segregation			-418.2	-294.8
			(109.2)	(71.2)
B-W fr. black: Residual portion			182.2	93.5
			(78.4)	(50.1)
Fr. black "main effect:" Fr. black in white & black SAT-takers' schools	-215.9	-58.0	61.6	122.2
	(108.2)	(54.8)	(105.9)	(82.7)
B-W background index (census)	1.22	1.09	0.64	0.48
	(0.52)	(0.70)	(0.58)	(0.79)
B-W mother's education (census)		13.4		19.5
		(11.8)		(10.2)
B-W father's education (census)		6.69		5.95
		(5.64)		(6.73)
B-W family income (Census; $1,000s)		-0.2		-0.2
		(0.4)		(0.6)
N	152	150	152	150
R-squared	0.73	0.76	0.81	0.82

Notes: All models are weighted by the geometric mean of the city's white and black populations, and include controls for the log of the city population and for the city land area, fraction Hispanic, fraction with some college and BAs, log mean HH income, gini coefficient, and census division effects. Dependent variable and segregation indices are constructed from trimmed SAT samples in which scores are dropped for all but the highest-scoring 30% of white students and 20% of black students; MSAs with white SAT-taking rates below 30% or black SAT-taking rates below 20% are not used. Fitted values of school segregation index are from first-stage regressions on residential segregation index plus all control variables included in the model here. All standard errors are clustered on the CMSA and computed using a version of the Murphy-Topel (1985) estimator.

Table 8. Estimates of residential and non-residential school segregation effects on black-white differences in alternative outcome measures

	Employed	In school	Employed or in school	Finished HS or in school
	(A)	(B)	(C)	(D)
B-W fr. black: Fitted value using residential segregation	-16.4	-8.0	-11.6	-9.1
	(5.0)	(4.9)	(4.7)	(3.7)
B-W fr. black: Residual portion	7.8	-1.0	6.6	0.4
	(4.4)	(5.4)	(4.1)	(3.2)
Fr. black "main effect:" Fr. black in white & black SAT-takers' schools	22.0	2.5	12.9	7.1
	(5.3)	(5.6)	(5.3)	(4.1)
B-W background index (census)	0.09	-0.07	0.01	0.17
	(0.10)	(0.13)	(0.11)	(0.07)
B-W mother's education (census)	-0.70	-0.05	-1.01	-0.55
	(1.49)	(1.33)	(1.21)	(1.06)
B-W father's education (census)	0.56	1.85	1.77	-0.11
	(0.79)	(1.29)	(1.01)	(0.64)
B-W family income (Census; $1,000s)	-0.028	0.139	0.089	-0.003
	(0.062)	(0.087)	(0.067)	(0.045)
N	319	319	319	319
R-squared	0.37	0.32	0.31	0.29

Notes: All models are weighted by the geometric mean of the city's white and black populations, and include controls for the city fraction Hispanic, fraction with some college and BAs, log mean HH income, gini coefficient, and census division effects. Fitted values of school segregation index are from first-stage regressions on residential segregation index plus all control variables included in the model here. All standard errors are clustered on the CMSA and computed using a version of the Murphy-Topel (1985) estimator. Outcome variables are differences between blacks and whites in percentage with the characteristic, and so range in principle from -100 to 100. Each is measured over 16-24 year olds in the 2000 census, assigned to the metropolitan area where they lived in 1995.

Table 9. Three-level decompositions using Welch & Light desegregation data as an additional instrument

	First Stage models for fr. Black at Blacks' - Whites' schools				Control function Estimates	
	All Public HS Students		SAT-takers schools			
	(A)	**(B)**	**(C)**	**(D)**	**(E)**	**(F)**
Instruments						
Black-white difference: Fr. Black in census tract	0.832 (0.134)	0.736 (0.112)	0.790 (0.154)	0.672 (0.112)		
Change in dissimilarity index induced by major desegregation plans (/100)		0.241 (0.077)		0.229 (0.073)		
Endogenous variable: Predicted values and residual						
B-W fr. black: Fitted value using residential segregation					-132.5 (48.9)	-132.5 (57.3)
B-W fr. black: Additional portion predicted from desegregation effect						64.5 (104.9)
B-W fr. black: Residual portion					7.1 (37.4)	-8.7 (34.8)
Control variables						
Fr. black "main effect:" Fr. Black in whites' & blacks' schools	0.48 (0.17)	0.43 (0.14)	0.50 (0.22)	0.53 (0.16)	62.9 (56.5)	62.9 (71.0)
W-B inverse Mills ratio			-0.33 (0.15)	-0.33 (0.13)	-4.0 (38.2)	-4.0 (41.9)
N	60	60	60	60	60	60
R-squared	0.897	0.917	0.899	0.921	0.819	0.822

Notes: Sample is metropolitan areas in SAT states containing at least one district in the Welch and Light (1987) sample. All models are weighted by the geometric mean of the city's white and black populations, and include controls for the black-white gap in family background, mother's education, father's education, and family income, all measured from census data; for the log of the city population; and for the city land area, fraction Hispanic, fraction with some college and BAs, log mean HH income, gini coefficient, and census division effects. The Fr. Black "main effect" is averaged over all HS students' schools in columns A and B, and over SAT-takers' schools in C-F. All standard errors are clustered on the CMSA and computed using a version of the Murphy-Topel (1985) estimator.

Table 10. Decompositions using historical residential segregation

Sample	All cities	Cities with 1990 data		Cities with 1980 data		Cities with 1970 data		Cities with 1960 data		Cities with 1950 data	
	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)
Fr. black "main effect:" Fr. black in	103.4	102.2	69.2	94.5	90.5	74.4	104.1	52.9	85.8	-20.7	9.8
white & black SAT-takers' schools	(40.6)	(43.0)	(40.7)	(42.1)	(42.3)	(48.0)	(49.8)	(60.1)	(69.2)	(72.9)	(113.8)
Portion of school segregation predicted from residential segregation											
Using 2000 residential segregation	-177.6	-174.6		-154.8		-144.6		-94.5		-84.2	
	(38.3)	(38.3)		(38.5)		(40.1)		(49.9)		(33.7)	
Using 1990 residential segregation			-141.0								
			(39.1)								
Using 1980 residential segregation					-150.6						
					(43.4)						
Using 1970 residential segregation							-173.9				
							(45.9)				
Using 1960 residential segregation									-125.7		
									(54.9)		
Using 1950 residential segregation											-120.9
											(126.7)
Residual school segregation	-3.8	-5.7	-39.4	-10.6	-41.6	-11.1	-34.0	-31.0	-34.9	6.4	-39.5
	(33.1)	(31.7)	(27.6)	(32.6)	(28.8)	(37.2)	(32.9)	(43.5)	(42.1)	(54.0)	(27.4)
N	184	170	170	157	157	112	112	81	81	37	37
R-squared	0.67	0.69	0.66	0.68	0.66	0.74	0.73	0.76	0.77	0.89	0.89

Notes: All models are weighted by the geometric mean of the city's white and black populations, and include controls for the black-white gap in SAT-taking inverse Mills ratios; for the black-white gap in family background, mother's education, father's education, and family income, all measured from census data; for the log of the city population; and for the city land area, fraction Hispanic, fraction with some college and BAs, log mean HH income, gini coefficient, and census division effects. Instruments in models using historical (1990 and prior) residential segregation measures are the isolation and dissimilarity indices calculated by Cutler, Glaeser, and Vigdor (1999). All standard errors are clustered on the CMSA and computed using a version of the Murphy-Topel (1985) estimator.

Table 11. Decomposition model estimates of residential and school segregation's effects on black-white differences in school resources and teacher characteristics

	School resources (CCD)					Teacher characteristics (SASS)			
	PP Expenditures ($1,000s)	PP Instructional expenditures ($1,000s)	PP Revenue ($1,000s)	Teacher / pupil ratio * 1000	Fraction white	Avg. salary ($1,000s)	Avg. experience	Certif. in main assignment (HS only)	BA: Educ. Major
	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)
B-W fr. black: Fitted value using residential segregation	2.25	1.11	2.45	1.14	-0.63	-2.28	2.01	-0.16	-0.32
	(0.82)	(0.28)	(0.84)	(1.08)	(0.14)	(5.80)	(2.94)	(0.14)	(0.13)
B-W fr. black: Residual portion	0.24	0.52	1.28	1.36	-0.29	-1.39	3.43	-0.12	0.05
	(0.76)	(0.35)	(0.83)	(0.99)	(0.14)	(4.93)	(2.97)	(0.13)	(0.12)
Fr. black "main effect:" Fr. black in white & black SAT-takers' schools	-3.11	-1.49	-3.40	-4.19	-0.17	4.30	4.13	0.41	0.23
	(0.97)	(0.41)	(1.09)	(1.46)	(0.16)	(6.22)	(3.31)	(0.16)	(0.14)
B-W background index (census)	-0.01	-0.01	-0.01	0.01	0.00	0.07	-0.01	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.06)	(0.04)	(0.00)	(0.00)
B-W mother's education (census)	0.00	0.01	0.07	-0.37	-0.01	-1.54	0.10	-0.01	-0.02
	(0.13)	(0.07)	(0.14)	(0.24)	(0.02)	(0.84)	(0.53)	(0.03)	(0.02)
B-W father's education (census)	0.07	0.04	0.06	0.02	-0.01	-0.39	-0.18	-0.03	0.01
	(0.09)	(0.03)	(0.08)	(0.12)	(0.02)	(0.60)	(0.34)	(0.02)	(0.02)
B-W family income (Census; $1,000s)	0.002	0.000	0.001	0.003	0.000	-0.020	0.019	0.002	-0.002
	(0.007)	(0.003)	(0.006)	(0.008)	(0.002)	(0.036)	(0.033)	(0.002)	(0.001)
N	319	319	319	301	316	316	316	303	316
R-squared	0.31	0.44	0.38	0.28	0.58	0.10	0.12	0.23	0.12

Notes: Dependent variable in each column is the difference between the average of the indicated variable in black students' schools (districts in A-C) and that in white students' schools. All models are weighted by the geometric mean of the city's white and black populations, and include controls for the city fraction Hispanic, fraction with some college and BAs, log mean HH income, gini coefficient, and census division effects. Fitted values of school segregation index are from first-stage regressions on residential segregation index plus all control variables included in the model here. All standard errors are clustered on the CMSA and computed using a version of the Murphy-Topel (1985) estimator.

Table 12. Residential and non-residential school segregation's effects on measures of tracking and honors course-taking

Dependent variable:	SASS HS teachers	SAT-takers			
	Class assigned based on student ability	=100 if honors courses in		=100 if plan to claim adv. / exempt status in	
		Math	English	Any subject	Math or English
	(A)	(B)	(C)	(D)	(E)
Panel A: Avg. at all high schools					
B-W fr. Black: Fitted value using residential segregation	0.14				
	(0.13)				
B-W fr. Black: Residual portion	0.01				
	(0.13)				
Panel B: Avg. among black SAT-takers					
B-W fr. Black: Fitted value using residential segregation		-6.21	-8.15	3.17	0.62
		(7.39)	(11.19)	(6.11)	(6.62)
B-W fr. Black: Residual portion		9.37	-6.47	1.41	-1.56
		(6.86)	(10.75)	(5.59)	(5.23)
Inverse Mills' ratio in MSA black SAT-taking rate		14.54	16.61	0.57	4.72
		(3.33)	(5.41)	(3.10)	(2.60)
Panel C: Avg. among white SAT-takers					
B-W fr. Black: Fitted value using residential segregation		12.42	13.46	0.39	3.85
		(7.44)	(9.82)	(7.04)	(6.22)
B-W fr. Black: Residual portion		-16.15	-23.76	-18.68	-17.30
		(9.55)	(12.00)	(6.81)	(6.02)
Inverse Mills' ratio in MSA white SAT-taking rate		19.67	26.07	0.61	4.64
		(5.96)	(7.00)	(5.08)	(4.82)
Panel D: Difference between black and white averages					
B-W fr. Black: Fitted value using residential segregation		6.47	5.08	6.65	5.02
		(6.24)	(7.18)	(5.69)	(4.86)
B-W fr. Black: Residual portion		17.69	9.62	21.64	16.08
		(5.91)	(5.43)	(4.24)	(3.53)
B-W Inverse Mills' ratio		-5.52	-3.24	-3.49	-3.04
		(4.11)	(4.00)	(3.02)	(2.41)
Panel E: Difference between black and white averages within schools					
B-W fr. Black: Fitted value using residential segregation		-1.58	-5.63	-7.53	-1.13
		(4.83)	(12.10)	(6.08)	(5.88)
B-W fr. Black: Residual portion		8.34	16.22	9.19	6.44
		(4.31)	(8.23)	(4.37)	(5.44)
B-W Inverse Mills' ratio		-12.59	-14.25	-5.45	-14.45
		(3.53)	(5.27)	(2.70)	(3.42)

Notes: All models are weighted by the geometric mean of the city's white and black populations, and include controls for the city fraction Hispanic, fraction with some college and BAs, log mean HH income, gini coefficient, and census division effects, and for the average (Panel A), black average (Panel B), white average (C), or black-white gap (D, E) in family background, mother's education, father's education, and family income, all measured from census data. All standard errors are clustered on the CMSA and computed using a version of the Murphy-Topel (1985) estimator.