NBER WORKING PAPER SERIES

SEGREGATION AND THE BLACK-WHITE TEST SCORE GAP

Jacob Vigdor
Jens Ludwig

Working Paper 12988
http://www.nber.org/papers/w12988

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2007

Paper prepared for the conference "Stalled Progress: Inequality and the Black-White Test Score Gap," Russell Sage Foundation, November 16-17, 2006. We are grateful to conference participants and especially Adam Gamoran, Katherine Magnuson, and Jane Waldfogel for helpful comments on an earlier draft. Any errors and all opinions are of course our own. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Segregation and the Black-White Test Score Gap
Jacob Vigdor and Jens Ludwig
NBER Working Paper No. 12988
March 2007
JEL No. I2,J15,R2

ABSTRACT

The mid-1980s witnessed breaks in two important trends related to race and schooling. School segregation, which had been declining, began a period of relative stasis. Black-white test score gaps, which had also been declining, also stagnated. The notion that these two phenomena may be related is also supported by basic cross-sectional evidence. We review existing literature on the relationship between neighborhood- and school-level segregation and the test score gap. Several recent studies point to a statistically significant causal relationship between school segregation and the test score gap, though in many cases the magnitude of the relationship is small in economic terms. Experimental studies, as well as methodologically convincing non-experimental studies, suggest that there is little if any causal role for neighborhood segregation operating through a mechanism other than school segregation.

Jacob Vigdor
Terry Sanford Institute of Public Policy
Duke University
Durham, NC 27708
and NBER
jacob.vigdor@duke.edu

Jens Ludwig
Georgetown University
Public Policy Institute
3520 Prospect Street, NW
Washington, DC 20007
and NBER
ludwigj@georgetown.edu

I. Introduction

Disparities in educational outcomes between African-Americans and whites declined steadily for most of the 20th century, but this progress has halted or even reversed in recent years (Neal, 2005). Understanding why the black-white test score gap narrowed over time, and why this progress stalled during the 1990s, is crucial for designing policies capable of further reducing inequality in schooling outcomes in the United States. Given the widely documented association between educational outcomes and earnings, health and crime, successful efforts to further reduce the black-white gap in schooling would undoubtedly have far-reaching consequences for society as a whole.

This essay considers the role of school and neighborhood segregation in explaining trends in the black-white schooling gap. Key to the Supreme Court's transformational decision 50 years ago in *Brown v. Board of Education* is the assumption that racial segregation within the public schools contributes to black-white inequality in schooling outcomes for children. Attending a disproportionately minority school might affect the motivation of students and their perceptions of the larger opportunity structure that they face in society, as well as potentially affecting their exposure to high-quality school resources or even the academic climate in the school.

It is possible that the race of one's schoolmates might simply be a stand-in for their academic achievement level, which could affect the way or rate at which teachers present material or the productivity of student study groups, or as a stand-in for their socio-economic status (SES). In fact the influential report by James Coleman in 1966 argued that a school's SES composition was at least as important in explaining inequality in student achievement as is school racial composition (Coleman et al., 1966). Understanding the distinct influences of school racial versus social class composition is relevant because some policies focused on reducing school racial segregation may not have very large impacts on school socio-economic composition, and vice versa.

Measures of school segregation by either race or social class could matter primarily because they are proxies for the racial or class composition of the local neighborhood. Since most children attend public schools that draw students from the local area, in national data there will be a great deal of cross-

sectional correlation between school and neighborhood measures of race or SES segregation. The neighborhood social environment could matter above and beyond the composition of the local public school by shaping the youth social norms that help shape children's behavior, particularly since so much socializing occurs outside of school, as well as exposure to local role models (Wilson, 1987). Understanding the distinct influences of school versus neighborhood environments is important for policy because policies like public or private school choice have the potential to change school but not neighborhood social compositions for children, while some housing mobility interventions might generate larger changes in neighborhood than school characteristics (see for example Sanbonmatsu et al., 2006).

Figure 1 presents a basic conceptual framework outlining the potential causal mechanisms linking segregation to test score gaps. Neighborhood segregation could affect these outcomes directly or through its influence on school segregation. Direct links between neighborhood and test scores might be brought about through neighborhood-level deviant peer influence, role model effects, or impacts on parental income derived from spatial mismatch-type effects. Direct links between school segregation and test scores could be mediated by differences in school input quality, or by peer influence operating at the school or classroom level.

Thinking carefully about the particular mechanism through which social environment affects children is also important for understanding patterns in black-white student outcomes because different measures of segregation have been following different trajectories in recent years. Neighborhood racial and economic segregation actually declined during the 1990s, although these national trends mask important differences by region and so do not necessarily imply that neighborhood segregation is not relevant for understanding the slowdown in narrowing of the black-white test score gap (Glaeser and Vigdor, 2003; Jargowsky, 2003).

Controversy exists regarding whether school segregation has increased or decreased over this same time period. Increases reported in some studies (e.g. Clotfelter 2004, Orfield and Eaton 1996) confound the increasing diversity of the student body in American public schools with increases in the separation of blacks from students of other races (Logan, 2004). Using measures more commonly

accepted in the sociological and economic literature on segregation, there have been very slight decreases in the degree of segregation over the past two decades. Regardless of measure, it is clear that school segregation has declined more slowly than neighborhood segregation over the same time period. The failure of school segregation to track neighborhood segregation reflects a broad decline in governmental efforts to integrate public schools (Orfield and Eaton 1996, Clotfelter, Ladd and Vigdor 2005).

In our judgment there is stronger evidence to support effects on student achievement from school versus neighborhood segregation. The current evidence implicating school race segregation is stronger than the available research on the effects of school segregation by social class, although the specific mechanisms through which racial composition affects student outcomes remain poorly understood.

The best available evidence suggests that a 10% increase in the share of a school's student body that is black would reduce achievement test scores for black students by between .025 and around .08 standard deviations, and reduce test scores for whites by perhaps one-quarter to two-fifths as much. Thus, if school segregation had displayed a decrease in the 1990s commensurate with the observed decrease in neighborhood segregation, our best estimate is that the black-white test score gap would be roughly 0.01 to 0.02 standard deviations narrower.

II. Basic Facts

A complex set of factors determines the degree of interracial exposure witnessed by students of a given race, or the degree of economic diversity experienced by individuals of a given socioeconomic status. Residential segregation has a direct impact on the characteristics of a student's neighbors, and indirectly influences classmate and schoolmate characteristics by affecting the cost to a school district of achieving racial balance across campuses. Beyond this residential component, the structure of local government in a local area, particularly whether school districts serve large or small geographic areas, and a number of other policies implemented by districts themselves can influence classroom racial composition. The past decade has witnessed a general decline in residential segregation, coupled with decreased integration effort on the part of districts. On net, these two effects led to a modest increase in

interracial contact among the nation's black students. In 1987, the average black public school student attended a campus that was 51.5% black. By 2003, this number had declined to 48.3%. The exposure of non-Hispanic white students to black students remained nearly constant over the same time period, thus much of the change in the racial composition of the typical black student's school can be attributed to increasing Asian and Hispanic enrollments.

Black-nonblack residential segregation in the United States peaked around 1970 and has been declining ever since (Cutler, Glaeser and Vigdor, 1999; Glaeser and Vigdor, 2003). Figure 2 illustrates the time series pattern of residential segregation over the 20th century, using perhaps the most common measure of segregation, the dissimilarity index, which is defined as follows:

$$(1)\quad D = \frac{1}{2}\sum_i \left| \frac{black_i}{black_{total}} - \frac{non-black_i}{non-black_{total}} \right|,$$

where *i* indexes all the neighborhoods within a metropolitan area, or all the schools within a district, $black_i$ represents the black population in a neighborhood or enrollment in a school, $black_{total}$ represents the overall black population of the metropolitan area or black enrollment in the district, and $non\text{-}black_i$ and $non\text{-}black_{total}$ represent analogous counts for individuals who are not black. The dissimilarity index can be interpreted as the fraction of black individuals who would have to be moved between neighborhoods or schools in order to attain a perfectly even balance across the metropolitan area or district.[1] In 1970, the average black resided in a metropolitan area where nearly four-fifths of the group would have to switch neighborhoods to achieve evenness. By 2000, that fraction had declined to just over three-fifths. In the 1990s alone, the average black witnessed a decline in neighborhood percent black from 56% to 51% (Glaeser and Vigdor 2003). This trend was accompanied by a modest increase in the percent black in the neighborhood occupied by the average white.

The decline in residential segregation since 1970 has generally been attributed to the enactment and enforcement of fair housing laws in the 1960s, along with other measures that reduced the severity of

[1] The dissimilarity index has been criticized along a number of dimensions. It is preferred here primarily for its ease of computation and of interpretation. For a more complete discussion of segregation measures and their relative advantages, see Massey and Denton (1988) and Echenique and Fryer (2005).

discrimination in the housing and mortgage markets. For the most part, the decline in segregation was accomplished by the entry of modest numbers of black families into new suburban developments or into existing neighborhoods that had been entirely white (Cutler, Glaeser and Vigdor 1999). Declines in segregation have been steepest in growing metropolitan areas, where new developments make up a larger proportion of the housing stock. The vast majority of neighborhoods that were predominantly black as of 1970 remain predominantly minority; in many areas these neighborhoods have depopulated substantially over the past thirty years.

Declines in racial residential segregation since 1970 are particularly noteworthy because they have occurred both in periods of increasing and decreasing economic inequality. While the coexistence of decreasing segregation and increasing inequality may seem paradoxical at first, it is important to note that the racial segregation that existed prior to the Civil Rights movement was pervasive along all levels of socioeconomic status. As of 1970, black high school dropouts and blacks with at least some postsecondary education experienced nearly identical dissimilarity levels (Cutler, Glaeser and Vigdor 1999, Table 2). Reductions in segregation after this time period were most rapid for the most educated group of blacks – consistent with the notion that black suburbanization explains most of the decline in segregation. Less-educated blacks, who presumably have been harmed the most by broad increases in inequality, witnessed more modest declines in segregation after 1970.

The decline in residential segregation has been accompanied, for a good part of the same time period, by an increase in similar measures of socioeconomic segregation. Segregation by household income increased between 1970 and 1990, and fell slightly between 1990 and 2000 (Jargowsky, 1996; Jargowsky 2003; Watson 2006). Trends in economic segregation thus track changes in the income distribution much more closely than trends in racial segregation.

The literature on school segregation has developed quite independently from the literature on residential segregation. One consequence of this disconnect is a general tendency for studies of school segregation to use measures and indices that are relatively uncommon elsewhere in the literature. While there have been reports of increased school segregation since the late 1980s or early 1990s, much of the

reported increases are unique to a single segregation measure, the fraction of black students attending majority nonwhite schools (Clotfelter, 2004). Owing to the nation's increased racial diversity, students of all races now find themselves attending schools with a high fraction of non-black nonwhites (Logan 2004, Clotfelter, Ladd and Vigdor 2005). Other measures of segregation show little or no time trend over this period. There is at least some evidence that the retreat from court-ordered busing in the 1990s led to segregation levels being higher than they otherwise would have been, which can possibly explain why school segregation remained effectively constant even while residential segregation was declining.

Which is the "right" method of measuring segregation? While there have been some attempts to ground segregation indices in economic or other social scientific models of behavior (see, for example, Echenique and Fryer 2005), there is surprisingly little guidance in the existing research literature regarding the relative merits of different measures. The dissimilarity index is clearly a better measure for evaluating the degree of effort a district undertakes to achieve integration in its schools and classrooms; it measures the proportion of the black (or non-black) members of the student body that would need to be moved to achieve perfect evenness. The dissimilarity index can be criticized, however, because it does not map directly into implications for the composition of classrooms and schools occupied by typical students of any race. For a school district that is experiencing no trend in racial enrollment patterns, increases in dissimilarity imply that percent black in the average black student's school is increasing relative to the average white's school. Such implications cannot be drawn when comparing across districts or over time in a district undergoing demographic change.

At first glance, a measure of peer percent nonwhite, such as those commonly used in the school segregation literature, might seem to be the best measure to use in an analysis of the impact of segregation on racial test score disparities. A brief example casts doubt on this supposition. Suppose district A is 75% black, and all schools within it are 75% black. Dissimilarity equals zero, while measures of peer percent nonwhite will be high. District B is 5% black, but all the blacks are concentrated in one school where they form 25% of the student body. The dissimilarity index will be much higher in B, while the peer percent nonwhite measure will be lower. In which district would we expect wider black-white

disparities in test scores? The potential for disparities is clearly higher in district B, since it is in that district that students of different races systematically attend different schools.

As will be seen below, the majority of empirical studies of the effects of racial composition on student achievement look for evidence that the impact of peer racial composition on individual test scores is different for blacks and whites. If increases in peer percent black are detrimental for black students, but inconsequential for white students, then the aggregate achievement-maximizing distribution of students would involve integration rather than segregation. The traditional measure of percent nonwhite is inadequate for using results such as these to infer the impact of segregation on black-white gaps, *because it incorporates no mechanism for contrasting the experience of white students*. Making such an inference requires the incorporation of information on the average experience of white students. The dissimilarity index captures exactly such a contrast in experiences by race.

Another caution regarding the use of a peer percent nonwhite measure is that much of the time-series variation in peer percent nonwhite has been driven by increases in the share of non-black non-whites in the population. Cross-sectional studies of the impact of peer composition on achievement, by contrast, are driven largely by variation in percent black. To accurately judge whether increases in peer percent nonwhite might account for stalled progress in the black-white test score gap, it would be necessary to identify differential impacts of peer percent Asian and Hispanic on test scores for black and white students.

In summary, the dissimilarity index is most useful for analyzing variation in the effort required to achieve perfect integration in school districts or states, and for assessing the potential for racial disparities in school-level conditions within a school district. Peer percent nonwhite measures provide a better idea of the classroom composition actually faced by typical students in a district or state.

Table 1 presents basic information on school segregation levels in the United States as a whole and for a selection of large urban school districts, based on school membership counts found in the Common Core of Data for the school years 1987-88 and 2003-04. This table employs the dissimilarity index as a measure of segregation, thus these values are quite comparable to the neighborhood-level

segregation indices summarized in Figure 2. There are several striking patterns evident in the data on school segregation. First, dissimilarity levels are generally much lower at the school level than they are at the neighborhood level, which reflects the fact that many districts take at least some action to counteract neighborhood segregation in the manner they assign students to schools.[2] Second, there has been virtually no change in the segregation level experienced by the average black student over the 16-year time period covered here. The stability of this mean masks some quite dramatic changes in the patterns evident in individual school districts over time. Among the nation's largest school districts, there are several examples of districts exhibiting double-digit increases in dissimilarity over this period. While we lack comprehensive data on official district busing policies, it is clear that segregation increases witnessed in some of these districts, such as Charlotte-Mecklenburg, directly reflect the decisions of Federal courts.

The relatively low dissimilarity indices within districts do not necessarily imply that schools are integrated by any objective standard. In many parts of the country, a high degree of school segregation results from between-district differences in racial composition. Table 2 documents this pattern by reporting the degree of dissimilarity between school districts for a set of states reporting comprehensive data on racial enrollment patterns in 1987 and 2003. Dissimilarity values tend to be quite high for Northeastern and Midwestern states. This can be explained by the relatively high degrees of residential segregation in those states, coupled with the general tendency for school districts to serve single municipalities, rather than larger areas. While these dissimilarity levels are high, there is also some evidence that they have been decreasing over time in most states, consistent with the general decline in residential segregation witnessed during the same time period.

In the Appendix, we provide analogous tables documenting the degree of dissimilarity experienced by students eligible for the Federal free lunch program. These data reveal trends quite

[2] Direct comparison between segregation levels at the school and neighborhood level should be undertaken with caution, as segregation indices tend to rise as the size of the neighborhood or other unit of observation decreases. In this comparison, neighborhoods are proxied by census tracts, which have an average of 4,000 residents each, implying roughly 800 school-age children per tract. The average school in the CCD data serves about 500 children. Thus, the direct comparison between neighborhood and school segregation likely understates the impact of school district efforts on segregation levels.

similar to racial segregation in schools: within-district dissimilarity shows little overall trend between 1987 and 2003 even though some districts post large increases. Across-district dissimilarity shows evidence of a decline in many states.

Could the relative stability of school segregation in the 1990s, which contrasts with periods of rapid integration in preceding decades, explain the stalled progress in closing the black-white test score gap in the same period? On the face of it, the two time series mesh together well. Cross-sectional evidence is also consistent with a causal link between school segregation and the black-white test score gap, as is evident in Figure 3, which plots the state-level black-white gap in 4th grade NAEP mathematics test scores against state-level across-district dissimilarity for 2003. There is an unmistakable positive association between the two variables; the plotted least-squares regression line indicates that the predicted black-white test score gap is nearly 50% larger in states with the highest levels of across-district dissimilarity as in states with the lowest levels.

Of course, it is inappropriate to assign a causal interpretation to a simple correlation such as the one shown here. Segregation levels are the outcome of a large number of choices, made by households in their residential location decisions and by public school officials. These decisions are quite likely correlated with a number of underlying variables which could easily exert their own influence on test scores. The following sections discuss the proposed causal mechanisms linking segregation levels to educational outcomes, and the most reliable attempts to disentangle such causal channels from other processes.

III. Effects of Neighborhoods

Whether or how neighborhood context affects the long-term life chances of poor families independently of school context remains somewhat unclear. The best evidence currently available suggests that moving into a neighborhood that is relatively less segregated by race or social class has at most very modest effects on children's test scores or other schooling outcomes, at least through the first 5 years or so following the neighborhood change. These findings, together with the fact that on average

neighborhood class and race segregation *declined* somewhat during the 1990s, lead us to conclude that changes in neighborhood conditions are unlikely to have contributed much to the halt in narrowing of the black-white test score gap.

Decades of research throughout the social sciences have documented substantial variation across neighborhoods in adult or child outcomes, even after conditioning on observable individual or family attributes (Sampson et al., 2002; Kawachi and Berkman, 2003; Leventhal and Brooks-Gunn, 2000; Ellen and Turner, 2003; see Vigdor 2006a for a review of numerous recent non-experimental studies of neighborhoods and youth outcomes). A number of causal pathways have been proposed to explain these correlational patterns, including "spatial mismatch" effects whereby adults suffer from residing in a neighborhood with few job opportunities, peer contagion operating at the neighborhood level, and the importance of local role models on the goal orientation of youth. Neighborhoods have also been hypothesized to matter because they partially determine school characteristics; we examine these proposed pathways more exclusively in the next section. Yet these studies, by and large, fail to provide convincing evidence in favor of any of these pathways, because they cannot separate the effects of neighborhoods *per se* from those of unmeasured individual attributes associated with residential selection.

The most compelling evidence for "neighborhood effects" on youth outcomes has come from quasi-experimental studies that exploit neighborhood changes induced by housing programs. One of the best known of these quasi-experiments is Chicago's Gautreaux residential mobility program (Rubinowitz and Rosenbaum, 2001), which resulted from a 1976 Supreme Court decree finding that the Chicago Housing Authority and HUD had engaged in "systematic and illegal segregation." Follow-up interviews with Gautreaux families suggested improved outcomes for those who moved to the suburbs rather than other parts of Chicago, particularly for youth. For example only 5% of suburban movers dropped out of high school compared to 20% of city movers, with an impressively large suburban advantage in college attendance as well (54% versus 21%) (Rubinowitz and Rosenbaum, 2001, p. 163). In principle these findings are relevant for present purposes because suburban communities are relatively more affluent and with a higher proportion of white families. However, more recent studies of Gautreaux show smaller

impacts (Mendenhall, Duncan and Deluca, 2004; Keels, Rosenbaum and Duncan, 2004). Moreover Gautreaux has important limitations: The degree to which the program breaks the link between family preferences and mobility outcomes is unclear, since convincing documentation of the voucher-offer and acceptance process remains difficult to reconstruct. In addition, Gautreaux studies cannot compare outcomes for program movers versus non-movers or non-participants.

Motivated by the encouraging findings from Gautreaux, the U.S. Department of Housing and Urban development (HUD) launched the Moving to Opportunity (MTO) randomized housing-mobility intervention to identify the causal effects of residential mobility interventions on families. MTO has been in operation since 1994 in 5 cities (Baltimore, Boston, Chicago, Los Angeles and New York). Eligibility for the voluntary MTO program was limited to low-income families with children living in public housing or Section 8 project-based housing located within designated high-poverty census tracts in these cities, with poverty rates of 40% or more. Families were recruited through fliers, tenant associations and other means. Almost all of the 4,600 households that signed up were headed by a female, nearly two-thirds of whom were African-American and most of the rest were Hispanic. Three-quarters of household heads were on welfare at baseline, and fewer than half had graduated from high school. While we do not know the immigration status of MTO participants, most speak English.

By random lottery, some families were assigned to a *Section 8 group* that was offered unrestricted Section 8 subsidies (vouchers)[3] that they could use to move to a new private-market apartment of their choice. Other families were assigned to an *Experimental group* that was offered housing search assistance and Section 8 subsidies that could only be used to relocate to a low-poverty Census tract (with a 1990 poverty rate below 10 percent). After one year in a low-poverty tract, Experimental families could use their subsidies to move elsewhere. Families in both groups were given four to six months (depending on the site) to submit a request for approval of an eligible apartment they would like to lease, and the

[3] The Section 8 program provides housing assistance through rental certificates or vouchers. Income eligibility is usually set to 50 percent of local median income, with subsidies equal to the difference between 30% of tenant income and an area-wide threshold established by HUD. Since MTO started the voucher program has been renamed the Housing Choice Voucher (HCV) program, although we will use the terminology in effect at the start of MTO.

apartment then had to pass a quality inspection. Families could continue to receive rental subsidies so long as they remained income eligible. A third *Control group* received no special assistance from MTO but remained eligible for public or project-based housing and other social programs.

Of those households assigned to the Experimental group 47 percent used a MTO voucher to relocate to a low-poverty Census tract, while 62 percent of those assigned to the Section 8 group relocated through MTO. The fact that the "compliance rate" is higher for the Section 8 group is presumably due in large part to the fact that the vouchers available to them through MTO were not geographically restricted. In both treatment groups, compared to non-compliers the compliers tend to be younger, relatively more dissatisfied with their baseline neighborhoods, and have fewer children. (For details see Shroder, 2002 and Feins and Shroder, 2005).

The explicit goal of MTO was to help move families into less economically distressed communities, and as shown in Table 3, by this measure MTO was successful. One year after random assignment families in the two MTO treatment groups live in Census tracts with average poverty rates 11-13 percentage points (25-30%) below those of the Control group. The gap declines somewhat over time in part because of subsequent mobility among all groups. But even 6 years out the treatment-control differences in tract poverty equal 7-8 percentage points (20% of the control mean), while the differences in cumulative exposure to neighborhood poverty (duration weighted averages) are 9-10 percentage points (20-25%). The MTO Experimental treatment in particular also increases households' exposure to affluent (college-educated) neighbors, which Duncan et al. (1994) suggest have distinct effects from exposure to poor neighbors. But unlike with Gautreaux, MTO generates surprisingly modest changes in neighborhood racial integration. Evidence also suggests that the treatment-control differences in the characteristics of public schools attended were also very modest. Thus, evidence of treatment-control differences in youth outcomes from a study such as MTO would point more clearly to causal pathways that did not involve school characteristics.

While the post-randomization mobility of Control families somewhat reduces over time MTO's "treatment dose" on neighborhood characteristics, these patterns may also make the demonstration results

somewhat more policy relevant: The counterfactual neighborhood conditions for families in the MTO Experimental and Section 8 groups is no longer restricted to just the highest-poverty census tracts in which a relatively small set of public housing families reside.

Short-term findings for MTO (2-3 years after assignment) are generally consistent with the predictions of previous research suggesting that moving to a less distressed community will on net improve adult and child outcomes (Katz, Kling and Liebman, 2001; Ludwig, Duncan and Hirschfield, 2001, Ludwig, Ladd and Duncan, 2001, Leventhal and Brooks-Gunn, 2004).

However the interim MTO evaluation conducted 4-7 years after random assignment finds a pattern of results that is more complicated than less segregated neighborhoods, better youth outcomes. Table 4 summarizes results for MTO youth, drawing on results reported in Orr et al. (2003), Sanbonmatsu et al. (forthcoming), Kling, Ludwig and Katz (2005), and Kling, Liebman and Katz (2006). These data yield no evidence for statistically significant differences across MTO-assigned groups in reading or math achievement scores measured using the Woodcock-Johnson-Revised tests. While a growing body of research in developmental psychology, neuroscience and economics suggest that children may be more sensitive to environmental changes early in life (Shonkoff and Phillips, 2000, Carniero and Heckman, 2003, Knudsen et al., 2006), Sanbonmatsu et al. (forthcoming) find no evidence that treatment-control differences in test scores vary by age at random assignment, at least within the range of ages found among those MTO children who were administered tests. Those MTO participants who were very young children at baseline were not tested as part of the interim study, and so it remains possible that MTO effects on their achievement may be more pronounced.

There is a hint in the MTO data that female youth assigned to the experimental group may be somewhat more likely to graduate from high school than those assigned to the control group, although the point estimate is not statistically significant and the estimated experimental-treatment effect on graduation for male youth is *negative* (and also not statistically significant). More generally the interim MTO evaluation finds that treatment-group assignment improves a variety of behavioral outcomes for female youth but on balance has detrimental effects on behavior for male youth.

One way to reconcile the MTO and Gautreaux findings is that the latter engendered more neighborhood racial segregation than the former, and so perhaps it is the racial rather than social class composition of a community that matters for youth schooling outcomes. In fact one recent study suggests that exposure to more affluent neighbors may actually have deleterious effects on happiness (Luttmer, 2005), which raises the possibility that any beneficial effects from MTO moves on behavior due to reductions in neighborhood racial segregation could be masked in the overall program impact estimates by offsetting deleterious effects from reductions in class segregation.

This possibility can be tested by exploiting variation across MTO sites in treatment effects on mobility outcomes to use site-treatment group interactions as instruments for specific neighborhood characteristics (see Kling, Liebman and Katz, 2007). Ludwig and Kling (2007) find that census tract minority composition is a stronger predictor of youth violent criminal behavior than is either the tract poverty rate or local-area crime rate. However in unpublished results they find no evidence for an association between tract minority composition and children's achievement test scores.

While MTO represents in our view perhaps the best evidence available to date about neighborhood effects on youth outcomes, it is important to keep in mind that the MTO program population consists of those very low-income minority families living in some of the nation's worst housing projects who volunteered to participate in this mobility demonstration. In principle neighborhood effects may be different on other populations. But there is a plausible case to be made that the MTO families are those who expect to benefit the most from moving, which might suggest that MTO estimates represent an upper bound for the effects of neighborhood change on other, similarly disadvantaged families.[4]

An alternative concern about the MTO findings is that the demonstration examines the effects of neighborhood change, and so in principle might confound the effects of mobility with those of neighborhood composition per se. One argument against this concern comes from Liebman, Katz and

[4] In addition Page, Solon and Duncan (2000) find relatively modest correlations in outcomes for youth living in the same primary sampling unit within the Panel Study of Income Dynamics (PSID).

Kling (2004), who exploit variation across MTO demonstration sites in treatment effects on neighborhood poverty rates and find evidence of a "treatment dose-response" relationship between changes in tract poverty and MTO participant outcomes.

Another data point against the "moving hypothesis" comes from Kling, Ludwig and Katz (2005), who examine longitudinal arrest histories for MTO youth and find that in the short run assignment to the experimental rather than control group reduces male arrests for violent crime, and leads to a detrimental experimental effect on property crime arrests only starting around three or four years after random assignment. This temporal pattern in MTO impacts runs contrary to what we would expect if the disruption of moving is driving behavioral problems among experimental group males. More generally, given that the U.S. population became only modestly less poor during the 1990s and experienced an increase in the share minority, national declines in neighborhood racial and class segregation must be driven more by re-sorting of families across neighborhoods than by changes in the characteristics of persistent residents of a given neighborhood. In this case any consequences of social churning and the difficulties of developing new social ties among neighborhood residents that is at play in the MTO data would seem to be relevant for understanding the consequences for the black-white gap from changes in neighborhood segregation in the U.S. as a whole.

Perhaps the most important qualification to the MTO findings is that the interim evaluation measures outcomes 4-7 years after random assignment, and in principle the effects of neighborhood changes on children could increase over time as they become more socially integrated into their new communities, or as children's exposure to more developmentally productive "inputs" accumulates. The interim MTO evaluation did find positive effects on some family inputs, particularly parent mental health, that could in principle improve student learning over the long term. However even if there are long-term effects of neighborhood mobility on youth outcomes that cannot be detected by MTO, this sort of lag would seem to rule out a very important role for neighborhood changes in the 1990s to contribute to a halt in narrowing of the black-white test score gap – particularly because these neighborhood changes of the

1990s moved in the opposite direction to what most existing theories in this area predict to improve child outcomes.

Evidence of at most a small role for neighborhood socioeconomic characteristics from the MTO study is corroborated by recent quasi-experimental studies of families randomly assigned to public housing units in Toronto (Oreopoulos, 2003) and families displaced by public housing demolition in Chicago (Jacob, 2004). Some studies, notably Aaronson's (1998) study of siblings in the Panel Study of Income Dynamics, and Harding's (2003) propensity score-matching model using the same dataset, do report significant associations between neighborhood characteristics and youth outcomes. Cutler and Glaeser (1997) find significant cross-sectional associations between segregation levels and black-white outcome disparities at the metropolitan level, using various instrumental variables for segregation. But each of these methodologies is subject to criticism, and none can exclude the possibility that schools are the main causal pathway between neighborhood characteristics and outcomes (see Vigdor 2006a for further discussion). The next section focuses more exclusively on arguments that racial disparities in school characteristics, made possible by school segregation, influence the magnitude of the black-white test score gap.

IV. Effects of Schools

The methodological equivalent of conducting an MTO-style experiment to assess the importance of school segregation in perpetuating black-white test score gaps would involve randomly assigning students to schools of varying racial composition, for potentially varying periods of time. To our knowledge, no such experiment has taken place at any point in history. Some efforts have been made to exploit variation in the timing of school desegregation orders in the 1960s and 1970s, as well as unitary status declarations in a later time period, to infer the impact of racial composition on outcomes. These studies rest on the assumption that variation in the timing of such court orders is idiosyncratic from the perspective of the individual student. Guryan (2004) finds evidence suggesting that court-ordered school desegregation plans reduce black dropout rates by 2-3 percentage points, with no detectable effect on

whites. Lutz (2005) finds qualitatively similar effects when he examines the impacts of termination of many of these desegregation plans during the 1990s. These court-ordered desegregation plans may also reduce violent crime victimizations among both white and black youth (Weiner, Lutz, and Ludwig, 2007).

Vigdor (2006b) uses data from the U.S. Census and the National Longitudinal Survey of Youth 1979 cohort to document that the black-white earnings gap was generally larger in the South than in the North for cohorts born before 1950, but displayed no association with region for cohorts born after that date. Cohorts born in the South after 1950 would have witnessed at least some effective school integration prior to their 18th birthday, as the South made its transition from having the nation's most segregated schools to having the most integrated schools. The NLSY evidence shows that the black-white gap among a group born between 1958 and 1965 is significantly narrower for those who resided in the South at age 14, controlling independently for region of birth and region of residence. This group would have reached age 14 between 1972 and 1979, a period of time when school segregation levels were significantly lower in the South than in other regions of the country.

To date, these studies of the impact of broad policy changes provide perhaps the most convincing evidence linking school segregation to changes in outcomes. None of these studies, however, considers the potential impact of segregation on test scores. An ideal study would couple comprehensive test score data with information on the timing and extent of school desegregation orders. To our knowledge, no such study has ever been conducted.

Non-experimental studies of the potential impact of segregation on test score gaps follow one of two methodologies. The first examines longitudinal patterns in test score gaps. If trends in the gap in test scores between blacks and whites who attend the same school are identical to the overall trend, it is difficult to argue that differential assignment to schools explains the overall gap. The second method directly relates individual test scores to measures of school or classroom composition. Studies using this second method are often referred to as "peer effects" studies.

Prominent studies of between- and within-school growth in the achievement gap include Fryer and Levitt (2004) and Hanushek and Rivkin (2006). These studies utilize the same data – the Early

Childhood Longitudinal Survey – but come to conflicting conclusions. Fryer and Levitt report that most of the growth in the black-white test score gap between kindergarten and grade 3 occurs within rather than between schools. Hanushek and Rivkin find fault with Fryer and Levitt's decomposition strategy, and report that "virtually all of the grade-to-grade increases in the overall gap occur between schools." (p.10)

Even accepting the finding that black-white gaps grow because blacks disproportionately attend schools where all students fare worse over time, it is not necessarily true that altering attendance patterns would reduce the gap. Students predisposed to poor growth in test scores for reasons unrelated to school quality might tend to congregate in schools serving a predominantly black population. This weakness with the first non-experimental methodology is addressed to some extent with the second method, which generally introduces at least some controls for student background characteristics when attempting to infer the impact of school or classroom composition on achievement.

The greatest weakness in these studies is the inherently imperfect nature of controlling for student background. Many studies go beyond controlling for student characteristics, adding classroom- or school-level covariates in a hierarchical framework. Some go so far as to use classroom-level variation in peer composition and employ school fixed-effects. While this strategy has the advantage of eliminating omitted variable bias associated with non-random sorting into schools, it has the disadvantage of eliminating several potential causal mechanisms linking racial composition to achievement. Predominantly black schools may find it more difficult to recruit high-quality teachers, for example, but the variation in teacher quality across classrooms within a school is probably smaller and less correlated with race. Classroom racial composition in year *t* may also be poorly correlated with classroom racial composition in prior years. Studies that employ school fixed-effects are thus ill-suited to test dose response hypotheses.

With these caveats in mind, it is worthwhile surveying studies linking school or classroom composition to achievement. Among studies examining school-level variation, which are subject to concerns regarding nonrandom sorting across schools, the Coleman Report of 1966 is perhaps the best-

known. This study found that a school's socioeconomic class composition was among the more important predictors of individual student achievement levels, certainly much more important than school racial composition. Recent studies that use cross-sectional variation in school segregation often generate similar findings – school racial composition seems to be weakly if at all related to student outcomes, while SES of the school's student body may have a stronger association with student achievement (Rivkin, 2000, Cook and Evans 2000, Rumberger and Palardy, 2005).

Card and Rothstein (2006) document a correlation similar to that in Figure 3, substituting SAT scores for NAEP test scores. In a contrast with much of the existing literature, the authors go on to attribute much of the apparent relationship to causal mechanisms other than those involving school quality or school segregation. The pattern of reported results would also be consistent with a model where long-term rather than instantaneous measures of segregation matter more for test score outcomes, and long-term school segregation is more highly correlated with neighborhood segregation than with instantaneous measures of school segregation.

Hanushek, Rivkin and Kain (2004) exploit a rich panel dataset of student-level observations from Texas public schools, which allows them to infer the impact of racial composition by comparing variation across cohorts of students who attend the same school. Since their preferred specifications employ school fixed effects in order to address concerns regarding endogenous sorting into schools, the results are subject to the criticism that they exclude some potential causal mechanisms and render inferences regarding nonlinear dose response difficult. With these caveats in mind, the study finds that a 10% reduction in percent black in school would increase test scores for blacks by around .025 standard deviations (from their Table 1) and increase test scores for whites by around .01 standard deviations (not statistically significant). These effects are about twice as large for black students at the top of the achievement distribution compared to those at the bottom of the distribution. Importantly, their finding holds even after controlling for either the average achievement level or social class composition of other students in the school. These authors control for a variety of potential confounding factors, including student-specific rates of change in achievement test scores and hard-to-measure factors that vary at the

level of the school-by-grade or even attendance zone-by-year. The authors argue that the magnitude of their estimates suggests that equalizing the racial composition of all schools in Texas would reduce achievement gaps in that state by roughly one-quarter. In the end, it is difficult to assess whether the estimates understate or overstate the true impact of segregation on test score gaps.

Hoxby (2000) uses plausibly random variation across cohorts in student demographic composition in Texas and finds that a 10% increase in the share of one's classmates that are black reduces achievement test scores for blacks by around .1 standard deviations in reading and around .06 standard deviations in math. She finds some evidence that the effects of changes in segregation are strongest in schools that are composed of at least one-third black students already. The effect of changes in racial segregation in the school on white students is of the same sign as for black students but only around one-quarter as large in magnitude. As with the Hanushek, Kain and Rivkin study, these effects may understate the impact of segregation because they exclude school-level causal mechanisms, or may overstate them if the restriction to cross-cohort variation in peer group composition fails to eliminate all potential sources of omitted variable bias.

Vigdor and Nechyba (2007) use administrative data on North Carolina public schools to examine the relationship between peer composition and test scores at both the school and classroom levels. In contrast with the Texas studies, their results fail to show any negative association between peer percent black and the relative performance of black students. Relative to the Texas administrative data, the North Carolina data include more detailed family background information and permit the matching of students to classrooms and teachers in elementary grades. The authors also analyze variation in peer composition associated with the opening of new schools in rapidly growing districts, by examining how achievement scores change for students that have a substantial change in peer group composition (without moving schools themselves). The absence of peer effects could be attributed to differences in methodology, including the focus on elementary rather than middle school grades.

Cooley (2006) uses the same North Carolina administrative data to estimate a structural model of the relationship between peer effort and achievement. She allows the impact of peer effort on individual

achievement to vary by race of the individual and the peer. The results are used to forecast the potential impact of integrating two neighboring school districts with varying racial composition, and eliminating within-district segregation in each. In this simulation, the black-white achievement gap is reduced by roughly 3% of a standard deviation.

Overall, the literature on school segregation and the black-white test score gap, while clearly not unanimous in its findings, taken together provides more substantial support than the complementary literature on the impact of neighborhood-level segregation on comparable outcomes. By several estimates, the impact of segregation is modest; extrapolations of even the most supportive studies would imply a black-white gap of sizable magnitude in the absence of any school segregation. There are several unexplored questions within this literature, as well, which seem ripe for further explanation. First, it is plausible that the cumulative effect of spending several years in a segregated environment is greater than the linear extrapolation of short-run effects. Few, if any, studies have examined the impact of past peer racial composition on current outcomes. Second, as discussed in section II above, the most striking change in school racial composition patterns has been the increase in non-black non-white students, and few if any studies have documented whether black students fare particularly poorly in schools with higher shares of Hispanic or Asian students. Finally, the further exploitation of true experiments or quasi-experiments creating exogenous variation in school racial composition also seems like a promising avenue for further research.

How do we reconcile evidence that school racial composition matters for student outcomes with findings from MTO of no detectable impacts on achievement test scores, dropout or other outcomes? As noted above, MTO had relatively modest effects on neighborhood racial segregation, and even more modest impacts on school racial segregation – in part because a fair number of children assigned to the MTO treatment groups remained in their old schools. The result is that the change in proportion school that is minority for MTO participants assigned to the experimental group is about half as large as the experimental treatment effect on proportion census tract that is minority (Sanbonmatsu et al., forthcoming, Table 2), or about 3 or 4 percentage points. The results from Hanushek et al. (2004) and

Hoxby (2000) suggest the result should be an increase in student achievement test scores on the order of .01 to .03 standard deviations, which would be too small to be detected in the MTO data.

V. Conclusion

Can segregation explain the stalled progress toward closing the black-white test score gap since 1990? If so, is there some reason for hope that progress will resume in the near future? As discussed above, there are a number of reasons to think that segregation influences black-white test score gaps. Some incorporate school-specific characteristics as a mediating factor, others do not. The most reliable empirical evidence, derived from the MTO demonstration program and other quasi-experimental studies, indicates that any links between neighborhood and outcomes not mediated by school factors must be small. This conclusion is supported by general evidence on the downward trend in residential segregation, which continued unabated in the 1990s.

The downward trend in residential segregation in the 1990s was not matched by a commensurate downward trend in school segregation. Percent black in the neighborhood occupied by the average black declined five percentage points in the 1990s, while percent black in the public school attended by the average black student declined only three percentage points between 1987 and 2003. The more modest decline in school segregation can be explained largely as the result of declining efforts to integrate public schools. While schools themselves remain more integrated than the communities they serve, school segregation is clearly converging toward the degree of neighborhood segregation.

While the literature on the impact of school segregation suffers from the lack of a school-level analogue to the MTO experiment, a number of studies offer evidence to support some causal component to the basic correlational relationship documented in Figure 3. Studies have documented relative gains to black students associated with discrete changes in school segregation, both at the regional and local levels. The gaps between black and white students who attend the same school tend to be small relative to the overall gap. Some additional studies exploiting idiosyncratic variation in racial composition apparent in administrative public school data have found a significant link with achievement test scores,

in spite of the fact that they frequently employ school fixed effects, which eliminate potentially important causal mechanisms linking racial composition to outcomes.

To be sure, the magnitude of the estimated school racial composition effects coupled with the slight increase in school segregation relative to residential segregation imply that the relationships identified here can explain only a small portion of what occurred in the 1990s. Accepting the largest available point estimate, the additional 2% reduction in percent black experienced by the average black that would have transpired if school segregation exactly tracked residential segregation would have resulted in a narrowing of the black-white test score gap by about 0.02 standard deviations – an effect too small to be detected except in samples of extraordinary size.

It is also clear, however, that these estimates suggest that the future trend of school segregation could have a larger impact on the black-white test score gap. A complete elimination in district efforts to integrate public schools could raise test score gaps significantly in some cases. In other cases, school integration is a moot point at the district level because entire school systems are nearly 100% nonwhite. Should the long-run trend toward lower levels of residential segregation continue in the 21st century, however, we might expect to see renewed progress towards closing the gap. Indeed, one reading of the evidence during the 1990s is that school segregation levels eased upwards in response to court rulings, but that future trends will be dominated by the secular trend towards greater neighborhood integration that is now nearing its fifth decade.

References

Aaronson, D. (1998) "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes." *Journal of Human Resources* 33: 915-946.

Carneiro, Pedro and James J. Heckman (2003) "Human Capital Policy." In *Inequality in America: What Role for Human Capital Policies?* Edited by James J. Heckman and Alan B. Krueger. Cambridge, MA: MIT Press. pp. 77-239.

Card, David and Jesse Rothstein (2006) "Racial Segregation and the Black-White Test Score Gap." NBER Working Paper No. 12078.

Case, Anne and Lawrence Katz (1991) "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths." NBER Working Paper No. 3705.

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

Clotfelter, Charles T., Helen F. Ladd and Jacob L. Vigdor (2005) "Federal Oversight, Local Control, and the Specter of 'Resegregation' in Southern Schools." NBER Working Paper No. 11086.

Coleman, James S. (1988) "Social Capital in the Creation of Human Capital." *American Journal of Sociology*. 94:S95-S120.

Cook, Michael D. and William N. Evans (2000) "Families or Schools? Explaining the Covergence in White and Black Academic Performance." *Journal of Labor Economics* 18(5):729-754.

Cook, Philip J. and Kristen Goss (1996) "A selective review of the social-contagion literature." Working Paper, Sanford Institute of Public Policy, Duke University.

Cooley, Jane (2006) "Desegregation and the Achievement Gap: do Diverse Peers Help?" Unpublished manuscript, University of Wisconsin-Madison.

Crane, Jonathan (1991) "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing." *American Journal of Sociology*. 96(5): 1226-1259.

Cutler, David M. and Edward L. Glaeser (1997) "Are Ghettos Good or Bad?" *Quarterly Journal of Economics* 112: 827-872.

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.

Duncan, Greg, J. Brooks-Gunn, and J. L. Aber, eds. (1997) *Neighborhood Poverty: Context and Consequences for Children*. Volumes I and II. New York: Russell Sage Foundation

Duncan, G.J., Jeanne Brooks-Gunn, and P.K. Klebanov (1994) "Economic deprivation and early childhood development." *Child Development.* 65(2): 296-318.

Echenique, Federico and Roland Fryer (2005) "On the Measurement of Segregation." NBER Working Paper No. 11258.

Ellen, Ingrid Gould and Margery Austin Turner (2003) "Do Neighborhoods Matter and Why?" In *Choosing a Better Life? Evaluating the Moving to Opportunity Social Experiment*. Edited by John Goering and Judith D. Feins. Washington, DC: Urban Institute Press. pp. 313-338.

Feins, Judith D., and Mark D. Shroder (2005) "Moving to Opportunity: The demonstration's design and its effects on mobility." *Urban Studies.* 42(8):1275-1299.

Fryer, R. and S. Levitt (2004) "Understanding the Black-White Test Score Ga in the First Two Years of School" *Review of Economics and Statistics* 86(2):447-464.

Glaeser, Edward L. and Jacob Vigdor (2003) "Racial Segregation: Promising News." In Bruce Katz and Robert E. Lang, eds., *Redefining Urban and Suburban America: Evidence from Census 2000, Volume 1.* Washington: Brookings Institution Press.

Goering, John and Judie Feins, Eds. (2003) *Choosing a Better Life: Evaluating the Moving to Opportunity Social Experiment.* Washington, D.C.: The Urban Institute Press.

Guryan, Jonathan (2004) "Desegregation and Black Dropout Rates." American Economic Review. 94(4): 919-943.

Hanushek, Eric A., and Steven G. Rivkin (2006) "School Quality and the Black-White Achievement Gap." NBER Working Paper #12651.

Hanushek, Eric A., John F. Kain and Steven G. Rivkin (2004) "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement." Working Paper, Hoover Institution, Stanford University.

Harding, D.J. "Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on Dropping Out and Teenage Pregnancy." *American Journal of Sociology* 109: 676-719.

Hoxby, Caroline (2000) "Peer Effects in the Classroom: Learning from Gender and Race Variation." Cambridge, MA: NBER Working Paper 7867.
Jacob, Brian A. (2004) "Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago." *American Economic Review* 94: 233-258.
Jargowsky, Paul (2003) "Stunning Progress, Hidden Problems: The Dramatic Decline of Concentrated Poverty in the 1990s." Washington, DC: Brookings Institution, Center on Urban and Metropolitan Policy.
Jargowsky, Paul (1996) "Take the Money and Run: Economic Segregation in U.S. Metropolitan Areas." American Sociological Review 61():984-998.
Jencks, Christopher, and Susan E. Mayer (1990) "The Social Consequences of Growing Up in a Poor Neighborhood." In *Inner-City Poverty in the United States*. Edited by L.E. Lynn, Jr. and M.G.H. McGeary. Washington, DC: National Academy Press. pp. 111-186.
Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman (2001) "Moving To Opportunity In Boston: Early Results of a Randomized Mobility Experiment." *Quarterly Journal of Economics*. 116: 607-654.
Keels, Micere, Greg J. Duncan, Stefanie Deluca, Ruby Mendenhall and James Rosenbaum (2005) "Fifteen Years Later: Can Residential Mobility Programs Provide a Long-Term Escape from Neighborhood Segregation, Crime and Poverty?" *Demography*. 42(1): 51-73.
Keels, Micere, Jim Rosenbaum, and Greg Duncan (2004) "Second Generation Effects of Chicago's Gautreaux Residential Mobility Program on Crime." Paper presented at October 28-30 Annual APPAM Research Conference on Creating and Using Evidence in Public Policy Analysis and Management. Atlanta, GA.
Kling, Jeffrey R., Jeffrey B. Liebman and Lawrence F. Katz (2005) "Bullets Don't Got No Name: Consequences of Fear in the Ghetto." In *Discovering Successful Pathways in Children's Development: Mixed Methods in the Study of Childhood and Family Life*. Edited by Thomas S. Weisner. Chicago: University of Chicago Press.
Kling, Jeffrey R. and Jeffrey B. Liebman (2004) "Experimental Analysis of Neighborhood Effects on Youth." Princeton University Industrial Relations Working Paper # 483. Revised May 2004. Revision requested by *Econometrica*.
Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz (2006) "Experimental Analysis of Neighborhood Effects." *Econometrica*. Forthcoming.
Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz (2005) "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment." *Quarterly Journal of Economics*. 120(1):87-130.
Knudsen, Eric I., James J. Heckman, Judy Cameron and Jack P. Shonkoff (2006) "Building America's future workforce: Economic, neurobiological and behavioral perspectives on investment in human skill development." *Proceedings of the National Academy of Sciences* (in press).
Leventhal, Tama, and Jeanne Brooks-Gunn (2000) "The Neighborhoods They Live In: The Effects of Neighborhood Residence on Child and Adolescent Outcomes." *Psychological Bulletin*. 126(2):309-337.
Leventhal, Tama and Jeanne Brooks-Gunn (2004) "A Randomized Study of Neighborhood Effects on Low-Income Children's Educational Outcomes." *Developmental Psychology*. 40(4): 488-507.
Logan, John (2004) "Resegregation in American Public Schools? Not in the 1990s." Mumford Center for Comparative Urban and Regional Research, SUNY-Albany.
Ludwig, Jens, Greg J. Duncan, and Paul Hirschfield (2001) "Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-mobility Experiment." *Quarterly Journal of Economics*. 116(2): 655-680.
Ludwig, Jens, Greg J. Duncan and Joshua C. Pinkston (2005) "Housing Mobility Programs and Economic Self-Sufficiency: Evidence from a Randomized Experiment." *Journal of Public Economics*. 89(1): 131-156.
Ludwig, Jens, Helen Ladd, and Greg J. Duncan (2001) "The Effects of Urban Poverty on Educational Outcomes: Evidence from a Randomized Experiment." In *Urban Poverty and Educational Outcomes*. Brookings-Wharton Papers on Urban Affairs. Edited by William Gale and Janet Rothenberg Pack. Washington, DC: Brookings Institution Press. pp.147-201.
Ludwig, Jens and Jeffrey R. Kling (2007) "Is Crime Contagious?" Forthcoming, *Journal of Law and Economics*.
Luttmer, Erzo F.P. (2005) "Neighbors as Negatives: Relative Earnings and Well-Being." *Quarterly Journal of Economics*. 120(3): 963-1002.
Lutz, Byron F. (2005) "Post Brown vs. the Board of Education: The Effects of the End of Court-Ordered Desegregation." Federal Reserve Board, Finance and Economics Discussion Series Working Paper 2005-64.
Manski, Charles F. (1995) *Identification Problems in the Social Sciences*. Cambridge: Harvard University Press.
Manski, Charles F. (2000) "Economic Analysis of Social Interactions." *Journal of Economic Perspectives*. XIV:115-136.

Massey, Douglas, and Nancy Denton (1988) “The Dimensions of Residential Segregation.” Social Forces 67: 281-315.
Mendenhall, Ruby, Greg Duncan and Stefanie DeLuca (2004) "Neighborhood Resources and Economic Mobility: Results from the Gautreaux Program." Unpublished paper.
Moffitt, Robert A. (2001) “Policy Interventions, Low-level Equilibria, and Social Interactions.” In *Social Dynamics*. Edited by Steven Durlauf and Peyton Young. Cambridge, MA: MIT Press.
Neal, Derek A. (2005) “Why has black-white skill convergence stopped?” Cambridge, MA: NBER Working Paper 11090.
Oreopoulos, P. (2003) “The Long-Run Consequences of Living in a Poor Neighborhood.” *Quarterly Journal of Economics* 118: 1533-1575.
Orfield, Gary and Susan Eaton (1996) *Dismantling Desegregation: The Quiet Reversal of Brown v. Board of Education*. New York: The New Press.
Orr, Larry, Judith D. Feins, Robin Jacob, Erik Beecroft, Lisa Sanbonmatsu, Lawrence F. Katz, Jeffrey B. Liebman, and Jeffrey R. Kling (2003) *Moving to Opportunity Interim Impacts Evaluation*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.
Orr, Larry (1999) *Social Experiments*. New York: Russell Sage.
Quigley, John M. (2000) “A Decent Home: Housing Policy in Perspective.” In *Brookings-Wharton Papers on Urban Affairs*. Edited by William G. Gale and Janet Rothenberg Pack. Washington, DC: Brookings Institution Press. pp. 53-100.
Rivkin, Steven G. (2000) “School Desegregation, Academic Attainment and Earnings.” *Journal of Human Resources*. 35(2): 333-346.
Rivkin, Steven G., Eric A. Hanushek and John F. Kain (2005) “Teachers, Schools, and Academic Achievement.” *Econometrica*. 73(2): 417-458.
Rubinowitz, Leonard S. and James E. Rosenbaum (2001) *Crossing the Class and Color Lines: From Public Housing to White Suburbia*. Chicago: University of Chicago Press.
Rumberger, Russell W. and Gregory J. Palardy (2005) “Does segregation still matter? The impact of student composition on academic achievement in high school.” *Teachers College Record*. 107(9): 1999-2045.
Sampson, R. J., S. Raudenbush, and F. Earls (1997) “Neighborhoods and Violent Crime: A Multi-level Study of Collective Efficacy.” *Science.* 277 (August 15):918-924.
Sampson, Robert J., Jeffrey D. Morenoff, and Thomas Gannon-Rowley (2002) “Assessing `Neighborhood Effects’: Social Processes and New Directions in Research.” *Annual Review of Sociology*. 28: 443-478.
Sanbonmatsu, Lisa, Jeffrey R. Kling, Greg J. Duncan and Jeanne Brooks-Gunn (2006) “Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment.” *Journal of Human Resources.* Forthcoming.
Shonkoff, Jack P. and Deborah A. Phillips (2000) *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, DC: National Academy Press.
Shroder, Mark (2002) “Locational Constraint, Housing Counseling, and Successful Lease-Up in a Randomized Housing Voucher Experiment.” *Journal of Urban Economics*. 51(2): 315-338.
Shroder, Mark, and Arthur Reiger (2000) “Vouchers Versus Production Revisited.” *Journal of Housing Research.* 11(1): 91-107.
Vigdor, Jacob L. (2006a) “Peer Effects in Neighborhoods and Housing.” In K. Dodge, T. Dishion and J.Lansford, eds., *Deviant Peer Influences in Programs for Youth: Problems and Solutions*. New York: Guilford Press.
Vigdor, Jacob L. (2006b) “The new promised land: Black-white convergence in the American South, 1960-2000.” Cambridge, MA: NBER Working Paper 12143.
Vigdor, Jacob L. and Thomas Nechyba (2007) “Peer Effects in North Carolina Public Schools” In P.E. Peterson and L. Woessmann, eds., *Schools and the Equal Opportunity Problem*. Cambridge: MIT Press.
Watson, Tara. (2006) “Metropolitan Growth, Inequality, and Neighborhood Segregation by Income.” Williams College manuscript.
Weiner, David A., Byron F. Lutz, and Jens Ludwig (2007) “The Effects of School Desegregation on Crime.” Paper presented at the annual meetings of the American Economic Association, Chicago, IL.
Wilson, William Julius (1987) *The Truly Disadvantaged: The Inner City, The Underclass, and Public Policy.* Chicago: University of Chicago Press.

Table 1
Summary of school-level black/nonblack dissimilarity, US and selected large school districts

Geographic entity	1987/88	2003/04	Change
US, LEAs weighted by enrollment	0.306	0.286	-0.02
	(0.177)	(0.172)	
	(N=7,939)	(N=13,416)	
US, LEAs weighted by black enrollment	0.358	0.357	-0.001
	(0.193)	(0.202)	
	(N=7,939)	(N=13,416)	
Los Angeles Unified	0.603	0.554	-0.049
City of Chicago	0.561	0.698	+0.137
Dade County FL	0.604	0.656	+0.052
Houston ISD	0.599	0.564	-0.035
Philadelphia	0.617	0.618	+0.001
Detroit	0.611	0.734	+0.121
Broward County FL	0.544	0.522	-0.022
Dallas ISD	0.584	0.531	-0.053
San Diego City	0.359	0.298	-0.061
Clark County, NV	0.267	0.260	-0.007
Jefferson County KY	0.157	0.204	+0.047
Milwaukee	0.303	0.596	+0.293
Palm Beach County FL	0.546	0.502	-0.044
Orange County FL	0.462	0.480	+0.018
Pinellas County FL	0.270	0.384	+0.114
Charlotte-Mecklenburg	0.217	0.418	+0.211
Cleveland City	0.277	0.723	+0.446

Note: Data source is the Common Core of Data. Indices are reported for the largest school districts with data on school racial composition available for both school years.

Table 2
Summary of district-level black/nonblack dissimilarity, by state

State	1987/88	2003/04	Change
Arizona	0.438	0.302	-0.136
Arkansas	0.607	0.675	+0.068
California	0.487	0.406	-0.081
Colorado	0.621	0.547	-0.074
Connecticut	0.609	0.572	-0.037
Delaware	0.079	0.219	+0.140
Florida	0.222	0.246	+0.024
Illinois	0.626	0.639	+0.013
Indiana	0.673	0.680	+0.007
Iowa	0.644	0.533	-0.111
Kansas	0.600	0.536	-0.064
Kentucky	0.574	0.576	+0.002
Massachusetts	0.620	0.586	-0.034
Michigan	0.819	0.758	-0.061
Minnesota	0.705	0.571	-0.134
Nebraska	0.728	0.602	-0.126
Nevada	0.292	0.254	-0.048
New Jersey	0.653	0.597	-0.056
North Carolina	0.354	0.335	-0.019
North Dakota	0.672	0.408	-0.264
Ohio	0.717	0.717	---
Oklahoma	0.576	0.551	-0.025
Oregon	0.650	0.470	-0.180
Pennsylvania	0.746	0.687	-0.059
Rhode Island	0.577	0.503	-0.074
South Carolina	0.357	0.353	-0.004
Texas	0.489	0.459	-0.030
Utah	0.508	0.285	-0.223
Washington	0.536	0.465	-0.071
Wisconsin	0.717	0.725	+0.008

Note: Data source is Common Core of Data. Indices are reported for those states reporting racial composition of schools reliably in both years.

Table 3 – Mobility Characteristics of MTO Households[5]

	Controls	Experimental Group		S8-Only Group	
	All	All	Program Movers Only	All	Program Movers Only
# Moves since RA	1.345	1.741	2.290	1.864	2.416
	(.060)	(.050)	(.061)	(.087)	(.099)
Tract Share Poor:					
1 Year Post RA	.454	.328	.194	.344	.286
	(.007)	(.008)	(.009)	(.011)	(.011)
6 Years Post RA	.376	.296	.206	.307	.283
	(.008)	(.007)	(.007)	(.010)	(.010)
Average through 6 years post RA*	.422	.316	.203	.335	.293
	(.006)	(.007)	(.005)	(.008)	(.008)
Tract Share College Ed.:					
1 Year Post RA	.147	.211	.278	.169	.178
	(.006)	(.006)	(.008)	(.007)	(.009)
6 Years Post RA	.156	.197	.236	.176	.176
	(.006)	(.005)	(.007)	(.007)	(.009)
Average through 6 years post RA*	.144	.196	.246	.167	.173
	(.004)	(.004)	(.006)	(.006)	(.007)
Tract Share Minority:					
1 Year Post RA	.889	.800	.676	.875	.841
	(.009)	(.011)	(.017)	(.013)	(.017)
6 Years Post RA	.885	.821	.744	.859	.859
	(.010)	(.010)	(.015)	(.015)	(.015)
Average through 6 years post RA*	.886	.801	.700	.855	.832
	(.008)	(.009)	(.013)	(.012)	(.015)
N	426	701	361	263	172

* = Duration-weighted post-randomization neighborhood averages.

[5] *Note: RA = random assignment. Mean values shown with standard errors in parentheses. Sample consists of adults interviewed during the MTO interim evaluation and randomized by 12/31/1995 (results 1 and 4 years post RA are similar for full sample of MTO adults). Tract characteristics are from the 2000 Census. Number of moves was calculated using the adult's address history. * = Average post-RA neighborhood attributes calculated using address duration as weights.*

Table 4
Impact of MTO on Selected Child and Youth Outcomes, Interim Evaluation[6]

	Females		Males	
	C Mean	ITT	Controls	ITT
Experimental vs Control Group				
Reading test z-scores [6-20]	.103	.060 (.038)	-.096	-.002 (.045)
Graduated or still enrolled [15-20]	.772	.064 (.036)	.759	-.044 (.037)
Used pot last 30 days [15-20]	.131	-.059* (.028)	.118	.053 (.030)
# arrests violent crime [15-25]	.241	-.077* (.031)	.537	-.045 (.051)
# arrests property crime [15-25]	.164	-.057* (.026)	.474	.150* (.055)
Psych distress K6, z-score [15-20]	.268	-.246* (.091)	-.162	.069 (.091)
Has fair or poor health [15-20]	.101	.021 (.027)	.045	.033 (.019)
Section 8 vs Control Group				
Reading test z-scores [6-20]	.103	.012 (.043)	-.096	.033 (.046)
Graduated or still enrolled [15-20]	.772	.049 (.037)	.759	-.040 (.041)
Used pot last 30 days [15-20]	.131	-.052 (.030)	.118	.075* (.035)
# arrests violent crime [15-25]	.241	-.079* (.036)	.537	.024 (.062)
# arrests property crime [15-25]	.164	.031 (.039)	.474	.072 (.059)
Psych distress K6, z-score [15-20]	.268	-.133 (.104)	-.162	-.027 (.096)
Has fair or poor health [15-20]	.101	-.003 (.027)	.045	.033 (.023)

[6] Age range of analytic sample as of December 31, 2001 is reported in parentheses next to each variable. For more details see Kling, Ludwig and Katz (2005), Kling, Liebman and Katz (2006) and Sanbonmatsu et al. (2006).

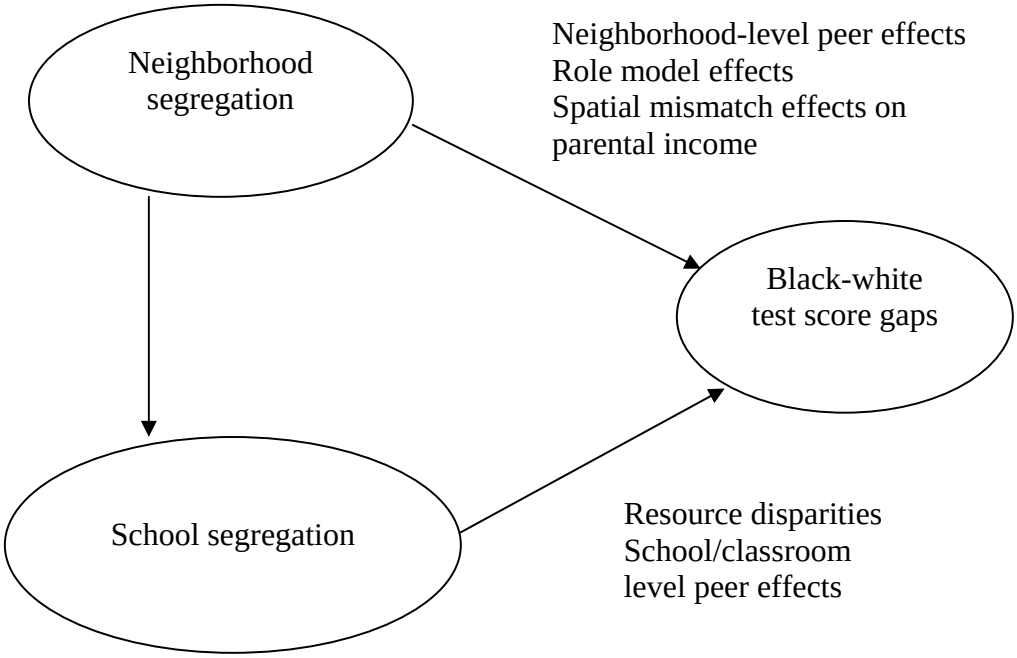


Figure 1: Conceptual Framework

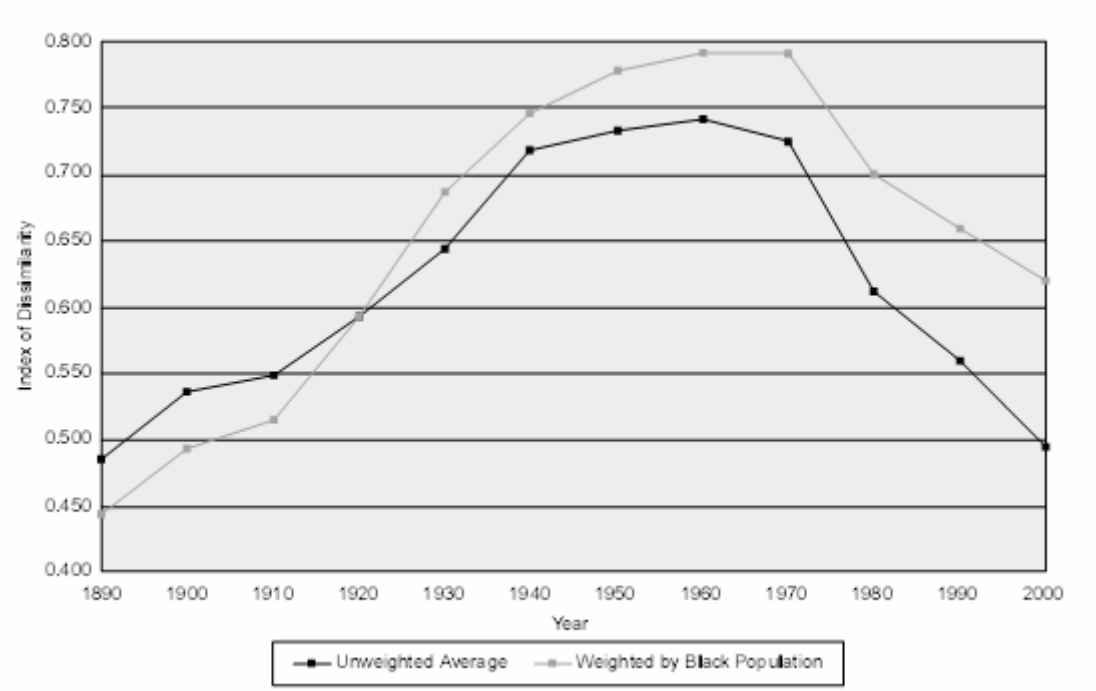
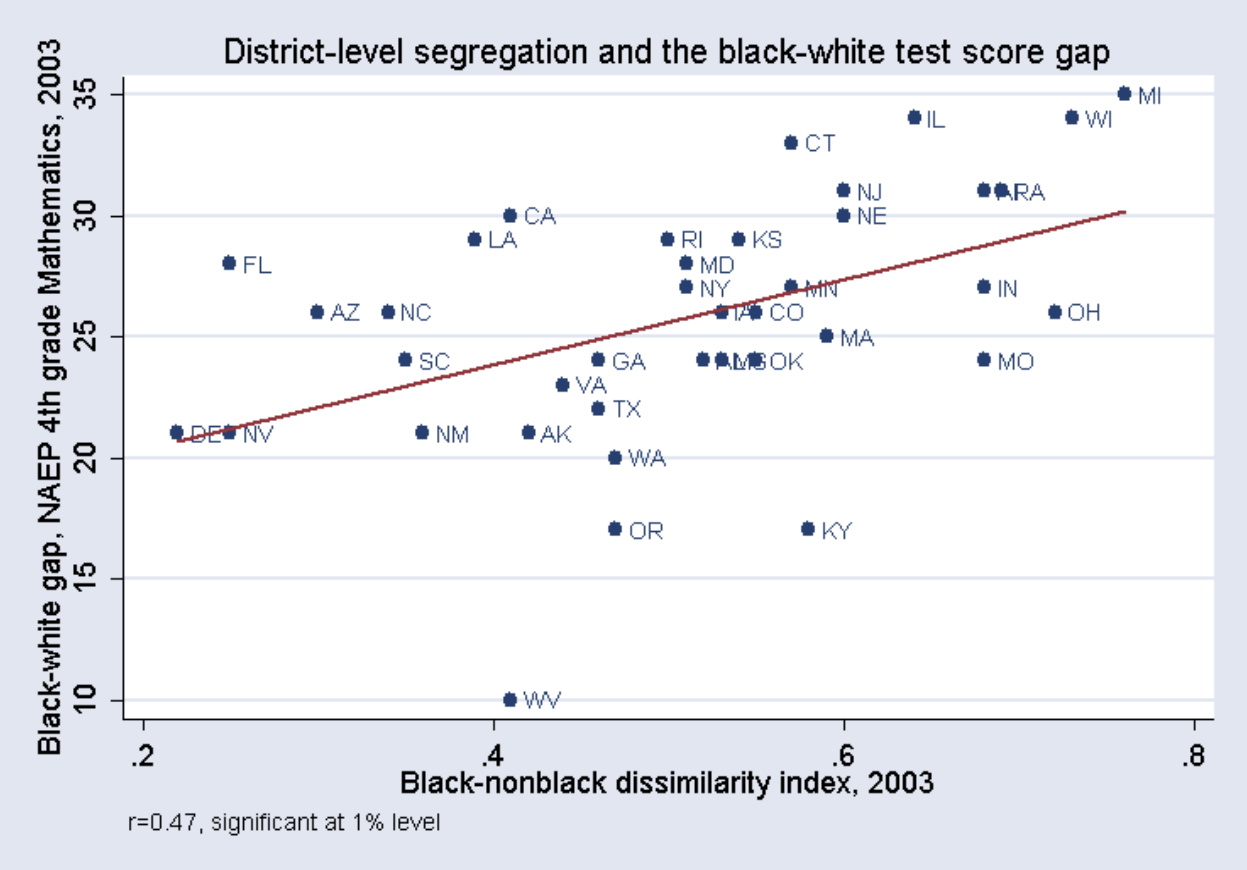


Figure 2: Mean residential dissimilarity for US Metropolitan Areas, 1890-2000. Source: Glaeser and Vigdor (2003).


District-level segregation and the black-white test score gap
Black-white gap, NAEP 4th grade Mathematics, 2003
10
15
20
25
30
35
.2
.4
.6
.8
Black-nonblack dissimilarity index, 2003
r=0.47, significant at 1% level
MI
IL
WI
CT
NJ
NE
CA
LA
RI
KS
MD
NY
MN
IN
FL
AZ
NC
CO
OH
MA
SC
GA
OK
MO
VA
TX
NV
NM
AK
WA
OR
KY
WV
Figure 3

Appendix Table 1

Summary of school-level free lunch eligible/ineligible dissimilarity, US and large school districts

Geographic entity	1987/88	2003/04	Change
US, LEAs weighted by enrollment	0.269	0.279	+0.010
	(0.143)	(0.131)	
	(N=3,198)	(N=12,976)	
US, LEAs weighted by free lunch-eligible enrollment	0.294	0.291	-0.003
	(0.131)	(0.127)	
	(N=3,198)	(N=12,976)	
Dade County FL	0.474	0.418	-0.056
Broward County FL	0.477	0.448	-0.029
Hillsborough County FL	0.352	0.372	+0.020
Duval County FL	0.385	0.381	-0.004
Palm Beach County FL	0.510	0.509	-0.001
Orange County FL	0.409	0.343	-0.066
Pinellas County FL	0.331	0.386	+0.055
District of Columbia	0.522	0.387	-0.135
Orleans Parish	0.411	0.447	+0.036
Charlotte-Mecklenburg	0.297	0.460	+0.163
DeKalb County GA	0.455	0.383	-0.072
Atlanta City	0.443	0.409	-0.034
Cobb County GA	0.523	0.511	-0.012
Polk County FL	0.345	0.333	-0.012
Wake County NC	0.257	0.286	+0.029
Gwinnett County GA	0.369	0.407	+0.038
Jefferson Parish	0.379	0.325	-0.054

Appendix Table 2
Summary of district-level free lunch eligible/ineligible dissimilarity, by state

State	1987/88	2003/04	Change
Alaska	0.310	0.252	-0.058
Delaware	0.101	0.170	+0.069
Florida	0.132	0.176	+0.044
Georgia	0.365	0.298	-0.067
Idaho	0.327	0.195	-0.132
Indiana	0.327	0.326	-0.001
Iowa	0.212	0.263	+0.051
Louisiana	0.276	0.230	-0.046
Minnesota	0.374	0.359	-0.015
Nebraska	0.347	0.288	-0.059
New Hampshire	0.289	0.313	+0.024
North Carolina	0.262	0.191	-0.071
Oregon	0.210	0.203	-0.007
Rhode Island	0.421	0.564	+0.143
Vermont	0.370	0.324	-0.046