Stanford cepa | Center for Education Policy Analysis

CEPA Working Paper No. 16-10

The Geography of Racial/Ethnic Test Score Gaps

AUTHORS

Sean F. Reardon
Stanford University

Demetra Kalogrides
Stanford University

Kenneth Shores
University of Pennsylvania

ABSTRACT

We estimate racial/ethnic achievement gaps in several hundred metropolitan areas and several thousand school districts in the United States using the results of roughly 200 million standardized math and reading tests administered to public school students from 2009-2013. We show that achievement gaps vary substantially, ranging from nearly 0 in some places to larger than 1.2 standard deviations in others. Economic, demographic, segregation and schooling characteristics explain roughly three-quarters of the geographic variation in these gaps. The strongest correlates of achievement gaps are local racial/ethnic differences in parental income, local average parental education levels, and patterns of racial/ethnic segregation, consistent with a theoretical model in which family socioeconomic factors affect educational opportunity partly though residential and school segregation patterns.

VERSION

December 2017

Suggested citation: Reardon, S.F., Kalogrides, D., & Shores, K. (2017). The Geography of Racial/Ethnic Test Score Gaps (CEPA Working Paper No.16-10). Retrieved from Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp16-10

The Geography of Racial/Ethnic Test Score Gaps

Sean F. Reardon

Demetra Kalogrides

Kenneth Shores

Stanford University

Version: December 2017

The research described here was supported by grants from the Institute of Education Sciences (R305D110018), the Spencer Foundation (Award #201500058), and the William T. Grant Foundation (Award #186173). The paper would not have been possible without the assistance of Ross Santy, Michael Hawes, and Marilyn Seastrom, who facilitated access to the ED*Facts* data. This paper benefitted substantially from ongoing collaboration with Andrew Ho, Erin Fahle, and Ben Shear. Some of the data used in this paper were provided by the National Center for Education Statistics (NCES). The opinions expressed here are our own and do not represent views of NCES, the Institute of Education Sciences, the Spencer Foundation, the William T. Grant Foundation, or the U.S. Department of Education. Direct correspondence and comments to Sean F. Reardon, sean.reardon@stanford.edu, 520 CERAS Building #526, Stanford University, Stanford, CA 94305.

The Geography of Racial/Ethnic Test Score Gaps

Abstract

We estimate racial/ethnic achievement gaps in several hundred metropolitan areas and several thousand school districts in the United States using the results of roughly 200 million standardized math and reading tests administered to public school students from 2009-2013. We show that achievement gaps vary substantially, ranging from nearly 0 in some places to larger than 1.2 standard deviations in others. Economic, demographic, segregation and schooling characteristics explain roughly three-quarters of the geographic variation in these gaps. The strongest correlates of achievement gaps are local racial/ethnic differences in parental income, local average parental education levels, and patterns of racial/ethnic segregation, consistent with a theoretical model in which family socioeconomic factors affect educational opportunity partly though residential and school segregation patterns.

The Geography of Racial/Ethnic Test Score Gaps

Introduction

Racial and ethnic disparities in children's academic performance are a stubborn feature of the US educational landscape, an indicator of continued racial inequality of educational opportunity. Though these achievement gaps are substantially smaller than they were 40 years ago, they remain quite large, on the order of two-thirds to three-quarters of a standard deviation (Neal 2006; Reardon, Robinson-Cimpian and Weathers 2015). They are large when children enter kindergarten and remain large through high school (Fryer and Levitt 2004; Hemphill, Vanneman and Rahman 2011; Phillips, Crouse and Ralph 1998; Reardon and Galindo 2009; Vanneman et al. 2009).

The size and trends of these gaps vary among states (Hemphill, Vanneman and Rahman 2011; Reardon 2015; Vanneman et al. 2009), though in no state are they near zero. National- and state-level patterns, however, may mask considerable variation in academic achievement patterns at smaller geographic scales. Metropolitan areas and counties, for example, vary widely in demographic composition, patterns of racial socioeconomic inequality and racial segregation, and in the structure of their schooling systems. The roughly 12,000 school districts in the U.S. likewise differ substantially in their demographics, patterns of inequality and segregation, and educational resources. They also each have autonomy over some—but not all—important features of the schooling system, including their curricula, their student and teacher assignment policies, and how resources are distributed among and within schools. These demographic and institutional factors may lead to significant variation in the size of achievement gaps among both metropolitan areas and school districts.

In this paper, we provide a detailed descriptive analysis of the patterns of white-black and white-Hispanic academic achievement gaps across US metropolitan areas and school districts. We use new data to estimate achievement gaps in almost every metropolitan area and school district in the US with a significant population of black or Hispanic students. The precision and detail of these estimates—which are based on the results of roughly 200 million standardized math and reading tests administered to elementary and middle school students from 2009-2013—far surpasses that of any previously available data. Using these estimates, we first describe the geographic patterns of racial/ethnic achievement gaps among metropolitan areas and school districts in the US. We then examine the extent to which these gaps are correlated—in both bivariate and multivariate models—with socioeconomic characteristics of the white, black, and Hispanic populations, with patterns of residential and school segregation, and with local features of the educational system.

We find that racial/ethnic achievement gaps within school districts and metropolitan areas average roughly 0.5 to 0.7 standard deviations. There is substantial geographic variation in the magnitude of achievement gaps, ranging from nearly 0 in some places to larger than 1.2 standard deviations in others. A vector of economic, demographic, segregation and policy variables explain between 64 and 82 percent of the geographic variation in these gaps. The strongest correlates of achievement gaps are racial/ethnic differences in parental education, racial/ethnic segregation, the overall level of parental education. After adjusting for variation among places in racial socioeconomic inequality and segregation, many school districts and metropolitan areas have larger or smaller achievement gaps than predicted, suggesting that other forces are at work as well.

The purpose of this paper is not to estimate the causal effect of any one particular feature of children's environments or of schools on achievement gaps, but rather to provide descriptive analyses that may help to build intuition and generate hypotheses regarding the causes of the observed achievement gaps. Think of this as a necessary, but not sufficient, analysis for understanding the causes of academic achievement gaps—providing a detailed description of the "stylized facts" regarding racial/ethnic achievement gaps.

Finally, although the ostensible focus of these descriptive analyses is achievement gaps, it is perhaps more appropriate to think of an observed achievement gap as a proxy measure of local racial inequalities in educational opportunity. That is, unless one posits large innate racial differences in academic potential (a poistion supported by no credible theory or evidence; for a review, see Nisbett et al. 2012), differences in average test scores must be understood to represent local racial differences in the average availability of opportunities to learn the tested material. These differences in opportunity may be present in early childhood, in children's home environments and experiences, in their neighborhoods, their child care and preschool centers, and/or their elementary and middle schools. The sociological (and social policy) questions of interest, then, are questions about the relative contribution of these various contexts and experiences in shaping racial disparities in educational opportunity and academic outcomes. With this proviso, for consistency with prior literature, we will refer to these differences in opportunity as achievement or test score gaps.

The Geographic Scale of Racial/Ethnic Test Score Gaps

The best evidence on racial/ethnic achievement gaps in the US comes from the National Assessment of Educational Progress (NAEP), a set of reading and math assessments that have been administered to large, nationally-representative samples of students since 1971. Since 1990, NAEP assessments have also been administered to state-representative samples of students. The NAEP assessments indicate that the white-black achievement reading and math gaps were both over one standard deviation in the 1970s; by 2012, those gaps had shrunk to roughly 0.60 and 0.80 standard deviations, respectively. The white-Hispanic gaps in reading and math have historically been slightly smaller than white-black gaps, but have followed a similar trend over the last 4 decades, and now are roughly 0.50 and 0.60 standard deviations in reading and math, respectively (National Center for Education Statistics 2013; Reardon 2015; Reardon, Robinson-Cimpian and Weathers 2015).

At the state level, most states' white-black achievement gaps in the last decade are between 0.75

and 1.10 standard deviations, though in states with small black populations, the gaps are generally smaller, in some cases less than 0.50 standard deviations. State white-Hispanic gaps generally range from 0.50 to 1.0 standard deviations in this same period. On average, state achievement gaps have narrowed slightly in the last two decades, though this varies among states (Hemphill, Vanneman and Rahman 2011; Reardon 2015; Vanneman et al. 2009). Reardon (2015) shows that state-level achievement gaps are correlated with state racial socioeconomic disparities: achievement gaps are largest, on average, in states with large racial differences in family income, poverty rates, educational attainment, and unemployment rates.

Evidence about the national and state-level patterns and trends of achievement gaps are useful as descriptors of overall patterns and trends of inequality in educational outcomes in the U.S., but they reveal little about local patterns of racial inequality. Thus, these aggregate data are, by themselves, relatively uninformative regarding the processes that produce and sustain achievement gaps. Large national and state-level achievement gaps do not necessarily imply that gaps are large in most school districts. If most black and Hispanic students are in school districts where all students—white students included—perform poorly on standardized tests, and most white students are in school districts where all students—including black and Hispanic students—perform well, then most students would encounter little racial achievement inequality in their own district, even while state and national achievement gaps are large.

Such patterns would suggest that the forces producing achievement gaps do not operate primarily *within* schools and districts, but *between* school districts. The primary candidates for between-district mechanisms are residential segregation and inequality among school districts in resources and quality. Conversely, if gaps are large within individual school districts, between-district forces cannot fully account for achievement gaps; instead, within-district factors—such as racial socioeconomic inequality, between-school segregation, and the unequal distribution of resources and opportunities to learn within schools—are more likely suspects. In the absence of achievement gap data for small geographical units, such as school districts and metropolitan areas, it is not possible to identify those areas where achievement gaps are largest, much less arbitrate these competing sociological theories.

To date, we have little systematic information about achievement gaps in districts or metropolitan areas. The NAEP Trial Urban District Assessment (TUDA) provides district-level estimates of achievement gaps, but only for 21 large school districts.[1] And although many states report race-specific proficiency rates by school district, gaps in these proficiency rates yield distorted measures of achievement gaps (Ho 2008; Ho, Lewis and Farris 2009) that cannot be meaningfully compared across states, grades, subjects, and years, because the tests and proficiency standards used differ widely. As a result, we know relatively little about how large and variable racial/ethnic achievement gaps are among school districts and what local factors contribute to racial/ethnic disparities in educational outcomes.

In this paper, we remedy this need. We compute and describe white-black and white-Hispanic achievement gaps for several thousand school districts, comprising 92 - 93% of the black and Hispanic populations attending public schools in grades three through eight for years 2009 to 2013. We also compute and describe achievement gaps in almost every metropolitan area in the US. We focus on both school districts and metropolitan areas for complementary reasons. School districts are a key organizational unit of the U.S. public school system. They have a large—though not complete—degree of autonomy over curricula, instruction, student assignment, teacher hiring, and the distribution of resources among schools. Thus, there is reason to think that school districts may vary substantially in practices that affect between- and within-school disparities in educational opportunities. Districts are also organizational units with clear geographic boundaries and relatively well-known "brands" (based on easily observable features, such as average test scores and student body composition), which means that

[1] Information about the Trial Urban District Assessment can be found here: https://nces.ed.gov/nationsreportcard/about/district.aspx.

families with sufficient resources can choose to live in the most desirable districts. This leads to relatively high levels of socioeconomic variation among school districts (Owens 2016; Owens, Reardon and Jencks 2016). Finally, districts vary enormously in size; in large districts, there is far more possibility of between-school segregation and between-school differences in school quality and opportunities. These factors suggest that the conditions that lead to achievement gaps may differ markedly among school districts.

A focus on school districts may obscure patterns of inequality evident at larger geographic scales, however. Because housing prices differ markedly among school districts, socioeconomic and racial differences among districts may be larger than those within districts. For example, roughly two-thirds of all racial/ethnic school segregation is due to between-district patterns of segregation (Reardon, Yun and Eitle 2000; Stroub and Richards 2013). Racial socioeconomic disparities are likewise smaller within school districts than in the population at large. We examine achievement gaps within metropolitan areas in order to account for these between-district sorting processes. Although metropolitan areas are not part of the formal organizational structure of the public schooling system, they encompass much of the relevant ecosystem for studying residential and school segregation, in part because most of the residential segregation relevant to inequality occurs within metropolitan areas, not between them (Cutler and Glaeser 1997). As a result, a great deal of social science research treats metropolitan areas as a key geographic unit for studying the patterns and consequences of segregation (see, for example, Card and Rothstein 2007; Logan, Oakley and Stowell 2008; Owens 2015; Reardon, Yun and Eitle 2000). We follow this tradition in including metropolitan area achievement gaps in our analyses.

Causes and Correlates of Racial/Ethnic Test Score Gaps

One of the central sets of questions in the sociology of education for the last 50 years—since the publication of the Coleman Report (Coleman, et al., 1966)—concerns the primary causes of racial and ethnic achievement gaps and disparities in educational outcomes more generally. To what extent are these disparities the result of racial/ethnic differences in socioeconomic family background and circumstances, and to what extent are they the result of racial/ethnic differences in school quality? Put differently, to what extent should racial/ethnic disparities in educational outcomes be attributed to institutional features of the US educational system—features that may be malleable through changes in organizational, institutional, and policy features of schooling—and to what extent should they be attributed to factors outside the school system's control, such as racial/ethnic disparities in socioeconomic family and neighborhood conditions?

Framed this way—as if inequalities inside the school system are distinct from inequalities outside of schools—the question implies a false dichotomy. Differences in socioeconomic conditions are not fully separable from disparities in educational conditions. Socioeconomic inequality may lead to inequality between and within schools, as communities with greater resources are able to better fund their local schools (in taxes and other ways). Parents in such communities may also use their greater social capital to secure better educational opportunities (better teachers, smaller classrooms, for example) for their children than less advantaged children within the same schools. Moreover, school systems react to social inequalities in ways that may reduce or exacerbate these inequalities. In most states, for example, when federal, state and local revenues are added up, per pupil expenditures are greater, on average, in districts enrolling large proportions of low-income students than in districts enrolling few poor students. This may attenuate differences in out-of-school opportunities (Aud et al. 2010; National Center for Education Statistics 2012). Conversely, school systems may also reinforce social inequalities by segregating children from low-income families into less demanding academic programs and/or into high poverty schools or by providing fewer resources to the classrooms and schools that enroll low-income students.

[Figure 1 here]

Figure 1 provides a stylized sociological model of the complex relationships between schooling and non-schooling factors that might affect achievement gaps. On the left of the figure are two primary categories of distal influences on achievement gaps. First are racial family socioeconomic disparities (i.e., racial differences in family income, parental education, and other forms of social and economic resources). These disparities are quite large in the U.S. For example, the median incomes of black and Hispanic families are 38% and 36% lower, respectively, than that of white, non-Hispanic families;[2] median black and Hispanic household wealth are less than 10% as large as median white wealth (Sullivan et al. 2015; Wolff 2014); and only 22% of black adults and 15% of Hispanic adults hold a bachelor's degree, compared to 36% of white adults.[3] Second are education policies and structures (such as school finance policies, student assignment policies, and the like); we discuss these at more length below. Both of these factors may lead to academic achievement gaps through multiple pathways.

On the center right of the figure are four categories of potential proximal sources of academic achievement gaps: racial differences in children's home environments; racial differences in children's neighborhood contexts (distinct from home and school environments); between-school racial differences in schooling experiences and opportunities; and within-school racial differences in schooling experiences and opportunities. Each of these might encompass many potential mechanisms.

First, racial differences in children's home environments include differences in opportunities for learning at home—differences in the amount of time parents have to read to their children; in children's access to computers, libraries, and museums; in parental investments in tutoring and other educational activities; in parental human and social capital; and differences in parental stress and depression. All of these experiences are affected by family socioeconomic status; high-income and highly-educated parents have, on average, more resources to foster and support their children's academic skills outside of school (Bassok et al. 2016; Bradley et al. 2001; Chin and Phillips 2004; Lareau 2003; Phillips 2011). To the extent

[2] http://www2.census.gov/library/publications/2011/compendia/statab/131ed/tables/12s0697.pdf; retrieved January 7, 2017.
[3] See Table 3 at http://www.census.gov/hhes/socdemo/education/data/cps/2014/tables.html; retrieved September 3, 2015.

that these affect students' academic achievement, it follows that racial differences in socioeconomic status would then lead to racial differences in academic achievement, net of other factors. Recent studies indicate this is the case; income affects children's academic achievement (Dahl and Lochner 2012; Duncan, Morris and Rodrigues 2011), though the exact pathways through which these effects operate are not clear. Moreover, racial differences in family socioeconomic conditions explain a large portion of racial achievement gaps present when children enter kindergarten (Fryer and Levitt 2004; Fryer and Levitt 2006; Reardon and Galindo 2009; Rothstein and Wozny 2013). There is less clarity about whether racial achievement gaps grow in ways unrelated to socioeconomic background differences as children progress through school.

Second, racial differences in family economic circumstances affect residential segregation patterns (though housing discrimination and racial preferences shape segregation patterns as well; for a review, see Lareau and Goyette 2014). This means that black and Hispanic children live, on average, in poorer neighborhoods than white children. In fact, black and Hispanic children live in much poorer neighborhoods, relative to white children, than would be expected based on their family income (Logan 2011; Pattillo 2013; Reardon, Fox and Townsend 2015; Sharkey 2014). Poorer neighborhoods typically have higher violent crime rates (Sampson, Raudenbush and Earls 1997) and weaker non-school social institutions (such as availability of high-quality child care and pre-school programs; safe parks and playgrounds; and constructive after-school activities, such as clubs and sports teams (Small 2006). These and other factors have long been hypothesized to affect schooling outcomes (Jencks and Mayer 1990; Leventhal and Brooks-Gunn 2000; Sampson 1998), and new evidence from the MTO experiment and other studies confirms that neighborhood conditions affect educational attainment (Burdick-Will et al. 2011; Chetty, Hendren and Katz 2015; Sampson, Sharkey and Raudenbush 2008; Sharkey 2010; Wodtke, Harding and Elwert 2011). This implies that residential segregation patterns may lead to disparities in educational outcomes (see, for example, Ananat 2009; Card and Rothstein 2007; Cutler and Glaeser

1997).

While the top two boxes at the right of Figure 1 describe potential out-of-school influences on racial achievement gaps, the bottom two describe potential school-related influences. These are divided into within- and between-school factors. The key to both is that achievement gaps may be caused, in part, by racial differences in school experiences and opportunities. These differences in experiences and opportunities may result from students attending different schools (between-school segregation) or they may occur even among students attending the same school. Between-school segregation is a necessary (though not sufficient) condition for between-school differences in educational experiences and opportunities to contribute to achievement gaps; if black, Hispanic, and white students are equally represented in each school, then each group will experience the same average level of (and the same variation in) school quality. In the presence of segregation, however, if school racial composition is correlated with school resources (including the ability to attract and retain skilled teachers; teacher/student ratios; the quality of instructional materials, equipment, and facilities; the availability of support staff; and less tangible factors like school climate), then black and Hispanic students will, on average, experience fewer opportunities for learning than their white peers. Although the effects of school segregation are difficult to estimate, the best available research suggests that school segregation tends to widen racial educational disparities in achievement and educational attainment, as well as adult income (Ashenfelter, Collins and Yoon 2005; Card and Rothstein 2007; Guryan 2004; Johnson 2011; Reardon 2016).

Historically, fewer resources were available to school districts serving large proportions of black, Hispanic, and poor children compared to those serving predominantly white and middle-class students; however, this pattern has been eliminated or reversed in many states. As a result of state school financing reforms enacted by state legislatures or ordered by courts, per-pupil revenues are now modestly *positively* correlated with districts' enrollment rates of poor and minority students within most states

(Cornman 2015). This means that in most states—conventional wisdom notwithstanding—poor and minority students are enrolled in districts with higher per-pupil spending than white and middle-class students, although there are notable exceptions. Cost-adjustments can affect this inference, as high poverty school districts have greater costs than low poverty districts (Bifulco 2005). Given recent evidence indicating that school spending positively affects student achievement, high school and college completion, and adult earnings (Candelaria and Shores 2017; Jackson, Johnson and Persico 2016; Lafortune, Rothstein and Schanzenbach 2016), this suggests that school policies affecting the distribution of resources among school districts may have important effects on achievement gaps.

Despite the fact that, in some states, school districts serving predominantly poor students spend more per pupil than those serving higher-income students, low-income and non-white students are, on average, more likely to have inexperienced teachers and greater teacher turnover (Clotfelter, Ladd and Vigdor 2005; Lankford, Loeb and Wycoff 2002; Scafidi, Sjoquist and Stinebrickner 2007), some of which may be due to the fact that high poverty districts must pay teachers more to attract them (Clotfelter et al. 2008). Higher salaries are thought to be necessary because teachers value working conditions that tend to be correlated with the demographic composition of schools such as safety, proximity of the school to their place or residence, leadership stability, availability of support staff (Boyd et al. 2011; Boyd et al. 2005a; Boyd et al. 2005b).

The reason that the influence of out-of-school family socioeconomic disparities cannot be cleanly distinguished from the role of schooling policies and practices in producing achievement gaps is that school segregation is shaped by both. Segregation is also shaped by a number of other forces, including housing policy, housing discrimination and preferences, private school enrollment patterns, and governmental policy (Lareau and Goyette 2014; Rothstein 2017). Moreover, the extent to which school segregation is linked to between-school racial disparities is dependent on educational policies and practices. If education policy were successful at achieving the “separate but equal” standard articulated in

Plessy v. Ferguson, school segregation would not be linked to between-school differences in the quality of educational experiences. While there is no evidence that this has ever been, or is likely to be, achieved, education policy may nonetheless moderate the relationship between segregation and unequal school quality. Policies that provide extra resources to schools serving large proportions of poor and minority students, for example, may weaken the link between school racial and socioeconomic composition and school quality. The effect of such policies is signified by the dashed line in Figure 1.

The processes sketched in Figure 1 suggest that the factors that produce academic achievement disparities cannot be neatly separated into inequalities in family socioeconomic background and inequalities in schooling experiences. Rather there are three sets of forces at work—1) differences in children's home and neighborhood environments that are due to family socioeconomic resources; 2) differences in children's schooling and neighborhood experiences that are due to education and (past and present) social policy, rather than family socioeconomic differences; and 3) differences in children's schooling experiences that are jointly produced by racial disparities in family resources—which lead to school segregation—and by educational policies and practices which more or less tightly link school segregation to patterns of unequal school quality.

Figure 1 highlights how differences between racial/ethnic groups along a variety of dimensions contribute to achievement gaps. However, these are not necessarily the only factors at play. For example, within-school racial differences in experiences and opportunities to learn may also play a role. Within any school, teachers' skills vary, as do the curricula, instructional practices, and peer composition of different classrooms. If these differences are patterned by student race—because of tracking, differences in teacher expectations, differences in parents' effectiveness at advocating for their children, or other reasons—then these within-school racial differences in educational opportunities and experiences may lead to achievement disparities.

In addition, although Figure 1 highlights racial disparities in socioeconomic, neighborhood, and school conditions as contributors to racial achievement gaps, achievement gaps may also co-vary with average socioeconomic conditions. Some existing scholarship notes that the white-black achievement gap is often large even in relatively affluent, racially diverse communities (Lewis and Diamond 2015; Ogbu 2003), and suggests that this may be due to processes within schools that provide more opportunities to white students than minority students, even in contexts of relative affluence (Tyson 2011). It is not clear, however, whether the achievement gaps in such communities are larger or smaller than in poorer communities with similar levels of racial socioeconomic disparity and segregation. For example, racial socioeconomic disparities and patterns of segregation may affect educational opportunities less in contexts of relative advantage than do comparable disparities in disadvantaged communities. Resource and context differences may be more salient when there are few resources to go around. On the other hand, given the sometimes-competitive focus on academic success in affluent communities, racial socioeconomic disparities may be particularly salient, as economic and social capital may matter more in such contexts. Hanushek and Rivkin (2009), for example, show that school segregation appears most harmful to high-achieving minority students, possibly because differences in access to the best schools particularly limits high-achieving students' educational opportunities. This is consistent with a substantial body of ethnographic and social psychological work that illustrates how subtle structural and exclusionary processes may limit minority students' opportunities and advantage white students, even in (or perhaps particularly in) schools enrolling largely middle- and high-income white and minority students (Carter 2012; Lewis and Diamond 2015; Ogbu 2003).

Finally, it is worth noting that Figure 1 describes a set of uni-directional relationships between family and educational effects on achievement gaps. Racial family socioeconomic disparities and education policies affect achievement gaps, in this stylized model, but not vice-versa. However, over a longer time period, there are certainly processes that work in the other direction as well. First, cross-district stratification is both a cause and consequence of the choices individuals make within a context of neighborhood segregation and inequality (Sampson 2008). Second, racial achievement gaps in one generation shape racial disparities in the next generation's parental educational attainment and family income (Carneiro, Heckman and Masterov 2003; Neal and Johnson 1996) and public policy reactions (such as school desegregation or changes in school funding policies) to persistent racial achievement disparities may shape children's schooling environments. A full model of the dynamic associations among socioeconomic inequality, schooling conditions, and achievement gaps would take these feedback processes into account, but that is beyond the scope of our analyses here.

The conceptual model illustrated in Figure 1 suggests that racial achievement gaps are dependent partly on local racial socioeconomic conditions and disparities, segregation patterns, and school policies, practices, and conditions. To the extent that these factors vary geographically, our model predicts corresponding variation in achievement gaps. Drawing on this model, we address three research questions. First, how large and varied are racial/ethnic achievement gaps among districts and metropolitan areas in the United States? In what districts and metropolitan areas are racial/ethnic achievement gaps largest and smallest? Second, what are the correlates of these achievement gaps? Are achievement gaps larger, as the model suggests, in areas where segregation, racial/ethnic socioeconomic differences, and racial/ethnic school resource differences are greater? Third and finally, how much of the variation in racial/ethnic achievement gaps can be attributed to characteristics of school districts and metropolitan areas? Which local conditions account for the most variation in achievement gaps?

Racial/Ethnic Test Score Gaps among School Districts and Metropolitan Areas

Data Sources and Estimation

To measure achievement gaps, we use data from the federal ED*Facts* data collection system, which were provided to us by the National Center for Education Statistics under a restricted data use license. The data include, for each public school in the United States, counts of students scoring in each of several academic proficiency levels (often labeled something like "Below Basic," "Basic," "Proficient," and "Advanced"). These counts are disaggregated by race (we use counts of non-Hispanic white, non-Hispanic Black, and Hispanic students in this paper), grade (grades 3-8), test subject (math and ELA), and year (school years 2008-09 through 2012-12). We aggregate the school-level counts to the district level. We combine the proficiency counts of charter schools with those of the public school district in which they are formally chartered or, if not chartered by a district, in the district in which they are physically located. Thus, a "school district" is conceptualized as a geographic catchment area that includes students in all local charter schools as well as in traditional public schools. Virtual schools—online schools that do not enroll students from any well-defined geographic area—are dropped from the sample.[4]

For metropolitan areas, we aggregate data from all public schools and charter schools within a given metropolitan area. Because districts in different states use different achievement tests, proficiency categories in different states are not comparable, so we cannot construct aggregated data for metropolitan areas that cross state boundaries. Instead, for the 45 (out of 384) metropolitan areas that cross state lines, we include only the portion of the metropolitan area that is in the state containing the largest number of the metropolitan area's student population.

We measure achievement gaps using the V-statistic (Ho 2009; Ho and Haertel 2006), which measures the non-overlap of two distributions (see Technical Appendix A1 and A2 for details).[5] The V-statistic has three useful properties for our purposes. First, it is readily interpretable as an effect size

[4] Virtual enrollments constitute less than one-half of one percent of all public school enrollments (based on our calculations using 2014-15 data from the Common Core of Data). Therefore, this restriction only excludes a small fraction of students.

[5] In Appendix A1, we review different statistical measures of achievement gaps and provide details about the measurement of V. Appendix A2 (with results in Table A1) includes a set of validation exercises comparing state-level achievement gaps estimated from state accountability tests (which differ among states) to achievement gaps estimated for those same states, grades, and years from the National Assessment of Educational Progress (NAEP), which is identical across states, within a grade, year, and subject. The correlation between the gap estimates from the two data sources is above 0.90 in most cases (ranging between 0.85 and 0.97 depending on year, grade and subject), indicating that different state tests order students similarly enough that the V-statistic can be used to compare achievement gaps across a wide range of state and NAEP tests.

(similar to the standardized mean difference in achievement between groups of students). Second, V is invariant to monotonic transformations of test scales: if a test metric is transformed by any non-linear monotonic transformation, V will be unchanged (unlike gaps computed as between-group differences in average standardized test scores). Third, the V-statistic can be estimated very reliably from aggregated coarsened test score data—counts of students of each group in each of several (at least three) proficiency categories (Ho and Reardon 2012; Reardon and Ho 2015). Because data from ED*Facts* come in this form, it is possible to easily estimate achievement gaps based on state accountability tests in each district/metropolitan area-year-grade-subject for which subgroup-specific proficiency category counts are available. Finally, note that gaps must be understood as relative comparisons of achievement among racial/ethnic groups. While it is important to conceptually distinguish between districts with small racial gaps at low levels of achievement versus districts with small racial gaps at high levels of achievement, the estimates we use here do not permit us to make this distinction empirically, given the differences in tests and definitions of proficiency across states, grades, and years.

We estimate white-black achievement gaps in at least one grade-year for 2,878 and 2,854 districts in ELA and math, respectively, and for 378 metropolitan areas. These 2,878 and 2,854 districts include 93% of black and 92-93% of Hispanic public school students in grades 3-8 in the US. The metropolitan area analytic sample includes 96% of black and 98% of Hispanic public school students in grades 3-8 that attend public schools in metropolitan areas; the remaining black and Hispanic students are in the excluded portions of one of the 45 metropolitan areas that cross state lines. White-Hispanic achievement gaps are available in at least one grade-year for 3,632 and 3,642 districts, and 377 metropolitan areas, in ELA and math respectively. The number of districts for which we can estimate achievement gaps in at least one subject-year-grade is 2,899 (for white-black gaps) or 3,689 (for white-Hispanic gaps); for metropolitan areas, the corresponding numbers are 378 and 377, respectively. Table 1 provides information about availability of these achievement gap data.

[Table 1 here]

Though there are roughly 12,000 school districts serving grades 3-8 in the United States, we can estimate achievement gaps for only approximately one-quarter of these districts, because most school districts have either too few white or too few black or Hispanic students to make estimation of achievement gaps possible. We compute both the white-black and white-Hispanic achievement gap for each cell in which there are at least 20 white and 20 black/Hispanic students. We cannot compute achievement gaps for districts and metropolitan areas in several states in particular years or grades because of insufficient data in the Ed*Facts* system or because states did not use a common test across all districts.[6] These inclusion criteria result in the analytic sample described above.

Given the 5 years, 6 grades, and 2 subjects for which achievement data are available, up to 60 gaps can be estimated for a district or metropolitan area. On average, in the districts and metropolitan areas in our sample (those for which we can estimate the achievement gap in at least one subject-grade-year), the average number of cells with estimated gaps available is 54 (for white-black and -Hispanic gaps in metropolitan areas) and 44 and 40 (for white-black and -Hispanic gaps in districts). In total, at the district level, there are 125,380 estimated white-black achievement gaps (in ELA or math) and 146,494 white-Hispanic achievement gaps; at the metropolitan area level, there are 21,507 white-black and 20,520 white-Hispanic achievement gap estimates.

The Magnitude and Variation of Racial/Ethnic Test Score Gaps

As stated, for both white-black and white-Hispanic gaps, we have up to 60 estimated achievement gaps (two subjects, six grades, and five school years) in each geographic unit (school district

[6] We cannot compute achievement gaps in Florida, Colorado, and Wyoming in several years because those states did not report sufficient data to the ED*Facts* system in all years (i.e., in some years they reported test scores in only two proficiency categories, which are too few to estimate the achievement gap statistic. See details about estimation below and in Technical Appendix A1 and A2.). We also do not compute achievement gaps for California and Virginia in 7th and 8th grade math in any school year or for Nebraska in all grades in ELA in the 2008-09 school year and in all grades in math in the 2008-09 and 2009-10 school years because districts administered different tests to students during these subjects, grades, and years.

or metropolitan area). For parsimony, we prefer to pool these estimates to create a single average gap measure for each unit. We pool the estimates using a multilevel regression model. But first we estimate models to examine whether there is significant variation in gaps between subjects, grades, and years to understand whether the pooling obscures any meaningful variation. We fit a model in the following form:

$$
\begin{aligned}
\hat{G}_{usgy} &= \left[\gamma_{m0} + \gamma_{m1}\left(g_g - 5.5\right) + \gamma_{m2}\left(c_y - 2005.5\right) + v_{mu}\right]M_s \\
&\quad + \left[\gamma_{e0} + \gamma_{e1}\left(g_g - 5.5\right) + \gamma_{e2}\left(c_y - 2005.5\right) + v_{eu}\right]E_s + e_{usgy} + \epsilon_{usgy} \\
\epsilon_{usgy} &\sim N\left[0, \hat{\omega}^2_{usgy}\right] \\
e_{usgy} &\sim N[0, \sigma^2] \\
\begin{bmatrix} v_{mu} \\ v_{eu} \end{bmatrix} &\sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{mm} & \tau_{me} \\ \tau_{me} & \tau_{ee} \end{pmatrix}\right]
\end{aligned}
\tag{1}
$$

Here, $\hat{G}_{usgy}$ is the estimated achievement gap for unit u in subject s in grade g and cohort c, where cohort is defined as the year in which a given group of students were expected to be in spring of first grade (for example, eighth graders in 2009 are in cohort 2001); its estimated standard error is $\hat{\omega}_{usgy}$. The variables M_s and E_s are indicator variables indicating, respectively, whether $\hat{G}_{usgy}$ describes a math or ELA gap. Grade (g_g) and cohort (c_y) are centered around 5.5 and 2005.5, respectively, so that the math and ELA intercepts (γ_{m0} and γ_{e0}, respectively) describe the average achievement gap at the midpoint of the cohorts (2001 to 2010) and grades (3 to 8) represented in our data. The error terms in the model indicate that estimated gaps may vary in three ways, net of linear subject-specific grade and cohort trends. First, estimated gaps may differ from their true values because of sampling variance; this is indicated by the error term ϵ_{usgy}, which is assumed to be normally distributed with a known variance equal to $\hat{\omega}^2_{usgy} = var\left(\hat{G}_{usgy}\right)$. Second, gaps may differ from their unit-specific grade-cohort-subject predicted value because of within-unit variation not captured by the subject-specific grade and cohort trends; this is indicated by the error term e_{usgy}, which is assumed normally distributed with a constant variance σ^2 that must be estimated. Third, unit-specific average math and ELA gaps may deviate from the mean math and ELA gaps among units (which are denoted by γ_{m0} and γ_{e0}, respectively). We allow these deviations (v_{mu} and v_{eu}) to differ for math and ELA; they are assumed multivariate normal with a variance matrix $\tau = \begin{pmatrix} \tau_{mm} & \tau_{me} \\ \tau_{me} & \tau_{ee} \end{pmatrix}$ that must be estimated.

The key parameters of interest here are the variance components σ^2 and τ. If σ^2 is small compared to τ_{mm} and τ_{ee} (that is, if the intraclass correlations $\rho_m = \frac{\tau_{mm}}{\tau_{mm}+\sigma^2}$ and $\rho_e = \frac{\tau_{ee}}{\tau_{ee}+\sigma^2}$ are large), then there is little variation within units relative to the variation in gaps among units. In addition, if the correlation between v_{mu} and v_{eu}, $\rho_{me} = \frac{\tau_{me}}{\sqrt{\tau_{mm} \cdot \tau_{ee}}}$, is high, then the math and ELA gaps in a unit contain little unique information. Taken together, if there is little variation within geographic unit—either across subjects, grades or cohorts—then pooling the data so that there is one observation in each geographic unit will be an effective way to simplify presentation with little loss of information.

The estimates from this model are shown in Panel A of Table 2. We fit model (1) separately for white-black and white-Hispanic gaps and for metropolitan areas and school districts. In the metropolitan area models, the intraclass correlations of achievement gaps are between 0.92 and 0.95; in the district models, they are between 0.88 and 0.92. In each case, then, roughly 90% of the within-subject variation in achievement gaps is between metropolitan areas or districts. The correlations between average math and ELA gaps within geographic units are also quite high, ranging from 0.91 to 0.95. Moreover, results from Panel A in Table 2 indicate that the pooled estimates are sufficiently precise that we can very reliably distinguish among geographic units. The subject-specific estimated reliabilities range between 0.96-0.98; for district level estimates it is 0.88-0.89. These results indicate that we can pool estimated gaps within geographic units with little loss of information and with sufficient precision to distinguish estimates among geographic locales.

[Table 2 here]

Given this evidence, we pool the estimated achievement gaps within each metropolitan area and district as follows. For each district or metropolitan area u, we fit the model:

$$\hat{G}_{usgy} = \gamma_{u0} + \gamma_{u1}(g_g - 5.5) + \gamma_{u2}(c_y - 2005.5) + \gamma_{u3}(M_s - 0.5) + e_{usgy} + \epsilon_{usgy}$$
$$\epsilon_{usgy} \sim N[0, \hat{\omega}^2_{usgy}]$$
$$e_{usgy} \sim N[0, \sigma^2] \tag{2}$$

We take $\hat{\gamma}_{u0}$, the OLS—or "unshrunken" estimated fitted value of γ_{u0}—as our pooled estimate. This is an estimate of the pooled math and ELA achievement gap in a geographic unit u in the middle grade and cohort of our data based on the (up to) 60 gap estimates available in each district or metropolitan area. We use these pooled estimates for the remainder of our analysis.

We show the results from the pooling model in Panel B of Table 2. In this pooled model, intra-class correlations are similar in magnitude compared to the subject-specific models, ranging between 0.93 to 0.95 for metropolitan statistical areas and for school districts. The reliability of metropolitan-level intercept random effects ranges between 0.97 and 0.99 for white-black and white-Hispanic gaps, respectively; for district level estimates, the reliabilities are a little smaller, ranging between 0.91 and 0.92 for white-black and white-Hispanic gaps, respectively.

Panel B of Table 2 shows that average achievement gaps are large and vary considerably in magnitude across the U.S. Across school districts, the average white-black and white-Hispanic gaps are 0.66 and 0.50 standard deviations, respectively, with standard deviations of 0.22 and 0.23, respectively. The means (0.75 and 0.55) and their standard deviations (0.20 and 0.21, respectively) are similar at the metropolitan level. Analyses (not shown here) indicate that only 11-13% of the variance in district-level achievement gaps is due to between-state variation. That is, within a given state, achievement gaps typically vary almost as much as they do nationwide. Thus, analyses of state-level achievement gaps (see, for example, Hemphill, Vanneman and Rahman 2011; Vanneman et al. 2009) miss almost all of the geographic variation in achievement gap patterns.

Districts and Metropolitan Areas with the Smallest and Largest Racial/Ethnic Test Score Gaps

One of the benefits of these data is that we can identify districts and metropolitan areas with the largest and smallest achievement gaps. Figures 2-5 identify the 20 school districts and metropolitan areas with the largest and smallest estimated achievement gaps in the US. Figures 2 and 3 list the 20 school districts, while Figures 4 and 5 list the 20 metropolitan areas, indicating the estimated mean and 95% confidence interval for each area. School districts and metropolitan areas are ranked based on a "shrunken" Empirical Bayes (EB) estimate of the achievement gap, so that units with few students of a given race and imprecisely estimated gaps do not show up as the places with the most extreme gaps simply because of sampling error. These EB estimates are shrunken toward the predicted value of the gap.[7]

[Figures 2 and 3 here]

The lists of districts with large white-black achievement gaps includes several large and medium-sized school districts (Atlanta, GA; Washington DC; Orleans Parish, LA; Madison, WI; Charleston, SC; Tuscaloosa, AL; Minneapolis, MN; Oakland, CA), most of which are in the South and are generally highly segregated, with large white-black socioeconomic disparities. But it also includes a number of smaller school districts that are home to prominent universities (Berkeley, CA; Chapel Hill, NC; Evanston, IL; University City, MO) as well as a set of small, relatively affluent suburban/exurban school districts (Shaker Heights, OH; LaGrange, IL; Huntington Union, NY). The set of districts with the smallest white-black achievement gaps includes a number of districts with relatively small black populations as well as several large, racially diverse, poor school districts (notably Detroit, MI, and Clayton County, GA).

[7] Specifically, the EB estimates are derived from a version of Model (2) that includes the full set of covariates in Δ_u and X_u used in Model (3) described below (as well as the state fixed effects λ_s for districts). The inclusion of the covariates and state fixed effects provides more information for the EB estimates when precision is low and has little impact on EB estimates when precision is high. Any district or metropolitan area that is not reliably estimated will be shrunken towards the population mean and will not show up here in the rankings of areas with large or small gaps.

Many of the districts with the largest white-black achievement gaps also appear on the list of places with the largest white-Hispanic gaps (Atlanta, GA; Chapel Hill, NC; Evanston, IL; Berkeley, CA; Washington, DC; Minneapolis, MN), suggesting either that the local forces producing racial/ethnic inequality are not specific to one race/ethnic group or that the gaps are large because of particularly high performance of white students (rather than particularly low performance of black or Hispanic students). Many of the districts with the largest white-Hispanic gaps are in the Bay Area in California (San Rafael; Berkeley; Mountain View; Cabrillo Unified (Half Moon Bay); Menlo Park), where white-Hispanic socioeconomic inequality and segregation are very high. Among the districts with the smallest white-Hispanic achievement gaps, many are in small, relatively low-income school districts in Texas and California.

[Figures 4 and 5 here]

Among metropolitan areas, those with the largest gaps all have relatively large black or Hispanic populations and large racial/ethnic socioeconomic disparities and relatively high levels of segregation. Ten metropolitan areas are on both the list of places with the 20 largest white-black gaps and the list of those with the 20 largest white-Hispanic gaps (notably, Bridgeport, CT and San Francisco, CA are among the top three on both lists). These metropolitan areas generally have very affluent white populations that are substantially segregated from very poor minority populations.

Associations Between Racial/Ethnic Test Score Gaps and District and Metropolitan Area Characteristics

Data on Local Contextual Characteristics

Given the substantial variation in achievement gaps among districts and metropolitan areas, we next examine the extent to which the gaps are associated with five sets of characteristics of school districts and metropolitan areas: 1) average socioeconomic conditions, 2) racial/ethnic composition, 3) racial socioeconomic disparities, 4) patterns of residential and school segregation, and 5) school system characteristics. Average socioeconomic conditions and racial composition describe school districts' aggregate levels of socioeconomic status and racial composition. The other three covariate sets correspond to three key elements of the conceptual framework laid out in Figure 1: racial socioeconomic disparities, residential and school segregation patterns, and school policies, practices, and conditions.

The covariates come primarily from two data sources. First is the American Community Survey (ACS) profile tables for years 2006-2010, available for download from the Education Demographic and Geographic Estimates (EDGE) web portal.[8] The EDGE data come from a special school district-level tabulation of the ACS. The data include tabulations of demographic and socioeconomic characteristics of families who live in each school district in the U.S. and who have children enrolled in public school, the same population represented in the ED*Facts* achievement data. We use estimates from the 2006-2010 pooled file, which combines data ACS data from 5 survey years.[9]

Second is the Common Core of Data (CCD) universe surveys and finance files for years 2009-2013.[10] The CCD Public Elementary/Secondary School Universe is an annual survey of all public elementary and secondary schools in the United States. The data include basic descriptive information on schools and school districts, including staff and enrollment counts. The CCD Local Education Agency (School District) Finance Survey (F-33) Data contains district level expenditures data by year for all districts in the United States. We compute the measures discussed below using CCD data from 2009-2013 (to overlap with the years of our achievement data) and then take the mean across years to construct a single value for each district. Below we briefly discuss the measures we use; more information on how each variable is constructed is available in Appendix A5.

[8] The ACS EDGE data tables are available for download at http://nces.ed.gov/programs/edge/demographicACS.aspx.
[9] Annual estimates are available only for school districts with 65,000 or more persons; the 5-year pooled estimates are available for all school districts, regardless of size. We use the 2006-2010 ACS tables because the available more recent tables do not include the specific tabulations describing parents who have children enrolled in public schools. For a useful description of the respective strengths and weaknesses of the different surveys, see http://www.census.gov/programs-surveys/acs/guidance/estimates.html.
[10] CCD universe surveys and finance files are available for download at https://nces.ed.gov/ccd/ccddata.asp.

Measures of Local Contextual Characteristics

We construct several cross-sectional indicators of school district and metropolitan areas' average socioeconomic characteristics. These include (logged) median family income, the proportion of families in which a parent has a bachelor's degree, poverty rates, unemployment rates, SNAP receipt rates, and single mother headed household rates. Each of these are tabulated in the EDGE data for the population of families with children enrolled in public schools. We also construct measures of the white-black and white-Hispanic disparities of each of these characteristics. Additionally, we construct a measure of the free-lunch eligibility rate in each school district from the CCD data; this measure is not available by race.

To ease interpretation and reduce multi-collinearity in our multivariate models, we generate a socioeconomic status (SES) composite that is the first principal component of the following variables: the log of median family income, the proportion of adults with a bachelor's degree, the poverty rate, the unemployment rate, the SNAP receipt rate, and the single mother headed household rate. The SES composite variable is based on data from all 11,582 districts in the United States with sufficient EDGE data and is standardized within this set of districts. Since the several thousand districts in our analytic sample of districts with available achievement gap estimates are slightly poorer than the full population of US districts, on average, the mean of the composite is below 0 in our sample. We construct race-specific SES composites for whites, blacks and Hispanics using the factor loadings for the full population. A racial SES composite gap is then constructed by subtracting these race-specific composites. Descriptive statistics for the aggregate SES composite and factor loadings are shown in Table 3.

[Table 3 here]

We compute the proportion of students in public schools who are black and Hispanic using the CCD and the proportion of Hispanics fluent in English using the EDGE data. We compute three types of measures of segregation: measures of racial segregation, measures of income segregation, and measures of racial differences in exposure to poverty. The last is of particular interest, given evidence that the racial difference in average school poverty rates is the measure of segregation most highly correlated with racial achievement gaps (Reardon 2016). For each of these, we compute measures of both school segregation (computed based on school enrollment data from CCD), and residential segregation (computed from EDGE data and available only for metropolitan areas).

The final set of covariates describes characteristics of the school system in school districts and metropolitan areas. These include several crude measures of school resources (per-pupil expenditures and student-teacher ratios), and the proportion of students attending charter schools, which we include because there is evidence that some—though not all—charter schools have been shown to be particularly effective at raising minority students' achievement (Bifulco and Bulkley 2015; Center for Research on Education Outcomes (CREDO) 2015; Gleason et al. 2010). These data are constructed from the CCD. We also compute measures of within-district or –metro racial/ethnic disparities in these characteristics—that is, measures of the extent to which white, black, and Hispanic students' schools or districts differ, on average, on these measures (note that racial differences in average per-pupil district expenditure data are constructed only for metropolitan areas because school-level expenditure data are not available).

Because these data come from multiple sources and because of minimum group size requirements for reporting (e.g., income is not reported for districts or tracts with small populations), the analytic sample is restricted to those districts and metropolitan areas for which there are no missing data for any of the covariates. For the metropolitan area analyses, the analytic samples for white-black and white-Hispanic gaps include 361 and 371 metropolitan areas (96 and 98 percent of metropolitan areas for which gap estimates are available), respectively. For the district analyses, the analytic samples include 2,341 and 3,038 districts (85 and 88 percent of the districts for which white-black and white-Hispanic gap estimates are available), respectively.

Table 4 reports means and standard deviations for these variables. There is considerable variation among school districts in many of the socioeconomic and demographic characteristics, with less variation among metropolitan areas. This is evident in a comparison of the standard deviations of the socioeconomic characteristics among school districts and among metropolitan areas. One consequence of this sorting is that between-race differences in socioeconomic status are smaller in districts, on average, than in metropolitan areas. In other words, districts tend to be more socioeconomically and racially homogeneous, on average, than metropolitan areas.

[Table 4 Here]

A Descriptive Model of Racial/Ethnic Test Score Gaps

We examine both bivariate and multivariate associations between achievement gaps and district and metropolitan area characteristics. Both the bivariate and multivariate associations are estimated using models of the form

$$\hat{\gamma}_{u0} = \alpha + (\mathbf{X}_u - \overline{\mathbf{X}})\mathbf{B} + \mathbf{\Delta}_u\mathbf{\Gamma} + \lambda_s + e_u + v_u; \;\; e_u \sim N[0, \tau]; \;\; v_u \sim N[0, \widehat{\omega}_u^2], \tag{3}$$

where $\hat{\gamma}_{u0}$ is the estimated (pooled) white-black or white-Hispanic achievement gap in a district or metropolitan area u (obtained from Model 2 above); $\mathbf{X}_u$ is a vector of district or metropolitan area covariates (including average socioeconomic conditions, racial/ethnic composition, and measures of educational policies and practices); and $\mathbf{\Delta}_u$ is a vector of covariates describing racial/ethnic *differences* in context or experience (including racial disparities in family resources and measures of residential and school segregation). $\mathbf{X}_u$ and $\mathbf{\Delta}_u$ can be distinguished in that $\mathbf{\Delta}_u$ has a natural interpretation at zero—there is no racial inequality or segregation in the geographic unit—whereas $\mathbf{X}_u$ corresponds to the average value of the variable in the geographic unit. We center $\mathbf{X}_u$ at the sample mean, but leave $\mathbf{\Delta}_u$ uncentered. This is done so that the intercept α can be interpreted as the average achievement gap in a district with average values of $\mathbf{X}$ and in which white and black/Hispanic students experience equal values of the contextual factors contained in $\mathbf{\Delta}$.

We include state fixed effects, denoted λ_s, in the district models but not the metropolitan area models. The residual error term e_u is assumed to have constant variance τ; the error term v_u is the sampling error in $\hat{\gamma}^*_{u0}$ (that is, $v_u = \hat{\gamma}^*_{u0} - \gamma_{u0}$) and is assumed to have known error variance ω_u^2 (where ω_u is the estimated standard error of $\hat{\gamma}^*_{u0}$). The two error terms are assumed independent of one another. Because e_u may not be independent of $\mathbf{X}_u$ and $\boldsymbol{\Delta}_u$, however, we cannot interpret the estimated coefficient vectors $\widehat{\mathbf{B}}$ and $\widehat{\boldsymbol{\Gamma}}$ in causal terms.

Models with the error structure of Model (3) are sometimes referred to as meta-analytic regression models or precision-weighted random effects models. Such models are appropriate when the outcome variable for each observation represents an estimated value (with known error variance) of a parameter from a different site, and where the true values of that parameter are assumed to vary among sites. We fit these models using Stata's `–metareg–` command (Harbord and Higgins 2008).

Correlates of Racial/Ethnic Test Score Gaps

The lists of school districts and metropolitan areas with the largest and smallest achievement gaps in Figures 2-5 suggest a few patterns: achievement gaps are largest in places with large racial/ethnic differences in socioeconomic status, in more segregated places, and in more affluent or socioeconomically advantaged places; they are smallest in smaller, poorer school districts where socioeconomic disparities are relatively small or where there are few minority students. Using data from all school districts and metropolitan areas in the analytic sample, the bivariate analyses reported in Table 4 examine whether these patterns hold more generally.

Table 5 reports the pairwise correlations between achievement gaps and a set of district and metropolitan area characteristics. These characteristics are organized into the five categories described earlier: 1) average socioeconomic conditions, 2) racial/ethnic composition, 3) racial socioeconomic disparities, 4) residential and school segregation, and 5) school system characteristics.

[Table 5 here]

Results from Table 5 suggest five general patterns. First, achievement gaps are larger in more affluent areas. Districts and metropolitan areas with high SES composite scores, higher median incomes, higher rates of adults with bachelor's degrees, lower rates of students qualifying for free lunch, and lower rates of single mother headed households have larger differences in achievement between white and minority students. The correlation between the SES composite and achievement gaps, for example, ranges between .26 to .27 in school districts and .26 to .47. in metropolitan areas. Further, because district average achievement is highly correlated with district socioeconomic status, we obtain similar correlations between achievement gaps and average district achievement (.23 to .33).

Second, white-black and white-Hispanic gaps tend to be larger in metropolitan areas with larger black and Hispanic public school enrollments, respectively. Racial/ethnic composition is largely unrelated to achievement gaps at the district level, however. The Hispanic population's English proficiency is negatively associated with the white-Hispanic achievement gap in both metropolitan areas and school districts. Third, areas with larger racial socioeconomic disparities have larger achievement gaps. The white-minority SES composite gap is correlated with achievement gaps between .35 to .39 in school districts and .51 to .52 in metropolitan areas. Correlations between racial income and adult education differences and achievement gaps range from .40 to .62, depending on population. This is not surprising given the large body of research describing the relationship between socioeconomic status and academic achievement.

Fourth, districts and metropolitan areas with higher levels of racial and economic segregation have larger achievement gaps, on average, than less segregated places. Although all the measures of segregation we examine are significantly correlated with achievement gaps, the most highly correlated measure is the difference in the extent to which white and minority students have schoolmates who are eligible for free lunch, consistent with evidence in Reardon (2016). In metropolitan areas and districts where black and Hispanic students attend schools with higher average poverty rates than white students, achievement gaps are larger, on average, than in places with smaller differences in exposure to school poverty (r=0.65 to 0.71 in metropolitan areas and r=0.38 to 0.42 in districts for white-black and white-Hispanic differences, respectively).

Fifth, although some measures of educational policies and practices are correlated with achievement gaps, these correlations are generally small and inconsistent across groups and geographic units of analysis. The most consistent relationship between schooling characteristics and achievement gaps is the positive association between per pupil instructional expenditures and achievement gaps; in areas with greater spending, even after controlling for between state differences (via the state fixed effects in the district models), achievement gaps tend to be larger than in those with lower spending. Of course, this association is not causal, and does not imply that increasing spending increases achievement gaps. It may be that places with greater spending have greater shares of higher income, non-minority families, or that school spending is modestly responsive to academic inequality. Relatedly, the table shows a negative and weak correlation between achievement gaps and racial differences in per pupil expenditures. This negative correlation indicates that achievement gaps are smaller, on average, in metropolitan areas where per pupil expenditures are higher in white students' districts than in minority students' districts. Again, this should not be interpreted causally. The correlations here do not indicate whether spending differences cause differences in achievement gaps or whether differences in achievement induce spending differences (or whether both are driven by some third factor).

Socioeconomic disparities and racial/ethnic achievement gaps

Given the relatively strong correlations between each of the racial/ethnic socioeconomic disparities measures and achievement gaps, we next investigate a) how much variation in achievement gaps can be accounted for by variation in racial socioeconomic disparities and in racial differences in exposure to poverty; and b) whether achievement gaps are zero, on average, in places with no racial/ethnic disparities on these measures. To answer these questions, we fit two versions of the model shown in Equation (3) including either (i) a vector $\mathbf{\Delta}_u$ of racial/ethnic differences in family income, parental education, occupational status, unemployment rates, poverty rates, SNAP receipt rates, single mother headed household rates, rental rates, and rates of residential mobility or (ii) racial differences in exposure to free lunch-eligible schoolmates. From this model, we compute $\mathbf{\Delta}_u\hat{\mathbf{\Gamma}}$, the predicted contribution of racial disparities to the local academic achievement gap. We also compute the EB estimate of the achievement gap from this model, $\hat{\gamma}^*_{u0}$, and then plot $\hat{\gamma}^*_{u0}$ against $\mathbf{\Delta}_u\hat{\mathbf{\Gamma}}$ (Figures 6-9). Each figure also includes the fitted line $\hat{\gamma}^*_{u0} = \hat{\alpha} + \mathbf{\Delta}_u\hat{\mathbf{\Gamma}}$. Note that the intercept of this line indicates the estimated average achievement gap in districts or metropolitan areas in which there are no racial/ethnic disparities in socioeconomic status or exposure to poverty.

Figures 6 and 7 display the association between district level white-black and white-Hispanic achievement gaps to the corresponding socioeconomic disparities (Figure 6) and to the disparities in exposure to school poverty (Figure 7). Each point in the figures corresponds to a school district; the size of each point is proportional to the average number of black or Hispanic students in the district. Several patterns are evident in the figures. First, as noted in Table 2, there is considerable variation among school districts in the magnitude of racial/ethnic socioeconomic disparities. For white-black gaps, nearly 11 percent of districts (with 7 percent of the black public school population) have racial/ethnic socioeconomic disparities that are less than or equal to zero (though many of these districts are small or have small numbers of black families, so that their white-black SES differences are imprecisely estimated). For white-Hispanic gaps, only 3 percent of districts (with less than 1 percent of the Hispanic public school population) have racial/ethnic socioeconomic disparities that are less than or equal to zero. In the vast majority of districts, the disparities are greater than zero, and in some the disparities are large—large enough to correspond to two thirds or more of a standard deviation of the achievement gap (recall that the horizontal axis is scaled in the same units as the achievement gap, since $\mathbf{\Delta}_u\hat{\mathbf{\Gamma}}$ is the magnitude of the contribution of socioeconomic disparities to achievement gaps).

Second, there is a moderate association between socioeconomic disparities and achievement gaps: the R^2's from the models in Figure 6 are 0.41 and 0.40 for white-black and white-Hispanic gaps, respectively (implying that the correlation between district-level achievement gaps and an index of racial socioeconomic differences ($\boldsymbol{\Delta}_u\hat{\boldsymbol{\Gamma}}$) is roughly 0.64). Racial differences in exposure to school poverty also explain a portion of achievement gaps: the R^2's from these models in Figure 7 are .15, implying that the correlation between district-level achievement gaps and racial differences in exposure to poverty is .39.

In other words, racial socioeconomic disparities and racial differences in exposure to poverty are moderately strong predictors of academic achievement gaps. There nonetheless remains considerable variation in achievement gaps, even conditional on racial socioeconomic disparities. In Figure 6, the conditional standard deviation of achievement gaps around the fitted line is roughly 0.17; districts with similar socioeconomic disparities vary in some cases by as much as half a standard deviation in their achievement gaps.

Third, in Figures 6 and 7 the intercept of the fitted line is 0.50 for the white-black models, and in the white-Hispanic models, the intercept is 0.25 in Figure 6 and .40 in Figure 7. That is, even in the relatively few districts where white and minority students have similar socioeconomic backgrounds and levels of economic isolation, the average district-level achievement gap is well above zero. Thus, racial/ethnic socioeconomic disparities alone do not account for the large racial achievement gaps, despite being highly predictive of the magnitudes of the gaps.

[Figures 6-7 here]

Figures 8 and 9 are similar to Figures 6 and 7, but illustrate the patterns for metropolitan areas rather than school districts. First note that in only a few metropolitan areas (with less than 0.1 percent of the black population and less than 0.2 percent of the Hispanic population) is there racial/ethnic socioeconomic equality. More generally, however, Figures 8 and 9 show that the associations between metropolitan area achievement gaps and racial/ethnic socioeconomic and exposure to poverty disparities are very similar to the district-level associations shown in Figure 6 and 7. First, in Figure 8 the R^2's are 0.38 and 0.57 (implying that the correlation between metropolitan area achievement gaps and an index of racial socioeconomic differences is roughly 0.62-0.75); in Figure 9 the R^2's are .42 and .50 (implying that the correlation between metropolitan area achievement gaps and racial differences in exposure to poverty are .65 and .71 Second, while racial/ethnic socioeconomic disparities and racial differences in exposure to poverty are strong predictors of academic achievement gaps, there is considerable variation in the magnitude of achievement gaps, even among metropolitan areas with similar degrees of racial socioeconomic inequality. Finally, even when white-black and white-Hispanic socioeconomic disparities are zero, white-black and white-Hispanic achievement gaps are roughly 0.43 and 0.13, respectively. Moreover, when blacks and Hispanics have similar exposure to school poverty as whites, the white-black and white-Hispanic achievement gaps are .60 and .40, respectively.

[Figures 8-9 here]

Multivariate regression model results

In order to describe the partial associations between achievement gaps and each covariate, net of other measures, we present a set of multivariate regression models that include the a restricted set of covariates ($\mathbf{X}_u$ and $\mathbf{\Delta}_u$), including the SES composite, racial/ethnic differences in the SES composite, white-minority free lunch rate gap (i.e., the difference in the extent to which white and minority students have schoolmates who are eligible for free lunch), and racial composition variables. The coefficients described here should not be interpreted causally. Rather, they simply indicate which covariates are the most robust predictors of achievement gaps. Results are shown in Table 6.

[Table 6 here]

Several patterns stand out in Table 6. First, even after controlling for racial socioeconomic differences, average socioeconomic status remains a significant predictor of achievement gaps.

Achievement gaps are larger in both school districts and metropolitan areas with higher levels of socioeconomic status, net of other characteristics. Less surprisingly, racial differences in socioeconomic status, as measured by the white-minority SES composite difference, remain a significant predictor in the model, net of other characteristics.

Second, the segregation and composition variables also consistently predict achievement gaps, net of other characteristics. In districts and metropolitan areas with higher white-minority free lunch rate differences, both white-black and white-Hispanic achievement gaps are larger. In districts and metropolitan areas with larger concentrations of black students, net of other characteristics, white-black achievement gaps are larger. In districts and metropolitan areas with larger concentrations of Hispanic students, net of other characteristics, white-Hispanic gaps are larger. Areas where the Hispanic population is more fluent in English also have lower white-Hispanic achievement gaps. In keeping with Reardon (2016), we find here that racial segregation that yields large racial disparities in exposure to poor schoolmates is strongly associated with larger achievement gaps. Collectively, the 6 variables included in these models explain between 39 and 72 percent of the variation in achievement gaps.

In the second set of models in Table 6, we include the set of school characteristics described above from Tables 4 and 5 (per pupil expenditures, class size, and charter enrollment). Here, we find no consistent association between achievement gaps and the measures of educational policies and practices, net of the other variables in the model. All of these measures have either very small coefficients or imprecise coefficients that sometimes even change signs across subgroups. Finally, the model R-squared increases by only a trivial amount when these additional school variables are included (ranging from an increase of 0.008 to 0.042).

Discussion

Several key findings emerge from these descriptive analyses. First, there is considerable variation in white-black and white-Hispanic achievement gaps across school districts and metropolitan areas. Yet most of this variation appears to be driven by local, rather than state-level forces: almost 90 percent of the variation in district achievement gaps lies within, rather than between, states. Although average levels of academic performance vary substantially among states, district racial achievement gaps do not differ much, on average, among states, at least not in comparison to how much they differ within states. Local forces dominate state-level processes in shaping patterns of racial/ethnic academic achievement gaps.

Second, of the several thousand school districts we analyze, which enroll over 90% of all black and Hispanic students in the U.S., there are but a handful where the achievement gap is near zero. With the notable exceptions of the Detroit, MI and Clayton County, GA school districts, these tend to be districts that enroll few minority students and where achievement gaps are very imprecisely estimated even in our large data set. And while Detroit and Clayton County do have achievement gaps near zero, this does not appear, at least in the case of Detroit, to be a desirable form of equity: Census data show that both white and black families in Detroit are very poor, on average, and given how low average test scores are in Detroit, the absence of an achievement gap implies that both black and white students are equally low scoring.[11] In other words, there is no school district in the U.S. that serves a moderately large number of black or Hispanic students where achievement is even moderately high and where achievement gaps are near zero.

Third, between roughly 40 - 60% of the variance in local achievement gaps can be explained by racial/ethnic disparities in socioeconomic status. The bivariate associations between achievement gaps and racial/ethnic differences in family income and parental education are strong. This is not surprising

[11] Average district NAEP data are available from the Trial Urban District Assessment (TUDA). Generally, subgroup achievement data are available from TUDA but sample sizes for white students in Detroit are too small to meet NAEP reporting guidelines. Average and black achievement data for Detroit are available for years 2009, 2011 and 2013 in ELA and math; data for Detroit and other large urban districts can be downloaded at http://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx.

given that many studies show a strong association between individual socioeconomic background and test performance and that racial/ethnic differences in socioeconomic status explain a substantial proportion of achievement gaps (Fryer and Levitt 2004; Fryer and Levitt 2006; Reardon and Galindo 2009; Rothstein and Wozny 2013). Nonetheless, the evidence here clearly indicates that those same associations appear to account for a significant proportion of geographic variation in achievement gaps as well.

Fourth, although racial/ethnic differences in socioeconomic status explain much of the variation in achievement gaps, socioeconomic disparities are far from determinative. Achievement gaps vary substantially even among places with similar socioeconomic disparities, and remain large even where white and minority students come from relatively similar socioeconomic backgrounds. There are clearly factors other than racial/ethnic socioeconomic disparities at play in generating academic achievement gaps.

Chief among these factors is racial segregation. Our analyses indicate that, net of socioeconomic disparities, general socioeconomic levels, demographic composition, and several (admittedly crude) measures of school quality and school quality disparities, segregation is a significant predictor of achievement gaps. In particular, while many of our measures of segregation are correlated with achievement gaps, the one that is a consistently significant predictor in our multivariate models is racial differences in exposure to poverty. Racial achievement gaps are larger, all else equal, in places where black and Hispanic students attend higher poverty schools than their white peers (see Reardon 2016 for a more detailed empirical argument on this point). This suggests that racial isolation, *per se*, is not the causal factor linking segregation to worse outcomes for minority students. Rather, racial isolation is correlated with other negative conditions such as exposure to more low-income peers, more crime, fewer positive role models, schools with fewer resources, etc. Because low-income students enter school with below-average academic skills, the curriculum and instructional practices in high poverty schools may, on average, target lower-level academic skills than in low-poverty schools. These differences in context and opportunities to learn between high- and low-poverty schools may explain why school poverty serves as a proxy for school quality, and why racial differences in school poverty lead to larger racial achievement gaps. Although our descriptive evidence provides only suggestive support for this claim, it is consistent with other recent studies. Specifically, rigorous causal identification strategies have been used to show that long-term exposure to poverty can have negative effects on cognitive and educational outcomes (Chetty, Hendren and Katz 2015; Sampson, Sharkey and Raudenbush 2008; Wodtke, Harding and Elwert 2011).

Our descriptive analyses reveal one additional somewhat puzzling pattern. Achievement gaps are larger, on average, in districts and metropolitan areas with higher levels of socioeconomic status, even after we control for many other variables, including racial socioeconomic disparities and segregation. As noted above, one possible explanation for this is the possibility that socioeconomic disparities—and corresponding disparities in social capital, social networks, and access to school district leaders—are more salient in competitive, high resource communities. Another possibility is that social psychological processes that inhibit minority students' performance, such as stereotype threat, are particularly strong in the most affluent places where academic performance is seen as a particularly important marker of intelligence and success and where minority students often comprise only a small share of school district enrollment (Steele 1997). A third possibility is that our socioeconomic measures understate the true racial resource disparities in the most economically advantaged places. Notably, our socioeconomic measures do not include wealth disparities (our only available proxy for wealth is median house value, but that is not available separately by race). If racial wealth disparities are particularly high in more advantaged communities, and if wealth is associated with academic performance after controlling for income (see, for example, Orr 2003; Yeung and Conley 2008), then unmeasured wealth disparities may account for the association we observe between parental education levels and racial achievement gaps.

It is also worth noting that our measures of school characteristics explain a very small amount of the variance in achievement gaps. Moreover, in some cases, the bivariate correlations (in Table 4) between racial disparities in school quality and achievement gaps have the opposite of the expected sign. There are several possible explanations for these patterns. First, as has been documented as far back as the Coleman Report (1966), the effect of schools on academic achievement may be relatively small relative to the impact of families. Second, it is possible that our measures of school quality do not capture the important features of school systems that affect racial achievement gaps. While rigorous studies of class size (one of our measures) and school spending indicate that both smaller classes and increased spending lead to higher achievement (Finn and Achilles 1990; Finn and Achilles 1999; Jackson, Johnson and Persico 2016; Lafortune, Rothstein and Schanzenbach 2016; Nye, Hedges and Konstantopoulos 2000), the racial differences in class size and average per pupil spending are not large (average class sizes are 1 to 4 percent larger in black and Hispanic students' schools than in white students' schools; average per pupil spending is roughly 3 percent greater in black and Hispanic students' school districts than in white students' districts; see Table 3). Moreover, these differences vary relatively little among metropolitan areas and school districts, so they have little power to explain the variation in racial achievement gaps. Finally, class size, spending, and charter school enrollment patterns are likely correlated with other unobserved characteristics of communities. For example, educational spending may be higher in communities with greater needs; there may be more demand for charter schools among minority students in communities where the local public schools are particularly low quality. Such patterns might give rise to bivariate correlations with the opposite of the expected sign—for example, the negative association between achievement gaps and white/minority per pupil instructional expenditures or the positive association between achievement gaps and minority-white differences in charter school enrollment rates (see Table 4).

Conclusion

We focus in this paper on achievement *gaps*, rather than achievement *levels*, because we are interested specifically in understanding the social processes that produce or ameliorate educational *inequality*. Educational success is both an absolute and a positional good; one's academic skills and one's rank in the distribution of educational outcomes both shape later educational and economic opportunities and outcomes. Our analyses here are designed to describe and analyze relative degrees of educational success. But that is not to say that questions about the social processes that produce higher levels of educational achievement and attainment are not equally interesting. We hope that the data we have used here can also be used to investigate the social and educational conditions that lead to not just to equal academic outcomes, but to equally high levels of academic outcomes.

Both sociological research and the design of effective social policies are enhanced by detailed data. And as administrative data systems grow more robust, sociologists increasingly have access to population-level data of the kind we use here. These data allow us to map the patterns of racial inequality in educational opportunity at a level of geographic detail not previously possible.

These data can also yield new theoretical insights and be used to test theoretical predications. We began with the standard sociological model that posits that racial academic achievement gaps are shaped by two key distal factors: racial differences in family socioeconomic circumstances and resources, and educational policies and practices. We hypothesized that patterns of residential and school segregation are key mediators through which both family resource disparities and educational policy affect achievement gaps. But because segregation patterns are jointly shaped by family resource patterns and social and educational policy, the relative contribution of these three factors to achievement gaps are not cleanly identifiable. That is, the old sociological question posed by Coleman and the generation of sociologists of education who followed him—"how much are the disparities in academic achievement due to family background, and how much are they due to inequalities in school environments?"—may not be answerable.

Instead, we set out to answer a different (and simpler) set of questions: “to what extent do racial achievement gaps vary across the U.S., and what are the strongest correlates of these gaps?” The analyses here confirm that family resource differences and segregation patterns are strongly associated with racial achievement gaps in school districts and metropolitan areas. In all of our analyses, racial socioeconomic disparities and segregation patterns are consistently the strongest predictors of racial achievement gaps. This is qualitatively consistent with our conceptual model, in which residential and school segregation play a key role in linking both family background and education policy to achievement gaps. And while school characteristics are correlated with achievement gaps, these correlations largely disappear once we control for segregation patterns and racial socioeconomic disparities. This may be because segregation patterns affect academic achievement through pathways other than those measured by our limited set of school quality variables.

In general, then, our analyses describe a set of distal factors that are strongly predictive of racial achievement gaps, but they do not identify the more proximal mechanisms that produce these gaps. Family resources and segregation patterns are likely linked to achievement gaps through a complex set of processes, including a host of racial differences in the opportunities and experiences of students in their homes, neighborhoods and schools. We would need better measures of such processes, as well as a research design capable of identifying their effects, to adjudicate among competing explanations of the proximal sources of achievement gaps.

References

Ananat, Elizabeth Oltmans. 2009. "The wrong side(s) of the tracks: The causal effects of racial segregation on urban poverty and inequality." *American Economic Journal: Applied Economics* 3(April):34-66.

Ashenfelter, Orley, William J. Collins, and Albert Yoon. 2005. "Evaluating the role of *Brown vs. Board of Education* in school equalization, desegregation and the income of African Americans." Vanderbilt University.

Aud, Susan, William Hussar, Michael Planty, Thomas Snyder, Kevin Bianco, Mary Ann Fox, Lauren Frolich, Jana Kemp, and Lauren Drake. 2010. "The Condition of Education 2010." Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Bassok, Daphna, Jenna E. Finch, RaeHyuck Lee, Sean F. Reardon, and Jane Waldfogel. 2016. "Socioeconomic Gaps in Early Childhood Experiences, 1998 to 2010." *AERA Open* 2(3):1-22.

Bifulco, Robert. 2005. "District-level Black-White funding disparities in the United States, 1987-2002." *Journal of Education Finance* 31(2):172-94.

Bifulco, Robert, and Katrina Bulkley. 2015. "Charter Schools." Pp. 423-43 in *Handbook of Research in Education Finance and Policy*, edited by Helen F. Ladd and Margaret E. Goertz. New York: Routledge.

Boyd, Don, Hamp Lankford, Susanna Loeb, Matt Ronfeldt, and Jim Wyckoff. 2011. "The effect of school neighborhoods on teachers' career decisions." *Whither opportunity*:377-96.

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2005a. "The draw of home: How teachers' preferences for proximity disadvantage urban schools." *Journal of Policy Analysis and Management* 24(1):113-32.

—. 2005b. "Explaining the short careers of high-achieving teachers in schools with low-performing students." *The American Economic Review* 95(2):166-71.

Bradley, Robert H., Robert F. Corwyn, Harriette Pipes McAdoo, and Cynthia Garcia Coll. 2001. "The home environments of children in the United States part I: Variations by age, ethnicity, and poverty status." *Child Development* 72(6):1844-67.

Burdick-Will, Julia, Jens Ludwig, Stephen W. Raudenbush, Robert J. Sampson, Lisa Sanbonmatsu, and Patrick Sharkey. 2011. "Converging Evidence for Neighborhood Effects on Children's Test Scores: An Experimental, Quasi-experimental, and Observational Comparison." Pp. 255-76 in *Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children*, edited by Greg J. Duncan and Richard J. Murnane. New York: Russell Sage Foundation.

Candelaria, Christopher A., and Kenneth A. Shores. 2017. "Court-Ordered Finance Reforms in The Adequacy Era: Heterogeneous Causal Effects and Sensitivity." *Education Finance and Policy*.

Card, David, and Jesse Rothstein. 2007. "Racial segregation and the black–white test score gap." *Journal of Public Economics* 91(11):2158-84.

Carneiro, Pedro, James J. Heckman, and Dimitry V. Masterov. 2003. "Labor market discrimination and racial differences in premarket factors " in *NBER working paper*. Cambridge, MA: National Bureau of Economic Research.

Carter, Prudence L. 2012. *Stubborn Roots: Race, Culture, and Inequality in U.S. and South African Schools*. New York: Oxford University Press.

Center for Research on Education Outcomes (CREDO). 2015. "National Charter School Study." Stanford, CA: Stanford University.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. 2015. "The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment." Harvard University.

Chin, Tiffani, and Meredith Phillips. 2004. "Social reproduction and child-rearing practices: Social class, children's agency, and the summer activity gap." *Sociology of Education* 77(3):185-210.

Clotfelter, Charles, Elizabeth Glennie, Helen Ladd, and Jacob Vigdor. 2008. "Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina." *Journal of Public Economics* 92(5):1352-70.

Clotfelter, Charles T, Helen F Ladd, and Jacob Vigdor. 2005. "Who teaches whom? Race and the distribution of novice teachers." *Economics of Education Review* 24(4):377-92.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederick D. Weinfeld, and Robert L. York. 1966. *Equality of educational opportunity*. Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education.

Cornman, Stephen Q. 2015. "Revenues and Expenditures for Public Elementary and Secondary Education: School Year 2011-12 (Fiscal Year 2012). First Look. NCES 2014-301." *National Center for Education Statistics*.

Cutler, David M., and Edward L. Glaeser. 1997. "Are ghettos good or bad?" *Quarterly Journal of Economics* 112(3):827-71.

Dahl, Gordon, and Lance Lochner. 2012. "The impact of family income on child achievement: Evidence from the Earned Income Tax Credit." *American Economic Review* 102(5):1927-56.

Duncan, Greg J., Pamela A. Morris, and Chris Rodrigues. 2011. "Does Money Really Matter? Estimating Impacts of Family Income on Young Children's Achievement with Data from Random-Assignment Experiments." *Developmental Psychology* 47(5):1263-79.

Finn, Jeremy D, and Charles M Achilles. 1990. "Answers and questions about class size: A statewide experiment." *American Educational Research Journal* 27(3):557-77.

—. 1999. "Tennessee's class size study: Findings, implications, misconceptions." *Educational Evaluation and Policy Analysis* 21(2):97-109.

Fryer, Roland G., and Stephen D. Levitt. 2004. "Understanding the black-white test score gap in the first two years of school." *The Review of Economics and Statistics* 86(2):447-64.

—. 2006. "The Black-White Test Score Gap through Third Grade." *American Law and Economics Review* 8(2):249-81.

Gleason, Philip, Melissa Clark, Christina Clark Tuttle, Emily Dwoyer, and Marsha Silverberg. 2010. "The Evaluation of Charter School Impacts: Final Report." Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Guryan, Jonathan. 2004. "Desegregation and Black dropout rates." *American Economic Review* 94(4):919-43.

Hanushek, Eric A, and Steven G Rivkin. 2009. "Harming the best: How schools affect the black-white achievement gap." *Journal of Policy Analysis and Management* 28(3):366-93.

Harbord, Roger M., and Julian P.T. Higgins. 2008. "Meta-Regression in Stata." *The Stata Journal* 8(4):493-519.

Hedges, L. V., and I. Olkin. 1985. *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hemphill, F. Cadelle, Alan Vanneman, and Taslima Rahman. 2011. "Achievement Gaps: How Hispanic and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress." Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Ho, Andrew D. 2008. "The problem with "proficiency": Limitations of statistics and policy under No Child Left Behind." *Educational Researcher* 37(6):351-60.

—. 2009. "A nonparametric framework for comparing trends and gaps across tests." *Journal of Educational and Behavioral Statistics* 34:201-28.

Ho, Andrew D., and Edward H. Haertel. 2006. "Metric-Free Measures of Test Score Trends and Gaps with Policy-Relevant Examples." Los Angeles, CA: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies.

Ho, Andrew D., D. M. Lewis, and J. L. M. Farris. 2009. "The dependence of growth-model results on proficiency cut scores." *Educational Measurement: Issues and Practice* 28(4):15-26.

Ho, Andrew D., and Sean F. Reardon. 2012. "Estimating Achievement Gaps From Test Scores Reported in Ordinal 'Proficiency' Categories." *Journal of Educational and Behavioral Statistics* 37(4):489-517.

Jackson, C Kirabo, Rucker C. Johnson, and Claudia Persico. 2016. "The effects of school spending on educational and economic outcomes: Evidence from school finance reforms." *Quarterly Journal of Economics* 131(1):157-218.

Jencks, Christopher, and Susan Mayer. 1990. "The social consequences of growing up in a poor neighborhood." Pp. 111-86 in *Inner-city poverty in the United States*, edited by Laurence E. Lynn, Jr. and Michael G. H. McGeary. Washington, DC: National Academy Press.

Johnson, Rucker C. 2011. "Long-Run Impacts of School Desegregation and School Quality on Adult Attainments." *National Bureau of Economic Research Working Papers* 16664.

Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach. 2016. "School Finance Reform and the Distribution of Student Achievement." *National Bureau of Economic Research Working Paper* (22011).

Lankford, Hamilton, Susanna Loeb, and James Wycoff. 2002. "Teacher sorting and the plight of urban schools: A descriptive analysis." *Educational Evaluation and Policy Analysis* 24(1):37-62.

Lareau, Annette. 2003. *Unequal Childhoods: Class, Race, and Family Life*. Berkeley, CA: University of California Press.

Lareau, Annette, and Kimberly Goyette. 2014. "Choosing homes, choosing schools: residential segregation and the search for a good school." New York: Russell Sage Foundation.

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):309-37.

Lewis, Amanda E., and John B. Diamond. 2015. *Despite the best intentions: How racial inequality thrives in good schools*. Oxford: Oxford University Press.

Logan, John R. 2011. "Separate and Unequal: The Neighborhood Gap for Blacks, Hispanics and Asians in Metropolitan America." in *US2010*, edited by John R. Logan: US2010 Project, Brown University.

Logan, John R., Deirdre Oakley, and Jacob Stowell. 2008. "School Segregation in Metropolitan Regions, 1970–2000: The Impacts of Policy Choices on Public Education." *American Journal of Sociology* 113(6):1611-44.

National Center for Education Statistics. 2013. "The Nation's Report Card: Trends in Academic Progress 2012." Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics, Education Finance Statistics Center. 2012. "Current Expenditures Per Pupil in Membership, by Poverty Quartile and State [no adjustment for students in poverty], 2012." Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Neal, Derek A. 2006. "Why has Black-White skill convergence stopped?" Pp. 511-76 in *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. New York: Elsevier.

Neal, Derek A., and William R. Johnson. 1996. "The role of premarket factors in black-white wage differences." *The Journal of Political Economy* 104(5):869-95.

Nisbett, Richard E, Joshua Aronson, Clancy Blair, William Dickens, James Flynn, Diane F Halpern, and Eric Turkheimer. 2012. "Intelligence: new findings and theoretical developments." *American Psychologist* 67(2):130.

Nye, Barbara, Larry V Hedges, and Spyros Konstantopoulos. 2000. "The effects of small classes on academic achievement: The results of the Tennessee class size experiment." *American Educational Research Journal* 37(1):123-51.

Ogbu, John U. 2003. *Black American Students in an Affluent Suburb: A Study of Academic Disengagement*. Mahwah, NJ: Lawrence Erlbaum Associates.
Orr, Amy J. 2003. "Black-White Differences in Achievement: The Importance of Wealth." *Sociology of Education* 76(4):281-304.
Owens, Ann. 2015. "Economic segregation betweeen neighborhoods and schools." edited by Unpublished manuscript.
—. 2016. "Inequality in Children's Contexts Income Segregation of Households with and without Children." *American Sociological Review* 81(3):549-74.
Owens, Ann, Sean F Reardon, and Christopher Jencks. 2016. "Income Segregation Between Schools and School Districts." *American Educational Research Journal*.
Pattillo, Mary. 2013. *Black picket fences: Privilege and peril among the black middle class*: University of Chicago Press.
Phillips, Meredith. 2011. "Parenting, time use, and disparities in academic outcomes." *Whither opportunity*:207-28.
Phillips, Meredith, James Crouse, and James Ralph. 1998. "Does the black-white test score gap widen after children enter school?" Pp. 229-72 in *The black-white test score gap*, edited by Christopher Jencks and Meredith Phillips. Washington, DC: Brookings Institution Press.
Reardon, Sean F. 2015. "Education." Pp. 43-48 in *State of the States: The Poverty and Inequality Report*, edited by David Grusky, Charles Varner, and Marybeth Mattingly. Stanford, CA: Stanford Center on Poverty and Inequality.
—. 2016. "School segregation and racial academic achievement gaps." *The Russell Sage Foundation Journal of the Social Sciences* 2(5):34-57.
Reardon, Sean F., Lindsay Fox, and Joseph Townsend. 2015. "Neighborhood Income Composition by Race and Income, 1990-2009." *The Annals of the American Academy of Political and Social Science* 660:78-97.
Reardon, Sean F., and Claudia Galindo. 2009. "The Hispanic-White Achievement Gap in Math and Reading in the Elementary Grades." *American Educational Research Journal* 46(3):853-91.
Reardon, Sean F., and Andrew D. Ho. 2015. "Practical Issues in Estimating Achievement Gaps From Coarsened Data." *Journal of Educational and Behavioral Statistics* 40(2):158-89.
Reardon, Sean F., Joseph P. Robinson-Cimpian, and Ericka S. Weathers. 2015. "Patterns and Trends in Racial/Ethnic and Socioeconomic Academic Achievement Gaps." Pp. 491-509 in *Handbook of Research in Education Finance and Policy*, edited by Helen Ladd and Margaret Goertz: Lawrence Erlbaum.
Reardon, Sean F., John T. Yun, and Tamela McNulty Eitle. 2000. "The changing structure of school segregation: Measurement and evidence of multi-racial metropolitan area school segregation, 1989-1995." *Demography* 37(3):351-64.
Rothstein, Jesse, and Nathan Wozny. 2013. "Permanent Income and the Black-White Test Score Gap." *The Journal of Human Resources* 48(3):509-44.
Rothstein, Richard. 2017. *The color of law: A forgotten history of how our government segregated America*. New York: Liveright Publishing Corporation.
Sampson, Robert J. 1998. "How do communities undergird or undermine human development? Relevant contexts and social mechanisms." in *Does it take a village? Community effects on children, adolescents, and families*. Pennsylvania State University.
—. 2008. "Moving to Inequality: Neighborhood Effects and Experiments Meet Social Structure." *American Journal of Sociology* 114(1):189-231.
Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and violent crime: A multilevel study of collective efficacy." *Science* 277:918-24.

Sampson, Robert J., Patrick Sharkey, and Stephen W. Raudenbush. 2008. "Durable effects of concentrated disadvantage on verbal ability among African-American children." *Proceedings of the National Academy of Sciences* 105(3):845-52.

Scafidi, Benjamin, David L Sjoquist, and Todd R Stinebrickner. 2007. "Race, poverty, and teacher mobility." *Economics of Education Review* 26(2):145-59.

Sharkey, Patrick. 2010. "The acute effect of local homicides on children's cognitive performance." *Proceedings of the National Academy of Sciences* 107(26):11733-38.

—. 2014. "Spatial Segmentation and the Black Middle Class." *American Journal of Sociology* 119(4):903-54.

Small, Mario Luis. 2006. "Neighborhood institutions as resource brokers: Childcare centers, interorganizational ties, and resource access among the poor." *Social Problems* 53(2):274-92.

Steele, Claude M. 1997. "A threat in the air: How stereotypes shape intellectual identity and performance." *American Psychologist* 52(6):613-29.

Stroub, Kori J., and Meredith P. Richards. 2013. "From Resegregation to Reintegration: Trends in the Racial/Ethnic Segregation of Metropolitan Public Schools, 1993–2009." *American Educational Research Journal* 50(3):497-531.

Sullivan, Laura, Tatjana Meschede, Lars Dietrich, Thomas Shapiro, Amy Traub, Catherine Ruetschlin, and Taara Draut. 2015. "The racial wealth gap: Why policy matters." New York, NY: Demos.

Tyson, Karolyn. 2011. *Integration Interrupted: Tracking, Black Students, and Acting White after Brown*. New York: Oxford University Press.

Vanneman, A., L. Hamilton, J. Baldwin Anderson, and T. Rahman. 2009. "Achievement Gaps: How Black and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress." Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Wodtke, Geoffrey T., David J. Harding, and Felix Elwert. 2011. "Neighborhood Effects in Temporal Perspective: The Impact of Long-Term Exposure to Concentrated Disadvantage on High School Graduation." *American Sociological Review* 76(5):713-36.

Wolff, Edward N. 2014. "The middle class: losing ground, losing wealth." Pp. 60-104 in *Diversity and Disparities: America Enters a New Century*, edited by John R. Logan. New York, NY: Russell Sage Foundation.

Yeung, W. Jean, and Dalton Conley. 2008. "Black–White Achievement Gap and Family Wealth." *Child Development* 79(2):303-24.

Figure 1

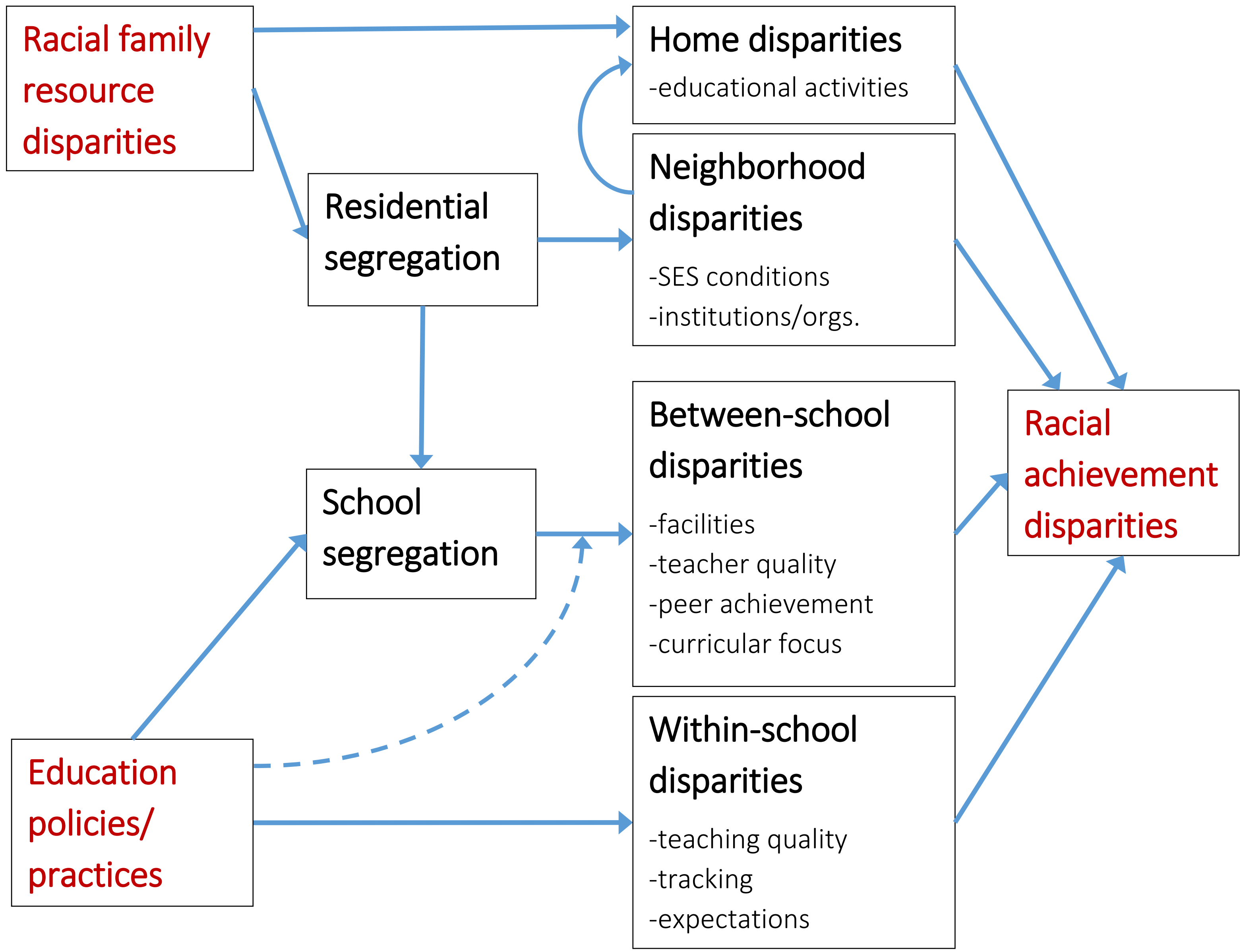

Racial family resource disparities
Home disparities
-educational activities
Neighborhood disparities
-SES conditions
-institutions/orgs.
Residential segregation
School segregation
Between-school disparities
-facilities
-teacher quality
-peer achievement
-curricular focus
Racial achievement disparities
Within-school disparities
-teaching quality
-tracking
-expectations
Education policies/ practices

Figure 2

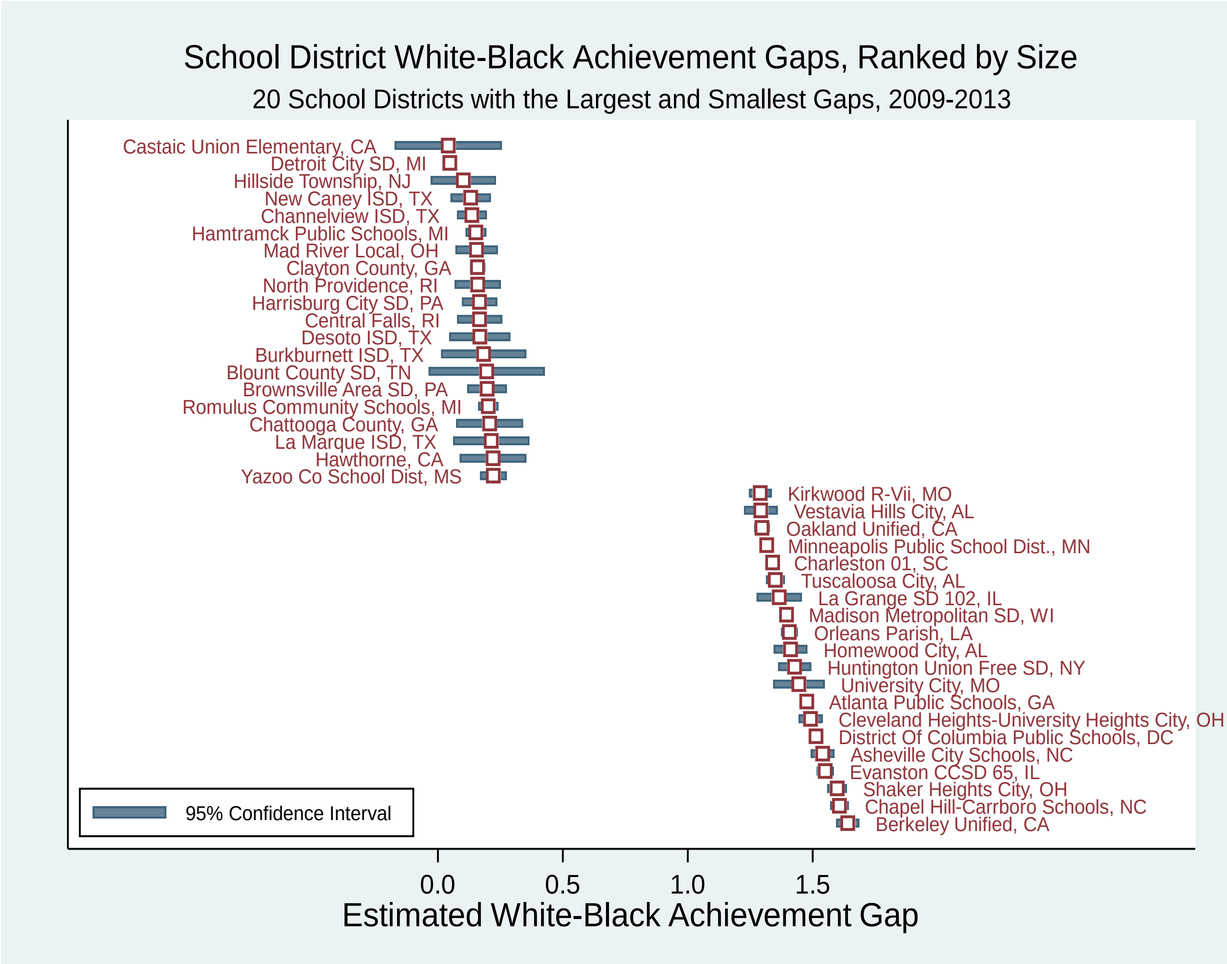
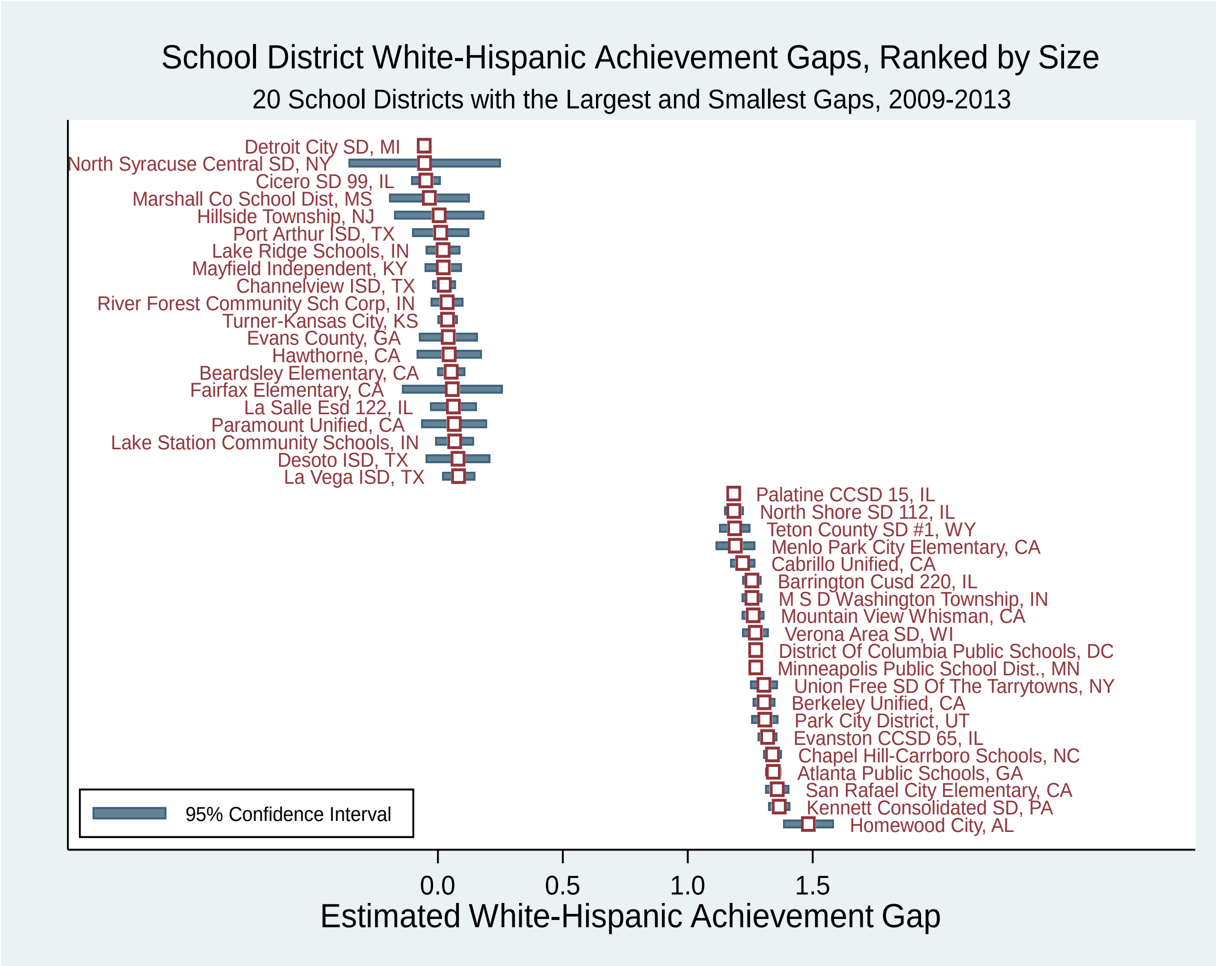
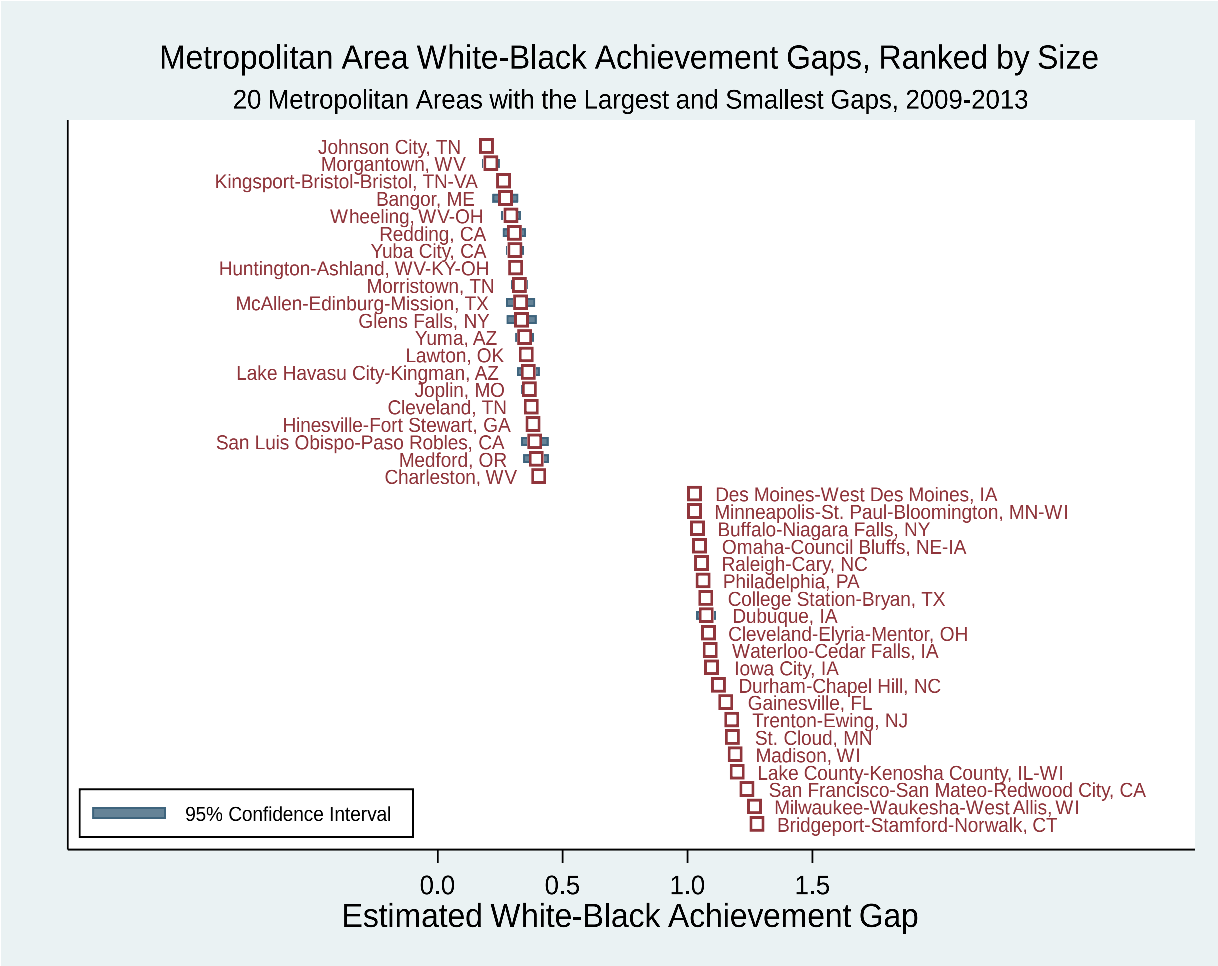

School District White-Black Achievement Gaps, Ranked by Size
20 School Districts with the Largest and Smallest Gaps, 2009-2013
Castaic Union Elementary, CA
Detroit City SD, MI
Hillside Township, NJ
New Caney ISD, TX
Channelview ISD, TX
Hamtramck Public Schools, MI
Mad River Local, OH
Clayton County, GA
North Providence, RI
Harrisburg City SD, PA
Central Falls, RI
Desoto ISD, TX
Burkburnett ISD, TX
Blount County SD, TN
Brownsville Area SD, PA
Romulus Community Schools, MI
Chattooga County, GA
La Marque ISD, TX
Hawthorne, CA
Yazoo Co School Dist, MS
Kirkwood R-Vii, MO
Vestavia Hills City, AL
Oakland Unified, CA
Minneapolis Public School Dist., MN
Charleston 01, SC
Tuscaloosa City, AL
La Grange SD 102, IL
Madison Metropolitan SD, WI
Orleans Parish, LA
Homewood City, AL
Huntington Union Free SD, NY
University City, MO
Atlanta Public Schools, GA
Cleveland Heights-University Heights City, OH
District Of Columbia Public Schools, DC
Asheville City Schools, NC
Evanston CCSD 65, IL
Shaker Heights City, OH
Chapel Hill-Carrboro Schools, NC
Berkeley Unified, CA
95% Confidence Interval
0.0
0.5
1.0
1.5
Estimated White-Black Achievement Gap

Figure 3



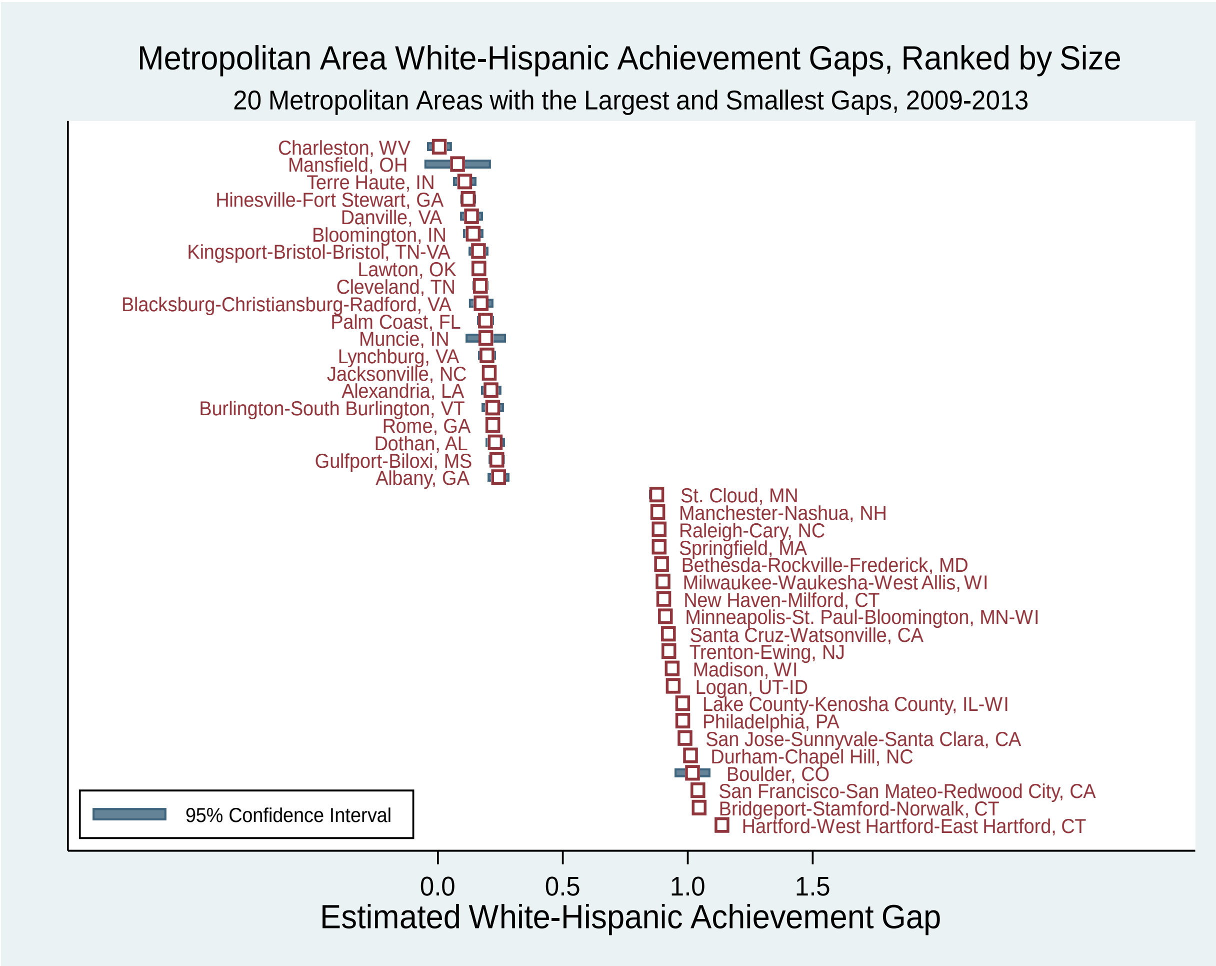
School District White-Hispanic Achievement Gaps, Ranked by Size
20 School Districts with the Largest and Smallest Gaps, 2009-2013
Detroit City SD, MI
North Syracuse Central SD, NY
Cicero SD 99, IL
Marshall Co School Dist, MS
Hillside Township, NJ
Port Arthur ISD, TX
Lake Ridge Schools, IN
Mayfield Independent, KY
Channelview ISD, TX
River Forest Community Sch Corp, IN
Turner-Kansas City, KS
Evans County, GA
Hawthorne, CA
Beardsley Elementary, CA
Fairfax Elementary, CA
La Salle Esd 122, IL
Paramount Unified, CA
Lake Station Community Schools, IN
Desoto ISD, TX
La Vega ISD, TX
Palatine CCSD 15, IL
North Shore SD 112, IL
Teton County SD #1, WY
Menlo Park City Elementary, CA
Cabrillo Unified, CA
Barrington Cusd 220, IL
M S D Washington Township, IN
Mountain View Whisman, CA
Verona Area SD, WI
District Of Columbia Public Schools, DC
Minneapolis Public School Dist., MN
Union Free SD Of The Tarrytowns, NY
Berkeley Unified, CA
Park City District, UT
Evanston CCSD 65, IL
Chapel Hill-Carrboro Schools, NC
Atlanta Public Schools, GA
San Rafael City Elementary, CA
Kennett Consolidated SD, PA
Homewood City, AL
95% Confidence Interval
0.0
0.5
1.0
1.5
Estimated White-Hispanic Achievement Gap

Figure 4



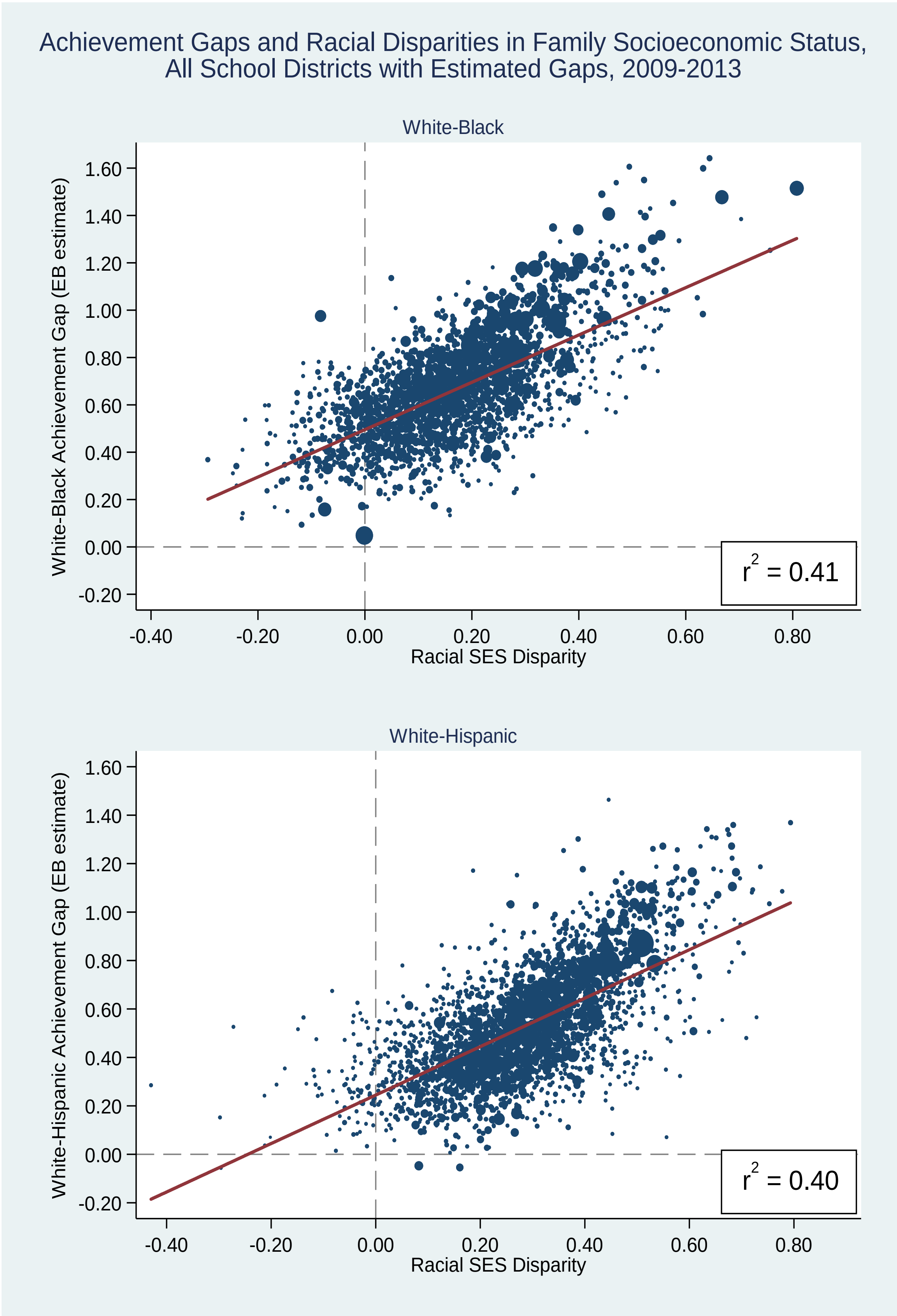
Metropolitan Area White-Black Achievement Gaps, Ranked by Size
20 Metropolitan Areas with the Largest and Smallest Gaps, 2009-2013
Johnson City, TN
Morgantown, WV
Kingsport-Bristol-Bristol, TN-VA
Bangor, ME
Wheeling, WV-OH
Redding, CA
Yuba City, CA
Huntington-Ashland, WV-KY-OH
Morristown, TN
McAllen-Edinburg-Mission, TX
Glens Falls, NY
Yuma, AZ
Lawton, OK
Lake Havasu City-Kingman, AZ
Joplin, MO
Cleveland, TN
Hinesville-Fort Stewart, GA
San Luis Obispo-Paso Robles, CA
Medford, OR
Charleston, WV
Des Moines-West Des Moines, IA
Minneapolis-St. Paul-Bloomington, MN-WI
Buffalo-Niagara Falls, NY
Omaha-Council Bluffs, NE-IA
Raleigh-Cary, NC
Philadelphia, PA
College Station-Bryan, TX
Dubuque, IA
Cleveland-Elyria-Mentor, OH
Waterloo-Cedar Falls, IA
Iowa City, IA
Durham-Chapel Hill, NC
Gainesville, FL
Trenton-Ewing, NJ
St. Cloud, MN
Madison, WI
Lake County-Kenosha County, IL-WI
San Francisco-San Mateo-Redwood City, CA
Milwaukee-Waukesha-West Allis, WI
Bridgeport-Stamford-Norwalk, CT
95% Confidence Interval
0.0
0.5
1.0
1.5
Estimated White-Black Achievement Gap

Figure 5



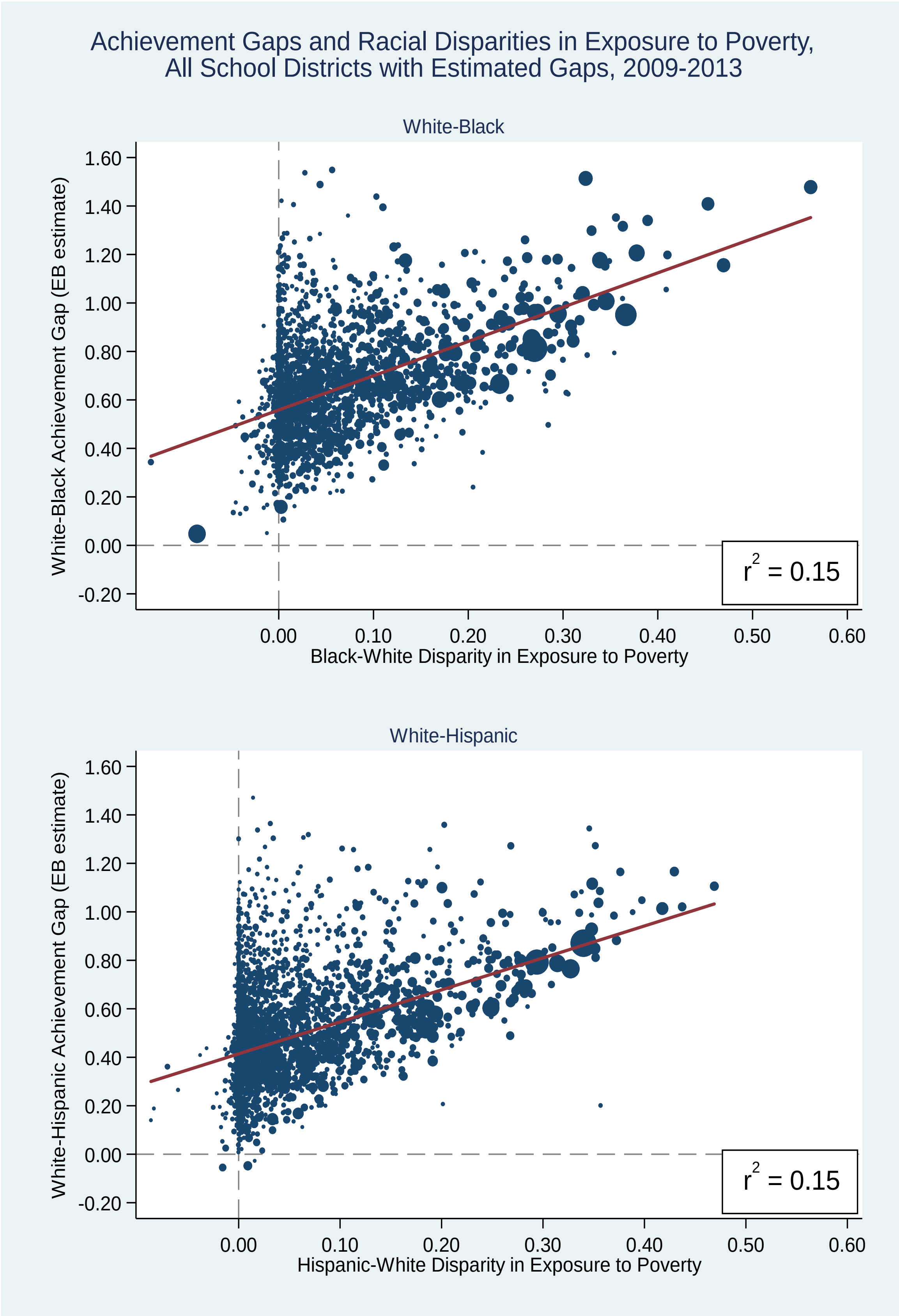
Metropolitan Area White-Hispanic Achievement Gaps, Ranked by Size
20 Metropolitan Areas with the Largest and Smallest Gaps, 2009-2013
Charleston, WV
Mansfield, OH
Terre Haute, IN
Hinesville-Fort Stewart, GA
Danville, VA
Bloomington, IN
Kingsport-Bristol-Bristol, TN-VA
Lawton, OK
Cleveland, TN
Blacksburg-Christiansburg-Radford, VA
Palm Coast, FL
Muncie, IN
Lynchburg, VA
Jacksonville, NC
Alexandria, LA
Burlington-South Burlington, VT
Rome, GA
Dothan, AL
Gulfport-Biloxi, MS
Albany, GA
St. Cloud, MN
Manchester-Nashua, NH
Raleigh-Cary, NC
Springfield, MA
Bethesda-Rockville-Frederick, MD
Milwaukee-Waukesha-West Allis, WI
New Haven-Milford, CT
Minneapolis-St. Paul-Bloomington, MN-WI
Santa Cruz-Watsonville, CA
Trenton-Ewing, NJ
Madison, WI
Logan, UT-ID
Lake County-Kenosha County, IL-WI
Philadelphia, PA
San Jose-Sunnyvale-Santa Clara, CA
Durham-Chapel Hill, NC
Boulder, CO
San Francisco-San Mateo-Redwood City, CA
Bridgeport-Stamford-Norwalk, CT
Hartford-West Hartford-East Hartford, CT
95% Confidence Interval
0.0
0.5
1.0
1.5
Estimated White-Hispanic Achievement Gap

Figure 6



Achievement Gaps and Racial Disparities in Family Socioeconomic Status,
All School Districts with Estimated Gaps, 2009-2013
White-Black
White-Black Achievement Gap (EB estimate)
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.20
0.00
0.20
0.40
0.60
0.80
Racial SES Disparity
$r^2 = 0.41$
White-Hispanic
White-Hispanic Achievement Gap (EB estimate)
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.20
0.00
0.20
0.40
0.60
0.80
Racial SES Disparity
$r^2 = 0.40$

Figure 7



Achievement Gaps and Racial Disparities in Exposure to Poverty,
All School Districts with Estimated Gaps, 2009-2013
White-Black
White-Black Achievement Gap (EB estimate)
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
0.00
0.10
0.20
0.30
0.40
0.50
0.60
Black-White Disparity in Exposure to Poverty
$r^2 = 0.15$
White-Hispanic
White-Hispanic Achievement Gap (EB estimate)
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
0.00
0.10
0.20
0.30
0.40
0.50
0.60
Hispanic-White Disparity in Exposure to Poverty
$r^2 = 0.15$

Figure 8

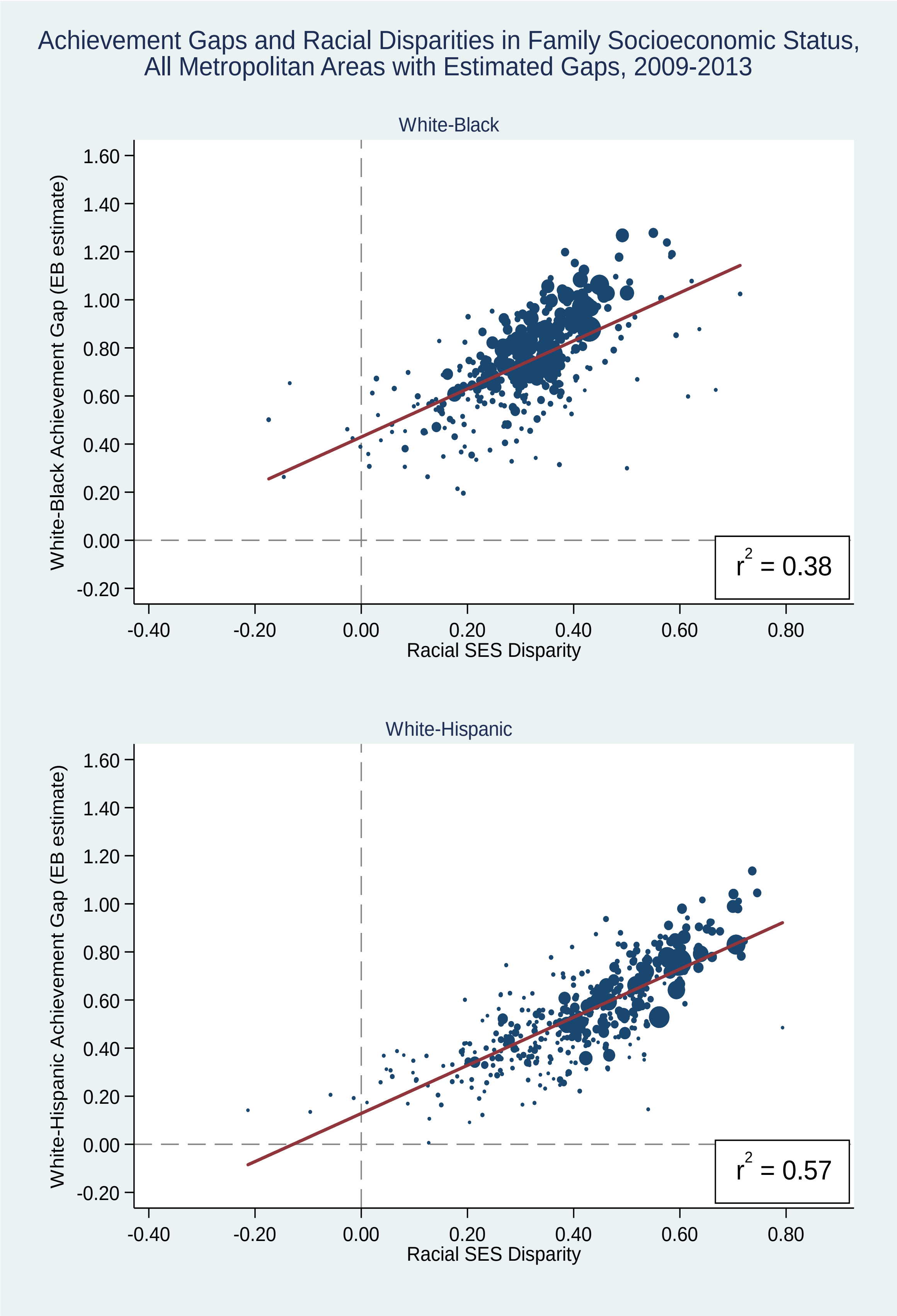

Achievement Gaps and Racial Disparities in Family Socioeconomic Status,
All Metropolitan Areas with Estimated Gaps, 2009-2013
White-Black
White-Black Achievement Gap (EB estimate)
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.20
0.00
0.20
0.40
0.60
0.80
Racial SES Disparity
r2 = 0.38
White-Hispanic
White-Hispanic Achievement Gap (EB estimate)
r2 = 0.57

Figure 9

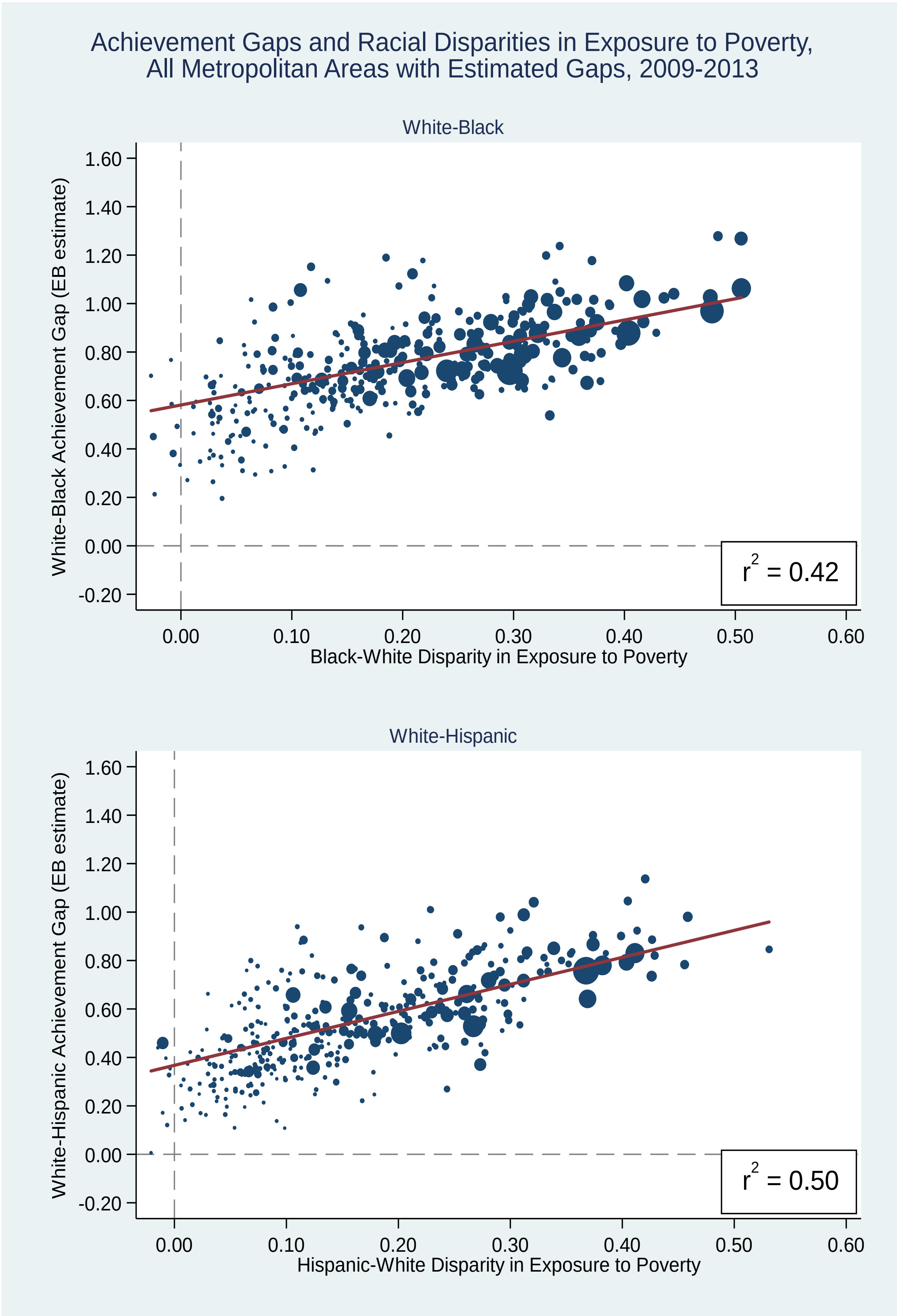

Achievement Gaps and Racial Disparities in Exposure to Poverty,
All Metropolitan Areas with Estimated Gaps, 2009-2013
White-Black
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
White-Black Achievement Gap (EB estimate)
r² = 0.42
0.00
0.10
0.20
0.30
0.40
0.50
0.60
Black-White Disparity in Exposure to Poverty
White-Hispanic
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
White-Hispanic Achievement Gap (EB estimate)
r² = 0.50
0.00
0.10
0.20
0.30
0.40
0.50
0.60
Hispanic-White Disparity in Exposure to Poverty

Table 1. Achievement Gap Data Population Coverage by Subject, Group and Geographic Unit

	White-Black Gaps			White-Hispanic Gaps		
	ELA	Math	Pooled	ELA	Math	Pooled
Number of geographic units with available achievement gap estimates in at least one grade-x-year						
Metropolitan Areas	378	378	378	377	377	377
School Districts	2,878	2,854	2,899	3,632	3,642	3,689
Mean number of grade-x-year cells with achievement gap estimates, among those with at least 1 (maximum=30 by subjects, 60 for pooled)						
Metropolitan Areas	27.4	26.6	54.0	27.7	26.8	54.5
School Districts	22.2	21.8	43.7	20.7	19.8	40.1
Proportion of geographic units with achievement gap data available for 20 (40 for pooled) or more grade-x-year cells, among those with at least 1						
Metropolitan Areas	0.93	0.92	0.93	0.94	0.93	0.93
School Districts	0.71	0.70	0.70	0.64	0.63	0.63
Proportion of minority students in metropolitan area analytic sample, among minority students in metropolitan area public schools						
2009	0.96	0.96	0.96	0.98	0.98	0.98
2010	0.96	0.96	0.96	0.98	0.98	0.98
2011	0.96	0.96	0.96	0.98	0.98	0.98
2012	0.96	0.96	0.96	0.98	0.98	0.98
2013	0.96	0.96	0.96	0.98	0.98	0.98
Proportion of minority students in school district analytic sample, among minority students in all public schools						
2009	0.93	0.93	0.93	0.93	0.92	0.93
2010	0.93	0.93	0.93	0.92	0.92	0.93
2011	0.93	0.93	0.93	0.93	0.92	0.93
2012	0.93	0.93	0.93	0.93	0.92	0.93
2013	0.93	0.93	0.93	0.92	0.92	0.93

Table 2. Comparison of Metro/District Gaps in Math and Reading

Panel A: Subject-Specific Random Intercepts

	Metropolitan Areas				School Districts			
	White-Black		White-Hispanic		White-Black		White-Hispanic	
	Math	Reading	Math	Reading	Math	Reading	Math	Reading
Average Gap	0.773 ***	0.730 ***	0.524 ***	0.580 ***	0.682 ***	0.635 ***	0.475 ***	0.532 ***
	(0.011)	(0.010)	(0.011)	(0.011)	(0.004)	(0.004)	(0.004)	(0.004)
Average Grade Trend	0.014 ***	0.015 ***	-0.007 ***	-0.011 ***	-0.004 ***	0.005 ***	-0.014 ***	-0.015 ***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Average Cohort Trend	0.008 ***	0.012 ***	-0.008 ***	-0.006 ***	0.001	0.007 ***	-0.009 ***	-0.008 ***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Within-Unit Standard Deviation	0.064	0.052	0.057	0.051	0.083	0.071	0.076	0.070
Between-Unit Standard Deviation	0.212	0.201	0.207	0.217	0.223	0.229	0.224	0.233
Reliability	0.963	0.963	0.976	0.979	0.876	0.883	0.885	0.891
Intraclass Correlation	0.915	0.938	0.930	0.948	0.877	0.911	0.898	0.918
p-value: Math Intercept = Reading Intercept	0.000		0.000		0.000		0.000	
Correlation: Math Intercept, Reading Intercept	0.942		0.910		0.950		0.928	

Panel B: Pooled across Subjects

	Pooled Subjects	Pooled Subjects	Pooled Subjects	Pooled Subjects
Average Gap	0.752***	0.551***	0.658***	0.503***
	(0.011)	(0.011)	(0.004)	(0.004)
Average Grade Trend	0.015***	-0.009***	0.001	-0.014***
	(0.002)	(0.002)	(0.001)	(0.001)
Average Cohort Trend	0.010***	-0.007***	0.004***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
Within-Unit Standard Deviation	0.056	0.052	0.056	0.051
Between-Unit Standard Deviation	0.203	0.207	0.224	0.225
Reliability	0.972	0.987	0.914	0.919
Intraclass Correlation	0.929	0.941	0.942	0.952
p-value: Math Intercept = Reading Intercept	0.000	0.000	0.000	0.000
Total Observations	20402	20544	125380	146494
Metro/District Observations	378	377	2899	3689

***p<=.001; **p<=.01; *p<=.05

* p-value: Math Intercept = Reading Intercept is a test of equality among the reading and math intercepts in Panel A and the reported p-value fo the Math coefficient from Panel B.

Table 3. Component Loadings for Socioeconomic Status Composite

	Standardized Component Loadings	Correlations between Indicators and Composite	Unstandardized Component Loadings	Mean	SD
Median Income	NA	NA	NA	$62,509	$26,565
Log of Median Income	0.22	0.96	0.40	15.82	0.56
Proportion of Adults, Aged 25+ with a Bachelor's Degree or Higher	0.16	0.67	1.08	0.29	0.14
Poverty Rate, Households with 5-17 Year Olds	-0.22	-0.94	-2.21	0.16	0.10
Unemployment Rate	-0.16	-0.68	-8.97	0.05	0.02
Proportion of Households Receiving Food Stamps or SNAP	-0.22	-0.93	-2.01	0.16	0.11
Proportion Single Mother Headed Households	-0.19	-0.83	-1.75	0.27	0.11

Based on 11,582 districts in ACS with non-missing data on all 6 measures and avergae per grade enrollment. All values are weighted by district enrollment. The log (in base 2) of median income is used in the construction of the SES composite, but the mean and standard deviation of median income are also shown above for ease of interpretation. Unstandardized component loadings are coefficients from a model that regresses the (standardized) SES composite on the (unstandardized) six variables used to construct the composite; the intercept from this model is -5.06. These coefficients and intercept are used to construct race-specific SES composites.

Table 4. Means and Standard Deviations of Local Characteristics

	White-Black				White-Hispanic			
	Metros		Districts		Metros		Districts	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Average Achievement Gap	0.732	0.195	0.658	0.232	0.535	0.202	0.518	0.225
Socioeconomic Composition								
SES Composite	-0.141	0.971	-0.172	0.996	-0.127	0.970	0.095	0.930
Median Income (in $100,000)	0.579	0.123	0.590	0.239	0.579	0.122	0.642	0.266
Proportion of Adults, Aged 25+ with a Bachelor's Degree or Higher	0.275	0.086	0.268	0.145	0.276	0.087	0.276	0.149
Proportion Receiving Free Lunches in Public Schools	0.419	0.112	0.465	0.200	0.419	0.112	0.424	0.199
Single Mother Headed Household Rate	0.272	0.056	0.291	0.106	0.271	0.056	0.246	0.097
Racial/Ethnic Composition								
Proportion Black in Public Schools	0.158	0.143	0.224	0.198	0.153	0.144	0.114	0.152
Proportion Hispanic in Public Schools	0.174	0.193	0.187	0.202	0.180	0.198	0.279	0.227
Proportion of Hispanics Who Report Speaking English Well or Very Well	0.864	0.074	0.864	0.144	0.862	0.074	0.857	0.112
Racial Socioeconomic Disparities								
White-Minority SES Composite Gap	4.267	1.707	1.968	1.238	2.710	1.290	1.136	0.949
White-Minority Income Gap	0.766	0.345	0.649	0.461	0.704	0.295	0.633	0.442
White-Minority Education Gap	0.244	0.222	0.177	0.354	0.677	0.323	0.726	0.406
Minority-White Single Mother Headed Household Rate Difference	0.288	0.144	0.272	0.173	0.069	0.114	0.037	0.157
Segregation								
Between School Racial Segregation	0.259	0.149	0.075	0.097	0.179	0.113	0.053	0.072
Between School Free Lunch/Not Free Lunch Segregation	0.154	0.077	0.055	0.061	0.152	0.077	0.047	0.058
Between Tract Racial Segregation	0.271	0.118			0.165	0.081		
Between Tract Poor-Non-Poor Segregation	0.106	0.039			0.105	0.039		
Minority-White Tract Poverty Rate Difference	0.086	0.049			0.056	0.038		
Minority-White School Free Lunch Rate Difference	0.188	0.114	0.051	0.071	0.159	0.106	0.045	0.066
School Characteristics								
Per Pupil Instructional Expenditures in Average Student's School (in $10,000)	0.602	0.155	0.633	0.210	0.595	0.154	0.635	0.236
Average Student-Teacher Ratio	16.414	2.794	16.835	20.259	16.544	2.826	16.949	4.562
Proportion Attending Charter Schools	0.029	0.037	0.029	0.071	0.029	0.038	0.027	0.068
White/Minority Per Pupil Instructional Expenditures Ratio	0.964	0.061			0.975	0.046		
Minority/White Student-Teacher Ratio Ratio	1.038	0.055	1.016	0.033	1.029	0.052	1.012	0.038
Minority-White Charter School Enrollment Rate Difference	0.019	0.053	0.002	0.066	0.000	0.035	-0.008	0.046
Average Achievement								
Average Pooled Math/ELA Achievement, National Standardization	0.000	0.194	-0.040	0.321	0.002	0.195	0.001	0.325
Sample Size	360		2,341		370		3,038	

Note: SD = standard deviation

Table 5. Pairwise Correlations Between Average Achievement Gaps and Local Characteristics

	White-Black Gaps		White-Hispanic Gaps	
	Metros	Districts	Metros	Districts
Socioeconomic Composition				
SES Composite	0.255 ***	0.268 ***	0.465 ***	0.260 ***
Median Income (in $100,000)	0.327 ***	0.275 ***	0.510 ***	0.219 ***
Proportion of Adults, Aged 25+ with a Bachelor's Degree or Higher	0.469 ***	0.587 ***	0.539 ***	0.486 ***
Proportion Receiving Free Lunches in Public Schools	-0.230 ***	-0.269 ***	-0.334 ***	-0.180 ***
Single Mother Headed Household Rate	0.059	-0.045 *	-0.307 ***	-0.092 ***
Racial/Ethnic Composition				
Proportion Black in Public Schools	0.242 ***	-0.020	-0.174 ***	-0.019
Proportion Hispanic in Public Schools	-0.122 **	-0.078 ***	0.235 ***	0.047 **
Proportion of Hispanics Who Report Speaking English Well or Very Well	-0.146 **	-0.054 **	-0.309 ***	-0.193 ***
Racial Socioeconomic Disparities				
White-Minority Gap in SES Composite	0.519 ***	0.385 ***	0.510 ***	0.352 ***
White-Minority Income Gap	0.561 ***	0.442 ***	0.622 ***	0.465 ***
White-Minority Education Gap	0.402 ***	0.583 ***	0.565 ***	0.466 ***
Minority-White Single Mother Headed Household Rate Difference	0.379 ***	0.259 ***	0.271 ***	0.137 ***
Segregation				
Between School Racial Segregation	0.530 ***	0.212 ***	0.644 ***	0.372 ***
Between School Free Lunch/Not Free Lunch Segregation	0.502 ***	0.293 ***	0.526 ***	0.324 ***
Between Tract Racial Segregation	0.443 ***		0.542 ***	
Between Tract Poor-Non-Poor Segregation	0.523 ***		0.368 ***	
Minority-White Tract Poverty Rate Difference	0.413 ***		0.421 ***	
Minority-White School Free Lunch Rate Difference	0.650 ***	0.380 ***	0.708 ***	0.419 ***
School Characteristics				
Per Pupil Instructional Expenditures in Average Student's School (in $10,000)	0.156 **	0.279 ***	0.241 ***	0.283 ***
Average Student-Teacher Ratio	-0.130 **	-0.012	0.075	-0.088 ***
Proportion Attending Charter Schools	0.104 *	-0.012	0.220 ***	0.072 ***
White/Minority Per Pupil Instructional Expenditures Ratio	-0.189 ***		-0.025	
Minority/White Student-Teacher Ratio Ratio	0.201 ***	0.215 ***	-0.012	0.212 ***
Minority-White Charter School Enrollment Rate Difference	0.372 ***	-0.018	-0.045	-0.055 **
Average Achievement				
Average Pooled Math/ELA Achievement, National Standardization	0.287 ***	0.330 ***	0.231 ***	0.238 ***
Sample Size	360	2,341	370	3,038

Notes: ***p<=.001; **p<=.01; *p<=.05 The correlation estimates for districts are based on models that include state fixed effects; they are interpreted as pooled within-state correlations. The estimated correlations for metros are simple bivariate correlations because there are not sufficient metropolitan areas within each state to support including the state fixed effects.

Table 6. Fitted Multivatiate Regression Model Parameter Estimates

	White-Black Gaps								White-Hispanic Gaps							
	Metros				Districts				Metros				Districts			
SES Composite	0.086	***	0.089	***	0.110	***	0.097	***	0.092	***	0.091	***	0.112	***	0.104	***
	(0.008)		(0.008)		(0.005)		(0.005)		(0.008)		(0.008)		(0.005)		(0.005)	
White-Minority Gap in SES Composite	0.040	***	0.038	***	0.075	***	0.070	***	0.035	***	0.039	***	0.068	***	0.064	***
	(0.004)		(0.004)		(0.003)		(0.003)		(0.005)		(0.005)		(0.004)		(0.003)	
Minority-White School Free Lunch Rate	0.764	***	0.748	***	0.841	***	0.876	***	0.866	***	0.869	***	1.044	***	1.013	***
Difference	(0.065)		(0.079)		(0.051)		(0.061)		(0.069)		(0.074)		(0.051)		(0.057)	
Proportion Black in Public Schools	0.323	***	0.295	***	0.199	***	0.134	***	0.079		0.086		0.140	***	0.092	**
	(0.061)		(0.061)		(0.025)		(0.026)		(0.055)		(0.056)		(0.029)		(0.029)	
Proportion Hispanic in Public Schools	0.148	***	0.167	***	0.057	*	-0.001		0.147	***	0.105	*	0.192	***	0.169	***
	(0.042)		(0.046)		(0.027)		(0.027)		(0.039)		(0.042)		(0.021)		(0.021)	
Proportion of Hispanics Who Report	-0.165		-0.178		-0.059	*	-0.056	*	-0.384	***	-0.363	***	-0.270	***	-0.246	***
Speaking English Well or Very Well	(0.094)		(0.094)		(0.025)		(0.024)		(0.087)		(0.090)		(0.031)		(0.030)	
Per Pupil Instructional Expenditures in			-0.106	*			0.388	***			-0.027				0.400	***
Average Student's School (in $10,000)			(0.050)				(0.038)				(0.047)				(0.034)	
Average Student-Teacher Ratio			-0.005				0.000				0.003				-0.004	***
			(0.003)				(0.000)				(0.003)				(0.001)	
Proportion Attending Charter Schools			0.354				-0.061				0.411	*			0.105	*
			(0.189)				(0.056)				(0.169)				(0.052)	
White/Minority Per Pupil Instructional			0.129								0.092					
Expenditures Ratio			(0.119)								(0.144)					
Minority/White Student-Teacher Ratio			0.453	***			-0.034				0.172				0.121	
Ratio			(0.124)				(0.122)				(0.133)				(0.095)	
Minority-White Charter School			0.180				0.006				-0.098				0.048	
Enrollment Rate Difference			(0.149)				(0.053)				(0.197)				(0.071)	
R^2	0.629		0.647		0.412		0.439		0.709		0.717		0.399		0.441	
Sample Size	360				2,341				370				3,038			

Notes: ***p<=.001; **p<=.01; *p<=.05. The district models include state fixed effects; the metropolitan area models do not.

Appendix A1: Estimating $\boldsymbol{V}$-statistic using Coarsened Proficiency Data

Achievement Gap Measure

Every state uses different standardized tests; within a state, these tests vary across subjects, grades, and often across years. Moreover, the ED*Facts* data do not include group-specific means and standard deviations, but instead include counts of students in a set of ordered proficiency categories whose definitions vary across states, grades, subjects, and sometimes years. Because these definitions vary, simple racial differences in proficiency rates do not provide measures of achievement gaps that are comparable across states, subjects, grades, or years (Ho 2008; Ho and Reardon 2012). Nonetheless, counts of students scoring in different proficiency categories can be used to estimate achievement gaps interpretable as effect sizes (Ho and Reardon 2012; Reardon and Ho 2015), as described below.

The most conventional measure of achievement gaps is the standardized mean difference in test scores between two groups, defined as

$$d = \frac{\mu_a - \mu_b}{\sigma_p}, \tag{A1}$$

where μ_a and μ_b are the mean test scores in groups a and b, respectively, and σ_p is the pooled standard deviation of test scores (the square root of the average of the test score variances in groups a and b):

$$\sigma_p = \sqrt{\frac{\sigma_a^2 + \sigma_b^2}{2}}. \tag{A2}$$

This measure, sometimes called Cohen's d (Hedges and Olkin 1985), is a measure of the relative difference in the test score distributions of two groups. It is relative in the sense that it measures gaps as the ratio of the difference in means to the average spread of the two distributions. It can be thought of, loosely, as a measure of the extent to which the distribution of scores in group a is higher than the distribution in group b.

Two factors complicate the use of d given the data and objectives of our analysis. First, computing d requires estimates of the mean and standard deviation of each district's test score distributions, by race; these statistics are not generally publicly available except from a few state websites in select years. Second, d is sensitive to the scale in which test scores are reported. Although d would be unchanged by any linear transformation of test scores (such a transformation would multiply both the difference in means and the pooled standard deviation by the same factor, leaving their ratio unchanged), it will be altered by a non-linear transformation of scores. Unless the metric in which achievement is measured is inherently meaningful, then, d is sensitive to arbitrary scaling decisions. In order to compare test score gaps across states, grades, subjects, and years in which different tests are used, it is necessary to use a gap measure that is not sensitive to differences in how test scores are scaled.

An alternate measure of the relative difference in distributions, one that is immune to scale transformations of the test score metric, is based on the probability that a randomly chosen observation from distribution a has a higher value than a randomly chosen observation from distribution b. Like d, this measure, denoted $P_{(a>b)}$, can be loosely thought of as a measure of the non-overlap of distributions a and b, or as a measure of the extent to which distribution a contains higher values than distribution b. The value of $P_{(a>b)}$ may range from 0 to 1, with values greater than 0.5 indicating that distribution a is higher than b, and vice versa. Applying a probit transformation to $P_{(a>b)}$ produces the V-statistic (Ho 2009; Ho and Haertel 2006; Ho and Reardon 2012):

$$V = \sqrt{2}\Phi^{-1}\left(P_{(a>b)}\right) \tag{A3}$$

The V-statistic has three useful properties for our purposes. First, it is readily interpretable as an effect size. Essentially, Equation (A3) converts $P_{(a>b)}$ to an effect size by computing the standardized difference between two normal distributions that would yield the observed value of $P_{(a>b)}$. As a result, if the test score distributions of groups a and b are both normal (regardless of whether they have equal variance), then V will be equal to Cohen's d (Ho and Reardon 2012). Thus, V can be thought of as measuring gaps in a familiar "effect size" metric.

Second, V is invariant to monotonic transformations of test scales: if a test metric is transformed by any non-linear monotonic transformation, Cohen's d will be changed, but V will not. Thus, V can be understood as the value of Cohen's d if the test score metric were transformed into a metric in which both groups' scores were normally distributed. This transformation-invariance property of V is particularly useful when comparing gaps measured using different tests. In order to compare gaps across tests using Cohen's d, we would have to assume that each test measures academic achievement in an interval-scaled metric (so that a score on any test can be written as a linear transformation of a score on any other test). To compare gaps using V, however, we need only to assume that each test measures achievement in a way that orders two groups the same way (so that the overlap between two groups' distributions would be the same in either test), a much more defensible assumption.[12]

A final advantage of the V-statistic is that it can be estimated very reliably either from student-level continuous test score data or from coarsened data indicating the number of students of each group in each of several (at least three) proficiency categories (Ho and Reardon 2012; Reardon and Ho 2015). That is, it is not necessary to know the means and standard deviations of each group's test score distribution; all that is needed are the counts of black, Hispanic, and white students who score "Far Below Basic," "Below Basic," "Basic," "Proficiency," and "Advanced," for example. This is the form of the achievement data available from ED*Facts*. Because V is estimable with such little ordinal information, it is

[12] In a set of validation exercises shown in Appendix A2 (with results in Table A1), state-level achievement gaps estimated from state accountability tests (which differ among states) are compared to achievement gaps estimated for those same states, grades, and years from the National Assessment of Educational Progress (NAEP), which is identical across states, within a grade, year, and subject. The correlation between the gap estimates from the two data sources is above 0.90 in most cases (ranging between 0.85 and 0.97 depending on year, grade and subject), indicating that different state tests order students similarly enough that the V-statistic can be used to compare achievement gaps across a wide range of state and NAEP tests.

possible to easily estimate achievement gaps based on state accountability tests in each district/metropolitan area-year-grade-subject for which subgroup-specific proficiency category counts are available.

We estimate V-gaps and their standard errors using the maximum likelihood (ML) algorithm described by Ho and Reardon (2012) and Reardon and Ho (2015) for each district/metropolitan area-grade-subject-year cell in which there are at least 20 white students and 20 black or Hispanic students tested. Ho and Reardon (2012) demonstrate that the ML estimator is unbiased under the assumption of respective normality, and is very nearly unbiased even under large departures from respective normality. Finally, in order to correct estimated achievement gaps and standard errors for measurement error, we disattenuate each of the estimated gaps and their standard errors by dividing both by the square root of the reliability of the test used.[13] The reliabilities of most state tests are about 0.90 (Reardon and Ho 2015), so the disattenuated gaps are generally about 5% larger than the unadjusted estimates.

For both white-black and white-Hispanic gaps, there are up to 30 estimated math and ELA achievement gap estimates (for each of six grades and 5 school years) in each geographic unit (school district or metropolitan area). We use these (up to) 60 gap estimates per unit in Models 1 and 2.

[13] Reliabilities were collected from state department of education websites.

Appendix A2: Using ED*Facts* Data to Estimate Achievement Gaps, NAEP Comparison

To assess whether state accountability test data from ED*Facts* (which is based on different tests in each state and grade) accurately describes achievement gaps, we compare estimates of the state-level V-statistics computed from ED*Facts* and from the National Assessment of Educational Progress (NAEP). We use two primary data sources to estimate state-level achievement gaps: NAEP[14] and state assessments. We use NAEP 4th and 8th grade math and reading test score data from 2009, 2011 and 2013, and categorical proficiency data (e.g., percentages of students scoring "Below Basic," "Basic," "Proficient," and "Advanced") from state-administered standardized math and reading tests compiled by ED*Facts*. Though ED*Facts* data are available for grades 3-8, we use data for grades 4 and 8 in order to match what is available from NAEP.

Across all grades, years, subjects and data source, the sample includes 1,062 white-black achievement gaps and 1,090 white-Hispanic achievement gaps. From these data, we compute state-level achievement gaps (V) for state assessments and the NAEP for white-black and white-Hispanic gaps in each state-year-grade-subject combination for which we have NAEP and state test data. In the case of Ed Facts, the V-statistic is estimated from coarsened proficiency data (see the section of the text titled *Achievement Gap Measure* for description); for the NAEP, the V-statistic can be calculated from the complete cumulative distribution function using student-level data.

We wish to test the extent to which V-gaps estimated from state assessment data depart from V-gaps estimated from NAEP. To do this, we estimate 24 precision-weighted random coefficients models, in which each model corresponds to a grade (4, 8), year (2009, 2011, 2013), subject (math, ELA), and gap (white-black, white-Hispanic). For each iteration, we include only observations for which we were able to estimate a gap in both the NAEP and ED*Facts*. Depending on the year, grade, subject and gap groups, the

[14] We use "State NAEP" data, based on math and reading assessments administered to representative samples of fourth- and eighth-graders roughly every two years in each of the 50 states. State NAEP sample sizes are roughly 2,500 students, from approximately 100 schools, in each state-grade-subject.

number of overlapping observations ranged between 70 to 98. The model takes the form:

$$\hat{G}_u = [\gamma_{n0} + v_{mu}]N_s + [\gamma_{e0} + v_{eu}]E_s + \epsilon_u$$
$$\epsilon_u \sim N[0, \hat{\omega}_u^2]$$
$$\begin{bmatrix} v_{nu} \\ v_{eu} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{nn} & \tau_{ne} \\ \tau_{ne} & \tau_{ee} \end{pmatrix}\right]$$

(A4)

Here, $\hat{G}_u$ is the estimated achievement gap for state u; its estimated standard error is $\hat{\omega}_u$. The variables N_s and E_s are dummy variables indicating, respectively, whether $\hat{G}_u$ describes a NAEP or ED*Facts* gap. The model is estimated separately for each grade, year and subject. The error terms in the model indicate that estimated gaps may vary in two ways. First, estimated gaps may differ from their true values because of sampling variance; this is indicated by the error term ϵ_u, which is assumed to be normally distributed with a known variance equal to $\hat{\omega}_u^2 = var(\hat{G}_u)$. Second, unit-specific average NAEP and ED*Facts* gaps may deviate from the mean NAEP and ED*Facts* gaps among states (which are denoted by γ_{n0} and γ_{e0}, respectively). We allow these deviations (v_{nu} and v_{eu}) to differ for NAEP and ED*Facts*; they are assumed multivariate normal with a variance matrix $\tau = \begin{pmatrix} \tau_{nn} & \tau_{ne} \\ \tau_{ne} & \tau_{ee} \end{pmatrix}$ that must be estimated.

The parameters of interest are the NAEP and ED*Facts* intercepts, γ_{n0} and γ_{e0}, respectively and the correlations between v_{nu} and v_{eu} ($\rho_{ne} = \frac{\tau_{ne}}{\sqrt{\tau_{nn} \cdot \tau_{ee}}}$). The intercept terms correspond to the precision-weighted average gap according to the NAEP and ED*Facts* in the corresponding grade, year and subject. We also construct NAEP/ED*Facts* average gap ratio ($\hat{\gamma}_{n0}/\hat{\gamma}_{e0}$); ratio values greater than one indicate the estimated NAEP gap is greater than the estimated ED*Facts* gap. The correlation of the NAEP and ED*Facts* state-specific random effects (ρ_{ne}) indicates how consistently states are ranked according to the different measures.

Estimates from model (A4) are shown in Table A1. Table A1 shows the precision-weighted NAEP and ED*Facts* intercepts for years 2009, 2011 and 2013; grades 4 and 8; subjects Math and ELA; and gap-groups White-Black and White-Hispanic. NAEP gaps are consistently larger than those from Ed Facts, on average, in math (the NAEP/ED*Facts* ratio is greater than one in all cases for math gaps, ranging between 1.11 and 1.27) and are closer, on average, to those from Ed Facts in ELA (the NAEP/ED*Facts* ratio ranges between 0.91 and 1.08). In most cases, the null hypothesis that $\gamma_{n0} = \gamma_{e0}$ is rejected (always in math, some of the time in ELA). Despite these differences, correlation coefficients on the NAEP and ED*Facts* random effects are very high, ranging between 0.85 to 0.97. These high correlations indicate that gaps estimated from state tests correspond well to gaps estimated from the NAEP, despite the fact that that states use different tests. Thus, comparing gaps based on different state tests appears largely valid.

[Table A1 here]

Appendix Table A1. Comparing State-Level ED*Facts* Gaps with NAEP Gaps for Overlapping Observations

Gap	Year	Grade	Subject	ED*Facts* Intercept	NAEP Intercept	NAEP/ ED*Facts* Ratio	NAEP- ED*Facts* Correlation	p-value from ttest that NAEP= ED*Facts*	N
White-Black	2009	4	Math	0.75	0.95	1.27	0.87	0.00	94
White-Black	2009	4	ELA	0.72	0.75	1.05	0.89	0.08	92
White-Black	2009	8	Math	0.81	0.92	1.14	0.94	0.00	98
White-Black	2009	8	ELA	0.71	0.76	1.08	0.96	0.00	98
White-Black	2011	4	Math	0.77	0.95	1.22	0.92	0.00	82
White-Black	2011	4	ELA	0.73	0.76	1.04	0.94	0.11	74
White-Black	2011	8	Math	0.81	0.92	1.14	0.87	0.00	94
White-Black	2011	8	ELA	0.74	0.76	1.03	0.91	0.25	90
White-Black	2013	4	Math	0.79	0.94	1.19	0.95	0.00	98
White-Black	2013	4	ELA	0.77	0.76	0.99	0.92	0.60	96
White-Black	2013	8	Math	0.82	0.92	1.11	0.89	0.00	76
White-Black	2013	8	ELA	0.77	0.76	0.99	0.93	0.80	70
White-Hispanic	2009	4	Math	0.59	0.73	1.24	0.87	0.00	96
White-Hispanic	2009	4	ELA	0.65	0.65	1.00	0.96	0.93	92
White-Hispanic	2009	8	Math	0.63	0.71	1.12	0.91	0.00	98
White-Hispanic	2009	8	ELA	0.66	0.60	0.92	0.85	0.02	98
White-Hispanic	2011	4	Math	0.55	0.70	1.27	0.89	0.00	86
White-Hispanic	2011	4	ELA	0.60	0.65	1.07	0.97	0.00	80
White-Hispanic	2011	8	Math	0.58	0.69	1.20	0.97	0.00	94
White-Hispanic	2011	8	ELA	0.61	0.58	0.96	0.93	0.16	92
White-Hispanic	2013	4	Math	0.54	0.65	1.20	0.93	0.00	98
White-Hispanic	2013	4	ELA	0.64	0.62	0.98	0.94	0.37	96
White-Hispanic	2013	8	Math	0.54	0.61	1.12	0.90	0.00	84
White-Hispanic	2013	8	ELA	0.60	0.54	0.91	0.91	0.00	76
Averages									
Math Average	All	All	Math			1.18	0.91		
ELA Average	All	All	ELA			1.00	0.93		
Grade 4 Average	All	4	All			1.13	0.92		
Grade 8 Average	All	8	All			1.06	0.91		
Math Grade 4 Average	All	4	Math			1.23	0.90		
Math Grade 8 Average	All	8	Math			1.14	0.91		
ELA Grade 4 Average	All	4	ELA			1.02	0.94		
ELA Grade 8 Average	All	8	ELA			0.98	0.91		