Title

The Long Road to Equality: A Meta-Regression Analysis of Changes in the Black Test Score Gap Over Time*

Journal

Social Science Quarterly, 99(3)

ISSN

0038-4941

Authors

Huntington-Klein, Nick
Ackert, Elizabeth

Publication Date

2018-09-01

DOI

10.1111/ssqu.12483

Peer reviewed

HHS Public Access

Author manuscript

Soc Sci Q. Author manuscript; available in PMC 2019 September 01.

Published in final edited form as:

Soc Sci Q. 2018 September ; 99(3): 1119–1133. doi:10.1111/ssqu.12483.

The Long Road to Equality: A Meta-Regression Analysis of Changes in the Black Test Score Gap over Time*

Nick Huntington-Klein and
California State University, Fullerton

Elizabeth Ackert
The University of Texas at Austin

Abstract

Objective—We analyze changes in test score gaps between Black students and their peers from 1979 to 2010 and examine how observable factors contribute to the gap.

Methods—Using meta-regression, we examine the relationship between African American racial status and achievement in United States K-12 education in 165 published studies.

Results—The absolute relationship between Black status and achievement decreased during the 1980s and early 1990s, but was stagnant from the late 1990s through 2010. Socioeconomic status explained more than half of the gap, and the influence of socioeconomic status on the gap did not changed significantly over time. Schooling characteristics explained relatively little of the gap, but school-level factors increased in importance over time.

Conclusions—Black test score gap closure stagnated in an era when federal education policy sought to close racial achievement gaps. Observable factors explain more of the gap than has been previously recognized.

Keywords

Achievement gap; African American Students; Meta-analysis; Education Policy

Introduction

Despite policy efforts to weaken the influence of ascriptive background characteristics on student achievement, race continues to be a defining factor shaping patterns of educational inequality among students in U.S. schools. One expression of the influence of race is the Black-White test score gap, which remains sizeable (Hedges & Nowell, 1999; Jencks &

*The authors will share all data and coding for replication purposes. Appendices and supplementary material can be found at http://nickchk.com/Huntington-Klein_Ackert_SSQ_Appendices.pdf. Research funds and support for this project have been provided by Collaborative Researchers for Education Sciences Training, Institute of Education Sciences Grant R305B090012, the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) grant, R24 HD042828, and the Shanahan Endowment Fellowship and a Eunice Kennedy Shriver NICHD training grant, T32 HD007543. We thank Pasita Chaijaroen and Andre Stephens for their research assistance and Ashley Jochim, Sean Reardon, Stew Tolnay, William Zumeta, anonymous reviewers, and session participants at the Association for Education Finance and Policy conference for comments.

Direct all correspondence to Nick Huntington-Klein, nhuntington-klein@fullerton.edu, Department of Economics, California State University, Fullerton, 800 N State College Blvd., Fullerton, CA, 92831.

Phillips, 1998; Magnuson & Waldfogel, 2008; Reardon, Robinson-Cimpian, & Weathers, 2014; Reardon, Valentino, Kalogrides, Shores, & Greenberg, 2013). Our study characterizes changes in the Black test score gap over time from the period of 1979 to 2010 using meta-regression.

Previous work shows an overall shrinking trend in the Black test score gap from the mid-1960s through 2010 (Berends & Peñaloza, 2008; Grissmer, Flanagan, & Williamson, 1998; Hedges & Nowell, 1999; Magnuson & Waldfogel, 2008; Reardon et al., 2014, 2013), but questions remain about the rate of gap closure and the existence of periods of stall or widening in the gap, especially into the accountability era of the 21st century under No Child Left Behind (NCLB). Some studies show a pattern of what Magnuson and Waldfogel (2008) refer to as "steady gains and stalled progress," with a period of gap closure in the 1970s to 1980s and stagnation thereafter, while others show a slow but steady decline in the gap from the mid-1990s through the 2000s (Reardon et al., 2014, 2013). Our study provides a look into time trends in the test score gap over a wide body of samples, students, and assessments.

In addition to the overall size of the gap, scholars have debated the roles that families, schools, public policy, and even genetics play in shaping Black achievement disparities with peers (Herrnstein & Murray, 1994; Jencks & Phillips, 1998). Prior research has provided estimates of the contribution of socioeconomic factors to the gap (Hedges & Nowell, 1999; Herrnstein & Murray, 1994; Phillips, Brooks-Gunn, Duncan, Klebanov, & Crane, 1998; Quinn, 2015; Rothstein & Wozny, 2013) and has decomposed time trends in the gap into changes in household versus schooling characteristics (Berends & Peñaloza, 2008). Our research takes a different approach to determining which observable factors most influence the gap. We use meta-regression to calculate how the estimated gap grows or shrinks in the presence of observable control variables in published studies, and we also analyze whether these observable confounding factors have a varying or constant influence on the gap over time.

Background

A major question in the literature on educational stratification is whether black students are closing achievement gaps with white students over time. Previous studies of time trends in the Black-White test score gap show that the Black test score gap with peers has declined relative to the 1960s (Berends & Peñaloza, 2008; Grissmer et al., 1998; Hedges & Nowell, 1999; Herrnstein & Murray, 1994; Jencks & Phillips, 1998; Magnuson & Waldfogel, 2008; Reardon et al., 2014, 2013). There is disagreement, however, regarding the rate of gap closure and whether there have been periods of stall or widening in the gap. Studies using the National Assessment of Educational Progress (NAEP) demonstrate that Black-White test score gaps have been decreasing slowly over time, but have detected a period of stagnation or even widening in the gap beginning in the late 1980s and early 1990s (Grissmer et al., 1998; Magnuson & Waldfogel, 2008). Other studies using state NAEP and state testing data, however, have instead found a slow but steady decline in the gap over time (Reardon, Greenberg, Kalogrides, Shores, & Valentino, 2012; Reardon et al., 2013; Reardon & Owens, 2014).

Characterizing temporal trends in the Black-White test score gap into the 21st century using meta-regression can inform debates over whether the test gap is declining or stagnating, and can also provide insight into whether accountability pressures linked to the No Child Left Behind Act of 2001 are associated with gap closure over time. An explicit goal of NCLB was to close racial/ethnic achievement gaps through regular testing and performance accountability by subgroup. A number of studies find that the accountability era has been associated with increased average student achievement levels on both state and national tests (Jennings & Lauen, 2016; Lauen & Gaddis, 2012), and a reduction in Black-white achievement gaps in states such as North Carolina and California (Gaddis & Lauen, 2014; Strunk & McEachin, 2014).

Some studies, however, find limited evidence that black student achievement is increasing amidst accountability pressures (Jennings & Lauen, 2016) or that increasing student achievement levels are associated with Black-White gap closure (Hanushek & Raymond, 2005). Additionally, improvement during the accountability era may be due to unrelated trends rather than NCLB (Lee & Reeves, 2012; Reardon, Greenberg, et al., 2012; Reardon et al., 2014).

In addition to uncertainty about the trajectory of the Black test score gap into the first period of the 21st century, there is disagreement among researchers about the proportion of the test score gap that is due to non-race observable factors. The role of socioeconomic status in explaining the Black-White test score gap represents a case in point. Hedges and Nowell (1999), Herrnstein and Murray (1994), and Phillips, Brooks-Gunn, Dunkan, Klebanov, and Crane (1998) have estimated that household socioeconomic characteristics such as parental education and income explain about one-third of the Black-White test score gap. Phillips et al. (1998), however, concluded that a wider set of socioeconomic indicators, such as grandparents' educational attainment, mothers' school quality, birth weight, and parenting practices explained about two-thirds of the Black-White test score gap. Rothstein and Wozny (2013) find similar results by adding permanent income as an explanatory variable.

The contribution of schooling characteristics to the gap is another case in point. Despite declining neighborhood Black-White neighborhood segregation over time (Logan & Stults, 2011; Logan, Stults, & Farley, 2004), black students are often in minority- and poverty-concentrated schools (Orfield, Kucsera, & Siegel-Hawley, 2012) as a result of resegregation as courts have rescinded formal desegregation orders and struck down the use of race in school assignment (Orfield & Eaton, 1996; Reardon, Grewal, Kalogrides, & Greenberg, 2012). Attending minority- or black-concentrated schools is negatively associated with achievement in general and for black students specifically (Hanushek, Kain, & Rivkin, 2009; Mickelson, Bottia, & Lambert, 2013). Racial and school segregation are both correlated negatively with black achievement growth and may exacerbate test score gaps (Card & Rothstein, 2007; Condron, Tope, Steidl, & Freeman, 2013; Rumberger & Palardy, 2005; Rumberger & Willms, 1992).

Most previous studies have also not considered whether the influence of observable variables on the racial gap has changed over time. The foremost dataset used to analyze achievement gaps over time, NAEP, lacks a wide range of background measures, particularly

those related to household socioeconomic status (Berends & Peñaloza, 2008; National Center for Education Statistics, 2012). This limits the ability of researchers to examine time-varying associations between socioeconomic status and the test score gap. Using state NAEP data, Reardon et al. (2013) find that differences in gap levels across states can be attributed to observable differences in student socioeconomic status and racial school segregation, but changes over time in gaps across states could not be explained by these factors. Berends and Peñaloza (2008) estimate that changes in family background characteristics accounted for 35–62% of reductions in Black-White test score gap, and changes in minority school composition accounted for a 65% increase in the gap, from 1972 to 2004. Neither Berends and Penaloza nor Reardon et al. determine whether the associations between background factors and schooling characteristics remained constant or had time-varying influences on the gap.

Our study contributes to research on time trends in the Black test score gap in three ways. First, we widen the scope of assessments used to summarize time trends in test score gaps. The school accountability movement has spurred the development and implementation of a myriad of tests to measure achievement, and the magnitude of educational gaps may vary according to the assessment administered. Second, we extend the time horizon for analyzing gap closure through 2010, the end of the first decade of the implementation of NCLB, which has sought to close racial/ethnic achievement gaps.

Finally, we leverage a novel analytical technique—meta-regression—to estimate the contribution of observable confounding variables to the Black test score gap. By aggregating information about the coefficient on Black racial status in published studies with achievement outcomes, we can gather evidence on the gap in a wide array of contexts, and estimate the contribution of observable confounding variables to the Black test score gap and can determine whether these observable factors have had constant or varying influences on the gap over time.

Meta-Regression Data

We collected information on 1,196 regression coefficients published in 165 journal articles.[1] Our data collection process consisted of (a) establishing inclusion criteria, (b) creating an article search pool based on inclusion criteria, (c) screening studies in the article search pool for inclusion in the study, and (d) coding qualifying articles for selected attributes.

We required that studies were of U.S. K-12 students and had at least one regression coefficient on African American status from an analysis using individual-level test scores. Importantly, we have included studies that incidentally calculated the test score gap without focusing the study on race. These studies mimic “gray literature”—studies that have not yet been published—to account for the possibility of how publication bias might affect meta-analysis (McAuley, Pham, Tugwell, & Moher, 2000). Among studies that focus on racial achievement gaps, the publication decision may be linked to the size or significance of racial

[1]This dataset is available upon request from the authors.

effects. These publication bias pressures are less likely to apply to non-race-focused studies.[2]

We narrowed our inclusion criteria to articles published between January 1990 and June 2013 in one of the top 50 journals in economics, sociology, and education, 150 journals in total.[3] These fields commonly make use of regression analysis rather than experiments that do not require race controls. The time frame began in 1990 to ensure that most datasets would have been collected in recent decades, and ended in June 2013 when candidate articles were selected.

We used Google Scholar to create the initial article search pool. Our 94 search terms (listed in the Electronic Supplementary Material) related to standardized exams in general, as well as the specific names of national and state standardized exams (NAEP, CBEST, etc.). We identified 9,062 articles containing our search terms, and screened them in a randomized order. We set a goal of finding 165 eligible studies, in order to have acceptable statistical power for estimating primary effects. We found 165 eligible articles after screening 6,730 articles.

The total sample consists of 165 articles. Per our search process, publication dates are between 1990 and 2013. The number of observations by year is available in Appendix A1, and the number of articles from each journal is listed in the Electronic Supplementary Material. The appendices and electronic supplementary material can both be found at http://nickchk.com/Huntington-Klein_Ackert_SSQ_Appendices.pdf.

Table 1 provides summary statistics. The final sample contains 1,157 observations from 165 studies. Each of the 1,157 observations represents a different statistical analysis. The median sample size is 6,818 students, with a mean of 13.2 covariates in the estimated model. Students cover the entire K-12 range, with a mean sample age of about 13 years. Approximately 14.1% of studies compare Black students to a multiracial reference group, rather than a White reference group. Results are robust if these studies are dropped, and we control for reference group.

We are interested in evaluating how confounding variables and other aspects of the regressions explain variation in the test score gap. We record the presence or absence of controls for confounding variables that are discussed widely in the literature on racial achievement gaps and could increase or decrease the magnitude of the test score gap. The presence of controls for socioeconomic status is a central focus. We distinguish between parental education and all other forms of socioeconomic status, including direct measures of income and free and reduced-price lunch eligibility. We refer to the latter simply as SES, keeping in mind that parental education is also an indicator of socioeconomic status. We also code whether the study controls for certain other confounding factors that have been shown to influence Black-White achievement gaps: parental marriage, urbanicity, English fluency,

[2]Only 15.8% of the articles in our sample were “race-focused.” We recorded an article as race-focused if the article mentioned race as a motivating topic in the title, abstract, or introduction.

[3]Top journals were determined by the quality-weighted citation index at Journal-Ranking.com on May 2, 2013. Journal-Ranking.com is defunct, but their methodology is in Lim et al. (2007).

prior academic achievement (prior grades or test scores),[4] school characteristics (school composition, school location, etc.), and teacher characteristics (years of experience, teacher race, etc.).

As Table 1 indicates, controls for SES are the most prevalent across regressions in our sample. Few studies include controls for urbanicity, English fluency, or school or teacher characteristics. Most exam scores reported are for math or reading/English exams. The Other exam score category primarily consists of spatial-mechanical reasoning tests. Results are not affected if the analysis is limited to only reading, math, and composite tests.

We additionally record whether or not the studies use exams that are administered on a nationally representative scale. Non-nationally representative exams include state and local assessments. Typically, exams given to nationally representative samples of students are low-stakes for students and teachers because they are not tied to intervention or sanctions. Approximately 56.8% of our sample uses nationally representative exams.

Meta-Regression Estimation

Following Ringquist (2013), we calculate a generalized partial correlation r between African American racial status and standardized test scores, which accounts for differences in testing levels and the spread of exam scores. We use partial correlations rather than other measures of effect size, for example d-indices, because they allow for the gap to be adjusted for the presence of confounding variables, provide an easily interpretable estimate of the gap that is comparable across studies, and can be computed from nearly any regression table without further information such as the standard deviation of the variables. This is the same approach taken by Mickelson et al. (2013), who focus on school composition as a confounding variable.

We extract one partial correlation r_{is} from each regression result i in each study s, which we put through a Fisher transformation (Fisher, 1921; Hotelling, 1953) to get Θ_{is} which has a normal distribution with variance $v_{is} = 1/(n_{is} - 3)$. Θ_{is} can be used to estimate the regression

$$\Theta_{is} = \alpha + F(\gamma, t_{is}) + \beta X_{is} + \varepsilon_{is} \quad (1)$$

where $F(\gamma, t_{is})$ is a function of the year in which the regression sample was collected t_{is} and a parameter vector γ, and X_{is} is a vector describing the original regression, such as what other variables were controlled for (e.g. prior academic achievement).

The intercept in the model is adjusted such that it can be interpreted as the estimated effect size at the mean of the data X_{is} and t_{is}. The presence or absence of confounding variables

[4]The control for prior academic achievement separates raw achievement gaps from gaps in achievement growth. We include studies that control for prior achievement to situate the gap properly as a part of a developmental and cumulative process, and then control for the presence of the confounding variable in the regression.

X_{is} helps to explain variation in the partial correlation Θ_{is}. For example, in a regression that does not feature a control for parental education, Θ_{is} contains the part of the gap associated with racial differences in parental education. The coefficient on "includes a control for parental education" measures the contribution of parental education to the measured gap.

In estimates of our equation of interest, observations cannot be considered independent, and the number of observations varies between studies. We estimate equation 2 using Generalized Estimating Equations (GEE) (Liang & Zeger, 1986). We assume that the covariance matrix of ε_{is} is block-diagonal, with zero covariance between studies and a covariance of ρ between observations in the same study. The diagonal terms are the sum of a common variance parameter τ^2, a study-specific error variance σ_s^2, and observation variance v_{is}. We estimate $\hat{\tau}^2$ using WLS where observations are weighted by $v_{is}^{-1/2}$, allowing GEE to incorporate differences in the number of observations by study into parameter and variance estimates (Ringquist, 2013).

We do not model dependence between studies, which may be nonzero if the same dataset or sample setting is studied multiple times. Between-study correlation exists but likely provokes only a small bias; overall 114 different samples were used and most repeated samples studied outcomes on different exams or at different times, used meaningfully different sample restrictions, or incorporated additional data sources.

Results

Test Score Gaps over Time

We begin with a meta-regression model in which the test score gap is only allowed to vary by year. Figure 1 shows predicted partial correlations between African American status and test score, based on the sample year. Values closer to zero represent a narrower gap between scores of African Americans and those of the reference group, most often non-Latino Whites.[5] In Figure 1, two of the time trend lines treat year as a continuous variable, and estimate a quadratic and a cubic form of year. The horizontal lines represent alternative models with time span dummies, in groups of three and five years each. Finally, we estimate a model introducing a linear term for year which is allowed to vary before and after 1995, the midpoint for the years of interest in our study. Raw regression coefficients can be found in Appendix Table A1.1.

There was non-linear change in the racial test gap from 1979 to 2010. Consistent with prior work, we find significant improvements in the test gap in the early 1980s. However, all analyses suggest hardly any decrease in the magnitude of the gap after 1995. The linear rate after 1995 is negative and not statistically distinguishable from zero, but distinguishable from the positive slope before 1995 at the 95% confidence level. The three- and five-year

[5]These estimates are corrected for sample size and the intra-correlation of effects within studies, as outlined in the methods section. Figure 1 does not take into account confounding variables. Results are robust to the inclusion of other confounding variables.

dummy results confirm that closure of the gap has slowed considerably over time, and that there is little progress in gap closure in recent years.

The time span dummies allow for a test of the change in the estimated gap after the implementation of No Child Left Behind (NCLB), although we cannot claim that our data are capable of estimating a causal effect of NCLB. Most change in the racial test gap occurred before NCLB was put into effect. The widest gap between the residual time effect in 1998-2001 (before implementation), 2002-2004 (during), or 2005-2007 and 2008-2010 (after) is only significant with $p = .118$ (2002-2004 vs. 2005-2007). We cannot conclude from these results that NCLB has had no effect on the test score gap, because it is possible that NCLB's effects were countervailed by unobserved factors that reversed the effect of the policy. However, NCLB has not been able to supersede any factors that could have widened the gap.

The Test Gap and Confounding Variables

We next use meta-regression to examine how confounding variables and other features of the regression models in the studies in our sample change the reported size of the test gap. Table 2 displays the results of three meta-regressions. A positive coefficient of .05 suggests that the inclusion of that control variable in a regression models shrinks the published gap by .05 in partial correlation units, or roughly .15 (three times larger) test score standard deviations.[6]

In Model 1a, the variables that most significantly reduce the test gap are parental education, SES, and prior academic achievement. Surprisingly, controls for teacher characteristics widen the estimated test gap, but this appears to be due to an association with school characteristics and socioeconomic status; with these dropped the coefficient on the teacher characteristics control becomes positive but insignificant. There are also some noteworthy null results in Model 1a. School characteristics and parental marriage have little influence on the gap, net of other confounding variables. Also, there is not a significant level of difference in the test gap between nationally representative and non-nationally representative exams or across different exam subjects, with the exception of social science exams.

Situating our results in the literature, we estimate the proportion of the gap explained by each confounding variable. Column 1b in Table 2 displays the percentage of the gap that is explained when each control variable is added to the baseline model, compared to the -.172 gap predicted when no confounding variables are included. For example, the inclusion of socioeconomic status as the only control variable explains .075/.172 = 43.6% of the gap.

We estimate that parental education and other forms of socioeconomic status explain 30.8% and 43.6% of the achievement gap, respectively. To predict the share of the gap explained by socioeconomic status as a whole, we estimate a new model with an interaction term to account for any overlap in the portion of the gap explained by the two measures. In this model, socioeconomic status as a whole explains 74.5% of the gap. Prior achievement also

[6]To see this, take the example of a reported standardized coefficient, which gives the effect of a one-standard deviation change in African American status on standardized exam scores. Dividing the coefficient by the standard deviation of racial status produces the gap in standard deviation units. If 10% of the sample is black, this suggests multiplying by $(.1 \times .9)^{.5}/.1 = 3$.

plays an important role in driving achievement disparities, and explains over one-quarter of the African American test score gap. In contrast, parental marriage, urbanicity, school, and teacher characteristics each explain less than 15% of the achievement gap.

Model 2 in Table 2 includes only studies that use nationally representative exams such as NAEP, which are most often analyzed in prior work on the black test score gap. In nationally representative studies, controlling for parental education and prior achievement leads to a larger reduction in the measured gap than in studies in the full sample. Controlling for SES in nationally representative studies reduces the gap to a lesser extent than controlling for this variable in the full sample of studies. This disparity may be due to the wider range of socioeconomic measures observed in the full sample. The results in Model 2 also demonstrate that nationally representative studies are capturing certain parts of the African American test gap, but they do not capture the same variation that is found in a broader sample of studies.

A common finding in the test gap literature is that the racial gap is larger for older students than for younger students (e.g. Mickelson et al., 2013). In Model 3 we also find a larger gap for older students, with the partial correlation becoming .025 more negative from age 13 to age 18, but the difference is not statistically significant.

The Influence of Confounding Variables on the Gap over Time

In Table 3 we evaluate how the influence of confounding variables on the size of the gap may have changed from 1979 to 2010. We use a linear specification for year because the sample size is not large enough to precisely estimate confounding variables interacted with three-year time span dummies. The influence of several confounding variables on the gap changes significantly over time, with differences in urbanicity and teacher characteristics having a weaker influence on the gap over time, and school characteristics having a stronger influence. We do not find statistically significant changes within the studies over time for the impact of socioeconomic status, parental marriage, or prior academic achievement on the test score gap.

Focusing even further on parental education and SES, Figure 2 displays the proportion of the gap that is explained by the introduction of a control for either SES, parental education, or both.[7] We allow this proportion to change over time, using both a quadratic specification for time and time dummies that span a three-year window. The proportion of the gap that is explained by socioeconomic status is rising, but the growth is slow and is not statistically significant, as can be seen in the confidence bands on the three-year dummies.

Conclusion

Black test score gaps with peers remain a pervasive component of the U.S. education system. The persistence of these achievement gaps will have long-term implications for

[7]The proportion is calculated using a model in which “controls for parental education, SES, or both” is included as a predictor. The total proportion is consistently smaller than the 74.5% reported in the previous section due to the use of a single time-varying indicator, as opposed to including both measures and their interaction independently. This approach is taken to avoid including too many interaction terms for the sample size to appropriately estimate.

racial inequality in multiple domains of social life, including patterns of income inequality and health disparities. Using meta-regression, we find an overall reduction in the Black achievement gap during the period of 1979 to 2010, mainly concentrated in the 1980s. This finding aligns with other studies that show that the gap narrowed directly after the major reforms of the Civil Rights movement in the 1970s and 1980s (Berends & Peñaloza, 2008; Grissmer et al., 1998; Hedges & Nowell, 1999), but stalled in the period of the 1990s and early 2000s (Magnuson & Waldfogel, 2008; Reardon et al., 2013). Our meta-regression analysis fails to find any improvement in racial achievement disparities during the era of increased school accountability pressures associated with the implementation of NCLB, in the late 1990s to 2010. We do not aim to provide a causal estimate of the effect of NCLB, but, we can infer that if No Child Left Behind reduced the gap, these positive effects must have been countered by other unobserved changes.

In an article predicting patterns of educational inequality in the 21st century, Gamoran speculated that, under conditions of stability, the United States would see "less inequality by race" in the 21st century (Gamoran, 2001 pp. 148). Two key questions, however, are "how much decline in racial inequality will be observed?" and, "in what period of time?" While we do find that the gap is smaller than it used to be, it is not going away any time soon. If this paper had been written in 1994, using observations from that time, we would have predicted the complete closure of the background-controlled gap in 1998. This did not occur. Our full sample predicts that the background-controlled gap will close entirely around 2046. This is to say nothing of the raw uncontrolled gap. A structural return to whatever mechanisms were closing the gap in the 1980s and early 1990s, assuming those same mechanisms are still as impactful, would be necessary in order to expect a closing of the gap within any reasonable time frame.

Importantly, our study highlights how a substantial portion of Black achievement gaps with peers can be explained by observable confounding variables. Some previous studies have found that differences in socioeconomic status only account for only approximately one-third of the Black-White test score gap (Hedges & Nowell, 1999; Herrnstein & Murray, 1994). We show that broad combined measures of socioeconomic status explain almost three-quarters of the partial correlation between Black racial status and achievement, which is consistent with Phillips et al. (1998) and Rothstein and Wozny (2013).

Despite an increasing emphasis on school and teacher accountability in education reform movements, we do not find that a significant proportion of the gap is explained by school and teacher characteristics from 1979 to 2010. Rather, controls for socioeconomic status appear to be the main mediators of the published Black test score gap. These results do not imply that "schools don't matter." Given that prior achievement explains a noticeable portion of the gap, it is likely that school- and classroom-level factors in early childhood influence the test score gap via academic development trajectories. We also find that the influence of schools on the gap is increasing over time, which could be the result of increased accountability pressures.

Our work suggests that socioeconomic status continues to be a key determinant of Black-White test score gaps, which is often underappreciated in current education policy

discussions. If education policies are to ameliorate achievement gaps, then more robust tactics must be taken to further increase the impact of schools on gap closure relative to socioeconomic factors. Additionally, any policy that boosts household socioeconomic resources can be viewed as a potential lever to change the Black achievement gap.

Our results are subject to the limitations of meta-analytic techniques, which are well-documented by methodologists in many fields (Cooper, Hedges, & Valentine, 2009; Franke, 2001; Hedges, 1992). One limitation is that our results only generalize to the population of studies in our sample. If these studies are consistently biased representations of reality, so too will be our results. Other limitations include varying quality in the studies included in the sample (Moher et al., 1998), comparability of outcome measures across studies, variability in the precision of findings across studies, and the non-independence of observations within studies. We cannot address all of these and they may affect findings, but we minimize heterogeneity in study quality within our sample by focusing on studies in top peer-reviewed journals and by including many studies, thus reducing the impact of any one low-quality paper.

Despite these limitations, we see the potential for using meta-regression in future social science research to understand social inequalities. Our approach takes advantage of a large number of research results that would be ignored in a typical literature review or meta-analysis. The inclusion of studies not focused on racial gaps as “gray literature” minimizes the influence of publication bias on results and allows us to understand the gap over a wider array of contexts than is standard. Our approach could be expanded to the study of inequality in outcomes between other groups, such as between immigrant and non-immigrant students.

Overall, our results suggest that African American students face a “long road” to achieving equal test score outcomes with their peers. We see avenues for improvement and ways in which policy could help, but recent improvements have been hard to come by. Our findings suggest that policies seeking to ameliorate the Black-White test score gap need to target multiple domains of African American students’ lives, but improving the socioeconomic status of the households where Black students live would go a long way to accelerating gap closure.

References

Berends M, Peñaloza RV. Changes in Families, Schools, and the Test Score Gap. In: Magnuson K, Waldfogel J, editorsSteady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap. New York: Russell Sage Foundation; 2008. 66–109.

Card D, Rothstein J. Racial Segregation and the Black–White Test Score Gap. Journal of Public Economics. 2007; 91(11):2158–2184.

Condron DJ, Tope D, Steidl CR, Freeman KJ. Racial Segregation and the Black/White Achievement Gap, 1992 to 2009. The Sociological Quarterly. 2013; 54(1):130–157.

Cooper H, Hedges LV, Valentine JC. The Handbook of Research Synthesis and Meta-analysis. New York City, NY: Russell Sage Foundation; 2009.

Fisher RA. On the “Probable Error” of a Coefficient of Correlation Deduced from a Small Sample. Metron. 1921; 1:3–32.

Franke GR. Applications of Meta-analysis for Marketing and Public Policy: A Review. Journal of Public Policy & Marketing. 2001; 20(2):186–200.

Gaddis SM, Lauen DL. School Accountability and the Black-White Test Score Gap. Social Science Research. 2014; 44:15–31. [PubMed: 24468431]

Gamoran A. American Schooling and Educational Inequality: A Forecast for the 21st Century. Sociology of Education. 2001; 74(2001):135–153.

Grissmer D, Flanagan A, Williamson S. Why did the Black-White Score Gap Narrow in the 1970s and 1980s?. In: Jencks C, Phillips M, editorsThe Black White Test Score Gap. Washington, D.C: The Brookings Institution; 1998. 182–228.

Hanushek EA, Kain JF, Rivkin SG. New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement. Journal of Labor Economics. 2009; 27(3): 349–383.

Hanushek EA, Raymond ME. Does School Accountability Lead to Improved Student Performance? Journal of Policy Analysis and Management. 2005; 24(2):297–327.

Hedges LV. Meta-analysis. Journal of Educational and Behavioral Statistics. 1992; 17(4):279–296.

Hedges LV, Nowell A. Changes in the Black-White Gap in Achievement Test Scores. Sociology of Education. 1999; 72(2):111–35.

Herrnstein RJ, Murray C. The Bell Curve: The Reshaping of American Life by Differences in Intelligence. New York City: Free Press; 1994.

Hotelling H. New Light on the Correlation Coefficient and its Transforms. Journal of the Royal Statistical Society Series B (Methodological). 1953; 15(2):193–232.

Jencks C, Phillips M. The Black-White Test Score Gap. Washington, D.C: Brookings Institution Press; 1998.

Jennings JL, Lauen DL. Accountability, Inequality, and Achievement: The Effects of the No Child Left Behind Act on Multiple Measures of Student Learning. RSF: The Russell Sage Foundation Journal of the Social Sciences. 2016; 2(5):220–241.

Lauen DL, Gaddis SM. Shining a Light or Fumbling in the Dark? The Effects of NCLB's Subgroup-Specific Accountability on Student Achievement. Educational Evaluation and Policy Analysis. 2012; 34(2):185–208.

Lee J, Reeves T. Revisiting the Impact of NCLB High-Stakes School Accountability, Capacity, and Resources: State NAEP 1990–2009 Reading and Math Achievement Gaps and Trends. Educational Evaluation and Policy Analysis. 2012; 34(2):209–231.

Liang KY, Zeger SL. Longitudinal Data Analysis using Generalized Linear Models. Biometrika. 1986; 73(1):13–22.

Lim A, Ma H, Wen Q, Xu Z, Cheang B, Tan B, Zhu W. Proceedings of the National Conference on Artificial Intelligence. Cambridge, MA: AIII Press; 2007. Journal-Ranking.com: An Online Interactive Journal Ranking System.

Logan JR, Stults B. The Persistence of Segregation in the Metropolis: New Findings from the 2010 Census. Census Brief prepared for Project US 2010. Washington, D.C: 2011.

Logan JR, Stults BJ, Farley R. Segregation of Minorities in the Metropolis: Two Decades of Change. Demography. 2004; 41(1):1–22. [PubMed: 15074122]

Magnuson K, Waldfogel J. Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap. New York City: Russell Sage; 2008. Introduction; 1–29.

McAuley L, Pham B, Tugwell P, Moher D. Does the Inclusion of Grey Literature Influence Estimates of Intervention Effectiveness Reported in Meta-analyses? The Lancet. 2000; 356(9237):1228–1231.

Mickelson RA, Bottia MC, Lambert R. Effects of School Racial Composition on K-12 Mathematics Outcomes: A Metaregression Analysis. Review of Educational Research. 2013; 83(1):121–158.

Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, Klassen TP. Does Quality of Reports of Randomised Trials Affect Estimates of Intervention Efficacy Reported in Meta-analyses? The Lancet. 1998; 352(9128):609–613.

National Center for Education Statistics. Improving the Measurement of Socioeconomic Status for the National Assessment of Educational Progress: A Theoretical Foundation. Washington, D.C: 2012.

Orfield G, Eaton SE. Dismantling Desegragation: The Quiet Reversal of Brown v Board of Education. New York City, NY: The New Press; 1996.

Orfield G, Kucsera J, Siegel-Hawley G. “E Pluribus”… Separation: Deepening Double Segregation for More Students. Los Angeles, CA: Proyecto Derechos Civiles; 2012.

Phillips M, Brooks-Gunn J, Duncan GJ, Klebanov P, Crane J. Family Background, Parenting Practices, and the Black-White Test Score Gap. In: Jencks C, Phillips M, editorsThe Black White Test Score Gap. Washington, D.C: Brookings Institution Press; 1998. 103–150.

Quinn DM. Kindergarten Black-White Test Score Gaps: Re-examining the Roles of Socioeconomic Status and School Quality with New Data. Sociology of Education. 2015; 88(2):120–139.

Reardon SF, Greenberg EH, Kalogrides D, Shores KA, Valentino RA. Trends in Academic Achievement Gaps in the Era of No Child Left Behind. Society for Research on Educational Effectiveness. 2012

Reardon SF, Grewal ET, Kalogrides D, Greenberg E. Brown Fades: The End of Court-Ordered School Desegregation and the Resegregation of American Public Schools. Journal of Policy Analysis and Management. 2012; 31(4):876–904.

Reardon SF, Owens A. 60 Years After Brown: Trends and Consequences of School Segregation. Annual Review of Sociology. 2014; 40:199–218.

Reardon SF, Robinson-Cimpian JP, Weathers ES. Patterns and Trends in Racial/Ethnic and Socioeconomic Academic Achievement Gaps. In: Ladd HA, Goertz ME, Erlbaum L, editorsHandbook of Research in Education Finance. New York City, NY: Routledge; 2014.

Reardon SF, Valentino RA, Kalogrides D, Shores KA, Greenberg EH. Patterns and Trends in Racial Academic Achievement Gaps Among States, 1999-2011. 2013

Ringquist E. Meta-analysis for Public Management and Policy. San Francisco, CA: Jossey-Bass; 2013.

Rothstein J, Wozny N. Permanent Income and the Black-White Test Score Gap. Journal of Human Resources. 2013; 48(3):509–544.

Rumberger RW, Palardy GJ. Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High School. Teachers College Record. 2005; 107(9):1999–2045.

Rumberger RW, Willms JD. The Impact of Racial and Ethnic Segregation on the Achievement Gap in California High Schools. Educational Evaluation and Policy Analysis. 1992; 14(4):377–396.

Strunk KO, McEachin A. More than Sanctions: Closing Achievement Gaps through California’s Use of Intensive Technical Assistance. Educational Evaluation and Policy Analysis. 2014; 36(3):281–306.

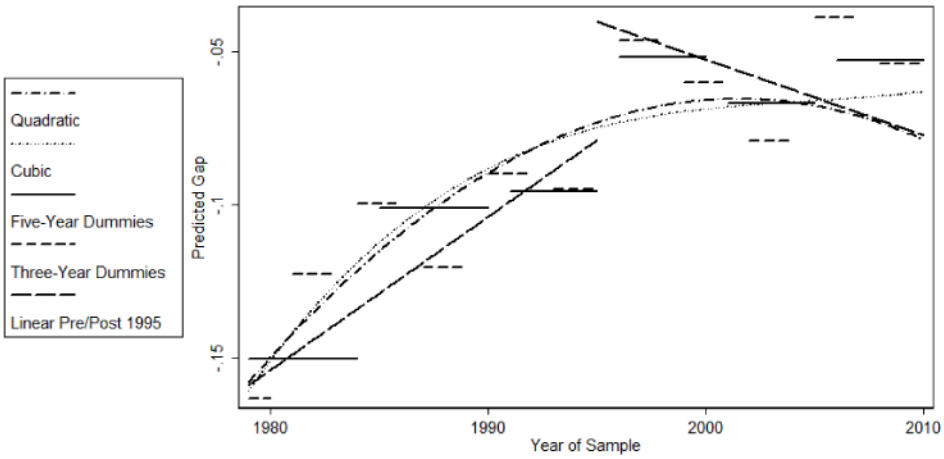

Figure 1.
Predicted Black Test Score Gap, 1979-2010, With Differing Functional Forms for Time

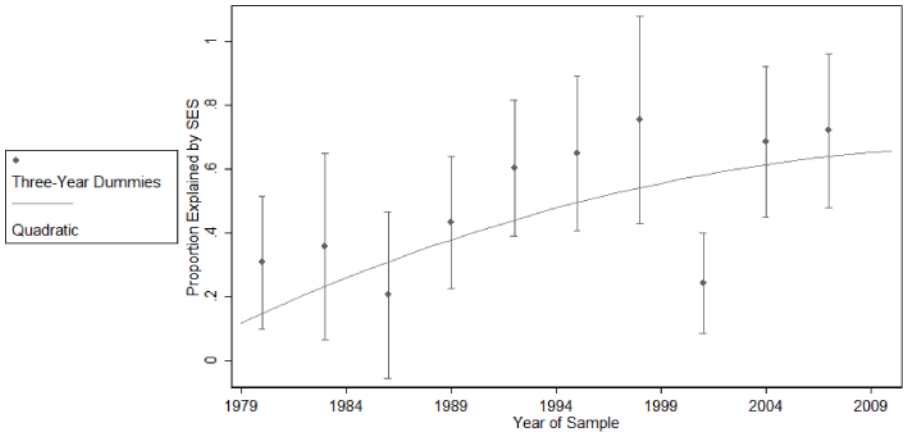

Figure 2.
Proportion of Black Test Score Gap Explained by Parental Education and SES over Time.

Table 1

Summary Statistics, Meta-Regression Sample from 1,157 Regressions in 165 Peer-Reviewed Journals

Variable	Mean	Std. Dev.
Adjusted Partial Correlation Θ_{is}	−.103	.204
Coefficient on African American is:		
Negative and significant at α = .95	.765	
Insignificant at α = .95	.204	
Positive and significant at α = .95	.031	
Confounding variable controlled for:		
Parental education	.362	
Socioeconomic status	.742	
Parental marriage	.264	
Urbanicity	.143	
English fluency	.221	
Prior academic achievement	.444	
School characteristics	.175	
Teacher characteristics	.086	
Sample size (median)	6,818	
Number of regression covariates	13.204	15.144
Age of students	13.198	4.305
White comparison group	.859	
Non-Black comparison group	.141	
Exam types:		
Math	.347	
Reading/English	.356	
Science	.114	
Social science	.029	
Composite	.135	
Other	.019	
Nationally representative	.568	

Table 2

Effects of Confounding Variables on the Black Test Score Gap

Model:	**(1a)**		**(1b)**	**(2)**		**(3)**	
	Coef.	**(s.e.)**	**Pct.**	**Coef.**	**(s.e.)**	**Coef.**	**(s.e.)**
Controls included:							
Parental ed.	.053***	(.017)	30.8	.077***	(.019)		
SES	.075***	(.020)	43.6	.036***	(.013)		
Parent marriage	−.009	(.017)	5.2	.006	(.012)		
Urbanicity	−.023	(.017)	13.4	−.015*	(.009)		
English fluency	.037***	(.013)	21.5	−.022	(.016)		
Prior achieve.	.048***	(.013)	27.9	.067***	(.014)		
School chars.	.021	(.013)	12.2	.001	(.009)		
Teacher chars.	−.009*	(.005)	5.2	.010	(.007)		
Exam types:							
Composite (omitted)							
Math	−.002	(.010)		.003	(.011)		
Reading/English	−.001	(.010)		.004	(.014)		
Science	−.006	(.006)		−.001	(.007)		
Social science	.046*	(.026)		.057**	(.027)		
Other	−.003	(.015)		.004	(.012)		
Year	.004	(.005)		.002	(.004)	.002	(.004)
$Year^2$	−.0001	(.0002)		−.0003*	(.0002)	−.0001	(.0001)
$Year^3$	-1.48×10^{-7}	(2.37×10^{5})		-2.94×10^{-6}	(4.65×10^{-5})	5.68×10^{-6}	(1.92×10^{-5})
Age						−.003	(.002)
Age^2						−.0004	(.0005)
Comparison= White	−.053**	(.024)		.046	(.036)	−.023	(.022)
Nat. Rep. Exam	.012	(.019)				.025	(.023)
Constant	.088***	(.017)		.067***	(.013)	.071***	(.017)

Model:	**(1a)**		**(1b)**	**(2)**		**(3)**	
	Coef.	**(s.e.)**	**Pct.**	**Coef.**	**(s.e.)**	**Coef.**	**(s.e.)**
n	1,157			657		1,157	

*/**/*** indicates statistical significance at the 10%/5%/1% level.

Table 3

Influence of Observable Background Characteristics over Time

	Coefficient	(s.e.)
Controls included:		
Parental education	.039**	(.017)
Socioeconomic status	.070***	(.016)
Parental marriage	−.001	(.019)
Urbanicity	−.028**	(.011)
English fluency	−.029**	(.013)
Prior academic achievement	.048***	(.012)
School characteristics	.027**	(.013)
Teacher characteristics	−.008**	(.004)
Year interaction:		
Parental education *Year	.003	(.002)
Socioeconomic status *Year	.002	(.002)
Parental marriage *Year	.003	(.002)
Urbanicity *Year	−.004**	(.002)
English fluency *Year	−.003**	(.002)
Prior academic achievement *Year	−.002	(.001)
School characteristics *Year	.005*	(.003)
Teacher characteristics *Year	−.001**	(.0003)
Comparison = White	−.057**	(.026)
Nationally Rep. Exam	.010	(.018)
Constant	−.086***	(.019)
Exam type controls	Yes	
Cubic year controls	Yes	

Notes:

*/**/*** indicates statistical significance at the 10%/5%/1% level.