NBER WORKING PAPER SERIES

HAVE INCOME-BASED ACHIEVEMENT GAPS WIDENED OR NARROWED?

Shirin A. Hashim
Thomas J. Kane
Thomas Kelley-Kemple
Mary E. Laski
Douglas O. Staiger

Working Paper 27714
http://www.nber.org/papers/w27714

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2020

A grant from the Walton Family Foundation supported Mary Laski and Shirin Hashim. Pre-doctoral training grant R305B150010-16 from the Institute of Education Sciences at the U.S. Department of Education supported Thomas Kelley-Kemple. We thank Aimee Chen, Janet Currie, David Deming, John Friedman, David Grissmer, Eric Hanushek, Caroline Hoxby, Dan Koretz, Eric Nielsen, Paul Peterson, Steve Raudenbush, Sarah Reber, Jesse Rothstein, Sarah Turner, Jim Wyckoff and seminar participants at the University of Virginia, Harvard Graduate School of Education and the 2020 NBER Summer Institute for helpful comments and suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w27714.ack

Have Income-Based Achievement Gaps Widened or Narrowed?
Shirin A. Hashim, Thomas J. Kane, Thomas Kelley-Kemple, Mary E. Laski, and Douglas O. Staiger
NBER Working Paper No. 27714
August 2020
JEL No. I21,I22,I24

ABSTRACT

Since 1990, U.S. policymakers have worked to close gaps in academic achievement by income and race (e.g. with school finance reform and school accountability systems) even as rising income inequality and income-based residential segregation have threatened to widen them. Using estimates of the mean and variance in household income for sampled schools, we reconstruct the student-level relationship between achievement and household income in the National Assessment of Educational Progress (NAEP) from 1990 to 2015. We find that achievement at all levels of parental income rose substantially in 4th and 8th grade. In contrast to Reardon (2011), we find that achievement gaps narrowed substantially in 4th grade reading and math and in 8th grade math, while the gaps remained stable in 8th grade reading. As a robustness check, we used the March Current Population survey to impute income for dependent children by race, mother's education, urbanicity and state and then calculated mean achievement for those same groups in the NAEP. Again, we found gaps in achievement narrowing between groups with high and low predicted mean household incomes. Our results challenge the prevailing understanding that income-based achievement gaps have widened in the United States over the last 30 years.

Shirin A. Hashim
255 W 94TH ST
APT 17F
10025
NEW YORK, NY
shirin_hashim@g.harvard.edu

Thomas J. Kane
Harvard Graduate School of Education
Center for Education Policy Research
50 Church St., 4th Floor
Cambridge, MA 02138
and NBER
kaneto@gse.harvard.edu

Thomas Kelley-Kemple
3436 30th St Apt 1
Astoria, NY 11106
tkelleykemple@g.harvard.edu

Mary E. Laski
marylaski@g.harvard.edu

Douglas O. Staiger
Dartmouth College
Department of Economics
HB6106, 301 Rockefeller Hall
Hanover, NH 03755-3514
and NBER
douglas.staiger@dartmouth.edu

I. Introduction

In 1848, Horace Mann famously described public education as "the great equalizer of the conditions of men" and the "balance wheel of the social machinery." In the face of rising income inequality and income-based housing segregation, policymakers and philanthropists have launched a series of efforts over the past three decades to preserve the balance wheel of schooling: raising educational standards, increasing school accountability, opening charter schools and increasing state aid to low-income districts. Yet, we know remarkably little about how the relationship between parental income and achievement has changed over time.

The chief obstacle has been the lack of data. In the United States, there is no nationally representative data source combining a consistent measure of academic achievement with the household income of individual students.

The best available evidence was provided in 2011, when Sean Reardon combined data from twelve nationally representative surveys, with different measures of academic achievement, administered to students at different ages, matched with different measures of student-reported or parent-reported income. After pooling the data for 12 cohorts of students, each with different measures of achievement and income, Reardon (2011) concluded that the income-based achievement gap had expanded substantially for the cohorts born between 1974 and 2001: the difference in mean achievement at the 90th and 10th percentiles of family income had risen by 40 to 50 percent in standard deviation units.

We take a different approach, supplementing the National Assessment of Educational Progress (NAEP) with neighborhood income measures from the U.S. Bureau of the Census. Specifically, we use the school-level measures of achievement and income to reconstruct the student-level relationships. To do so, we rely on the fact that the student-level least-squares coefficient of achievement on household income is a weighted average of the between-school and within-school relationships between achievement and income. We estimate the between-school coefficient of achievement on income directly, regressing school-level mean achievement on mean household income for families living in the surrounding census tracts. We then estimate the within-school slope by regressing the within-school variance in achievement on the variance in income. Intuitively, the schools with more unequal incomes should also have proportionally more variance in achievement (and the slope of that relationship should be equivalent to the squared value of the within-school slope). We then combine the within- and between-school coefficients with an estimate of the proportion of variance in income attributable to schools (a measure of the degree of income segregation by school) to estimate the *student-level* relationship between achievement and income.

Because census tract boundaries do not correspond with the boundaries of school attendance zones, our imputations of the mean and variance in income by school are subject to measurement error. However, because the correlation between our imputed income measure and school-reported measures of free and reduced-price lunch receipt are stable over time, our finding of a declining slope is unlikely to be driven by measurement error.

As a robustness check, we use the March Current Population Survey to impute household income using the traits which are available for individual students in the NAEP: race/ethnicity, mother's education, urbanicity and state. We then investigate how the gaps in achievement associated with each of these traits compare with the changing gaps in family income corresponding to those same traits.

In contrast to Reardon (2011), we see no evidence of widening of achievement gaps in 4th grade or 8th grade math or reading. In fact, we see achievement gaps *narrowing* in 4th grade math and reading as well as in 8th grade math. For example, between 1990 and 2015, we estimate that the achievement of students at the 10th percentile of income improved substantially: a full standard deviation in 4th grade math and .68 standard deviations in 8th grade math. Although achievement also rose for higher income families, the achievement gap between those at the 90th and 10th income percentile *closed* by .4 and .15 standard deviations respectively in 4th and 8th grade math.   When we plot the mean achievement of each subgroup of students by race-ethnicity, mother's education, urbanicity and state against the mean income from the March CPS, we see consistent patterns of narrowing or stable achievement gaps between high and low-income groups. Our results challenge the prevailing view that gaps in achievement by family income have been widening.   In fact, just the opposite appears to be true.

II. **Prior Literature**

In a widely cited paper, Reardon (2011) combined survey data on parental income and achievement for multiple cohorts of children.   He reports that reading and math achievement gaps at the 90th and 10th family income percentiles grew by 40 to 50 percent among cohorts born between 1974 and 2001.

However, each of the longitudinal surveys used a different measure of achievement and a different measure of income.   Thus, to make comparisons among the different surveys, Reardon was forced to make a number of adjustments:

- First, because each survey measured student achievement at different ages (some measuring students at ages 1-6 and others measuring students at age 18) Reardon adjusted each survey for age differences.

- Second, the achievement measures had varying reliabilities, ranging from .75 to .96. As a result, Reardon standardized scores and adjusted by the estimated reliability for each assessment (multiplying scores by $\frac{1}{\sqrt{r_{test}}}$).
- Third, each survey measured income differently. Some used parent-reported income; for others, the income measures were reported by students. Moreover, depending on the survey, the number of income categories respondents could choose ranged from five to fifteen categories. For each survey, Reardon estimated the mean achievement in each income category and then used the percentage of students in each income category to infer the estimated achievement gap at the 90th and 10th income percentile. Finally, he adjusted by $\frac{1}{\sqrt{r_{income}}}$).

Although the surveys prior to 1975 often included student-reported family income, all of the surveys for birth cohorts after 1975 (the period during which Reardon estimates a rapid growth in income-based inequality) rely on parent-reported income.[1] Reardon assumes that all the parent-reported income measures had a reliability of .86. For studies with student reported income, he used different reliabilities for students of different ages, assuming a reliability of 0.50 for 9th grade reports, 0.57 for 10th grade, 0.65 for 11th grade, and 0.72 for 12th grade.

Nielsen (2019) analyzed data from the National Longitudinal Survey of Youth in 1979 and 1997 (two out of the 13 surveys analyzed by Reardon). Using fully ordinal methods (which do not rely on Reardon's assumption of a consistent linear scaling across the various achievement measures), Nielsen finds gaps narrowing. When he uses Reardon's methods and similar definition of income, he finds that

[1] For the birth cohorts for whom Reardon finds widening gaps in achievement by income, the income measures are parent reported. Thus, it seems unlikely that the finding is driven by the shift from student-reported income measures (which tend to be less reliable) to parent-reported measures.

gaps were stable. However, Reardon (2011) also finds the gaps to have been stable for those two surveys. It was on the basis of the other surveys he analyzes that he finds gaps widening.

In a recent paper, Hanushek, Peterson, Talpey, and Woessmann (2020) take a different approach, using achievement data from a different set of national and international tests (PISA, TIMMS, and the NAEP). Because none of the surveys include student-level measures of family income, the authors create an SES index-- using indicators of parental education and the number of home possessions as reported by students. Hanushek and colleagues find no change in reading or math achievement gaps between top and bottom quartiles of the SES distribution for 13-17-year old students born between 1954 and 2001.

However, the findings in Hanushek et al. (2020) of stable gaps in achievement by SES are not necessarily inconsistent with Reardon's findings. Although Reardon focused primarily on income-based gaps, he also reported differences in achievement by parental education (the main factor in the Hanushek et al. SES index.) Like Hanushek et al., Reardon reported that achievement gaps by parental education have remained relatively stable over the same time period. (One of our contributions below is to reconcile the findings in Reardon and Hanushek et al., by rescaling achievement gaps associated with mother's education by the average household income for mothers of varying education levels and race-ethnicity.)

Often motivated by concerns over rising income inequality or lawsuits targeting inequities in school funding, U.S. policymakers have launched a series of policies over the last three decades to raise achievement in low-scoring and low-income schools. For instance, Dee and Jacob (2011) studied the effect of school accountability policies, which were adopted in 30 states between 1990 and 2001 and then expanded nationwide with the No Child Left Behind Act in 2002. The strongest incentives were focused on lower performing schools and on subgroups of students by race/ethnicity and economic disadvantage. Using state by year and subgroup means from the National Assessment of Educational Progress, the authors find that school accountability policies had positive effects on average achievement in 4th and 8th grade math, and weaker effects on 4th grade reading. The effects were larger for African American and Hispanic students and students eligible for Free or Reduced Price Lunch, especially in 4th and 8th grade math.

Lafortune, Rothstein and Schanzenbach (2018) investigated the impact of school finance reforms (SFRs) on gaps in district spending and on achievement on the NAEP overtime. They focused on 64 major reform events across 26 states between 1990 and 2011. Leveraging the variation in timing of these reform events in an event-study framework, they find that SFRs were associated with an average total revenue increase of nearly $700 per pupil in the lowest-income quintile school districts relative to top-income quintile districts. Moreover, they estimate that the impact of reform events on the test score gap between bottom and top-quintile districts was 0.008 standard deviations per year, which they extrapolate to mean that SFRs reduced between-district achievement gaps in those states by 10% within ten years.

As the authors emphasize, the majority of the variance in income is within-districts and within-schools. As a result, district-level equalization does not necessarily translate to a narrowing of income-based achievement gaps. For instance, the authors find no impact of state SFR's on achievement gaps by student's free lunch status.

Over the same time period during which Reardon found a sharp widening in achievement by income, other social policy initiatives not usually seen as connected with test scores—such as the expansion of health coverage or the Earned Income Tax Credit— were improving health outcomes for low income families and narrowing gaps in child health measures.[2] Currie and Duque (2019) review the

[2] We thank Janet Currie for pointing this out.

literature on the impact of the Medicaid expansions, which occurred in the 1980's and 1990's on children's health and mortality. For instance, Currie and Schwandt (2016) find that expansions in public health insurance coverage have dramatically reduced mortality among poor children. A 2019 consensus report from the National Academy of Sciences, Engineering and Mathematics concluded that the Earned Income Tax Credit (EITC) Program have improved children's health outcomes. The EITC was enacted in 1975, but then sharply expanded in 1986, 1990 and 1993. Reardon estimated a sharp steepening in income-based achievement gap for those born after 1975. If that were true, it would imply that gaps in children's health and educational outcomes were moving in opposite directions.

III. Empirical Strategy

We employ two different strategies to estimate the student-level relationship between income and achievement. First, we use school-level aggregates to estimate the student-level relationship between income and achievement. Second, as a robustness check, we use the March Current Population survey in each year to impute household income by students' race/ethnicity, parental education and geographic location. In doing so, we essentially scale the gaps in achievement by race, mother's education, urbanicity and state of residence by the gaps in mean household income associated with those same characteristics. We then test whether gaps in achievement grew by imputed income. We briefly describe each of those two approaches below.

Using School Level Aggregates to Estimate the Student-Level Parameters

Our goal is to estimate the student-level relationship between achievement and income:

$$(1) \quad Score_{ij} = \beta_0 + \beta_1 Y_{ij} + v_{ij}$$

Where $Score_{ij}$ is a student's scaled test score, Y_{ij} is a measure of log household income and i and j are subscripts for the student and the student's school respectively. (β_1 is a descriptive relationship, not a causal one.)

It can be shown (Raudenbush and Bryk, 2002, pp. 136-137) that the OLS estimator for β_1 is a weighted average of the between-school relationship of mean achievement on mean achievement ($\hat{\beta}_{between}$) and the within-school relationship between an individual students' household income and achievement ($\hat{\beta}_{within}$):

$$(2)\ \hat{\beta}_1^{OLS} = \hat{\beta}_{between} * ICC_Y + \hat{\beta}_{within} * (1 - ICC_Y)$$

Where ICC_Y is the percent of the total variance in log household income that is attributable to school-level differences in mean log income ($ICC_Y = \frac{\sigma^2_{\bar{Y}_j}}{\sigma^2_{Y_{ij}}}$). The statistic ICC_Y is the empirical analog of the intra-class correlation in log income by school.[3] With a value between 0 and 1, ICC_Y is a measure of the degree of income segregation by school. That is, if all of the variation in income was by school (every student in a given school had the same income) and different schools had different incomes, then the student level relationship between achievement and income would be equal to the between-school relationship. As implied by equation (2), the more segregated schools are by household income, the more the student-level relationship will reflect the between-school relationship.

[3] Bryk and Raudenbush (1992) refer to the statistic, ICC_Y, as "eta squared". In contrast to the traditional "intra-class correlation", or ICC, which uses an estimate of the variance of the true school mean incomes in the numerator, μ_{Y_j}, the statistic, ICC_Y, uses the variance in the sample means, $\bar{Y}_j$(which has more variance because $\bar{Y}_j$ is equal to μ_{Y_j} plus estimation error). In a setting with a large sample (or the entire population) for each school, the two measures are approximately equivalent.

With school-level measures of mean achievement and mean log income, $\overline{Score_j}$ and $\bar{Y}_j$, we can estimate the between-school relationship:

$$(3) \quad \overline{Score_j} = \beta_0 + \beta_{between}\bar{Y}_j + \varepsilon_j$$

Although we do not have the covariance of achievement and income at the student level, we can estimate the within-school slope using the school-level variance in achievement and income, $\hat{\sigma}^2_{score_j}$ and $\hat{\sigma}^2_{Y_j}$:

$$(4) \quad \hat{\sigma}^2_{score_j} = \sigma^2_v + \beta^2_{within}\hat{\sigma}^2_{Y_j} + \vartheta_j$$

The basic intuition behind equation (4) is that schools with higher income variance (i.e. more income inequality) should have proportionally higher variance in achievement-- with the proportion determined by the square of the within-school slope coefficient. In other words, although we do not see the covariance in income and achievement within schools, we can infer the within-school slope by studying the changing relationship between within-school income variance and within-school achievement variance.

In estimating equations (3) and (4), we weight by the number of tested students in each school. We also allow the errors to be clustered by the geographic primary sampling unit used by NAEP. Given that the error terms, ε_j and ϑ_j, could be related, we estimate equations (3) and (4) as seemingly unrelated regressions (SURE), thus estimating the covariance between $\beta_{between}$ and β^2_{within}. Finally, as described in equation (2), the student level slope is a nonlinear function of the coefficients in equations (3) and (4), $\beta_{between}$ and β^2_{within}. Thus, we use the delta method to estimate a standard error for our estimate of the student-level slope ($\hat{\beta}_1^{OLS}$).

In equation (4), we express β_{within} as a scalar, assuming that the within-school slope is the same in all schools. However, the above model could be generalized to allow the slope to be a random coefficient, varying by school. If the within-school slope did vary by school, but was independent of log income, then the coefficient on $\hat{\sigma}^2_{Y_j}$ in equation (4) would be interpreted as the sum of the expected value of β_{within} squared and the variance in β_{within} (that is, $E[\beta_{within}]^2 + \sigma^2_{\beta_{within}}$.) As long as the variance in β_{within} ($\sigma^2_{\beta_{within}}$) is constant or rising, then any increase in the expected value of β_{within} should lead to an increase in the coefficient on $\hat{\sigma}^2_{Y_j}$in equation (4). It's only if the variance in within school slopes is declining that we might not detect an increase in the within-school slope using our method.

Using Imputed Household Income for Individual Students

As a validity check to our census-based income measures, we produce a second set of estimates, using student-reported race and mother's education as well as school location (urbanicity and state) to impute log household income using the March Current Population Survey. We then estimate the student-level slope, as well as the between and within-school slopes, using imputed household income for each year. (We can estimate all three directly since we have imputed household income at the student-level.)

Students participating in the 8th grade NAEP assessments were asked to report education levels for their mother as well as their father. (The survey measures of parents' education were not consistently available for the 4th grade students in NAEP.) However, the NAEP did not consistently ask whether either parent lives in the same household with the student—which would have been helpful in imputing family income from the March CPS. Therefore, we assume that students in the NAEP lived in households where their mother was present in order to be able to impute household income by mother's education. (We do not use information about the father's education, given that we cannot know whether the father is present.) We estimate the following relationship for all households in the Current Population Survey with dependent children between age 5 and 18 with a mother present:

$$(5) \quad Y_{it} = \sum_{race/eth} \sum_{moth\,ed} \pi_{race,moth\,ed\,t} + \gamma_{urban,\,t} + \delta_{state,t} + \varphi_{it}$$

Where Y_{it} is a measure of log household income, $\pi_{race,moth\,ed\,t}$ represent a full set of interactions between four mother's education categories and four categories for race/ethnicity, $\gamma_{urban,\,t}$reflects differences in income by four urban categories and $\delta_{state,t}$ represents fixed effects by state and year. We assume that the student has the same race/ethnicity as the head of household.

Using the parameter estimates from equation (5), we generate a predicted household income for each 8th grade respondent with non-missing mother's education, $\hat{Y}_{it}$. Since we are using imputed income, $\hat{Y}_{it}$, to substitute for actual household income, we would rewrite equation (1) as:

$$(1') \quad Score_{it} = \beta_{0t} + \beta_{1t}\hat{Y}_{it} + (\beta_{1t}\hat{\varphi}_{it} + v_{it})$$

Where $Y_{it} = \hat{Y}_{it}$+$\hat{\varphi}_{it}$. Note that as long as $\hat{Y}_{it}$ is estimated by OLS, then $\hat{Y}_{it}$ and $\hat{\varphi}_{it}$ will be orthogonal by construction. In other words, as in two-stage least squares, even though imputed household income is an imperfect measure of true household income, our estimate of β_{1t} in equation (1') will not be biased due to measurement error (although we do not claim that it reflects a causal relationship). Many of the variables being used for imputation—mother's education, race, state, urbanicity—are likely to have their own direct effects on achievement beyond their influence on family income (that is, $\hat{Y}_{it}$ will be correlated with the error term from equation (1), v_{it}.) For that reason, they are unlikely to be valid instruments. However, in order to infer the change in the descriptive relationship between income and achievement, we must assume that the correlation between $\hat{Y}_{it}$ and v_{it} is constant over time and that any change in $\hat{\beta}_{1t}$ is due to the effect of income and not to changes in the direct effects of the variables used for the imputation.)

Just as with the usual 2SLS estimate, equation 1’ will estimate a LATE that may differ from that estimated by OLS. Recall from equation 2 that $\hat{\beta}_1^{OLS}$ is a weighted average of the within-school and between-school slopes, where the weight on the between-school slope is ICC_Y (the intra class correlation in log income by school). Similarly, we can restate the estimate using from equation 1’ using imputed log income (call this $\hat{\beta}_1^{\hat{Y}}$) as:

$$(2')\ \ \hat{\beta}_1^{\hat{Y}} = \hat{\beta}_{between}^{\hat{Y}} * ICC_{\hat{Y}} + \hat{\beta}_{within}^{\hat{Y}} * (1 - ICC_{\hat{Y}})$$

If we assume that $\overline{\hat{Y}_\iota}$ and $\overline{\hat{\varphi}_\iota}$ (the school-level means of $\hat{Y}_{it}$ and $\hat{\varphi}_{it}$) are orthogonal (an additional assumption not guaranteed by OLS), then it can be shown that $\hat{\beta}_{between}^{\hat{Y}} = \hat{\beta}_{between}, \hat{\beta}_{within}^{\hat{Y}} = \hat{\beta}_{within}$ even when $ICC_Y \neq ICC_{\hat{Y}}$. In other words, if the error in equation (5) is uncorrelated with the proxy measure of income at both the individual level (true by construction) and at the school level (an assumption), then the within and between slopes are the same for the true income data as for the imputed, proxy income measures. In this case, our estimate of the overall relationship between scores and log income using actual income (equation 2) differs from our estimate using imputed income (equation 2’) because the two estimates put different weight on the between and within slopes (since in general $ICC_Y \neq ICC_{\hat{Y}}$).

Therefore, although we estimate equation (1’) with student-level data, we also estimate the between-school and within-school relationships and the ICC using imputed student level data for comparison to the estimates above. We then evaluate whether any difference in our estimates of β_{1t} between our two methods are due to differences in the within- and between-school estimates (which should be similar) or simply differences in the weighting places on within versus between estimates due to differences in the ICC (which are expected).

Our imputation using mother’s education and race uses much of the same student-level information as that used by Hanushek et al. (2020). There are some important differences: in addition to mother's education and race, Hanushek et al. use several items reflecting household possessions (which we exclude given their absence from the CPS), while we use urbanicity and state. Moreover, while we scale the information on student-reported characteristics by the income differentials associated with those traits, Hanushek et al. (2020) scale the differences in student characteristics using loadings from a factor analysis. In doing so, they are focused on achievement differentials by socio-economic status, a latent construct. Because the measures of household possessions they use to construct the SES index varies by year, that construct will have an unknown—and potentially changing—relationship to wealth and income over time. Since Reardon (2011) focused on differentials in achievement by income, our approach presents a more direct test of his conclusions.

IV. Data

To measure student achievement over time, we use student level data from the trial state assessments in 1990, 1992 and 1994, and from the state-representative samples in subsequent years. Known as "the Nation's Report Card," the main NAEP assessment provides a scaled score which is designed to be comparable over time. The samples are intended to provide a representative sample of schools and roughly 30 students in each school. Prior to 2002, states were able to opt into participating and the number of participating states ranged from 38 to 46, depending upon the subject and grade. Since 2003, the NAEP has been administered in every state, every other year, in both reading and mathematics.

Balancing the goal of maintaining a consistent sample of states with the goal of remaining as close to nationally representative as possible, we included states that missed at most one year of assessment scores in a given grade-subject combination. Thus, in 4th grade math and reading, that left us with 45 and 41 states respectively, and 40 and 46 states in 8th grade math and reading. We report the number of schools and students by grade/subject and by year in Table 1. For example, in 2015, we included scores from 108,290 students attending 5,250 schools in 4th grade reading.

To add data on household income for public schools, we use information on school locations from the Common Core of Data. In 1990 and 2000, we rely on the decennial censuses. For 2009 and after, we use the American Community Survey. The NCES's Common Core of Data (CCD) provides student enrollments as well as the latitude and longitude for each school. Using each school's geographical coordinates, we rank the nearest block groups (as measured by block group centroid) by their distance from each school.[4] The census data report the number of related children attending public schools in grades 1-4 and in grades 5-8 in each block group. Combining the census data on number of children enrolled in various grade levels with data from the CCD on the enrollment of each school, we identify the K nearest block groups with a sufficient number of children to "fill up" the school's official enrollment. In other words, if N_k is the number of students enrolled in public schools from block group *k* according to the census, and N_j is the number of students enrolled in school *j* according to the CCD, then we associate the nearest *K* block groups that satisfy the inequality $\sum_{k=1}^{K-1} N_k < N_j \leq \sum_{k=1}^{K} N_k$ where block groups *k* are ordered in increasing distance from the school location.

In the census, detailed data on household income is only available at the census tract level—not the block-group level. As a result, we assume that the distribution of income in each block group matches the distribution in its associated census tract. We then weight the tract-level data by the counts of students in each of the K block groups associated with the school. Thus, the number of students in school *j* who come from a household with income in the i^{th} reported bin is given by

[4]Overlaying school attendance zones and block group boundaries for a set of 21 of the largest 22 districts, Saporito and Sohoni (2007) found that 53 percent of census block groups in 2000 lay entirely within a school boundary in 2007.

$$N_{ji} = \sum_{k=1}^{K} \frac{p_{ki} n_k}{n_j}$$

Where $p_{ki}$is the proportion of students in the corresponding census tract in income bin *i*. These counts can be thought of as a coarsened distribution of the household income for students in each school[5]. We merge our estimates for each school using a unique school identifier.

To infer the mean and variance in income from the binned income data, we assume the distribution of income for each school is approximately log normal. By definition, if the log of income, ln(Y), is distributed normally, and C_l is the upper range of the lth income bin (expressed in log 2016 dollars), then the proportion of the school enrollment with income below can be expressed as follows:

$$P_{C_l} = \Phi\,(\frac{\ln(C) - \mu_{Y_j}}{\sigma_{Y_j}})$$

Where μ_{Y_j} and $\sigma_{Y_j}{}^2$ represent the mean and variance of log income in school j.

Accordingly, using for the top limit in each income bin (C_l, where l=1,...,15 corresponding to 16 income bins) and the proportion of households, P_{C_l}, in the school with incomes below C_l, we estimate the following regression for each school:

$$\ln(C_l) = \mu_{Y_j} + \sigma_{Y_j} invnorm\left(P_{C_{jl}}\right) + \vartheta_{jl}$$

Where j subscripts the school and l subscripts one of 15 income bins.

For inter-censal years and for the years between 2000 and 2009 (when the American Community Survey became available), we use a linear interpolation of the mean and variance of income from 1990, 2000 and 2009.

[5] To calculate the mean and variance of this distribution we use a method outlined by vonHippel, Scarpino, & Drown (2016), fitting the cumulative CDF of each school's income bins to a log-Normal distribution

Our use of Census data to approximate the mean and variance in family income for individual public schools will suffer from two sources of measurement error: First, Census block group boundaries will not coincide with actual school attendance zones; second, not all children in a neighborhood will be attending public schools. According to the National Center for Education Statistics, 13.1 percent of students in grades Pre-K through 8 attended private schools in 1990 (the beginning of our period). The percentage had declined slightly to 10.8 percent by 2016 (National Center for Education Statistics 2020).

Nevertheless, the method performs reasonably well in replicating administrative data on race and income from the Common Core of Data (CCD). To check, we compare our Census-based estimates of poverty rates and race/ethnicity to the proportion of students receiving free or reduced price lunch (FRPL) and each race/ethnic group in the CCD. Those results are reported in Table 2. Unfortunately, the data on students receiving free or reduced price lunch was missing for more than half of schools in 1990 (52 percent). As a result, the first year we report the correlation is in 2000 (when 13 percent of schools were missing data). In 2000, the correlation between the poverty rates in nearby block-groups and the school's reported percentage of students receiving free or reduced price lunch was .715. By 2015, the correlation was roughly the same, .690.[6] The method performs even better for approximating a school's race-ethnicity; the correlation between our estimate of the percentage of students who are black in each school and the CCD report was roughly .9 in all years between 1990 and 2015. Although the census-based measure of school income is subject to error, there is no evidence that the relationship has diminished over time or that the reliability has gotten worse.

In order to impute household income for various student traits, we identify a subset of variables available in both the NAEP and the March Current Population Survey: mother's highest level of

---

[6] The imputation performs best in suburban areas, with an R^2 of approximately .75 with respect to FRPL status, but R^2 between .6 and .65 in the urban and rural areas as well as towns. The correlations for primary schools (those with 4th grade students) and middle schools (which we define as those with 8th grade students) were also similar and consistent over time.

educational attainment (5 categories for less than high school, high school graduate, some college, college graduate and unknown), race-ethnicity (6 categories for white, Black, Asian, Native American, Hispanic and other), urbanicity (3 categories for urban, suburban and rural schools) and state. We limit the CPS sample in each year to households with a related child aged 5 through 18 and a mother present. Table 3 reports the R^2 for each specification for each year. Although the traits account for a small share of the variance in household income in each year (R^2 in the combined model is approximately .22), the proportion of variance explained is stable over time.

V. Results Using School-Level Data to Estimate Student-Level Parameters

Before presenting parameter estimates, we start with a graphical summary of the between-school and within-school relationships in various years. Figure 1 reports the results of a locally weighted polynomial regression of school mean achievement on mean log income for 4th grade math and reading. The solid lines portray the relationship for the first year (1992) and the last year (2015) in our analyses; the dashed lines report the relationships for the years between 1992 and 2015. Upon inspection, the relationship is roughly linear in log income. Moreover, in both subjects—but especially in math—there has been a noticeable increase in the intercept and a decline in the slope. At least in terms of the differences in mean achievement between schools, there is no evidence of a sharp steepening of the relationship between mean achievement and mean income between schools.

Figure 2 reports similar relationships for 8th grade math and reading. In 4th grade and 8th grade math, there were sharp increases in mean achievement for schools at all levels of mean log income in math. Likewise, there were increases in achievement for schools all income levels in 4th grade reading. However, in 8th grade reading, the shift was much smaller than in the remaining three grade/subject combinations. In no subject-grade combination is there evidence of a steepening of the relationship mean achievement and mean log income at the school level.

Figure 3 reports the trend over time in the intra-class correlation in log income by school. Using the ICC metric, there's been a 20 percent increase in income-based school segregation over time (from .19 to .22.) Figure 3 is consistent with prior research on income-based residential segregation. Using a sample of the 100 largest metropolitan areas, Reardon and Bischoff (2011) find that income segregation grew rapidly between 1970 and 2000. (Watson (2009) finds consistent results using a different measure of income segregation.) Owens, Reardon, and Jencks (2016) find that between-district income segregation of families with children enrolled in public school increased by over 15% from 1990 to 2010. They also find that within-district segregation based on students' free-lunch status increased by over 40% between 1990 and 2012.

Figure 4 uses school-level data to plot the relationship between the within-school variances in achievement and within-school variances in log income for grade 4 math. Under the assumption that the within-school slope is the same across schools, the slope of the relationship in Figure 4 should be equivalent to the square of the within-school slope of achievement on log income. Sample-based estimates of variances are inherently noisy, so to summarize the central tendency, we have calculated the average variance in achievement and the average variance in income for 20 equal-sized bins arranged by variance in log income. We have plotted the mean variance in achievement for each bin and then fitted a line. The relationship does appear to be roughly linear, with an increase in achievement variance being approximately proportional to the rise in log income variance. However, there is again no evidence of a steepening slope. If anything, the slope appears to be declining.

Figures 5 through 7 report analogous figures for 4th grade reading as well as 8th grade math and reading. In all cases, the relationship between the variance in achievement and the variance in income is either flat or declining.

As reported above, there is no evidence of a steepening of the between-school relationship between mean achievement and mean log income. It would still be possible for the student-level slope to rise, if there were a sufficient increase in the ICC (increasing the weight on the between-school slope which is larger than the within-school slope) or if the within-school slope were rising. However, equation (1) lets us bound the magnitude of the improvement one would have to see. If there has been no change in the between-school slope and the ICC is roughly .2, a 40 percent increase in the overall slope that Reardon (2011) infers would have required a 50 percent increase in the within-school slope. [7] Even if we cannot estimate it precisely, we see no evidence of a 50 percent increase in the average within-school slope.

In Tables 4 through 7, we report our estimates of the between-school and within-school slope parameters and then use the estimated ICC to reconstruct the student-level relationship in each year. Table 4 contains our estimates for 4th grade math. As reported in column (1), between 1992 and 2015, our point estimates of the between-school slope decline from 27.9 to 20.9, approximately a 25 percent decline. That is a statistically significant difference. In column (2) we report estimates of the within-school slope parameter. The within-school coefficient also declines from 14.7 to 7.2. This decline too is statistically significant. As reported in column (3), the intra-class correlation in ln income, which is a measure of income-based segregation, rises from .19 to .23. Because the between-school slope is larger than the within-school slope, an increase in the ICC would have led to a small increase in the student-

[7] A 50 percent increase in the within-school slope with respect to log income would have required a more than doubling of the slope in Figures 4 through 7, since $1.5^2=2.25$.

level slope if the between- and within-school coefficients had both remained constant. However, the rise in ICC was not nearly large enough to offset the declining between-school and within-school slope.

In column 4, we combine the ICC, the between-school slope and the within-school slope to estimate the student-level slope parameter. Our results imply that the student-level slope declined from 17.2 to 10.3 between 1992 and 2015, a statistically significant difference. In other words, a 10 percentage point difference in family income was associated with a 1.7 point higher score in 4th grade math (.05 standard deviations) in 1992.[8] By 2015, that had declined to 1 point or .03 standard deviations.

In the remaining columns of Table 4, we report the implied change in mean achievement at the 10th and 90th percentile of log household income. (Recall that we do not actually observe any individual student at either percentile. Thus, we are estimating these by plugging in the relevant percentiles of household income from March CPS data and using our student-level slope and intercept parameters to estimate mean achievement.) Reardon (2011) makes a distinction between the student-level slope with respect to the level of log-income and the effect of rising inequality (that is, the widening gap in real incomes between those at the 10th and 90th percentiles.) He concludes that most of the change over time is due to a steepening in the slope rather than simply a widening gap in real incomes between those at the 90th and 10th percentiles. We also find that the declining slope was a more important driver of widening gaps between rich and poor than the widening income distribution itself. To isolate these effects, we estimate gaps holding the 10th and 90th percentiles of log income constant at 1990 levels, thus isolating the effect of the changing relationship between income and achievement. We then compare this to the test score gap between students at these percentiles for the year that they took the

[8] The standard deviation in 4th grade math was 32 points in the 1992 NAEP.

NAEP. This estimate combines the effects of the expanding income distribution and the changing relationship between income and test scores.

In column (5) we report the estimated mean achievement at the 10^{th} percentile of family income in 1990. We estimate that there has been a 30.2 point increase in the mean 4^{th} grade NAEP score for students from families who had incomes that placed them in 10^{th} percentile in 1990, from 199.7 to 229.9. This is equivalent to approximately a full standard deviation increase in 4^{th} grade math achievement.

In column (6) we report the difference between students at the 90^{th} and 10^{th} percentile of 1990 household income. Using the estimated slope and intercept for 1992, the gap in achievement from students at the 90^{th} and 10^{th} percentile of incomes from 1990 was 35.3 points or 1.1 standard deviations. By 2015, the implied gap at the 90^{th} and 10^{th} percentile of income in 1990 had shrunk to 21.2 points—or .71 standard deviations.

In column (7) we combine the effects of the changing income distribution and the changing slope by allowing the 90^{th} and 10^{th} income percentiles to vary by year. (The 90-10 gaps in income were widening as incomes were becoming more unequal.) In 2015, the gaps would have been 2.5 scaled score points wider using the 2015 income percentiles rather than the 1990 income percentiles. In other words, the primary driver of the closing of the income-based achievement gap has been the flattening relationship between income and test scores, which has only been partially offset by the widening in the income distribution.

Tables 5 through 7 report similar estimates for 4^{th} grade reading, 8^{th} grade math and 8^{th} grade reading. In 4^{th} grade reading, we estimated that the variance in achievement had a small negative slope with respect to the variance in log incomes in the three latter years—2011, 2013 and 2015. We interpret a negative slope at the boundary constraint of zero. (Because the estimated slope was negative, the square root was not identified.) However, the between-school slope for 4th grade reading declined by 7 percent between 1992 and 2015 (from 26.0 to 24.2) and by 22 percent relative to its peak in 2002). As reported in column 7, the difference in predicted achievement at the 90th and 10th percentiles of income in each year closed from 42.5 points (1.2 standard deviations) to 12.8 points (.35 standard deviations).

As reported in Table 6, the between-school slope with respect to log income declined by 11 percent in 8th grade math, and the within-school slope declined by 28 percent from 1990 to 2015. We also estimate that mean 8th grade math achievement at the 10th percentile of household income from 1990 increased by .7 standard deviations, from 240.2 to 264.5. Moreover, as reported in column 7, the gap in achievement between high and low income students declined from 42.9 to 37.5 points, a narrowing of .15 standard deviations.

For the early birth cohorts, our estimates of the income-based achievement gaps are similar in magnitude to those reported by Reardon. For instance, Reardon (2011) reports that size of the achievement gap at the 10th and 90th percentile of income was 1 standard deviation for the 1976 cohort. In Table 6, we estimate the 90-10 gap for the same cohort (the 8th graders in 1990) to be 42.9 scaled score points or 1.19 standard deviations in 8th grade math. However, between the 1976 and 2001 birth cohorts, Reardon estimates that the math achievement gap at the 90th and 10th percentiles grew sharply from 1 to 1.4 standard deviations, while our estimates suggest that the gap in 8th grade math achievement declined from 1.19 standard deviations to 1.01 standard deviations (37.5 points in the bottom row of Table 6).

Our estimates of the racial achievement gaps also start out comparable: Reardon estimates that the black-white gap was .8 standard deviations for the birth cohort of 1976; in the NAEP data, the black-white gap for 8th grade students in 1990 was scores was 34 points, or .94 standard deviations.

However, unlike with income, our findings on the subsequent changes in the black-white achievement gap largely correspond with those reported by Reardon. He finds that the racial achievement gap closed from .8 to .7 standard deviations for the birth cohorts born of 1976 and 2001, while the NAEP data show the gap closing from .94 to .83 standard deviations over the same period in 8th grade math.

Reardon highlighted his finding that gaps in achievement by income and by race appeared to be moving in opposite directions: widening by income and closing by race. It was a puzzle, given the well-established relationship between household income and race. However, while we find the same narrowing of the racial achievement gap, we find no such puzzle: our results suggest that the income gaps were moving in the same direction. In the next section, we explore further the change over time in achievement gaps with other traits correlated with income, such as mother's education.

VI. Results Using Imputed Household Income

Using the Census-based measure of schools' mean income and variance, our results above suggest that income-based achievement gaps have been narrowing or stable since 1990, even as mean achievement in 4th grade math and reading and 8th grade math have been rising. In this section, we investigate whether there has been any change in the achievement gaps associated with student-level characteristics available in the NAEP that are correlated with income: race/ethnicity, mother's education, urbanicity and state of residence. We use the March Current Population Survey to estimate mean family income for each combination of those traits (as described in equation (5)).

In Table 8, we report the coefficients on imputed income using each instrument separately (race/ethnicity, mother's education, urbanicity and state) and then combined (using interactions between race/ethnicity and mother's education.) Essentially, by instrumenting for income with the student level traits, we are scaling the differentials in achievement associated with each of these traits by the magnitude of the associated family income differential from the March CPS. Of course, if there is a direct effect of a given trait on achievement that operates outside of income, we will be overstating the impact of income on achievement in each year. (We are not claiming that these are valid instruments without a direct effect.) However, the trends over time should be consistent with the trends in underlying income effects as long as the direct effects of race, mother's education, urbanicity/state are constant. The implied slope with respect to log income for 8th grade math based on the race differentials declined by 18 percent between 1990 and 2015, from 51.1 to 42.1. Similarly, using mother's education as the instrument, the implied slope with respect to log income declined by 16 percent. Also, using state and urbanicity as the instruments, the implied slope declined by 30 percent, from 25.2 to 17.6—largely as many low-income states closed the gap with respect to higher income states. When we combine all of the measures, the implied slope declined by 13 percent, from 32.9 to 28.6. Recall from Table 6, the slope with respect to log income using our school-level measures of mean and variance in income, declined by 22 percent, from 20.9 to 16.4.

In Figure 8, we plot the mean 8th grade math score (from the NAEP) and the mean family income (from the March CPS) for each observed combination of race by mother's education by urbanicity by state. To provide a more concrete sense of what we are doing, we call out two specific points: children of Hispanic, high school drop-outs from rural New Mexico and children of white college-graduates from suburban New York. We then estimate the relationship between mean achievement and mean income for each of those subgroups. We report the same relationship in 1990 and 2015. Among 8th grade students, mean achievement rose at all income levels, but particularly for low income groups, with the implied relationship between achievement and income flattening.

In Figure 8, we also plot the points corresponding to mean achievement and mean income for black and white students as a whole. Reardon (2011) notes that the black and white achievement gap narrowed somewhat over the time period in which he estimated that the income slope coefficient rose sharply. He reconciles the two facts by noting that many black families have incomes below the median. In his estimates, the gap in achievement between the 90^{th} and 50^{th} percentile of income increased more than the gap between the 50^{th} and 10^{th} income percentile. However, as we find in Figure 8, the narrowing of gaps also seemed to occur between the highest income groups (e.g. the children of white, suburban college graduates in suburban New York) and the median household, as well as between the lower income groups (e.g. the children of Hispanic high school drop-outs in rural New Mexico.) There is no evidence that the slope increased within either income range.

In Figure 9, we report analogous results for 8^{th} grade reading. The time span is somewhat different—the first year of 8^{th} grade reading scores in the state trial assessment occurred in 1998. Nevertheless, the results are similar to those reported earlier—the slope with respect to income remained fairly constant, with no evidence of a sharp widening.

VII. Comparing the Coefficients from the Two Methods

In Figures 10 and 11, we report the trend in the student-level relationships for the two sets of estimates in 8^{th} grade math and reading: using the census-derived school income measures and the imputed income measures, using race/ethnicity, mother's education, urbanicity and state. (Recall that the imputed income measures are only available in 8^{th} grade.) The grey lines represent the trend in the overall slope (combining the within and between-school slopes). The slope estimates based on imputed income are larger than those based on the census-derived school income measures—possibly because the coefficients using imputed income are not attenuated by measurement error. However, the time trend is very similar: a stable slope during much of the 1990s, with a decline starting in 2005. Although the series for 8^{th} grade reading does not start until 1998, the pattern is very similar, with the slopes based on imputed income exceeding those based on census-derived income, with both gradually declining over time.

Above, we highlighted the strong assumption required to estimate the within-school slope coefficient with the census-derived income measure, namely that schools have the same within-school slopes. As a robustness check, we estimated the between-school and within-school slope coefficients using the imputed family income measures. We wanted to know if we saw the same pattern of flattening in both the between school and within-school coefficients. We find a similar time trend in the within-school relationships using imputed income as we found with the variance on variance regressions: the within-school relationships between income and achievement were stable during the 1990's and declining after 2005.

VII. Private Schools

Thus far, our analysis has been limited to public schools. Yet, as reported by Murnane and Reardon (2018), children in families at the 90th income percentile are roughly 4 times as likely to enroll in a private school than students at the 10th percentile of family income (roughly 18 percent versus 4 percent during the time period we are studying.) So, the question is, might the rise in the gap in achievement by income that Reardon (2011) reported be driven by a sharp widening of the gap in achievement between public schools and private schools?

There's little reason to believe that our conclusions would be any different if we had been able to include student-level data on income and achievement from private schools. In Figure 12, we report the trend in achievement for the average public school and the average private school for grade 4 reading and math. We also report the achievement of the 90th percentile private school, recognizing that high-income families are unlikely to attend the average private school. For each grade-subject combination, the public-private gap in mean achievement has been narrowing over time. In addition, the gap between the average public school and the 90th percentile private school has also been narrowing. Figure 13 reports similar results for 8th grade students. Again, there is no evidence of widening.

VIII. Conclusion

The ideal data for measuring the change in income-based achievement gaps would combine a consistent measure of student achievement with a reliable measure of each student's parental income over a long period of time. Unfortunately, the federal government has not collected such data. Thus any effort to shed light on changes in income-based achievement gaps will necessarily involve compromises.

Reardon (2011) combined surveys from different time periods with different measures of academic achievement and different measures of income over time. Although the methods he uses are reasonable, there are obvious hazards in trying to make comparisons over time with measures that are so different. The achievement measures varied in terms of content coverage, age group and reliability. The income measures varied as well, containing different numbers of categories, with some reported by students and some reported by parents.

Our approach to estimating the student-level relationship with school-level aggregates has its own shortcomings. For instance, we have had to rely on block-group boundaries and not actual school attendance zones to infer the mean and variance in income in each school. Although the reliability seems not to have changed over time, we know that the Census-based measures are subject to error. Moreover, we have assumed that the within-school relationship between income and achievement is the same across schools in any given year (although our evidence would also be consistent with a declining mean slope with random coefficients.)[9]

However, our two sets of findings— first, using school-level aggregates from the Census to estimate the student-level relationship and, second, scaling the student-level traits by differences in household income (race/ethnicity, mother's education, urbanicity and state)—are consistent. In contrast, Reardon's finding of sharply widening achievement gaps must be reconciled with the narrowing gaps by race/ethnicity, mother's education, urbanicity and state.

Reardon (2011) noted the potential anomaly and offered an explanation. He finds that the steepening slopes with respect to income occurred primarily at the top end, for those with above-median income. He points out that most African American families have incomes below the median. However, that explanation cannot reconcile declining gaps for those with high and low levels of mother's education, or for those living in high- and low-income states, whose mean incomes span the median. In contrast, the time trend in our estimates of the student-level relationships matches very well with the time trend for the imputed gaps by race, mother's education, urbanicity and state. In our two approaches to measuring changes in income-based achievement differentials, there is no difference to reconcile. Using both methods, we find that the income-based gaps are narrowing in three out of four grade/subject combinations we study. In the remaining grade/subject, 8th grade reading, the income-based gap is constant.

In some grades and subjects, the implied achievement gains at the 10th percentile of the family income distribution have been remarkable: a full standard deviation in 4th grade math and .67 standard deviation in 8th grade math. We cannot say how much of the improvement has been due to accountability reform, school finance reform or other changes, such as students becoming more

---

[9] This would be true as long as there was not a more than offsetting decline in the variance in coefficients.

comfortable with standardized testing. Although the previous literature has suggested that both accountability and school finance reforms have had an effect, they have been separate literatures—one focused on the effect of accountability laws (e.g. Dee and Jacob (2011), Hanushek and Raymond (2005)) and another focused on school finance reforms (Lafortune, Rothstein and Schanzenbach (2018) and Jackson, Johnson and Persico (2016)). However, many of the same states were implementing both school finance reforms and accountability at the same time. In fact, the accountability reforms and the school finance reforms were often explicitly linked as a way of ensuring that the additional dollars were to be spent well. In future work, we will be using the estimated shifts in intercepts and slopes of the student level-relationship between achievement and income to sort out the relative effect of different reforms in explaining the substantial increases in achievement since 1990.

Despite the substantial increases in 4th and 8th grade math and in 4th grade reading achievement, Blagg and Chingos (2016) document that such progress has not translated into improved 12th grade scores for the same cohorts. On one hand, the trend in 12th grade NAEP scores—like the trend in SAT scores-- is likely biased downward by the rise in high school graduation rates over time. On the other hand, compulsory schooling laws typically require students to remain enrolled in school until age 16 and, yet, the scores of U.S. 15 year-olds have essentially been unchanged since 2000 on the OECD's Programme in International Student Assessment (PISA). It is possible that the No Child Left Behind Act—which focused on student achievement in grades 3 through 8 and only required schools to test in one grade in high school—simply shifted educational investments to earlier grades, affecting the timing of learning, rather than improving the stock of math and literacy skills students acquire before entering college or the labor market. Thus, in future work, we will study whether improvements in mean math and reading scores in 4th and 8th grade by state, gender and race/ethnicity translated into improved earnings among these same cohorts as adults.

References:

Blagg, Kristin and Matthew M. Chingos, "Varsity Blues: Are High School Students Being Left Behind?" Urban Institute Research Report, May 2016

Bryk, Anthony S., and Stephen W. Raudenbush. *Hierarchical linear models: Applications and data analysis methods*. Sage Publications, Inc, 1992.

Currie, Janet and Valentina Duque. "Medicaid: What Does It Do, and Can We Do It Better?" *The ANNALS of the American Academy of Political and Social Science* (2019) *686*(1), 148–179.

Currie, Janet and Hannes Schwandt. "Inequality in mortality decreased among the young while increasing for older adults, 1990–2010," Science, Vol. 352, No. 6286, April 2016, pp. 708-712.

Dee, Thomas S. and Brian Jacob. "The impact of No Child Left Behind on student achievement." *Journal of Policy Analysis and Management (2011) 30*(3), 418-446.

Hanushek, Eric A. and Margaret E. Raymond. "Does Accountability Lead to Improved Student Performance?" *Journal of Policy Analysis and Management* (2005) 24, pp. 297-327.

Hanushek, Eric A., Paul E. Peterson, Laura Talpey and Ludger Woessmann. "The Unwavering SES Achievement Gap: Trends in U.S. Student Performance." NBER Working Paper No. 25648, March 2019 (Updated March 2020).

Jackson, C. Kirabo, Rucker C. Johnson and Claudia Persico. "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms." *The Quarterly Journal of Economics*, Volume 131, Issue 1, February 2016, Pages 157–218.

Lafortune, Julien, Jesse Rothstein and Diane Whitmore Schanzenbach. "School Finance Reform and the Distribution of Student Achievement." *American Economic Journal: Applied Economics* 2018, Vol. 10, No. 2, pp. 1-26.

Murnane, Richard J. and Sean F. Reardon. "Long-Term Trends in Private School Enrollments by Family Income" *AERA Open* January-March 2018, Vol. 4, No. 1, pp. 1–24.

National Academies of Sciences, Engineering, and Medicine. *A Roadmap to Reducing Child Poverty* Washington, DC: The National Academies Press, 2019.

National Center for Education Statistics. (2020). Digest of Education Statistics 2018: Chapter 2 - Elementary and Secondary Education; 2020 ASI 4824-2.25005.

Nielsen, Eric. "The Income-Achievement Gap and Adult Outcome Inequality." Unpublished paper, Federal Reserve Board, May 13, 2019.

Owens, Ann, Sean F Reardon, and Christopher Jencks. "Income Segregation Between Schools and School Districts." *American Educational Research Journal 53*, no. 4 (2016): 1159-197.

Raudenbush, Stephen W. and Anthony S. Bryk. Hierarchical linear models: Applications and data analysis methods. Vol. 1. sage, 2002.

Reardon, Sean F. “The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations.” in Greg J. Duncan and Richard J. Murnane (eds.) *Whither Opportunity?* (New York: Russell Sage Foundation, 2011.)

Reardon, Sean F, and Kendra Bischoff. "Income Inequality and Income Segregation." *AJS; American Journal of Sociology 116*, no. 4 (2011): 1092-153.

Saporito, Salvatore and Deenesh Sohoni “Mapping Educational Inequality: Concentrations of Poverty among Poor and Minority Students in Public Schools” Social *Forces*, Volume 85, Number 3, March 2007.

Watson, Tara. "Inequality and the Measurement of Residential Segregation by Income." *Review of Income and Wealth 55*, no. 3 (2009): 820-44.

Table 1: NAEP Administration Schedule and State Sample Size

Year	Grade 4 Reading		Grade 8 Reading		Grade 4 Math		Grade 8 Math	
	Schools	Students	Schools	Students	Schools	Students	Schools	Students
1990							2,820	82,270
1992	3,930	97,470			4,200	105,030	3,170	95,240
1994	3,560	97,060						
1996					3,720	95,530	2,900	78,070
1998	2,980	75,930	2,840	77,110				
2000							2,560	67,170
2002	4,060	115,560	3,840	106,270				
2003	5,000	142,310	4,580	129,690	5,500	160,720	3,920	112,840
2005	6,280	126,340	5,280	133,870	6,880	143,100	4,570	118,330
2007	5,280	145,490	5,280	136,790	5,810	164,140	4,500	113,540
2009	6,460	136,640	5,330	137,730	7,080	140,560	4,500	118,800
2011	5,580	161,910	5,470	139,260	6,120	172,100	4,610	125,820
2013	5,500	147,070	5,160	148,880	6,050	156,990	4,410	127,630
2015	5,250	108,290	4,730	118,870	5,780	118,070	4,010	103,580
States in our sample	*41*		*46*		*45*		*40*	

Note: School counts include District of Columbia but excludes Department of Defense Schools, as well as Bureau of Indian's Affairs schools. States were included if they opted out of no more than one available NAEP administration during the period between 1990 and 2002. All counts were rounded to the nearest 10, per IES reporting requirements.

Source: U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Mathematics and Reading Assessments.

Table 2: Correlation Between Census-Derived School Characteristics and the Common Core of Data

	1990	2000	2009	2011	2013	2015
Percent of Students Receiving FRPL						
Correlation between CCD and Estimate	N.A.	0.715	0.669	0.655	0.694	0.690
Coefficient on Estimate in Predicting CCD	N.A.	1.042	0.938	0.890	0.926	0.918
N		*60,860*	*74,930*	*76,400*	*76,130*	*75,820*
Proportion of Students Black (including Black-Hispanic)						
Correlation between CCD and Estimate	0.853	0.903	0.900	0.899	0.901	0.898
Coefficient on Estimate in Predicting CCD	0.915	0.976	0.977	0.955	0.952	0.957
N	*59,470*	*67,970*	*76,610*	*76,770*	*76,630*	*76,560*
Proportion of Students White						
Correlation between CCD and Estimate	0.828	0.859	0.912	0.914	0.914	0.913
Coefficient on Estimate in Predicting CCD	1.006	1.053	0.982	0.966	0.963	0.967
N	*60,796*	*67,972*	*76,607*	*76,773*	*76,630*	*76,564*
Total # Schools in CCD	*68,300*	*69,790*	*76,610*	*76,780*	*76,630*	*76,560*

Note: 52% schools are missing FRPL information in the CCD in 1990, 13% are missing in 2000, and less than 2% are missing in 2009 and beyond. The CCD and Census handle race/ethnicity differently. For both datasets, the Black category includes both Black Hispanics and Black non-Hispanics. In the CCD, the white category always includes both white Hispanics and white non-Hispanics. The Census data codes both white Hispanics and white non-Hispanics as "white" in 1990 and 2000. From 2009 on, our Census-based estimates include only white non-Hispanics, but are still compared to the CCD proportions which include white Hispanics as well.

Table 3: R^2 for First-Stage Imputations of Household Income Using CPS

Year	R^2	CPS Sample Size
1990	0.268	32,826
1991	0.256	33,079
1992	0.267	32,230
1993	0.271	32,203
1994	0.287	31,879
1995	0.258	31,910
1996	0.250	28,287
1997	0.256	28,486
1998	0.259	28,513
1999	0.260	28,704
2000	0.255	28,695
2001	0.235	53,812
2002	0.246	53,339
2003	0.249	52,741
2004	0.243	51,707
2005	0.235	50,644
2006	0.247	49,797
2007	0.260	48,738
2008	0.256	48,067
2009	0.252	47,938
2010	0.261	47,689
2011	0.277	45,764
2012	0.261	44,794
2013	0.268	45,118
2014	0.243	44,154
2015	0.240	43,660

Note: In columns 2 and 3, each row presents the R2 and sample size for a separate regression predicting income in the CPS data in that year.

Table 4: Reconstructing the Relationship Between Student Achievement and Income using School-Level Aggregates (4th Grade Math)

					Using 1990 Income Percentiles		Using Actual Percentiles	
Year	Between-School Slope	Within-School Slope	ICC_Y	Implied Student Level Slope	Predicted Score at 10th Pctile	Predicted Difference at 90th vs. 10th Pctile	Predicted Difference at 90th vs. 10th Pctile	Difference between (6) and (7)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1990								
1992	27.898 (0.616)	14.671 (1.301)	0.191	17.204 (1.067)	199.699	35.313	36.314	1.001
1994								
1996	28.579 (0.669)	13.525 (1.442)	0.192	16.414 (1.184)	203.824	33.693	34.760	1.068
1998								
2000								
2002								
2003	25.626 (0.415)	10.274 (1.043)	0.211	13.517 (0.835)	218.987	27.746	28.338	0.593
2005	23.270 (0.364)	9.321 (1.221)	0.215	12.315 (0.971)	223.702	25.279	26.464	1.185
2007	23.182 (0.364)	6.715 (1.669)	0.222	10.364 (1.313)	227.802	21.274	22.420	1.146
2009	23.240 (0.339)	7.925 (1.174)	0.230	11.453 (0.911)	226.626	23.508	24.827	1.319
2011	21.834 (0.331)	7.438 (1.152)	0.229	10.736 (0.896)	228.747	22.036	24.282	2.245
2013	21.353 (0.325)	6.178 (1.346)	0.228	9.643 (1.045)	231.657	19.793	22.032	2.238
2015	20.861 (0.370)	7.179 (1.327)	0.230	10.326 (1.029)	229.925	21.196	23.663	2.467

Source: U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Mathematics and Reading Assessments. Mean and variance of income were derived from 1990 and 2000 Decennial Censuses, and the American Communities Survey from the Bureau of the Census. Income percentiles were calculated based on data of households with school aged children from IPUMS-CPS, University of Minnesota, www.ipums.org.

Table 5: Reconstructing the Relationship Between Student Achievement and Income using School-Level Aggregates (4th Grade Reading)

					Using 1990 Income Percentiles		Using Actual Percentiles	
Year	Between-School Slope	Within-School Slope	ICC_Y	Implied Student Level Slope	Predicted Score at 10th Pctile	Predicted Difference at 90th vs. 10th Pctile	Predicted Difference at 90th vs. 10th Pctile	Difference between (6) and (7)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1990								
1992	26.024 (0.752)	18.734 (1.254)	0.191	20.130 (1.042)	192.756	41.320	42.490	1.171
1994	29.858 (0.858)	21.867 (1.692)	0.190	23.389 (1.413)	186.157	48.009	51.715	3.707
1996								
1998	29.774 (0.833)	18.770 (1.719)	0.197	20.934 (1.431)	189.906	42.970	44.751	1.780
2000								
2002	30.964 (0.608)	11.720 (1.919)	0.210	15.753 (1.537)	198.972	32.336	33.043	0.706
2003	29.623 (0.538)	13.060 (1.660)	0.211	16.560 (1.330)	197.630	33.991	34.717	0.726
2005	27.082 (0.461)	8.229 (2.338)	0.215	12.276 (1.858)	203.387	25.198	26.379	1.181
2007	26.603 (0.466)	5.576 (3.437)	0.222	10.236 (2.698)	207.934	21.011	22.142	1.132
2009	25.948 (0.421)	4.830 (3.523)	0.230	9.695 (2.731)	208.432	19.901	21.018	1.116
2011	24.819 (0.415)	0* (0.000)	0.229	5.685 (0.093)	213.623	11.669	12.858	1.189
2013	24.960 (0.413)	0* (0.000)	0.228	5.700 (0.092)	214.823	11.699	13.023	1.323
2015	24.236 (0.435)	0* (0.000)	0.230	5.574 (0.098)	215.594	11.442	12.774	1.332

*Assigned to boundary of 0 because point estimate was negative.

Source: U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Mathematics and Reading Assessments. Mean and variance of income were derived from 1990 and 2000 Decennial Censuses, and the American Communities Survey from the Bureau of the Census. Income percentiles were calculated based on data of households with school aged children from IPUMS-CPS, University of Minnesota, www.ipums.org.

Table 6: Reconstructing the Relationship Between Student Achievement and Income using School-Level Aggregates (8th Grade Math)

Year	Between-School Slope	Within-School Slope	ICC_Y	Implied Student Level Slope	Using 1990 Income Percentiles: Predicted Score at 10th Pctile	Using 1990 Income Percentiles: Predicted Difference at 90th vs. 10th Pctile	Using Actual Percentiles: Predicted Difference at 90th vs. 10th Pctile	Difference between (6) and (7)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1990	29.079 (0.780)	18.913 (1.500)	0.194	20.890 (1.217)	240.231	42.880	42.880	0.000
1992	32.067 (0.768)	20.441 (1.418)	0.191	22.668 (1.151)	242.037	46.529	47.847	1.318
1994								
1996	33.978 (1.035)	20.094 (1.709)	0.192	22.759 (1.411)	244.807	46.716	48.196	1.480
1998								
2000	33.439 (1.002)	19.709 (2.188)	0.209	22.579 (1.778)	247.418	46.347	47.201	0.853
2002								
2003	32.455 (0.661)	16.812 (1.448)	0.211	20.117 (1.155)	253.177	41.293	42.175	0.882
2005	30.370 (0.647)	18.997 (1.144)	0.215	21.439 (0.905)	253.001	44.006	46.069	2.063
2007	28.264 (0.609)	15.473 (1.469)	0.222	18.308 (1.147)	259.483	37.579	39.603	2.024
2009	28.728 (0.639)	16.170 (1.154)	0.230	19.063 (0.895)	260.268	39.129	41.324	2.195
2011	25.717 (0.642)	12.750 (1.381)	0.229	15.720 (1.064)	265.423	32.268	35.556	3.288
2013	25.890 (0.545)	13.253 (1.191)	0.228	16.139 (0.924)	266.508	33.127	36.873	3.746
2015	25.818 (0.588)	13.526 (1.229)	0.230	16.353 (0.952)	264.489	33.568	37.475	3.907

Source: U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Mathematics and Reading Assessments. Mean and variance of income were derived from 1990 and 2000 Decennial Censuses, and the American Communities Survey from the Bureau of the Census. Income percentiles were calculated based on data of households with school aged children from IPUMS-CPS, University of Minnesota, www.ipums.org.

Table 7: Reconstructing the Relationship Between Student Achievement and Income using School-Level Aggregates (8th Grade Reading)

					Using 1990 Income Percentiles		Using Actual Percentiles	
Year	Between-School Slope	Within-School Slope	ICC_Y	Implied Student Level Slope	Predicted Score at 10th Pctile	Predicted Difference at 90th vs. 10th Pctile	Predicted Difference at 90th vs. 10th Pctile	Difference between (6) and (7)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1990								
1992								
1994								
1996								
1998	24.526 (0.861)	15.025 (2.537)	0.197	16.894 (2.090)	242.201	34.676	36.113	1.437
2000								
2002	25.457 (0.652)	16.583 (1.474)	0.210	18.443 (1.182)	241.614	37.857	38.685	0.827
2003	27.241 (0.614)	12.816 (1.885)	0.211	15.864 (1.512)	243.447	32.562	33.258	0.695
2005	24.529 (0.570)	14.479 (1.469)	0.215	16.637 (1.170)	241.632	34.149	35.750	1.601
2007	24.176 (0.552)	10.047 (2.024)	0.222	13.179 (1.589)	246.477	27.051	28.508	1.457
2009	23.661 (0.530)	9.975 (1.766)	0.230	13.128 (1.367)	247.881	26.947	28.459	1.512
2011	22.435 (0.577)	8.975 (1.756)	0.229	12.058 (1.362)	250.648	24.751	27.273	2.522
2013	22.982 (0.501)	8.904 (1.566)	0.228	12.119 (1.221)	253.448	24.876	27.689	2.813
2015	22.292 (0.525)	6.915 (2.155)	0.230	10.452 (1.670)	253.432	21.453	23.951	2.497

Source: U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Mathematics and Reading Assessments. Mean and variance of income were derived from 1990 and 2000 Decennial Censuses, and the American Communities Survey from the Bureau of the Census. Income percentiles were calculated based on data of households with school aged children from IPUMS-CPS, University of Minnesota, www.ipums.org.

Table 8: The Relationship Between Achievement and Alternative Ways of Imputing Income (8th Grade Math)

	Instruments Used			
Year	Race/ Ethnicity	Mother's Education	Urbanicity and State	Race*Moth ed, Urbanicity, State
1990	51.091	28.458	25.156	32.883
	(0.804)	(0.518)	(1.131)	(0.459)
1992	51.602	26.848	28.489	33.683
	(0.892)	(0.560)	(1.607)	(0.543)
1994				
1996	47.946	26.010	27.243	33.778
	(1.027)	(0.602)	(1.383)	(0.544)
1998				
2000	50.154	25.915	20.473	33.633
	(1.446)	(1.120)	(1.607)	(0.963)
2002				
2003	48.341	23.664	16.696	31.144
	(0.642)	(0.453)	(1.358)	(0.362)
2005	47.539	24.551	14.555	32.831
	(0.569)	(0.391)	(1.123)	(0.364)
2007	43.961	24.093	9.497	29.417
	(0.605)	(0.382)	(1.234)	(0.352)
2009	44.560	23.954	14.206	28.380
	(0.617)	(0.460)	(1.053)	(0.394)
2011	39.734	22.796	16.332	27.399
	(0.639)	(0.433)	(1.014)	(0.370)
2013	40.226	23.343	13.753	27.636
	(0.532)	(0.391)	(1.061)	(0.309)
2015	42.141	23.999	17.552	28.606
	(0.685)	(0.389)	(1.335)	(0.356)

Note: Each column reports the slope of a regession of student test score on income with income imputed using the variables indicated at the top of the column.

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on 8th Grade Mathematics Assessments. Imputed income based on data from IPUMS-CPS, University of Minnesota, www.ipums.org.

Table 9: The Within- and Between-School Relationship Between Achievement and Imputed Income (8th Grade Math)

Year	Beween-School Slope	Within-School Slope	ICC_Y	Student Level Slope	Number of Students	Number of Schools
1990	37.388 (0.875)	27.129 (0.557)	0.464	32.880 (0.459)	69,000	2,630
1992	45.851 (0.848)	24.050 (0.467)	0.405	33.680 (0.543)	73,770	2,820
1994						
1996	46.691 (0.886)	23.155 (0.497)	0.416	33.780 (0.544)	62,380	2,670
1998						
2000	51.261 (1.123)	21.826 (0.650)	0.370	33.630 (0.963)	55,140	2,400
2002						
2003	52.476 (0.811)	21.679 (0.401)	0.336	31.140 (0.362)	93,250	3,640
2005	51.602 (0.768)	23.066 (0.322)	0.337	32.830 (0.364)	104,890	4,420
2007	44.317 (0.715)	21.027 (0.332)	0.353	29.420 (0.352)	97,470	4,230
2009	42.768 (0.649)	20.763 (0.391)	0.388	28.380 (0.394)	100,940	4,220
2011	42.161 (0.614)	19.770 (0.375)	0.376	27.400 (0.370)	106,690	4,350
2013	41.857 (0.623)	20.120 (0.371)	0.379	27.640 (0.309)	106,600	4,150
2015	43.788 (0.669)	19.544 (0.422)	0.389	28.610 (0.356)	88,690	3,840

Note: Column 1 was estimated using mean achievement and mean imputed income by school. Income was imputed based on the interaction between race and mother's education as well as dummies for urbanicity and state. Column 2 was calculated regressing students' NAEP score on their imputed income and including school fixed effects. Column 3 was estimated with school random effects and imputed income as the outcome.

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on 8th Grade Mathematics Assessments. Imputed income based on data from IPUMS-CPS, University of Minnesota, www.ipums.org.

Table 10: The Relationship Between Achievement and Alternative Ways of Imputing Income (8th Grade Reading)

	Instruments Used			
Year	Race/ Ethnicity	Mother's Education	Urbanicity and State	Race*Moth ed, Urbanicity, State
1990				
1992				
1994				
1996				
1998	36.667	21.220	16.504	26.354
	(1.122)	(0.550)	(1.626)	(0.609)
2000				
2002	40.628	22.157	12.523	27.243
	(0.869)	(0.497)	(1.213)	(0.499)
2003	40.576	22.290	13.460	27.711
	(0.702)	(0.337)	(1.177)	(0.376)
2005	40.790	21.131	15.200	27.983
	(0.503)	(0.279)	(1.177)	(0.284)
2007	38.972	21.535	8.372	25.969
	(0.533)	(0.338)	(1.079)	(0.326)
2009	39.532	21.228	15.880	25.257
	(0.606)	(0.367)	(0.931)	(0.379)
2011	34.939	20.642	19.220	24.302
	(0.464)	(0.384)	(0.943)	(0.335)
2013	35.685	20.848	15.117	24.686
	(0.553)	(0.350)	(1.028)	(0.328)
2015	37.828	20.912	16.746	24.648
	(0.646)	(0.447)	(1.113)	(0.386)

Note: Each column reports the slope of a regession of student test score on income with income imputed using the variables indicated at the top of the column.

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on 8th Grade Reading Assessments. Imputed income based on data from IPUMS-CPS, University of Minnesota, www.ipums.org.

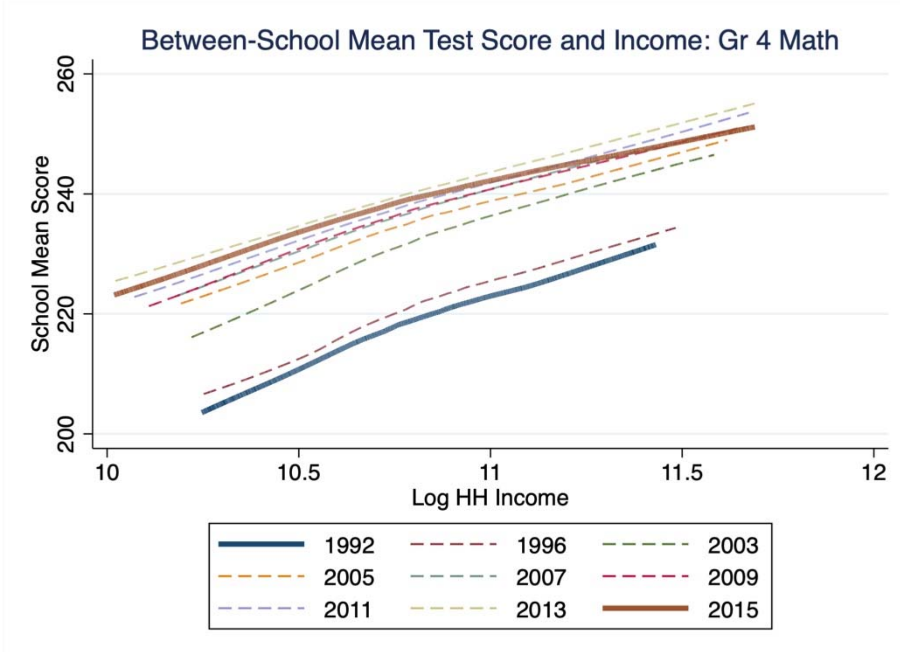
Table 11: The Within- and Between-School Relationship Between Achievement and Imputed Income (8th Grade Reading)

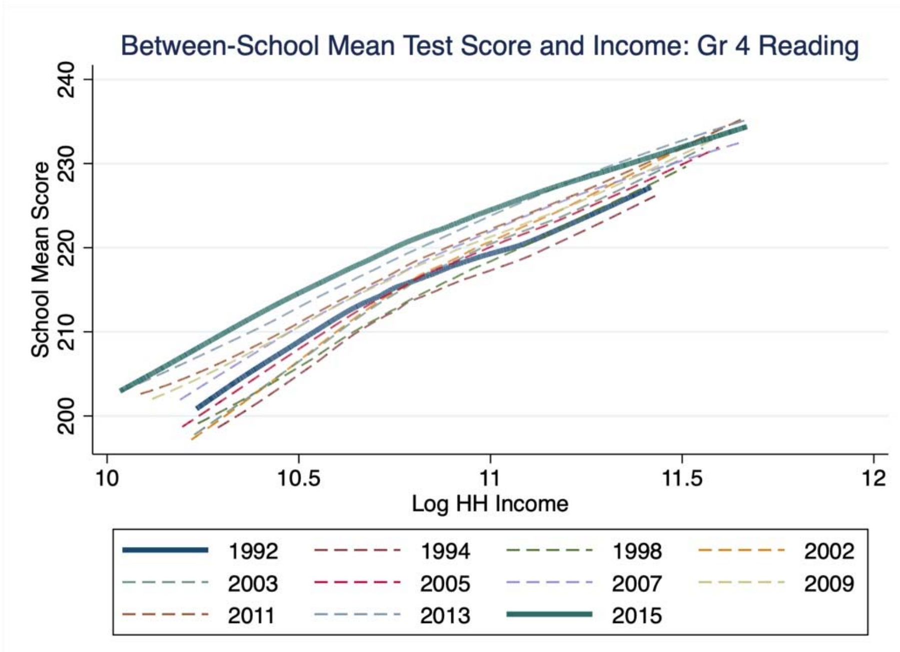
Year	Beween-School Slope	Within-School Slope	ICC_Y	Student Level Slope	Number of Students	Number of Schools
1990						
1992						
1994						
1996						
1998	36.031 (0.823)	19.61 (0.466)	0.382	26.354 (0.609)	2,620	62,270
2000						
2002	41.873 (0.780)	19.31 (0.412)	0.332	27.243 (0.499)	3,570	87,460
2003	46.527 (0.740)	19.07 (0.367)	0.322	27.711 (0.376)	4,290	108,230
2005	43.912 (0.695)	19.47 (0.303)	0.324	27.983 (0.284)	5,090	117,880
2007	41.302 (0.648)	18.18 (0.301)	0.332	25.969 (0.326)	4,990	117,510
2009	38.054 (0.558)	17.82 (0.362)	0.375	25.257 (0.379)	5,030	116,330
2011	38.989 (0.534)	17.60 (0.341)	0.361	24.302 (0.335)	5,170	117,460
2013	39.283 (0.530)	17.47 (0.321)	0.364	24.686 (0.328)	4,870	123,650
2015	37.821 (0.588)	16.66 (0.365)	0.379	24.648 (0.386)	4,530	101,040

Note: Column 1 was estimated using mean achievement and mean imputed income by school. Income was imputed based on the interaction between race and mother's education as well as dummies for urbanicity and state. Column 2 was calculated regressing students' NAEP score on their imputed income and including school fixed effects. Column 3 was estimated with school random effects and imputed income as the outcome.

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on 8th Grade Reading Assessments. Imputed income based on data from IPUMS-CPS, University of Minnesota, www.ipums.org.

Figure 1: Non-Parametric Relationship Between School Mean Achievement and Mean Log Income by Year (4th Grade Math and Reading)

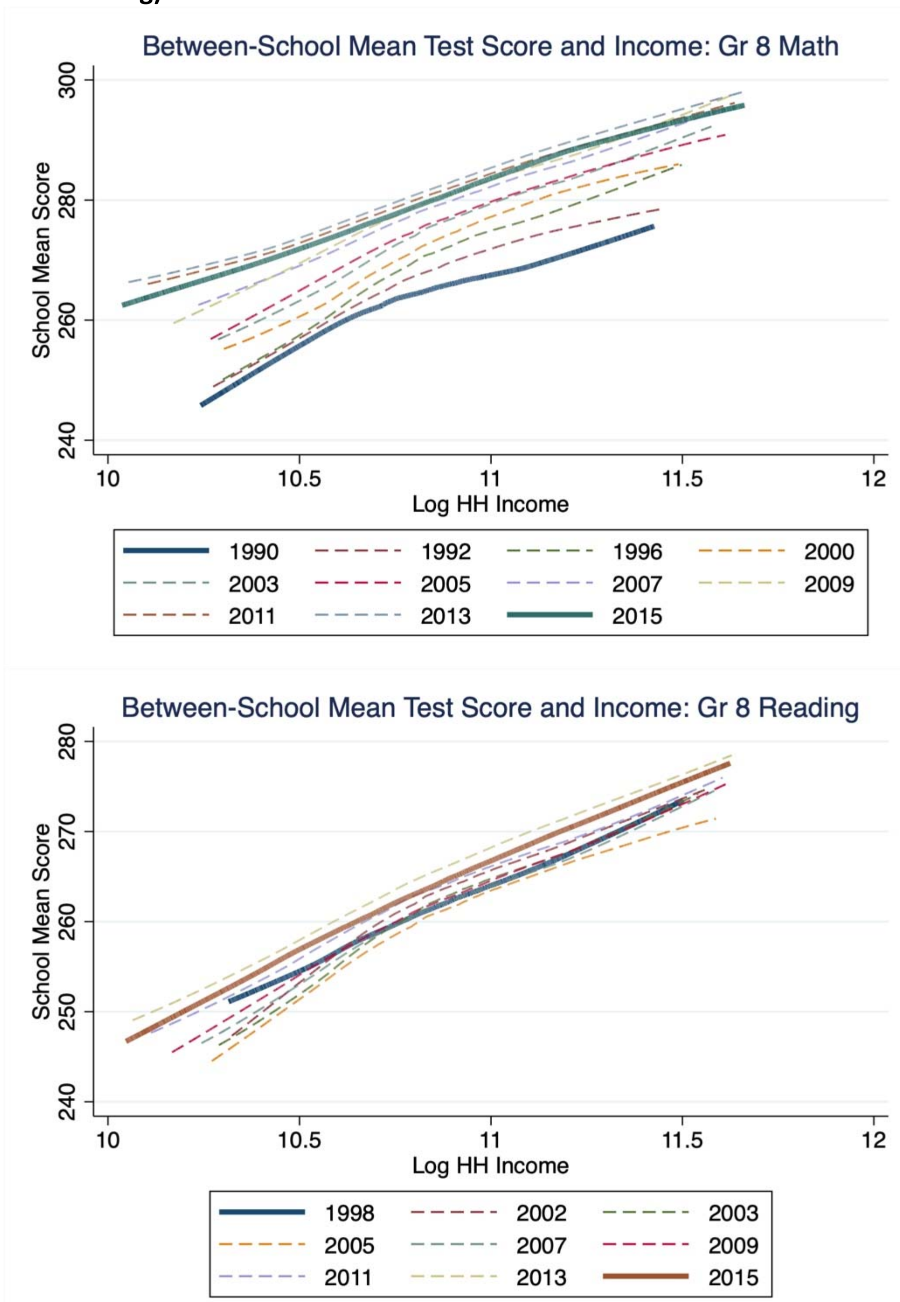




Note: Each line represents a separate lowess regression run at the school level. School-level weighted averages are calculated using NAEP's student-level weights. Schools with scores or estimated income in the top or bottom 5% of the distribution are excluded. Schools with estimated income variance in the top 5% of the distribution are also excluded. Sample is limited to consistent state sample (see Data section for details).

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Mathematics and Reading Assessments. Mean household income data are constructed using data from the U.S. Bureau of the Census.

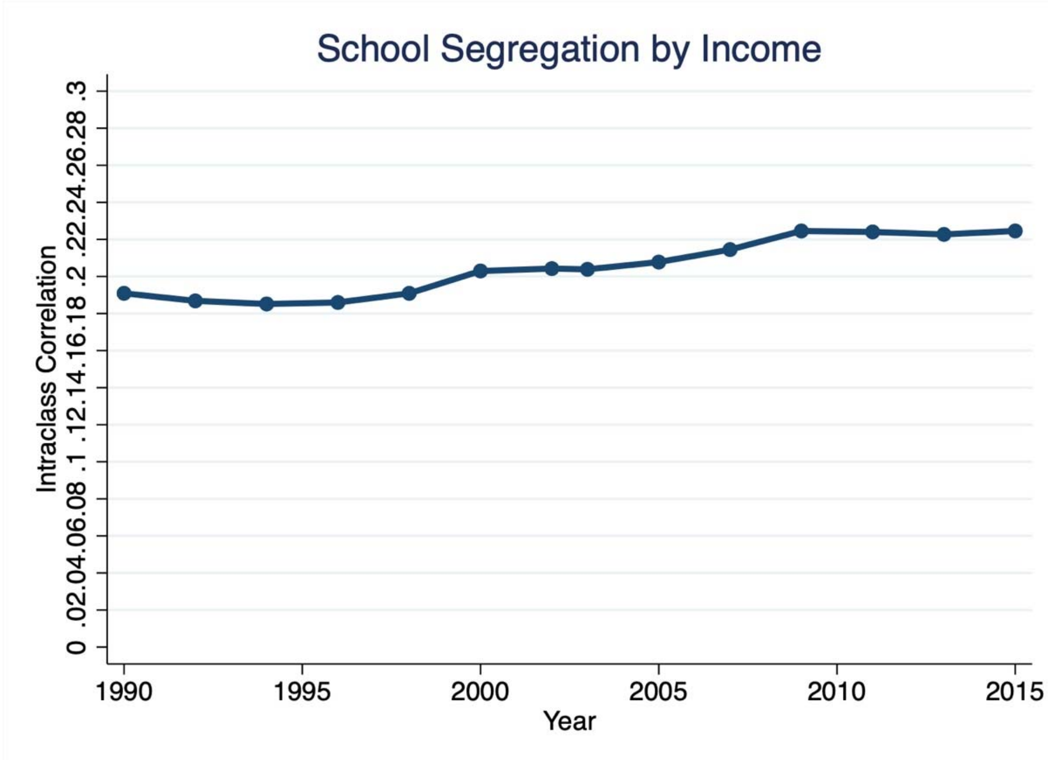
Figure 2: Non-Parametric Relationship Between School Mean Achievement and Mean Log Income by Year (8th Grade Math and Reading)



Note: Each line represents a separate lowess regression run at the school level. School-level weighted averages are calculated using NAEP's student-level weights. Schools with scores or estimated income in the top or bottom 5% of the distribution are excluded. Schools with estimated income variance in the top 5% of the distribution are also excluded. Sample is limited to consistent state sample (see Data section for details).

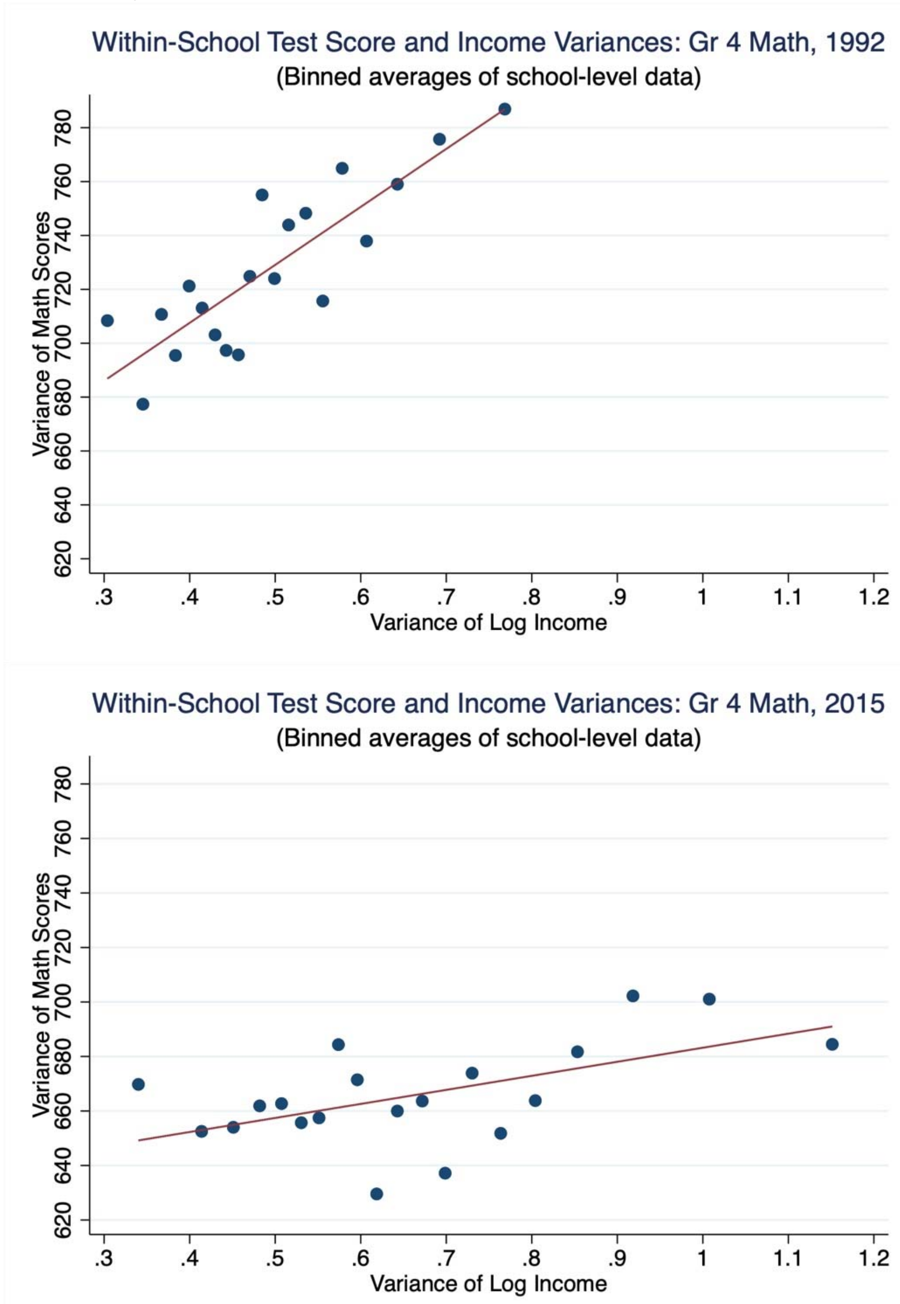
Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Mathematics and Reading Assessments. Mean household income data are constructed using data from the U.S. Bureau of the Census.

Figure 3: Intra-Class Correlation in School Household Income by Year



Note: ICC is calculated using means and variances of school household income calculated from census data. Schools whose variance fell above the 97th percentile were excluded from this calculation.

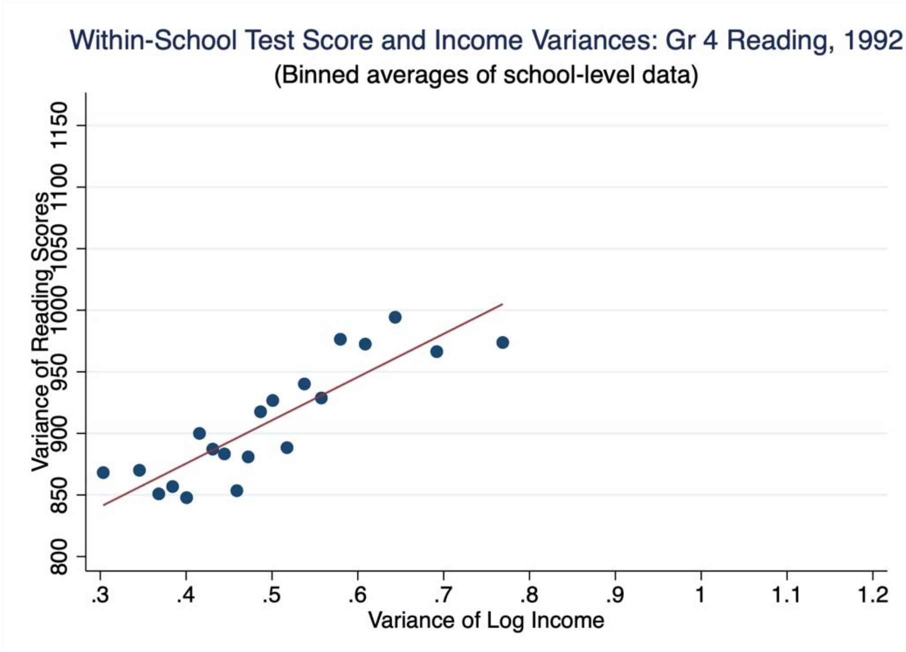
Figure 4: Within-School Variance in Achievement by Variance in Log Income by Year (Binned Averages, 4th Grade Math, 1992 and 2015)

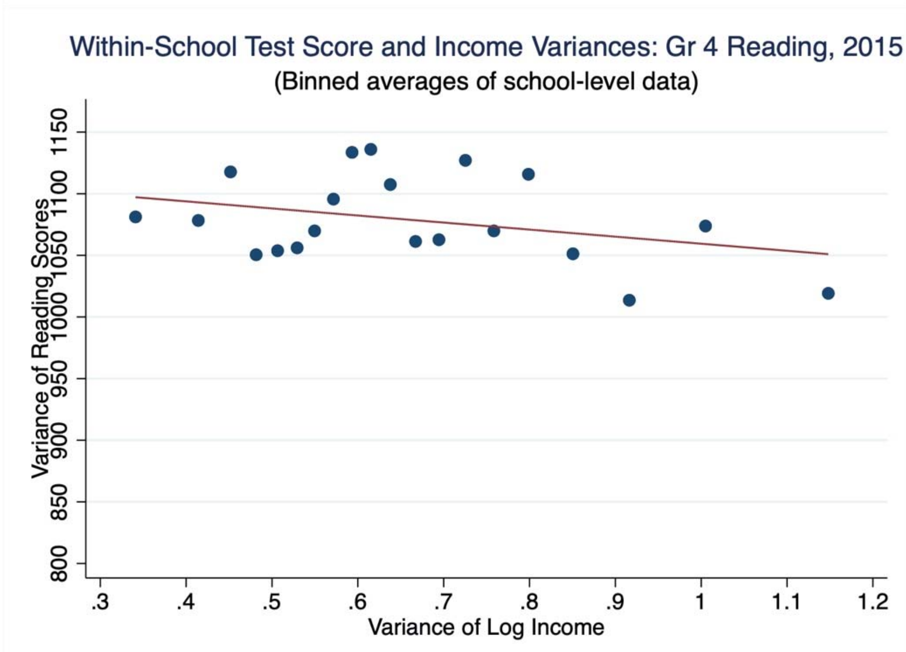


Note: Each point is a binned average of school-level data. Bin averages were weighted by the number of student scores in each school. Schools with estimated income variance in the top 5% of the distribution are excluded. Sample is limited to consistent state sample (see Data section for details).

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Mathematics and Reading Assessments. The variance in ln income in each school was estimated using data from the U.S. Bureau of the Census.

Figure 5: Within-School Variance in Achievement by Variance in Log Income By Year (Binned Averages, 4th Grade Reading, 1992 and 2015)

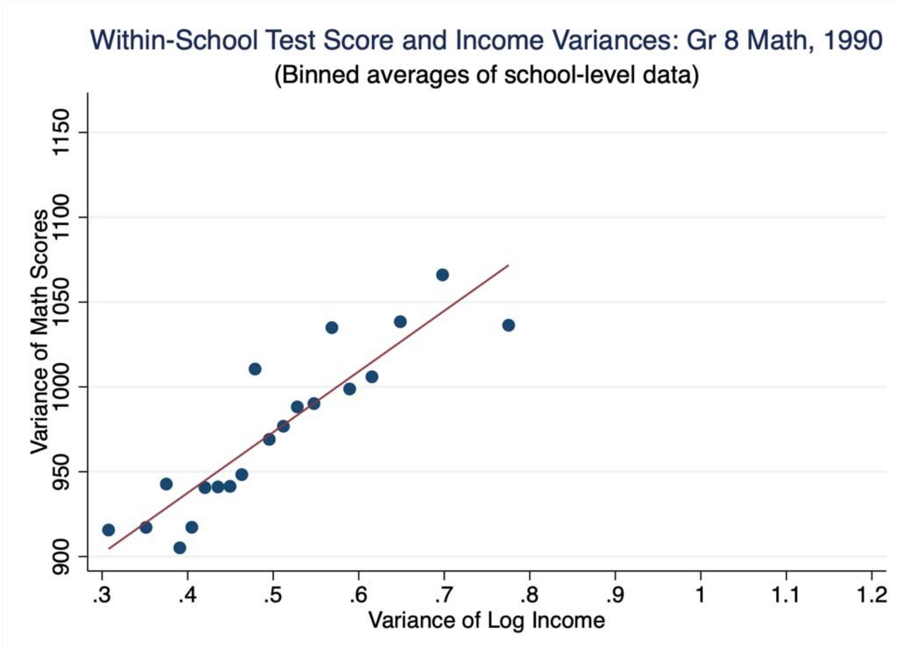


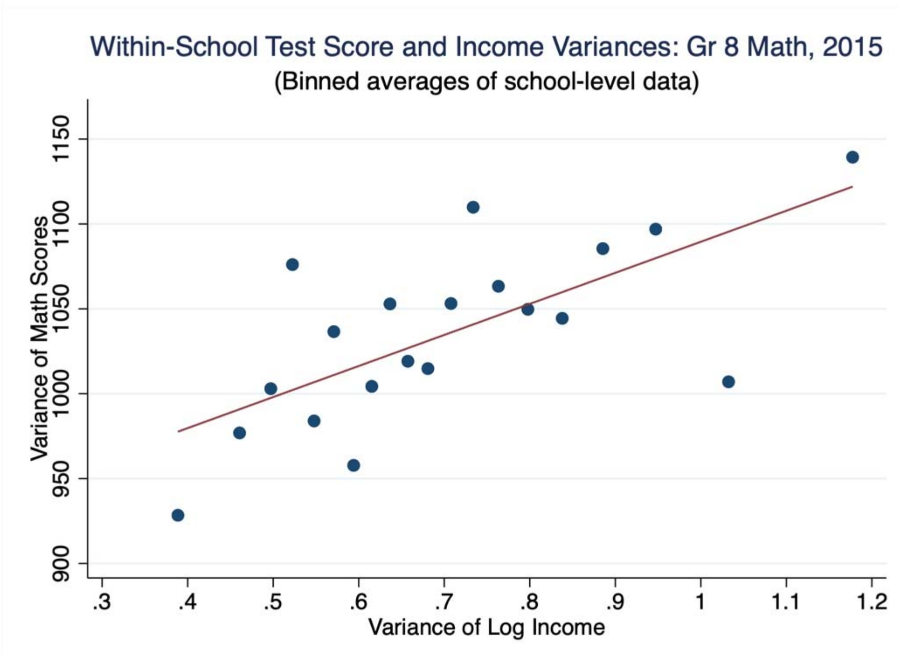


Note: Each point is a binned average of school-level data. Bin averages were weighted by the number of student scores in each school. Schools with estimated income variance in the top 5% of the distribution are excluded. Sample is limited to consistent state sample (see Data section for details).

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Mathematics and Reading Assessments. The variance in ln income in each school was estimated using data from the U.S. Bureau of the Census.

Figure 6: Within-School Variance in Achievement by Variance in Log Income by Year (Binned Averages, 8th Grade Math, 1990 and 2015)

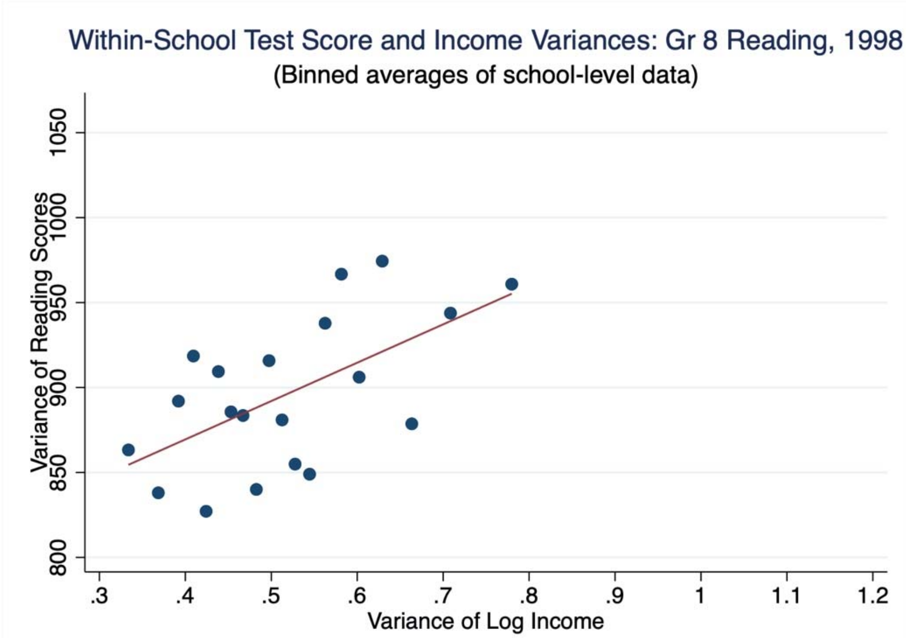


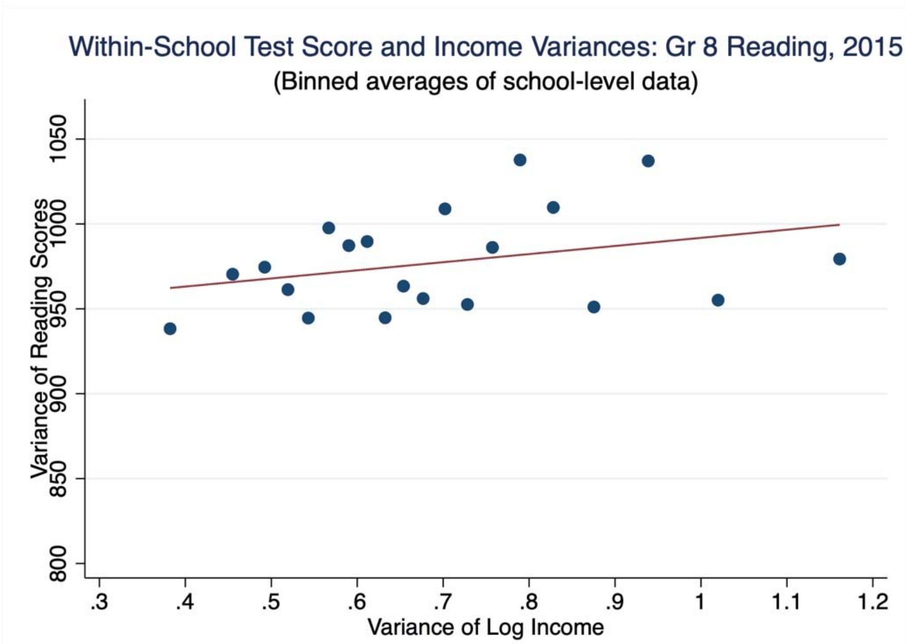


Note: Each point is a binned average of school-level data. Bin averages were weighted by the number of student scores in each school. Schools with estimated income variance in the top 5% of the distribution are excluded. Sample is limited to consistent state sample (see Data section for details).

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Mathematics and Reading Assessments. The variance in ln income in each school was estimated using data from the U.S. Bureau of the Census.

Figure 7: Within-School Variance in Achievement by Variance in Log Income by Year (Binned Averages, 8th Grade Reading, 1998 and 2015)

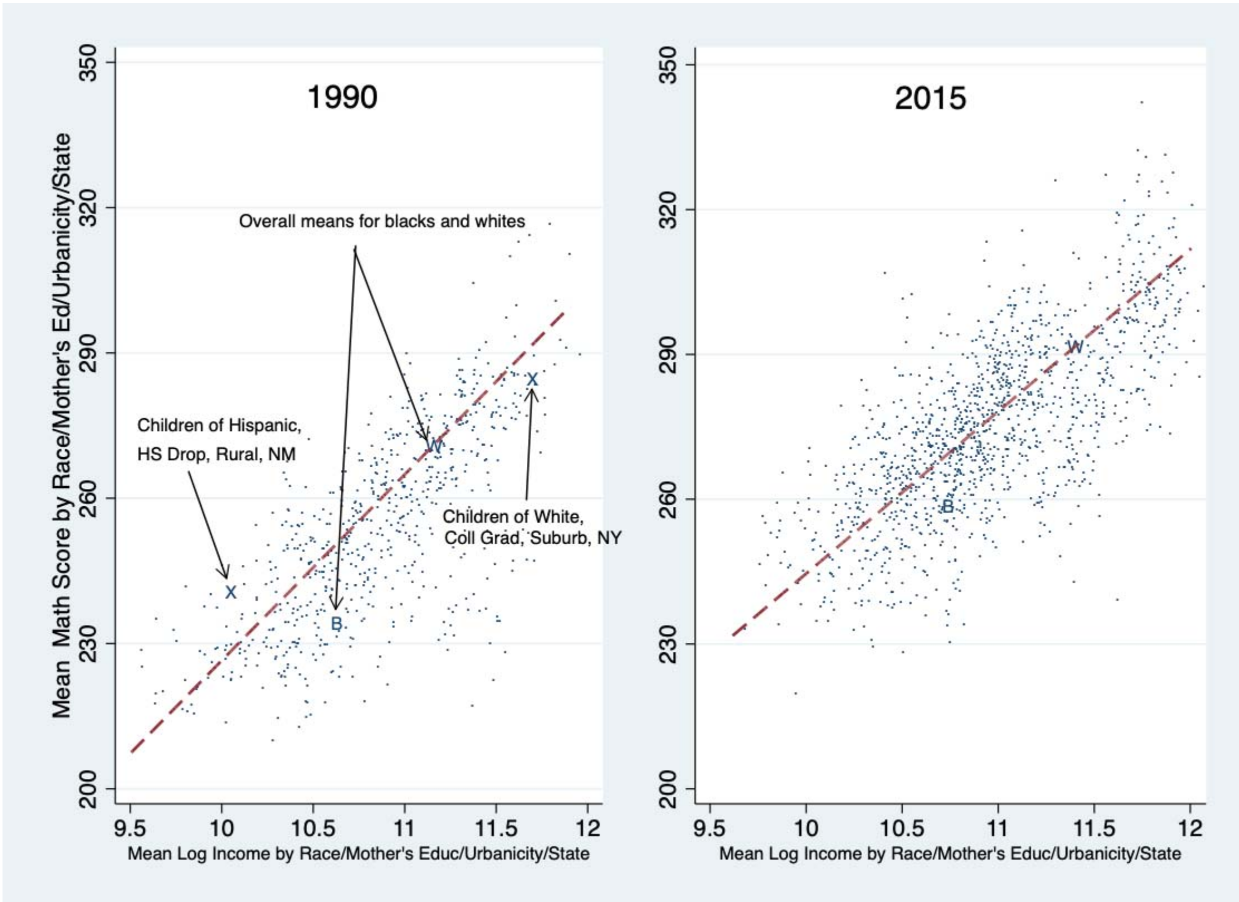




Note: Each point is a binned average of school-level data. Bin averages were weighted by the number of student scores in each school. Schools with estimated income variance in the top 5% of the distribution are excluded. Sample is limited to consistent state sample (see Data section for details).

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Mathematics and Reading Assessments. The variance in ln income in each school was estimated using data from the U.S. Bureau of the Census.

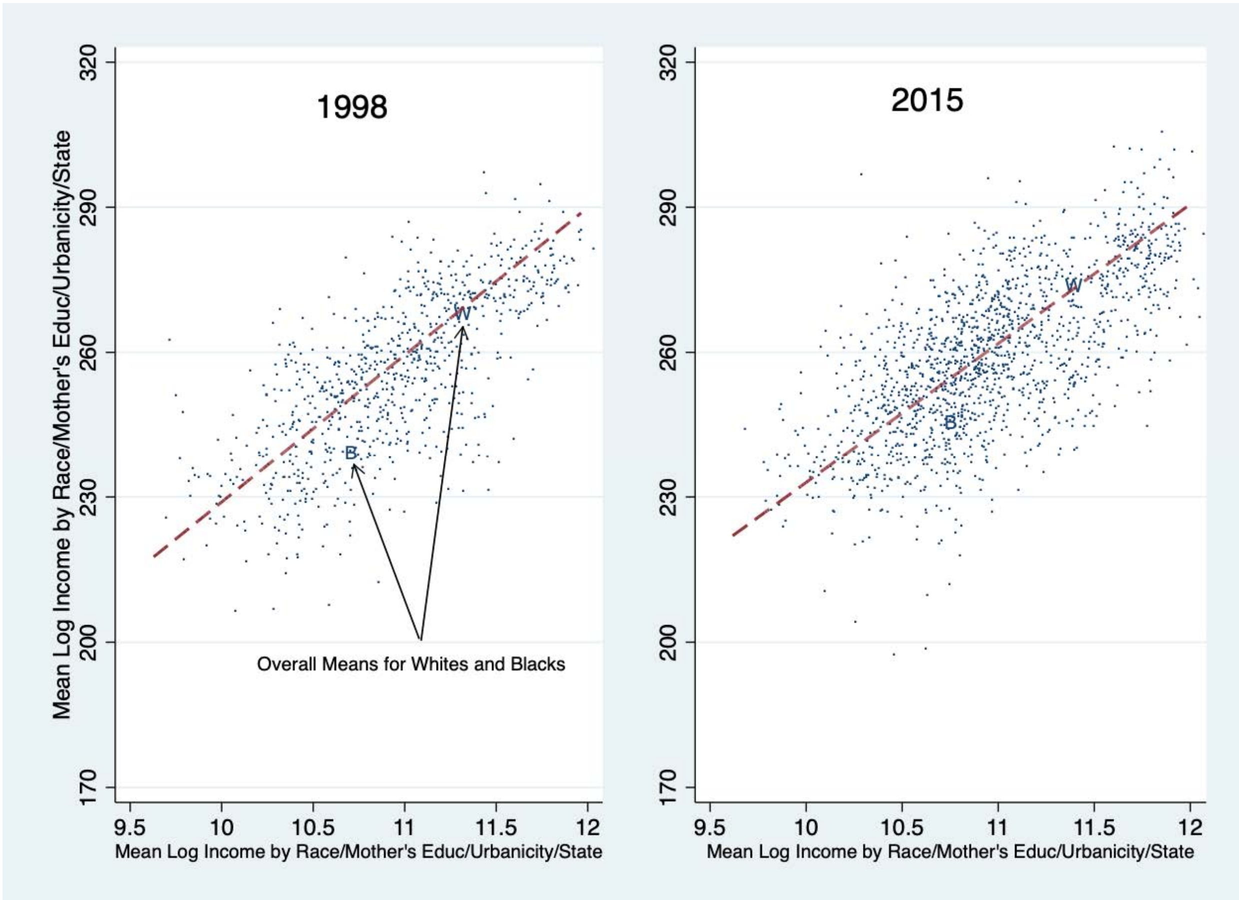
Figure 8: Mean Test Scores and Ln Incomes by Race, Mother's Education, Urbanicity and State Subgroup (8th grade Math)



Note: Each point represents the mean NAEP score and mean ln income for a subgroup defined by race, mother's education, urbanicity and state, weighted by the inverse of students' probability of being selected, adjusted for non-response.

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on 8th Grade Mathematics Assessments. Imputed income based on data from IPUMS-CPS, University of Minnesota, www.ipums.org.

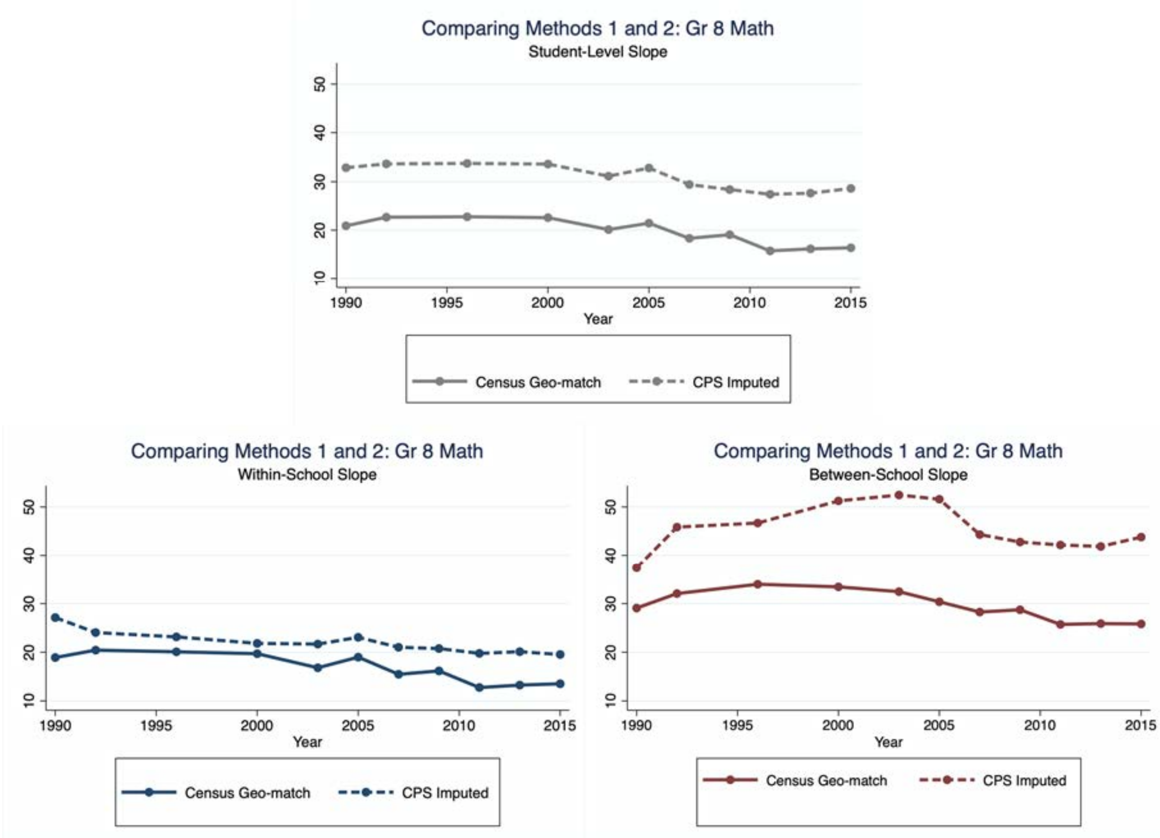
Figure 9: Mean Test Scores and Ln Incomes by Race, Mother's Education, Urbanicity and State Subgroup (8th grade Reading)



Note: Each point represents the mean NAEP score and mean ln income for a subgroup defined by race, mother's education, urbanicity and state, weighted by the inverse of students' probability of being selected, adjusted for non-response.

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on 8th Grade Mathematics Assessments. Imputed income based on data from IPUMS-CPS, University of Minnesota, www.ipums.org.

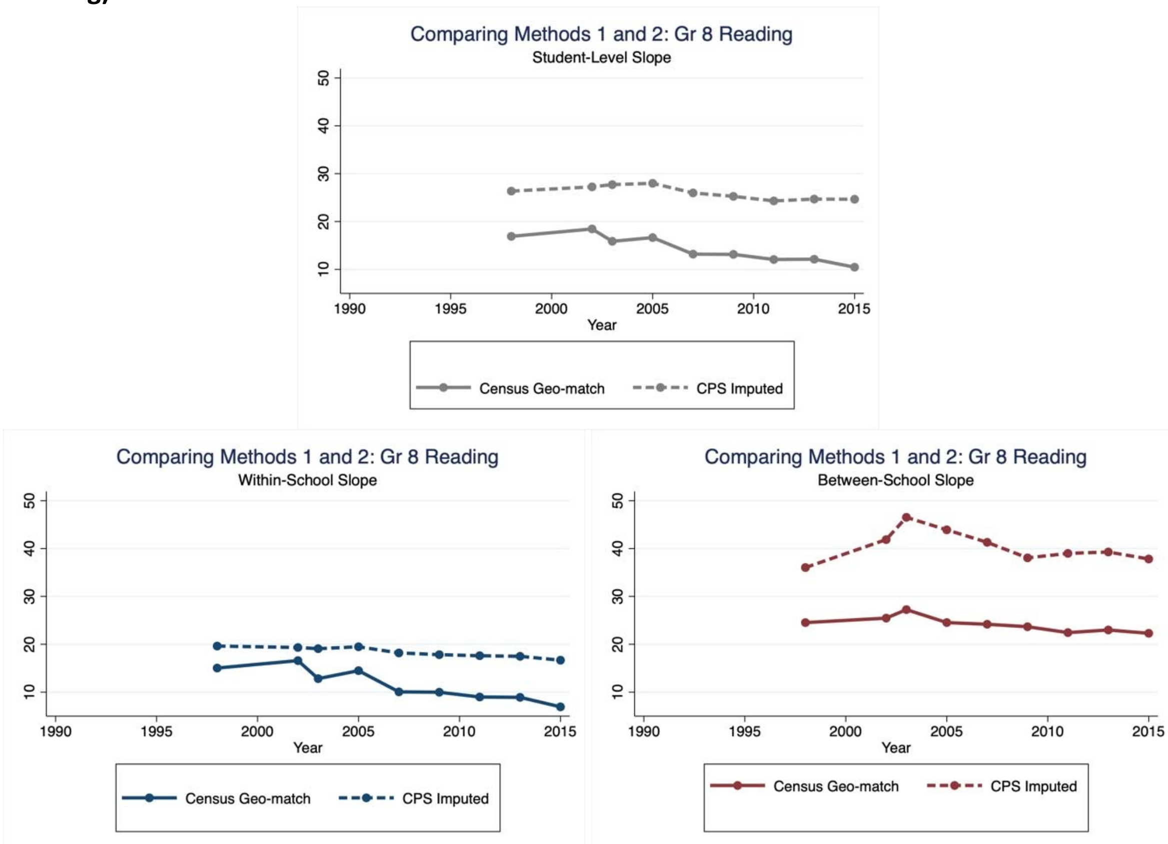
Figure 10: Comparing Estimates Based on Neighborhood Incomes and Imputed Student-Level Income (8th Grade Math)



Note: The line graphs are portraying the estimated coefficient of NAEP scaled score on log household income. The census-based estimates use the mean and variance in log household income in Census tracts near the schools. The CPS-imputed estimates use the March Current Population Survey to impute household income based on student-level characteristics found in the NAEP: race/ethnicity, mother's education, urbanicity, and state.

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on 8th Grade Mathematics Assessments. Mean household income data are constructed using data from the U.S. Bureau of the Census or imputed based on data from IPUMS-CPS, University of Minnesota, www.ipums.org.

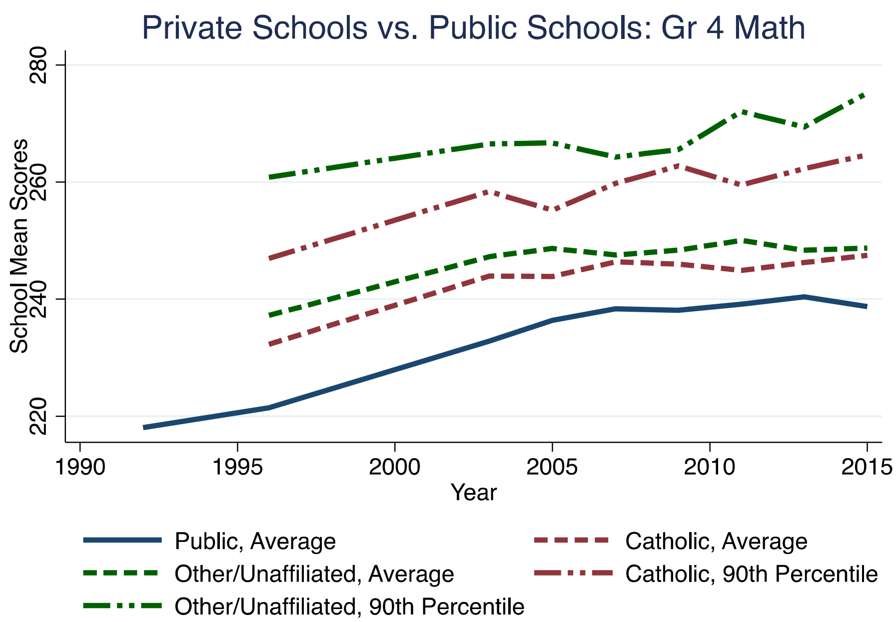
Figure 11: Comparing Estimates Based on Neighborhood Incomes and Imputed Student-Level Income (8th Grade Reading)

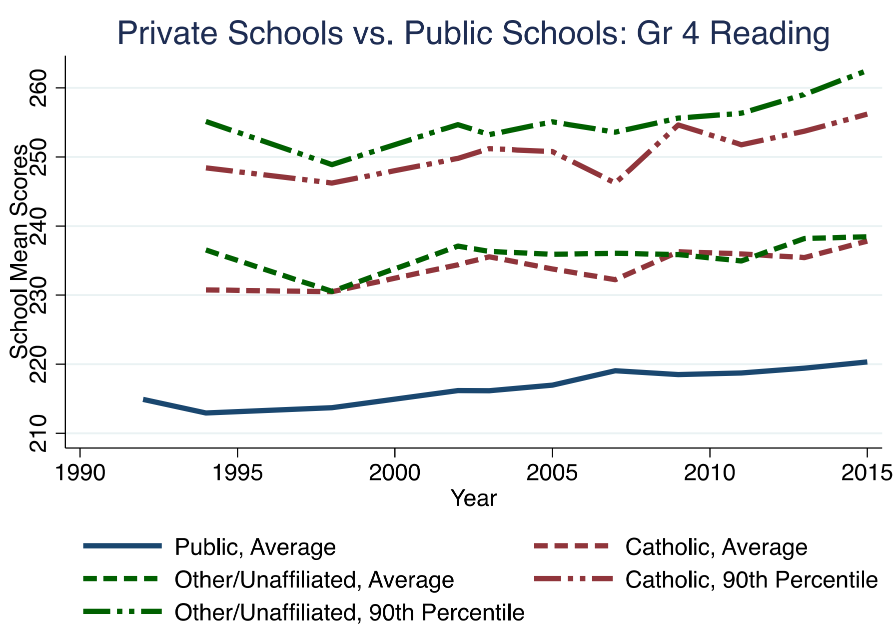


Note: The line graphs are portraying the estimated coefficient of NAEP scaled score on log household income. The census-based estimates use the mean and variance in log household income in Census tracts near the schools. The CPS-imputed estimates use the March Current Population Survey to impute household income based on student-level characteristics found in the NAEP: race/ethnicity, mother's education, urbanicity, and state.

Source: Student test score data are from U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on 8th Grade Reading Assessments. Mean household income data are constructed using data from the U.S. Bureau of the Census or imputed based on data from IPUMS-CPS, University of Minnesota, www.ipums.org.

Figure 12: Trend in Achievement in Private Schools (Grade 4)

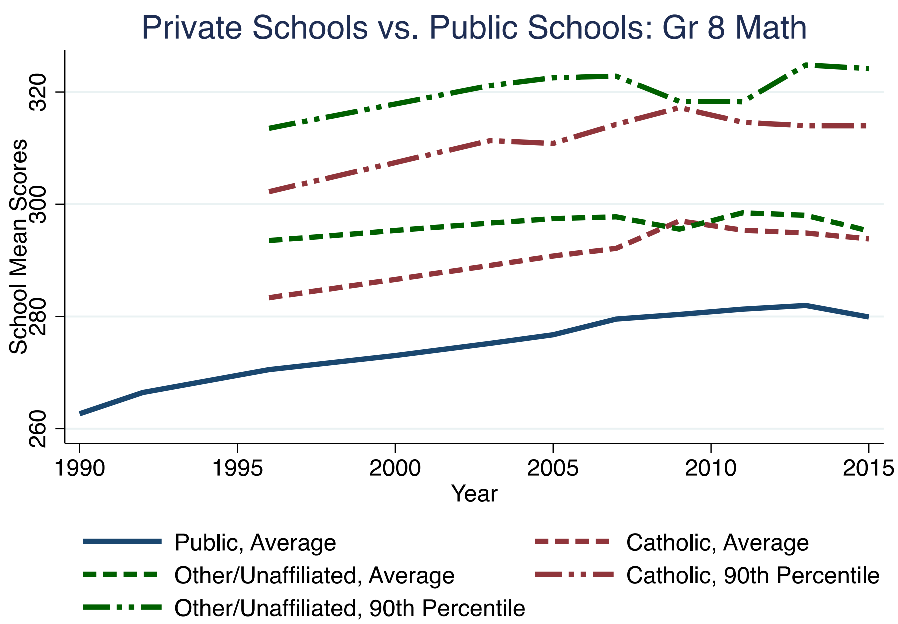


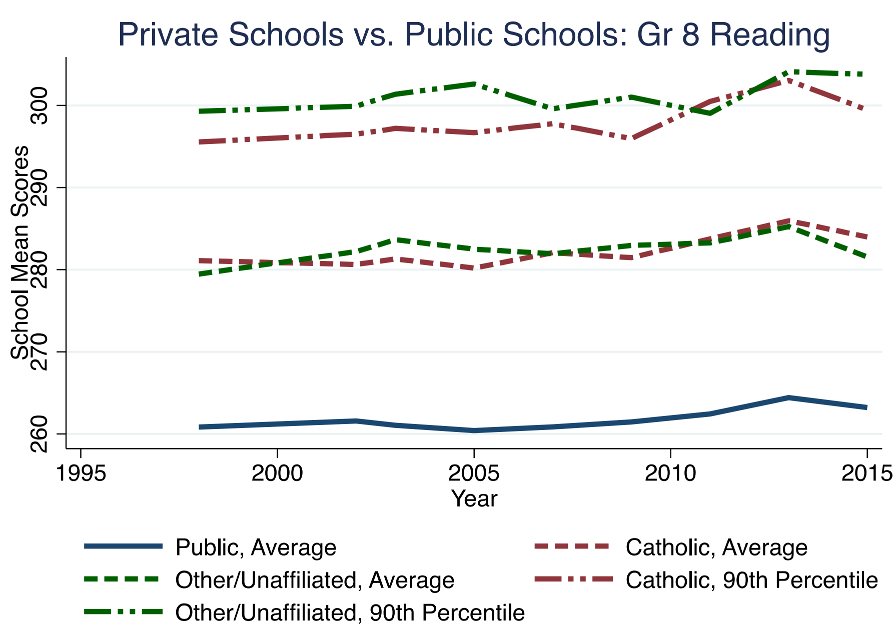


Note: For each school type, school-level weighted averages are calculated using NAEP's student-level weights. The average and 90th percentile of these school-level estimates are plotted for each year where school type is available.

Source: U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Reading and Mathematics Assessments.

Figure 13: Trend in Achievement in Private Schools (Grade 8)





Note: For each school type, school-level weighted averages are calculated using NAEP's student-level weights. The average and 90th percentile of these school-level estimates are plotted for each year where school type is available.

Source: U.S. Department of Education, National Center for Education Statistics, Main State National Assessment of Educational Progress (NAEP) various years 1990 on Reading and Mathematics Assessments.