Program on Education Policy and Governance Working Papers Series

A Half Century of Progress in U. S. Student Achievement: Ethnic and SES Differences; Agency and Flynn Effects

M. Danish Shakeel and Paul E. Peterson [1]

PEPG 21-01

Harvard Kennedy School
79 JFK Street, Taubman 304
Cambridge, MA 02138
Tel: 617-495-7976 Fax: 617-496-4428
www.hks.harvard.edu/pepg/

[1] Shakeel: Harvard University; Peterson: Harvard University and Hoover Institution/Stanford University.

Program on Education Policy and Governance Working Papers Series

A Half Century of Progress in U. S. Student Achievement:

Ethnic and SES Differences; Agency and Flynn Effects

M. Danish Shakeel and Paul E. Peterson

PEPG 21-01

Harvard Kennedy School
79 JFK Street, Taubman 304 Cambridge, MA 02138
Tel: 617-495-7976 Fax: 617-496-4428
www.hks.harvard.edu/pepg/

M. Danish Shakeel (danish_shakeel@hks.harvard.edu)

Paul E. Peterson (ppeterso@gov.harvard.edu)

April 20, 2021

Prior versions of this paper were presented at the 2020 and 2021 International School Choice and Reform Conferences, the 2021 faculty seminar at Harvard Kennedy School, and the 2021 conference of the Association for Education Finance and Policy. We received valuable feedback from Patrick Wolf, Benjamin Scafidi, Robert Maranto, Jonathan Wai, Jonathan Mills, Martin West, Anna Egalite, Sean Reardon, Benjamin Arold, and Jennifer Hochschild. Antonio Wendland provided exceptional administrative support. Kathleen Wolf provided valuable editorial review. We thank the U.S. Department of Education for providing restricted-use data for the surveys analyzed in this paper (license #001006519E). We also thank the staff at the National Center for Education Statistics (NCES) for answering several questions related to these datasets. All errors are ours.

Abstract

Principals (policy makers) have debated the progress in U. S. student performance for a half century or more. Informing these conversations, survey agents have administered seven million psychometrically linked tests in math and reading in 160 waves to national probability samples of selected cohorts born between 1954 and 2007. This study is the first to assess consistency of results by agency. We find results vary by agent, but consistent with Flynn effects, gains are larger in math than reading, except for the most recent period. Non-whites progress at a faster pace. Socio-economically disadvantaged white, black, and Hispanic students make greater progress when tested in elementary school, but that advantage attenuates and reverses itself as students age. We discuss potential moderators.

Keywords: Achievement levels; Flynn Effect; LTT; NAEP; TIMSS; PIRLS; PISA; ethnicity, SES; family structure.

Policy makers and public commentators, considered here as principals, have expressed concern about the rate of educational progress made by U. S. students. To measure the direction and rate of change, agents have administered over the past fifty years 160 waves of 17 psychometrically linked test surveys of student achievement in math and reading to national probability samples of U. S. student cohorts. Research in the sociology and economics of education on the rate of cohort progress, as measured by these survey waves, has been partial, bifurcated, and inattentive to variations in agent purpose, survey design, and test content (see Online Appendix, Table A.1). Meanwhile, a separate, psychometric literature has observed a growth in intelligence of 3 IQ points per decade over the past one hundred years, a trend named after James Flynn, the New Zealand political scientist who first called attention to this phenomenon. Recent meta-analyses of the Flynn effect by psychometricians find greater growth in one of the two main types of intelligence, fluid reasoning (ability to analyze abstract relationships), than in the other, crystallized knowledge (understanding the empirical world) (Pietschnig & Voracek, 2015; Trahan, Stuebing, Fletcher, & Hiscock, 2014).

We use individual level restricted-use student data that the National Center for Education Statistics (NCES) has provided in order to offer a more comprehensive look at the progress cohorts of U. S. students have made in math and reading achievement. Although we find substantial differences in estimates of progress by agent, most surveys show an upward trend. However, one agent reports a downward trend in the performance of students at age 15.

Trends are more favorable in math than in reading except for the most recent period. Cohorts of 4th grade students in other countries have also made more progress in math than in reading. The difference in trends for the two subjects resemble Flynn effects observed for fluid reasoning and crystallized knowledge (Pietschnig & Voracek, 2015). Also consistent with Flynn effects, we find steeper upward trends for students when tested at a younger age.

We observe heterogeneities. Young boys are making slightly more progress than young girls, but gender differences disappear when students are tested at an older age. Non-white students progress at a moderately faster pace than white ones. When tested at a younger age, the advance for those from households in the lowest quartile of the socio-economic status (SES) distribution is steeper than for those in the highest quartile. That advantage attenuates and reverses itself at older ages for white, black, and, to a lesser extent, Hispanic students, but not for Asian ones. Moderators that may account for trends remain uncertain, but family-school interactions could account for ethnic and SES differentials in the rate of progress. Greater gains in math than in reading, especially at younger ages could be due to improvements in neo-natal and early childhood health and well-being.

The remainder of the paper is organized in the following fashion: 1) a discussion of the principal-agent problem in research on trends in student achievement; 2) a review of trend research in the sociology and economics of education and in psychometric studies of intelligence; 3) data collection and organization; 4) methodology; 5) the results; and, 6) a discussion of the findings.

Policy makers and survey agents: A Principal-Agent problem

Policy makers and influential commentators have for decades expressed concern about the educational progress students in the United States have made. A highly publicized report in 1983 that the U. S. National Commission on Educational Excellence (NCEE) released, claimed that "the educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and a people (NCEE, 1983, p. 5)." According to Hirsch (1987, p. 7), the National Assessment of Educational Progress (NAEP) provides persuasive "evidence for the decline in shared knowledge." Others have published books with such disturbing titles as *The Literacy Hoax* (Copperman (1979), *Dumbing Down our Kids* (Sykes, 1995), *The Decline of Intelligence in America* (Itzkoff, 1994), and *The Dumbest Generation* (Bauerlein, 2008). The other side claims that schools are improving, students are making progress, and that claims of deterioration are false (Berliner & Biddle, 1996; Bracey, 1992).

These writers are addressing an important question. The well-being of the next generation—and the country as a whole—depends upon continuing progress in student math and reading achievement. Higher levels of achievement have positive impacts on college attainment, future earnings, teenage pregnancy rates, physical and mental health, political participation, and many other life outcomes (Borghans et al., 2016; Chetty, Friedman, & Rockoff, 2014). Nations that show higher average levels of student achievement enjoy steeper rates of economic growth (Hanushek & Woessmann, 2008, 2012).

Given the importance of the topic, policy makers have sought information on student progress. Their efforts to gather this information confront challenges which may be conceptualized as instances of the principal-agent problem (Bovens, Goodin, & Schillemans, 2014). Lacking the requisite technical skills, principals (policymakers) select one or more agents (survey organizations) with the technical capacity to gather the data (Braun & Guston, 2003). The assignment of the task to an agent creates moral hazards, as agents may shirk certain tasks or have goals that differ from those of the principal, or they may have difficulty interpreting the preferences of principals. In order to monitor agents, the principal can assign the task to multiple agencies. If agency preferences differ, the data they collect may generate alternative estimates of performance trends. To see whether agency effects are likely, we shall briefly describe the goals and practices of the four agencies that policy-making principals have used to gather this information.

Agents

Four agents have administered 160 waves of 17 temporally linked surveys of achievement in math and/or reading to nationally representative cohorts of U. S. students for various portions of the past half century: 1) the Long-Term Trend (LTT) version of NAEP, 2) the Main version of NAEP, 3) The International Association for the Evaluation of Educational Achievement (IEA) which administers Trends in International Math and Science Survey (TIMSS) and Progress in International Reading Literacy Study (PIRLS), and 4) the Program on Individual Student Assessment (PISA). Each survey has specific purposes, sampling frames, and test content that differentiates it from the others. Although our analysis is focused on U. S. student achievement, we also make use of some TIMMS-PIRLS and PISA trend data on math and reading achievement in other countries. Although several agents administer waves of tests in other subjects, we limit this analysis to trends in math and reading in order to ascertain consistency of results by agent.

LTT

LTT is the first of two surveys that the National Assessment Governing Board (NAGB) administers which is located within the U.S. Department of Education. LTT has been administered periodically to waves of nationally representative samples of cohorts born as early as 1954 in reading and 1961 in math (Tables 1, A.2). The most recent cohort for which information is available in this analysis was born in 2002. Because the main purpose of LTT is to track trends in the performances of students nationwide, LTT collects only a few thousand observations in each wave of the survey (Table 1). Samples are drawn by age, at 9, 13, and 17 years, which then standardizes the population sampled across waves of the survey. Still, LTT may not be a perfectly constant measure of student performance across survey waves. As one commentator observes, "The objectives and test items became more closely aligned with school curricula; . . . its sampling, which had included young adults and out-of-school youth, was narrowed to those in school (Stedman, 2009, p. 3)." But others fault LTT for maintaining consistent testing frameworks that "become increasingly irrelevant by failing to reflect changes in curricula and instructional practice (Pellegrino, Jones, & Mitchell, 1999, p. 78)."

NAEP

As a supplement to LTT, NAGB, beginning in 1990, initiated NAEP, a second series of psychometrically linked tests, as a supplement to LTT. We analyze NAEP cohorts born between 1973 and 2007. Although NAEP administered tests prior to 1990, the surveys for this period cannot be temporally linked to later ones. Although NAGB is responsible for both surveys, we treat them as separate agents because the two surveys differ in purpose, sampling frame, and test content. Unlike its predecessor, NAEP is designed to help states hold school districts accountable for their students' performance by providing a common metric of student achievement across the country. For this purpose, it is necessary to administer tests to representative samples of students in every state as well as in the United States as a whole. Accordingly, the number of observations is much larger than the LTT number (Table 1). The students are tested in 4^{th}, 8^{th}, and 12^{th} grade (rather than by age) in order to assess student performance on material aligned to the curriculum expected to have been taught by that grade. If the age composition of a grade varies, results may not be strictly comparable over time. Further, NAEP, unlike LTT, revises its tests to adjust to changes in its understanding of the curriculum being taught in U. S. schools, although it also attempts to maintain consistency in tests over time by using “bridging questions” that overlap survey waves. NAEP’s commitment to assessing performance against the curriculum draws criticism from those who prefer a test more tied to the external world that a child will eventually face. “It is not enough for students to receive high scores on a math test,” says one critic; “in addition, we want them to be comfortable with math and readily use it in the real world (Stedman, 2009, p. 31).”

Trends in NAEP math for students tested in grade 12 are included only for students born after 1988. We exclude results for cohorts born between 1973 and 1983, because NAGB, after administering four waves, withheld further testing for five years until a revised version not psychometrically linked to prior ones was given. This decision suggests NAGB was dissatisfied with the reliability and validity of the earlier version of its 12th grade math test.

TIMSS-PIRLS

The IEA, a private organization established in the 1960s, administers tests in math and science to students in grades 4 and 8 and, as discussed below, in literacy in grade 4. In 1995, IEA decided that going forward, TIMSS tests would be psychometrically linked so that performances within individual countries could be tracked over time. TIMSS is administered at four-year intervals for cohorts born between 1981 and 2005, the last year for which data are included in this analysis of U. S. trends. Its main purpose is to provide a common framework for assessing student performance across countries in order to see what features of educational systems critically affect student performance on a math and science curriculum the organization expects to be taught universally. Test items are drawn from frameworks “organized around two dimensions: a content dimension specifying the content to be assessed and a cognitive dimension specifying the thinking processes to be assessed." TIMSS says that "the majority of" its items "assess students' applying and reasoning skills" (IEA, 2021). Like NAEP, the questions TIMSS asks are "knowledge oriented;" its questions are "direct and abstract." It asks students, "What do you know?" (Hutchinson and Schagen, 2007, p. 54).

For years, IEA did not administer a literacy test on the grounds that languages differ in difficulty, but in 2001 it began the administration of PIRLS to students in grade 4, who were born in 1991. The survey's purpose is to assess student performance against an understanding of the literacy children should be expected to achieve by a specific grade. On its website, PIRLS states that it assesses the following four processes of comprehension: 1) a student's ability to "retrieve . . . information, 2) make straightforward inferences, 3) interpret and integrate ideas and information, and 4) evaluate and critique content" (Mullis & Martin, 2015). Results are available for cohorts born between 1991 and 2006. Information on student ethnicity is missing for 2001 and 2016, as the restricted-use data set is not available from NCES for these years.

PISA

This survey is administered by the Organization for Economic Co-operation and Development (OECD), a 37 member international organization of industrialized countries that collects statistics and information on the economic and social well-being of member countries. It relies upon the cooperation of member states for the administration of a survey of student performance at age 15 in math and reading every three years. In each administration of the test, it selects a subject for special emphasis, asking a broader range of questions. Although some researchers prefer to use only data from the subject tested more intensively, we include all results the agency deemed accurate enough to report. Ever since its first administration to cohorts born in 1985, the United States has fully cooperated with PISA, although results in reading in 2006 were not reported due to an error in test administration. The latest cohort analyzed in this paper was born in 2000.

PISA assessments are designed to measure the ability to apply knowledge to practical matters, not comprehension of the curriculum taught as of a particular grade. In math, for example, PISA says that the math literacy which they assess is "an individual's capacity to identify and understand the role that mathematics plays in the world, make well-founded judgments, and use and engage with mathematics in ways that meet one's needs as a constructive, concerned, and reflective citizen (U. S. Department of Education, 2014)." Given this purpose, its tests differ from those that NAEP and TIMSS administer to students in 8th grade. In a thoughtful comparison of TIMSS and PISA, Hutchison and Schagen (2007, p. 254) state the difference in these words:

> PISA's goals generate a different set of test questions than those offered by LTT, NAEP and TIMSS. PISA items are aimed at life skills while TIMSS items are more knowledge oriented. Where TIMSS questions are more direct and abstract, PISA questions are more lengthy and wordy. TIMSS asks, "What do you know?" while PISA asks "What can you do?"

Summary

Using agents with varying goals and procedures, principals were able to track student progress. Although all four agents link their surveys psychometrically in ways that allow for tracking trends in student performance over time, each has its own purposes and design. LTT appears to be most closely aligned to the ostensible purposes of principals, the tracking of long-term trends, but even it has modified its sampling design and test content. NAEP, which is designed to facilitate state accountability practices, adjusts tests to changing curricular objectives. TIMSS-PIRLS are designed to reflect international understanding of what is expected to be learned by a certain grade. And PISA estimates the level of student competence to participate in the social and economic world at the age of 15 when many countries end compulsory schooling. Because of agency differences in their interpretation of principal objectives, estimates of trends in student performance could vary from one survey to the next.

Scholarly research

Research on progress in student achievement is bifurcated between research in the sociology and economics of education, on the one hand, and intelligence studies that psychometricians have undertaken, on the other. This selected review brings together key results from two traditions that have generally ignored one another (see Table A.1 for a summary of findings).

Sociology and economics of education

Numerous studies have made intensive use of survey data that the agents have conducted which is discussed above, but the literature has yet to assemble estimates of progress in student achievement in math and reading from all available linked survey waves that all four agencies have administered. Most studies draw upon surveys from one or more agencies to estimate trends in socio-economic, gender, ethnic, or school sector divides, although they may touch upon growth in student performance as part of their larger project. Some studies find modest gains in student achievement (Hanushek, Peterson, & Woessmann, 2012), chiefly in math (Campbell et al., 1996). Reardon, Valentino, and Shores (2012, p. 23) reports steeper upward trends in math than in reading on the LTT. Studies on ethnic achievement gaps report substantial closing of the black-white gap, but a flat trend in the most recent period (Jencks & Phillips, 1998; Magnuson &

Waldfogel, 2008). Hedges & Nowell (1998) report closing of gaps between white and non-white students over a similar time period. Miller (1995) finds closing of Hispanic-white gaps and Asian-white gaps. Nowell and Hedges (1995) find more males than females among high-scoring individuals in math, but not in reading. Others have looked at influence of family structure on changes in student achievement (Grissmer et al., 1994) and the black-white gap (Phillips, Crouse, & Ralph, 1998). Bloom et al. (2008) find academic growth declines as students move from early grades to later grades. Elsewhere, Reardon (2011) draws upon surveys which are not psychometrically linked and finds income achievement gaps to have widened between students from households in the top and bottom 10 percent of the income distribution. Other studies find no clear trend, flat, or declining trends in income or SES achievement gaps in the United States (Broer, Bai, & Fonseca, 2019; Chmielewski, 2019; Hanushek et al., 2020; Hashim et al., 2020; Hedges & Nowell, 1998; OECD, 2018). Both Hanushek et al. (2020) and Hashim et al. (2020) find steeper upward trends for those tested at a younger age. Shakeel and Peterson (2020) show greater increases in math and reading achievement on NAEP at charter schools than at district schools for those tested between 2005 and 2017. Matheny et al. (2021) report growing gaps on state tests for school districts comprised of students from contrasting social and ethnic backgrounds.

Intelligence studies

These studies are complemented by research that has estimated changes in intelligence over time known as the "Flynn effect" (Herrnstein & Murray, 1994). See Table A.1. James Flynn discerned a positive trend in the intelligence quotient (IQ) of about three points per decade among those who took the Stanford-Binet intelligence tests administered during the 1930s and for another sample tested in the 1970s (Flynn, 1984). Rodgers and Wänström (2007) show a

Flynn effect for math (but not for reading) in the Children of the National Longitudinal Survey of Youth (NLSYC) PIAT-Math at each age between 5 and 13. Using the same data set, Ang, Rodgers and Wänström (2010) find Flynn effects for subgroups differentiated by gender, race, maternal education, income, and locale. A Flynn effect has also been identified among high ability students who took the SAT and ACT (Wai & Putallaz, 2011). Pietschnig & Voracek (2015), hereinafter PV, conduct a meta-analysis of 219 multiple administrations of the same (or similar) intelligence tests over the period 1909-2013 to cohorts of representative individuals from similar populations throughout the world. In general, they confirm Flynn's findings of average gains in IQ, gains of approximately 3 points every decade during this period. However, they find variation by region, and they find that the size of IQ gains has diminished in recent decades in industrialized countries. They also report for the most recent decades that there are larger IQ gains in Asia and Africa than in Europe and the Americas (Argentina, Brazil, Canada, Dominica, and the United States). Similarly, Meisenberg and Woodley (2013) report higher gains on the PISA and the TIMSS tests in countries with lower test scores at the beginning of the time period, from which they conclude that intelligence gains fade as a country's level of economic and social development reaches higher levels.

Like other psychometricians, PV distinguish between two subtypes of intelligence—fluid reasoning and crystallized knowledge (also, see Pinker (2018, pp 240-245)). The former tests the ability to perform "reasoning-based tasks that can be solved with (virtually) no prior knowledge," while the latter "consists of knowledge based questions that cannot be solved by reasoning" alone (PV, 2015, p. 284). Put another way, measures of fluid reasoning estimate the ability to perform analytical tasks (such as finding a number in an arithmetic progression), and measures of crystallized knowledge estimate the ability to answer successfully empirical questions (such as the name of a capital or the meaning of a word). Surprisingly, PV report that it is fluid reasoning, not crystallized knowledge, that has increased at the more rapid rate in recent decades. Between 1985 and 2013, the gains in fluid reasoning have increased by approximately 2.2 IQ points per decade, down from 4.3 IQ points between 1952 and 1985 (p. 285). Still, these recent gains in fluid reasoning, if less than previous gains, are considerably larger than the gains of 0.04 IQ points per decade in crystallized knowledge between 1987 and 2011 (p. 285), down from 3 points per decade between 1962 and 1987. Others report similar differentials in the rate of progress for these two subtypes of intelligence in Anglo-American countries (Jensen, 1998, pp. 319–320; Lynn, 2009a; Nisbett et al., 2012). Finn et al. (2014) find a Pearson correlation of 0.53 between fluid reasoning and math scores, but only a 0.36 correlation with reading performance, for a selected population of Massachusetts students. Rindermann and Thompson (2013) search for Flynn effects across waves of LTT surveys. They show mean increases in math performance of 2.37 IQ points per decade but only 0.54 points in reading, a finding that accords with PV's results that identify steeper upward trends in fluid reasoning than in crystallized knowledge.

In sum, the existing literature on student progress in math and reading is bifurcated and fragmentary, although most studies suggest that progress has been made on one or more surveys at least for some students at certain ages in some subjects during some of the period. What remains to be considered is the consistency in the estimates of progress in both subjects for students at various ages and for important subgroups. Thus, we report: 1) consistency of estimates policy makers received from agents asked to survey student progress; 2) differences in the rate of progress in math and reading; 3) variation in trends by age at which a child is tested; 4) changes in the rate of progress over time, giving special attention to the most recent period for which intelligence studies find a slowdown in IQ growth in industrialized societies; 5) heterogeneities by gender, ethnicity, and SES, as well as by SES quartile for each ethnic group.

Data

To measure cohort progress, the four agents periodically administered 160 waves of tests in math and reading to national probability samples of U. S. cohorts born between 1954 and 2007 (for the dates of each test administration, see Table A.2). The period that each survey covers as well as the number of observations for each set of waves are shown in Table 1. Individual level student data from the restricted-use data set that the National Center for Education Statistics (NCES) has provided are available to qualified researchers for all waves except for PIRLS 2001, PIRLS, 2016, PISA 2018 and TIMSS 2019 – the latter two are only used in cross country comparisons. Approximately seven million student-level observations of U. S. students and approximately four million observations of non-U. S. students are used in this analysis. Key features of the various test surveys are discussed in the principal-agent section above.

Ethnicity

The ethnic composition of those tested changed substantially over the 50-year period (Table A.3). The percentage of white students taking the LTT math test declined from 81 percent in its first wave in 1978 to 55 percent for its latest wave in 2012. The percentage of black students ticked only slightly upward (from 13 percent to 14 percent) over the period, but the percentage of Hispanic students increased dramatically from 5 percent to 23 percent, and the percentage of Asian (including those from the Pacific islands) students jumped from 1 percent to 6 percent. Despite marginal agency differences in ethnicity definitions (Table A.4), similar trends are observed across surveys.

Parental Education

Other than in the TIMSS and PIRLS 4th grade survey waves, estimates of parents' educational attainment are available. Although students are asked about their parents' education in different ways (Table A.5), all surveys allow for classification of the educational variable into one of four categories—whether or not the parent with the higher attainment level has: 1) a four-year college degree; 2) at least some post-secondary education, 3) no more than a high school diploma; or 4) no high school diploma. Parental education indicators are available for 15 surveys.

Possession Index

None of the agencies ask children about the annual income of the household. Instead, they inquire about various items that may or may not be available in the child's home. The possession index is a simple count of the number of items reported to be in the household. This measure provides a crude proxy for household permanent income. The correlation between the number of household possessions and household annual income is 0.36 in the Education Longitudinal Study (ELS), 2002, which contains information on both variables. Although admittedly lower than desirable, it provides a rough measure of permanent income that provides an alternative to eligibility for free or reduced lunch, an annual income indicator discussed in the Appendix. Reporting errors may be particularly large when younger students are asked about items with which they are not familiar. Also, the meaning of any home possession may vary with time. For example, a computer is rare among cohorts at the beginning of this period but common for those at the end of it. TIMSS and PISA inquire about many more possessions than do LTT and NAEP (Table A.6), so it is important to estimate relationships between possessions and achievement separately for each survey. The number of possessions also varies within surveys.

For the index, we count only the items which are available across all waves of each survey. Furthermore, we measure the variables consistently across all waves within a survey. The possession index is available for 17 surveys.

SES Index

The SES index, our preferred measure of SES, is estimated by extracting the first component from a principal component analysis obtained from the possession index and the original scale ranking of the parental education categories. To obtain as consistent a measure of SES as possible within each survey, we use the same measure of parental education and the same measure of household possessions across survey waves, even though on some waves additional information is available. The SES index is available for 15 surveys.

Methodology

We use the standard deviation (sd) of initial cohorts to calculate the distance (in sd) of the test-score distribution for all subsequent cohorts tested in each survey, subject, and age or grade. For each of these distributions, we estimate trends in mean performances over time by calculating the distance (in sd) of the test score distribution for each cohort's performance in each survey, subject, and age or grade level from the means of the initial cohort observations, which are set to zero. The assessments do not administer the full test to any one student. Instead, they estimate the performance of students on the test from their performance on the section administered to them. Each assessment provides various plausible values of that performance. Our empirical models use the second plausible value from each wave of the assessments. Results are robust to estimating the effects with the first five plausible values, and then taking their averages. We employ survey weighs in all estimations.

Following Reardon (2011) and Hanushek et al. (2020), we extract the performance trend with a quadratic function of the birth cohort for each year of the test administered with the following equations:

$$O_{sg\gamma i}^{t} = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \varepsilon_{sg\gamma ti} \quad \text{, where} \qquad (1)$$

O is the achievement score for student i, by subject s, assessment g, grade/age γ, and cohort t; α is the achievement trend, and ε is the error term. The results are robust to estimations from linear and cubic functions (Figures A.7-A.9).

We estimate the overall change and its associated standard error from the fitted point estimate and standard errors of the start and end points for each model (Feiveson, 1999; Gould, 1996; Oehlert, 1992; Philips & Park, 1988). We base our estimates on the delta method that calculates the variance, standard error, and Wald test statistic (z-test) from the nonlinear transformations of the estimated parameter vector from the fitted model. We repeat the above process separately for each subgroup analysis. We carry out all analyses in STATA 16.

Results

Unless otherwise indicated, all results are reported in sd per decade, which allows for direct comparisons with the Flynn effect originally estimated to be 3 IQ points or 0.21sd per decade. To avoid giving undue weight to outliers, we report median rather than the mean estimates of rates of cohort progress when summarizing results from several surveys. When summarizing results by the age a student is tested, we group surveys as follows: younger (age 9 and 4th grade); early adolescence (ages 13-15 and 8th grade); and older (age 17 and 12th grade). In estimates of differences in trends between dichotomously defined subgroups, a positive valence indicates more rapid progress by the group generally thought to be disadvantaged. A negative valence indicates the more advantaged group is making the greater progress, not that the trend is downward.

Table 1 displays the total amount of change and the rate of change in sd per decade observed by all the waves in each survey for each subject and age or grade level. The statistical significance of each trend line is shown in columns 5 and 9. But, since the waves for each survey differ both in number and the period covered, we show in columns 6 and 10 the change per decade in sd, which allows for more direct comparisons across surveys. Inspection of these columns reveals that, except for the PISA estimates, all have a positive sign indicating at least some progress in student achievement.

Agency Effects

Principal-agent problem

Principals receive different estimates of student progress from the different agents chosen to gather this information (Table 1). NAEP and TIMSS-PIRLS estimate the steepest upward trends. NAEP estimates that the median trend is 0.27sd per decade in math and 0.08sd in reading; TIMSS-PIRLS estimates that the median trend is 0.25sd in math and 0.09sd in reading (Tables 2 and 3). Note that both agencies adjust tests to fit the curriculum which the agency assumes is being taught in public schools at the time the test is administered. The median estimate trend provided by LTT is a more modest 0.15sd per decade in math and just 0.03sd in reading. Note that LTT ordinarily does not adjust for curricular changes. PISA, which estimates student preparation for participation in the post-secondary world at age 15, estimates a sizeable negative median trend of -0.10sd in math and a marginally negative trend of -0.02sd in reading. In other words, the median estimate of student progress in math can vary by as much as 0.37sd per decade, depending on the agency. It is thus risky to rely upon surveys by any one agency to reach conclusions about progress in student achievement. The large divergence of trends across surveys discourages us from combining them or providing average rather than median results.

PISA math deviation

The negative trend in PISA math (-0.10sd per decade) stands out as an especially clear exception to the positive estimates provided by the other agencies. Since PISA testing only began with the cohort born in 1985, it might be thought that PISA math has identified a recent downturn in performance at age 15. But other surveys, although at times showing diminished progress for cohorts born after 1990, do not show retrogression, as does PISA math (Table 4). Another possible explanation could be the reliability of tests administered at age 15. Some have argued that contemporary high school students, even those as young as age 15, are taking low-stakes tests less seriously than cohorts did in the past (Rindermann & Thompson, 2013). As testing has become increasingly pervasive and controversial in popular discourse, older students may increasingly see PISA math as a test that can be treated less seriously than other high-stakes tests (SAT, ACT, Advanced Placement and the like). This phenomenon could generate the appearance of deterioration in student achievement. However, if the cause is increasing test skepticism, it should appear on LTT and NAEP tests administered to still older students. It is true that these older students do not display the same growth as when students are tested at a younger age, but neither do they show the pronounced negative trend that PISA reports.

The content of the PISA math test itself is a possible explanation for some of the differences between its math trend and those registered in other surveys. PISA math assesses “an individual’s capacity to identify and understand the role that mathematics plays in the world (NCES, 2021a).” The assessment gives preference to questions that give the student a chance to display an ability to apply math to real world contexts rather than demonstrate facility in algebra, geometry and other analytic skills. For that reason, PISA math very likely requires the use of crystallized knowledge as much as fluid reasoning. Thus, PISA math and reading trends resemble one another in ways that stand in contrast with sharp differences in trends between the two subjects that other agents observe.

Flynn Effects

Differential trends by subject and by the age at which students are tested are consistent with PV's analyses of differential trends in fluid reasoning and crystallized knowledge intelligence. Subject and age differences are detected not only in the U. S. data but also in international data sets that the TIMSS and the PIRLS have assembled.

Math-reading differentials in the United States

The median decadal trend in math comes to 0.19sd, but it is only 0.04sd in reading or in literacy (Tables 2 and 3). The largest difference between the two subjects that NAEP has observed is for students in 4th grade, which shows a rate of change per decade of 0.39sd for math but only 0.08sd for reading. The result parallels the steeper upward trend for fluid reasoning than the crystallized knowledge in intelligence research.

Math-reading differentials in other countries

The differential rate of progress for the two subjects observed in the United States appears in other countries as well on the TIMMS-PIRLS surveys, although not on the PISA (Tables A.8-A.11). The latter survey finds no difference in trends for the two subjects on average for all the jurisdictions it surveys (-0.01sd per decade in math and reading). That could well be a function of the demands PISA math places on crystallized knowledge. However, the 4th grade

TIMSS-PIRLS surveys that IEA has administered show larger cohort advances in math (0.24sd) than in literacy (0.14sd). When one restricts the TIMMS-PIRLS samples to members of the OECD, even larger differences are observed between math (0.20sd) and literacy (0.03sd). The pattern holds in an attenuated form for non-OECD countries as well, with progress in math (0.32sd) moderately outpacing gains in literacy (0.28sd).

When one looks at specific countries in which 4th grade students have taken both TIMSS and PIRLS tests, progress in math is usually greater than in literacy. In the United Kingdom, decadal change on the math test is 0.36sd but only 0.06sd in literacy. In New Zealand, these numbers are 0.17sd, and -0.04sd, respectively. In Iran, the numbers are 0.33sd and 0.19sd, and in Hungary, they are 0.13sd and 0.04sd. In the United States, it is 0.20sd for math, and 0.08sd for literacy. Only in Singapore do we find literacy gains outpacing those in math. For Singapore, it is 0.26sd in math, and 0.33sd in reading. It is clearly possible for a country to educate its citizens in literacy at a rate that outpaces the math trend, as Singapore shows, but in general, students are making more progress in math than in reading.

Differential trends by age

In math, the steepest upward trend is observed among students tested at a younger age. For those age 9 or in the 4th grade, the decadal rate of change in math varies between 0.23sd on the LTT to 0.39sd on the NAEP. For those tested in early adolescence, the decadal rate of change is more modest, varying between a negative trend on the PISA (-0.10sd) to positive trends on the LTT (0.15sd), the TIMSS (0.19sd) and the NAEP (0.27sd). For older students, the progress proceeds at a rate of no more than 0.05sd on the NAEP and 0.06sd on the LTT. The international comparisons reveal a similar pattern of larger rate of progress in grade 4 than grade 8 math on the TIMSS (Tables A.8-A.9).

On reading tests, it is less clear that cohort advances in student achievement vary by the age at which a student is tested, mainly because there is little progress detected for any age group. For those at age 9 or in grade 4, the decadal rate of progress is 0.05sd on the LTT test, 0.08sd on the NAEP test and 0.09sd on the PIRLS test. For those in early adolescence, the rate of change is the same as for the younger students on the LTT test (0.03sd) and the NAEP test (0.12sd), but negative on the PISA test (-0.02sd). For older students, both estimates show little change by decade— 0.03sd on the NAEP test and hardly any change at all on the LTT test. In sum, progress in reading appears to have been modest for young students as well as those in early adolescence ages, but minimal among those tested at an older age.

Progress persistence

To see whether the rate of progress has persisted over time, we divide cohorts between those born before and after 1990. As shown in Table 4, the median difference in the math trends before and after this date comes to -0.08sd per decade, an indication of slowing progress. Yet there is considerable variation across surveys. The climb in math performances on the LTT test for students at age 9 slides from 0.22sd per decade for tests administered to cohorts born before 1990 to 0.16sd per decade for those born afterwards. But at age 13, the trend picks up from 0.13sd for those born before that date to 0.21sd for those born subsequently. At age 17, LTT math test performance declines from 0.10sd for cohorts born before 1990 to 0.02sd after that date.

NAEP math trends also vary by age. In the 4th grade, the trend shifts upward at a substantial rate of 0.65sd per decade prior to 1990 but slows to 0.21sd after that date. In the 8th grade, the progress slips from 0.44sd per decade for those born before 1990 to 0.08sd afterwards.

8th grade students also perform better on the TIMMS math test if born before (0.35sd) rather than after (0.14sd) that date. But on the PISA test, the steep decline for those born before 1990 (-0.36sd) slows among those born subsequently (-0.06sd). Summing up, math performance fails to persist at the same rate among cohorts born after 1990 in five of the seven surveys for which a comparison can be made. The PISA math test proves to be an exception to the dominant pattern, as are performances at age 13 on the LTT test.

In reading, progress continues and even accelerates for cohorts born after 1990. The median difference in trends before and after that date is 0.08sd per decade. Once again, surveys vary. On the LTT test administered to those at age 9, progress is greater after (0.15sd) than before (0.01sd) 1990. The same is true for the 4th grade NAEP test: before (-0.04sd); after (0.07sd). In early adolescence, the LTT test again shows greater progress for the more recent period (before 0.01sd; after 0.22sd). Although the NAEP test does not: (before 0.22sd; after 0.10sd), the PISA test points in the positive direction: (before -0.31sd; after -0.04sd). Among older students, the LTT test (before, 0.00sd; after 0.09sd) shows a marked increase, but the NAEP test (before, 0.05sd; after -0.07sd) shows quite the opposite. In sum, progress persistence in reading is observed in five of seven surveys. The NAEP tests administered to those in 8th and 12th grades are the outliers.

In summary, birth cohorts born after 1990, as compared to those born before that date, show a median trend that is a negative -0.08sd per decade in math but a positive 0.08sd per decade in reading. Five of seven surveys reveal a more gradual trend upward in math after 1990, and five of the seven surveys show a steeper one in reading after that date. The pattern is consistent with research on intelligence, which suggests a moderation of growth in fluid reasoning but not necessarily any change in crystallized knowledge. Still, the number of comparisons is few, and results are not uniform across surveys (Figures A.1-A.6). Moreover, in most cases we are unable to improve model fits significantly by using quadratic or cubic models rather than a linear one that assumes constant growth over time (Figures A.7-A.9). That implies that change has occurred steadily over the past fifty years undisturbed by the many apparent perturbations that might have disrupted the learning process.

The exceptions to this generalization are the NAEP surveys in both 4th grade and 8th grade math, which show a steep upward trend for birth cohorts born in the 1970s and 1980s, with only a modest increment subsequently. This exception would be easier to interpret were it not contradicted by the steady but less dramatic increase in the LTT throughout the entire period. Altogether, the signals are uncertain.

Heterogeneities

Gender, ethnic, and SES heterogeneities are observed. For gender, the median trend reveals a male advantage at an early age, but that fades by the end of secondary school. For ethnicity, we detect moderately slower gains for white than for Asian, African American, and Hispanic students. For SES, we observe somewhat larger gains for the lowest quartile as compared to the top quartile of the distribution when students are tested at a younger age, but that advantage attenuates for white, African American and, to a lesser extent, for Hispanic students. That pattern is not observed for Asian students, however.

Gender

There are not large differences in performance trends by gender. Across all surveys, the median differences in the achievement advantage of males trends downward in math (-0.02sd), but in reading (0.02sd) it moves in the opposite direction (Tables A.12-A.14). The male advantage is more noticeable when students are tested at a younger age (0.03sd in math, 0.04sd in reading) than in early adolescence (zero in math, 0.01sd in reading). The male achievement advantage disappears (-0.01sd in math and zero in reading) when students are tested at an older age.

Ethnicity

In nearly all surveys, trends in performances of non-white cohorts shift upward more rapidly than those for white ones (Tables 5-7). Only small differences are observed between reading and math.

The median difference in Hispanic-white and black-white progress on all math surveys is 0.11sd per decade. The median Asian-white difference is 0.09sd. Among those of elementary age, the median difference is largest between white and Asian (0.18sd). On the four tests given in math to students in early adolescence, the median differences in the upward trend between white students and Asian, Hispanic, and black students are 0.12sd, 0.11sd and 0.09sd per decade, respectively (Tables 2 and 3). The relative position of white students usually continues to lag in high school. On the LTT math test, the decadal change for white students (0.09sd) trails that for black (0.15sd) and Hispanic (0.16sd) students, but it outpaces the decadal change that Asian cohorts register (0.05sd). On the NAEP math test, white progress is 0.09sd per decade, behind those of both Hispanic (0.16sd) and Asian (0.21sd) students, but beyond the decadal change of African American ones (0.07sd).

Differences in trends by ethnicity are about the same for reading as for math performances. In the former subject, the largest difference is between Asians and whites, a median difference for all surveys of 0.15sd per decade. The median white-black and white-Hispanic differences in reading are 0.09sd and 0.06sd per decade.

SES

When cohort progress is estimated by quartiles of the SES distribution, 10 of the 15 estimates show marginally larger upward trends for those in the lowest quartile than for those in the top quartile (Table 8). Four show more progress for those from the highest SES backgrounds, and one shows no difference. The trend lines vary by the point in the life cycle when the student is tested, and the differences are more favorable for disadvantaged students tested at a younger age. For those tested in math at the youngest age, the median trend is 0.08sd larger for the bottom quartile, for those tested in early adolescence, it is 0.03sd, and for the oldest cohort, the top quartile shows more progress than the lowest quartile (-0.03sd) (Table 2). The median differential in reading is large for students tested at the youngest age (0.15), lesser for those tested in early adolescence (0.11sd) and for those tested toward the end of secondary education, the trends are more favorable for the top quartile (-0.04sd) (Table 3). The PISA test, which is an outlier, shows a differential at age 15 in favor of the lowest SES quartile of 0.24sd in math and 0.29sd in reading.

Consistency of SES effects across ethnic groups

NAEP has enough observations in each wave to estimate with precision achievement progress by SES quartile for each ethnic group. We report the differences between the highest and lowest SES quartiles on NAEP tests for each ethnic group in Tables 2 and 3 (see Appendix Tables A.15 – A.18 for detailed analyses). To avoid dependence on results from any one agent, we also report median trends from all surveys, even when estimations are less precise. As elsewhere, differentials are given a positive valence if the lowest SES quartile shows greater progress than the highest quartile.

Results from both the NAEP test and from the median of all surveys vary by age. When white and African American students are tested at a young age, gains that cohorts of students made in the lowest SES quartile are substantially larger than those that the cohorts made in the highest SES quartile. But the trend toward greater achievement equality attenuates by early adolescence, and reverses itself by the time students reach the end of secondary schooling, with the top SES quartile among the older students showing a steeper upward trend than the lowest SES quartile. Attenuation is less apparent for Hispanic students and especially so for Asian students.

The downward shift in progress made by low relative to high SES students is quite pronounced for white students. On the NAEP, the lowest SES quartile of white students tested at a younger age outpaces the top quartile in math by 0.06sd per decade, but the advantage shifts to -0.01sd in early adolescence and increases to -0.11sd by the end of secondary education. On the median for all surveys, these numbers are 0.04sd, 0.05sd, and -0.04sd, respectively. On reading tests, the rate of progress on the NAEP declines from 0.20sd to 0.04sd, and then reverses to -0.14sd for the older students; the median for all surveys declines from an upward trend of 0.15sd to 0.00sd, and then reverses itself to -0.10sd.

For African American students, a similar pattern is apparent. The rate of cohort progress among the lowest quartile of students tested on NAEP's math exam tested at the younger age outpaces the highest quartile by 0.14sd, but that advantage fades to 0.08sd in early adolescence and then dramatically reverses itself (-0.22sd). The median trend numbers for all surveys are less dramatic but otherwise quite similar: 0.11sd, -0.03sd, and -0.09sd, respectively. On the NAEP reading test, the relative advantage of the lowest quartile in the rate of progress once again slows as students age: 0.15sd, 0.06sd, and -0.02sd. The median for all surveys shows a similar pattern: 0.12sd, 0.06sd and 0.01sd.

Among Hispanic students, a similar pattern is apparent on the NAEP tests but less so on median results from all surveys. Those students in the lowest quartile tested on the NAEP exam in math at a younger age show gains relative to those in the highest quartile of 0.22sd. The advantage fades to 0.10sd in early adolescence and turns negative (-0.06sd) at the end of secondary education. On the reading test, the numbers are 0.28sd, 0.09sd, and 0.18sd. The median trends for all surveys do not vary as much in either math (0.10sd, 0.03sd, and 0.01sd) or in reading (0.17sd, 0.17sd, and 0.11sd).

The progress shown by the lowest SES quartile among Asian students shows a somewhat different pattern than for white and African American students. On the NAEP test, students at a younger age show gains in math registered by the lowest quartile that exceeds the top quartile by 0.08sd and in reading 0.24sd per decade. In early adolescence, the lowest quartile trails the top quartile in math (-0.03sd), but exceeds it in reading (0.12sd). Among older students, the lowest SES quartile continues to outpace the highest quartile on both math (0.02sd) and reading (0.12sd) tests. Median trends from all surveys in both subjects show greater growth for the lowest quartile across all three age groups—in math 0.19sd, 0.08sd, and 0.09sd, in reading 0.06sd, 0.11sd, and 0.12sd, respectively. In sum, there is little evidence of a downward shift in the degree of progress made as Asian students age.

Consistency of SES effects across different constructs

As mentioned, the SES index is constructed from student reports of parental education and possessions in the home. To ascertain whether we obtain similar results with alternative SES indicators, we estimate trends by parental education, possessions in the home, eligibility for free and reduced lunch, and family structure. For all of these alternative ways of measuring SES, we generally see more progress for students from less advantaged backgrounds than from those who are more advantaged (see text in Appendix, Tables A.7 and A.19 – A.27 for details).

Discussion

Limitations and Agency Effects

This study is descriptive in nature and is unable to estimate causal effects. However, it does address an important question—how much progress in math and reading are cohorts of U. S. students making?—This question is of interest to principals for the very good reason that the rate of student progress is critical both for future generations and for the nation's well-being.

It is no easy task to answer this question. The four agents asked to track progress in student achievement are to be heralded for their ambition, objectivity, resourcefulness, and endurance. But the answer to the basic question—how much progress? —remains elusive. At early adolescence, the median reading estimate varies from -0.02sd per decade on the PISA test to 0.12sd per decade on the NAEP test, a range of 0.14sd per decade or 0.70sd, if extrapolated to the entire fifty-year period. In math, the median estimate ranges wildly from -0.10sd per decade retrogression on the PISA math test to 0.27 on the NAEP math test, a range of 0.37sd per decade or, by extrapolation, 1.85sd over the fifty-year period. Depending on which agency's estimate is taken as authoritative, the public commentator can either lament the downward dive in American education or celebrate extraordinary advances forward.

Agency effects are large in part because the information obtained is subject to more than just sampling error. The practice of linking waves by "bridging questions"—items repeated in subsequent waves—is not an exact science. Furthermore, agencies have chosen questions with different purposes in mind and with varying commitments to linking the waves of their surveys. Two surveys—the NAEP and the TIMMS-PIRLS that show the most progress for students at younger age and at early adolescence—are designed to test knowledge and skills against the contemporary curriculum. Both agencies could be administering tests that correctly observe student achievement by asking them questions on topics they are studying. But the tests may be over-estimating progress by shaping the exams to a curriculum that has lower expectations for students than was once the case. In other words, both NAEP and IEA (the agency that administers TIMMS-PIRLS) could have adjusted the difficulty of their tests downward in the process of adjusting questions to the contemporary curriculum. Meanwhile, the PISA test shows regression in student performance, especially among those at the upper end of the SES distributions. The PISA test could be the one and only honest broker, telling the truth about the "best and the brightest" in the United States. Or the PISA test could be shifting the bar upwards. It could be setting even higher standards for what it takes to be prepared for the larger society, especially for those at the upper end of the SES spectrum, while lowering the bar for those at the other end of the distribution. There is no reason to think that PISA is deliberately adjusting its tests in this manner. Still, the number of non-OECD countries participating in PISA has increased over time (NCES, 2021b), a trend that could be facilitated by adjusting the performance bar downward for students at the lower end of the distribution. Also, PISA, more than the other agencies, has cut the link between curriculum and testing, leaving the agency to imagine for itself what is needed to be ready for life at age 15. More certain is the danger of ignoring agency effects when estimating trends, as studies do when they cobble trends together from surveys designed by multiple agents without controlling for agency effects. Given the size of the agency effects shown here, such studies should be treated with unabashed skepticism.

Student background variables used in our study are measured in alternative ways. Although we standardize education and possession indicators used in the SES index across waves of each survey, questions used by agents to obtain this information vary in number and type (see Tables A.4 – A.6). Meanwhile, the meaning of categories may be changing. A high school diploma has a different meaning by the end of the period than it does at its beginning. To minimize dependence on any specific data collection strategy, we focus on consistency of results by agency and, to avoid excessive dependence on outliers, report median—not mean—results when summarizing information across surveys.

Summary of results

A number of findings are nonetheless robust to most estimations. Most important, progress in student achievement by cohorts of students in the United States is observed in 15 of 17 surveys. Only the two PISA surveys suggest retrogression. Much greater progress is observed in math (median estimate: 0.19sd) than in reading (0.04sd). A similar difference is observed in other countries for 4th grade performance on the TIMSS math and the PIRLS literacy tests.

The pace of the upward trend may have slowed in math even while it has increased in reading for cohorts born since 1990, although some exceptions make these findings less than certain. Math trends are larger for students tested at a younger age than for those tested in the middle years or in high school, but reading gains, always small, do not vary much by the age when students are tested. The trend is more steeply upward for cohorts of nonwhite students than for white ones in both reading and math. The largest group difference is the median trend disparity between Asian and white cohorts (0.09sd in math and 0.15sd in reading). These numbers are 0.11sd in math and 0.06sd in reading between Hispanic and white cohorts, and 0.11sd in math and 0.09sd in reading between black and white cohorts. When tested at an earlier age, the median upward trend is greater for white, black, and Hispanic students in the lowest SES categories than for those in the highest ones, but that fades as students age and the highest SES quartile shows the greater rate of progress when students are tested as they are about to leave secondary schooling. The reversals by SES as students age is also detected, separately, among white, black, and Hispanic students. Among Asian students, the lowest quartile shows greater progress (except for the 8th grade LTT reading test and the NAEP math test) no matter at what age the student is tested.

Potential moderators

Ethnic and SES heterogeneities appear to be driven by changes in families and schools over the half century. But differential changes by subject matter, and perhaps also by the age when students are tested, suggest broad changes in nutrition, reduction in contagious diseases and greater protection against environmental risks, factors not always considered in the sociology and economics of education literature, which may also moderate the relationship between birth cohort and student achievement.

Family-school interactions

Changes within families and schools may moderate shifts in ethnic group differences during this period. The increasingly large investments that Asian American families make in their children's education very likely account for the group's more rapid rate of growth in reading (0.15sd) and in math (0.09) than for white students. The change of 0.15sd per decade in reading extrapolates to 0.75sd over the past half century. Schools are not likely to be the driving force for a change of this magnitude, given the fact that Asians and whites attend much the same public schools and that special education programs only occasionally serve Asian students. The disproportionate Hispanic (0.11sd, 0.06sd) and African American (0.11sd, 0.09sd) student rates of progress relative to the white student rate in math and in reading, respectively, could be a function of changes in both families and schools. Hispanic and African American parents are increasingly well educated, earn higher incomes, and have fewer children, all of which correlate with higher achievement levels. At the same time, school interventions have often focused on redressing differentials between whites and these two ethnic groups. School desegregation (Rivkin, 2016; Rivkin & Welch, 2006; Welch & Light, 1987), class size reduction (Krueger, 2003), Head Start (Morris et al. (2018)), equity law suits (Jackson, Johnson, & Persico, 2016; Lafortune, Rothstein, & Schanzenbach, 2018), English Language Learner programming (Shin, 2018), school accountability requirements (Dee & Jacobs, 2011), and greater school choice opportunities (Cheng & Peterson, 2020) have all been interpreted as having disproportionately positive impacts on disadvantaged minorities in certain places and at particular times over the past half century.

Many equal opportunity programs have focused on preschool and elementary education. It may not surprise us, then, to see greater progress on reading tests (0.15sd) and on math tests

(0.08) for children in the lowest (as compared to the highest) SES quartile when tested in elementary school. Unfortunately, these strongly positive gains apparent in elementary school fade when children enter early adolescence and then turn negative (-0.04sd in reading, -0.03sd in math) by the time students near the end of their secondary education. This pattern appears for low SES students, regardless of whether their ethnic background is white, black, or Hispanic. It is not clear whether attenuation and reversal of positive trends can be attributed to a changing peer group culture or to inadequacies within the school system, or a combination thereof. But efforts to equalize opportunities by socio-economic background may consider placing increased emphasis on the educational experiences of disadvantaged students in secondary school.

Broader social forces

Whatever the role families and schools have played in the moderation of heterogeneities across ethnic and SES groups, they probably do not account for the much greater progress in math (0.19sd) than in reading (0.04sd) per decade, a near five-fold difference. Extrapolated to a half century, the rate of change for the subject that places the most demands on fluid reasoning has been 0.95sd as compared to just 0.20sd for the subject that places a greater demand on crystallized knowledge.

The original Flynn effect observes a steady increase of 0.21sd per decade. PV refine this result by showing in a meta-analysis of all sequential intelligence studies that in recent decades in industrialized societies that fluid reasoning—the ability to comprehend and manipulate abstract relationships—has trended upward at a rate of 0.15sd per decade, a much steeper rate than the 0.03sd upward trend for crystallized knowledge, the ability to understand and interpret the physical and social environment. Notably, those rates of change in the two types of intelligence closely approximate the decadal rates of progress in math (0.19sd) and reading

(0.04sd) in the median survey reported here. Studies in neurobiology and brain imaging confirm the distinction between fluid and crystallized intelligence (Blair, 2006; Horn & McArdle, 2007). The former is associated with the brain's prefrontal cortex (PFC). Damage to the PFC appears to have little effect on crystallized intelligence (Waltz et al., 1999). PV say the different rate of change for the two subtypes of intelligence could be due to improved nutrition and health care (Lynn, 2009b), disease containment (Eppig, Fincher, & Thornhill, 2010; Van Panhuis et al., 2013), and reductions in environmental risks (for example, lead poisoning (Kaufman et al., 2014) and air pollution (Chay & Greenstone, 2003)). Each may enhance the brain's analytical capacities during prenatal and infant stages of the life cycle. For example, "fending off aversive pathogens necessitates considerable amounts of energy, thereby removing important resources from brain development in early childhood. (p.293)." If student performances on math tests (other than the PISA test) depend more on the subtype of intelligence known as fluid reasoning than on the subtype known as crystallized knowledge, then the differential progress in the two subjects suggests that progress in math is primarily due to factors operating in early childhood and even prior to birth when brain capacity is most malleable. That math gains are registered most clearly among children tested at a younger age and the upward trend is not as consistently observed among older students adds weight to the suggestion that decisive influences are operating in very early childhood, perhaps even prior to birth. Similarly, Bloom et al. (2008) show that average annual gains on seven nationally normed tests are largest in elementary grades, then there are declines at early adolescence, and then those gains reach a minimum at grade 12. Further, median math gains diminish by 0.08sd for those born after 1990, perhaps a sign that the returns to improvements in nutrition and health have begun to diminish.

The focus on families and schools in the sociology and economics of education may have distracted attention away from broader social forces that may be at least as important. Recently, school closings in response to the Covid pandemic seem to have had a negative impact on learning for an entire generation of students (Kuhfeld et al., 2020) much as children suffered educational setbacks from school closures during wars (Ichino and Winter-Ebmer, 2004), strikes (Jaume and Willén, 2019; Belot and Webbink, 2010), and weather events (Goodman, 2015; Sacerdote, 2012). Indeed, PV (p. 285) detect a slowdown in intellectual growth during World War II, a likely byproduct of both school closures and worldwide disruptions of economic and social progress. Gradual alterations of the economic, social, and physical environment over the past half century may well have had their own consequences on student achievement. It may be that drivers of change in student achievement over the past fifty years stem in part from broad changes in social and economic well-being. The steady, small pace of achievement progress over decades, together with the larger rate of change in math than in reading achievement, suggests a larger social context shaping the way in which families and schools affect learning. This finding does not mean that families and schools are not moderators as well. The reading gains of 0.08sd per decade for cohorts born after 1990 (even as math progress was abating) may be a function of enhanced effectiveness of families and schools in the most recent period. It is worth noting that many school reforms mentioned above—school accountability, equitable financing, English Language Learner policies, and school choice—have had their greatest impact on cohorts born since 1990.

Families and schools matter. There is much to be said for finding ways to improve their effectiveness, particularly for students during their adolescent years. But if the sharp difference between math and reading implies that the broader social, economic, and physical environments are no less important, then policies should also focus on the earliest possible interventions. A coalition of educators called “The Bolder, Broader Approach to Education” says the country should ensure “that every student arrives at kindergarten with the benefit of high-quality early learning and necessary health, wellness, and family support services from birth.” (BBA, 2021). Our findings are consistent with this recommendation. It is reasonable to infer from our research that universally beneficial policies for “every student . . . from birth” could have as much impact on student achievement, especially in math, as focused interventions attempted later on.

References

Ang, S., Rodgers, J. L., & Wänström, L. (2010). The Flynn Effect within subgroups in the US: Gender, race, income, education, and urbanization differences in the NLSY-Children data. *Intelligence*, *38*(4), 367-384.

Bauerlein, M. (2008). *The dumbest generation: How the digital age stupefies young Americans and jeopardizes our future.* New York, NY: Jeremy P. Tarcher/Penguin.

BBA. (2021). Out-of-school experiences. Early childhood education. We need a broader, bolder approach. Retrieved from: https://www.boldapproach.org/topic/early-childhood-education/index.html.

Belot, M., & Webbink, D. (2010). Do teacher strikes harm educational attainment of students?. *Labour*, *24*(4), 391-406.

Berliner, D. and Biddle, B. (1996). Making molehills out of molehills: Reply to Lawrence Stedman's review of the manufactured crisis. *Education Policy Analysis Archives* 4(3).

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences*, *29*(2), 109.

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, *1*(4), 289-328.

Borghans, L., Golsteyn, B. H., Heckman, J. J., & Humphries, J. E. (2016). What grades and achievement tests measure. *Proceedings of the National Academy of Sciences*, *113*(47), 13354-13359.

Bovens, M., Goodin, R. E., & Schillemans, T. (eds) (2014). *Oxford handbook of public accountability*. Oxford, UK: Oxford University Press.

Bracey, G. (1992). The second Bracey report on the condition of public education. *Phi Delta Kappan 74*(2): 104–117.

Braun, D., & Guston, D. H. (2003). Principal-agent theory and research policy: An introduction. *Science and public policy*, *30*(5), 302-308.

Broer, M., Bai, Y., & Fonseca, F. (2019). *Socioeconomic inequality and educational outcomes: Evidence from twenty years of TIMSS.* IEA Research for Education and Springer Open.

Campbell, J. R., Reese, C. M., O'Sullivan, C., & Dossey, J. A. (1996). *NAEP 1994 trends in academic progress*. Washington, DC: U.S. Department of Education.

Chay, K. Y., & Greenstone, M. (2003). The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. *The quarterly journal of economics*, *118*(3), 1121-1167.

Cheng, A., & Peterson, P. E. (2020). Experimentally estimated impacts of school vouchers on educational attainments of moderately and severely disadvantaged students. (EdWorkingPaper: 20-221). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/622r-tk70

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American economic review*, *104*(9), 2633-79.

Chmielewski, A. K. (2019). The global increase in the socioeconomic achievement gap, 1964 to 2015. *American Sociological Review*, *84*(3), 517-544.

Copperman, P. (1978). *The literacy hoax: The decline of reading, writing, and learning in the public schools and what we can do about it*. New York, NY: William Morrow & Company.

Dee, T. S., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and management*, *30*(3), 418-446.

Eppig, C., Fincher, C. L., & Thornhill, R. (2010). Parasite prevalence and the worldwide distribution of cognitive ability. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1701), 3801-3808.

Feiveson, A. H. (1999). FAQ: What is the delta method and how is it used to estimate the standard error of a transformed parameter? http://www.stata.com/support/faqs/stat/deltam.html.

Finn, A. S., Kraft, M. A., West, M. R., Leonard, J. A., Bish, C. E., Martin, R. E., . . . Gabrieli, J. D. E. (2014). Cognitive skills, student achievement tests, and schools. *Psychological Science, 25*, 736–744.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological bulletin*, *95*(1), 29.

Goodman, J. (2015). In defense of snow days: Students who stay home when school is in session are a much larger problem. *Education Next, 15*(3), 64-69.

Gould, W. W. (1996). crc43: Wald test of nonlinear hypotheses after model estimation. Stata Technical Bulletin 29: 2–4. Reprinted in Stata Technical Bulletin Reprints, vol. 5, pp. 15–18. College Station, TX: Stata Press.

Grissmer, D. W., Kirby, S. N., Berends, M., & Williamson, S. (1994). Student Achievement and the changing American family, Santa Monica, CA: RAND. *Institute for Education and Training*.

Hanushek, E. A., Peterson, P. E., Talpey, L. M., & Woessmann, L. (2020). *Long-run trends in the US SES-Achievement gap* (No. w26764). National Bureau of Economic Research.

Hanushek, E., Peterson, P. E., & Woessman, L. (2012). Achievement growth: International and U.S. state trends in student performance. PEPG Report 12–03, Program on Education Policy and Governance, Harvard University.

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature, 46*(3), 607–668.

Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of economic growth*, *17*(4), 267-321.

Hashim, S. A., Kane, T. J., Kelley-Temple, T., Laski, M. E., & Staiger, D. O. (2020). *Have income-based achievement gaps widened or narrowed?* (No. w27714). National Bureau of Economic Research.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, *269*(5220), 41-45.

Hedges, L.V., Nowell, A. (1998). Black-White test score convergence since 1965. In Jencks, C., Phillips, M. (Eds.), *The Black-White test score gap* (pp. 149–181). Washington, DC: Brookings Institution Press.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York, NY: Free Press.

Hirsch, E. D. (1987). *Cultural literacy.* Boston, MA: Houghton Mifflin Company.

Horn, J. L., & McArdle, J. J. (2007). Understanding human intelligence since Spearman. In R. Cudeck & R. MacCallum (Eds.), *Factor analysis at 100 years* (pp. 205–247). Mahwah, NJ: Erlbaum.

Hutchison, D., & Schagen, I. (2007). Comparisons between PISA and TIMSS—Are we the man with two watches? In T. Loveless (Ed.), *Lessons Learned: What international assessments tell us about math achievement* (pp. 227–262). Washington, DC: Brookings Institute Press.

Ichino, A., & Winter-Ebmer, R. (2004). The long-run educational cost of World War II. *Journal of Labor Economics*, *22*(1), 57-87.

IEA (2021). About TIMSS 2015. Retrieved from: http://timss2015.org/timss-2015/about-timss-2015/

Itzkoff, S. W. (1994). *The decline of intelligence in America: A strategy for national renewal*. Westport, CT: Praeger.

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

Jaume, D., & Willén, A. (2019). The long-run effects of teacher strikes: Evidence from Argentina. *Journal of Labor Economics*, *37*(4), 1097-1139.

Jencks, C. & Phillips, M. (Eds.). (1998). *The Black-White test score gap*. Washington, DC: Brookings Institution Press.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Kaufman, A. S., Zhou, X., Reynolds, M. R., Kaufman, N. L., Green, G. P., & Weiss, L. G. (2014). The possible societal impact of the decrease in US blood lead levels on adult IQ. *Environmental research*, *132*, 413-420.

Krueger, A. (2003). Economic considerations and class size. *Economic Journal, 113*,34–63.

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher*, *49*(8), 549-565.

Lafortune, J., Rothstein, J., & Schanzenbach, D. W. (2018). School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics*, *10*(2), 1-26.

Lynn, R. (2009a). Fluid intelligence but not vocabulary has increased in Britain, 1979-2008. *Intelligence*, *37*, 249–255.

Lynn, R. (2009b). What has caused the Flynn effect? Secular increases in the development quotients of infants. *Intelligence*, *37*(1), 16-24.

Magnuson, K., & Waldfogel, J. (2008). *Steady gains and stalled progress: Inequality and the black-white test score gap*. New York, NY: Russell Sage.

Matheny, K. T., Thompson, M. E., Flores, C. T. & Reardon, S. F. (2021). Uneven progress: Recent trends in academic performance among U.S. school districts. Retrieved from: https://edopportunity.org/papers/seda%20district%20trends%20paper.pdf

Meisenberg, G., & Woodley, M. A. (2013). Are cognitive differences between countries diminishing? Evidence from TIMSS and PISA. *Intelligence*, *41*(6), 808-816.

Miller, L. S. (1995). *An American imperative: Accelerating minority educational advancement*. Yale University Press.

Morris, P. A., Connors, M., Friedman-Krauss, A., McCoy, D. C., Weiland, C., Feller, A., ... & Yoshikawa, H. (2018). New findings on impact variation from the Head Start Impact Study: Informing the scale-up of early childhood programs. *AERA Open*, *4*(2), 2332858418769287.

Mullis, I. V. S., & Martin, M. O. (Eds.). (2015). *PIRLS 2016 Assessment Framework* (2nd ed.). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/pirls2016/framework.html

National Center for Education Statistics. (2021a). NCES handbook of survey methods technical report. *National Center for Education Statistics*. Retrieved from: https://nces.ed.gov/statprog/handbook/pisa_keyconcepts.asp

National Center for Education Statistics. (2021b). Participation in PISA by year. Retrieved from: https://nces.ed.gov/surveys/pisa/countries.asp

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington, DC: U.S. Government Printing Office.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American Psychologist*, *67*, 130–159.

OECD. (2018). *Equity in education: Breaking down barriers to social mobility*. Paris, France: OECD Publishing.

Oehlert, G. W. (1992). A note on the delta method. *American Statistician, 46*(1): 27–29.

Pellegrino, J. W., Jones, L. R., & Mitchell, K. J. (Eds.). (1999). *Grading the nation's report card: Evaluating NAEP and transforming the assessment of educational progress*. Washington, DC: National Academy Press.

Phillips, M. Crouse, J. & Ralph, J. (1998). Does the black-white test score gap widen after children enter school?. In Jencks, C., Phillips, M. (Eds.), *The Black-White test score gap* (pp. 229–272). Washington, DC: Brookings Institution Press.

Phillips, P. C. B., & Park. J. Y. (1988). On the formulation of Wald tests of nonlinear restrictions. *Econometrica 56*: 1065–1083.

Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the Flynn effect (1909–2013). *Perspectives on Psychological Science*, *10*(3), 282-306.

Pinker, S. (2018). *Enlightenment now: The case for reason, science, humanism, and progress*. New York, NY: Viking.

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–115). New York, NY: Russell Sage.

Reardon, S. F., Valentino, R. A., & Shores, K. A. (2012). Patterns of literacy among US students. *The Future of Children*, 17-37.

Rindermann, H., & Thompson, J. (2013). Ability rise in NAEP and narrowing ethnic gaps?. *Intelligence*, *41*(6), 821-831.

Rivkin, S., & Welch, F. (2006). Has school desegregation improved academic and economic outcomes for blacks?. *Handbook of the Economics of Education*, *2*, 1019-1049.

Rivkin, S. (2016). Desegregation since the Coleman report. *Education Next,16*(2), 28–37.

Rodgers, J. L., & Wänström, L. (2007). Identification of a Flynn Effect in the NLSY: Moving from the center to the boundaries. *Intelligence*, *35*(2), 187-196.

Sacerdote, B. (2012). When the saints go marching out: Long-term outcomes for student evacuees from Hurricanes Katrina and Rita. *American Economic Journal: Applied Economics*, *4*(1), 109-35.

Shakeel, M. D., & Peterson, P. E. (2020). Changes in the performance of students in charter and district sectors of US education: An analysis of nationwide trends. *Journal of School Choice, 14*(4), 604-632.

Shin, N. (2018). The effects of the initial English language learner classification on students' later academic outcomes. *Educational Evaluation and Policy Analysis*, *40*(2), 175-195.

Stedman, L. (2009). *The NAEP long-term trend assessment: A review of its transformation, use, and findings*. Washington DC: National Assessment Governing Board.

Sykes, C. J. (1995). Dumbing down our kids: *Why American children feelgood about themselves but can't read, write, or add*. New York: St. Martin's Press Griffin.

Trahan, L. H., Stuebing, K. K., Fletcher, J. M., & Hiscock, M. (2014). The Flynn effect: a meta-analysis. *Psychological bulletin*, *140*(5), 1332.

U. S. Department of Education. (2014). Program for International Student Assessment (PISA). Retrieved from: https://nces.ed.gov/statprog/handbook/pisa.asp

Van Panhuis, W. G., Grefenstette, J., Jung, S. Y., Chok, N. S., Cross, A., Eng, H., ... & Burke, D. S. (2013). Contagious diseases in the United States from 1888 to the present. *The New England journal of medicine*, *369*(22), 2152.

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, Thomas, C. R., & Miller, B. L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological science*, *10*(2), 119-125.

Wai, J., & Putallaz, M. (2011). The Flynn effect puzzle: A 30-year examination from the right tail of the ability distribution provides some missing pieces. *Intelligence*, *39*(6), 443-455.

Welch, F., & Light, A. (1987). *New evidence on school desegregation*. Washington, DC: U.S. Commission on Civil Rights.

Table 1: Change in achievement levels in math and reading by grade/age and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Math				Reading			
Survey	Grade/Age	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
LTT	Age 9	1968-2002	89,170	0.783** (0.012)	0.23	1961-2002	137,750	0.222** (0.011)	0.05
NAEP	Grade 4	1980-2007	1,455,630	1.051** (0.002)	0.39	1980-2007	1,606,030	0.227** (0.002)	0.08
TIMSS	Grade 4	1985-2005	51,440	0.610** (0.011)	0.31				
PIRLS	Grade 4					1991-2006	26,100	0.129** (0.015)	0.09
LTT	Age 13	1964-1998	101,200	0.493** (0.011)	0.15	1957-1998	142,060	0.126** (0.012)	0.03
NAEP	Grade 8	1976-2003	1,301,960	0.716** (0.003)	0.27	1976-2003	1,423,560	0.325** (0.002)	0.12
TIMSS	Grade 8	1981-2001	57,030	0.374** (0.011)	0.19				
PISA	Age 15	1985-2000	29,130	-0.145** (0.016)	-0.10	1985-2000	25,230	-0.031** (0.016)	-0.02
LTT	Age 17	1961-1995	92,460	0.218** (0.013)	0.06	1954-1995	137,350	0.013** (0.011)	0.00
NAEP	Grade 12	1988-1998	71,500	0.048** (0.009)	0.05	1973-1998	145,160	0.071** (0.008)	0.03

Note: Table shows change in achievement levels by grade/age and survey. Normalized achievement is measured in standard deviations (s.d.). The s.d. is the difference between the year test was administered and the starting year for a specific test series. The changes have been estimated from a quadratic fit. Change per decade (dc) in achievement levels is also displayed. Standard error are in parenthesis. *p <0.05, **p<0.01.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 1990-2017 Main NAEP, 1971-2012 LTT NAEP, 1995-2015 TIMSS, 2000-2015 PISA and 2001-2016 PIRLS.

Table 2: Medians of change/decade in achievement levels in math by subgroups and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Base	Subgroup	Young	Diff.	E. A.	Diff.	Older	Diff.	All age	Diff.
	All	0.31		0.17		0.06		0.19	
	LTT	0.23		0.15		0.06		0.15	
	NAEP	0.39		0.27		0.05		0.27	
	TIMSS	0.31		0.19				0.25	
	PISA			-0.10				-0.10	
Bir yr: Until 1990	Until 1990	0.44		0.24		0.10		0.22	
	Sin. 1990	0.19	-0.25	0.11	-0.13	0.02	-0.08	0.14	-0.08
Gender: Female	Female	0.29		0.17		0.06		0.20	
	Male	0.32	0.03	0.17	0.00	0.05	-0.01	0.18	-0.02
Ethnicity: White	White	0.28		0.17		0.09		0.18	
	Asian	0.46	0.18	0.29	0.12	0.13	0.04	0.27	0.09
	Black	0.36	0.08	0.26	0.09	0.11	0.02	0.29	0.11
	Hispanic	0.29	0.01	0.28	0.11	0.16	0.07	0.29	0.11
SES quart: Top	Top	0.42		0.16		0.09		0.13	
	Second	0.45	0.03	0.17	0.01	0.06	-0.03	0.17	0.04
	Third	0.44	0.02	0.19	0.03	0.03	-0.06	0.16	0.03
	Bottom	0.50	0.08	0.19	0.03	0.06	-0.03	0.16	0.03
Asian SES: Top	NAEP top	0.46		0.34		0.18		0.34	
	Bottom	0.54	0.08	0.31	-0.03	0.20	0.02	0.31	-0.03
	All top	0.24		0.21		0.09		0.14	
	Bottom	0.42	0.19	0.28	0.08	0.19	0.09	0.28	0.15
Black SES: Top	NAEP top	0.65		0.42		0.27		0.42	
	Bottom	0.79	0.14	0.50	0.08	0.05	-0.22	0.50	0.07
	All top	0.48		0.27		0.19		0.29	
	Bottom	0.59	0.11	0.24	-0.03	0.11	-0.09	0.24	-0.05
Hispan. SES: Top	NAEP top	0.64		0.37		0.22		0.37	
	Bottom	0.86	0.22	0.47	0.10	0.16	-0.06	0.47	0.10
	All top	0.42		0.26		0.18		0.21	
	Bottom	0.51	0.10	0.29	0.03	0.19	0.01	0.23	0.02
White SES: Top	NAEP top	0.64		0.40		0.17		0.40	
	Bottom	0.70	0.06	0.39	-0.01	0.06	-0.11	0.39	0.00
	All top	0.45		0.15		0.11		0.17	
	Bottom	0.50	0.04	0.20	0.05	0.07	-0.04	0.20	0.03

Notes & Source: See Table 1. Table displays medians of change/decade in standard deviations in achievement levels in math by subgroups and survey displayed in Tables 1, 4-8 and A.12 – A.18. E. A. = Early Adolescence. Bir yr = Birth year, Sin. = Since, All = LTT, NAEP, TIMSS and PISA. Birth years differ across subgroups, depending on the availability of data (see tables and appendix for details). Differences between the base category and other categories are also displayed.

Table 3: Medians of change/decade in achievement levels in reading by subgroups and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Base	Subgroup	Young	Diff.	E. A.	Diff.	Older	Diff.	All age	Diff.
	All	0.08		0.03		0.02		0.04	
	LTT	0.05		0.03		0.00		0.03	
	NAEP	0.08		0.12		0.03		0.08	
	PIRLS	0.09						0.09	
	PISA			-0.02				-0.02	
Bir yr: Until 1990	Until 1990	-0.02		0.01		0.03		0.01	
	Sin. 1990	0.11	0.13	0.10	0.09	0.01	-0.02	0.09	0.08
Gender: Female	Female	0.06		0.03		0.02		0.04	
	Male	0.10	0.04	0.04	0.01	0.02	0.00	0.06	0.02
Ethnicity: White	White	0.09		0.05		0.03		0.06	
	Asian	0.28	0.19	0.13	0.08	0.21	0.18	0.21	0.15
	Black	0.19	0.10	0.15	0.10	0.11	0.08	0.15	0.09
	Hispanic	0.13	0.04	0.21	0.16	0.06	0.03	0.12	0.06
SES quart: Top	Top	-0.06		0.04		0.06		0.00	
	Second	-0.05	0.01	0.02	-0.02	0.01	-0.05	0.00	0.00
	Third	-0.04	0.02	0.04	0.00	0.02	-0.04	0.00	0.00
	Bottom	0.09	0.15	0.15	0.11	0.02	-0.04	0.03	0.03
Asian SES: Top	NAEP top	0.02		-0.02		0.07		0.02	
	Bottom	0.26	0.24	0.10	0.12	0.19	0.12	0.16	0.14
	All top	0.03		-0.02		0.05		0.03	
	Bottom	0.09	0.06	0.10	0.11	0.17	0.12	0.16	0.13
Black SES: Top	NAEP top	0.09		0.13		0.07		0.10	
	Bottom	0.24	0.15	0.19	0.06	0.05	-0.02	0.19	0.09
	All top	0.01		0.07		0.08		0.07	
	Bottom	0.13	0.12	0.13	0.06	0.09	0.01	0.12	0.05
Hispan. SES: Top	NAEP top	-0.07		0.19		0.04		0.05	
	Bottom	0.21	0.28	0.28	0.09	0.22	0.18	0.21	0.16
	All top	0.01		0.11		0.05		0.05	
	Bottom	0.18	0.17	0.28	0.17	0.16	0.11	0.21	0.16
White SES: Top	NAEP top	-0.04		0.17		0.12		0.03	
	Bottom	0.16	0.20	0.21	0.04	-0.02	-0.14	0.16	0.13
	All top	-0.05		0.05		0.08		0.03	
	Bottom	0.10	0.15	0.05	0.00	-0.02	-0.10	0.05	0.01

Notes & Source: See Tables 1 and 2.

Table 4: Comparing the change in achievement levels in surveys until and since birth year 1990.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Math				Reading			
		Until birthyr 1990		Since birthyr 1990		Until birthyr 1990		Since birthyr 1990	
Survey	Grade/Age	Change	Change/dc	Change	Change/dc	Change	Change/dc	Change	Change/dc
LTT	Age 9	0.466**	0.22	0.125**	0.16	0.026*	0.01	0.120**	0.15
		(0.013)		(0.015)		(0.012)		(0.014)	
NAEP	Grade 4	0.645**	0.65	0.355**	0.21	-0.037**	-0.04	0.116**	0.07
		(0.015)		(0.003)		(0.013)		(0.002)	
LTT	Age 13	0.343**	0.13	0.164**	0.21	0.043**	0.01	0.178**	0.22
		(0.013)		(0.014)		(0.013)		(0.016)	
NAEP	Grade 8	0.571**	0.44	0.100**	0.08	0.292**	0.22	0.122**	0.10
		(0.006)		(0.003)		(0.004)		(0.003)	
TIMSS	Grade 8	0.276**	0.35	0.113**	0.14				
		(0.014)		(0.013)					
PISA	Age 15	-0.108**	-0.36	-0.057**	-0.06	-0.092**	-0.31	-0.024	-0.04
		(0.023)		(0.017)		(0.021)		(0.018)	
LTT	Age 17	0.254**	0.10	0.008	0.02	0.016	0.00	0.034*	0.09
		(0.015)		(0.013)		(0.012)		(0.014)	
NAEP	Grade 12					0.073**	0.05	-0.039**	-0.07
						(0.010)		(0.008)	

Notes & Source: See Table 1. The table compares the changes in achievement levels in LTT, NAEP and TIMSS until and since birth year (birthyr) 1990 (or the closest available; see Table 1 for details).

Table 5: Change in achievement levels at younger age by ethnicity and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Math			Reading		
Survey	Ethnicity	N	Change	Change/dc	N	Change	Change/dc
LTT	Asian	2,520	0.928**	0.27	3,490	1.131**	0.31
			(0.054)			(0.044)	
	Black	14,070	0.974**	0.29	22,020	0.622**	0.15
			(0.028)			(0.025)	
	Hispanic	14,190	0.977**	0.29	14,080	0.442**	0.11
			(0.023)			(0.024)	
	White	56,420	0.853**	0.25	96,260	0.267**	0.07
			(0.015)			(0.014)	
NAEP	Asian	69,280	1.230**	0.46	76,960	0.470**	0.17
			(0.012)			(0.011)	
	Black	250,190	1.305**	0.48	278,640	0.509**	0.19
			(0.006)			(0.005)	
	Hispanic	262,040	1.286**	0.48	271,750	0.352**	0.13
			(0.006)			(0.005)	
	White	810,110	1.077**	0.40	907,840	0.255**	0.09
			(0.003)			(0.003)	
TIMSS	Asian	2,330	1.274**	0.64			
			(0.050)				
	Black	7,700	0.724**	0.36			
			(0.027)				
	Hispanic	11,530	0.589**	0.29			
			(0.023)				
	White	24,820	0.559**	0.28			
			(0.015)				
PIRLS	Asian				840	0.140*	0.28
						(0.066)	
	Black				2,410	0.233**	0.47
						(0.035)	
	Hispanic				4,500	0.160**	0.32
						(0.026)	
	White				8,980	0.168**	0.34
						(0.018)	

Notes & Source: See Table 1. See Table A.1 for details of ethnic coding.

Table 6: Change in achievement levels at early adolescence by ethnicity and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Math			Reading		
Survey	Ethnicity	N	Change	Change/dc	N	Change	Change/dc
LTT	Asian	3,150	0.806**	0.24	3,530	1.152**	0.31
			(0.046)			(0.053)	
	Black	15,880	0.739**	0.22	21,730	0.606**	0.15
			(0.026)			(0.028)	
	Hispanic	12,820	0.735**	0.22	12,300	0.147**	0.04
			(0.022)			(0.031)	
	White	68,070	0.553**	0.16	102,760	0.192**	0.05
			(0.014)			(0.014)	
NAEP	Asian	64,080	0.972**	0.36	69,690	0.349**	0.13
			(0.014)			(0.012)	
	Black	223,150	0.947**	0.35	242,490	0.411**	0.15
			(0.006)			(0.006)	
	Hispanic	213,560	0.959**	0.36	222,510	0.562**	0.21
			(0.007)			(0.006)	
	White	750,100	0.728**	0.27	833,220	0.360**	0.13
			(0.003)			(0.003)	
TIMSS	Asian	2,660	0.668**	0.33			
			(0.052)				
	Black	8,000	0.605**	0.30			
			(0.028)				
	Hispanic	11,610	0.660**	0.33			
			(0.023)				
	White	30,630	0.369**	0.18			
			(0.014)				
PISA	Asian	1,180	-0.090	-0.06	1,030	-0.084	-0.06
			(0.086)			(0.082)	
	Black	4,040	0.027**	0.02	3,580	0.057	0.04
			(0.037)			(0.038)	
	Hispanic	6,480	0.153**	0.10	5,820	0.332**	0.22
			(0.032)			(0.032)	
	White	15,340	-0.213**	-0.14	12,920	-0.101**	-0.07
			(0.021)			(0.021)	

Notes & Source: See Table 1.

Table 7: Change in achievement levels at older age by ethnicity and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Math			Reading		
Survey	Ethnicity	N	Change	Change/dc	N	Change	Change/dc
LTT	Asian	2,780	0.161**	0.05	3,410	0.840**	0.23
			(0.050)			(0.045)	
	Black	13,360	0.520**	0.15	19,200	0.689**	0.17
			(0.029)			(0.026)	
	Hispanic	9,260	0.544**	0.16	10,400	0.137**	0.03
			(0.026)			(0.029)	
	White	66,110	0.289**	0.09	103,050	0.044**	0.01
			(0.016)			(0.013)	
NAEP	Asian	3,530	0.212**	0.21	6,810	0.459**	0.18
			(0.044)			(0.039)	
	Black	10,120	0.072**	0.07	21,560	0.088**	0.04
			(0.022)			(0.020)	
	Hispanic	11,160	0.162**	0.16	21,140	0.213**	0.09
			(0.022)			(0.020)	
	White	45,470	0.089**	0.09	92,850	0.097**	0.04
			(0.011)			(0.010)	

Notes & Source: See Table 1.

Table 8: Change in achievement levels at by SES quartiles, age categories, and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Math			Reading		
Quartile	Survey	N	Change	Change/dc	N	Change	Change/dc
Younger							
Top	LTT	46,410	0.506**	0.24	30,860	-0.131**	-0.07
			(0.030)			(0.034)	
Second			0.544**	0.26		-0.068	-0.04
			(0.030)			(0.035)	
Third			0.501**	0.24		-0.179**	-0.09
			(0.031)			(0.036)	
Bottom			0.582**	0.28		0.046	0.02
			(0.032)			(0.038)	
Top	NAEP	147,430	0.761**	0.59	161,090	-0.049**	-0.04
			(0.014)			(0.012)	
Second			0.822**	0.63		-0.060**	-0.05
			(0.014)			(0.013)	
Third			0.825**	0.63		0.027*	0.02
			(0.014)			(0.013)	
Bottom			0.930**	0.72		0.211**	0.16
			(0.014)			(0.013)	
Early Adolescence							
Top	LTT	90,010	0.454**	0.13	64,070	0.125**	0.04
			(0.022)			(0.025)	
Second			0.568**	0.17		0.078**	0.02
			(0.022)			(0.025)	
Third			0.531**	0.16		0.113**	0.04
			(0.022)			(0.025)	
Bottom			0.543**	0.16		0.098**	0.03
			(0.023)			(0.026)	
Top	NAEP	587,020	0.709**	0.37	703,770	0.283**	0.15
			(0.007)			(0.006)	
Second			0.655**	0.34		0.196**	0.10
			(0.007)			(0.006)	
Third			0.693**	0.36		0.330**	0.17
			(0.007)			(0.006)	
Bottom			0.711**	0.37		0.312**	0.16
			(0.007)			(0.006)	

Table 8 (Cont'd): Change in achievement levels by SES quartiles, age categories, and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Math			Reading		
Quartile	Survey	N	Change	Change/dc	N	Change	Change/dc
Top	TIMSS	46,470	0.355**	0.18			
			(0.023)				
Second			0.348**	0.17			
			(0.023)				
Third			0.440**	0.22			
			(0.023)				
Bottom			0.439**	0.22			
			(0.022)				
Top	PISA	28,370	-0.296**	-0.20	24,400	-0.203**	-0.14
			(0.031)			(0.029)	
Second			-0.200**	-0.13		-0.135**	-0.09
			(0.031)			(0.030)	
Third			-0.258**	-0.17		-0.098**	-0.07
			(0.029)			(0.029)	
Bottom			0.067*	0.04		0.218**	0.15
			(0.030)			(0.030)	
Older							
Top	LTT	86,640	0.132**	0.04	65,820	0.046*	0.01
			(0.024)			(0.021)	
Second			0.251**	0.07		-0.033	-0.01
			(0.024)			(0.022)	
Third			0.249**	0.07		-0.064**	-0.02
			(0.024)			(0.022)	
Bottom			0.293**	0.09		0.007	0.00
			(0.024)			(0.023)	
Top	NAEP	68,320	0.133**	0.13	121,800	0.199**	0.10
			(0.017)			(0.015)	
Second			0.050**	0.05		0.043**	0.02
			(0.017)			(0.015)	
Third			-0.013	-0.01		0.100**	0.05
			(0.017)			(0.016)	
Bottom			0.020	0.02		0.068**	0.04
			(0.017)			(0.016)	

Notes & Source: See Table 1. SES has been constructed from constant principal component analysis of parents' education and possession index (see Tables A.4 and A.5). Birth cohorts are: younger age (LTT math 1968-1989 and reading 1970-1989 and NAEP 1980-1993), early adolescence (LTT math 1964-1998 and reading 1966-1998, NAEP 1976-1995, TIMSS 1981-2001 and PISA 1985-2000) and older age (LTT math 1961-1995 and reading 1963-1995 and NAEP math 1988-1998 and reading 1973-1992).

Online Appendix

Consistency of SES effects across different constructs

Parental education

In 12 of the 15 surveys, cohorts of students who have parents with the least amount of education (no high school diploma) outpace those who have the most (college degree or more) educated parents. The differences vary somewhat by the age a student is tested. In math, the median difference is only 0.01sd for those tested at a younger age, 0.10sd for those tested in early adolescence, and 0.05sd for those tested just before completing secondary education. In reading, these numbers are 0.01sd, 0.03sd and -0.05sd. Interestingly, those in the top and bottom categories both make greater progress than do those in the two middle categories (high school and some college) in twelve of the fifteen comparisons (Tables A.11-A.13). This U-shape distribution is evident in both math and reading.

Possession Index

The performance of students in the lowest quartile of the permanent income distribution, as indicated by the number of home possessions, outpaces that of those in the top quartile in nine of seventeen comparisons (Tables A.14-A.16). The top quartile outperforms the bottom quartile in seven comparisons and there is no difference in one survey. The median differential between the top and bottom quartiles across all surveys in both math and in reading is 0.05sd. Nine of the 17 surveys show a U-shaped distribution, with the top and bottom quartiles showing greater progress than the middle two quartiles.

Income (eligibility for free or reduced lunch)

Two surveys—LTT and NAEP—obtain from administrative records a blunt measure of annual income, the eligibility of students for participation in the free or reduced lunch program, which provides subsidized meals to students with incomes below 185 percent of the poverty line. Eligibility in recent years has been gradually extended to all students within some districts if most students are deemed to be eligible. In part for this reason is that the percentage of children eligible for program participation increased between 37 percent in the school year ending in 1999 and 52 percent in the school year ending in 2015 (Chingos, 2016; Greenberg, 2018). Given the broadening of the eligibility definition, the variable is not a constant indicator of annual family income. But as it is an alternative to the possession index, a crude measure of permanent income, we report the rate of progress for those deemed eligible (and ineligible) as reported in twelve LTT and NAEP surveys.

No consistent differentials in the rate of progress are evident for students from higher and lower income households in the twelve surveys for which information is available as to student eligibility for free or reduced lunch. (Table A.17). In six comparisons the ineligible group is making more progress; in four the greater progress is being made by those who are eligible, and in two cases differences are statistically insignificant. Even when significant differences are detected, they are small. Either this indicator of income is too blunt to capture the role of annual income in rates of cohort progress, or the rate of progress is not conditional on annual household income, or the changes in eligibility requirements cloud this measure.

Family structure

In eleven surveys, students report information on the structure of the household in which they live. The surveys categorize household structure in various ways (Table A.6); to facilitate comparisons within and across surveys we simply dichotomize responses into two broad categories: 1) household where the students say they live with their two parents (not necessarily biological), and 2) all other households. Although family structure is not a precise indicator of social class, two-parent households tend to have higher levels of income (Berger & McLanahan, 2015). The meaning of this indicator may shift with changes in divorce and separation rates.

We find some, but not consistent, signs that the rate of progress is greater for those from households without two parents in the eleven surveys for which this information is available (Table A.18). In seven surveys, the rate of progress for these groups is greater, but in three the opposite is true, and in one survey no significant difference is detected. It could be that the advantage of living in households with two parents has declined over time, or that household formation patterns have changed, or it could be that family structure is not critical for achievement.

References

Berger, L. M., & McLanahan, S. S. (2015). Income, relationship quality, and parenting: Associations with child development in two-parent families. *Journal of Marriage and Family*, *77*(4), 996-1015.

Chingos, M. M. (2016). *No more free lunch for education policymakers and researchers*. Washington, DC: Brookings Institution. Retrieved from: http://www.brookings.edu/research/reports/2016/06/30-no-more-free-lunch-for-education-policy-makers-and-researchers-chingos

Greenberg, E. (2018). *New measures of student poverty: Replacing free and reduced-price lunch status based on household forms with direct certification* [Education policy program policy brief]. Urban Institute.

Table A.1: Selected studies on sociology and economics of education and intelligence.

Study	Datasets	Years	Primary focus	Brief findings
Sociology and economics of education				
Miller (1995)	LTT, other	1971-1990	Ethnic gaps	Some closing of Asian-white and Hispanic-white gaps
Jencks & Phillips (1998)	LTT, other	1971-1996	Black-white gap	Black-white gap closes substantially
Campbell et al. (1996)	LTT, NAEP	1971-1994	Achievement levels	Improvement in math. Not much change in reading
Hedges & Nowell (1995)	NAEP, other	1960-1992	Gender gap	Male advantage in math but not in reading. Gender differences are stable
Hedges & Nowell (1998)	Other	1965-1992	Income, SES, and ethnic gaps	Black-white gap closes. Unclear trend in SES gaps
Grissmer et al. (1994)	Other	1970-1990	Family structure and achievement	Black-white and Hispanic-white gaps close
Phillips, Crouse, & Ralph (1998)	LTT, other	1965-1996	Black-white gap	Gap increases with grades in math, but not in reading
Bloom et al. (2008)	Other	1991-2002	Effect size benchmarks	Growth declines as students move from early grades to later grades
Magnuson, Rosenbaum, & Waldfogel (2008)	LTT	1975-2004	Black-white gap	Black-white gap closes substantially, but little further closing after 1990
Reardon (2011)	Other	1960-2007	90-10 income gap	Large increases in gaps
Hanushek, Peterson, & Woessmann (2012)	NAEP, TIMSS, PIRLS, PISA	1995-2009	Achievement levels	Growth rate is continuous, but modest
Reardon, Valentino, & Shores (2012)	LTT, PIRLS, PISA, other	1971-2009	Literacy patterns	Steeper upward trends in math than reading on the LTT
OECD (2018)	PISA	2000-2015	SES gap	Gaps close modestly in U. S.
Broer, Bai, & Fonseca (2019)	TIMSS	1995-2015	72-25 SES gap	No significant change for gaps in math in U. S.
Chmielewski (2019)	TIMSS, PIRLS, PISA, other	1964-2015	90-10 SES gap	No significant trend in gaps in U. S.

Table A.1 (Cont'd): Selected studies on sociology and economics of education and intelligence.

Study	Datasets	Years	Primary focus	Brief findings
Hanushek et al. (2020)	LTT, NAEP, TIMSS, PISA	1971-2015	75-25 SES gap	Not much change in gaps
Hashim et al. (2020)	NAEP	1990-2015	Income gap	Achievement levels increase. Income gaps close modestly
Shakeel & Peterson (2020)	NAEP	2005-2017	Achievement levels	Larger growth in achievement in charters than at district schools
Matheny et al. (2021)	Other	2009-2018	School district gaps	Gaps across school districts grow based on student SES and ethnic composition
Intelligence studies				
Flynn (1984)	Other	1932-1978	IQ gains	Gains of 3 IQ points per decade
Rodgers & Wänström (2007)	Other	1986-2000	IQ gains	Flynn effect for PIAT-math, but not reading
Ang, Rodgers & Wänström (2010)	Other	1986-2000	IQ gains	Flynn effects for subgroups
Wai & Putallaz (2011)	Other	1981-2000	IQ gains	Flynn effect among high ability students
Pietschnig & Voracek (2015)	Other	1909-2013	IQ gains	Confirm Flynn Effect. Larger increases in fluid reasoning that crystallized knowledge
Rindermann & Thompson (2013)	LTT	1971-2008	IQ gains and ethnic gaps	Larger gains for younger students, and for ethnic minorities
Meisenberg & Woodley (2013)	TIMSS, PISA	1995-2011	IQ gains and cross-country cognitive differences	Higher gains in countries with lower test scores at the beginning of the time period

Note: "Other" datasets include a variety of tests usually administered once or a few times. For details on subjects, grades, exact years for each dataset and samples size, please refer to the original studies.

Table A.2: Test administrations included in analysis, by survey, age, and year of administration.

	LTT (age)		NAEP (grade)		TIMSS (grade)	PIRLS (grade)	PISA (age)
Year	Math	Read.	Math	Read.	Math	Read.	Math/Read.
1971		9,13,17					
1975		9,13,17					
1978	9,13,17						
1980		9,13,17					
1982	9,13,17						
1984		9,13,17					
1986	9,13,17						
1988		9,13,17					
1990	9,13,17	9,13,17	4,8	4,8,12			
1992	9,13,17	9,13,17	4,8	4,8,12			
1994	9,13,17	9,13,17		4,8,12			
1995					4,8		
1996	9,13,17	9,13,17	4,8				
1998				4,8,12			
1999	9,13,17	9,13,17			8		
2000			4,8	4			15
2001						4	
2002				4,8,12			
2003			4,8	4,8	4,8		15
2004	9,13,17	9,13,17					
2005			4,8,12	4,8,12			
2006						4	15*
2007			4,8	4,8	4,8		
2008	9,13,17	9,13,17					
2009			4,8,12	4,8,12			15
2011			4,8	4,8	4,8	4	
2012	9,13,17	9,13,17					15
2013			4,8	4,8			
2015			4,8,12	4,8,12	4,8		15
2016						4	
2017			4,8	4,8			

Note: Table displays math and reading tests available from U.S. Department of Education for each survey. The cells list age/grade of assessment (9,14,15,17;4,8,12). Grade 12 math in NAEP excludes 1990, 1992, 1996 and 2000 as the test format changed in 2005. Student-level LTT math data for 1971 and 1973 could not be obtained. TIMSS was only administered at grade 8 in 1999. We used rescaled scores for the 1995 administration of TIMSS available from https://timssandpirls.bc.edu/timss1995i/Database.html and https://timssandpirls.bc.edu/timss1995i/newscale/layout.txt * PISA test in reading was excluded in 2006 by the administering body. PIRLS 2001 and 2016 waves did not have restricted use versions (information on student ethnicity is missing).
Source: See Table 1.

Table A.3: Ethnic representation in surveys by testing year in math.

Survey	Year	White	Black	Hispanic	Asian	Others
LTT	1978	81	13	5	1	0
	1982	80	13	5	2	0
	1986	73	14	9	1	2
	1990	71	15	10	3	2
	1992	71	15	10	2	1
	1994	70	15	11	3	2
	1996	69	14	12	3	1
	1999	68	14	13	3	2
	2004	63	15	16	4	2
	2008	57	15	21	5	2
	2012	55	14	23	6	2
NAEP	1990	70	15	11	2	2
	1992	69	16	10	3	2
	1996	69	14	12	3	2
	2000	68	14	13	4	2
	2003	61	16	16	4	2
	2005	60	16	17	5	2
	2007	58	16	19	5	2
	2009	57	15	20	5	3
	2011	54	15	22	5	3
	2013	53	15	23	5	3
	2015	51	15	25	6	4
	2017	50	15	25	6	4
TIMSS	1995	66	14	10	4	6
	1999	63	15	12	5	5
	2003	62	15	17	3	3
	2007	55	13	23	4	6
	2011	52	12	25	5	7
	2015	48	12	27	5	8
PISA	2000	59	14	18	3	5
	2003	59	16	17	3	5
	2006	59	13	18	4	6
	2009	56	12	23	5	4
	2012	51	13	25	5	6
	2015	45	13	30	4	6

Note. Table shows weighted ethnic representation in the math test by the year tested. Grades 4 and 8 are combined for TIMSS. Grades 4, 8 and 12 are combined for Main NAEP. Ages 9, 13 and 17 are combined for LTT NAEP. White and Black categories are additionally defined as "non-Hispanic" in TIMSS and PISA. Asians also comprise Pacific Islanders, and we coded it consistently across the surveys. For details see Table A.4.
Source: See Table 1.

Table A.4: Ethnic coding in the test administrations included in analysis, by survey, age, and year of administration.

Survey	Year	Grade/Age	Coding
LTT* and NAEP	1971-2017	9,13,17 4,8,12	1 White 2 Black 3 Hispanic 4 Asian American/Pacific Islander 5 American Indian/Alaskan Native
TIMSS	1995	4	1 White (not Hispanic) 2 Black (not Hispanic) 3 Hispanic 4 Asian/Pacific Islander 5 American Indian/Alaskan Native 6 Other
		8	1 White 2 Black 3 Hispanic 4 Asian/Pacific Islander 5 Native American 6 Other
	1999	8	1 White 2 Black 3 Hispanic 4 Asian/Pacific Islander 5 Native American 6 Other 8 Not admin. 9 missing
	2003	4,8	0 Missing 1 White 2 Black 3 Hispanic 4 Asian 5 American Indian/Native Hawaiian/Pacific Islander
	2007, 2011, 2015	4,8	1 White, not Hispanic 2 Black, not Hispanic 3 Hispanic 4 Asian 5 Native American 6 Pacific Islander 7 2 or more races 9 omitted
PIRLS	2006	4	1 Hispanic 2 White, not Hispanic 3 Black, not Hispanic 4 Asian, not Hispanic 5 American Indian/Alaskan, not Hispanic 6 Multiple races, not Hispanic 9 omitted
	2011		1 White, not Hispanic 2 Black, not Hispanic 3 Hispanic 4 Asian 5 Native American 6 Pacific islander 7 2 or more races 9 omitted or invalid
PISA	2000	15	1 Hispanic 2 White 3 Black 4 Asian 5 American Indian 6 Hawaiian 7 Multirace 9 Missing
	2003		0 Missing 1 White 2 Black 3 Hispanic 4 Asian 5 Other 6 Multirace
	2006		1 Hispanic 2 White, not Hispanic 3 Black, not Hispanic 4 Asian, not Hispanic 5 Multiracial, not Hispanic 6 American Indian, not Hispanic 7 Pacific Islander, not Hispanic 9 Missing
	2009		1 White 2 Black 3 Hispanic 4 Asian 5 American Indian/Alaska Native 6 Native Hawaiian/Other Pacific Islander 7 More than one race 8 N/A 9 Miss
	2012		1 White 2 Black or African American 3 Hispanic 4 Asian 5 American Indian or Alaska Native 6 Native Hawaiian or Other Pacific Islander 7 More than one race 97 N/A 99 Miss

Table A.4 (Cont'd): Ethnic coding in the test administrations included in analysis, by survey, age, and year of administration.

Survey	Year	Grade/Age	Coding
PISA	2015	15	1 White, not Hispanic 2 Black or African American 3 Hispanic or Latino 4 Asian 5 American Indian or Alaska Native 6 Native Hawaiian or Other Pacific 7 Multi-Racial 99 No Response

Note. Table displays codes and labels for ethnicity across the surveys. Some datasets in NAEP do no mention "Pacific Islander" and "Alaskan Native" in the variable label. * For 1971 LTT reading does not have Asian and Other categories. Restricted-use data on ethnic identification for PIRLS 2001 and 2016 was not available from NCES.
Source: See Table 1.

Table A.5: Parent's education coding in the test administrations included in analysis, by survey, age, and year of administration.

Survey	Years	Notes
LTT	1971-2012	At age 9, parent's education is only reported for 1971-1999. Parent's education is coded as 1 Did not finish High School 2 Graduated High School 3 Post High School 4 Graduated college. In 1971 and 1975, the parent's education variable PARED4 was constructed from a variable on parent's highest education PEDH (values 0,1,2,3,5,9). The tabulations of the two variables mismatch for 1971 and 1975 (graduated college category in PARED4 comprises of both 0 and 9 values in PEDH). Documentation is lacking to resolve these discrepancies. Thus, PARED4 is poorly constructed. Hence, we excluded data on parent's education for 1971 and 1975.
NAEP	1990-2017	For grade 4, parent's education is only reported for 1990-1998, 2000 Reading and 2003. Parent's education is coded as 1 Did not finish High School 2 Graduated High School 3 Post High School 4 Graduated college.
TIMSS	1995-2015	Parent's education is not reported for grade 4. Parent's education is coded as 1 No more than primary schooling 2 Finished lower secondary schooling 3 Finished upper secondary schooling 4 Finish post-secondary, vocational/technical education but no university 5 Finished university/equivalent or higher. Parent's education can be recoded to compare to Main NAEP and LTT NAEP. For parent's education the coding would be (2=1) (3=2) (4=3) (5=4).
PISA	2000-2015	Parent's education is coded as 0 None 1 ISCED 1 i.e. elementary school 2 ISCED 2 i.e. middle/junior high school 3 ISCED 3B, C i.e. high school equivalent/GED 4 ISCED 3A, ISCED 4 i.e. high school diploma/ vocational or other certificate/ diploma of high school 5 ISCED 5B i.e. an associate degree 6 ISCED 5A, 6 i.e. bachelor, master, doctorate, professional degree, law, or medicine. Parent's education can be recoded to compare to Main NAEP and LTT NAEP. For parent's education the coding would be (0 2 =1) (3 4 =2) (5=3) (6=4).

Source: See Table 1.

Table A.6: Possession index coding in the test administrations included in analysis, by survey, age, and year of administration.

Survey	Years	Items	Total items	Notes
LTT	1971-2012	Books	4	Books is binary for 1971-1999. The binary measure has coding 0 “<=25 books” 1 “>25 books.” Books has four categories for 2004-2012. The four categories are 1 "0-10 i.e. few” 2 "11-25 i.e. one shelf” 3 "26-100 i.e. one bookcase 4 ">100 i.e. > one bookcase.” To maintain a consistent measurement, we recoded books as binary across all the years within LTT.
	1971-2012	Newspaper, magazines, encyclopedia		Encyclopedia is missing in 1984.
NAEP	1990-2017	Books	4	Books is binary for 1990-2000. The binary measure has coding 0 “<=25 books” 1 “>25 books.” Books has four categories after the year 2000. The four categories are 1 "0-10 i.e. few” 2 "11-25 i.e. one shelf” 3 "26-100 i.e. one bookcase 4 ">100 i.e. > one bookcase.” To maintain a consistent measurement, we recoded books as binary across all the years within NAEP.
	1990-2010	Newspaper, magazines, encyclopedia		Magazine and encyclopedia exist in 2011
	2005-2015	Computer	2	This variable was only used for grade 12 Math.
TIMSS	1995-2015	Books	7	Books has five categories. The five categories are 1 "0-10" 2 "11-25" 3 "26-100" 4 "101-200" 5 ">200."
	1995-2015	Study desk, computer		Using own computer as proxy for computer in 2015
PIRLS	2001-2016	Books	8	Books has five categories. The five categories are 1 "0-10" 2 "11-25" 3 "26-100" 4 "101-200" 5 ">200."
	2001-2016	Study desk, computer, own books		

Table A.6 (Cont'd): Possession index coding in the test administrations included in analysis, by survey, age, and year of administration.

Survey	Years	Items	Total items	Notes
PISA	2000-2015	Books	25	Books has six categories. The six categories are 1 "0-10" 2 "11-25" 3 "26-100" 4 "101-200" 5 "251-500" 6"More than 500." PISA 2000 has seven categories (1 "none" 2 "0-10" 3"11-50" 4 "51-100" 5 "101-200" 6 "251-500" 7"More than 500"). We assumed that the weighted proportional distribution of #books remains same for 2000 and 2003, and recoded books as (2=1) (3=2) (4=3) (5=4) (6=5) (7=6).
	2000-2015	Study desk, own room, study place, computer, educational software, internet, classic literature, poetry books, art works (e.g. paintings), textbooks, dictionary, TV, car		

Note: All items in LTT and NAEP are dichotomous. In TIMSS and PIRLS, except for books at home, all other items are dichotomous. In PISA, books are categorical, and the following items are polytomous and coded 0-3: computer, TV, and car. All remaining items are dichotomous.
Source: See Table 1.

Table A.7: Family structure coding in the test administrations included in analysis, by survey, age, and year of administration.

Survey	Year	Grade/Age	Coding
LTT	1986, 1990, 1992, 1994, 1996, 1999	9,13, 17	1 2 parents at home 2 1 parent at home 3 Neither 8 Omitted 9 Not reached
NAEP	1990,1992, 1994	4,8,12	1 2 Parents at home 2 1 Parent at home 3 Neither Parent Home 8 Omitted
	1996*		Does (mother/father) or (stepmother/stepfather) live at home with you? 0 Multiple Response 1 Yes 2 No 8 Omitted
	2013*	4, 8	Lives in home: (mother/ stepmother/ father/ stepfather)? 1 Yes 8 Omitted
	2015*	4, 8, 12	
	2017*	4, 8	Lives in home: (mother/ stepmother/ father/ stepfather)? 1 Yes 2 No 8 Omitted
TIMSS	1995*	4,8	Student lives with (mother/ stepmother/ father/ stepfather)? 1 yes 2 no 8 not admin. 9 missing
	1999*	8	
PISA	2000, 2003	15	1 Single parent family 2 Nuclear family 3 Mixed family 4 Other 9 Missing
	2009		1 Single parent (natural or otherwise) 2 Two parents (natural or otherwise) 3 Other 7 N/A 9 Missing
	2012		1 Single parent (natural or otherwise) 2 Two parents (natural or otherwise) 3 Other 9 Missing

Note: Table displays the available data for family structure. Data on family structure was only collected for Math in LTT. For the years marked with an asterisk (*), family structure variable was coded from questions about who lives in the home.
Source: See Table 1.

Table A.8: Change in achievement levels per decade at younger age and early adolescence in cross-country surveys.

(1)	(2)	(3)	(4)	(5)	(6)
Survey	Sample	Math	Reading	Math	Reading
		Younger age (grade 4)		Early adolescence (grade 8/age 15)	
TIMSS	Overall	0.24		0.11	
	OECD	0.20		0.09	
	Non-OECD	0.32		0.16	
N		340,130		386,650	
PIRLS	Overall		0.14		
	OECD		0.03		
	Non-OECD		0.28		
N			307,980		
PISA	Overall			-0.01	-0.01
	OECD			-0.04	-0.01
	Non-OECD			0.04	-0.01
N				1,728,980	1,767,990

Note: Table shows average of changes in achievement levels per decade for 10 regions in TIMSS, 32 regions in PISA and 16 regions in PIRLS. Normalized achievement is measured in standard deviations (s.d.). The s.d. is the difference between the year test was administered and the starting year for a specific set of countries that were consistently tested for all administrations of a test between the available years. The changes have been estimated from a quadratic fit. See Tables A.9-A.11 for cross-country comparisons.

For TIMSS, 10 countries were tested both at grades 4 and 8 between 1995-2019 and these countries also has test data for the initial and final years (https://nces.ed.gov/timss/participation.asp). For PISA, 32 regions were tested in all years between 2001-2018 (https://nces.ed.gov/surveys/pisa/countries.asp). For PIRLS, 16 regions were tested for all years between 2001-2016 (https://nces.ed.gov/surveys/pirls/countries.asp). We used rescaled scores for the 1995 administration of TIMSS available from https://timssandpirls.bc.edu/timss1995i/Database.html and https://timssandpirls.bc.edu/timss1995i/newscale/layout.txt PISA test in reading was excluded in 2006 in U. S. by the administering body.

OECD represents regions that are affiliated with the Organisation for Economic Co-operation and Development.

SOURCE: The International Association for the Evaluation of Educational Achievement (IEA), TIMSS 1995-2019, PISA 2000-2018 and PIRLS 2001-2016.

Table A.9: Change in achievement levels at younger age and early adolescence in cross-country TIMSS survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Grade 4			Grade 8		
Country	Change	s.e.	Change/dc	Change	s.e.	Change/dc
Australia	0.296**	(0.012)	0.12	0.029**	(0.011)	0.01
Hong Kong-China*	0.709**	(0.012)	0.30	0.189**	(0.014)	0.08
Hungary	0.321**	(0.014)	0.13	-0.066**	(0.014)	-0.03
Iran*	0.788**	(0.014)	0.33	0.398**	(0.013)	0.17
Japan	0.517**	(0.011)	0.22	0.213**	(0.013)	0.09
New Zealand	0.408**	(0.014)	0.17	-0.010	(0.013)	0.00
Norway	1.164**	(0.014)	0.49	0.350**	(0.012)	0.15
Singapore*	0.616**	(0.011)	0.26	0.232**	(0.012)	0.10
United Kingdom	0.868**	(0.015)	0.36	0.322**	(0.015)	0.13
United States	0.476**	(0.009)	0.20	0.323**	(0.010)	0.13

Notes & Source: See Table A.8. Table displays cross-country changes in achievement levels in TIMSS between 1995-2019. The changes have been estimated from a quadratic fit. Change per decade (dc) in achievement levels is also displayed. Standard errors are in parenthesis. *p <0.05, **p<0.01.* represents Non-OECD region. Hong Kong is not an independent country and therefore is not discussed in text.

Table A.10: Change in achievement levels at younger age in cross-country PIRLS survey.

(1)	(2)	(3)	(4)
Country	Change	s.e.	Change/dc
Bulgaria*	-0.038*	(0.019)	-0.03
France	-0.144**	(0.015)	-0.10
Germany	-0.043**	(0.013)	-0.03
Hong Kong-China*	0.425**	(0.014)	0.28
Hungary	0.067**	(0.015)	0.04
Iran*	0.278**	(0.018)	0.19
Italy	0.054**	(0.016)	0.04
Lithuania*	0.038**	(0.015)	0.03
Netherlands	-0.087**	(0.013)	-0.06
New Zealand	-0.062**	(0.019)	-0.04
Russian Federation*	0.546**	(0.014)	0.36
Singapore*	0.501**	(0.015)	0.33
Slovakia*	0.187**	(0.015)	0.12
Slovenia*	0.426**	(0.015)	0.28
United Kingdom	0.094**	(0.019)	0.06
United States	0.116**	(0.014)	0.08

Notes & Source: See Table A.8. Table displays cross-country changes in achievement levels in PIRLS between 2001-2016. The changes have been estimated from a quadratic fit. Change per decade (dc) in achievement levels is also displayed. Standard errors are in parenthesis. *p <0.05, **p<0.01.* represents Non-OECD region. Hong Kong is not an independent country and therefore is not discussed in text.

Table A.11: Change in achievement levels at early adolescence in cross-country PISA survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Math			Reading		
Country	Change	s.e.	Change/dc	Change	s.e.	Change/dc
Australia	-0.375**	(0.008)	-0.21	-0.249**	(0.009)	-0.14
Belgium	-0.150**	(0.011)	-0.08	-0.120**	(0.012)	-0.07
Brazil*	0.380**	(0.007)	0.21	0.155**	(0.009)	0.09
Canada	-0.186**	(0.006)	-0.10	-0.105**	(0.006)	-0.06
Czech Republic	-0.096**	(0.012)	-0.05	-0.011	(0.013)	-0.01
Denmark	-0.077**	(0.011)	-0.04	0.056**	(0.012)	0.03
Finland	-0.319**	(0.011)	-0.18	-0.278**	(0.012)	-0.15
France	-0.190**	(0.013)	-0.11	-0.029	(0.015)	-0.02
Germany	0.090**	(0.013)	0.05	0.162**	(0.015)	0.09
Greece	0.034**	(0.013)	0.02	-0.105**	(0.014)	-0.06
Hong Kong-China*	-0.025	(0.014)	-0.01	0.060**	(0.014)	0.03
Hungary	-0.103**	(0.013)	-0.06	-0.075**	(0.014)	-0.04
Iceland	-0.224**	(0.015)	-0.12	-0.241**	(0.018)	-0.13
Indonesia*	0.122**	(0.009)	0.07	0.064**	(0.008)	0.04
Ireland	-0.012	(0.012)	-0.01	-0.022	(0.014)	-0.01
Italy	0.288**	(0.007)	0.16	0.004	(0.008)	0.00
Korea	-0.173**	(0.013)	-0.10	-0.171**	(0.013)	-0.10
Latvia	0.158**	(0.013)	0.09	0.156**	(0.015)	0.09
Luxembourg	0.184**	(0.014)	0.10	0.221**	(0.016)	0.12
Mexico	0.216**	(0.005)	0.12	0.098**	(0.006)	0.05
Netherlands	-0.338**	(0.014)	-0.19	-0.333**	(0.015)	-0.19
New Zealand	-0.364**	(0.014)	-0.20	-0.201**	(0.016)	-0.11
Norway	0.016	(0.013)	0.01	0.061**	(0.016)	0.03
Poland	0.337**	(0.014)	0.19	0.262**	(0.015)	0.15
Portugal	0.334**	(0.012)	0.19	0.247**	(0.013)	0.14
Russian Federation*	0.147**	(0.012)	0.08	0.367**	(0.013)	0.20
Spain	0.044**	(0.006)	0.02	0.283**	(0.012)	0.16
Sweden	-0.146**	(0.014)	-0.08	-0.161**	(0.015)	-0.09
Switzerland	-0.096**	(0.010)	-0.05	-0.086**	(0.011)	-0.05
Thailand*	-0.059**	(0.010)	-0.03	-0.221**	(0.010)	-0.12
United Kingdom	-0.212**	(0.008)	-0.12	-0.147**	(0.010)	-0.08
United States	-0.106**	(0.013)	-0.06	0.020	(0.015)	0.01

Notes & Source: See Table A.8. Table displays cross-country changes in achievement levels in PISA between 2000-2018. PISA test in reading was excluded in 2006 by the administering body. The changes have been estimated from a quadratic fit. Change per decade (dc) in achievement levels is also displayed. Standard errors are in parenthesis. *p <0.05, **p<0.01. * represents Non-OECD country. Hong Kong is not an independent country and therefore is not discussed in text.

Table A.12: Change in achievement levels at younger age by gender and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Math			Reading		
Survey	Gender	N	Change	Change/dc	N	Change	Change/dc
LTT	Male	44,340	0.822**	0.24	68,780	0.298**	0.07
			(0.017)			(0.015)	
	Female	44,830	0.745**	0.22	68,970	0.147**	0.04
			(0.016)			(0.015)	
NAEP	Male	740,100	1.054**	0.39	812,600	0.273**	0.10
			(0.004)			(0.003)	
	Female	715,530	1.050**	0.39	793,430	0.170**	0.06
			(0.003)			(0.003)	
TIMSS	Male	25,490	0.640**	0.32			
			(0.016)				
	Female	25,950	0.581**	0.29			
			(0.015)				
PIRLS	Male				12,980	0.176**	0.12
						(0.022)	
	Female				13,130	0.086**	0.06
						(0.021)	

Notes & Source: See Table 1.

Table A.13: Change in achievement levels at early adolescence by gender and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Math			Reading		
Survey	Gender	N	Change	Change/dc	N	Change	Change/dc
LTT	Male	50,300	0.527**	0.16	70,930	0.152**	0.04
			(0.017)			(0.016)	
	Female	50,900	0.457**	0.13	71,140	0.105**	0.03
			(0.016)			(0.016)	
NAEP	Male	658,950	0.730**	0.27	716,270	0.391**	0.14
			(0.004)			(0.003)	
	Female	643,000	0.701**	0.26	707,290	0.247**	0.09
			(0.004)			(0.003)	
TIMSS	Male	28,110	0.359**	0.18			
			(0.016)				
	Female	28,920	0.390**	0.20			
			(0.015)				
PISA	Male	14,680	-0.140**	-0.09	12,640	0.003	0.00
			(0.024)			(0.023)	
	Female	14,450	-0.152**	-0.10	12,590	-0.059**	-0.04
			(0.022)			(0.021)	

Notes & Source: See Table 1.

Table A.14: Change in achievement levels at older age by gender and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Math			Reading		
Survey	Gender	N	Change	Change/dc	N	Change	Change/dc
LTT	Male	45,420	0.180**	0.05	68,380	0.003	0.00
			(0.018)			(0.015)	
	Female	47,050	0.250**	0.07	68,980	0.031*	0.01
			(0.018)			(0.015)	
NAEP	Male	35,590	0.052**	0.05	71,860	0.081**	0.03
			(0.014)			(0.012)	
	Female	35,900	0.039**	0.04	73,300	0.053**	0.02
			(0.013)			(0.011)	

Notes & Source: See Table 1.

Table A.15: Medians of change/decade in achievement levels in math and reading by SES quartile for each ethnic group.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Ethnicity	Base	SES quart	Young.	Diff.	E.A.	Diff.	Older	Diff.	All ages	Diff.
Math										
Asian	Top	Top	0.24		0.21		0.09		0.14	
		Second	0.42	0.18	0.26	0.06	0.05	-0.05	0.24	0.10
		Third	0.48	0.25	0.30	0.09	0.15	0.06	0.24	0.10
		Bottom	0.42	0.19	0.28	0.08	0.19	0.09	0.28	0.15
Black	Top	Top	0.48		0.27		0.19		0.29	
		Second	0.55	0.07	0.26	-0.01	0.13	-0.06	0.26	-0.03
		Third	0.56	0.08	0.28	0.02	0.08	-0.12	0.28	-0.01
		Bottom	0.59	0.11	0.24	-0.03	0.11	-0.09	0.24	-0.05
Hispanic	Top	Top	0.42		0.26		0.18		0.21	
		Second	0.40	-0.02	0.28	0.02	0.10	-0.08	0.17	-0.04
		Third	0.48	0.06	0.26	0.00	0.18	0.00	0.21	0.00
		Bottom	0.51	0.10	0.29	0.03	0.19	0.01	0.23	0.02
White	Top	Top	0.45		0.15		0.11		0.17	
		Second	0.46	0.01	0.18	0.03	0.11	-0.01	0.18	0.01
		Third	0.49	0.04	0.20	0.05	0.06	-0.05	0.20	0.03
		Bottom	0.50	0.04	0.20	0.05	0.07	-0.04	0.20	0.03
Reading										
Asian	Top	Top	0.03		-0.02		0.05		0.03	
		Second	-0.04	-0.07	-0.01	0.01	0.16	0.11	0.05	0.01
		Third	0.15	0.12	0.05	0.06	0.13	0.08	0.08	0.05
		Bottom	0.09	0.06	0.10	0.11	0.17	0.12	0.16	0.13
Black	Top	Top	0.01		0.07		0.08		0.07	
		Second	0.05	0.04	0.09	0.02	0.13	0.05	0.10	0.02
		Third	0.07	0.06	0.13	0.05	0.09	0.01	0.13	0.05
		Bottom	0.13	0.12	0.13	0.06	0.09	0.01	0.12	0.05
Hispanic	Top	Top	0.01		0.11		0.05		0.05	
		Second	-0.02	-0.03	0.14	0.03	0.04	0.00	0.05	0.00
		Third	0.07	0.06	0.16	0.05	0.07	0.02	0.09	0.03
		Bottom	0.18	0.17	0.28	0.17	0.16	0.11	0.21	0.16
White	Top	Top	-0.05		0.05		0.08		0.03	
		Second	0.00	0.04	0.04	-0.01	0.03	-0.05	0.03	0.00
		Third	-0.01	0.04	0.04	-0.01	0.02	-0.05	0.02	-0.01
		Bottom	0.10	0.15	0.05	0.00	-0.02	-0.10	0.05	0.01

Notes & Source: See Table 1. Table displays medians of change/decade in achievement levels by SES quartile for each ethnic group displayed in Tables A.16 - A.18. E. A. = Early Adolescence. Birth years differ by age categories, depending on the availability of data (see Table 8 for details). Differences between the base category and other categories are also displayed.

Table A.16: Change in achievement levels at younger age by SES quartile and survey for each ethnic group.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
			Math			Read.		
Ethnicity	Quartile	Survey	N	Change	Change/dc	N	Change	Change/dc
Asian	Top	LTT	800	0.024	0.01	540	0.082	0.04
				(0.204)			(0.201)	
	Second			0.541**	0.26		-0.244	-0.13
				(0.183)			(0.173)	
	Third			0.254	0.12		0.363	0.19
				(0.177)			(0.189)	
	Bottom			0.642**	0.31		-0.137	-0.07
				(0.205)			(0.227)	
Black	Top		7,690	0.656**	0.31	4,700	-0.129	-0.07
				(0.069)			(0.083)	
	Second			0.725**	0.35		0.020	0.01
				(0.072)			(0.086)	
	Third			0.812**	0.39		-0.007	0.00
				(0.073)			(0.088)	
	Bottom			0.817**	0.39		0.052	0.03
				(0.072)			(0.092)	
Hispanic	Top		5,510	0.413**	0.20	3,630	0.159	0.08
				(0.076)			(0.090)	
	Second			0.327**	0.16		-0.018	-0.01
				(0.079)			(0.086)	
	Third			0.389**	0.19		0.129	0.07
				(0.078)			(0.091)	
	Bottom			0.337**	0.16		0.262**	0.14
				(0.076)			(0.094)	
White	Top		31,520	0.571**	0.27	21,470	-0.100*	-0.05
				(0.036)			(0.041)	
	Second			0.658**	0.31		0.055	0.03
				(0.036)			(0.042)	
	Third			0.650**	0.31		-0.082	-0.04
				(0.037)			(0.043)	
	Bottom			0.615**	0.29		0.089	0.05
				(0.039)			(0.045)	

Table A.16 (Cont'd): Change in achievement levels at younger age by SES quartile and survey for each ethnic group.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
			Math			Read.		
Ethnicity	Quartile	Survey	N	Change	Change/dc	N	Change	Change/dc
Asian	Top	NAEP	5,620	0.598** (0.063)	0.46	6,030	0.021 (0.058)	0.02
	Second			0.747** (0.065)	0.57		0.059 (0.057)	0.05
	Third			1.098** (0.065)	0.84		0.131* (0.065)	0.10
	Bottom			0.700** (0.059)	0.54		0.336** (0.060)	0.26
Black	Top		28,180	0.849** (0.030)	0.65	30,530	0.122** (0.028)	0.09
	Second			0.984** (0.027)	0.76		0.124** (0.028)	0.10
	Third			0.948** (0.030)	0.73		0.182** (0.029)	0.14
	Bottom			1.029** (0.029)	0.79		0.307** (0.029)	0.24
Hispanic	Top		16,130	0.826** (0.037)	0.64	18,210	-0.085* (0.039)	-0.07
	Second			0.825** (0.036)	0.63		-0.031 (0.039)	-0.02
	Third			0.996** (0.036)	0.77		0.083* (0.039)	0.06
	Bottom			1.124** (0.036)	0.86		0.279** (0.038)	0.21
White	Top		93,670	0.826** (0.016)	0.64	101,990	-0.049** (0.014)	-0.04
	Second			0.799** (0.016)	0.61		-0.041** (0.015)	-0.03
	Third			0.875** (0.016)	0.67		0.030* (0.015)	0.02
	Bottom			0.913** (0.016)	0.70		0.211** (0.015)	0.16

Notes & Source: See Table 1 and Table 8. Also see Table 8 for birth cohorts.

Table A.17: Change in achievement levels at early adolescence by SES quartile and survey for each ethnic group.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
			Math			Read.		
Ethnicity	Quartile	Survey	N	Change	Change/dc	N	Change	Change/dc
Asian	Top	LTT	2,630	0.323**	0.10	2,000	0.433**	0.14
				(0.081)			(0.126)	
	Second			1.043**	0.31		0.371**	0.12
				(0.089)			(0.127)	
	Third			1.129**	0.33		0.252*	0.08
				(0.094)			(0.122)	
	Bottom			0.878**	0.26		0.312*	0.10
				(0.096)			(0.129)	
Black	Top		13,600	0.672**	0.20	9,310	0.228**	0.07
				(0.053)			(0.063)	
	Second			0.739**	0.22		0.302**	0.09
				(0.053)			(0.064)	
	Third			0.767**	0.23		0.403**	0.13
				(0.053)			(0.065)	
	Bottom			0.730**	0.21		0.417**	0.13
				(0.053)			(0.067)	
Hispanic	Top		10,010	0.681**	0.20	7,550	0.074	0.02
				(0.048)			(0.062)	
	Second			0.637**	0.19		0.265**	0.08
				(0.048)			(0.061)	
	Third			0.793**	0.23		0.369**	0.12
				(0.047)			(0.064)	
	Bottom			0.800**	0.24		0.405**	0.13
				(0.046)			(0.063)	
White	Top		62,630	0.441**	0.13	44,120	0.158**	0.05
				(0.027)			(0.031)	
	Second			0.622**	0.18		0.125**	0.04
				(0.027)			(0.031)	
	Third			0.641**	0.19		0.133**	0.04
				(0.028)			(0.031)	
	Bottom			0.608**	0.18		0.156**	0.05
				(0.029)			(0.033)	

Table A.17 (Cont'd): Change in achievement levels at early adolescence by SES quartile and survey for each ethnic group.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
			Math			Read.		
Ethnicity	Quartile	Survey	N	Change	Change/dc	N	Change	Change/dc
Asian	Top	NAEP	25,220	0.644**	0.34	30,330	-0.030	-0.02
				(0.036)			(0.032)	
	Second			0.771**	0.41		-0.014	-0.01
				(0.039)			(0.031)	
	Third			0.493**	0.26		0.093**	0.05
				(0.037)			(0.031)	
	Bottom			0.590**	0.31		0.181**	0.10
				(0.034)			(0.029)	
Black	Top		98,830	0.804**	0.42	117,810	0.251**	0.13
				(0.017)			(0.014)	
	Second			0.971**	0.51		0.338**	0.18
				(0.017)			(0.015)	
	Third			0.832**	0.44		0.353**	0.19
				(0.017)			(0.014)	
	Bottom			0.944**	0.50		0.365**	0.19
				(0.016)			(0.014)	
Hispanic	Top		70,950	0.712**	0.37	80,080	0.359**	0.19
				(0.020)			(0.018)	
	Second			0.773**	0.41		0.363**	0.19
				(0.020)			(0.018)	
	Third			0.789**	0.42		0.340**	0.18
				(0.019)			(0.018)	
	Bottom			0.899**	0.47		0.523**	0.28
				(0.019)			(0.018)	
White	Top		375,350	0.752**	0.40	454,720	0.324**	0.17
				(0.008)			(0.007)	
	Second			0.663**	0.35		0.239**	0.13
				(0.008)			(0.007)	
	Third			0.705**	0.37		0.314**	0.17
				(0.008)			(0.007)	
	Bottom			0.747**	0.39		0.398**	0.21
				(0.008)			(0.007)	

Table A.17 (Cont'd): Change in achievement levels at early adolescence by SES quartile and survey for each ethnic group.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
			Math			Read.		
Ethnicity	Quartile	Survey	N	Change	Change/dc	N	Change	Change/dc
Asian	Top	TIMSS	1,990	0.644**	0.32			
				(0.112)				
	Second			0.442**	0.22			
				(0.115)				
	Third			0.784**	0.39			
				(0.108)				
	Bottom			0.706**	0.35			
				(0.113)				
Black	Top		6,560	0.674**	0.34			
				(0.065)				
	Second			0.613**	0.31			
				(0.062)				
	Third			0.687**	0.34			
				(0.060)				
	Bottom			0.524**	0.26			
				(0.058)				
Hispanic	Top		8,960	0.650**	0.33			
				(0.053)				
	Second			0.745**	0.37			
				(0.051)				
	Third			0.574**	0.29			
				(0.048)				
	Bottom			0.692**	0.35			
				(0.047)				
White	Top		26,090	0.355**	0.18			
				(0.029)				
	Second			0.358**	0.18			
				(0.029)				
	Third			0.425**	0.21			
				(0.028)				
	Bottom			0.454**	0.23			
				(0.028)				

Table A.17 (Cont'd): Change in achievement levels at early adolescence by SES quartile and survey for each ethnic group.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
			Math			Read.		
Ethnicity	Quartile	Survey	N	Change	Change/dc	N	Change	Change/dc
Asian	Top	PISA	1,160	0.139	0.09	1,010	-0.103	-0.07
				(0.151)			(0.139)	
	Second			-0.588**	-0.39		-0.283	-0.19
				(0.173)			(0.159)	
	Third			-0.065	-0.04		-0.248	-0.17
				(0.168)			(0.158)	
	Bottom			0.196	0.13		0.288	0.19
				(0.168)			(0.170)	
Black	Top		3,940	0.027	0.02	3,490	-0.017	-0.01
				(0.075)			(0.078)	
	Second			-0.036	-0.02		0.017	0.01
				(0.073)			(0.077)	
	Third			0.029	0.02		0.113	0.08
				(0.069)			(0.070)	
	Bottom			0.047	0.03		0.093	0.06
				(0.067)			(0.072)	
Hispanic	Top		6,310	0.056	0.04	5,620	0.158*	0.11
				(0.069)			(0.064)	
	Second			0.074	0.05		0.210**	0.14
				(0.062)			(0.062)	
	Third			0.137*	0.09		0.236**	0.16
				(0.062)			(0.062)	
	Bottom			0.236**	0.16		0.533**	0.36
				(0.061)			(0.061)	
White	Top		15,200	-0.302**	-0.20	12,770	-0.207**	-0.14
				(0.040)			(0.039)	
	Second			-0.200**	-0.13		-0.105**	-0.07
				(0.040)			(0.039)	
	Third			-0.234**	-0.16		-0.094*	-0.06
				(0.038)			(0.038)	
	Bottom			-0.158**	-0.11		-0.035	-0.02
				(0.040)			(0.040)	

Notes & Source: See Table 1 and Table 8. Also see Table 8 for birth cohorts.

Table A.18: Change in achievement levels at older age by SES quartile and survey for each ethnic group.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
			Math			Read.		
Ethnicity	Quartile	Survey	N	Change	Change/dc	N	Change	Change/dc
Asian	Top	LTT	2,560	0.036	0.01	2,290	0.105	0.03
				(0.095)			(0.107)	
	Second			-0.056	-0.01		0.223*	0.07
				(0.097)			(0.110)	
	Third			0.288**	0.08		0.156	0.05
				(0.095)			(0.105)	
	Bottom			0.578**	0.17		0.508**	0.16
				(0.097)			(0.114)	
Black	Top		12,130	0.408**	0.12	9,000	0.315**	0.10
				(0.059)			(0.054)	
	Second			0.591**	0.17		0.433**	0.14
				(0.059)			(0.056)	
	Third			0.592**	0.17		0.409**	0.13
				(0.058)			(0.059)	
	Bottom			0.553**	0.16		0.396**	0.12
				(0.056)			(0.058)	
Hispanic	Top		8,320	0.473**	0.14	7,110	0.162**	0.05
				(0.053)			(0.056)	
	Second			0.467**	0.14		0.115*	0.04
				(0.052)			(0.058)	
	Third			0.586**	0.17		0.273**	0.09
				(0.051)			(0.060)	
	Bottom			0.730**	0.21		0.329**	0.10
				(0.051)			(0.059)	
White	Top		62,750	0.188**	0.06	46,520	0.108**	0.03
				(0.029)			(0.027)	
	Second			0.326**	0.10		0.002	0.00
				(0.030)			(0.027)	
	Third			0.366**	0.11		0.018	0.01
				(0.030)			(0.027)	
	Bottom			0.309**	0.09		-0.077**	-0.02
				(0.031)			(0.029)	

Table A.18 (Cont'd): Change in achievement levels at older age by SES quartile and survey for each ethnic group.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
			Math			Read.		
Ethnicity	Quartile	Survey	N	Change	Change/dc	N	Change	Change/dc
Asian	Top	NAEP	3,290	0.178*	0.18	5,250	0.130	0.07
				(0.087)			(0.074)	
	Second			0.106	0.11		0.465**	0.24
				(0.079)			(0.073)	
	Third			0.216*	0.22		0.410**	0.22
				(0.083)			(0.077)	
	Bottom			0.201*	0.20		0.360**	0.19
				(0.089)			(0.081)	
Black	Top		9,540	0.267**	0.27	17,570	0.128**	0.07
				(0.046)			(0.039)	
	Second			0.086*	0.09		0.252**	0.13
				(0.044)			(0.038)	
	Third			-0.024	-0.02		0.112**	0.06
				(0.042)			(0.038)	
	Bottom			0.052	0.05		0.099*	0.05
				(0.043)			(0.039)	
Hispanic	Top		10,180	0.219**	0.22	15,180	0.080	0.04
				(0.044)			(0.042)	
	Second			0.061	0.06		0.099*	0.05
				(0.045)			(0.042)	
	Third			0.178**	0.18		0.107*	0.06
				(0.045)			(0.043)	
	Bottom			0.161**	0.16		0.410**	0.22
				(0.042)			(0.045)	
White	Top		44,170	0.170**	0.17	81,710	0.225**	0.12
				(0.021)			(0.018)	
	Second			0.115**	0.12		0.114**	0.06
				(0.020)			(0.018)	
	Third			0.019	0.02		0.083**	0.04
				(0.020)			(0.019)	
	Bottom			0.058**	0.06		-0.030	-0.02
				(0.020)			(0.020)	

Notes & Source: See Table 1 and Table 8. Also see Table 8 for birth cohorts.

Table A.19: Medians of change/decade in achievement levels in math and reading for different SES constructs by subgroups and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Base	Subgroup	Young	Diff.	E. A.	Diff.	Older	Diff.	All age	Diff.
Math									
Pared.: Grad. coll.	Grad. coll.	0.44		0.11		0.05		0.11	
	Post h.s.	0.38	-0.06	0.08	-0.03	0.01	-0.04	0.08	-0.03
	Grad. h.s.	0.35	-0.09	0.12	0.01	0.02	-0.03	0.09	-0.02
	< h.s.	0.45	0.01	0.21	0.10	0.10	0.05	0.21	0.10
Poss. quart: Top	Top	0.31		0.18		0.14		0.19	
	Second	0.28	-0.03	0.17	-0.01	0.10	-0.04	0.17	-0.02
	Third	0.26	-0.05	0.15	-0.03	-0.11	-0.25	0.15	-0.04
	Bottom	0.36	0.05	0.24	0.06	0.08	-0.06	0.24	0.05
FRL: No	No	0.32		0.27		0.17		0.26	
	Yes	0.31	-0.01	0.27	0.00	0.15	-0.02	0.25	-0.01
Two parents: Yes	Yes	0.29		0.20		0.12		0.20	
	No	0.32	0.03	0.13	-0.07	0.17	0.05	0.22	0.02
Reading									
Pared.: Grad. coll.	Grad. coll.	0.05		0.10		0.08		0.10	
	Post h.s.	-0.08	-0.13	0.01	-0.09	-0.05	-0.13	-0.02	-0.12
	Grad. h.s.	0.03	-0.02	0.08	-0.02	-0.04	-0.12	-0.02	-0.12
	< h.s.	0.06	0.01	0.13	0.03	0.03	-0.05	0.06	-0.04
Poss. quart: Top	Top	0.05		0.10		0.08		0.06	
	Second	0.09	0.04	0.07	-0.03	0.03	-0.05	0.05	-0.01
	Third	0.05	0.00	0.05	-0.05	-0.03	-0.11	0.02	-0.04
	Bottom	0.11	0.06	0.09	-0.01	-0.01	-0.09	0.11	0.05
FRL: No	No	0.18		0.24		0.15		0.18	
	Yes	0.19	0.01	0.17	-0.07	0.16	0.01	0.19	0.01
Two parents: Yes	Yes	0.05		-0.03		0.01		0.03	
	No	0.12	0.07	0.07	0.10	-0.02	-0.03	0.07	0.04

Notes & Source: See Table 1. Table displays medians of change/decade in achievement levels in math and reading for different SES constructs by subgroups and survey displayed in Tables A.20-A.27. E. A. = Early Adolescence. Bir yr = Birth year, Pared. = Parental education, Grad. = Graduated, coll. = college, h.s. = high school, Poss. = Possession index, quart=quartile and FRL = Free or reduced lunch. Birth years differ across subgroups, depending on the availability of data (see tables and appendix for details). Differences between the base category and other categories are also displayed.

Table A.20: Change in achievement levels in math and reading at younger age by parent's education levels and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Math				Reading			
Parent's Educ.	Survey	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
Grad. coll.	LTT	1968-1989	25,800	0.544**	0.26	1961-1989	37,830	0.451**	0.16
				(0.020)				(0.018)	
Post h.s.			5,060	0.399**	0.19		19,000	-0.380**	-0.14
				(0.050)				(0.054)	
Grad. h.s.			11,860	0.202**	0.10		24,050	-0.123**	-0.04
				(0.036)				(0.027)	
< h.s.			3,850	0.596**	0.28		8,630	0.168**	0.06
				(0.065)				(0.055)	
Grad. coll.	NAEP	1980-1993	101,260	0.787**	0.61	1980-1993	111,090	-0.072**	-0.06
				(0.008)				(0.008)	
Post h.s.			16,060	0.735**	0.57		18,440	-0.028	-0.02
				(0.022)				(0.020)	
Grad. h.s.			22,420	0.763**	0.59		23,660	0.113**	0.09
				(0.020)				(0.018)	
< h.s.			7,750	0.811**	0.62		8,000	0.073*	0.06
				(0.032)				(0.029)	

Notes & Source: See Table 1. Table displays change in achievement levels at younger age by parent's education levels and survey. Students report their parent's education levels. For this table, the parent's education levels were recoded in four comparable categories (graduated college, post high school, graduated high school and less than high school) across all surveys (see notes in Table A.5).

Table A.21: Change in achievement levels in math and reading at early adolescence by parent's education levels and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Math				Reading			
Parent's Educ.	Survey	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
Grad. coll.	LTT	1964-1998	40,410	0.433** (0.015)	0.13	1957-1998	45,280	0.759** (0.016)	0.19
Post h.s.			15,080	0.325** (0.028)	0.10		29,780	-0.187** (0.031)	-0.05
Grad. h.s.			26,100	0.247** (0.027)	0.07		41,110	-0.182** (0.026)	-0.04
< h.s.			8,850	0.502** (0.045)	0.15		15,150	0.097* (0.045)	0.02
Grad. coll.	NAEP	1976-2003	603,560	0.668** (0.004)	0.25	1976-2003	661,970	0.276** (0.003)	0.10
Post h.s.			203,910	0.470** (0.007)	0.17		231,080	0.171** (0.006)	0.06
Grad. h.s.			216,380	0.542** (0.007)	0.20		235,590	0.219** (0.006)	0.08
< h.s.			90,140	0.731** (0.010)	0.27		95,870	0.356** (0.009)	0.13
Grad. coll.	TIMSS	1985-2001	18,470	0.124** (0.020)	0.08				
Post h.s.		1981-2001	11,850	0.095** (0.031)	0.05				
Grad. h.s.			9,790	0.323** (0.026)	0.16				
< h.s.			6,450	0.410** (0.035)	0.21				

Table A.21 (Cont'd): Change in achievement levels in math and reading at early adolescence by parent's education levels and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Math				Reading			
Parent's Educ.	Survey	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
Grad. coll.	PISA	1985-2000	11,660	-0.235**	-0.16	1985-2000	10,190	-0.161**	-0.11
				(0.026)				(0.024)	
Post h.s.			4,970	-0.174**	-0.12		4,620	0.012	0.01
				(0.035)				(0.036)	
Grad. h.s.			9,630	0.012	0.01		7,800	0.126**	0.08
				(0.032)				(0.032)	
< h.s.			2,160	0.309**	0.21		1,810	0.545**	0.36
				(0.062)				(0.061)	

Notes & Source: See Table 1 and Table A.20.

Table A.22: Change in achievement levels in math and reading at older age by parent's education levels and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Math				Reading			
Parent's Educ.	Survey	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
Grad. coll.	LTT	1961-1995	36,230	0.112**	0.03	1954-1995	40,000	0.530**	0.13
				(0.017)				(0.016)	
Post h.s.			18,090	0.039	0.01		35,340	-0.334**	-0.08
				(0.025)				(0.023)	
Grad. h.s.			24,440	0.017	0.01		40,080	-0.242**	-0.06
				(0.030)				(0.023)	
< h.s.			10,350	0.363**	0.11		18,000	0.082*	0.02
				(0.040)				(0.036)	
Grad. coll.	NAEP	1988-1998	34,520	0.064**	0.06	1973-1998	67,470	0.082**	0.03
				(0.013)				(0.016)	
Post h.s.			15,410	0.002	0.00		33,240	-0.049**	-0.02
				(0.017)				(0.016)	
Grad. h.s.			12,770	0.021	0.02		27,180	-0.039*	-0.02
				(0.020)				(0.019)	
< h.s.			5,620	0.084**	0.08		11,520	0.098**	0.04
				(0.029)				(0.026)	

Notes & Source: See Table 1 and Table A.20.

Table A.23: Change in achievement levels in math and reading at younger age by possession index quartiles and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Math				Reading			
Quartile	Survey	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
Top	LTT	1968-2002	88,570	0.737**	0.22	1961-2002	116,440	0.127**	0.03
				(0.021)				(0.021)	
Second				0.665**	0.20			0.130**	0.03
				(0.022)				(0.021)	
Third				0.716**	0.21			0.195**	0.05
				(0.022)				(0.022)	
Bottom				0.999**	0.29			0.410**	0.10
				(0.022)				(0.022)	
Top	NAEP	1980-2007	1,408,140	0.943**	0.35	1980-2007	1,557,550	0.131**	0.05
				(0.005)				(0.004)	
Second				1.168**	0.43			0.281**	0.10
				(0.005)				(0.004)	
Third				1.073**	0.40			0.236**	0.09
				(0.005)				(0.004)	
Bottom				1.084**	0.40			0.291**	0.11
				(0.005)				(0.004)	
Top	TIMSS	1985-2005	51,020	0.624**	0.31				
				(0.022)					
Second				0.555**	0.28				
				(0.021)					
Third				0.526**	0.26				
				(0.020)					
Bottom				0.716**	0.36				
				(0.020)					

Table A.23 (Cont'd): Change in achievement levels in math and reading at younger age by possession index quartiles and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Math				Reading			
Quartile	Survey	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
Top	PIRLS					1991-2006	25,840	0.154**	0.10
								(0.029)	
Second								0.130**	0.09
								(0.028)	
Third								0.035	0.02
								(0.029)	
Bottom								0.170**	0.11
								(0.030)	

Notes & Source: See Table 1. Table displays change in achievement levels at younger age by home item quartiles and survey. Students report items in their homes. For this table, we included items that were consistently available and similarly measured for all administrations within a subject, age/grade, and survey. (see notes in Table A.6).

Table A.24: Change in achievement levels in math and reading at early adolescence by possession index quartiles and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Math				Reading			
Quartile	Survey	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
Top	LTT	1964-1998	99,920	0.578**	0.17	1957-1998	120,820	0.246**	0.06
				(0.021)				(0.022)	
Second				0.391**	0.12			0.008	0.00
				(0.021)				(0.022)	
Third				0.472**	0.14			0.084**	0.02
				(0.022)				(0.024)	
Bottom				0.499**	0.15			0.141**	0.03
				(0.022)				(0.025)	
Top	NAEP	1976-2003	1,259,180	0.770**	0.29	1976-2003	1,380,970	0.368**	0.14
				(0.005)				(0.005)	
Second				0.759**	0.28			0.364**	0.13
				(0.006)				(0.005)	
Third				0.619**	0.23			0.200**	0.07
				(0.006)				(0.005)	
Bottom				0.719**	0.27			0.366**	0.14
				(0.006)				(0.005)	
Top	TIMSS	1981-2001	56,260	0.351**	0.18				
				(0.022)					
Second				0.343**	0.17				
				(0.021)					
Third				0.303**	0.15				
				(0.021)					
Bottom				0.488**	0.24				
				(0.020)					

Table A.24 (Cont'd): Change in achievement levels in math and reading at early adolescence by possession index quartiles and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Math				Reading			
Quartile	Survey	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
Top	PISA	1985-2000	28,800	-0.315**	-0.21	1985-2000	24,860	-0.214**	-0.14
				(0.030)				(0.029)	
Second				-0.224**	-0.15			-0.136**	-0.09
				(0.030)				(0.029)	
Third				-0.164**	-0.11			-0.006	0.00
				(0.030)				(0.029)	
Bottom				0.114**	0.08			0.234**	0.16
				(0.030)				(0.030)	

Notes & Source: See Table 1 and Table A.23.

Table A.25: Change in achievement levels in math and reading at older age by possession index quartiles and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Math				Reading			
Quartile	Survey	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
Top	LTT	1961-1995	89,470	0.313**	0.09	1954-1995	113,910	0.164**	0.04
				(0.024)				(0.021)	
Second				0.125**	0.04			-0.051*	-0.01
				(0.024)				(0.021)	
Third				0.138**	0.04			-0.110**	-0.03
				(0.025)				(0.022)	
Bottom				0.224**	0.07			0.023	0.01
				(0.025)				(0.024)	
Top	NAEP	1988-1998	70,740	0.189**	0.19	1973-1998	143,960	0.264**	0.11
				(0.018)				(0.015)	
Second				0.155**	0.16			0.174**	0.07
				(0.017)				(0.015)	
Third				-0.261**	-0.26			-0.080**	-0.03
				(0.018)				(0.016)	
Bottom				0.078**	0.08			-0.072**	-0.03
				(0.017)				(0.017)	

Notes & Source: See Table 1 and Table A.23.

Table A.26: Change in achievement levels in math and reading by eligibility for free or reduced lunch and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
			Math				Reading			
FRL elig.	Survey	Grade/Age	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
No	LTT	Age 9	1994-2002	11,080	0.212** (0.021)	0.27	1994-2002	11,180	0.187** (0.019)	0.23
Yes				10,970	0.182** (0.023)	0.23		10,830	0.187** (0.020)	0.23
No	NAEP	Grade 4	1986-2007	669,150	0.747** (0.004)	0.36	1988-2007	744,280	0.226** (0.003)	0.12
Yes				705,110	0.800** (0.004)	0.38		759,620	0.271** (0.003)	0.14
No	LTT	Age 13	1990-1998	12,570	0.222** (0.018)	0.28	1990-1998	12870	0.277** (0.022)	0.35
Yes				9,490	0.205** (0.021)	0.26		9,770	0.185** (0.024)	0.23
No	NAEP	Grade 8	1982-2003	655,400	0.535** (0.004)	0.25	1984-2003	716,580	0.236** (0.003)	0.12
Yes				566,760	0.557** (0.005)	0.27		599,240	0.204** (0.004)	0.11
No	LTT	Age 17	1987-1995	14,420	0.138** (0.017)	0.17	1987-1995	14,950	0.194** (0.019)	0.24
Yes				7,620	0.104** (0.025)	0.13		7,950	0.230** (0.026)	0.29
No	NAEP	Grade 12	1988-1998	45,140	0.162** (0.011)	0.16	1981-1998	65,960	0.088** (0.012)	0.05
Yes				21,230	0.164** (0.016)	0.16		29,420	0.052** (0.019)	0.03

Notes & Source: See Table 1. Table displays change in achievement levels eligibility for free or reduced lunch and survey.

Table A.27: Change in achievement levels in math and reading by family structure and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Subject		Math				Reading			
#Parents	Survey	Grade/Age	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
Two	LTT	Age 9	1976-1989	28,950	0.264** (0.016)	0.20				
Other				8,470	0.300** (0.030)	0.23				
Two	NAEP	Grade 4	1980-2007	365,080	0.995** (0.004)	0.37	1980-2007	369,400	0.144** (0.003)	0.05
Other				118,300	1.104** (0.007)	0.41		118,160	0.329** (0.006)	0.12
Two	LTT	Age 13	1972-1985	25,400	0.189** (0.014)	0.15				
Other				7,360	0.056* (0.027)	0.04				
Two	NAEP	Grade 8	1976-2003	334,850	0.640** (0.004)	0.24	1976-2003	341,210	0.259** (0.004)	0.10
Other				124,320	0.587** (0.007)	0.22		124,050	0.281** (0.006)	0.10
Two	TIMSS	Grade 8	1981-1985	14,550	0.290** (0.017)	0.73				
Other				4,430	0.329** (0.031)	0.82				
Two	PISA	Age 15	1985-1997	11,130	-0.235** (0.022)	-0.20	1985-1997	11,940	-0.187** (0.020)	-0.16
Other				5,790	-0.006 (0.033)	-0.01		6,580	0.044 (0.032)	0.04

Table A.27 (Cont'd): Change in achievement levels in math and reading by family structure and survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Subject		Math				Reading			
#Parents	Survey	Grade/Age	Birth Cohorts	N	Change	Change/dc	Birth Cohorts	N	Change	Change/dc
Two	LTT	Age 17	1969-1982	15,030	0.153** (0.020)	0.12				
Other				4,510	0.225** (0.036)	0.17				
Two	NAEP	Grade 12					1973-1998	37,950	0.029* (0.014)	0.01
Other								14,610	-0.038 (0.022)	-0.02

Notes & Source: See Table 1. Table displays change in achievement levels by family structure and survey (reported by students: two parent family and others; surveys do not ask if the parents are biological parents) (see details in Table A.7).

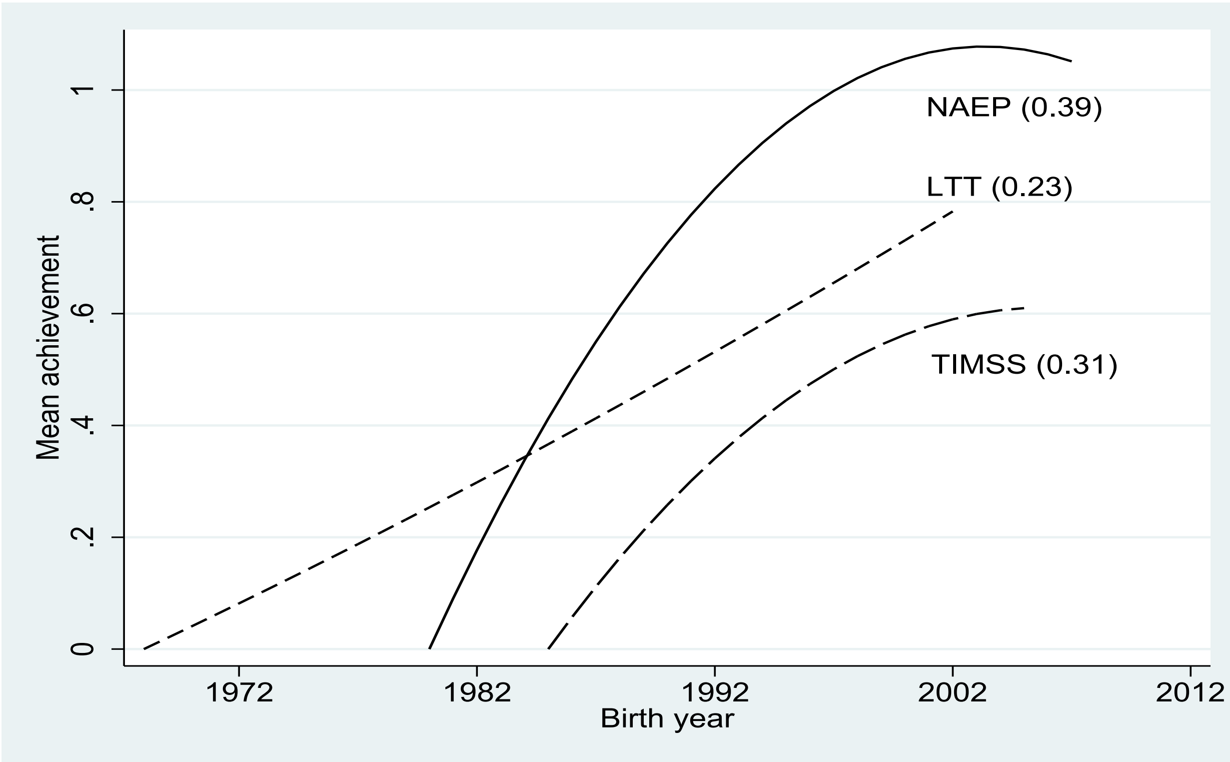

Figure A.1. Change in achievement levels in math at younger age by survey: Birth Cohort 1968-2007.
Note: Figure shows change in achievement levels at younger age by survey. Normalized achievement is measured in standard deviations (s.d.). The s.d. is the difference between the year test was administered and the starting year for a specific test series. The lines represent a quadratic fit. Each line is forced to begin at zero. Magnitude of the rate of change in achievement levels per decade is displayed in parenthesis for each line.
Source: See Table 1.

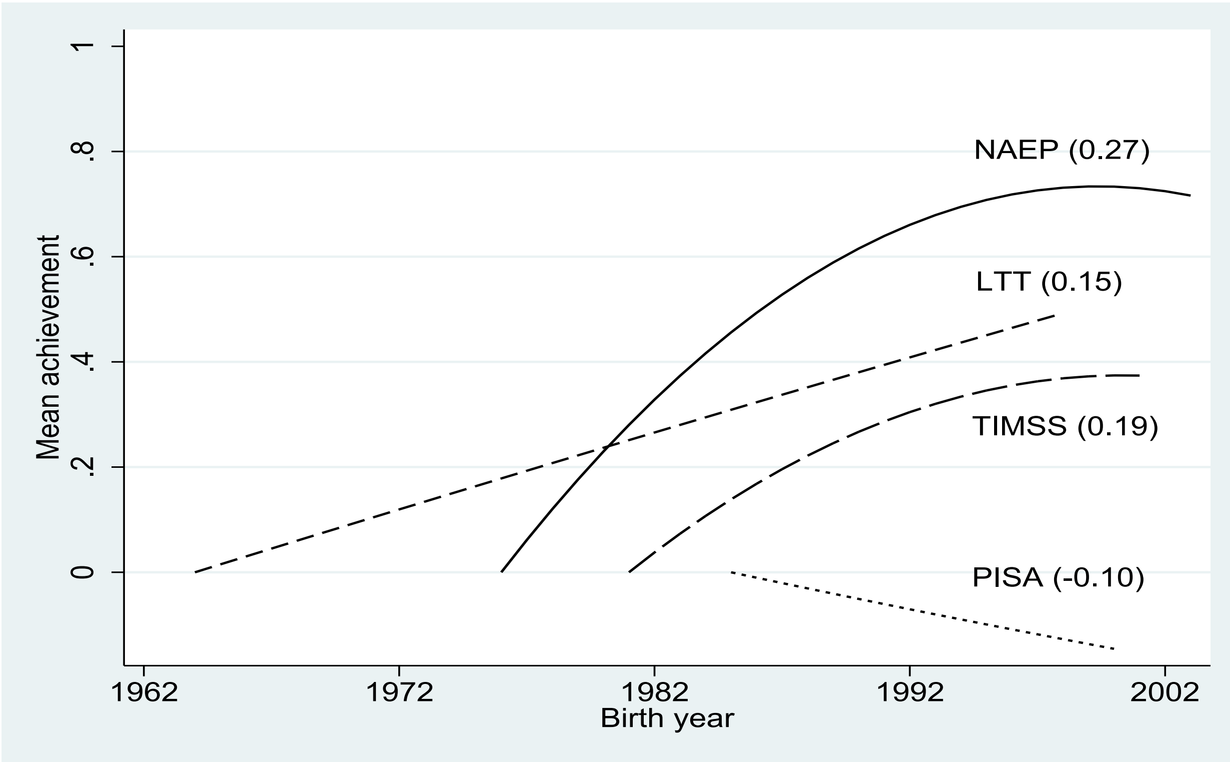

Figure A.2. Change in achievement levels in math at early adolescence by survey: Birth Cohort 1964-2003.
Notes & Source: See Figure A.1 and Table 1.

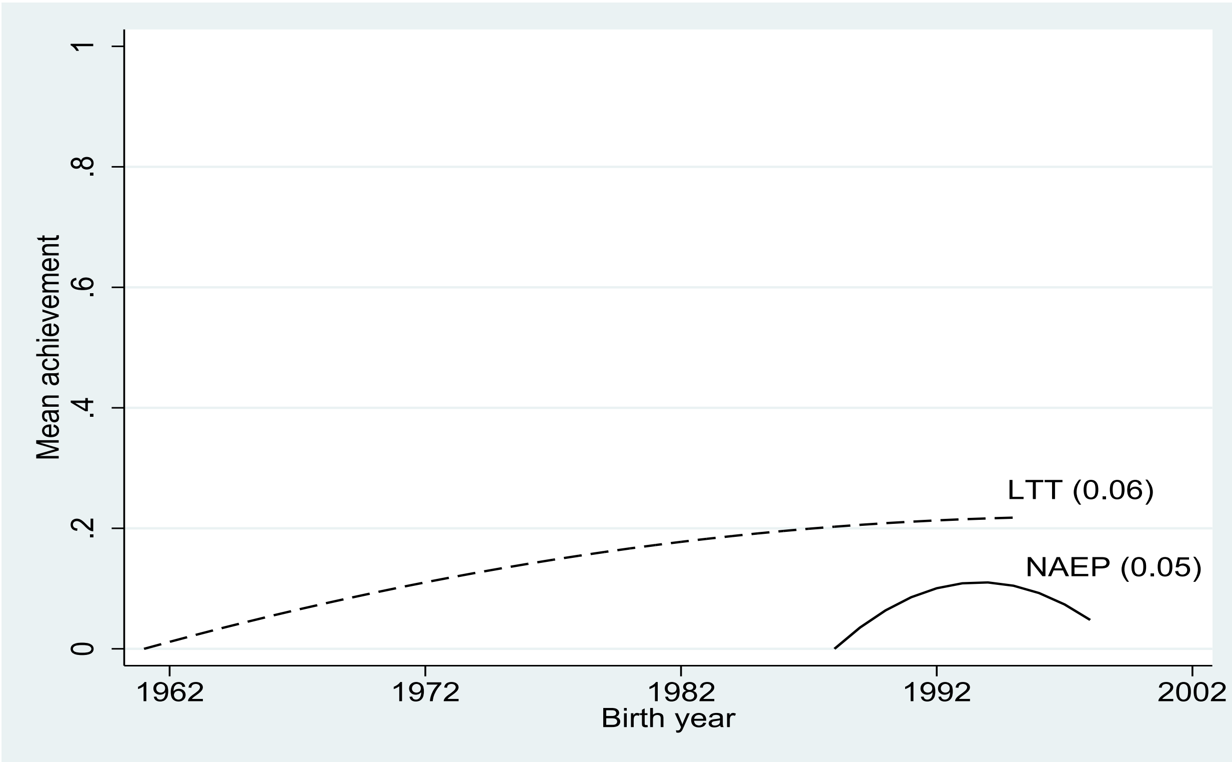

Figure A.3. Change in achievement levels in math at older age by survey: Birth Cohort 1961-1998.
Notes & Source: See Figure A.1 and Table 1.

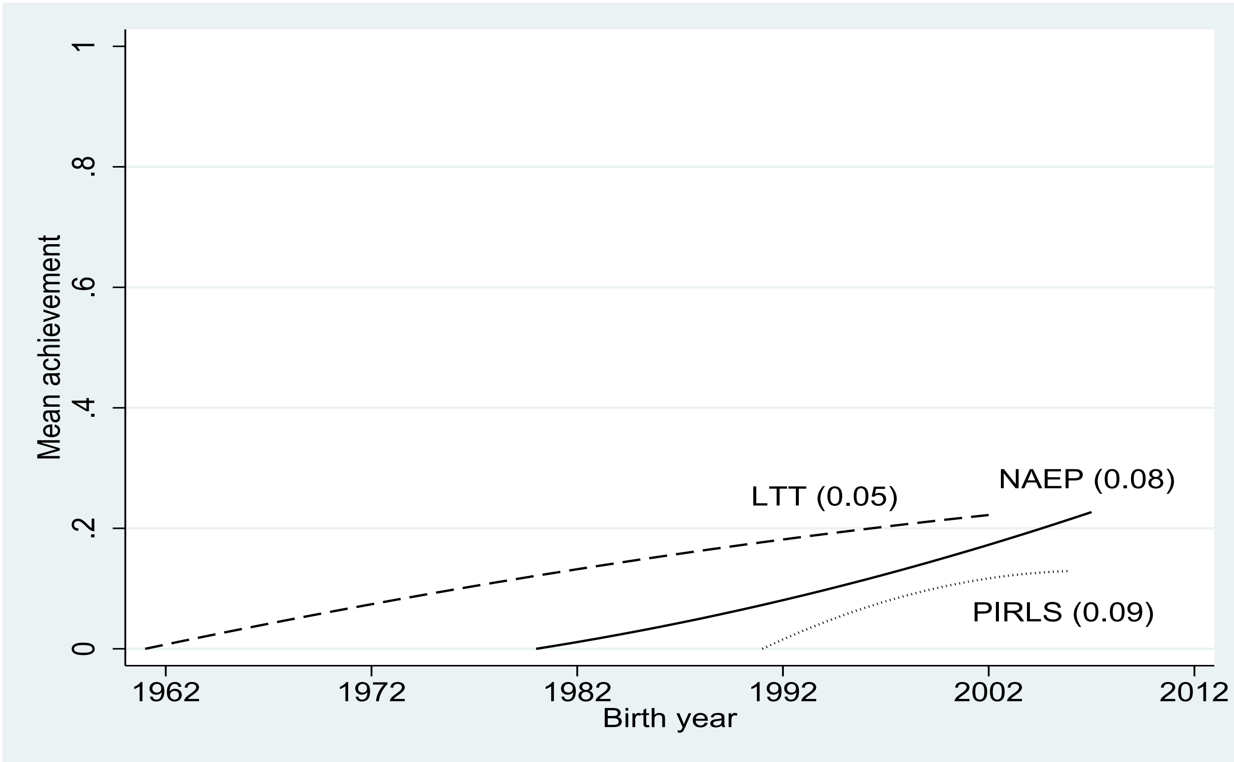

Figure A.4. Change in achievement levels in reading at young age by survey: Birth Cohort 1961-2007.
Notes & Source: See Figure A.1 and Table 1.

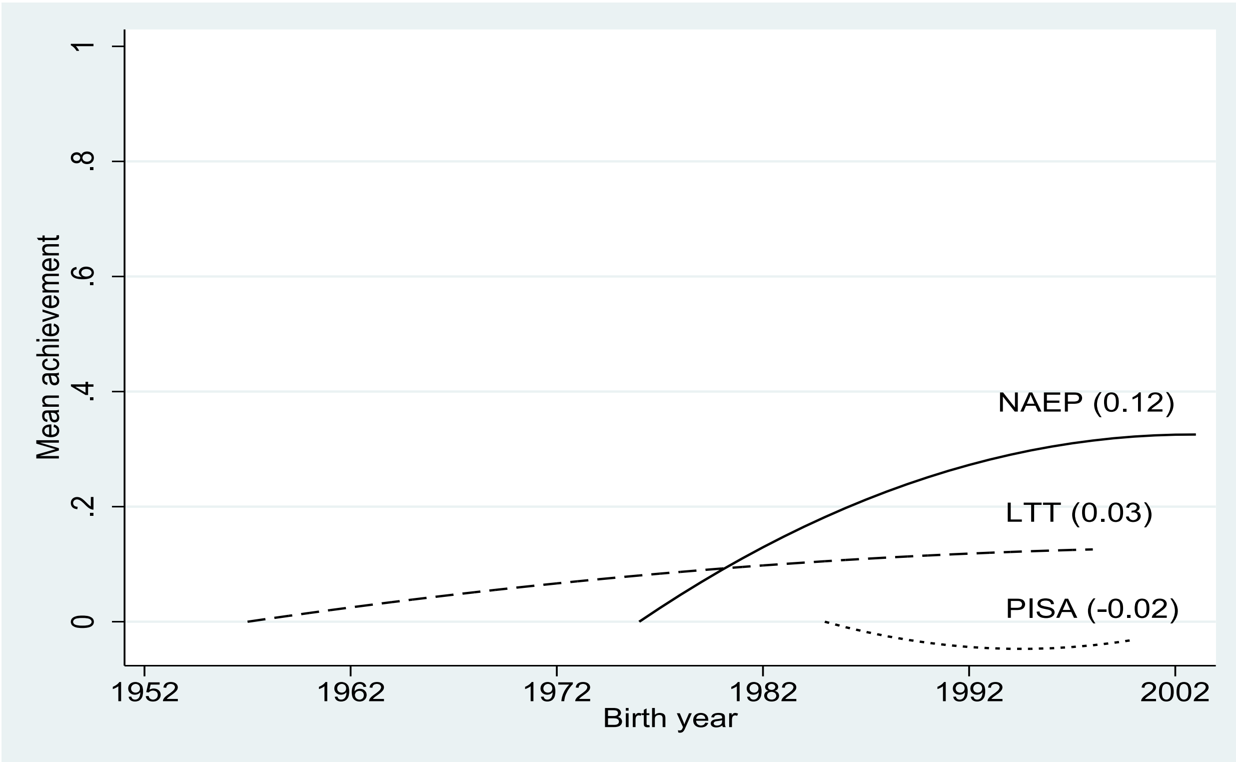

Figure A.5. Change in achievement levels in reading at early adolescence by survey: Birth Cohort 1957-2003.
Notes & Source: See Figure A.1 and Table 1.

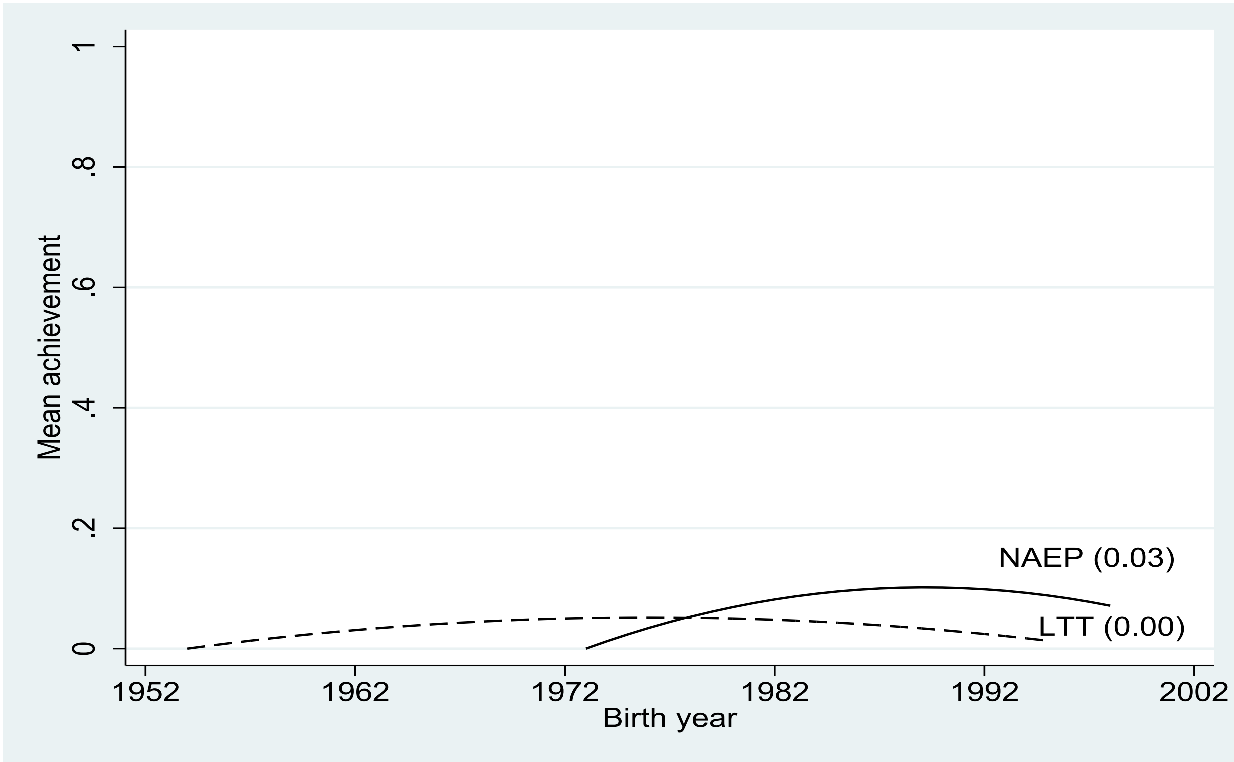

Figure A.6. Change in achievement levels in reading at older age by survey: Birth Cohort 1954-1998.
Notes & Source: See Figure A.1 and Table 1.

Figure A.7. Change in achievement levels in math at age 13 in LTT for linear, quadratic, and cubic fit. These changes are 0.50, 0.49 and 0.54 standard deviations.
Notes & Source: See Figure A.1 and Table 1.

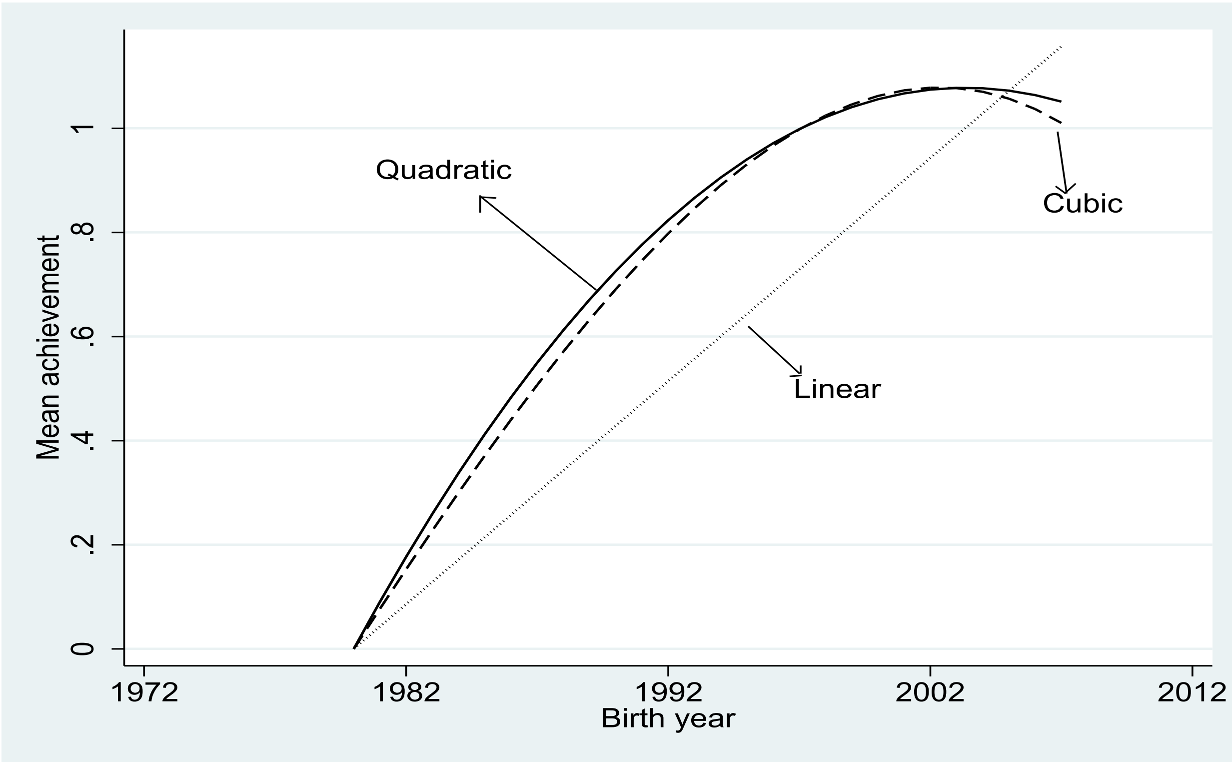

Figure A.8. Change in achievement levels in math at grade 4 in NAEP for linear, quadratic, and cubic fit. These changes are 1.16, 1.05 and 1.01 standard deviations.

Notes & Source: See Figure A.1 and Table 1.

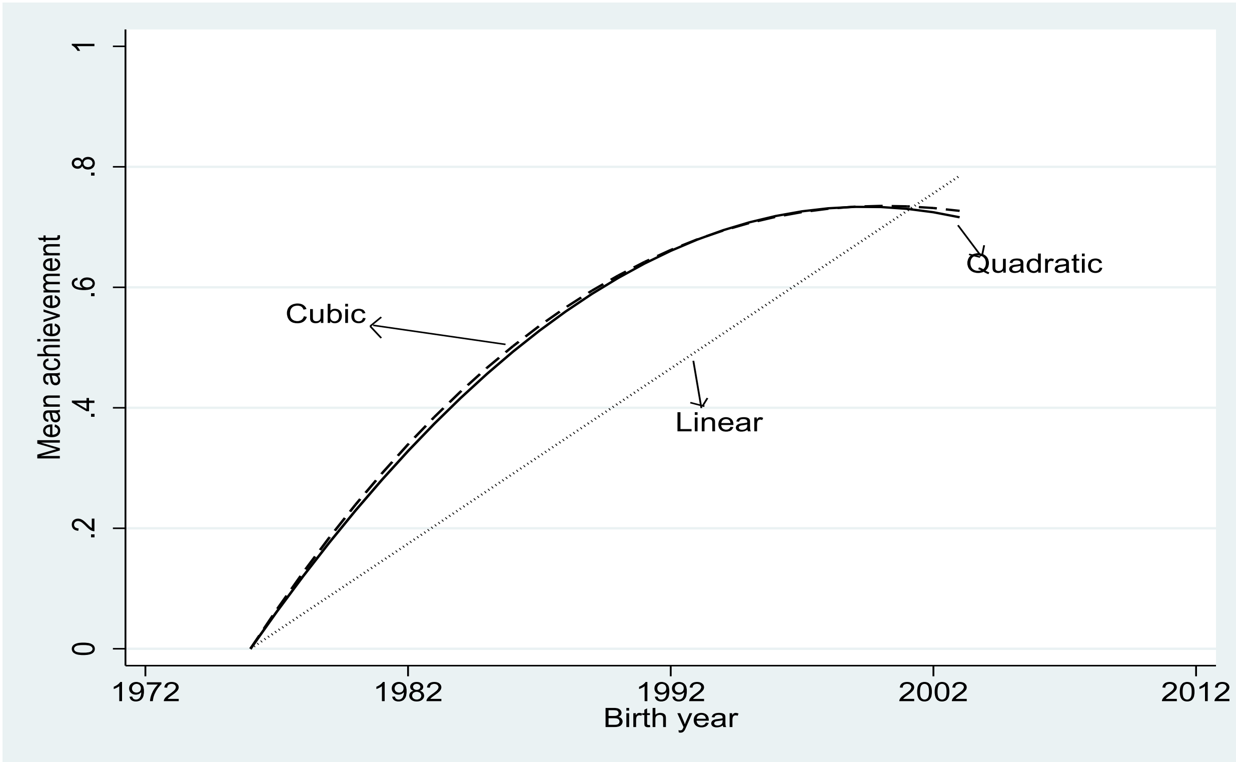

Figure A.9. Change in achievement levels in math at grade 8 in NAEP for linear, quadratic, and cubic fit. These changes are 0.79, 0.72 and 0.73 standard deviations.
Notes & Source: See Figure A.1 and Table 1.