American Sociological Review
2019, Vol. 84(3) 517–544

DOI: 10.1177/0003122419847165
journals.sagepub.com/home/asr

The Global Increase in the Socioeconomic Achievement Gap, 1964 to 2015

Anna K. Chmielewski[a]

Abstract

The "socioeconomic achievement gap"—the disparity in academic achievement between students from high- and low-socioeconomic status (SES) backgrounds—is well-known in the sociology of education. The SES achievement gap has been documented across a wide range of countries. Yet in most countries, we do not know whether the SES achievement gap has been *changing* over time. This study combines 30 international large-scale assessments over 50 years, representing 100 countries and about 5.8 million students. SES achievement gaps are computed between the 90th and 10th percentiles of three available measures of family SES: parents' education, parents' occupation, and the number of books in the home. Results indicate that, for each of the three SES variables examined, achievement gaps increased in a majority of sample countries. Yet there is substantial cross-national variation in the size of increases in SES achievement gaps. The largest increases are observed in countries with rapidly increasing school enrollments, implying that expanding access reveals educational inequality that was previously hidden outside the school system. However, gaps also increased in many countries with consistently high enrollments, suggesting that cognitive skills are an increasingly important dimension of educational stratification worldwide.

Keywords

education, international comparison, academic achievement, socioeconomic inequality

The existence of a "socioeconomic achievement gap"—a disparity in scores on tests of academic achievement between students from high- and low-socioeconomic status (SES) backgrounds—is well-known in sociology and education research. International assessments show that SES achievement gaps are present across a wide range of countries (Mullis et al. 2016; OECD 2016). This suggests that, in most societies, low-SES children do not receive the same learning experiences in or out of school as do their high-SES counterparts. Across many countries, SES achievement gaps impede upward mobility (Jackson 2013). This contradicts the traditional view of education in the United States as a "great equalizer" (Downey and Condron 2016), but it may be less surprising in societies that historically have not viewed themselves as meritocracies (Janmaat 2013).

Recently, there has been heightened interest in whether the SES achievement gap might be *changing* over time. Studies have

[a]University of Toronto

Corresponding Author:
Anna K. Chmielewski, University of Toronto, Ontario Institute for Studies in Education (OISE), 252 Bloor Street West, Toronto, ON M5S1V6, Canada
Email: ak.chmielewski@utoronto.ca

found increasing SES achievement gaps in the United States (Reardon 2011b), South Korea (Byun and Kim 2010), and Malaysia (Saw 2016). Reports by the organizations that administer two major international assessments—the Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS)—document wide cross-national variation in trends in SES achievement gaps across the years of each assessment (Broer, Bai, and Fonseca forthcoming; OECD 2018).

This article aims to provide the most comprehensive picture to date of cross-national trends in the SES achievement gap. I use evidence from 51 years (55 cohort birth-years) of international large-scale assessments, dating from the First International Mathematics Study (FIMS) in 1964 to recent data from PISA, TIMSS, and the Progress in International Reading Literacy Study (PIRLS). I draw on 30 datasets across 100 countries representing some 5.8 million students, and I describe the global trend and cross-national variation in SES achievement gaps, as well as identify the possible causes of this variation.

EVIDENCE ON TRENDS IN THE SES ACHIEVEMENT GAP

It is difficult to draw conclusions about the global trend in SES achievement gaps based on prior research, as different studies have used different data sources, different SES and achievement measures, and covered different time periods. Some early international evidence suggests that SES achievement gaps may have increased in a number of countries between the 1970s and 1990s. The associations between science achievement and SES measures (parent education, parent occupation, and household books) increased between the First International Science Study (FISS) of 1970 and the Second International Science Study (SISS) of 1984 (Keeves 1992). The authors of the SISS report wrote that this increase might be partly attributable to increased validity of home background measures but was likely also related to "increased polarization in society and in the benefits that flow from education" (Keeves 1992:11). Baker, Goesling, and LeTendre (2002) show that in developing countries between the 1970s and 1995, the importance of family SES grew relative to school resources in predicting students' achievement, a change they attribute to expanding school access and standardization of school quality.

More recent single-country studies also suggest increasing SES achievement gaps, but they have produced some contradictory evidence. Using 19 nationally representative U.S. studies, Reardon (2011b) shows that gaps in reading and math achievement between students from families at the 90th and 10th income percentiles grew by about 40 percent between children born in the 1970s and the 1990s. However, the U.S. gap appears to have narrowed slightly for children born in the subsequent decade (Reardon and Portilla 2016). In contrast, using data from PISA, TIMSS, and the National Assessment of Educational Progress (NAEP) for students born between the 1950s and 2000, Hanushek and colleagues (2019) find no change in gaps in reading or math achievement between the 90th and 10th percentiles of an index of SES (including parent education and household possessions). In South Korean subsamples from three waves of TIMSS (corresponding to birth years 1985 to 1993), Byun and Kim (2010) find a strengthening association between math achievement and an index of SES (including parent education and household possessions). Using Malaysian subsamples from four waves of TIMSS (corresponding to birth years 1985 to 1997), Saw (2016) observes rapid growth in math and science achievement gaps between students whose parents attended postsecondary education and those who did not.

Two recent reports on trends in SES achievement gaps for a larger set of countries across waves of PISA and TIMSS also produce inconsistent evidence. Looking at PISA 2015 and one earlier wave (corresponding approximately to birth years 1985 to 2000), associations between reading, math, and science achievement and an SES index (including parent education, parent occupation, and household possessions) declined in a majority of the 60 participating countries (OECD 2018). In contrast, for TIMSS 1995 and 2015 (corresponding to birth years 1981 and 2001), achievement gaps in math and science between the top and bottom quartiles of an SES index (including parent education and household possessions) increased in about half of the 13 countries (Broer et al. forthcoming). Several countries or jurisdictions have trends in different directions in the PISA and TIMSS reports, including Hong Kong, Hungary, Korea, New Zealand, Norway, and Slovenia.

Thus, the evidence on international trends in SES achievement gaps is mixed, calling for a more comprehensive analysis that measures SES achievement gaps consistently across countries and years. Furthermore, all previous research finds wide cross-national variation in the size and direction of changes in SES achievement gaps. What could explain cross-national differences in SES achievement gap trends?

EXPLANATIONS FOR TRENDS IN SES ACHIEVEMENT GAPS

The authors of the three single-country studies described in the previous section offer a number of potential explanations for growing achievement gaps, including rising income inequality, increasing school choice, and growing inequality in parental investments in children (Byun and Kim 2010; Reardon 2011b; Saw 2016). However, it is difficult to adjudicate among different explanations in a single-country study, where multiple causes may be occurring simultaneously. A large body of international comparative research shows which country characteristics are associated with larger SES achievement gaps, but most of this research is cross-sectional—conducted at a single point in time. With such a design, it is difficult to isolate the causes of gaps, as differences between countries may be due to a wide variety of cultural and historical factors. Thus, examining changes in gaps over time across a large number of countries improves upon prior single-country and cross-sectional evidence on the causes of SES achievement gaps.

Previous research suggests several candidates for trends that could drive increasing SES achievement gaps in many countries. First, the population of students enrolled in schools has become more diverse. Primary and lower-secondary school enrollment has become virtually universal in developed countries and has increased dramatically in less developed countries (Baker et al. 2002). Because the target population of international assessments includes only students currently enrolled in school, countries with the most rapidly expanding school access may appear to have growing SES achievement gaps due to the inclusion of relatively disadvantaged populations. Additionally, increasing global migration has led to a larger share of immigrant students enrolled in schools in many countries, which could lead to growing SES achievement gaps in those countries, to the extent that immigrant students are lower-achieving and lower-SES than native-born students (Andon, Thompson, and Becker 2014).

Second, economic trends could be responsible for growing SES achievement gaps. The level of economic development is rising in most countries that participate in international assessments, implying rising standards of living and a greater capacity for public and private investment in education and child well-being. However, it is not clear that a higher level of development leads to smaller SES achievement gaps; in fact, the reverse may be true. Comparing countries cross-sectionally at a single point in time (the 1970s), Heyneman and Loxley (1983) found that family SES was a more important predictor of student achievement in more developed countries, a correlation that still appears weakly present in PISA 2015 results (OECD 2016). When looking at changes over time, Baker and colleagues (2002) suggest that the importance of SES grew more in developing countries. These past findings imply that SES achievement gaps may increase more in lower-income than in higher-income countries, and gaps may increase more in countries experiencing more rapid growth in economic development. Another important economic trend, rising income inequality, was a suggested explanation for rising SES achievement gaps in both the United States and South Korea (Byun and Kim 2010; Reardon 2011b). Income inequality is increasing in many other countries as well, particularly in Europe and Asia (although income inequality appears to be decreasing in many Latin American and African countries) (OECD 2015; UNDP 2013). Although cross-sectional research shows that country income inequality is not strongly related to SES achievement gaps (Dupriez and Dumay 2006; Duru-Bellat and Suchaut 2005; Marks 2005), there is little published evidence on whether *changes* in income inequality within countries over time predict *changes* in SES achievement gaps. Countries with increasing income inequality might experience increasing SES achievement gaps due to increasing disparities in the material resources of low- and high-SES families, as well as possible corresponding increases in neighborhood segregation by income (Musterd et al. 2017; Reardon and Bischoff 2011).

Third, changing educational institutions could cause rising SES achievement gaps. A strong and consistent finding in cross-sectional comparative research is that countries with more rigid systems of curricular differentiation tend to have larger SES achievement gaps. In these studies, highly differentiated systems are those (primarily European) countries that select students at relatively young ages into academic and vocational tracks or schools (for a review, see van de Werfhorst and Mijs 2010). According to this work, we would expect countries that increase the rigidity of curricular differentiation or begin tracking at younger ages to experience increasing SES achievement gaps. However, it is not clear that such changes in tracking systems can explain increasing SES achievement gaps in many countries. Although Byun and Kim (2010) identify increased tracking as a potential explanation for increasing SES achievement gaps in South Korea, in most other countries participating in international assessments, reforms have been toward *de*-tracking, such as delaying the onset of tracking or enrolling a greater share of students in the academic track (Ariga et al. 2005; Benavot 1983; Manning and Pischke 2006). Moreover, results from two over-time studies comparing SES achievement gaps within countries across cohorts that were subject to different tracking policies provide inconclusive evidence. Van de Werfhorst (2018) finds that, among nine countries participating in both FIMS in 1964 and the Second International Mathematics Study (SIMS) in 1980, on average, the countries that implemented de-tracking reforms experienced declines in SES achievement gaps. In contrast, Brunello and Checchi (2007) find that SES origin gaps in literacy measured in adulthood are *larger* in cohorts educated after de-tracking reforms.

Although formal stratification by curricular tracks is declining globally, more informal stratification among schools by market forces may be increasing. School choice and privatization have increased in recent decades in many countries around the world (Musset 2012; UNESCO 2015). Research in several countries has found that rising school choice is associated with increasing SES segregation among schools (Bohlmark and Lindahl 2007; Byun, Kim, and Park 2012; Söderström and Uusitalo 2010; Valenzuela, Bellei, and Ríos 2014). However, other scholars argue that the relationship between school choice and segregation in certain countries may not be causal (Gorard 2014; Lindbom 2010). Nevertheless, if in most countries marketization of school attendance policies increases segregation, then such policies may cause students of different SES backgrounds to experience increasingly differentiated learning environments. Thus, countries with increasing school choice or private school enrollment are expected to experience increasing SES achievement gaps.

Finally, increasing SES achievement gaps could be due to increasing disparities in parental investments of time and money in children. Private household expenditures on children, such as childcare, school tuition, and private tutoring, appear to be growing dramatically and unequally between SES groups in a number of countries (Aurini, Davies, and Dierkes 2013; Kornrich, Gauthier, and Furstenberg 2011; Park et al. 2016). Likewise, parental time-use surveys across a range of countries show increasing time spent on childcare and increasing SES disparities in childcare time (Dotti Sani and Treas 2016; Gauthier, Smeeding, and Furstenberg 2004). Lareau's (2003) description of the "concerted cultivation" parenting style of the U.S. middle and upper class is echoed by a growing international qualitative literature on "intensive parenting" and the "parentocracy" (Brown 1990; Chang 2014; Dumont, Klinge, and Maaz 2019; Faircloth, Hoffman, and Layne 2013; Gomez Espino 2013; Hays 1996; Karsten 2015; Katartzi 2017; Liu 2016; Quirke 2006; Tan 2017). In the United States, these trends have been attributed to increasingly competitive college admissions (Alon 2009; Ramey and Ramey 2010; Schaub 2010). In other countries, competition may similarly increase after de-tracking reforms leave a growing share of students potentially eligible for university admission. Thus, a possible proxy for intensified parenting pressures is increasing higher-education aspirations; countries experiencing this trend should see larger increases in SES achievement gaps.

EMPIRICAL APPROACH

No study has yet taken advantage of the full history of international assessments to study global changes in SES inequality. A small number of economics studies combine modern and historical international assessments to study changes in the *level* of achievement over time (e.g., Altinok, Diebolt, and Demeulemeester 2014; Falch and Fischer 2012; Hanushek and Wößmann 2012), and two sociological studies use these data to compare changes in gender achievement gaps (Wiseman et al. 2009) and the effects of tracking reforms on SES achievement gaps (van de Werfhorst 2018). The strength of an over-time design is twofold. It allows investigation of the understudied question of *changes* in SES achievement gaps, rather than the size of gaps at only a single point in time. Moreover, in predicting which national characteristics and policies are associated with SES achievement gaps, an over-time design allows each country to "be its own control," ruling out observed and unobserved historical and cultural differences that often confound cross-sectional international comparisons. Such a design allows us to investigate, first, whether increasing SES achievement gaps are a global phenomenon; second, whether some countries have avoided the trend; and third, whether increasing SES achievement gaps can be explained by changing educational and social policies and conditions.

DATA

The data for this study are derived from 30 international large-scale assessments of math, science, and reading: FIMS 1964, SIMS 1980, FISS 1970, SISS 1984, the first international reading comprehension study (FIRCS 1970), the Reading Literacy Study (RLS 1991), and multiple years of TIMSS (1995 to 2015), PIRLS (2001 to 2011), and PISA (2000 to 2015). All studies were conducted by the International Association for the Evaluation of Educational Achievement (IEA) except PISA, which was conducted by the Organization for Economic Cooperation and Development (OECD). Together, the studies represent 109 countries and about 5.8 million students. All country samples are intended to be nationally representative, although full population coverage was not achieved in every country-study-year. Because population coverage information is inconsistently provided in early studies, I retain all available data in all analyses to avoid possibly biasing results by inappropriately excluding data.[1]

The unit of analysis in the current investigation is the country-study-achievement gap. For each country-study, I calculate SES achievement gaps in each subject for each available SES variable. After limiting the sample to countries that participated in at least two different studies in different years, the final sample is 5,541 country-study-gaps within 1,026 country-studies within 100 countries. Countries participating in international assessments tend to be high- or middle-income; the mean GDP per capita in 2015 for countries in the analytic sample was $30,366.69, compared to the world GDP per capita of $15,546.30.[2] A full list of included countries appears in Part A of the online supplement.

Variables

Achievement. Full descriptions of the math, science, and reading skills assessed in each study are available from the IEA's and OECD's official published reports. Different tests of the same subject have similarities, but only the scores from multiple years of the TIMSS, PIRLS, and PISA studies are strictly comparable. Because each test is on a different scale, in the main models that combine different studies, I standardize all scores to a mean of 0 and standard deviation of 1 within each country-study-year-subject before calculating each SES achievement gap. In standardizing scores within country-study-year-subject, I assume achievement matters as a positional good, consistent with previous research using achievement as a predictor of status attainment (e.g., Breen and Goldthorpe 2001; Mare 1980).[3] The validity of gap estimates based on standardized achievement then depends on the assumptions that all tests are interval scaled and that different tests rank students similarly.[4]

Subject. The main models pool math, science, and reading gaps and include dummy variables indicating whether a gap was estimated using math (35.1 percent of observations) or science (37.1 percent) achievement versus reading achievement (reference category; 27.9 percent).

SES. In each dataset, at least one of the following three measures of family socioeconomic status is available: parents' education, parents' occupation, and the number of books in the household. For parents' education and occupation, I use the higher of the two parents.[5] All SES variables are reported in ordered categories; the number of categories varies somewhat by study and by country. Parent education was generally six to eight categories, such as (1) none, (2) primary, (3) lower secondary, (4) vocational upper secondary, (5) academic upper secondary, (6) postsecondary vocational certificate, (7) associate's degree, and (8) bachelor's degree or more. Parent occupation was generally nine to ten categories corresponding to one-digit ISCO codes, reordered by average occupational status (Ganzeboom and Treiman 1996). In order of lowest to highest status, they are (1) laborers, (2) agricultural, (3) plant operators, (4) craft/trade, (5) service, (6) clerk, (7) business, (8) technician, (9) managerial, and (10) professional. Books in the household were usually reported in five to six categories, such as (1) 0 to 10 books, (2) 11 to 25 books, (3) 26 to 100 books, (4) 101 to 200 books, (5) 201 to 500 books, and (6) more than 500 books. In the final sample, 34.7 percent of country-study-gaps are based on parent education as the SES measure, 25.8 percent are based on parent occupation, and 39.5 percent are based on household books.

Although the percentile method I use to calculate SES achievement gaps (described in the Methods section) addresses some issues of comparability in the measurement of SES in different studies and countries, it may not fully account for differences in data quality. Thus, the main models include the following four variables to control for the quality of SES variables.

Parent versus student reporting. Most SES variables are student-reported, except for eight recent studies where they are parent-reported in some countries: PIRLS 2001, 2006, and 2011; TIMSS 2011 and 2015 4th grade; and PISA 2006, 2009, and 2012. Because students typically report SES less reliably than parents, gaps will tend to be attenuated due to measurement error when SES is reported by students. In addition to adjusting each SES achievement gap for estimated SES reliability (described in the Methods section), I also include a dummy variable indicating whether each gap was based on student-reported SES (81.4 percent of country-study-gaps) or parent-reported SES (18.6 percent). I interact this variable with gap type (parent education, parent occupation, or household books), because students' and parents' relative accuracy depends on the SES variable they are reporting (Jerrim and Micklewright 2014).

Number of categories. The percentile method used to calculate SES achievement gaps (described in the Methods section) requires only that categories be ordered, not an equal number of categories with consistent meanings or distributions across years or countries, so I retain the maximum possible SES categories for each country-study-gap.[6] However, gap estimates computed from a greater number of SES categories may tend to be larger due to the higher resolution of the data. Therefore, I include a control for the number of categories of the SES variable, ranging from 3 to 26, which I center at its median of seven categories.

20 percent or more students in the bottom SES category. The percentile method may not perform as well when more than 20 percent of observations are in the bottom or top SES category (Reardon 2011a). I include a dummy variable indicating whether 20 percent or more of students fall into the bottom category (14.8 percent of country-study gaps) versus less than 20 percent in the bottom category (85.2 percent).

20 percent or more students in the top SES category. I also include a dummy variable indicating whether 20 percent or more of students fall into the top category (38.7 percent of country-study gaps) versus less than 20 percent in the top category (61.3 percent).[7]

Cohort birth-year. I compute the mean birth year for each country-study from student reports either of birth year/month or age in years and months, relative to the known year and month of testing in each country. I use survey weights when calculating means. Birth year ranges from 1949.86 in the England FIMS 1964 sample to 2005.78 in the New Zealand TIMSS 2015 4th-grade sample. In the models, I set birth year to 0 in 1989, producing a range from −39.14 to 16.78.

Age at testing. Students are either in 4th grade/age 10 (FISS, FIRCS, SISS, RLS, TIMSS, and PIRLS), 8th grade/age 14 (FIMS, FISS, FIRCS, SIMS, SISS, RLS, and TIMSS), or age 15 (PISA).[8] The main models include dummy variables indicating age 10 (20.7 percent of observations) or age 15 (56.3 percent) versus age 14 (reference category; 23.0 percent).

The following time-varying country covariates are all measured at the country-study-year level. Unfortunately, due to low availability of comparable data across a large number of countries and long span of years, not all hypothesized causes of increasing achievement gaps can be included, and some covariates are relatively weak proxies of the intended concepts. Country covariates are drawn from a variety of sources, as noted. For country-level indicators not collected annually, I linearly interpolate missing years.

Level of school enrollment. Net proportion of the age-cohort enrolled in school in the year of testing comes from the World Bank. For 4th-grade testing cohorts, I use the proportion enrolled in primary school in the testing year; for 8th-grade and 15-year-old cohorts, I use the proportion enrolled in secondary school.

Proportion immigrant background. I compute the proportion of students reporting first- or second-generation immigrant status from the microdata.

GDP per capita. I obtain gross domestic product per capita converted to 2012 international dollars using purchasing power parity (PPP) rates from the World Bank. I average over the lifetime of each testing cohort from birth to test year.

Income inequality. Gini coefficients measured on a scale from 0 (perfect equality) to 1 (perfect inequality) come from the World Bank for less-developed countries and from the Luxembourg Income Study or OECD for wealthier countries.[9] I average over the lifetime of each testing cohort from birth to test year.

Age when tracking begins. Consistent with prior international comparative research, I define "tracking" as selection into overarching programs with academically or vocationally oriented curricula. I code the age when this selection occurred in a given country in each testing year, using a variety of sources: Brunello and Checchi (2007), UNESCO/International Bureau of Education (IBE) National Reports, the OECD's PISA reports, and the *International Encyclopedia of National Systems of Education* (Postlethwaite 1995). Age of track selection ranges from 10 to 16. Countries such as the United States that did not practice this type of tracking between 1964 and 2015 I code as age 16 in all years.

Proportion in private schooling. Students enrolled in privately-managed institutions (regardless of funding source) as a proportion of total enrollment comes from the World Bank for less-developed countries and from the OECD for wealthier countries. I average over all years when the testing cohort was school-aged, using primary-school private-enrollment figures in the years when the cohort was age 6 to 12, and secondary-school private-enrollment figures for age 13 to 15 (as applicable, up until the age at testing).

Proportion expecting higher education. Competition for higher-education admission is operationalized as the proportion of students expecting to attend higher education in the test year, estimated from the microdata. Higher education refers to any tertiary program (short or long cycle, i.e., ISCED 1997 5B or 5A) or more.[10]

METHODS

First, missing data for all student-level variables except achievement are imputed using multiple imputation by iterative chained equations, creating five imputed datasets for each country-study.[11, 12] Next, I draw 1,000 bootstrap samples from each of the five imputed datasets. In each sample, for each subject-SES variable combination, the SES achievement gap is computed as the gap in standardized achievement between the 90th and 10th percentiles of the country's distribution of that SES variable, following Reardon's (2011b) method for income achievement gaps. That is, within each country-study-year-subject, achievement Y is standardized to a mean of 0 and standard deviation of 1 (using student sample weights); for each SES variable within each country-study-year-subject, mean achievement $\overline{Y}$ and standard error are calculated for each SES category k (using student sample weights); each SES category is assigned a percentile θ_k corresponding to the middle percentile of the category within the country-study-year-specific SES distribution (using student sample weights); and a cubic function estimating the association between Y and θ is fit using weighted least squares (weighting by the inverse squared standard error of $\overline{Y}_k$).[13] This yields a fitted curve:

$$\hat{Y} = \hat{a} + \hat{b}(\theta) + \hat{c}(\theta^2) + \hat{d}(\theta^3) \tag{1}$$

Using this fitted curve, the estimated 90/10, 90/50, and 50/10 achievement gaps are as follows (Reardon 2011b):

$$\begin{aligned}\hat{\delta}^{90/10} &= \left[\hat{Y} \mid \theta = .9\right] - \left[\hat{Y} \mid \theta = .1\right] \\ &= .8\hat{b} + .8\hat{c} + .728\hat{d}\end{aligned} \tag{2}$$

$$\begin{aligned}\hat{\delta}^{90/50} &= \left[\hat{Y} \mid \theta = .9\right] - \left[\hat{Y} \mid \theta = .5\right] \\ &= .4\hat{b} + .56\hat{c} + .604\hat{d}\end{aligned} \tag{3}$$

$$\begin{aligned}\hat{\delta}^{50/10} &= \left[\hat{Y} \mid \theta = .5\right] - \left[\hat{Y} \mid \theta = .1\right] \\ &= .4\hat{b} + .24\hat{c} + .124\hat{d}\end{aligned} \tag{4}$$

As mentioned in the Data section, gaps tend to be attenuated in country-studies where SES is less reliably measured. Due to the standardization of achievement described earlier, gaps will also be attenuated in country-studies where achievement is less reliably measured. Therefore, gaps are adjusted according to each country's test reliability for each study, as published in the corresponding technical reports, as well as according to the estimated reliability of each SES report. For studies where both students and parents reported the same SES variable, reliability can be calculated from the microdata. These reliabilities are then applied to all other years.[14] Next, I use the 1,000 bootstrap sample gaps to estimate the error variances for each gap and error covariances among different gap types within each country-study-year. Finally, gaps are averaged across the five imputed datasets, and bootstrap error variances and covariances are adjusted for imputation variance, using formulas in Schomaker and Heumann's (2016) "MI Boot" method.[15] The plausible values of achievement included in some datasets (PISA, TIMSS, and PIRLS) can also be understood in a multiple imputation framework, and therefore are included in this procedure.[16]

The 90/10 percentile method compares students at the same relative position within the SES distribution of their respective country birth-cohorts, even as shifting SES distributions cause the absolute meanings of these positions to change. Thus, the analyses here assume that family SES is a positional rather than an absolute good in terms of the advantages it confers to children.[17] In the procedures described earlier, gaps are estimated separately for each SES variable in each country-study, rather than constructing an SES index, to avoid loss of information because not all SES variables are available in every dataset. The models below then pool gaps based on all three SES variables and test whether results differ depending on the SES variable used.[18]

Because each observation in the data is an achievement gap for a given test subject and SES variable (level 1), nested within study-years (level 2) and within countries (level 3), I use a three-level hierarchical growth curve model to estimate how gaps change across cohorts. Each study-year has up to nine different outcomes (gaps based on three SES variables × three subjects), each gap is measured with error, and errors are correlated across different gaps within a given country-study-year, so I implement this model using a multivariate variance-known model. The model was originally developed for use in meta-analysis with multiple outcomes, but it can be applied in the present setting where I am reanalyzing microdata and have multiple gaps in each study, along with estimated sampling error variances and covariances among gaps, computed via bootstrapping.[19] Following Kalaian and Raudenbush (1996), I fit a model that, instead of estimating a single constant, enters gap-type indicators (parent education, parent occupation, and household books) with no omitted category, meaning the model estimates a different intercept for each gap type. This multivariate specification allows more straightforward formal tests of whether the three different gap types exhibit similar cohort trends, both on average globally and within countries. The model is estimated as follows:

$$
\begin{aligned}
\hat{G}_{pjk} &= \boldsymbol{\delta}\left(\mathbf{T}_{pjk}\right) + \boldsymbol{\alpha}\left(\mathbf{S}_{pjk}\right) + \boldsymbol{\lambda}\left(\mathbf{A}_{jk}\right) \\
&\quad + \boldsymbol{\gamma}\left(\mathbf{T}_{pjk} C_{jk}\right) + \boldsymbol{u}_k\left(\mathbf{T}_{pjk}\right) \\
&\quad + \boldsymbol{w}_k\left(\mathbf{T}_{pjk} C_{jk}\right) + \boldsymbol{r}_{jk}\left(\mathbf{T}_{pjk}\right) + \epsilon_{pjk}, \\
&\begin{bmatrix} \epsilon_{1jk} \\ \vdots \\ \epsilon_{9jk} \end{bmatrix} \sim MVN\left[\mathbf{0}, \mathbf{V}_{jk}\right]; \\
&\begin{bmatrix} r_{1jk} \\ r_{2jk} \\ r_{3jk} \end{bmatrix} \sim MVN\left[\mathbf{0}, \Sigma\right]; \\
&\begin{bmatrix} u_{1k} \\ u_{2k} \\ u_{3k} \\ w_{1k} \\ w_{2k} \\ w_{3k} \end{bmatrix} \sim MVN\left[\mathbf{0}, \tau\right]
\end{aligned}
\qquad (5)
$$

where $\hat{G}_{pjk}$ is the pth observed gap (level 1) in study-year j (level 2) in country k (level 3), $\boldsymbol{\delta}$ is a vector of the true gaps conditional on all covariates in the model, $\mathbf{T}_{pjk}$ is a vector of dummy variables indicating gap type (parent education, parent occupation, or household books), $\boldsymbol{\alpha}$ is a vector of coefficents on control variables $\mathbf{S}_{pjk}$ for test subject (math, reading, or science) and SES variable quality measures for country-study-year-gap pjk, $\boldsymbol{\lambda}$ is a vector of coefficients on dummy variables $\mathbf{A}_{jk}$ indicating age at testing (10, 14, or 15) in country-study-year jk, $\boldsymbol{\gamma}$ is a vector of coefficients on interactions between gap type $\mathbf{T}_{pjk}$ and cohort birth-year C_{jk}, $\boldsymbol{u}_k$ is a vector of three country-level random intercepts for each gap type $\mathbf{T}_{pjk}$, $\boldsymbol{w}_k$ is a vector of three country-level random slopes on $\mathbf{T}_{pjk} C_{jk}$ interactions, $\boldsymbol{r}_{jk}$ is a vector of three study-year-level random intercepts for each gap type $\mathbf{T}_{pjk}$, ε_{pjk} is a level-1 error term, $\boldsymbol{\Sigma}$ and $\boldsymbol{\tau}$ are the within-country and between-country covariance matrices among the true gaps, and $\mathbf{V}_{jk}$ is the known sampling error variance-covariance matrix among the observed gap estimates $\hat{G}_{pjk}$ within study-year-country *jk*. Note that cohort birth-year and age at testing are not collinear because observations come from a wide range of years. Model estimates are reported with robust Huber-White standard errors.

The coefficients γ for the interactions between gap type and cohort birth-year represent the average trends in gaps over time across countries for each SES variable. If gaps are increasing on average, we would expect these coefficients to be positive. To further explore patterns in these trends, I estimate several additional models of a similar form. Model 2 estimates a single slope on cohort birth-year rather than different slopes for each gap type. Model 3 estimates interactions between cohort slope and world region. Model 4 estimates interactions between cohort slope and an indicator of country income-level in 1980 (above or below a GDP per capita of $6,000). Model 5 estimates quadratic growth curves by entering a squared cohort term. Models 6 and 7 predict 90/50 and 50/10 rather than 90/10 gaps. Finally, Model 8 attempts to explain changes in gaps; I remove the cohort terms and enter a series of study-year dummies and time-varying country covariates (mean-centered within countries) at level 2, and country mean covariates at level 3. Thus, Model 8 can be interpreted similarly to a model with country and study-year fixed effects. The coefficients for time-varying country covariates represent the associations between changes in covariates and changes in gaps within countries over time, after accounting for secular trends across study-years.

RESULTS

Figure 1 shows an example of an estimated trend in the 90/10 SES achievement gap for one country (the United States) for one SES variable (parent education). Each data point is the estimated achievement gap between students at the 90th and 10th percentiles of parent education in the U.S. subsample of a particular international assessment. The gaps are plotted against the birth year of sampled students, which runs from approximately 1950, corresponding to 14-year-old students tested in FIMS 1964, to approximately 2001, corresponding to 14-year-olds tested in TIMSS 2015. A quadratic fit line is estimated using weighted least squares to describe the trend in gaps across birth cohorts. The parent education achievement gap declined slightly in the United States over the past 50 years, from about 1.2 SDs of achievement in the 1950 birth cohort to about 1.1 SDs in the 2001 cohort, a decline that is not statistically significant. This result is consistent with Reardon's (2011b) study, which, in contrast with a substantial increase in the U.S. achievement gap based on *income*, did not find any significant change in the achievement gap based on *parent education*. This result is also similar to Hanushek and colleagues' (2019) finding of no change in U.S. achievement gaps between the 90th and 10th percentiles of an index of SES (including parent education and household possessions) and to Broer and colleagues' (forthcoming) finding of a small decline in U.S. SES achievement gaps across recent waves of TIMSS. However, the slight decline in Figure 1 is less pronounced than the more marked decline in U.S. SES achievement gaps reported for recent waves of PISA (OECD 2018) (I discuss possible methodological reasons for this discrepancy below).

My estimates for U.S. trends for achievement gaps based on the other two SES variables, parent occupation and household books (not shown), are broadly similar to the trend in the parent education achievement gap. All gap types are relatively stable over the full 51-year period, although the parent

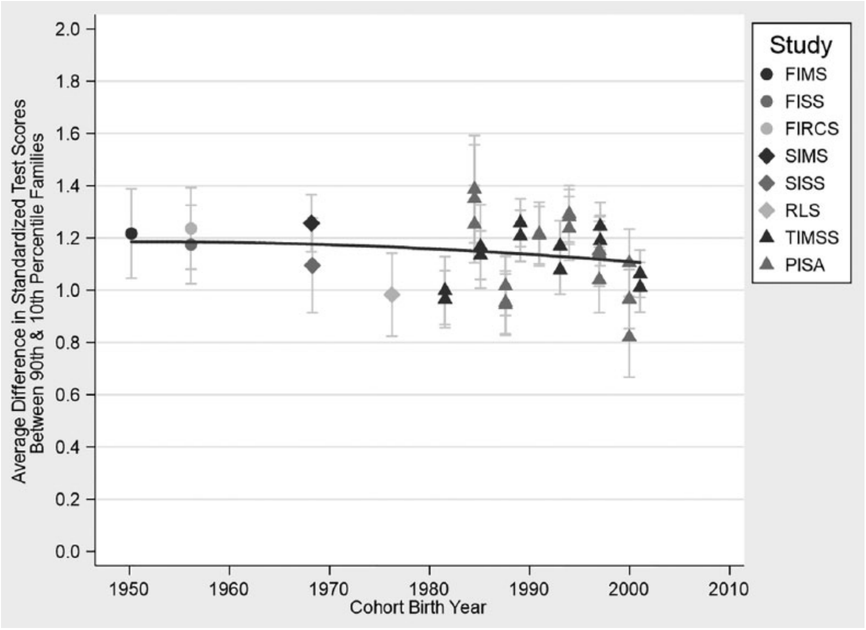


Figure 1. Trend in 90/10 Parent Education Achievement Gaps, United States, 1950 to 2001 Cohorts
Note: Gaps and quadratic fit line adjusted for age of testing and subject. Gray brackets are 95 percent confidence intervals.

occupation gap shows a slight decline like the parent education gap, whereas the household books gap shows a slight increase. In the most recent years of data, books gaps are substantially larger than parent education and occupation gaps. The different trend for achievement gaps based on books may imply that household books are gaining salience relative to parent education and occupation in predicting children's academic achievement. However, the discrepancy also likely reflects differences in data quality. In later years, large proportions of U.S. students fall into the top categories of parent education and occupation, making it difficult to precisely estimate achievement at the 90th percentile of SES. This issue affects the United States and several other high-income countries, and it appears to cause achievement gaps based on parent education and occupation, but not books, to be underestimated in later years (discussed further below).

Table 1 puts the U.S. results into global context by reporting on hierarchical growth curve models summarizing average global trends in SES achievement gaps as well as cross-national variation across all available countries. Models pool SES achievement gaps across all test subjects and SES variables and predict the size of each gap based on the cohort birth-year variable and controls. Conceptually, by pooling all gap types to estimate trends, I assume that, although different gap types do not have identical meanings, any observed *trend* in gaps across cohorts is driven by the same underlying process. Methodologically, the multivariate variance-known model allows a formal test of the assumption that trends in gaps do not significantly differ depending on the SES variable used. Practically, pooling data prevents loss of information, because not all gap types are observed in all study-years (the variance-known model can also accommodate this unbalanced data structure).

Model 1 estimates a different cohort slope for each gap type (parent education, parent occupation, and household books) using interactions between cohort birth-year and gap-type indicators. As described in the Methods section, the multivariate variance-known

Table 1. Unstandardized Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th Percentiles of SES

	(1)		(2)	
	3 Cohort Slopes		1 Cohort Slope	
	Coef.	(se)	Coef.	(se)
Parent education gaps intercept	1.032***	(.030)	1.039***	(.030)
Parent occupation gaps intercept	.958***	(.030)	.964***	(.030)
Household books gaps intercept	1.299***	(.041)	1.294***	(.041)
Level 1 – Gaps				
Subject (ref. = reading)				
Math	.020**	(.007)	.020**	(.007)
Science	.034***	(.005)	.034***	(.005)
SES variable quality measures				
Parent-reported × Parent education	.132***	(.030)	.112***	(.031)
Parent-reported × Parent occupation	.075**	(.025)	.073**	(.024)
Parent-reported × Books	–.039	(.029)	–.017	(.026)
Number of categories (centered at 7)	.003	(.003)	.002	(.003)
≥ 20% in bottom category	–.065**	(.021)	–.063**	(.021)
≥ 20% in top category	–.135***	(.013)	–.146***	(.013)
Level 2 – Study-years				
Age at testing (ref. = 14)				
Age 10 at testing	–.170***	(.024)	–.168***	(.024)
Age 15 at testing	–.024	(.020)	–.023	(.020)
Cohort birth-year × Parent education	.007***	(.001)		
Cohort birth-year × Parent occupation	.007***	(.001)		
Cohort birth-year × Books	.008***	(.001)		
Cohort birth-year			.007***	(.001)
Random effects				
Level 2 – Residual variance between studies in . . .				
Parent education intercepts	.03736		.03831	
Parent occupation intercepts	.02322		.02284	
Books intercepts	.03698		.03823	
Level 3 – Residual variance between countries in . . .				
Parent education intercepts	.05426		.05362	
Parent occupation intercepts	.05227		.05330	
Books intercepts	.11590		.12149	
Parent education cohort slopes	.00004			
Parent occupation cohort slopes	.00003			
Books cohort slopes	.00007			
Cohort slopes			.00003	
N (Level 1 – gaps)	5,541		5,541	
N (Level 2 – study-years)	1,026		1,026	
N (Level 3 – countries)	100		100	

*$p < .05$; **$p < .01$; ***$p < .001$ (two-tailed tests).

models estimate a different intercept for each of the three gap types. Because cohort birth-year is centered at 1989, the parent education gaps intercept of 1.032 represents the average 90/10 parent education gap in reading for the 1989 birth cohort at age 14 (i.e., tested in 2003) when all SES variable quality measures are held at their reference categories. On average, parent occupation gaps tend to be slightly smaller than parent education gaps and books gaps substantially larger.

Turning to the control variables, on average, math achievement gaps are significantly larger than reading gaps, which is consistent with prior U.S. research (Reardon 2011b). Science gaps are also larger than reading gaps. However, supplemental analyses show that *trends* in gaps across cohorts are similar for all three test subjects; thus, the main models pool gaps for all subjects.[20] Even after the reliability adjustment described in the Methods section, parent education and occupation achievement gaps tend to be larger when estimated from parent-reported SES data. The difference is especially pronounced for parent education, consistent with Jerrim and Micklewright (2014), who note greater consistency in students' and parents' reports of occupation than education. There is no significant difference in the size of books gaps depending on whether books are reported by parents or students, after the reliability adjustment. Additional analyses show that trends in gaps across cohorts are similar for gaps based only on student- or on parent-reported SES.[21]

As expected, the number of categories of the SES variable is positively associated with the size of gaps, although the association is small and not significantly different from 0. More than 20 percent of students falling into the bottom or top category of the SES variable is associated with smaller estimated gaps. This suggests the 90/10 SES achievement gap method may systematically underestimate achievement gaps when the 90th or 10th percentile is extrapolated outside the SES data. For this reason, parent education and occupation gap increases may be more conservatively estimated in the United States and other wealthy countries that have large proportions of students in the top categories in later years. Nevertheless, when SES variable quality measures are omitted, results are similar.[22] SES achievement gaps tend to be smaller when they are estimated from students tested at age 10 than at age 14, and gaps estimated at age 15 are slightly smaller but not significantly different from those at age 14. Trends in gaps across cohorts are similar when age groups are analyzed separately.[23]

Most interesting are the coefficients for interactions between cohort birth-year and gap-type dummies, as they measure the average annual change in achievement gaps across all sample countries for each of the three SES variables. All three coefficients are positive and significant, indicating that on average across all sample countries, all three types of SES achievement gaps increased. Net of controls, 90/10 parent education and occupation gaps both increased at a rate of .007 SD of achievement per year, and 90/10 books gaps increased .008 SD per year. Although these annual increases are small, they correspond to quite large total gap increases across the full time-span of study-years: about .4 SD of achievement for all three gap types. As mentioned earlier, the model specification allows a formal test of whether gap trends differ depending on the SES variable used. A Wald test of the joint null hypothesis that all three coefficients are equal cannot be rejected ($p > .5$).

In addition to the three average cross-national trends, the model also provides evidence on whether the three gap types exhibit similar trends within countries—that is, whether countries with large increases in achievement gaps based on one SES variable also tend to have large increases in achievement gaps based on the other two SES variables. The correlation between country-specific random effects on cohort slopes for parent education and occupation gaps is .58, for parent education and books gaps it is .51, and for parent occupation and books gaps it is .90. Based on these moderate-to-strong positive correlations and the joint hypothesis tests, I conclude that, although achievement gaps by each SES variable do not have identical meanings, the *trends* in gaps across cohorts appear similar regardless of the SES variable used, suggesting they may be driven by a single underlying process. To the extent that there are small differences in trends by SES variable, it is not possible with the data available to adjudicate conclusively between substantive versus data quality/availability explanations. Therefore, in Model 2 (and all subsequent models), I estimate a single cohort birth-year coefficient, pooling across all gap types to summarize the general trend in SES achievement gaps.[24] In Model 2, this pooled cohort coefficient is estimated at .007. The coefficient estimates for all control variables are similar to Model 1.

The lower "random effects" panel of Table 1 estimates the cross-study and cross-national variability of results. Of particular interest are the cross-national variances of cohort slopes, as these summarize the degree to which countries deviate from the average global trend of increasing gaps described earlier. Chi-squared tests show that the cross-national variances of the cohort slopes for all three SES variables in Model 1, as well as the pooled cohort slopes in Model 2, are all significantly different from 0 ($p < .001$), meaning there is substantial cross-national variation in trends. Assuming (as the hierarchical growth curve model does) a normal distribution of country-specific cohort slope residuals, the estimated cohort slope variances imply that 95 percent of countries' parent education cohort slopes fall within the range (–.006, .019). The 95 percent plausible value ranges for parent occupation, books, and pooled cohort slopes are (–.003, .017), (–.009, .025), and (–.004, .019), respectively. Also implied is that the share of countries with trends greater than 0 is approximately 84 percent for parent education, 92 percent for parent occupation, 82 percent for books, and 90 percent for pooled gaps. Thus, although a large majority of countries experience increasing SES achievement gaps, the size of these increases varies widely, and gaps decline in about 8 to 18 percent of countries.

The models in Table 2 test for systematic patterns in the types of countries that experience larger increases in gaps by interacting cohort birth-year with world region and country income-level (see Table A1 in the online supplement for a list of countries by region and income-level). In Model 3, the main effect of cohort birth-year indicates that the average annual increase in gaps in Western countries (the reference category) is .008 SDs. The gap increase for African countries is larger but not significant; this trend is imprecisely estimated due to a small sample of African countries. Gap trends in Asian, Middle Eastern, and Eastern European countries are similar to trends in Western countries. The only region with a significantly different gap trend from the West is Latin America and the Caribbean, where gaps remained flat or even slightly declined over time. Model 4 interacts cohort birth-year with a dummy variable indicating that a country's GDP per capita in 1980 was below $6,000. Hereafter, these countries are referred to as "low-income" for brevity, recognizing that there are few truly low-income countries in the dataset (most are high- or middle-income). The interaction should be positive, as prior research suggests countries at lower levels of economic development experienced larger increases in SES achievement gaps between the 1970s and 1990s (Baker et al. 2002). The coefficient is indeed positive but not significant.

However, Baker and colleagues' (2002) findings pertain to cohorts born between approximately 1960 and 1980, a shorter time frame than in the present study. Model 5 includes a squared cohort birth-year term and interaction with country income, to estimate curvilinear trends and allow trends to differ by country income. The main effect for the squared term is positive but not significant, indicating that the gap trend for high-income countries curves very slightly upward. The interaction between the squared cohort term and the low-income country dummy is negative and significant, and the resulting point estimate is negative, indicating the gap trend for low-income countries curves downward.

Table 2. Unstandardized Coefficients from Hierarchical Growth Curve Models Predicting Achievement Gaps between 90th, 50th, and 10th Percentiles of SES, Adding Interactions by Country Region and Income-Level

	(3)		(4)		(5)		(6)		(7)	
	Region Interactions		Income Interaction		Quadratic		90/50 Gap		50/10 Gap	
	Coef.	(se)	Coef.	(se)	Coef.	(se)	Coef.	(se)	Coef.	(se)
Parent education gaps intercept	1.126***	(.027)	1.074***	(.035)	1.066***	(.037)	.548***	(.019)	.514***	(.022)
Parent occupation gaps intercept	1.111***	(.029)	1.054***	(.038)	1.046***	(.042)	.564***	(.020)	.481***	(.022)
Household books gaps intercept	1.563***	(.038)	1.434***	(.047)	1.426***	(.049)	.622***	(.022)	.806***	(.029)
Level 1 – Gaps										
Subject controls (ref. = reading)	yes		yes		yes		yes		yes	
SES variable quality measures	yes		yes		yes		yes		yes	
Level 2 – Study-years										
Age at testing controls (ref. = 14)	yes		yes		yes		yes		yes	
Cohort birth-year	.008***	(.001)	.007***	(.001)	.007***	(.001)	.001	(.001)	.006***	(.001)
Cohort birth-year^2					.00005	(.00005)				
Level 3 – Countries										
Region (ref. = Western) × Intercept interactions	yes									
Mid/low-income country × Intercept interactions			yes		yes		yes		yes	
Region (ref. = Western) × Cohort interactions										
Sub-Saharan Africa × Cohort	.004	(.004)								
East Asia & Pacific × Cohort	–.001	(.002)								
Middle East & N. Africa × Cohort	–.001	(.004)								
E. Europe & CIS × Cohort	–.001	(.002)								
Latin America & Caribbean × Cohort	–.009*	(.004)								
Mid/low-income country × Cohort			.002	(.002)	.001	(.002)	.003*	(.001)	–.001	(.001)
Mid/low-income country × Cohort^2					–.00022*	(.00010)				
N (Level 1 – gaps)	5,541		5,541		5,541		5,541		5,541	
N (Level 2 – study-years)	1,026		1,026		1,026		1,026		1,026	
N (Level 3 – countries)	100		100		100		100		100	

Note: "Middle/low-income" countries had GDPs per capita of less than $6,000 in 1980 (the reference category is high-income countries; see Table A1 in the online supplement for coding). Specification of control variables (subject, SES variable quality measures, and age) is identical to Models 1 and 2; coefficents are omitted due to space constraints. See Part L of the online supplement for full results.
$^*p < .05$; $^{**}p < .01$; $^{***}p < .001$ (two-tailed tests).

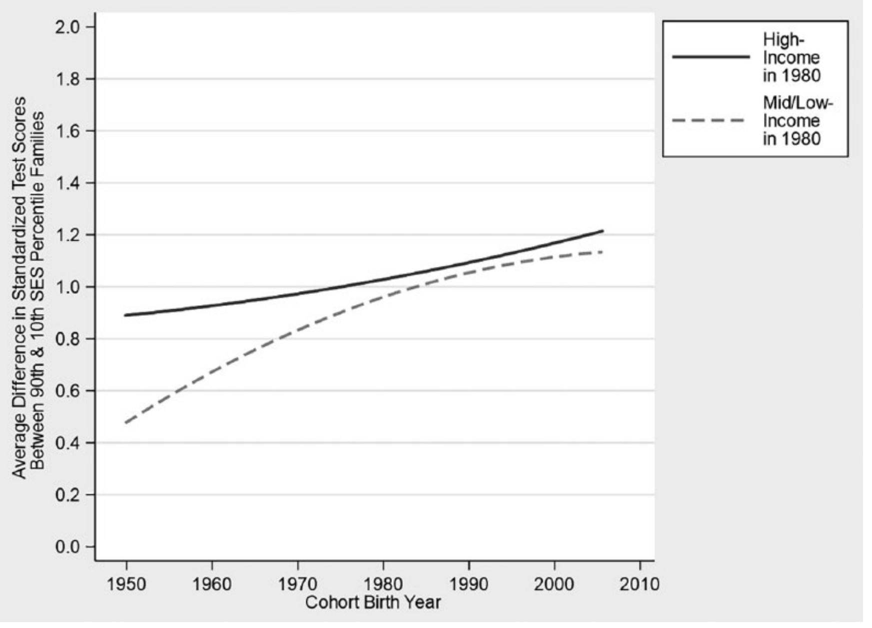


Figure 2. Estimated Quadratic Trends in 90/10 SES Achievement Gaps, by Country Income-Level, 1950 to 2005 Birth Cohorts
Note: "High-income" countries had GDPs per capita of at least $6,000 in 1980 (see Table A1 in the online supplement for coding). Trend lines are estimates from Model 5 (Table 2). Fixed values for control variables: SES = parent education, subject = math, all others = 0 or reference category.

Figure 2 illustrates Model 5. High-income countries' SES achievement gaps increased at a steady and nearly linear rate between the 1950 and 2005 birth cohorts, whereas low-income countries' gaps increased rapidly in early years and then at a slower rate. Thus, in early years, low-income countries experienced greater increases in gaps than did high-income countries, consistent with earlier results from Baker and colleagues (2002). Yet in recent years, this pattern reversed, and high-income countries experienced slightly greater increases in gaps than did low-income countries. Additional analyses show that the flattening trend in low-income countries is largely driven by Latin America and Caribbean countries, where gaps declined particularly in recent years.

To examine cross-national variation in gap trends in more detail, Figure 3 plots estimated quadratic trends for 24 countries. I selected countries with the most available data points over the longest time span that also provide some variation in region and country income-level. The trend lines are derived from coefficient estimates and country-specific shrunken empirical Bayes residuals from Model 5. Thus, they draw on data from all available gap types in each country and across the entire international sample to obtain the best estimate of the true trend in the SES achievement gap for each country. The number of study-years available for each country (i.e., the level-2 sample size) is in parentheses. The figure shows that, among countries with many years of data, most—although not all—experienced increases in SES gaps. This is consistent with results for the full sample of countries, as observed in the random slope estimates in Table 1. The countries without increasing gaps (e.g., England, Finland, Israel, Japan, and Scotland) tend to be high-income and, like the United States, already had large gaps in early cohorts and have large proportions of students in the top categories of parent education and occupation in later

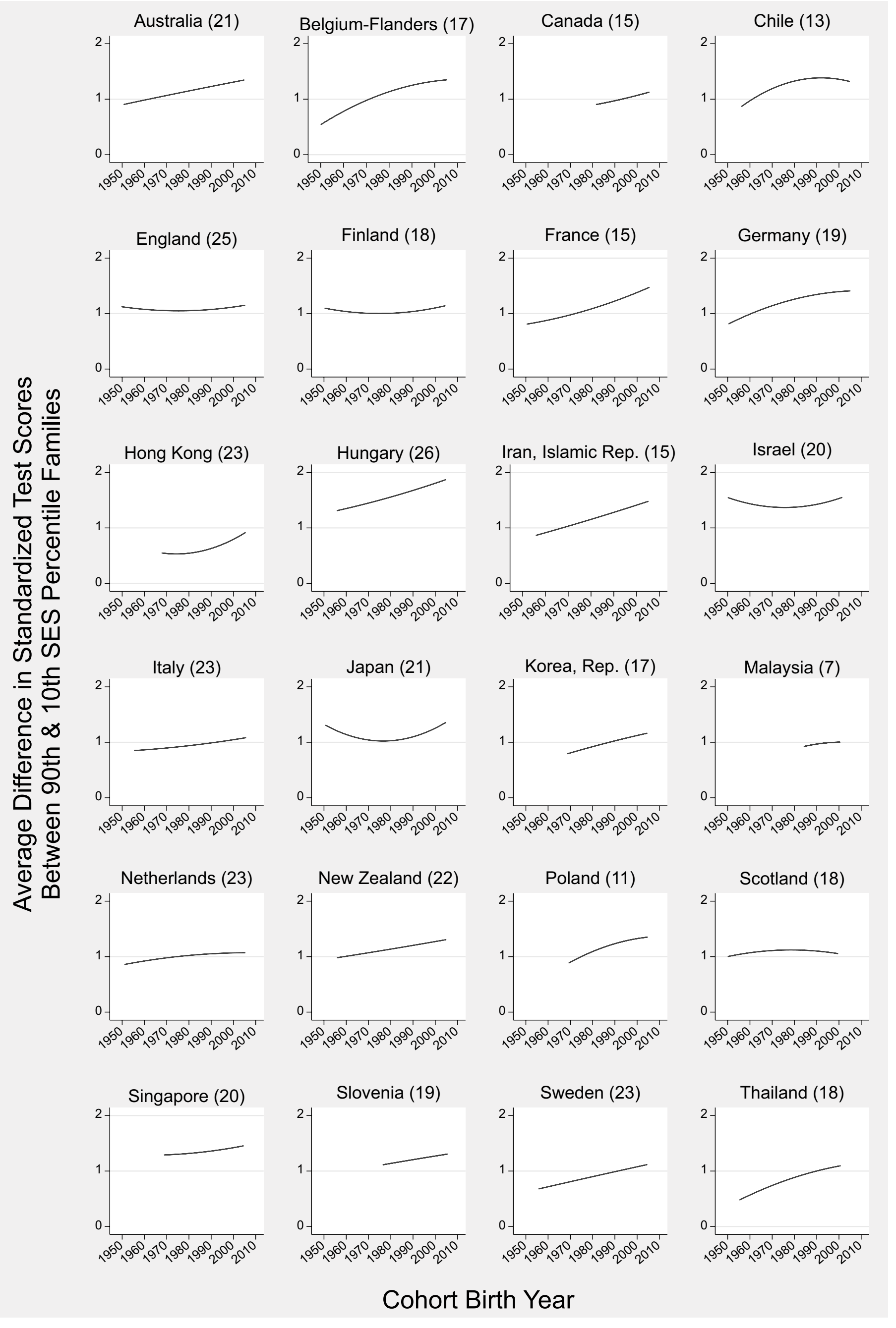


Figure 3. Estimated Quadratic Trends in 90/10 SES Achievement Gaps, Selected Countries (Number of Study-Years in Parentheses)

Note: Countries are sorted alphabetically. Trend lines are derived from shrunken empirical Bayes estimates from Model 5 (Table 2). Fixed values for control variables: SES = parent education, subject = math, all others = 0 or reference category.

years. Note, however, that several other high-income countries with large shares of students in the top education and occupation categories nevertheless experienced sizable increases in gaps (e.g., Australia, Canada, Norway, and Sweden).

Figure 3 also shows the estimated trends for several countries studied in prior research. The trend for South Korea is positive, consistent with the increasing SES achievement gap observed by Byun and Kim (2010). In contrast, the trend for Malaysia is nearly flat, inconsistent with the increasing gap described by Saw (2016). Both prior studies, however, use somewhat different data and measures than the current study. Byun and Kim (2010) use three waves of TIMSS and an SES index composed of parent education, household books, and other household possessions. Saw (2016) uses four waves of TIMSS and a dichotomous measure of parent education. The discrepancy in the Malaysian findings appears to be primarily due not to the difference in SES measures but the inclusion of more recent data, as the Malaysian parent education achievement gap declined markedly in TIMSS 2015. The Malaysian 90/10 gap trend estimated using data only up to 2011 is positive, consistent with Saw (2016). Consistent with Broer and colleagues' (forthcoming) report on TIMSS trends, I find increasing SES achievement gaps for Hungary, Iran, New Zealand, and Singapore. In line with the OECD's (2018) report on PISA trends, I find an increasing SES achievement gap for Singapore and a decreasing gap in recent years for Chile.

However, many of the other trend estimates in Figure 3 are inconsistent with the PISA and TIMSS reports. There appear to be several reasons for this, apart from the inclusion of more study-years in the current analysis. First, both PISA and TIMSS reports examine differences only between gaps in 2015 and one early assessment year, rather than estimating linear or curved trends using all study waves. Additionally, in the TIMSS report, Broer and colleagues (forthcoming) measure achievement in the original TIMSS scale rather than standardizing within waves. This produces declining achievement gaps in some countries where score variance decreases substantially, even as the relative relationship between SES and achievement grows stronger.[25] In contrast, the OECD (2018) measures achievement gaps as the R^2 of a model predicting achievement from SES and so captures only changes in the strength of the association. Finally, both reports measure SES using an index composed of parent education, household possessions (including books), and—for PISA only—parent occupation, whereas the trends in Figure 3 are estimated by pooling parent education, occupation, and books gaps in a multivariate variance-known model. This difference in the treatment of SES does not appear to contribute much to disparities between the current analysis and the TIMSS report, where parent education and books are weighted equally in the SES index, but it does produce different results in the PISA report, where books receive less weight than parent education. Supplemental analyses show that parent education may be poorly measured in later years of PISA.[26]

Models 6 and 7 examine changes in SES achievement gaps at the top and bottom of the SES distribution by predicting 90/50 and 50/10 gaps. Although 90/50 gaps increased very little in high-income countries, they increased significantly more in low-income countries. 50/10 gaps increased substantially in both high- and low-income countries, and the trends are not significantly different. Thus, the overall increase in the 90/10 gap in high-income countries is primarily concentrated at the bottom of the SES distribution; in other words, it is driven by the achievement of middle- and high-SES students pulling away from that of low-SES students.[27] Note, however, that the 90/50 gap in high-income countries is still substantial in recent years; in the 2005 cohort, the 90/50 parent education gap is estimated at about .56 SDs, or only slightly less than half of the overall 90/10 parent education gap of 1.19 SDs. In low-income countries, in contrast, the overall increase in the 90/10 SES achievement gap is more evenly spread across the entire SES distribution, with high-SES and low-SES students' achievement pulling away from middle-SES students at approximately equal rates.

Finally, Model 8 in Table 3 attempts to explain cross-national and over-time variability in 90/10 SES achievement gaps using country covariates. Model 2B shows that the average annual increase in SES achievement gaps is nearly identical in the analytic sample of countries with available covariate data to the full sample of countries in Model 2. In Model 8, the main predictors of interest are the time-varying covariates at level 2 (the study-year level), but the model also compares these over-time results to traditional cross-sectional associations by reporting the associations between country mean covariates and the size of gaps in the 1989 birth cohort conditional on controls, that is, the intercept of the model (cross-sectional associations are displayed in the lower "Level 3" portion of the table). All level-3 country mean covariates are grand-mean centered across the entire international sample. These country-level results mostly replicate the findings of previous cross-sectional comparative literature. Focusing on the coefficients that are significantly different from 0, the countries with the largest SES achievement gaps in the 1989 cohort tend to be those with a greater proportion of youth enrolled in school, higher GDPs per capita, and earlier tracking. Higher income inequality is also associated with larger SES achievement gaps, but this association is only marginally significant ($p < .1$). Although intuitively one might expect a strong association between income inequality and SES achievement gaps, this result is consistent with weak relationships found in prior cross-sectional research (Dupriez and Dumay 2006; Duru-Bellat and Suchaut 2005; Marks 2005). Keep in mind, however, that with a country-level sample size of only 78, the level-3 portion of the model may be overfit.

The level-2 within-country, over-time portion of the model improves on cross-sectional research and takes advantage of the unique long time-series dataset by examining associations between changes in country characteristics and changes in gaps. Time-varying country covariates are entered at level 2 (the study-year level) and are mean-centered within countries, meaning their coefficients can be interpreted similarly to a model with country fixed effects. The first two time-varying covariates pertain to the increasing diversity of the population of students included in international assessments. The coefficient for the proportion of the relevant age-cohort enrolled in school is positive, as expected, indicating that countries with increasing school access tend to experience increasing SES achievement gaps. This is not surprising, as increasing school access corresponds to increasing population coverage of international assessments, which sample only students enrolled in school. Controlling for other covariates, when the enrollment share increases by 10 percentage points, the SES achievement gap is expected to increase by .04 SD ($p < .001$). Also as expected, an increasing share of immigrant students is associated with increasing achievement gaps, although this relationship is not significant.

The next two covariates pertain to economic changes. As expected based on Baker and colleagues' (2002) research, countries with increasing GDPs per capita tend to experience increasing SES achievement gaps, although this association is not significant. Contrary to expectation, high-income countries with the largest increases in income inequality, all else equal, experience *declining* SES achievement gaps. Controlling for other covariates, an increase of .1 in the Gini coefficient is associated with a decrease in the SES achievement gap of .19 SD ($p < .05$). However, the opposite is true for low-income countries, whose income inequality coefficient is significantly more positive than that of high-income countries ($p < .05$). The point estimate for the income inequality coefficient for low-income countries is positive, indicating that among these countries, those with the largest increases in income inequality tend to experience increasing gaps, as expected.

Table 3. Unstandardized Coefficients from Hierarchical Models Predicting Achievement Gaps between 90th and 10th Percentiles of SES, Adding Country Covariates

	(2B)		(8)	
	Analytic Sample		Country Covariates	
	Coef.	(se)	Coef.	(se)
Parent education gaps intercept	1.066***	(.032)	.970***	(.046)
Parent occupation gaps intercept	.994***	(.031)	.969***	(.049)
Household books gaps intercept	1.338***	(.045)	1.405***	(.057)
Level 1 – Gaps				
Subject controls (ref. = reading)	yes		yes	
SES variable quality measures	yes		yes	
Level 2 – Study-years				
Age at texting controls (ref. = 14)	yes			
Cohort birth year	.007***	(.001)		
Study fixed effects (ref. = TIMSS 2003 grade 8)			yes	
School enrollment (proportion)			.486***	(.107)
Immigrant background (proportion)			.226	(.250)
GDP per capita (logged)			.055	(.059)
Income inequality (Gini)			–1.913*	(.887)
Mid/low-income country × Income inequality			2.539*	(1.129)
Age when tracking begins			–.037*	(.016)
Private school enrollment (proportion)			.240	(.249)
Expecting higher education (proportion)			–.029	(.094)
Level 3 – Countries				
Mid/low-income country × Intercept interactions			yes	
Mean school enrollment			.640*	(.317)
Mean proportion immigrant background			.134	(.261)
Mean GDP per capita (logged)			.142***	(.041)
Mean income inequality			.605	(.317)
Mean age when tracking begins			–.035**	(.011)
Mean private school enrollment			.043	(.099)
Mean proportion expecting higher education			–.255	(.169)
N (Level 1 – gaps)	4,604		4,604	
N (Level 2 – study-years)	855		855	
N (Level 3 – countries)	78		78	

Note: "Middle/low-income" countries had GDPs per capita of less than $6,000 in 1980 (the reference category is high-income countries; see Table A1 in the online supplement for coding). All level-2 time-varying country covariates are mean-centered within countries, meaning results can be interpreted similarly to a model with country fixed effects (as well as study-year fixed effects, included at level 2). All level-3 country mean covariates are grand-mean centered across the entire international sample. Coefficents for control variables (subject, SES variable quality measures, age, and study fixed effects) are omitted due to space constraints; see Part M of the online supplement for full results.
$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$ (two-tailed tests).

However, a joint hypothesis test shows that the positive income inequality coefficient for low-income countries is not significantly different from 0.

The following two covariates measure changes in educational institutions. As expected, increasing the age when tracking begins is associated with declining SES achievement gaps, consistent with cross-sectional results and with recent over-time findings by van de Werfhorst (2018) for a shorter period of time and a smaller number of countries. Controlling for other covariates, a one-year increase in the age when tracking begins is associated with nearly a .04 SD decline in the SES achievement gap ($p < .05$). As expected, an increasing share of students enrolled in private schools is associated with increasing SES achievement gaps, although this association is not significant. The last covariate pertains to increasing competition for higher-education admissions, measured as an increasing share of students expecting to attend higher education. Unexpectedly, increasing educational aspirations are associated with slightly *declining* SES achievement gaps, although this association is small and not significantly different from 0.

We can examine to what extent the country covariates explain variance in the size of SES achievement gaps over time by comparing the level-2 residual variances for this full model to a reduced model that includes study fixed effects and controls but no country covariates (not shown). Compared to a reduced model, the country covariates in Model 8 explain an additional 3 percent, 7 percent, and 8 percent of the within-country, between-study-year variance in SES achievement gaps based on parent education, occupation, and books, respectively. These percentages are small but indicate the covariates have some explanatory power, net of the secular time trend in gaps captured by the study fixed effects. That the variance explained is not greater is an indication that the time trend is very strong (the study fixed effects explain 15 to 35 percent of within-country variance in gaps), but also that some important causes of achievement gaps may be omitted from the model, the covariates included may be poorly measured, or there is cross-national heterogeneity in the causes of increasing gaps.[28]

I performed a number of robustness checks, which are reported in the online supplement. The results of these analyses show that global increases in SES gaps do not appear to be an artifact of increasing levels or narrowing variability of achievement or of SES, nor an artifact of declining measurement error in achievement or in SES.[29]

DISCUSSION

This study found strong and robust evidence of increasing SES achievement gaps over the past 50 years across the majority of countries examined. Gaps are consistently increasing for a variety of different model specifications and for three different measures of SES. Gaps based on parent education increased by about 50 percent, gaps based on parent occupation by about 55 percent, and gaps based on household books by about 40 percent. Results for all three variables are broadly consistent, lending support to the assumption that, even though different gap types do not have identical meanings and are generated through somewhat different processes, *trends* in gaps across cohorts appear to be driven by the same underlying process: a strengthening association between students' academic achievement and their family SES, broadly defined. This result appears to hold not only for two traditional measures of family SES—parent education and occupation—but also for the less traditional measure, household books. Although one might expect books would become a weaker proxy for SES in recent years if high-SES families can increasingly afford to substitute digital devices, supplemental analyses show that student-level correlations between books and both other SES variables are growing stronger over time.[30] Moreover, the results in Model 1 show that achievement gaps based on household books increased slightly more than those based on parent education or occupation in absolute terms. This small difference is driven mainly by high-income countries and may indicate that, with widespread access to digital devices, owning physical books increasingly captures not only economic but also cultural capital.

SES achievement gaps have increased in most countries, but the size of the increase varies widely, and in a substantial number of countries, gaps are stable or declining. The countries with the largest increases in gaps are a diverse set, including high-income countries such as Belgium (both the Flemish and French communities), Luxembourg, Ireland, and Norway, as well as middle- and low-income countries such as Poland, Hungary, Iran, and Thailand. The strongest and most significant predictor of increasing SES achievement gaps is increasing school enrollment, and indeed, several of these countries dramatically expanded enrollment. For example, Luxembourg and Ireland both increased secondary-school enrollment by over 20 percentage points over the years they participated in international assessments, and Thailand increased secondary enrollment by nearly 70 percentage points. The results for enrollment are consistent with Baker and colleagues' (2002) argument that growing SES achievement gaps are driven in part by expanding access and an increasingly diverse population of students included in schools and in international assessments. Also supporting this idea, in most countries, gaps are increasing more between the middle and bottom of the SES distribution (the 50/10 gap) than between the middle and top (the 90/50 gap). Thus, expanding access to school may not directly increase inequality but rather *reveal* inequality that was previously hidden outside the school system. However, gaps also increased in many countries with consistently high enrollment levels, such as Norway and Sweden, suggesting that increasing SES achievement gaps are driven by more than simply expanded population coverage of international assessments.

The countries with stable or declining gaps include several Latin American and Caribbean countries (e.g., Mexico, Brazil, and Trinidad and Tobago), as well as some wealthy countries, such as the United States, England, Finland, Israel, and Japan. The countries with declining gaps appear to drive the results for income inequality in the multivariate models. In low-income countries, increasing income inequality is positively associated with increasing SES achievement gaps, as expected, an association that is in part driven by declining income inequality in several Latin American countries and increasing income inequality in several post-Soviet countries. In contrast, in high-income countries, increasing income inequality is unexpectedly associated with *decreasing* gaps, driven by countries with increasing income inequality, very high levels of educational and occupational attainment, and stable or declining gaps, including the United States, England, Finland, Israel, and Japan. This latter result suggests that, in wealthy postindustrial economies with high levels of educational attainment and white-collar employment, many important gradations of inequality are not captured by educational degree and occupational categories (e.g., status hierarchies of educational institutions or fields of study and occupational sector). If household income better captures these gradations, this may explain why income achievement gaps but not parent education achievement gaps have increased in the United States (Reardon 2011b). The salience of income relative to other measures of SES may be growing in other societies as well; unfortunately, household income is not available in a large enough number of international assessments to examine this possibility in the present study. It may also be that declining gaps in some high-income countries represent true declines in educational inequality. Both Finland and England delayed the age when curricular tracking begins, a change associated with declining SES achievement gaps in the multivariate models, consistent with findings by van de Werfhorst (2018). However, changes in tracking policies cannot explain the secular global trend of increasing SES achievement gaps, as far more countries have moved the age of track selection later rather than earlier.

Thus, even as formal educational institutions have grown more equitable globally in terms of expanded access and less differentiation, other more informal, family-based inequalities may be driving increasing SES achievement gaps. This suggests that, in a growing number of countries, cognitive skills are an increasingly important dimension of education stratification. This is consistent with Alon's (2009) concept of "effectively expanding inequality" in the United States, in which higher social classes adapt to greater competition in higher-education admissions through an increased focus on their children's test scores. It also supports Baker's (2014) notion of a global "schooled society," in which cognitive skills are increasingly seen as the most important outcome of schooling and replace direct inheritance as the only legitimate source of social stratification. In such a society, all parents may equally recognize the importance of academic skills, but higher-SES families have greater resources and information about how to foster their children's achievement (Ishizuka 2018; Lareau 2000). Note that the driver of effectively expanding inequality highlighted by Alon—competition in higher-education admissions—was not found to predict increasing SES achievement gaps in the present study. However, it was not possible to measure higher-education competition in the same way as Alon for a large number of countries. It also may be that educational competition is not as strongly focused on the college transition in other countries as in the United States.

Although the multivariate analyses in the current study were not able to fully explain cross-national differences in trends in SES achievement gaps, the descriptive finding of a substantial average increase in the SES achievement gap worldwide, using a comprehensive long-term dataset, is an important starting point for future within-county and cross-national research. Growing SES achievement gaps raise serious concerns about equality of opportunity in many countries, as educational achievement (not on these particular tests—which are low-stakes—but on other national exams and in school grades) is an important predictor of higher educational attainment and life chances in adulthood. With broadening access to higher education, there is some evidence that the share of attainment inequality explained by achievement is declining in the United States and United Kingdom (Bailey and Dynarski 2011; Belley and Lochner 2007; Galindo-Rueda and Vignoles 2005). However, in the United States, the story changes when looking at *selective* university admissions, where the role of test scores appears to be increasing, meaning that SES gaps in enrollment are increasingly explained by SES achievement gaps (Alon and Tienda 2007; Bastedo and Jaquette 2011). International evidence also shows that SES achievement gaps explain a great deal of high-SES students' advantage in enrolling in high-status institutions in two other countries with highly stratified university systems, the United Kingdom and Australia (Jerrim, Chmielewski, and Parker 2015). Growing SES achievement gaps may also have political implications. Although belief in meritocracy is growing in many countries, this belief is strongly socioeconomically graded, particularly in countries with the highest income inequality (Mijs 2019; Roex, Huijts, and Sieben 2019). A growing awareness of increasing SES achievement gaps—coupled with cases of outright fraud, such as the recent U.S. college admissions bribery scandal (Smith 2019)—may contribute to increased socioeconomic polarization of trust in the legitimacy of educational institutions.

Finally, this study has important methodological implications. It implies that any future cross-cohort studies should take into account increasing SES achievement gaps, even when SES is merely a control variable, because SES is expected to explain larger amounts of variance in achievement over time in most countries around the world. It also demonstrates the power of examining data from a wide variety of countries, years, and sources. Unlike most prior cross-national evidence on the causes of SES achievement gaps, this study is not cross-sectional but instead examines changes over time within a large number of countries. Results from the multivariate models demonstrate that several key predictors have over-time relationships with SES achievement gaps that differ somewhat in size or direction from cross-sectional relationships. In addition, trends in SES achievement gaps are sometimes inconsistent when different international assessments are examined separately, as discussed earlier when comparing recent PISA and TIMSS reports (Broer et al. forthcoming; OECD 2018). In pooling gaps from different assessments (after harmonizing measures to the extent possible), I assume that the best estimate of the true average international trend should draw on all available data. However, results may still be confounded by discrepancies in testing frameworks and SES measures of different international assessments. Ultimately, the precise trends in the SES achievement gap for each individual country remain more uncertain than the overall average global trend.

Despite this uncertainty, the average global increase in SES achievement gaps is striking. However, the trend is not irreversible. Recent data show evidence of declining SES achievement gaps in some countries where they were previously increasing, including the United States, France, Hong Kong, and Russia (Broer et al. forthcoming; OECD 2018; Reardon and Portilla 2016). The large international dataset compiled for this study will be an important source of future evidence on a possible reversal of the global increase in SES achievement gaps, and it may point toward educational and social policies that could help mitigate disparities in learning opportunities for high- and low-SES children.

Acknowledgments

I would like to thank Sean Reardon, Barbara Schneider, Francisco Ramirez, Scott Davies, Elizabeth Dhuey, Hanna Dumont, Geoffrey Wodtke, Stephen Raudenbush, and Omar Khan. All errors are my own.

Funding

This project was partially funded by the Pathways to Adulthood program, sponsored by the Jacobs Foundation, and by a National Academy of Education/Spencer Foundation Postdoctoral Fellowship.

Notes

1. See Part B of the online supplement for models excluding country-study-years with low population coverage.
2. Gross domestic product per capita converted to current (2016) international dollars using purchasing power parity (PPP), obtained from the World Bank.
3. Supplemental analyses show that within-country variance in achievement is declining across waves in PISA, PIRLS, and TIMSS science (but not TIMSS math) (Part C of the online supplement). Yet SES gaps in unstandardized achievement are increasing, on average, for all test instruments except PISA math and science (Part B of the online supplement).
4. Supplemental analyses check the robustness of results by running models separately by subject (Part K of the online supplement); separately for TIMSS, PIRLS, and PISA (Part B); and computing gaps based on achievement rank (Part N). Results are similar for all analyses.
5. Additional detail on the treatment of mothers' and fathers' SES characteristics is reported in Part F of the online supplement.
6. I also ran models with categories harmonized across datasets, and results were similar (see Part B of the online supplement).
7. The large number of gaps with over 20 percent of students falling into the top SES category occur primarily in wealthy countries in recent years for parent education or parent occupation gaps, where large numbers of parents have university degrees or professional occupations. Because estimating the 90th percentile of SES in these cases requires extrapolation, 90/10 and 90/50 gaps may be poorly estimated. These gaps usually appear to be underestimated, as they increase less than gaps in household books in the same countries. Thus, including these poorly-estimated gaps likely yields a more conservative estimate of a smaller global increase in the SES achievement gap. However, I also ran models excluding gaps with 20 percent or more students falling into the top or bottom category, and results were similar (see Part H of the online supplement).
8. Assessments of 12th-grade students are omitted, as only a small proportion of the age-cohort remains in upper-secondary school in many countries, particularly in early cohorts.
9. World Bank, Luxembourg Income Study (LIS), and OECD data on income inequality (Gini coefficient) are not perfectly comparable. LIS and the OECD use disposable income (post-tax and transfer), whereas the World Bank uses household consumption in most countries (which I consider more comparable to disposable income) but gross income (pre-tax and transfer) in other countries (which I consider less comparable to disposable income). All three sources (the World Bank, LIS, and OECD) adjust by household size. Because I am interested here in comparing changes in time-varying covariates within countries over time, I only use one data source for each country. The validity of results thus

relies on the assumption that a one-unit change in each Gini measure is approximately equivalent, but not that the absolute levels of each measure are comparable.

10. Expected higher-education attendance is either student- or parent-reported, depending on the dataset. Student- and parent-reported expectations do not appear to differ in magnitude.
11. I also ran models using listwise deletion rather than multiple imputation of missing data, and results were similar (see Part L of the online supplement).
12. I use $m = 5$ rather than a greater number of imputed datasets, such as $m = 20$, out of practical consideration for computing time. With five imputed datasets, the main analyses required computing nearly 28 million gap estimates (5,541 SES achievement gaps $\times$ 1,000 bootstraps $\times$ 5 imputed datasets), in addition to several million more gap estimates for robustness checks reported in the online appendices. This required approximately 3,240 hours of computing time. The choice of $m = 5$ balances reasonable computing times with accuracy and efficiency of results. Using a smaller number of imputed datasets produces point estimates for coefficients that are unbiased and efficient. However, standard errors are unbiased but inefficient (von Hippel 2018). Thus, if I were to generate five new imputed datasets and rerun all analyses, the estimate for the trend in the SES achievement gap would likely remain similar. However, the standard error would likely change (it might either increase or decrease). Because the reported trend estimate is highly significant ($p <$ .001), I believe it is unlikely that a new estimate would fail to reach conventional significance levels. More importantly, I argue that the magnitude of the point estimate for the gap trend is large enough to be practically significant and theoretically meaningful for sociology of education research.
13. Cubic functions were chosen for consistency with Reardon (2011b). Quadratic or linear functions are used in country-years where there are insufficient SES categories. Linear functions are also used for country-years when over 20 percent of students fall into the top or bottom SES category, as linear functions can be estimated more reliably than cubic functions in these cases. I also ran models with all linear gaps, and results were similar (see Part L of the online supplement).
14. See Parts D and H of the online supplement for more information on the reliability adjustment. I also ran models without adjusting for reliability, and results were similar (see Part D of the online supplement).
15. Schomaker and Heumann (2016) show that "MI Boot" is unbiased but less efficient (produces more conservative confidence intervals) compared to "Boot MI." However, I prefer "MI Boot" because it requires far less computation time.
16. PISA 2015 used 10 rather than five plausible values of achievement. Thus, I generated 10 imputed datasets and combined them with the 10 plausible values of achievement.
17. This issue is discussed in more detail in Part E of the online supplement.
18. Models run separately for each SES variable are reported in Part J of the online supplement. Additional analyses of gaps computed from models including all three SES variables are reported in Part G of the online supplement.
19. I estimate this model in HLM7, which requires independent level-1 (within-study) errors. Therefore, following Kalaian and Raudenbush (1996), I implement the model by first transforming the within-study portion of the model using Cholesky factorization, yielding a level-1 error distribution of $\epsilon^*_{pjk} \sim N\left(0, \mathbf{I}_P\right)$, where $\mathbf{I}_P$ is the identity matrix of dimension P (the total number of gaps in country-study jk). I then estimate the model, constraining the level-1 variance to 1.
20. See Part K of the online supplement for estimates of separate gap trends by subject.
21. See Part H of the online supplement for a comparison of gaps based on student- and parent-reported SES.
22. See Part H of the online supplement for models omitting SES variable quality measures.
23. See Part K of the online supplement for estimates of separate gap trends by age.
24. See Part J of the online supplement for models run separately by SES variable.
25. See Part B of the online supplement for a comparison of results between Broer and colleagues (forthcoming) and the current study.
26. See Part H of the online supplement for more information on the quality of the parent education variable in PISA.
27. The 90/50 SES achievement gap may be somewhat underestimated in high-income countries due to the large number of students in the top parent education and occupation categories. However, results for 75/50 and 50/25 SES achievement gaps also show larger increases between the middle and bottom of the SES distribution than between the middle and top, even when the top is no longer as imprecisely estimated (see Part L of the online supplement).
28. It is likely that some covariates are not measured comparably across countries. For example, private school enrollment is very difficult to measure, as different organizational types are considered "private" in different countries. The private-school enrollment variable used in this study includes students enrolled in either privately- or publicly-funded private schools. This is a practical choice due to how the data are reported by the World Bank and OECD, but also a theoretical choice because the hypothesized mechanism behind the private-enrollment association includes not only tuition costs but also the stratifying effects of school choice more generally. But there are still inconsistencies across

countries in how publicly-funded private schools are counted. For example, charter schools in the United States are "public," but academy schools in the United Kingdom are "private"; publicly-funded Catholic schools are "private" in Belgium but "public" in Ontario, Canada.

29. Robustness checks pertaining to changing distributions of achievement and SES are available in Parts C and E of the online supplement, and those pertaining to changing measurement error are available in Parts D and H.
30. See Part G of the online supplement for analyses of trends in student-level correlations between different SES variables.

References

Alon, Sigal. 2009. "The Evolution of Class Inequality in Higher Education: Competition, Exclusion, and Adaptation." *American Sociological Review* 74(5):731–55.

Alon, Sigal, and Marta Tienda. 2007. "Diversity, Opportunity, and the Shifting Meritocracy in Higher Education." *American Sociological Review* 72(4):487.

Altinok, Nadir, Claude Diebolt, and Jean-Luc Demeulemeester. 2014. "A New International Database on Education Quality: 1965–2010." *Applied Economics* 46(11):1212–47.

Andon, Anabelle, Christopher G. Thompson, and Betsy J. Becker. 2014. "A Quantitative Synthesis of the Immigrant Achievement Gap across OECD Countries." *Large-Scale Assessments in Education* 2(7) (https://doi.org/10.1186/s40536-014-0007-2).

Ariga, Kenn, Giorgio Brunello, Roki Iwahashi, and Lorenzo Rocco. 2005. "Why Is the Timing of School Tracking So Heterogeneous?" Bonn, Germany: IZA Discussion Paper No. 1854.

Aurini, Janice, Scott Davies, and Julian Dierkes, eds. 2013. *Out of the Shadows: The Global Intensification of Supplementary Education*. Bingley, UK: Emerald Group Publishing.

Bailey, Martha J., and Susan M. Dynarski. 2011. "Inequality in Postsecondary Education." Pp. 117–32 in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, edited by R. J. Murnane and G. J. Duncan. New York: Russell Sage Foundation.

Baker, David P. 2014. *The Schooled Society: The Educational Transformation of Global Culture*. Stanford, CA: Stanford University Press.

Baker, David P., Brian Goesling, and Gerald K. LeTendre. 2002. "Socioeconomic Status, School Quality, and National Economic Development: A Cross-National Analysis of the 'Heyneman-Loxley Effect' on Mathematics and Science Achievement." *Comparative Education Review* 46(3):291–312.

Bastedo, Michael N., and Ozan Jaquette. 2011. "Running in Place: Low-Income Students and the Dynamics of Higher Education Stratification." *Educational Evaluation and Policy Analysis* 33(3):318–39.

Belley, Philippe, and Lance Lochner. 2007. "The Changing Role of Family Income and Ability in Determining Educational Achievement." *Journal of Human Capital* 1(1):37–89.

Benavot, Aaron. 1983. "The Rise and Decline of Vocational Education." *Sociology of Education* 56(2):63–76.

Bohlmark, Anders, and Mikael Lindahl. 2007. "The Impact of School Choice on Pupil Achievement, Segregation and Costs: Swedish Evidence." IZA Discussion Paper No. 2786.

Breen, Richard, and John H. Goldthorpe. 2001. "Class, Mobility and Merit: The Experience of Two British Birth Cohorts." *European Sociological Review* 17(2):81–101.

Broer, Markus, Yifan Bai, and Frank Fonseca. Forthcoming. *Socioeconomic Inequality and Educational Outcomes: Evidence from Twenty Years of TIMSS*. Vol. 5, *IEA Research for Education*. New York: Springer International Publishing (doi:https://www.springer.com/us/book/9783030119904).

Brown, Phillip. 1990. "The 'Third Wave': Education and the Ideology of Parentocracy." *British Journal of Sociology of Education* 11(1):65–86.

Brunello, Giorgio, and Daniele Checchi. 2007. "Does School Tracking Affect Equality of Opportunity? New International Evidence." *Economic Policy* 52:781–861.

Byun, Soo-yong, and Kyung-keun Kim. 2010. "Educational Inequality in South Korea: The Widening Socioeconomic Gap in Student Achievement." Pp. 155–82 in *Globalization, Changing Demographics, and Educational Challenges in East Asia*, Vol. 17, *Research in Sociology of Education*, edited by E. Hannum, H. Park, and Y. G. Butler. Bingley, UK: Emerald Group Publishing Limited.

Byun, Soo-yong, Kyung-keun Kim, and Hyunjoon Park. 2012. "School Choice and Educational Inequality in South Korea." *Journal of School Choice* 6(2):158–83.

Chang, Jason Chien-chen. 2014. "Parentocracy and the Life and Death of Secondary Education for All in Taiwan." Presented at the XVIII ISA World Congress of Sociology, July 13–19, Yokohama, Japan.

Dotti Sani, Giulia M., and Judith Treas. 2016. "Educational Gradients in Parents' Child-Care Time across Countries, 1965–2012." *Journal of Marriage and Family* 78(4):1083–96.

Downey, Douglas B., and Dennis J. Condron. 2016. "Fifty Years since the Coleman Report: Rethinking the Relationship between Schools and Inequality." *Sociology of Education* 89(3):207–20.

Dumont, Hanna, Denise Klinge, and Kai Maaz. 2019. "The Many (Subtle) Ways Parents Game the System: Mixed-Method Evidence on the Transition into Secondary-School Tracks in Germany." *Sociology of Education* 92(2):199–228.

Dupriez, Vincent, and Xavier Dumay. 2006. "Inequalities in School Systems: Effect of School Structure or of Society Structure?" *Comparative Education* 42(2):243–60.

Duru-Bellat, Marie, and Bruno Suchaut. 2005. "Organisation and Context, Efficiency and Equity of Educational Systems: What PISA Tells Us." *European Educational Research Journal* 4(3):181–94.

Faircloth, Charlotte, Diane M. Hoffman, and Linda L. Layne. 2013. *Parenting in Global Perspective: Negotiating Ideologies of Kinship, Self and Politics*. New York: Routledge.

Falch, Torberg, and Justina A. V. Fischer. 2012. "Public Sector Decentralization and School Performance: International Evidence." *Economics Letters* 114(3):276–79.

Galindo-Rueda, Fernando, and Anna Vignoles. 2005. "The Declining Relative Importance of Ability in Predicting Educational Attainment." *Journal of Human Resources* 40(2):335–53.

Ganzeboom, Harry B. G., and Donald J. Treiman. 1996. "Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations." *Social Science Research* 25(3):201–39.

Gauthier, Anne H., Timothy M. Smeeding, and Frank F. Furstenberg. 2004. "Are Parents Investing Less Time in Children? Trends in Selected Industrialized Countries." *Population and Development Review* 30(4):647–72.

Gomez Espino, Juan Miguel. 2013. "Two Sides of Intensive Parenting: Present and Future Dimensions in Contemporary Relations between Parents and Children in Spain." *Childhood* 20(1):22–36.

Gorard, Stephen. 2014. "The Link between Academies in England, Pupil Outcomes and Local Patterns of Socio-economic Segregation between Schools." *Research Papers in Education* 29(3):268–84.

Hanushek, Eric A., Paul E. Peterson, Laura M. Talpey, and Ludger Wößmann. 2019. "The Unwavering SES Achievement Gap: Trends in U.S. Student Performance." NBER Working Paper No. 25648.

Hanushek, Eric A., and Ludger Wößmann. 2012. "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation." *Journal of Economic Growth* 17(4):267–321.

Hays, Sharon. 1996. *The Cultural Contradictions of Motherhood*. New Haven, CT: Yale University Press.

Heyneman, Stephen P., and William A. Loxley. 1983. "The Effect of Primary-School Quality on Academic Achievement across Twenty-Nine High- and Low-Income Countries." *American Journal of Sociology* 88(6):1162–94.

Ishizuka, Patrick. 2018. "Social Class, Gender, and Contemporary Parenting Standards in the United States: Evidence from a National Survey Experiment." *Social Forces* (https://doi.org/10.1093/sf/soy107).

Jackson, Michelle. 2013. *Determined to Succeed? Performance Versus Choice in Educational Attainment*. Stanford, CA: Stanford University Press.

Janmaat, Jan Germen. 2013. "Subjective Inequality: A Review of International Comparative Studies on People's Views about Inequality." *European Journal of Sociology/Archives Européennes de Sociologie* 54(3):357–89.

Jerrim, John, Anna K. Chmielewski, and Phil Parker. 2015. "Socioeconomic Inequality in Access to High-Status Colleges: A Cross-Country Comparison." *Research in Social Stratification and Mobility* 42:20–32.

Jerrim, John, and John Micklewright. 2014. "Socio-economic Gradients in Children's Cognitive Skills: Are Cross-Country Comparisons Robust to Who Reports Family Background?" *European Sociological Review* 30(6):766–81.

Kalaian, Hripsime A., and Stephen W. Raudenbush. 1996. "A Multivariate Mixed Linear Model for Meta-analysis." *Psychological Methods* 1(3):227–35.

Karsten, Lia. 2015. "Middle-Class Childhood and Parenting Culture in High-Rise Hong Kong: On Scheduled Lives, the School Trap and a New Urban Idyll." *Children's Geographies* 13(5):556–70.

Katartzi, Eugenia. 2017. "Youth, Family and Education: Exploring the Greek Case of Parentocracy." *International Studies in Sociology of Education* 26(3):310–25.

Keeves, John P., ed. 1992. *The IEA Study of Science III: Changes in Science Education and Achievement: 1970 to 1984*. Oxford, UK: Pergamon Press.

Kornrich, Sabino, Anne H. Gauthier, and Frank F. Furstenberg. 2011. "Changes in Private Investments in Children across Three Liberal Welfare States: Australia, Canada, and the United States." Presented at the Annual Meeting of the Population Association of America, Washington, DC.

Lareau, Annette. 2000. *Home Advantage: Social Class and Parental Intervention in Elementary Education*, 2nd ed. Lanham, MD: Rowman & Littlefield Publishers.

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

Lindbom, Anders. 2010. "School Choice in Sweden: Effects on Student Performance, School Costs, and Segregation." *Scandinavian Journal of Educational Research* 54(6):615–30.

Liu, Fengshu. 2016. "The Rise of the 'Priceless' Child in China." *Comparative Education Review* 60(1):105–30.

Manning, Alan, and Jörn-Steffen Pischke. 2006. "Comprehensive versus Selective Schooling in England in Wales: What Do We know?" Bonn, Germany: IZA Discussion Paper No. 2072.

Mare, Robert D. 1980. "Social Background and School Continuation Decisions." *Journal of the American Statistical Association* 75(370):295–305.

Marks, Gary N. 2005. "Cross-National Differences and Accounting for Social Class Inequalities in Education." *International Sociology* 20(4):483–505.

Mijs, Jonathan J. B. 2019. "The Paradox of Inequality: Income Inequality and Belief in Meritocracy Go Hand in Hand." *Socio-Economic Review* (https://doi.org/10.1093/ser/mwy051).

Mullis, Ina V. S., Michael O. Martin, Pierre Foy, and Martin Hooper. 2016. "TIMSS 2015 International Results in Mathematics." Boston, MA: TIMSS & PIRLS International Study Center, Boston College.

Musset, Pauline. 2012. "School Choice and Equity: Current Policies in OECD Countries and a Literature Review." Paris: OECD Working Paper EDU/WKP(2012)3.

Musterd, Sako, Szymon Marcińczak, Maarten Van Ham, and Tiit Tammaru. 2017. "Socioeconomic Segregation in European Capital Cities: Increasing Separation between Poor and Rich." *Urban Geography* 38(7):1062–83.

OECD (Organization for Economic Cooperation and Development). 2015. "In It Together: Why Less Inequality Benefits All." Paris: OECD Publishing.

OECD (Organization for Economic Cooperation and Development). 2016. "PISA 2015 Results: Excellence and Equity in Education (Volume I)." Paris: OECD Publishing.

OECD (Organization for Economic Cooperation and Development). 2018. "Equity in Education: Breaking Down Barriers to Social Mobility." Paris: OECD Publishing.

Park, Hyunjoon, Claudia Buchmann, Jaesung Choi, and Joseph J. Merry. 2016. "Learning beyond the School Walls: Trends and Implications." *Annual Review of Sociology* 42:231–52.

Postlethwaite, T. Neville. 1995. *International Encyclopedia of National Systems of Education*. Oxford, UK: Pergamon.

Quirke, Linda. 2006. "'Keeping Young Minds Sharp': Children's Cognitive Stimulation and the Rise of Parenting Magazines, 1959–2003." *Canadian Review of Sociology/Revue Canadienne de Sociologie* 43(4):387–406.

Ramey, Garey, and Valerie A. Ramey. 2010. "The Rug Rat Race." *Brookings Papers on Economic Activity* 41(1):129–99.

Reardon, Sean F. 2011a. "Appendices for 'The Widening Academic Achievement Gap between the Rich and the Poor: New Evidence and Possible Explanations.'" Pp. 1–49 in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, edited by G. J. Duncan and R. J. Murnane. New York: Russell Sage Foundation.

Reardon, Sean F. 2011b. "The Widening Academic Achievement Gap between the Rich and the Poor: New Evidence and Possible Explanations." Pp. 91–115 in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, edited by G. J. Duncan and R. J. Murnane. New York: Russell Sage Foundation.

Reardon, Sean F., and Kendra Bischoff. 2011. "Income Inequality and Income Segregation." *American Journal of Sociology* 116(4):1092–153.

Reardon, Sean F., and Ximena A. Portilla. 2016. "Recent Trends in Income, Racial, and Ethnic School Readiness Gaps at Kindergarten Entry." *AERA Open* 2(3) (https://doi.org/10.1177/2332858416657343).

Roex, Karlijn L. A., Tim Huijts, and Inge Sieben. 2019. "Attitudes towards Income Inequality: 'Winners' versus 'Losers' of the Perceived Meritocracy." *Acta Sociologica* 62(1):47–63.

Saw, Guan Kung. 2016. "Patterns and Trends in Achievement Gaps in Malaysian Secondary Schools (1999–2011): Gender, Ethnicity, and Socioeconomic Status." *Educational Research for Policy and Practice* 15(1):41–54.

Schaub, Maryellen. 2010. "Parenting for Cognitive Development from 1950 to 2000: The Institutionalization of Mass Education and the Social Construction of Parenting in the United States." *Sociology of Education* 83(1):46–66.

Schomaker, Michael, and Christian Heumann. 2016. "Bootstrap Inference When Using Multiple Imputation." *arXiv* (https://arxiv.org/abs/1602.07933).

Smith, Laura, Special Agent FBI. 2019. "College Admissions Bribery Scheme Affidavit." *The Washington Post*, March 12.

Söderström, Martin, and Roope Uusitalo. 2010. "School Choice and Segregation: Evidence from an Admission Reform." *Scandinavian Journal of Economics* 112(1):55–76.

Tan, Charlene. 2017. "Private Supplementary Tutoring and Parentocracy in Singapore." *Interchange* 48(4):315–29.

UNDP (United Nations Development Programme). 2013. "Humanity Divided: Confronting Inequality in Developing Countries." New York: United Nations Development Program.

UNESCO. 2015. "Education for All 2000–2015: Achievements and Challenges." Paris: UNESCO.

Valenzuela, Juan Pablo, Cristian Bellei, and Danae De Los Ríos. 2014. "Socioeconomic School Segregation in a Market-Oriented Educational System: The Case of Chile." *Journal of Education Policy* 29(2):217–41.

Van de Werfhorst, Herman G. 2018. "Early Tracking and Socioeconomic Inequality in Academic Achievement: Studying Reforms in Nine Countries." *Research in Social Stratification and Mobility* 58:22–32.

Van de Werfhorst, Herman G., and Jonathan J. B. Mijs. 2010. "Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective." *Annual Review of Sociology* 36:407–28.

von Hippel, Paul T. 2018. "How Many Imputations Do You Need? A Two-Stage Calculation Using a Quadratic Rule." *Sociological Methods & Research*. Online first (https://doi.org/10.1177/0049124117747303).

Wiseman, Alexander W., David P. Baker, Catherine Riegle-Crumb, and Francisco O. Ramirez. 2009. "Shifting Gender Effects: Opportunity Structures, Institutionalized Mass Schooling, and Cross-National Achievement in Mathematics." Pp. 395–422 in *Gender, Equality and Education from International and Comparative Perspectives*, edited by D. P. Baker and A. W. Wiseman. Bingley, UK: Emerald Group Publishing Limited.

Anna K. Chmielewski is Assistant Professor of Educational Leadership and Policy at the Ontario Institute for Studies in Education (OISE) of the University of Toronto. Her research focuses on international comparisons of educational inequality; curriculum differentiation and tracking; school segregation; and international large-scale assessments.

The Global Increase in Socioeconomic Achievement Gaps, 1964 to 2015

Anna K. Chmielewski
University of Toronto

Online Appendices

These appendices report additional details of the results from the main text of the paper, as well as supplementary analyses undertaken to test the sensitivity of results to a number of different limitations of the data. The finding of global increases in SES achievement gaps is generally robust to differences across test instruments, changes in the distribution of achievement and of SES, and changes in the measurement error of achievement and of SES. The multivariate findings predicting changing country achievement gaps from changing country characteristics and policies are generally robust across a variety of model specifications.

Table of Contents

Appendix A. List of countries and datasets included in the study

Table A1. List of Included Countries and Datasets

code	country	region	mid/low-income in 1980	fims1964	fiss1970gr4	fiss1970gr8	fircs1970gr4	fircs1970gr8	sims1980	siss1984gr4	siss1984gr8	rls1991gr4	rls1991gr8	timss1995gr4	timss2003gr4	timss2007gr4	timss2011gr4	timss2015gr4	timss1995gr8	timss1999	timss2003gr8	timss2007gr8	timss2011gr8	timss2015gr8	pirls2001	pirls2006	pirls2011	pisa2000	pisa2003	pisa2006	pisa2009	pisa2012	pisa2015
ALB	Albania	E	x																									x			x	(x)	(x)
DZA	Algeria	M	x													x						x											x
ARG	Argentina	L	x																						x			x		x	x	x	
ARM	Armenia	E	x												x	(x)	x				x	x	x										
AUS	Australia	W		x		x				x	x			x	x	x	x	x	x	x	x	x	x	x			x	x	x	x	x	x	x
AUT	Austria	W												x		x	x		x							x	x	x	x	x	x	x	x
AZE	Azerbaijan	E	x														x										x			x	x		
BHR	Bahrain	M															(x)	x			x	x	x	x									
BEL	Belgium	W		x																													
BFL	Belgium-Flemish	W			x	x	x	x	x						x		x	x	x	x	x					x		x	x	x	x	x	x
BFR	Belgium-French	W			x	x	x	x	x			x	x						x							x	x	x	x	x	x	x	x
BLZ	Belize	L	x																						x								
BIH	Bosnia & Herzegovina	E	x																			x											
BWA	Botswana	Af	x										x				x				x	x	x	x			x						
BRA	Brazil	L	x																									x	x	x	x	x	x
BGR	Bulgaria	E	x															x	(m)	x	x	x			x	x	x	x		x	x	x	x
CAN	Canada	W								(m)	(m)			x		x		x	x	x		x		x		x	x	x	x	x	x	x	x
CHL	Chile	L	x		(x)	x	x	x									x	x		x	x		x	x				x		x	x	x	x
CHN	China	As	x								x																						
TWN	Chinese Taipei	As	x												x	x	x	x		x	x	x	x	x		x	x			x	x	x	x
COL	Colombia	L	x													x			x			x			x		x			x	x	x	x
CRI	Costa Rica	L	x																												x	x	x
HRV	Croatia	E															x	x									x			x	x	x	x
CYP	Cyprus	W	x									(x)	x	x	x			x	x	x	x	x			x								
CZE	Czech Rep.	E												x		x	x	x	x	x		x			x		x	x	x	x	x	x	x
DNK	Denmark	W										x	x			x	x	x	x							x	x	x	x	x	x	x	x
DOM	Dominican Republic	L	x																														x
EGY	Egypt	M	x																		x	x		x									
SLV	El Salvador	L	x													x						x											
ENG	England	W		x	x	x	x	x	x	x	x			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
EST	Estonia	E																			x									x	x	x	x
FIN	Finland	W		x	x	x	x	x	x	x	x	x	x				x	x		x			x				x	x	x	x	x	x	x
FRA	France	W		x					x			x	x					x	x						x	x	x	x	x	x	x	x	x
GEO	Georgia	E	x													x	x	x				x	x	x		x	x				x		x
DEU	Germany	W														x	x	x	x						x	x	x	x	x	x	x	x	x
GDR	Germany-East	W										x	x																				
FRG	Germany-West	W		x	x	x						x	x																				
GHA	Ghana	Af	x								x										x	x	x										
GRC	Greece	W										x	x	x					x						x			x	x	x	x	x	x
HND	Honduras	L	x														x						x				x						
HKG	Hong Kong	As							x	x	x	(x)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
HUN	Hungary	E	x		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
ISL	Iceland	W										x	x	x					x						x	x		x	x	x	x	x	x
IND	India	As	x		x	x	x	x																									
IDN	Indonesia	As	x									x		(m)				x		x	x	x	x			x	x	x	x	x	x	x	x
IRN	Iran	M	x		x	x	x	x						x	x	x	x	x	x	x	x	x	x	x	x	x	x						
IRL	Ireland	W										x	x	x			x	x	x					x			x	x	x	x	x	x	x
ISR	Israel	W		x			x	x	x	x	(m)			x					x	x	x	x	x	x	x	x	x	x		x	x	x	x
ITA	Italy	W			x	x	x	x		x	x	(x)	x		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
JPN	Japan	As		x	x	x			x	x	x			(m)	x	x	x	x	x	(m)	x	x	x	x				x	x	x	x	x	x
JOR	Jordan	M	x																	x	x	x	x	x						x	x	x	x
KAZ	Kazakhstan	E	x													x	x	x					x	x							x	x	
KOR	Korea, Rep.	As	x							x	x			x			x	x	x	x	x	x	x	x				x	x	x	x	x	x
KSV	Kosovo	E	x																														x
KWT	Kuwait	M												x		x	x	x	x			x		x	x	x	x						
KGZ	Kyrgyzstan	E	x																											x	x		
LVA	Latvia	E	x											x	x	x			x	x	x				x	x		x	x	x	x	x	x

Notes: Shaded countries were excluded because they participated in only one test. (m) denotes missing SES data; (x) denotes that gaps could not be computed, usually because of low-quality SES data. Regions include: Af=sub-Saharan Africa, As=east and southeast Asian and Pacific countries, M=Middle Eastern and North African countries, E=Eastern Europe and the Commonwealth of Independent States, L=Latin America and the Caribbean, and W=Western countries (Western Europe and Anglophone countries).

Table A1. (cont.) List of Included Countries and Datasets

code	country	region	mid/low-income in 1980	fims1964	fiss1970gr4	fiss1970gr8	fircs1970gr4	fircs1970gr8	sims1980	siss1984gr4	siss1984gr8	rls1991gr4	rls1991gr8	timss1995gr4	timss2003gr4	timss2007gr4	timss2011gr4	timss2015gr4	timss1995gr8	timss1999	timss2003gr8	timss2007gr8	timss2011gr8	timss2015gr8	pirls2001	pirls2006	pirls2011	pisa2000	pisa2003	pisa2006	pisa2009	pisa2012	pisa2015
LBN	Lebanon	M	x																		x	x	x	x									x
LIE	Liechtenstein	W																										x	x	x	x	x	
LTU	Lithuania	E													x	x	x	x	x	x	x	x	x	x	x	x	x	(m)		x	x	x	x
LUX	Luxembourg	W							x																	x		x	x	x	x	x	x
MAC	Macao-China	As																											x	x	x	x	x
MKD	Macedonia	E	x																	x	x		x		x	x		x					x
MYS	Malaysia	As	x																	x	x	x	x	x							x	x	
MLT	Malta	W	x														x					x		x			x				x		x
MUS	Mauritius	Af	x																												x		
MEX	Mexico	L	x											(m)														x	x	x	x	x	x
MDA	Moldova	E	x												x					x	x				x	x					x		x
MNG	Mongolia	As	x													x						x											
MNE	Montenegro	E																												x	x	x	x
MAR	Morocco	M	x												x	x	x	x		x	x	x	x	x	(x)	x	x						
NLD	Netherlands	W		x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x				x	x	x	x	x	x	x	x	x
NZL	New Zealand	W				x		x	x			x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x
NGA	Nigeria	Af	x						(x)	x	x		x																				
NIR	Northern Ireland	W															x	x									x						
NOR	Norway	W								x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x	x
OMN	Oman	M	x														x	x				x	x	x			x						
PAN	Panama	L	x																												x		
PNG	Papua New Guinea	As	x								x																						
PER	Peru	L	x																									x			x	x	x
PHL	Philippines	As	x							x	x		x		x				x	x	x												
POL	Poland	E	x							x	x						x	x								x	x	x	x	x	x	x	x
PRT	Portugal	W	x									x	x	x			x	x	x								x	x	x	x	x	x	x
QAT	Qatar	M														x	x	x				x	x	x		x	x			x	x	x	x
ROM	Romania	E	x														x		x	x	x	x	x		x	x	x	x		x	x	x	x
RUS	Russian Fed.	E													x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
SAU	Saudi Arabia	M															x	x			x	x	x	x			x						
SCO	Scotland	W		x	x	x	x	x	x					x	x	x			x		x	x			x	x		x	x	x	x	x	x
SRB	Serbia	E	x														x	x			x	x							x	x	x	x	
QCN	Shanghai-China	As	x																												x	x	
SGP	Singapore	As								x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				x	x	x
SVK	Slovak Rep.	E	x													x	x	x	x	x	x				x	x	x		x	x	x	x	x
SVN	Slovenia	E										x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x			x	x	x	x
ZAF	South Africa	Af	x																x	x	x		x	x		x	x						
ESP	Spain	W										x	x				x	x	x							x	x	x	x	x	x	x	x
SWZ	Swaziland	Af	x						x																								
SWE	Sweden	W		(m)	x	x	x	x	x	x	x	x	x			x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x	x
CHE	Switzerland	W										x	x						x									x	x	x	x	x	x
SYR	Syria	M	x																		x	x	x										
THA	Thailand	As	x		x	x			x		x		x	x			x		x	x		x	x	x				x	x	x	x	x	x
TTO	Trinidad and Tobago	L										x	x													x	x				x		x
TUN	Tunisia	M	x												x	x	x			x	x	x	x						x	x	x	x	x
TUR	Turkey	W	x														x	x		x		x	x	x	x				x	x	x	x	x
UKR	Ukraine	E	x													x						x	x										
ARE	United Arab Emirates	M															x	x					x	x			x				x	x	x
USA	United States	W		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	(x)	x	x	x	x	x	x	x
URY	Uruguay	L	x																										x	x	x	x	x
VEN	Venezuela	L	x									x	x																				
VNM	Vietnam	As	x																												(x)	x	x
PSE	West Bank and Gaza	M	x																		x	x	x										
YEM	Yemen	M	x												(x)	x	x																
ZWE	Zimbabwe	Af	x								x		x																				

Notes: Shaded countries were excluded because they participated in only one test. (m) denotes missing SES data; (x) denotes that gaps could not be computed, usually because of low-quality SES data. Regions include: Af=sub-Saharan Africa, As=east and southeast Asian and Pacific countries, M=Middle Eastern and North African countries, E=Eastern Europe and the Commonwealth of Independent States, L=Latin America and the Caribbean, and W=Western countries (Western Europe and Anglophone countries).

Appendix B. Combining different test instruments

In order to estimate international trends in SES achievement gaps over a 50-year period, this paper combines data from a variety of different international assessments of math, science, and reading. However, among tests of the same subject, a comparison of skills frameworks from the official study reports reveals differences. For example, the IEA math and science tests—FIMS, FISS, SIMS, SISS, and TIMSS—are curriculum-based, while the PISA math and science tests, as well as all reading tests—PISA, FIRCS, RLS, and PIRLS—are literacy-based. Though the early IEA tests contained anchor items to enable studying trends in achievement, the scores were not placed on common scales, and they did not have the advantage of improvements to testing methodology in the 1990s; thus, early and recent IEA tests are not strictly comparable. The analyses in the main text of the paper deal with this issue by standardizing achievement within each study and each country and assuming only that each test is interval-scaled and that different tests rank students similarly.

However, there are six recent studies that repeat the same test instrument to enable measuring trends over time: TIMSS 4th and 8th grade math and science; PIRLS; and PISA reading, math, and science. These trend studies allow us to investigate the sensitivity of gap trend results to differences across test instruments—but only over the recent 9-20 years that the studies have been conducted. TIMSS trends (for both grades and subjects) can be estimated for test years 1995 to 2015 (birth cohorts 1981 to 2001 for 8th grade and birth cohorts 1985 to 2005 for 4th grade); PIRLS trends for test years 2001 to 2011 (birth cohorts 1991 to 2001); PISA reading trends for test years 2000 to 2015 (birth cohorts 1985 to 2000); PISA math trends for test years 2003 to 2015 (birth cohorts 1988 to 2000); and PISA science trends for test years 2006 to 2015 (birth cohorts 1991 to 2000). In addition, because each instrument remains the same over time, it is not necessary to standardize achievement within studies or countries, meaning we can examine changes in SES achievement gaps in light of possible changes in the variance of skills (which will be addressed in Appendix C).

Tables B1-B4 compute trends in SES achievement gaps separately for each test instrument. For each of the eight test instruments, Model 1 is a hierarchical multivariate variance-known model that estimates a different cohort slope for each gap type (parent education, parent occupation, or household books) using interactions between cohort birth year and gap type indicators. Model 2 estimates a single pooled cohort slope for all gap types. Thus, the specifications of Models 1 and 2 are the same as Models 1 and 2 in the main text but without control variables (as the subject and age at testing dummies from the models in the main text do not vary within each test instrument). Tables B1 and B2 report trends in gaps estimated without standardizing achievement within each country-study-year. Tables B3 and B4 report gaps estimated when achievement is standardized within each county-study-year, as in the main text of the paper.[1,2] It can be seen from the reported country sample sizes that the number of participating countries varies widely across the different test instruments. It is not possible to reliably estimate trends across all test instruments for a core group of countries that has

[1] The Standardized Achievement models adjust each gap estimate for the reliabilities of test instruments and SES report; the Unstandardized Achievement models adjust each gap estimate only for the reliability of SES report. It is not necessary to adjust for test reliability in these models as gap estimates are not attenuated since they have not been divided by the test score variance.

[2] When computing these gaps, rather than using all available categories of each SES variable as in the main text of the paper, each SES variable was recoded to ensure that the SES instrument remained the same across test years. There were six categories for parent education and parent occupation and five categories for household books.

participated in every test, as there are too few countries that have done so. Thus, the reported trends for each test instrument should be interpreted only as a rough indication of the sensitivity of the general finding of increasing SES achievement gaps over time. The size of coefficients can be compared across different test instruments only for the models using standardized achievement, not for those using unstandardized achievement, as they are in different metrics and 1 point in PISA, for example, is not the same as 1 point in PIRLS or TIMSS. Significance levels should be interpreted with caution because of changing sample sizes and the large number of significance tests conducted; significance is reported only as a rough indication of the precision of each estimate.

Overall, the estimated gap trends are positive for most test instruments. Gaps are consistently increasing for all three SES variables in PIRLS and for both available SES variables in both subjects of TIMSS at both the 4^{th} and the 8^{th} grades. However, gap trends for PISA are more mixed. In particular, trends in gaps based on unstandardized math and science achievement are negative, while trends in gaps based on standardized math and science achievement and based on reading achievement (standardized or unstandardized) are more consistently positive. This is likely due to substantial declines in the variance of PISA math and science scores (see Appendix C). Additionally, for all three PISA subject tests, trends in gaps based on parent education are negative. Further analysis shows that this pattern may be due to problems in the measurement of parent education in the PISA student survey (see Appendix H). Therefore, with the exception of some unreliable trends, positive increases in SES achievement gaps over time are quite robust across the different test instruments that are combined in the main text of the paper.

Table B1. Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th Percentiles of SES, Run Separately by Test Instrument (Unstandardized Achievement) – PISA and PIRLS

Test instrument	PISA Math		PISA Reading		PISA Science		PIRLS Reading	
Test years	2003-2015		2000-2015		2006-2015		2001-2011	
Cohort birth years	1988-2000		1985-2000		1991-2000		1991-2001	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Parent education gaps intercept	88.580 ***	86.591 ***	85.725 ***	82.817 ***	92.806 ***	86.595 ***	76.698 ***	78.551 ***
	(3.696)	(3.381)	(3.337)	(3.130)	(4.189)	(3.472)	(3.715)	(3.840)
Parent occupation gaps intercept	86.131 ***	85.944 ***	83.782 ***	84.089 ***	84.459 ***	85.742 ***	68.610 ***	68.730 ***
	(3.328)	(3.278)	(3.180)	(3.193)	(3.602)	(3.620)	(3.579)	(3.447)
Household books gaps intercept	127.448 ***	127.860 ***	125.759 ***	128.238 ***	132.753 ***	132.500 ***	87.755 ***	86.356 ***
	(4.460)	(4.364)	(4.230)	(4.380)	(5.621)	(5.252)	(3.864)	(3.975)
Cohort birth year × Parent education	-0.491 *		-0.374 *		-0.865 **		0.769 +	
	(0.225)		(0.188)		(0.263)		(0.407)	
Cohort birth year × Parent occupation	-0.288 +		0.049		-0.124		0.507	
	(0.173)		(0.172)		(0.189)		(0.358)	
Cohort birth year × Books	-0.206		0.377		-0.261		0.224	
	(0.272)		(0.276)		(0.286)		(0.324)	
Cohort birth year		-0.266		0.027		-0.242		0.470
		(0.174)		(0.174)		(0.194)		(0.319)
Random effects								
Level 2 - Residual variance between studies in…								
Parent education intercepts	83.585	80.729	59.861	55.654	49.099	58.929	46.742	45.620
Parent occupation intercepts	34.517	35.085	32.022	31.198	29.029	29.485	35.950	40.488
Books intercepts	78.836	98.737	55.998	73.371	44.217	55.873	86.374	86.604
Level 3 - Residual variance between countries in…								
Parent education intercepts	629.437	601.335	605.245	560.178	713.887	564.316	408.697	465.914
Parent occupation intercepts	629.187	600.721	614.484	630.121	689.660	679.301	408.276	364.008
Books intercepts	1048.458	1152.394	1057.018	1245.643	1737.470	1668.747	433.859	491.341
Parent education cohort slopes	0.649		0.806		0.845		3.323	
Parent occupation cohort slopes	0.802		1.015		0.769		2.778	
Books cohort slopes	1.838		2.996		1.757		0.708	
Cohort slopes		0.737		1.198		0.684		1.671
% Parent education cohort slopes > 0	0.271		0.339		0.173		0.663	
% Parent occupation cohort slopes > 0	0.374		0.519		0.444		0.620	
% Books cohort slopes > 0	0.440		0.586		0.422		0.605	
% Cohort slopes > 0		0.378		0.510		0.385		0.642
N (Level 1 - gaps)	893	893	1030	1030	764	764	300	300
N (Level 2 - study-years)	298	298	344	344	255	255	103	103
N (Level 3 - countries)	70	70	72	72	70	70	41	41

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L). "% cohort slopes > 0" is computed by comparing cohort coefficient estimate and cohort slope variance component to a normal distribution.

Table B2. Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th Percentiles of SES, Run Separately by Test Instrument (Unstandardized Achievement) – TIMSS

Test instrument	TIMSS Grade 4 Math		TIMSS Grade 4 Science		TIMSS Grade 8 Math		TIMSS Grade 8 Science	
Test years	1995-2015		1995-2015		1995-2015		1995-2015	
Cohort birth years	1985-2005		1985-2005		1981-2001		1981-2001	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Parent education gaps intercept					70.929 ***	70.212 ***	70.426 ***	68.680 ***
					(3.629)	(3.561)	(4.224)	(3.998)
Household books gaps intercept	74.438 ***	74.438 ***	80.765 ***	80.765 ***	80.117 ***	80.820 ***	81.421 ***	83.011 ***
	(4.988)	(4.988)	(5.859)	(5.859)	(4.473)	(4.505)	(5.034)	(5.132)
Cohort birth year × Parent education					1.225 ***		1.267 ***	
					(0.255)		(0.269)	
Cohort birth year × Books	0.775 **		0.682 *		1.389 ***		1.591 ***	
	(0.275)		(0.312)		(0.309)		(0.271)	
Cohort birth year		0.775 **		0.682 *		1.318 ***		1.450 ***
		(0.275)		(0.312)		(0.257)		(0.240)
Random effects								
Level 2 - Residual variance between studies in…								
Parent education intercepts					127.994	122.721	157.985	163.727
Books intercepts	38.198	38.198	109.020	109.020	138.450	157.292	162.137	167.889
Level 3 - Residual variance between countries in…								
Parent education intercepts					441.812	460.884	670.049	624.631
Books intercepts	681.626	681.626	903.670	903.670	817.700	925.355	1116.629	1258.709
Parent education cohort slopes					1.501		1.572	
Books cohort slopes	1.526		1.513		2.974		1.590	
Cohort slopes		1.526		1.513		1.990		1.345
% Parent education cohort slopes > 0					0.841		0.844	
% Books cohort slopes > 0	0.735		0.710		0.790		0.897	
% Cohort slopes > 0		0.735		0.710		0.825		0.894
N (Level 1 - gaps)	163	163	165	165	485	485	485	485
N (Level 2 - study-years)	163	163	165	165	245	245	245	245
N (Level 3 - countries)	49	49	50	50	61	61	61	61

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L). "% cohort slopes > 0" is computed by comparing cohort coefficient estimate and cohort slope variance component to a normal distribution.

Table B3. Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th Percentiles of SES, Run Separately by Test Instrument (Standardized Achievement) – PISA and PIRLS

Test instrument	PISA Math		PISA Reading		PISA Science		PIRLS Reading	
Test years	2003-2015		2000-2015		2006-2015		2001-2011	
Cohort birth years	1988-2000		1985-2000		1991-2000		1991-2001	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Parent education gaps intercept	1.000 ***	0.972 ***	0.946 ***	0.912 ***	1.044 ***	0.980 ***	1.042 ***	1.066 ***
	(0.037)	(0.034)	(0.032)	(0.030)	(0.041)	(0.034)	(0.043)	(0.039)
Parent occupation gaps intercept	0.971 ***	0.964 ***	0.923 ***	0.925 ***	0.953 ***	0.968 ***	0.928 ***	0.934 ***
	(0.034)	(0.033)	(0.029)	(0.029)	(0.034)	(0.034)	(0.043)	(0.040)
Household books gaps intercept	1.438 ***	1.455 ***	1.401 ***	1.424 ***	1.506 ***	1.505 ***	1.209 ***	1.192 ***
	(0.044)	(0.043)	(0.040)	(0.042)	(0.051)	(0.048)	(0.041)	(0.046)
Cohort birth year × Parent education	-0.001		-0.002		-0.007 **		0.013 **	
	(0.002)		(0.002)		(0.003)		(0.004)	
Cohort birth year × Parent occupation	0.001		0.002		0.001		0.010 **	
	(0.002)		(0.002)		(0.002)		(0.004)	
Cohort birth year × Books	0.004		0.006 *		-0.001		0.006	
	(0.003)		(0.003)		(0.002)		(0.004)	
Cohort birth year		0.002		0.002 +		-0.001		0.009 **
		(0.002)		(0.001)		(0.002)		(0.003)
Random effects								
Level 2 - Residual variance between studies in…								
Parent education intercepts	0.00693	0.00662	0.00754	0.00721	0.00498	0.00513	0.00647	0.00963
Parent occupation intercepts	0.00217	0.00212	0.00227	0.00239	0.00251	0.00239	0.00457	0.00589
Books intercepts	0.00413	0.00534	0.00353	0.00472	0.00274	0.00288	0.01014	0.01203
Level 3 - Residual variance between countries in…								
Parent education intercepts	0.06126	0.05959	0.05010	0.04805	0.06348	0.05186	0.04715	0.03904
Parent occupation intercepts	0.06315	0.05965	0.04929	0.05019	0.05750	0.05809	0.06038	0.04729
Books intercepts	0.10118	0.11315	0.08929	0.11116	0.13984	0.13857	0.04441	0.06193
Parent education cohort slopes	0.00006		0.00004		0.00002		0.00021	
Parent occupation cohort slopes	0.00007		0.00005		0.00003		0.00023	
Books cohort slopes	0.00015		0.00019		0.00002		0.00007	
Cohort slopes		0.00007		0.00006		0.00002		0.00005
% Parent education cohort slopes > 0	0.432		0.354		0.060		0.818	
% Parent occupation cohort slopes > 0	0.550		0.629		0.555		0.750	
% Books cohort slopes > 0	0.624		0.660		0.431		0.776	
% Cohort slopes > 0		0.586		0.621		0.440		0.900
N (Level 1 - gaps)	893	893	1030	1030	764	764	300	300
N (Level 2 - study-years)	298	298	344	344	255	255	103	103
N (Level 3 - countries)	70	70	72	72	70	70	41	41

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L). "% cohort slopes > 0" is computed by comparing cohort coefficient estimate and cohort slope variance component to a normal distribution.

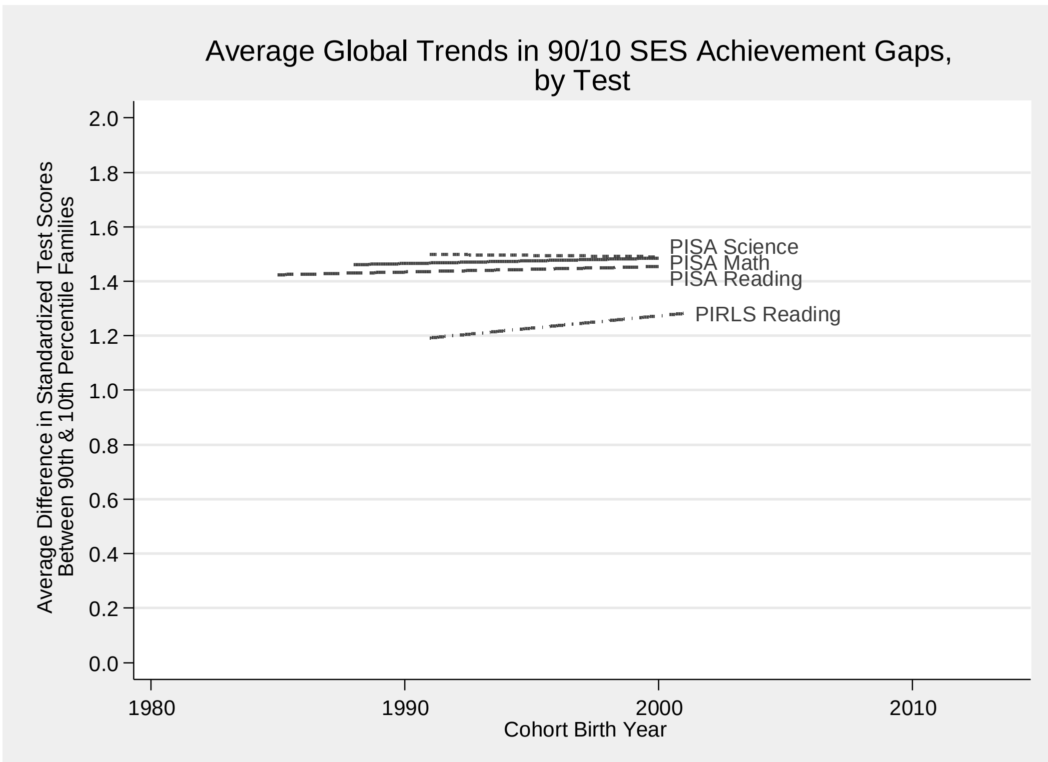
Table B4. Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th Percentiles of SES, Run Separately by Test Instrument (Standardized Achievement) – TIMSS

Test instrument	TIMSS Grade 4 Math		TIMSS Grade 4 Science		TIMSS Grade 8 Math		TIMSS Grade 8 Science	
Test years	1995-2015		1995-2015		1995-2015		1995-2015	
Cohort birth years	1985-2005		1985-2005		1981-2001		1981-2001	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Parent education gaps intercept					0.933 ***	0.934 ***	0.927 ***	0.918 ***
					(0.044)	(0.042)	(0.049)	(0.043)
Household books gaps intercept	1.011 ***	1.011 ***	1.046 ***	1.046 ***	1.066 ***	1.068 ***	1.100 ***	1.113 ***
	(0.068)	(0.068)	(0.077)	(0.077)	(0.052)	(0.053)	(0.058)	(0.061)
Cohort birth year × Parent education					0.012 ***		0.013 ***	
					(0.003)		(0.003)	
Cohort birth year × Books	0.013 ***		0.015 ***		0.013 ***		0.016 ***	
	(0.004)		(0.004)		(0.003)		(0.003)	
Cohort birth year		0.013 ***		0.015 ***		0.013 ***		0.014 ***
		(0.004)		(0.004)		(0.003)		(0.002)
Random effects								
Level 2 - Residual variance between studies in…								
Parent education intercepts					0.01337	0.01365	0.01571	0.01665
Books intercepts	0.00339	0.00339	0.01708	0.01708	0.01152	0.01331	0.01766	0.01863
Level 3 - Residual variance between countries in…								
Parent education intercepts					0.06882	0.06648	0.09139	0.07122
Books intercepts	0.14078	0.14078	0.16034	0.16034	0.11734	0.13454	0.15155	0.18369
Parent education cohort slopes					0.00011		0.00009	
Books cohort slopes	0.00026		0.00026		0.00034		0.00020	
Cohort slopes		0.00026		0.00026		0.00018		0.00009
% Parent education cohort slopes > 0					0.884		0.918	
% Books cohort slopes > 0	0.784		0.827		0.765		0.870	
% Cohort slopes > 0		0.784		0.827		0.829		0.931
N (Level 1 - gaps)	163	163	165	165	485	485	485	485
N (Level 2 - study-years)	163	163	165	165	245	245	245	245
N (Level 3 - countries)	49	49	50	50	61	61	61	61

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L). "% cohort slopes > 0" is computed by comparing cohort coefficient estimate and cohort slope variance component to a normal distribution.

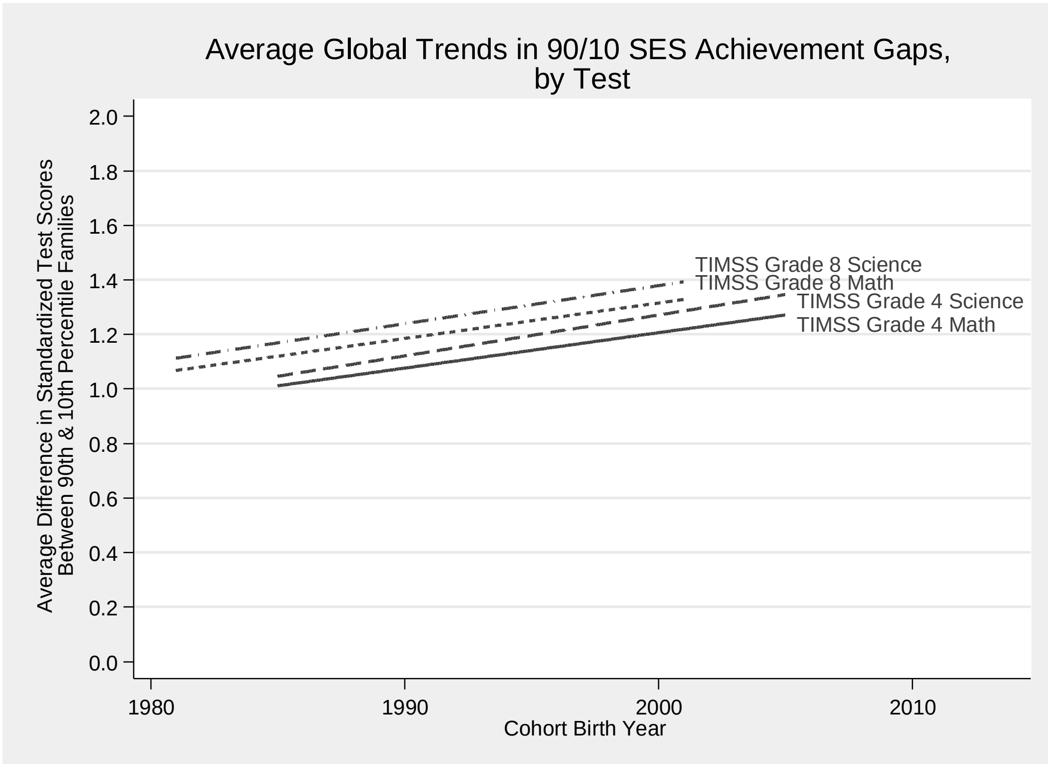
Figures B1 and B2 illustrate average global trends in 90/10 SES achievement gaps for each of the eight test instruments, estimated from the cohort birth year slope in Model 2 (based on standardized achievement). Figure B1 displays cohort slopes for the four instruments in Table B3, and Figure B2 displays cohort slopes for the four instruments in Table B4. The figures show that trends for all test instruments are positive except for PISA science, which is the PISA instrument that is available for the shortest span of cohort years. The trend in SES gaps for PISA science scores is very close to flat.

Figure B1



Note: Trend lines are estimates from Table B3, setting SES variable equal to parent education.

Figure B2



Note: Trend lines are estimates from Table B4, setting SES variable equal to parent education.

The lower 'Random effects' portions of Tables B1-B4 contain the residual country-level variance in cohort slopes, which quantifies how much country gap trends vary around the average global trend for each test instrument. By comparing the cohort coefficient estimate and cohort slope variance component from each model to a normal distribution, we can compute the estimated proportion of countries with positive gap trends. This is reported in each table below the random effects. Naturally, for all test instruments where average cohort coefficient estimates are positive, we estimate that a majority of countries have increasing gaps. Even for PISA, where negative average cohort coefficients indicate that a majority of countries have decreasing gaps, we still estimate that a substantial number of countries have increasing gaps.

To further understand cross-national variation by test instrument, Tables B5-B12 present trend estimates by country for each test. In each table, I present trends only for countries with at least 4 available study-years (except in PIRLS, where only 3 study-years are available). Even for this restricted sample, country trend estimates are highly unreliable due to the small number of observed years, which can be seen by the large discrepancies between OLS and shrunken empirical Bayes (EB) estimates. I prefer OLS estimates for these estimates based on very sparse country data because EB estimates tend to be biased toward the grand mean trend, and thus will underestimate cross-national variability in gap trends. Consistent with the model results reported above, for every test instrument except PISA science, a majority of countries have increasing gaps when achievement is standardized within country-study-year. When achievement is unstandardized, a smaller share of countries has positive gap trends, though still the majority for all tests except PISA.

Table B5. Estimated Change in 90/10 SES Achievement Gaps, PISA Math 2003-2015 (1988-2000 Birth Cohorts)

		Standardized Achievement			Unstandardized Achievement		
Country	N (study-years)	OLS cohort trend	EB cohort trend	(SE)	OLS cohort trend	EB cohort trend	(SE)
Mexico	5	-0.024	-0.011	(0.006)	-2.757	-1.470	(0.496)
Chile	4	-0.024	-0.006	(0.006)	-2.038	-0.919	(0.566)
Slovenia	4	-0.021	-0.002	(0.006)	-2.304	-0.395	(0.529)
Lithuania	4	-0.020	-0.003	(0.006)	-1.869	-0.423	(0.505)
Brazil	5	-0.019	-0.011	(0.005)	-3.255	-1.880	(0.469)
Germany	5	-0.019	-0.007	(0.005)	-3.114	-1.120	(0.470)
Netherlands	5	-0.017	-0.004	(0.005)	-1.540	-0.418	(0.474)
Liechtenstein	4	-0.014	-0.003	(0.007)	-1.749	-0.877	(0.627)
Qatar	4	-0.014	-0.006	(0.006)	-0.691	-1.130	(0.521)
Jordan	4	-0.013	-0.005	(0.006)	-1.339	-0.974	(0.524)
Tunisia	5	-0.013	-0.007	(0.006)	-1.391	-1.170	(0.525)
Thailand	5	-0.011	-0.004	(0.006)	-0.931	-0.809	(0.533)
Poland	5	-0.010	-0.002	(0.005)	-1.128	-0.444	(0.474)
Estonia	4	-0.009	-0.002	(0.006)	-1.333	-0.637	(0.494)
England	5	-0.008	0.000	(0.005)	-0.821	-0.182	(0.446)
Bulgaria	4	-0.005	-0.002	(0.007)	-1.440	-0.623	(0.593)
Norway	5	-0.005	0.002	(0.005)	-1.015	-0.191	(0.449)
United States	5	-0.002	0.003	(0.005)	-0.631	0.038	(0.439)
Montenegro	4	-0.001	0.002	(0.006)	-0.150	-0.180	(0.499)
Russian Fed.	5	-0.001	0.002	(0.005)	-0.458	-0.126	(0.448)
Portugal	5	-0.001	0.000	(0.005)	0.686	-0.160	(0.467)
Scotland	5	-0.001	0.004	(0.005)	-0.241	0.254	(0.462)
Slovak Rep.	5	0.000	0.003	(0.005)	0.354	0.086	(0.478)
Austria	5	0.001	0.004	(0.006)	-0.064	0.231	(0.525)
Switzerland	5	0.001	0.003	(0.005)	-0.411	-0.062	(0.500)
Japan	5	0.001	0.006	(0.005)	-0.569	0.246	(0.474)
Colombia	4	0.001	0.003	(0.006)	-0.705	-0.175	(0.502)
Turkey	5	0.002	0.007	(0.006)	-1.355	0.260	(0.493)
Greece	5	0.003	0.001	(0.005)	-0.291	-0.514	(0.456)
France	5	0.004	0.003	(0.005)	0.569	0.148	(0.452)
Romania	4	0.004	0.003	(0.007)	0.552	-0.005	(0.580)
Czech Rep.	5	0.005	0.002	(0.006)	-0.492	-0.425	(0.506)
Australia	5	0.006	0.005	(0.005)	0.390	0.057	(0.446)
Hong Kong	5	0.008	0.010	(0.005)	0.251	0.660	(0.466)
Canada	5	0.008	0.006	(0.005)	0.536	0.144	(0.423)
Ireland	5	0.009	0.008	(0.005)	0.212	0.350	(0.446)
Belgium-French	5	0.009	-0.001	(0.005)	-1.348	-1.170	(0.476)
Chinese Taipei	4	0.010	0.010	(0.006)	1.513	0.897	(0.516)
Uruguay	5	0.010	0.002	(0.005)	-0.665	-0.678	(0.482)
Finland	5	0.010	0.008	(0.005)	0.666	0.366	(0.438)
Serbia	4	0.011	0.004	(0.006)	1.657	0.013	(0.473)
Latvia	5	0.012	0.004	(0.005)	0.325	-0.302	(0.437)
Croatia	4	0.012	0.003	(0.006)	1.422	-0.209	(0.499)
Hungary	5	0.013	0.005	(0.006)	1.213	0.209	(0.499)
Indonesia	5	0.013	-0.002	(0.006)	0.424	-0.787	(0.485)
Sweden	5	0.014	0.007	(0.005)	0.740	0.111	(0.453)
Spain	5	0.014	0.009	(0.005)	0.786	0.468	(0.476)
Belgium-Flemish	5	0.014	0.001	(0.005)	0.826	-0.907	(0.485)
Korea, Rep.	5	0.015	0.013	(0.006)	1.925	1.200	(0.503)
Denmark	5	0.016	0.008	(0.005)	0.511	0.129	(0.450)
Iceland	5	0.016	0.011	(0.005)	1.329	0.702	(0.469)
Italy	5	0.019	0.012	(0.005)	1.365	0.602	(0.446)
New Zealand	5	0.020	0.006	(0.005)	1.343	-0.115	(0.493)
Israel	4	0.021	0.002	(0.006)	1.752	-0.669	(0.521)
Luxembourg	5	0.023	0.008	(0.005)	1.768	0.033	(0.470)
Macao-China	5	0.024	0.016	(0.005)	1.792	0.953	(0.459)

Note: Counties sorted by OLS cohort trend in gaps using standardized achievement. OLS and shrunken EB (empirical Bayes) cohort trends are derived from Model 2 in Table B1 (unstandardized achievement) and Table B3 (standardized achievement). Trends are shown only for countries with at least 4 available study-years.

Table B6. Estimated Change in 90/10 SES Achievement Gaps, PISA Reading 2000-2015 (1985-2000 Birth Cohorts)

		Standardized Achievement			Unstandardized Achievement		
Country	N (study-years)	OLS cohort trend	EB cohort trend	(SE)	OLS cohort trend	EB cohort trend	(SE)
Jordan	4	-0.026	-0.008	(0.006)	-2.469	-1.320	(0.772)
Mexico	6	-0.023	-0.014	(0.005)	-2.440	-1.950	(0.544)
Argentina	4	-0.021	-0.005	(0.006)	-3.797	-1.540	(0.728)
Chile	5	-0.017	-0.008	(0.005)	-2.358	-1.550	(0.617)
Scotland	6	-0.014	-0.004	(0.005)	-1.654	-0.917	(0.546)
Brazil	6	-0.013	-0.008	(0.005)	-1.297	-1.110	(0.563)
Germany	6	-0.012	-0.005	(0.005)	-2.927	-1.700	(0.599)
Slovenia	4	-0.011	0.000	(0.006)	-0.798	-0.141	(0.778)
England	6	-0.008	-0.003	(0.004)	-1.093	-0.649	(0.503)
Chinese Taipei	4	-0.007	0.004	(0.006)	0.572	0.828	(0.728)
United States	5	-0.007	-0.001	(0.005)	-1.197	-0.526	(0.534)
Tunisia	5	-0.007	-0.003	(0.006)	-1.123	-0.907	(0.678)
Poland	6	-0.006	-0.003	(0.005)	-1.653	-1.070	(0.546)
Switzerland	6	-0.006	-0.002	(0.005)	-0.965	-0.568	(0.570)
Netherlands	6	-0.003	0.001	(0.005)	0.687	0.531	(0.545)
Indonesia	6	-0.003	-0.003	(0.006)	-0.226	-0.427	(0.578)
Estonia	4	-0.002	-0.001	(0.006)	-0.468	-0.554	(0.736)
Thailand	6	-0.001	-0.002	(0.006)	0.118	-0.151	(0.628)
Belgium-French	6	-0.001	-0.002	(0.005)	-1.376	-1.160	(0.581)
Australia	6	-0.001	0.000	(0.004)	-0.129	-0.088	(0.492)
Serbia	4	0.001	0.002	(0.006)	1.565	0.494	(0.724)
Turkey	5	0.001	0.005	(0.006)	-0.805	-0.067	(0.670)
Czech Rep.	6	0.001	0.002	(0.005)	-0.204	-0.091	(0.604)
Lithuania	4	0.002	0.003	(0.006)	-0.249	0.067	(0.747)
Russian Fed.	6	0.003	0.003	(0.005)	0.162	0.171	(0.545)
Peru	4	0.003	-0.002	(0.005)	-0.211	-0.534	(0.585)
Liechtenstein	5	0.003	0.000	(0.007)	-0.845	-0.535	(0.829)
Korea, Rep.	6	0.003	0.006	(0.005)	1.218	1.120	(0.523)
Austria	6	0.004	0.005	(0.005)	0.482	0.518	(0.580)
Sweden	6	0.004	0.004	(0.005)	1.100	0.699	(0.577)
Qatar	4	0.005	-0.001	(0.006)	0.428	-0.443	(0.806)
Canada	6	0.005	0.005	(0.005)	0.322	0.393	(0.539)
Norway	6	0.005	0.005	(0.005)	0.006	0.181	(0.567)
Israel	5	0.005	0.001	(0.006)	0.557	-0.087	(0.691)
Hong Kong	6	0.006	0.007	(0.005)	0.398	0.570	(0.550)
France	6	0.007	0.005	(0.005)	2.238	1.600	(0.528)
Hungary	6	0.007	0.005	(0.005)	0.812	0.618	(0.581)
Spain	6	0.007	0.006	(0.004)	0.512	0.481	(0.492)
Luxembourg	6	0.008	0.005	(0.005)	1.322	0.801	(0.630)
Portugal	6	0.010	0.006	(0.005)	0.247	0.099	(0.537)
Japan	6	0.011	0.009	(0.005)	1.036	0.928	(0.563)
Slovak Rep.	5	0.011	0.006	(0.006)	2.651	1.210	(0.712)
New Zealand	6	0.012	0.006	(0.005)	0.931	0.440	(0.642)
Croatia	4	0.012	0.005	(0.006)	0.963	0.358	(0.727)
Belgium-Flemish	6	0.012	0.004	(0.005)	1.030	0.330	(0.614)
Greece	6	0.012	0.007	(0.005)	1.017	0.558	(0.530)
Finland	6	0.013	0.010	(0.005)	1.591	1.240	(0.527)
Iceland	6	0.013	0.010	(0.005)	1.238	1.040	(0.579)
Ireland	6	0.013	0.010	(0.004)	0.531	0.542	(0.499)
Colombia	4	0.014	0.006	(0.006)	0.078	0.084	(0.729)
Uruguay	5	0.014	0.005	(0.006)	-0.627	-0.497	(0.669)
Bulgaria	5	0.015	0.006	(0.006)	2.684	1.250	(0.690)
Denmark	6	0.016	0.011	(0.005)	0.762	0.618	(0.520)
Latvia	6	0.016	0.008	(0.005)	0.739	0.349	(0.542)
Italy	6	0.021	0.013	(0.005)	1.516	1.130	(0.550)
Montenegro	4	0.021	0.007	(0.006)	1.389	0.419	(0.740)
Macao-China	5	0.022	0.012	(0.006)	1.987	1.300	(0.646)
Romania	5	0.028	0.011	(0.006)	1.731	0.730	(0.707)

Note: Counties sorted by OLS cohort trend in gaps using standardized achievement. OLS and shrunken EB (empirical Bayes) cohort trends are derived from Model 2 in Table B1 (unstandardized achievement) and Table B3 (standardized achievement). Trends are shown only for countries with at least 4 available study-years.

Table B7. Estimated Change in 90/10 SES Achievement Gaps, PISA Science 2006-2015 (1991-2000 Birth Cohorts)

		Standardized Achievement			Unstandardized Achievement		
Country	N (study-years)	OLS cohort trend	EB cohort trend	(SE)	OLS cohort trend	EB cohort trend	(SE)
Mexico	4	-0.032	-0.008	(0.001)	-3.451	-1.530	(0.608)
Netherlands	4	-0.020	-0.001	(0.002)	-1.240	-0.496	(0.619)
United States	4	-0.019	0.002	(0.001)	-2.886	-0.628	(0.593)
Scotland	4	-0.018	0.004	(0.001)	-2.491	-0.448	(0.610)
Jordan	4	-0.017	-0.004	(0.002)	-1.981	-0.921	(0.613)
Iceland	4	-0.016	0.003	(0.002)	-1.953	-0.250	(0.667)
Chile	4	-0.015	-0.004	(0.002)	-1.558	-0.670	(0.647)
Germany	4	-0.014	-0.002	(0.002)	-1.790	-0.608	(0.615)
Thailand	4	-0.014	-0.004	(0.002)	-0.966	-0.658	(0.646)
Brazil	4	-0.014	-0.005	(0.002)	-1.431	-0.837	(0.632)
Uruguay	4	-0.012	-0.003	(0.002)	-1.978	-0.687	(0.624)
Poland	4	-0.009	0.000	(0.002)	-0.415	-0.050	(0.629)
Lithuania	4	-0.008	-0.001	(0.001)	-0.415	-0.244	(0.597)
Slovenia	4	-0.007	-0.001	(0.002)	-1.064	-0.397	(0.627)
Austria	4	-0.007	0.001	(0.002)	-1.094	-0.105	(0.653)
Bulgaria	4	-0.006	-0.004	(0.002)	-1.550	-0.500	(0.690)
England	4	-0.006	0.002	(0.001)	-1.354	-0.413	(0.612)
Estonia	4	-0.005	-0.002	(0.002)	-0.308	-0.599	(0.610)
Croatia	4	-0.004	-0.001	(0.002)	0.296	-0.134	(0.595)
Spain	4	-0.003	0.001	(0.002)	-0.472	-0.144	(0.627)
France	4	-0.003	0.000	(0.001)	-0.365	-0.085	(0.631)
Tunisia	4	-0.002	-0.006	(0.002)	-1.863	-1.090	(0.639)
Turkey	4	-0.001	0.005	(0.002)	-0.461	0.045	(0.622)
Sweden	4	-0.001	-0.001	(0.002)	0.730	-0.116	(0.633)
Australia	4	-0.001	0.002	(0.002)	0.099	0.104	(0.599)
Hong Kong	4	0.000	0.006	(0.001)	-0.868	0.014	(0.588)
Colombia	4	0.000	0.000	(0.002)	-0.194	-0.251	(0.596)
Latvia	4	0.000	-0.003	(0.002)	0.062	-0.523	(0.594)
Qatar	4	0.001	-0.005	(0.001)	1.014	-0.337	(0.620)
Norway	4	0.001	0.002	(0.002)	0.412	0.018	(0.641)
Denmark	4	0.001	0.000	(0.002)	-0.209	-0.159	(0.619)
Switzerland	4	0.001	-0.001	(0.002)	0.172	-0.106	(0.641)
Portugal	4	0.002	-0.003	(0.002)	0.637	-0.229	(0.587)
Romania	4	0.003	0.000	(0.002)	-0.342	-0.321	(0.638)
Greece	4	0.003	-0.003	(0.002)	0.095	-0.339	(0.612)
Luxembourg	4	0.005	-0.003	(0.002)	0.751	-0.080	(0.645)
Japan	4	0.005	0.006	(0.002)	-0.133	0.405	(0.614)
New Zealand	4	0.005	-0.003	(0.002)	-0.036	-0.438	(0.651)
Canada	4	0.006	0.004	(0.002)	0.391	0.186	(0.614)
Indonesia	4	0.006	-0.004	(0.002)	0.085	-0.412	(0.598)
Belgium-French	4	0.006	-0.004	(0.002)	-0.849	-0.808	(0.617)
Slovak Rep.	4	0.006	0.000	(0.002)	1.901	0.284	(0.637)
Montenegro	4	0.008	-0.001	(0.001)	0.664	-0.154	(0.592)
Chinese Taipei	4	0.009	0.005	(0.002)	1.045	0.786	(0.599)
Czech Rep.	4	0.011	-0.001	(0.002)	0.458	-0.141	(0.653)
Russian Fed.	4	0.012	0.000	(0.001)	0.190	-0.123	(0.582)
Ireland	4	0.013	0.004	(0.002)	0.444	0.324	(0.604)
Macao-China	4	0.014	0.007	(0.001)	1.033	0.411	(0.587)
Finland	4	0.015	0.003	(0.002)	2.442	0.609	(0.635)
Belgium-Flemish	4	0.015	-0.007	(0.001)	2.737	0.002	(0.615)
Korea, Rep.	4	0.018	0.007	(0.002)	1.899	0.825	(0.604)
Israel	4	0.020	-0.004	(0.002)	1.784	-0.064	(0.625)
Italy	4	0.021	0.003	(0.001)	1.494	0.481	(0.593)
Hungary	4	0.024	0.000	(0.002)	3.309	0.779	(0.627)

Note: Counties sorted by OLS cohort trend in gaps using standardized achievement. OLS and shrunken EB (empirical Bayes) cohort trends are derived from Model 2 in Table B1 (unstandardized achievement) and Table B3 (standardized achievement). Trends are shown only for countries with at least 4 available study-years.

Table B8. Estimated Change in 90/10 SES Achievement Gaps, PIRLS Reading 2001-2011 (1991-2001 Birth Cohorts)

		Standardized Achievement			Unstandardized Achievement		
Country	N (study-years)	OLS cohort trend	EB cohort trend	(SE)	OLS cohort trend	EB cohort trend	(SE)
Singapore	3	-0.018	0.007	(0.004)	-2.736	-0.694	(0.808)
Iran	3	-0.018	0.005	(0.005)	-2.327	-0.652	(0.843)
Netherlands	3	-0.012	0.006	(0.005)	-0.853	-0.283	(0.741)
Germany	3	-0.006	0.002	(0.005)	-0.263	-0.189	(0.796)
Norway	3	-0.002	0.005	(0.005)	-1.997	-1.090	(0.764)
Slovenia	3	0.005	0.012	(0.005)	-0.184	0.502	(0.782)
Italy	3	0.005	0.008	(0.005)	0.232	0.340	(0.757)
Hong Kong	3	0.006	0.010	(0.005)	0.476	0.469	(0.709)
Hungary	3	0.007	0.008	(0.005)	2.453	1.430	(0.850)
France	3	0.007	0.007	(0.005)	0.173	0.295	(0.773)
England	3	0.009	0.005	(0.005)	0.266	0.051	(0.993)
Bulgaria	3	0.013	0.010	(0.005)	1.002	0.764	(0.833)
Slovak Rep.	3	0.013	0.011	(0.005)	0.553	0.922	(0.833)
Sweden	3	0.014	0.006	(0.005)	0.725	0.232	(0.787)
Kuwait	3	0.018	0.019	(0.005)	3.356	2.390	(0.863)
New Zealand	3	0.020	0.003	(0.005)	1.190	0.047	(0.891)
Israel	3	0.021	0.019	(0.005)	0.911	1.660	(0.836)
Lithuania	3	0.022	0.009	(0.005)	1.697	0.924	(0.753)
Russian Fed.	3	0.042	0.014	(0.005)	2.711	1.520	(0.747)
Romania	3	0.053	0.009	(0.005)	4.896	1.990	(0.852)

Note: Counties sorted by OLS cohort trend in gaps using standardized achievement. OLS and shrunken EB (empirical Bayes) cohort trends are derived from Model 2 in Table B1 (unstandardized achievement) and Table B3 (standardized achievement). Trends are shown only for countries with at least 3 available study-years.

Table B9. Estimated Change in 90/10 SES Achievement Gaps, TIMSS Grade 4 Math 1995-2015 (1985-2005 Birth Cohorts)

		Standardized Achievement			Unstandardized Achievement		
Country	N (study-years)	OLS cohort trend	EB cohort trend	(SE)	OLS cohort trend	EB cohort trend	(SE)
New Zealand	5	-0.005	0.000	(0.009)	-0.076	0.319	(0.742)
Netherlands	5	-0.004	0.003	(0.011)	-0.678	-0.236	(0.723)
United States	5	-0.003	0.002	(0.010)	0.120	0.415	(0.795)
Chinese Taipei	4	-0.003	0.001	(0.008)	0.799	0.786	(0.696)
Singapore	5	0.002	0.003	(0.005)	-0.228	-0.055	(0.514)
Slovenia	5	0.009	0.009	(0.008)	-0.334	-0.043	(0.633)
Morocco	4	0.009	0.014	(0.014)	1.189	0.636	(1.075)
England	5	0.014	0.013	(0.009)	0.720	0.804	(0.769)
Czech Rep.	4	0.015	0.013	(0.012)	0.313	0.595	(0.915)
Norway	5	0.015	0.014	(0.009)	0.301	0.448	(0.715)
Russian Fed.	4	0.018	0.018	(0.008)	0.886	0.813	(0.705)
Lithuania	4	0.021	0.017	(0.011)	1.224	0.996	(0.872)
Australia	5	0.023	0.018	(0.011)	1.856	1.380	(0.847)
Iran	5	0.026	0.020	(0.009)	2.902	2.060	(0.767)
Japan	4	0.027	0.023	(0.008)	1.189	1.080	(0.676)
Hungary	5	0.033	0.029	(0.007)	3.884	3.160	(0.628)
Hong Kong	5	0.036	0.030	(0.009)	2.113	1.690	(0.652)
Italy	4	0.064	0.050	(0.009)	3.696	2.620	(0.725)

Note: Counties sorted by OLS cohort trend in gaps using standardized achievement. OLS and shrunken EB (empirical Bayes) cohort trends are derived from Model 2 in Table B2 (unstandardized achievement) and Table B4 (standardized achievement). Trends are shown only for countries with at least 4 available study-years.

Table B10. Estimated Change in 90/10 SES Achievement Gaps, TIMSS Grade 4 Science 1995-2015 (1985-2015 Birth Cohorts)

		Standardized Achievement			Unstandardized Achievement		
Country	N (study-years)	OLS cohort trend	EB cohort trend	(SE)	OLS cohort trend	EB cohort trend	(SE)
Netherlands	5	-0.013	0.000	(0.011)	-0.422	0.146	(0.762)
Singapore	5	-0.004	0.000	(0.009)	-1.110	-0.739	(0.682)
Slovenia	5	-0.002	0.005	(0.011)	-0.209	0.186	(0.778)
United States	5	0.004	0.008	(0.011)	0.019	0.210	(0.862)
Chinese Taipei	4	0.008	0.012	(0.011)	0.502	0.614	(0.874)
New Zealand	5	0.013	0.013	(0.011)	0.673	0.500	(0.869)
Czech Rep.	4	0.015	0.014	(0.012)	0.469	0.539	(0.848)
Australia	5	0.019	0.017	(0.012)	0.767	0.711	(0.945)
Iran	5	0.022	0.018	(0.010)	3.005	1.640	(0.792)
England	5	0.023	0.017	(0.011)	-0.062	0.138	(0.881)
Morocco	4	0.026	0.022	(0.014)	2.809	1.190	(1.090)
Norway	5	0.026	0.021	(0.011)	0.258	0.501	(0.852)
Hong Kong	5	0.032	0.027	(0.010)	2.170	1.810	(0.738)
Hungary	5	0.042	0.030	(0.010)	3.700	2.250	(0.774)
Russian Fed.	4	0.043	0.029	(0.012)	1.931	1.390	(0.906)
Lithuania	4	0.045	0.025	(0.013)	2.887	1.570	(0.974)
Japan	4	0.046	0.028	(0.012)	1.950	1.270	(0.897)
Italy	4	0.073	0.043	(0.012)	3.437	2.060	(0.887)

Note: Counties sorted by OLS cohort trend in gaps using standardized achievement. OLS and shrunken EB (empirical Bayes) cohort trends are derived from Model 2 in Table B2 (unstandardized achievement) and Table B4 (standardized achievement). Trends are shown only for countries with at least 4 available study-years.

Table B11. Estimated Change in 90/10 SES Achievement Gaps, TIMSS Grade 8 Math 1995-2015 (1981-2001 Birth Cohorts)

		Standardized Achievement			Unstandardized Achievement		
Country	N (study-years)	OLS cohort trend	EB cohort trend	(SE)	OLS cohort trend	EB cohort trend	(SE)
Chile	4	-0.024	-0.006	(0.009)	-2.150	-0.355	(0.841)
Russian Fed.	6	-0.018	-0.009	(0.007)	-1.083	-0.407	(0.729)
Jordan	5	-0.011	-0.003	(0.008)	-1.137	-0.043	(0.822)
Lebanon	4	-0.008	0.003	(0.010)	0.069	0.602	(0.900)
Cyprus	4	-0.004	0.006	(0.010)	-1.122	0.338	(1.077)
Botswana	4	-0.003	0.003	(0.011)	0.780	0.354	(1.022)
Saudi Arabia	4	-0.002	0.006	(0.011)	1.744	1.140	(0.965)
Slovenia	6	-0.002	0.003	(0.007)	-0.772	-0.280	(0.638)
Malaysia	5	0.000	0.006	(0.009)	0.767	0.922	(0.822)
Korea, Rep.	6	0.002	0.008	(0.008)	0.551	0.859	(0.786)
United States	6	0.003	0.007	(0.007)	-0.039	0.429	(0.695)
Indonesia	4	0.003	0.006	(0.011)	-0.784	0.209	(1.084)
Lithuania	6	0.005	0.007	(0.007)	0.285	0.670	(0.695)
Chinese Taipei	5	0.005	0.011	(0.009)	0.355	1.300	(0.896)
Sweden	5	0.012	0.014	(0.008)	0.800	0.853	(0.758)
Italy	6	0.016	0.015	(0.008)	0.103	0.446	(0.721)
Japan	5	0.017	0.015	(0.009)	2.211	1.960	(0.847)
South Africa	5	0.018	0.016	(0.008)	0.449	0.874	(0.813)
Hungary	6	0.018	0.016	(0.009)	2.577	2.380	(0.863)
Singapore	6	0.019	0.018	(0.007)	2.857	2.470	(0.687)
Bahrain	4	0.020	0.015	(0.009)	2.229	1.600	(0.906)
Norway	5	0.020	0.018	(0.008)	0.806	0.792	(0.741)
Australia	6	0.021	0.018	(0.008)	1.677	1.650	(0.738)
New Zealand	5	0.024	0.022	(0.008)	2.113	2.010	(0.742)
Thailand	5	0.024	0.019	(0.009)	2.169	1.870	(0.828)
Iran	6	0.024	0.021	(0.008)	3.290	2.710	(0.754)
Hong Kong	6	0.027	0.024	(0.007)	2.184	1.840	(0.679)
Israel	6	0.027	0.022	(0.007)	3.715	3.150	(0.734)
England	6	0.029	0.024	(0.008)	2.375	1.930	(0.802)
Tunisia	4	0.030	0.018	(0.011)	2.757	1.750	(0.980)
Canada	4	0.033	0.026	(0.007)	2.108	1.750	(0.714)
Morocco	5	0.034	0.021	(0.009)	2.576	1.570	(0.889)
Romania	5	0.046	0.033	(0.009)	5.213	3.830	(0.890)
Turkey	4	0.052	0.033	(0.009)	6.999	4.580	(0.927)

Note: Counties sorted by OLS cohort trend in gaps using standardized achievement. OLS and shrunken EB (empirical Bayes) cohort trends are derived from Model 2 in Table B2 (unstandardized achievement) and Table B4 (standardized achievement). Trends are shown only for countries with at least 4 available study-years.

Table B12. Estimated Change in 90/10 SES Achievement Gaps, TIMSS Grade 8 Science 1995-2015 (1982-2001 Birth Cohorts)

		Standardized Achievement			Unstandardized Achievement		
Country	N (study-years)	OLS cohort trend	EB cohort trend	(SE)	OLS cohort trend	EB cohort trend	(SE)
Lebanon	4	-0.017	0.007	(0.008)	-0.607	1.170	(0.931)
Cyprus	4	-0.017	0.006	(0.008)	-1.981	0.405	(0.954)
Botswana	4	-0.015	0.008	(0.009)	0.308	1.100	(0.982)
Chile	4	-0.013	0.004	(0.008)	-1.578	0.050	(0.803)
Jordan	5	-0.010	0.005	(0.008)	-0.600	0.575	(0.832)
Russian Fed.	6	0.000	0.007	(0.007)	-0.482	0.323	(0.741)
Korea, Rep.	6	0.001	0.011	(0.007)	0.013	0.504	(0.747)
Malaysia	5	0.002	0.010	(0.007)	2.537	1.970	(0.791)
United States	6	0.004	0.011	(0.007)	-0.469	0.272	(0.722)
Slovenia	6	0.004	0.009	(0.006)	0.423	0.747	(0.666)
Saudi Arabia	4	0.005	0.011	(0.008)	2.578	1.710	(0.917)
Indonesia	4	0.006	0.011	(0.009)	0.052	1.150	(0.962)
Singapore	6	0.008	0.011	(0.006)	1.732	1.640	(0.672)
Sweden	5	0.012	0.016	(0.008)	1.904	1.550	(0.785)
Chinese Taipei	5	0.012	0.015	(0.008)	1.014	1.220	(0.840)
Japan	5	0.014	0.015	(0.008)	1.383	1.340	(0.829)
South Africa	5	0.017	0.015	(0.007)	0.705	1.270	(0.815)
Hungary	6	0.017	0.014	(0.008)	1.983	1.720	(0.801)
Lithuania	6	0.019	0.016	(0.007)	1.419	1.440	(0.712)
Canada	4	0.019	0.018	(0.007)	0.937	1.100	(0.720)
Italy	6	0.020	0.018	(0.007)	0.809	0.997	(0.699)
Israel	6	0.020	0.016	(0.007)	2.496	2.160	(0.725)
New Zealand	5	0.021	0.019	(0.007)	2.141	1.840	(0.742)
Tunisia	4	0.023	0.016	(0.009)	1.479	1.470	(0.917)
Australia	6	0.024	0.019	(0.007)	1.530	1.520	(0.744)
Hong Kong	6	0.024	0.020	(0.007)	1.690	1.540	(0.689)
Norway	5	0.025	0.020	(0.007)	1.997	1.670	(0.768)
Morocco	5	0.025	0.017	(0.008)	2.092	1.620	(0.860)
England	6	0.027	0.020	(0.008)	1.716	1.390	(0.796)
Thailand	5	0.027	0.019	(0.008)	2.723	2.220	(0.747)
Bahrain	4	0.034	0.018	(0.008)	5.645	2.930	(0.911)
Iran	6	0.038	0.026	(0.007)	4.099	3.110	(0.705)
Turkey	4	0.047	0.024	(0.008)	5.975	3.490	(0.864)
Romania	5	0.048	0.027	(0.008)	2.854	2.060	(0.838)

Note: Counties sorted by OLS cohort trend in gaps using standardized achievement. OLS and shrunken EB (empirical Bayes) cohort trends are derived from Model 2 in Table B2 (unstandardized achievement) and Table B4 (standardized achievement). Trends are shown only for countries with at least 4 available study-years.

Country sampling practices in international large-scale assessments have generally improved over time, resulting in greater target population coverage in more recent study-years. This may tend to bias gap trend estimates upward, to the extent that early samples with poor coverage are positively selected, resulting in smaller gaps in older cohorts. As population coverage information is inconsistently provided in early studies, I retain all available data in all analyses in the main paper and throughout these appendices, as I do not wish to possibly bias results by inappropriately excluding data.

However, Tables B13, B14, and B15 report analyses to test the robustness of results to this decision. Table B13 shows gap trends for each PISA or PIRLS test instrument, after dropping 13 country-study-years in PISA and 8 country-study-years in PIRLS that the OECD or IEA consider not comparable for trend analysis. Table B14 shows gap trends for each TIMSS test instrument, after dropping 18 country-study-years in TIMSS 4th grade and 36 country-study-years in TIMSS 8th grade that the IEA considers not comparable for trend analysis. (In each model, I additionally drop any countries that have fewer than 2 remaining observations after noncomparable years are dropped.) The IEA states that noncomparability is usually due to improved translations or increased population coverage. I draw the list of noncomparable country-study-years primarily from tables in OECD and IEA reports listing which previous country-years are not considered comparable to the most recent cycle of each instrument: OECD (2016) Annex B1.4 Table I.4.4a for PISA, Mullis et al. (2012) Appendix A.1 for PIRLS, and Mullis et al. (2016) Appendix A.1 for TIMSS. For countries that did not participate in the most recent testing cycle, I supplement these lists by additionally dropping any country-study-years that were excluded from trend reports for previous cycles of each instrument.

Comparing Tables B13 and B14 to models for the full samples of PISA, PIRLS, and TIMSS country-study-years in Tables B3 and B4, results are similar, with large and significant positive cohort coefficients for PIRLS and TIMSS and smaller but usually positive cohort coefficients for PISA. If gap trends are biased upward by including early studies with poor population coverage, we would expect cohort coefficients to be smaller in Tables B13 and B14 than in Tables B3 and B4. However, this is not consistently the case. For PISA, PIRLS, and TIMSS 4th grade, coefficients are generally *larger* than those for the full samples in Tables B3 and B4 by -2% to +19% for PISA, +4% for PIRLS, and +12% to +17% for TIMSS 4th grade. This implies that excluding years with poor population coverage may bias trends for these instruments *downward.* Only for TIMSS 8th grade are coefficients for models excluding non-comparable years smaller than those for full samples, by -27% to -35%, implying that excluding years with poor population coverage may bias trends for TIMSS 8th grade upward.

In order to see how these changes may affect results from pooled models in the main paper, Table B15 pools all observations from Tables B13 and B14 in a multivariate variance-known model with the same specification as that in the main paper. Because the population coverage in early studies is poorly documented in some cases, I drop all studies from test years before 1995 (corresponding to cohort birth years before approximately 1981). For comparison, Models B1 and B2 include the full sample of all observations from test years in 1995 and later, while Models B3 and B4 drop any observations excluded from Tables B13 and B14. Model B1 estimates separate cohort birth year coefficients for each gap type (parent education, occupation, and books), using the same specification as Model 1 (Table 1) in the main paper, but including only recent cohorts rather than the full sample of cohort years (1950-2005) included in Model 1. Comparing Model B1 to Model 1 reveals that the parent occupation achievement gap trend in recent cohorts is identical to the trend over the full time span, while the trend for the books gap is increasing at a somewhat faster rate in recent cohorts and the parent education gap is increasing at a much slower rate. The smaller coefficient for parent education gaps is attributable to the large share of recent data that come from PISA, where the parent education gap is declining on average (see Table B3). Nevertheless, in Model B2, which estimates a single cohort birth year coefficient pooling all gap types, the pooled cohort birth year coefficient is 0.007, identical to the coefficient for the model including all cohort years in the main paper (Model 2). This suggests that, on average across all SES variables, the SES achievement gap increases at a roughly constant rate over the full span of cohort birth years.

We turn now to Models B3 and B4, which drop any observations the OECD or IEA consider non-comparable, and compare to Models B1 and B2. In Model B3, the parent education cohort coefficient remains the same as in Model B1, and the parent occupation and books cohort coefficients both decrease slightly. In Model B4, the pooled cohort coefficient also decreases slightly from 0.007 to 0.006. This 14% decline translates into a total gap increase over 24 cohort years (1981-2005) of 0.14 SD rather than 0.17 SD. If we assume a similar upward bias of the cohort coefficient over the full span of 55 cohort years (1950-2005), this implies a total gap increase of 0.33 SD rather than the 0.39 SD reported in the main paper. I consider 0.33 SD still a substantial increase in inequality. Thus, I conclude that, while trends reported in the main analyses of this paper are likely biased upward by the inclusion of some country-study-years with poor population coverage, the magnitude of the bias is likely small and does not substantially change the overall conclusion that the SES achievement gap has increased markedly on average across most countries. As it is difficult to apply a single exclusion rule consistently across the full sample of 1026 country-study-years due to inconsistencies in documentation—particularly in early years—I choose to retain all available data in all other analyses.

Table B13. Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th Percentiles of SES, Run Separately by Test Instrument (Standardized Achievement) – PISA and PIRLS, dropping country-study-years that the OECD or IEA consider not comparable for trend analysis

Test instrument	PISA Math		PISA Reading		PISA Science		PIRLS Reading	
Test years	2003-2015		2000-2015		2006-2015		2001-2011	
Cohort birth years	1988-2000		1985-2000		1991-2000		1991-2001	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Parent education gaps intercept	0.998 ***	0.970 ***	0.946 ***	0.907 ***	1.041 ***	0.978 ***	1.064 ***	1.074 ***
	(0.037)	(0.034)	(0.033)	(0.030)	(0.041)	(0.034)	(0.043)	(0.043)
Parent occupation gaps intercept	0.970 ***	0.964 ***	0.919 ***	0.922 ***	0.952 ***	0.967 ***	0.955 ***	0.947 ***
	(0.034)	(0.033)	(0.029)	(0.029)	(0.034)	(0.034)	(0.044)	(0.043)
Household books gaps intercept	1.444 ***	1.457 ***	1.407 ***	1.427 ***	1.505 ***	1.504 ***	1.229 ***	1.229 ***
	(0.045)	(0.043)	(0.041)	(0.042)	(0.050)	(0.048)	(0.045)	(0.043)
Cohort birth year × Parent education	-0.001		-0.002		-0.007 **		0.011 **	
	(0.002)		(0.002)		(0.003)		(0.004)	
Cohort birth year × Parent occupation	0.001		0.003 *		0.001		0.008 *	
	(0.002)		(0.001)		(0.002)		(0.003)	
Cohort birth year × Books	0.003		0.006 *		-0.001		0.010 *	
	(0.003)		(0.003)		(0.002)		(0.004)	
Cohort birth year		0.002		0.003 *		-0.001		0.009 **
		(0.002)		(0.001)		(0.002)		(0.003)
Random effects								
Level 2 - Residual variance between studies in…								
Parent education intercepts	0.00683	0.00659	0.00748	0.00745	0.00494	0.00510	0.01150	0.01205
Parent occupation intercepts	0.00216	0.00216	0.00241	0.00228	0.00251	0.00239	0.00672	0.00707
Books intercepts	0.00394	0.00519	0.00339	0.00459	0.00275	0.00289	0.01140	0.01319
Level 3 - Residual variance between countries in…								
Parent education intercepts	0.06179	0.06001	0.05165	0.04934	0.06203	0.05112	0.03823	0.04152
Parent occupation intercepts	0.06322	0.05965	0.04893	0.05062	0.05757	0.05803	0.05124	0.04873
Books intercepts	0.10804	0.11388	0.09177	0.11193	0.13861	0.13761	0.04356	0.04332
Parent education cohort slopes	0.00006		0.00004		0.00002		0.00004	
Parent occupation cohort slopes	0.00008		0.00003		0.00003		0.00006	
Books cohort slopes	0.00016		0.00020		0.00002		0.00011	
Cohort slopes		0.00007		0.00005		0.00002		0.00003
% Parent education cohort slopes > 0	0.433		0.365		0.058		0.963	
% Parent occupation cohort slopes > 0	0.553		0.702		0.557		0.868	
% Books cohort slopes > 0	0.604		0.652		0.438		0.829	
% Cohort slopes > 0		0.584		0.654		0.446		0.967
N (Level 1 - gaps)	881	881	1000	1000	761	761	264	264
N (Level 2 - study-years)	294	294	334	334	254	254	91	91
N (Level 3 - countries)	70	70	71	71	70	70	36	36

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L). "% cohort slopes > 0" is computed by comparing cohort coefficient estimate and cohort slope variance component to a normal distribution. See OECD (2016) and Mullis et al. (2012) Appendix A.1 for lists of excluded country-study-years.

Table B14. Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th Percentiles of SES, Run Separately by Test Instrument (Standardized Achievement) – TIMSS, dropping country-study-years that IEA considers not comparable for trend analysis

Test instrument	TIMSS Grade 4 Math		TIMSS Grade 4 Science		TIMSS Grade 8 Math		TIMSS Grade 8 Science	
Test years	1995-2015		1995-2015		1995-2015		1995-2015	
Cohort birth years	1985-2005		1985-2005		1981-2001		1981-2001	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Parent education gaps intercept					0.995 ***	1.010 ***	0.998 ***	0.987 ***
					(0.048)	(0.045)	(0.052)	(0.047)
Household books gaps intercept	0.989 ***	0.989 ***	1.020 ***	1.020 ***	1.158 ***	1.150 ***	1.153 ***	1.173 ***
	(0.071)	(0.071)	(0.082)	(0.082)	(0.049)	(0.051)	(0.060)	(0.061)
Cohort birth year × Parent education					0.009 ***		0.009 ***	
					(0.003)		(0.003)	
Cohort birth year × Books	0.014 ***		0.018 ***		0.008 *		0.013 ***	
	(0.003)		(0.004)		(0.004)		(0.003)	
Cohort birth year		0.014 ***		0.018 ***		0.008 **		0.010 ***
		(0.003)		(0.004)		(0.003)		(0.003)
Random effects								
Level 2 - Residual variance between studies in…								
Parent education intercepts					0.00973	0.01024	0.00885	0.00914
Books intercepts	0.00358	0.00358	0.01736	0.01736	0.00870	0.01029	0.01215	0.01345
Level 3 - Residual variance between countries in…								
Parent education intercepts					0.07064	0.06875	0.09248	0.07872
Books intercepts	0.14023	0.14023	0.16781	0.16781	0.08176	0.10544	0.13615	0.16855
Parent education cohort slopes					0.00007		0.00006	
Books cohort slopes	0.00020		0.00018		0.00034		0.00024	
Cohort slopes		0.00020		0.00018		0.00017		0.00012
% Parent education cohort slopes > 0					0.849		0.874	
% Books cohort slopes > 0	0.840		0.906		0.668		0.795	
% Cohort slopes > 0		0.840		0.906		0.739		0.832
N (Level 1 - gaps)	149	149	149	149	411	411	411	411
N (Level 2 - study-years)	149	149	149	149	208	208	208	208
N (Level 3 - countries)	46	46	46	46	58	58	58	58

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L). "% cohort slopes > 0" is computed by comparing cohort coefficient estimate and cohort slope variance component to a normal distribution. See Mullis et al. (2016) Appendix A.1 for list of excluded country-study-years.

Table B15. Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th Percentiles of SES, Test Years 1995-2015 (Cohort Birth Years 1981-2005)

	Since Test Year 1995 (Birth Year 1981) Full Sample				Since Test Year 1995 (Birth Year 1981) Omitting Study-Years with Low Coverage			
	(B1) 3 Cohort Slopes		(B2) 1 Cohort Slope		(B3) 3 Cohort Slopes		(B4) 1 Cohort Slope	
	coef	(se)	coef	(se)	coef	(se)	coef	(se)
Parent education gaps intercept	1.075	(0.033) ***	1.055	(0.032) ***	1.089	(0.034) ***	1.07	(0.032) ***
Parent occupation gaps intercept	0.928	(0.035) ***	0.954	(0.035) ***	0.943	(0.036) ***	0.968	(0.036) ***
Household books gaps intercept	1.379	(0.051) ***	1.361	(0.049) ***	1.409	(0.053) ***	1.395	(0.052) ***
Level 1 - Gaps								
Subject (ref=Reading):								
Math	0.022	(0.007) **	0.022	(0.007) **	0.023	(0.007) **	0.023	(0.007) **
Science	0.036	(0.005) ***	0.036	(0.005) ***	0.035	(0.005) ***	0.035	(0.005) ***
SES variable quality measures								
Parent-reported × Parent education	0.038	(0.009) ***	0.024	(0.008) **	0.039	(0.010) ***	0.026	(0.009) **
Parent-reported × Parent occupation	0.191	(0.030) ***	0.155	(0.030) ***	0.191	(0.035) ***	0.157	(0.033) ***
Parent-reported × Books	0.071	(0.024) **	0.073	(0.023) **	0.069	(0.024) **	0.072	(0.024) **
Number of categories (centered at 7)	-0.021	(0.027)	0.002	(0.026)	-0.048	(0.029) +	-0.027	(0.027)
≥ 20% in bottom category	-0.070	(0.024) **	-0.065	(0.026) *	-0.063	(0.025) *	-0.061	(0.027) *
≥ 20% in top category	-0.109	(0.013) ***	-0.116	(0.013) ***	-0.098	(0.014) ***	-0.102	(0.014) ***
Level 2 - Study-years								
Age at testing (ref=14)								
Age 10 at testing	-0.220	(0.028) ***	-0.210	(0.027) ***	-0.249	(0.032) ***	-0.241	(0.031) ***
Age 15 at testing	-0.068	(0.022) **	-0.060	(0.022) **	-0.083	(0.024) ***	-0.077	(0.024) **
Cohort birth year × Parent education	0.002	(0.002)			0.002	(0.002)		
Cohort birth year × Parent occupation	0.007	(0.001) ***			0.006	(0.002) ***		
Cohort birth year × Books	0.010	(0.002) ***			0.009	(0.002) ***		
Cohort birth year			0.007	(0.001) ***			0.006	(0.001) ***
Random effects								
Level 2 - Residual variance between studies in…								
Parent education intercepts	0.03097		0.03180		0.02727		0.02782	
Parent occupation intercepts	0.01240		0.01282		0.01125		0.01135	
Books intercepts	0.02825		0.03110		0.02221		0.02444	
Level 3 - Residual variance between countries in…								
Parent education intercepts	0.05239		0.05253		0.05474		0.05151	
Parent occupation intercepts	0.05556		0.05463		0.05603		0.05515	
Books intercepts	0.13397		0.13382		0.14643		0.14746	
Parent education cohort slopes	0.00006				0.00005			
Parent occupation cohort slopes	0.00006				0.00006			
Books cohort slopes	0.00013				0.00015			
Cohort slopes			0.00005				0.00007	
N (Level 1 - gaps)	5541		5541		4652		4652	
N (Level 2 - study-years)	1026		1026		752		752	
N (Level 3 - countries)	100		100		92		92	

+ $p<.1$, * $p<.05$, ** $p<.01$, *** $p < .001$. Note: All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper. See OECD (2016) and Mullis et al. (2012, 2016) Appendix A.1 for lists of excluded country-study-years in Models B3 and B4.

Tables B16 and B17 show a comparison of gap trend estimates from the models above and those from Broer et al. (forthcoming). There are a number of methodological differences between the analyses reported in the paper and those in Broer et al. (forthcoming). First, they classify trends as “increasing” or “decreasing” based only on the contrast between 1995 and 2015 gaps, rather than estimating a linear trend line. The column labeled “Broer diff./20” reproduces those reported differences, dividing by 20 to yield the estimated annual change in gaps. The column labeled “Broer coef.” contains estimates that I computed by regressing gap estimates from all waves reported by Broer et al. on test year. Third, they do not standardize achievement within waves. The columns labeled “Unstandardized Achievement” also keep scores in the original TIMSS scale. Fourth, they exclude two country-study-years that the IEA does not consider comparable for trend analysis: TIMSS 1999 for Australia and Slovenia. These two country-study-years are omitted from all trends reported in Tables B16 and B17. They also exclude TIMSS 2015 Lithuanian students who were assessed in languages other than Lithuanian. In Tables B16 and B17, I exclude the entire TIMSS 2015 Lithuanian sample from Broer et al.’s and my estimates. After making all the adjustments above, the trend estimates in the “Broer coef.” column and the “Unstandardized Achievement—OLS cohort trend” column are fairly similar. They have a correlation of 0.87 for both math and science. The remaining small discrepancies between these two columns are due to the different measurement of SES and computation of achievement gaps. Broer et al. compute gaps between the top and bottom quartile of an SES index composed from parent education, books, and two other household possessions (a computer and a desk for studying). Broer et al. also do not adjust gaps according to the estimated reliability of SES reports, as I do in this study.

Table B16. Comparing Estimated Change in SES Achievement Gaps in Broer et al. and Current Study, TIMSS Grade 8 Math 1995-2015 (1982-2001 Birth Cohorts)

				Unstandardized Achievement			Standardized Achievement		
Country	N (study-years)	Broer diff./20[1]	Broer coef.[2]	OLS cohort trend	EB cohort trend (SE)		OLS cohort trend	EB cohort trend (SE)	
Australia	5	-1.205	-0.284	1.255	1.420	(0.805)	0.016	0.015	(0.009)
Hong Kong	6	0.802	1.357	2.199	1.840	(0.684)	0.027	0.024	(0.007)
Hungary	6	3.403	1.721	2.584	2.390	(0.864)	0.018	0.015	(0.009)
Iran	6	5.108	3.414	3.285	2.690	(0.756)	0.024	0.020	(0.008)
Korea, Rep.	6	-0.856	0.271	0.549	0.861	(0.788)	0.002	0.008	(0.008)
Lithuania	5	-0.063	-0.325	0.100	0.687	(0.862)	0.002	0.006	(0.009)
New Zealand	5	1.212	0.363	2.119	2.010	(0.745)	0.024	0.022	(0.008)
Norway	5	-1.102	-0.432	0.803	0.789	(0.745)	0.020	0.018	(0.008)
Russian Fed.	6	-1.338	-0.571	-1.099	-0.402	(0.732)	-0.018	-0.009	(0.007)
Singapore	6	3.125	1.986	2.865	2.460	(0.691)	0.019	0.018	(0.007)
Slovenia	5	-1.183	-0.422	-0.719	-0.163	(0.680)	-0.005	0.001	(0.007)
Sweden	5	-0.116	-0.193	0.801	0.852	(0.762)	0.012	0.014	(0.008)
United States	6	-0.871	-0.679	-0.034	0.443	(0.700)	0.003	0.007	(0.007)

[1] Broer diff./20 is derived from difference estimates reported Broer et al. (forthcoming) Table 4.1., divided by 20. Lithuanian difference is divided by 16 (2011-1995).

[2] I computed Broer coefficients by regressing gap estimates reported in Broer et al. (forthcoming) Table 4.2 on test year.

Note: OLS and shrunken EB (empirical Bayes) cohort trends are derived from Model 2 in Table B2 (unstandardized achievement) and Table B4 (standardized achievement), after omitting TIMSS 1999 for Australia and Slovenia and TIMSS 2015 for Lithuania.

Table B17. Comparing Estimated Change in SES Achievement Gaps in Broer et al. and Current Study, TIMSS Grade 8 Science 1995-2015 (1982-2001 Birth Cohorts)

				Unstandardized Achievement			Standardized Achievement		
Country	N (study-years)	Broer diff./20[1]	Broer coef.[2]	OLS cohort trend	EB cohort trend	(SE)	OLS cohort trend	EB cohort trend	(SE)
Australia	5	-1.123	-0.226	1.398	1.460	(0.803)	0.021	0.018	(0.008)
Hong Kong	6	-0.288	0.607	1.690	1.530	(0.694)	0.024	0.020	(0.007)
Hungary	6	2.562	1.571	1.984	1.720	(0.805)	0.017	0.014	(0.008)
Iran	6	5.308	3.314	4.099	3.100	(0.710)	0.038	0.026	(0.007)
Korea, Rep.	6	-0.288	0.057	0.012	0.500	(0.751)	0.001	0.010	(0.007)
Lithuania	5	0.313	0.150	1.123	1.270	(0.837)	0.018	0.015	(0.008)
New Zealand	5	1.093	0.291	2.137	1.830	(0.746)	0.021	0.019	(0.007)
Norway	5	0.538	0.730	1.997	1.660	(0.774)	0.025	0.020	(0.008)
Russian Fed.	6	-1.074	-0.400	-0.482	0.325	(0.746)	0.000	0.007	(0.007)
Singapore	6	0.786	0.879	1.731	1.640	(0.677)	0.008	0.010	(0.007)
Slovenia	5	0.045	0.139	0.329	0.724	(0.713)	-0.002	0.007	(0.007)
Sweden	5	0.702	0.530	1.902	1.540	(0.790)	0.012	0.015	(0.008)
United States	6	-1.869	-1.079	-0.468	0.270	(0.727)	0.004	0.010	(0.007)

[1] Broer diff./20 is derived from difference estimates reported Broer et al. (forthcoming) Table 4.1., divided by 20. Lithuanian difference is divided by 16 (2011-1995).
[2] I computed Broer coefficients by regressing gap estimates reported in Broer et al. (forthcoming) Table 4.3 on test year.
Note: OLS and shrunken EB (empirical Bayes) cohort trends are derived from Model 2 in Table B2 (unstandardized achievement) and Table B4 (standardized achievement), after omitting TIMSS 1999 for Australia and Slovenia and TIMSS 2015 for Lithuania.

Appendix C. Changing distribution of achievement

The models in the main text of the paper standardize achievement within each country-study-year before computing SES achievement gaps, both out of necessity since they combine different test instruments and also out of a theoretical preference for treating achievement as a positional good. Thus, trends in gaps are unaffected by changes in the variance of student achievement; they reflect changes in the relative *strength* of the SES-achievement association rather than the absolute *size* of the association. However, it is also interesting to ask whether the absolute size of the SES-achievement association is changing over time, as gaps in unstandardized achievement are assumed to reflect gaps in meaningful academic skills, rather than gaps in students' position within the achievement distribution.

In contrast to the main analyses that pool different studies, in separate analyses of trend studies (PISA, TIMSS, PIRLS), it is possible to estimate whether the variance of student achievement has changed over time. Figures C1-C8 display score variance at the student, school/classroom, and country levels for each year of each trend study, estimated from separate three-level hierarchical models, as follows:

$$\hat{T}_{ij} = \gamma_{00} + v_k + u_{jk} + \epsilon_{ijk},$$
$$v_k \sim N(0, \tau_{000}); u_{jk} \sim N(0, \tau_{00}); \;\; \epsilon_{ijk} \sim N(0, \sigma^2),$$

[C1]

where $\hat{T}_{ijk}$ is the estimated test score of student i in school or classroom j in country k, τ_{000} is the between-country variance of scores, τ_{00} is the between-school variance of scores, and σ^2 is the within-school student-level variance of scores. Total student weights are applied at the student level, meaning all students are weighted in proportion to their probability of selection, and all countries are weighted in proportion to the size of their target population (i.e., more populous countries receive greater weight). Models are estimated once for each plausible value and averaged. Only countries that participated in all years of a given trend study are included. Samples of included countries vary depending on the study. After estimating the student-, school/classroom-, and country-level variances, all three are adjusted for estimated test reliability (α) for the relevant set of countries in each year, as follows:

$$\sigma^2_{True} = \; \alpha * \sigma^2_{Total}$$

[C2]

Figure C1

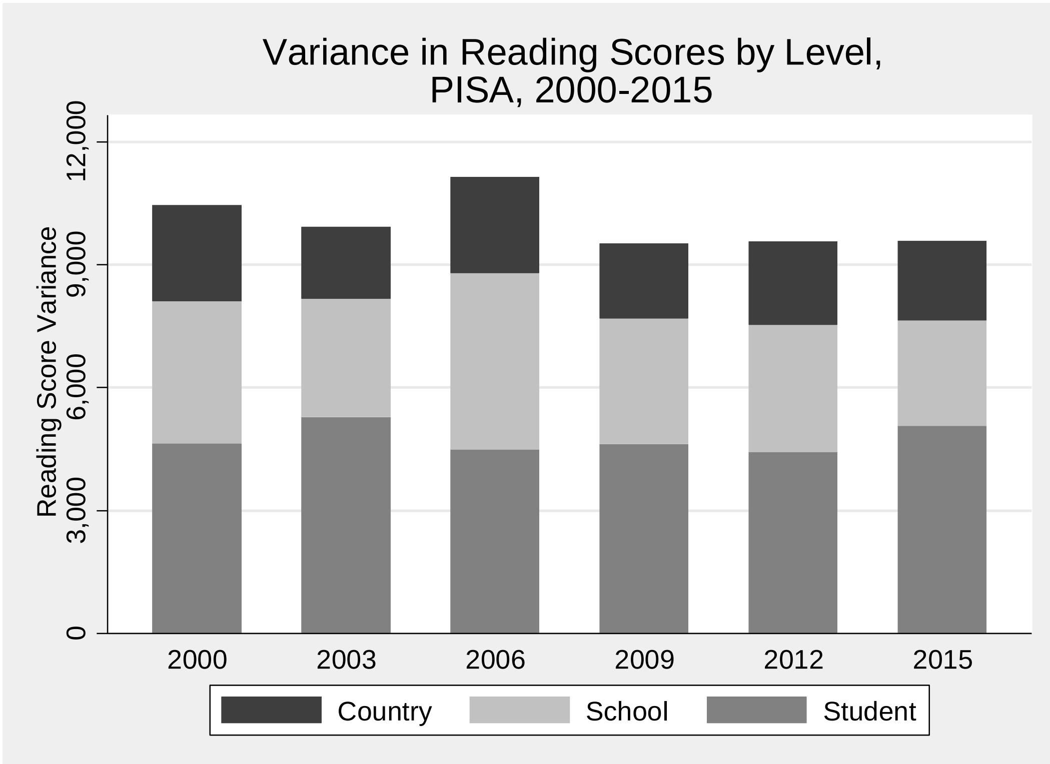


Figure C2

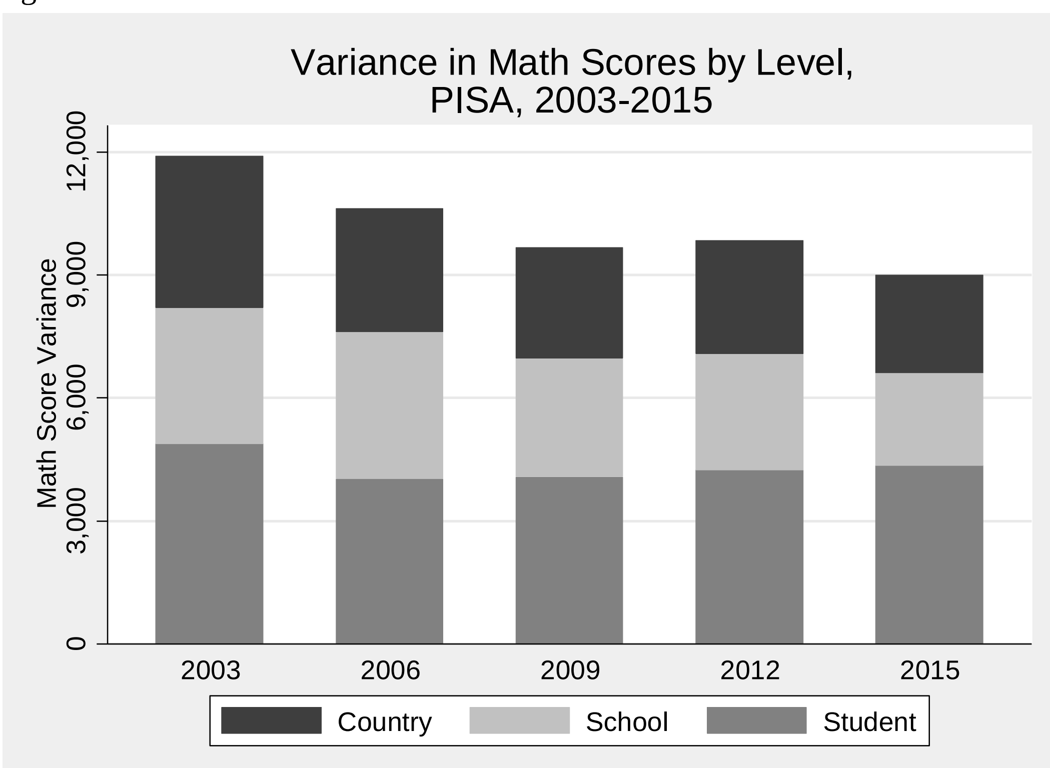


Figure C3

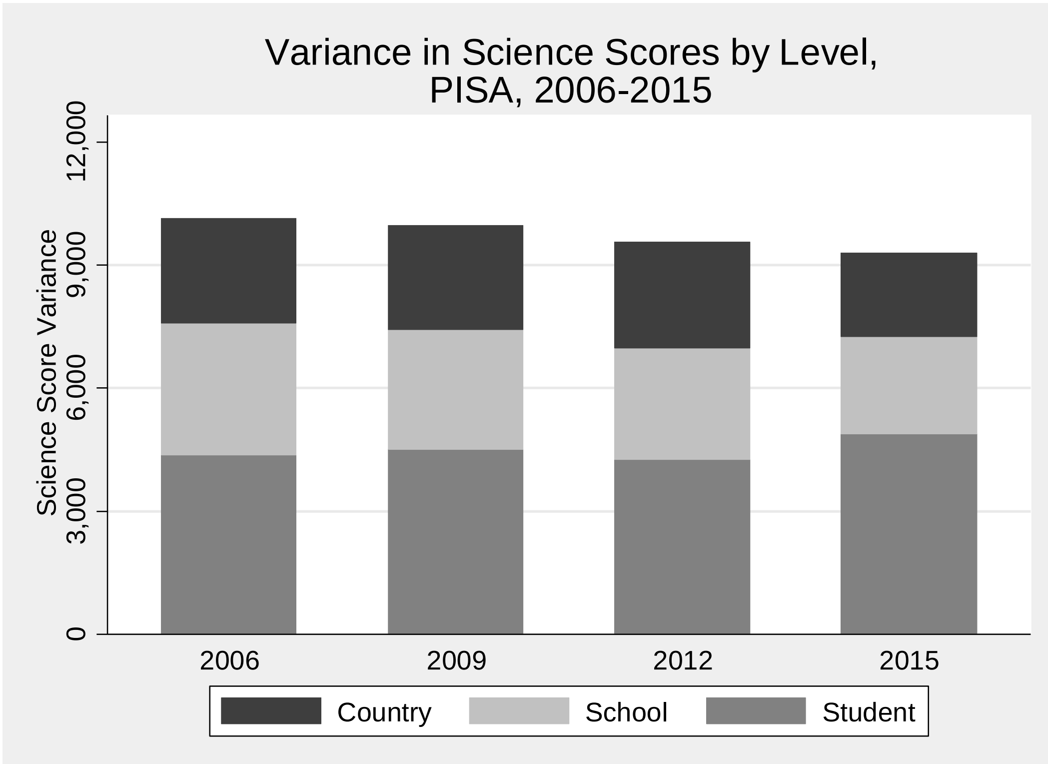


Figure C4

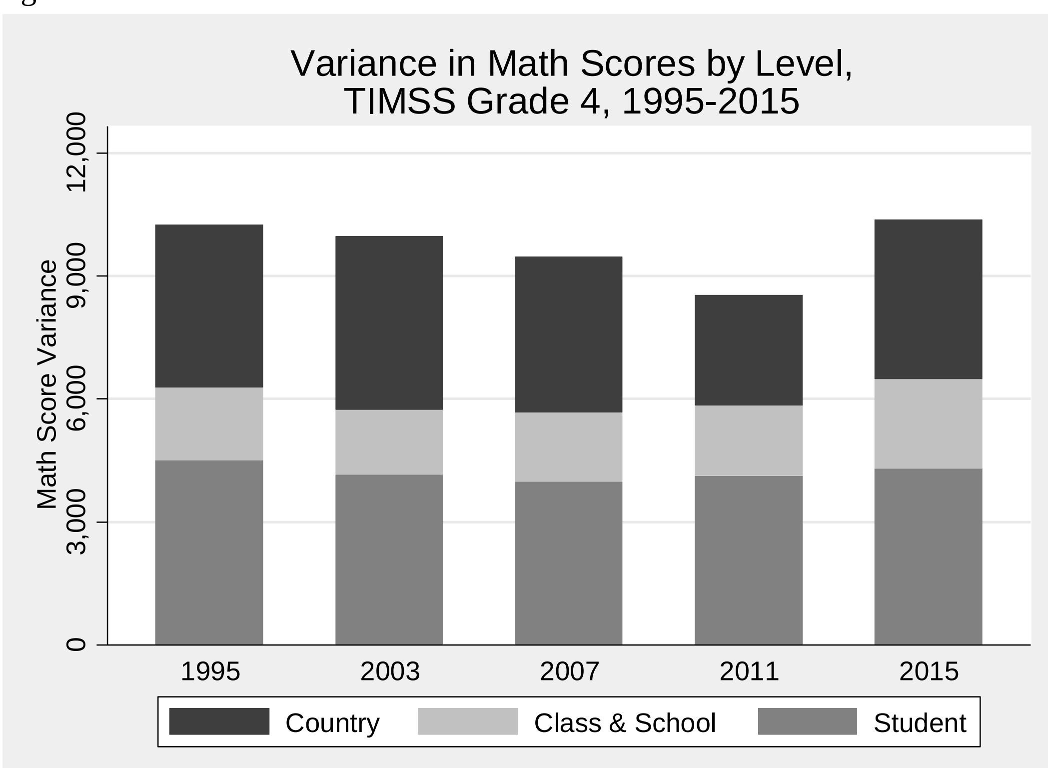


Figure C5

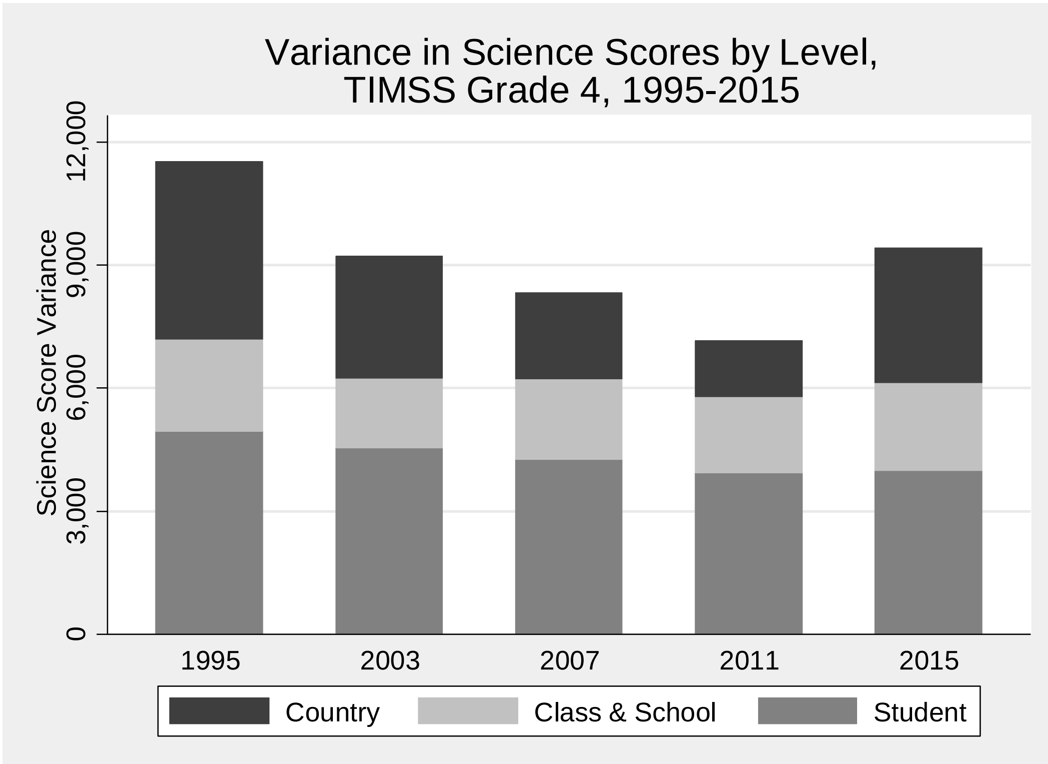


Figure C6

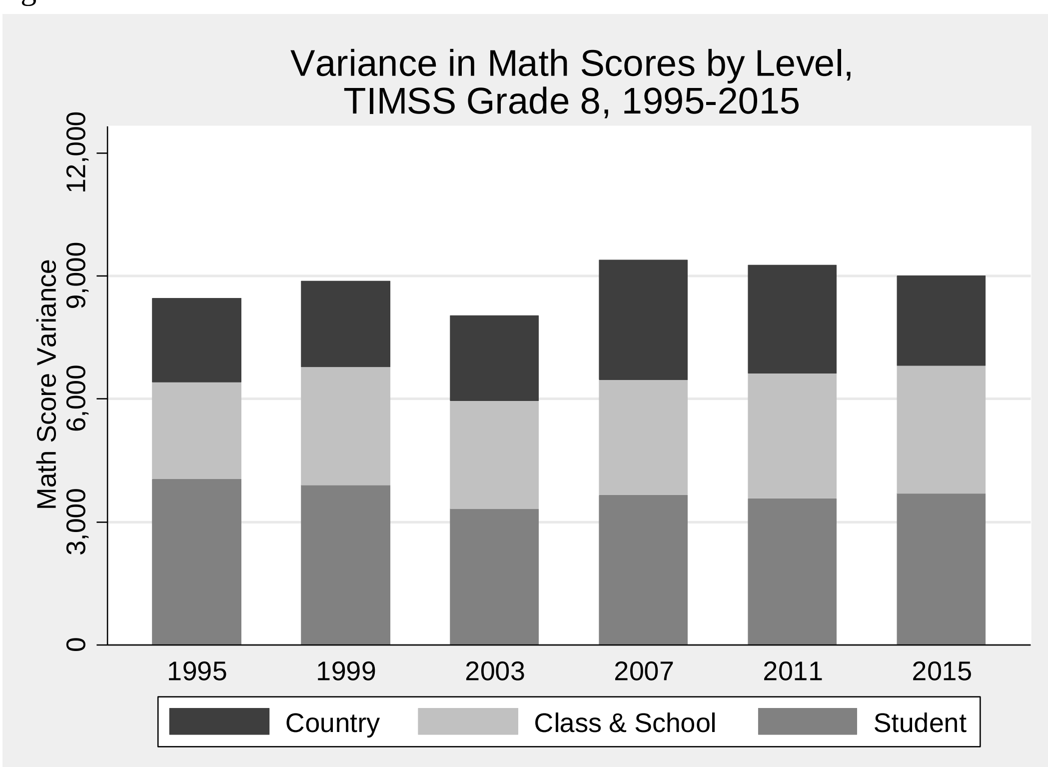


Figure C7

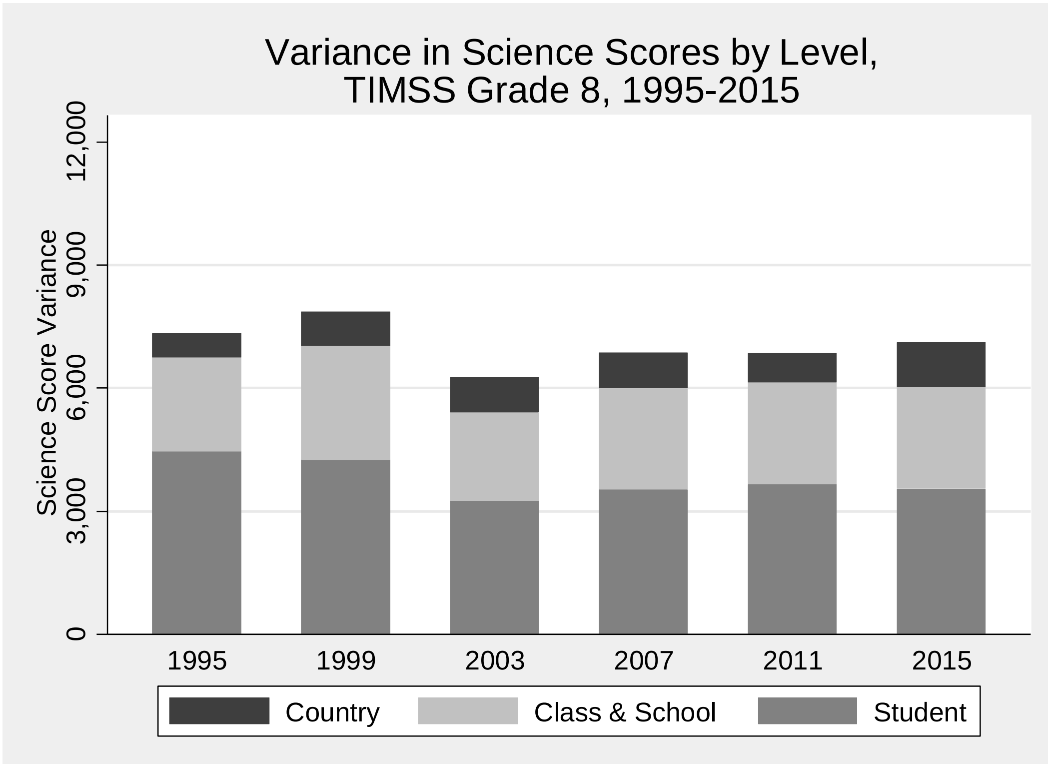
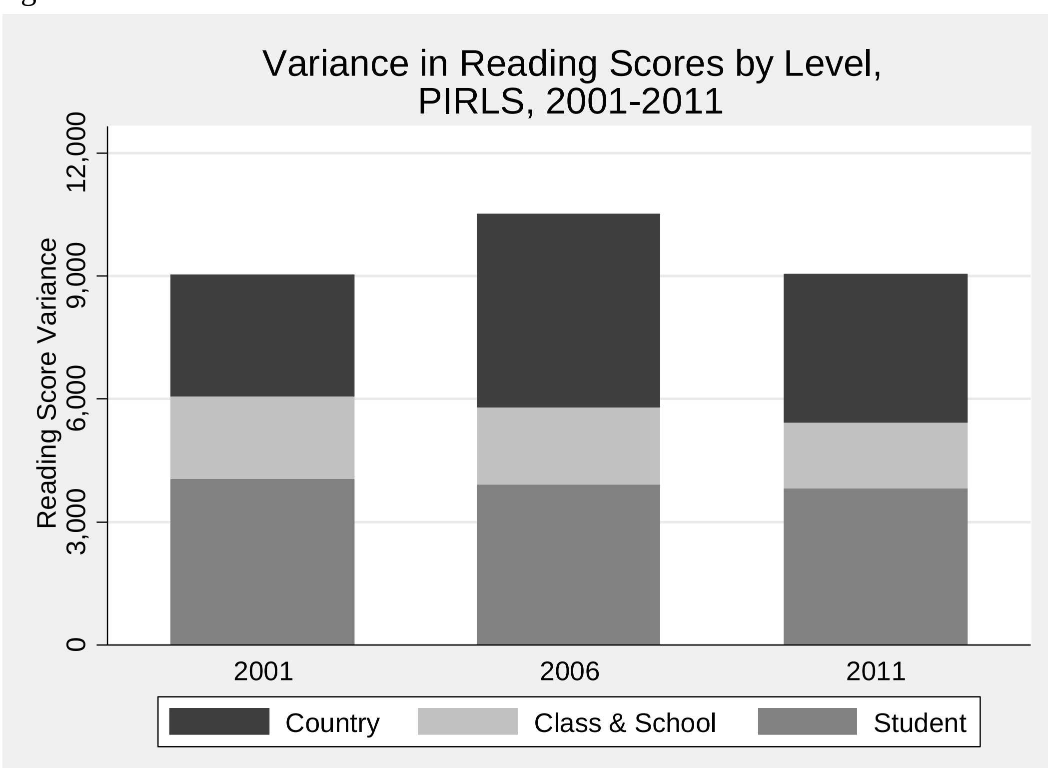


Figure C8



The hierarchal models reveal that within-country test score variance (the sum of the student- and school/classroom-level variances) has declined for all test instruments except TIMSS math (4th and 8th grades). These decreases in variance might lead us to question whether the absolute *size* of SES achievement gaps in terms of skills is declining even as their *strength* increases. However, the results for unstandardized achievement in Tables B1 and B2 show positive trends for most test instruments (except for PISA math and science). This indicates that, for the specific sets of reading, math, and science skills tested by the trend studies, differences in the degree to which high- and low-SES students have mastered those skills have mostly grown over the past 9-20 years.

The trend studies also allow us to examine not only whether achievement gaps have truly grown in size but also the changing *levels* of achievement for students of different SES. In other words, we can ask: Do SES achievement gaps increase because low-SES students' achievement is declining or because it is not rising as quickly as that of high-SES students? A series of models addressing this question, separately for each test instrument, are estimated as follows:

$$\hat{A}_{pjk} = \boldsymbol{\alpha}(\mathbf{T}_{pjk}) + \gamma_{100}(P10_{pjk}) + \gamma_{010}(C_{jk}) + \gamma_{110}(C_{jk}P10_{pjk}) + \boldsymbol{u}_k(\mathbf{T}_{pjk}) + w_{1k}(C_{jk}) + w_{2k}(C_{jk}P10_{pjk}) + \boldsymbol{r}_{jk}(\mathbf{T}_{pjk}) + s_{jk}(P10_{pjk}) + \epsilon_{pjk},$$

$$\epsilon_{pjk} \sim N\left(0, V_{jk}\right); \begin{bmatrix} r_{1jk} \\ r_{2jk} \\ r_{3jk} \\ s_{jk} \end{bmatrix} \sim MVN[\mathbf{0}, \boldsymbol{\Sigma}]; \begin{bmatrix} u_{1k} \\ u_{2k} \\ u_{3k} \\ w_{1k} \\ w_{2k} \end{bmatrix} \sim MVN[\mathbf{0}, \boldsymbol{\tau}] \quad \text{[C3]}$$

where $\hat{A}_{pjk}$ is the pth observed mean achievement of the 90th or the 10th SES percentile (level 1) in study-year j (level 2) in country k (level 3), $\boldsymbol{\alpha}$ is a vector of the true achievement means conditional on all covariates in the model, $\mathbf{T}_{pjk}$ is a vector of dummy variables indicating gap type (parent education, parent occupation, or household books), γ_{100} is the coefficient for the dummy variable $P10_{ijk}$ indicating whether the mean achievement p was estimated for the 10th (1) or the 90th (0) SES percentile, γ_{010} is the coefficient for cohort birth year C_{jk}, γ_{110} is the coefficient for the interaction between cohort birth year C_{jk} and the 10th percentile dummy $P10_{pjk}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\tau}$ are the within-country and between-country covariance matrices among the true gaps, and $V_{jk} = var\left(\hat{A}_{pjk}\right)$ is the sampling variance of $\hat{A}_{pjk}$, which I compute using the squared standard error of $\hat{A}_{pjk}$.[3]

Table C1 reports these models for each test instrument. The coefficient for Cohort birth year represents the trend in achievement for the reference category, the 90th SES percentile. The interaction between Cohort birth year and the 10th percentile dummy represents the difference in

[3] The level-1 known-variance portion of this model accounts only for the sampling variance of $\hat{A}_{pjk}$ and not for the sampling error covariances among different mean achievement estimates within the same country-study-year. Thus, the model assumes that all sampling error covariances among different mean achievement estimates are 0, which is unlikely to be true. However, the models in the main text yield very similar results when sampling error covariances are omitted (see Appendix L). Therefore, sampling error covariances are omitted from this and most appendix models due to the very long computation time of sampling error covariances via bootstrapping.

achievement trends for the two SES groups.[4] The sum of these two coefficients represents the trend in achievement for the 10^{th} SES percentile, which is reported at the bottom of the table, along with significance of a Wald joint test of the hypothesis that the sum of coefficients is equal to 0. On average, the achievement of both high- and low-SES students has increased in all test instruments. However, in most test instruments, the achievement of low-SES students has increased by a smaller amount than that of high-SES students. The two exceptions are PISA math and science, where the achievement of low-SES students has increased more than that of high-SES students, consistent with declining SES achievement gaps for these test instruments when achievement is unstandardized, seen in Table B1. Thus, the results for the majority of test instruments show that SES achievement gaps are increasing not because low-SES students' achievement is declining, but because it is not rising as quickly as that of high-SES students.

[4] Note that if the models in Tables B1-B4 and C1 were simple OLS regression models, the coefficients for the Cohort birth year*p10 interactions in Table C1 would be equal to the Cohort birth year trends in the gaps in Table B1-B4, but opposite in sign. This is not true in the reported models due to the precision weighting and random cohort slopes in the hierarchical growth models. However, estimated coefficients are generally similar in size and opposite in sign across Tables B1-B4 and C1.

Table C1. Models Predicting Achievement Levels for 90th and 10th SES Percentiles, Run Separately by Test Instrument

Test instrument	PISA Math	PISA Reading	PISA Science	PIRLS Reading	TIMSS Grade 4 Math	TIMSS Grade 4 Science	TIMSS Grade 8 Math	TIMSS Grade 8 Science
Test years	2003-2015	2000-2015	2006-2015	2001-2011	1995-2015	1995-2015	1995-2015	1995-2015
Cohort birth years	1988-2000	1985-2000	1991-2000	1991-2001	1985-2005	1985-2005	1981-2001	1981-2001
Parent education gaps intercept	507.127 ***	494.311 ***	513.236 ***	525.820 ***			482.550 ***	491.766 ***
	(8.761)	(7.744)	(8.354)	(10.582)			(11.904)	(10.809)
Parent occupation gaps intercept	506.873 ***	494.334 ***	512.779 ***	525.935 ***				
	(8.675)	(7.665)	(8.252)	(10.573)				
Household books gaps intercept	505.730 ***	492.764 ***	511.313 ***	525.379 ***	493.917 ***	497.314 ***	480.153 ***	489.396 ***
	(8.629)	(7.656)	(8.212)	(10.562)	(11.616)	(12.110)	(11.804)	(10.821)
p10	-81.823 ***	-78.787 ***	-83.919 ***	-63.540 ***	-49.020 ***	-55.704 ***	-54.590 ***	-54.802 ***
	(2.872)	(2.622)	(3.334)	(2.800)	(3.684)	(4.240)	(3.073)	(3.563)
Cohort birth year	0.167	0.841 ***	0.018	1.836 ***	1.444 ***	1.194 ***	1.297 ***	1.317 ***
	(0.252)	(0.241)	(0.254)	(0.531)	(0.323)	(0.327)	(0.374)	(0.362)
Cohort birth year × p10	0.294 +	-0.023	0.377 +	-0.530 *	-0.448 *	-0.273	-1.146 ***	-1.295 ***
	(0.166)	(0.158)	(0.201)	(0.248)	(0.209)	(0.232)	(0.198)	(0.185)
N (Level 1 - gaps)	1786	2060	1528	602	340	340	970	970
N (Level 2 - study-years)	298	344	255	104	170	170	245	245
N (Level 3 - countries)	70	72	70	41	51	51	61	61
	Wald test	Wald test	Wald test	Wald test	Wald test	Wald test	Wald test	Wald test
Estimated cohort birth year slope for p10	0.461 +	0.818 **	0.395	1.306 *	0.996 **	0.921 **	0.151	0.022
	(0.282)	(0.252)	(0.261)	(0.509)	(0.338)	(0.337)	(0.360)	(0.350)

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L).

Appendix D. Changing measurement error of achievement

Even assuming the true variance of achievement had remained constant over time, if measurement error of achievement declines over time (e.g., because of improvements in testing methodology), the SES achievement gap estimates in the main text of the paper will artificially appear to increase because they were attenuated in early years where measurement error was higher. This is because the method of standardizing achievement in each country-year involves dividing by the standard deviation of achievement, which will be inflated due to measurement error. That SES achievement gaps are also increasing for most of the unstandardized scores reported in Appendix B is evidence that findings are robust, even when not standardizing achievement.

Tables D1-D3 report median, minimum, and maximum test reliabilities by age group for math, reading, and science tests. Median test reliabilities have not consistently increased over time for all test subjects and age groups. Reliabilities have increased for 4th grade tests and for secondary science tests but appear to have declined somewhat for secondary math and reading tests. However, it should be kept in mind that the sample of countries participating in international assessments has become more diverse over time, and countries at a lower level of development often have lower test reliabilities.

Table D1. Median, Minimum, and Maximum Test Reliability for Math Tests

Study	Year	Median	Minimum	Maximum
		Grade 4 Math		
TIMSS	1995	0.84	0.74	0.88
TIMSS	2003	0.87	0.76	0.91
TIMSS	2007	0.83	0.55	0.88
TIMSS	2011	0.82	0.57	0.89
TIMSS	2015	0.88	0.78	0.92
		Grade 8 Math		
FIMS	1964	0.92	0.87	0.95
SIMS[a]	1980	0.85[a]	0.81[a]	0.85[a]
TIMSS	1995	0.89	0.73	0.92
TIMSS	1999	0.89	0.69	0.94
TIMSS	2003	0.89	0.51	0.94
TIMSS	2007	0.88	0.62	0.93
TIMSS	2011	0.87	0.66	0.94
TIMSS	2015	0.91	0.81	0.94
		Age 15 Math		
PISA	2000	0.88	0.82	0.92
PISA	2003	0.90	0.83	0.93
PISA	2006	0.88	0.83	0.93
PISA	2009	0.88	0.77	0.92
PISA	2012	0.92	0.84	0.94
PISA	2015	0.85	0.67	0.89

[a] SIMS test reliability was not reported in the available documentation and was estimated for each country using a model that included age, subject, year, and countries' level of development.

Table D2. Median, Minimum, and Maximum Test Reliability for Reading Tests

Study	Year	Median	Minimum	Maximum
Grade 4 Reading				
FIRCS	1970	0.85	0.74	0.89
RLS	1991	0.93	0.89	0.97
PIRLS	2001	0.88	0.83	0.91
PIRLS	2006	0.87	0.81	0.92
PIRLS	2011	0.88	0.79	0.93
Grade 8 Reading				
FIRCS	1970	0.85	0.64	0.90
RLS	1991	0.92	0.77	0.95
Age 15 Reading				
PISA	2000	0.92	0.87	0.94
PISA	2003	0.83	0.70	0.88
PISA	2006	0.88	0.80	0.93
PISA	2009	0.92	0.86	0.94
PISA	2012	0.89	0.81	0.93
PISA	2015	0.86	0.72	0.89

Table D3. Median, Minimum, and Maximum Test Reliability for Science Tests

Study	Year	Median	Minimum	Maximum
Grade 4 Science				
FISS	1970	0.82	0.68	0.87
SISS	1984	0.74	0.70	0.79
TIMSS	1995	0.77	0.70	0.83
TIMSS	2003	0.84	0.74	0.87
TIMSS	2007	0.80	0.69	0.88
TIMSS	2011	0.78	0.62	0.85
TIMSS	2015	0.85	0.81	0.90
Grade 8 Science				
FISS	1970	0.83	0.57	0.89
SISS	1984	0.75	0.60	0.80
TIMSS	1995	0.78	0.69	0.84
TIMSS	1999	0.80	0.62	0.86
TIMSS	2003	0.84	0.63	0.91
TIMSS	2007	0.84	0.65	0.91
TIMSS	2011	0.83	0.67	0.89
TIMSS	2015	0.89	0.81	0.92
Age 15 Science				
PISA	2000	0.87	0.75	0.92
PISA	2003	0.82	0.68	0.88
PISA	2006	0.91	0.84	0.94
PISA	2009	0.89	0.79	0.93
PISA	2012	0.89	0.80	0.93
PISA	2015	0.91	0.77	0.93

For the models in the main text of the paper, which pool different tests with different scales and must standardize achievement, all SES achievement gaps (and their standard errors) are adjusted according to each country's test reliability for each study, as published in the corresponding technical reports (as well as for the estimated reliability of SES reports, which is explained in more detail in Appendix H). The adjustment is computed as follows:

$$\widehat{Gap}_{adj} = \widehat{Gap}_{raw} * \frac{1}{\sqrt{\alpha_{ach}} * \sqrt{\alpha_{SES}}} \quad [D1]$$

Table D4 reports estimated trends in SES achievement gaps without adjusting those gaps for differences in test reliability (and also without adjusting for the reliability of SES report, which is discussed in more detail in Appendix H). Gap trends are positive and significant and are very similar to those reported in the models in the main text. Without adjusting for reliability, the increase in parent occupation gaps is nearly identical to that reported in the main text, and increases in parent education and books gaps are slightly smaller. Gaps adjusted for reliability are preferred, as they are likely more accurate.

Table D4. Estimated trends in 90/10 SES achievement gaps, without adjusting for test or SES reliability

	Adjusted for reliability		No adjustment	
	(1)	(2)	(1)	(2)
Parent education gaps intercept	1.032 ***	1.039 ***	0.741 ***	0.744 ***
	(0.030)	(0.030)	(0.022)	(0.022)
Parent occupation gaps intercept	0.958 ***	0.964 ***	0.775 ***	0.776 ***
	(0.030)	(0.030)	(0.024)	(0.024)
Household books gaps intercept	1.299 ***	1.294 ***	0.851 ***	0.850 ***
	(0.041)	(0.041)	(0.028)	(0.028)
Level 1 - Gaps				
Subject (ref=Reading):				
Math	0.020 **	0.020 **	0.017 **	0.017 **
	(0.007)	(0.007)	(0.006)	(0.006)
Science	0.034 ***	0.034 ***	0.025 ***	0.025 ***
	(0.005)	(0.005)	(0.004)	(0.004)
SES variable quality measures				
Parent-reported × Parent education	0.132 ***	0.112 ***	0.277 ***	0.265 ***
	(0.030)	(0.031)	(0.025)	(0.026)
Parent-reported × Parent occupation	0.075 **	0.073 **	0.082 ***	0.089 ***
	(0.025)	(0.024)	(0.021)	(0.020)
Parent-reported × Books	-0.039	-0.017	0.041	0.044 +
	(0.029)	(0.026)	(0.025)	(0.023)
Number of categories (centered at 7)	0.003	0.002	0.004	0.004 +
	(0.003)	(0.003)	(0.002)	(0.002)
≥ 20% in bottom category	-0.065 **	-0.063 **	-0.057 ***	-0.057 ***
	(0.021)	(0.021)	(0.015)	(0.015)
≥ 20% in top category	-0.135 ***	-0.146 ***	-0.099 ***	-0.101 ***
	(0.013)	(0.013)	(0.009)	(0.009)
Level 2 - Study-years				
Age at testing (ref=14)				
Age 10 at testing	-0.170 ***	-0.168 ***	-0.159 ***	-0.159 ***
	(0.024)	(0.024)	(0.017)	(0.017)
Age 15 at testing	-0.024	-0.023	-0.003	-0.002
	(0.020)	(0.020)	(0.014)	(0.014)
Cohort birth year × Parent education	0.007 ***		0.005 ***	
	(0.001)		(0.001)	
Cohort birth year × Parent occupation	0.007 ***		0.007 ***	
	(0.001)		(0.001)	
Cohort birth year × Books	0.008 ***		0.006 ***	
	(0.001)		(0.001)	
Cohort birth year		0.007 ***		0.006 ***
		(0.001)		(0.001)
Random effects				
Level 2 - Residual variance between studies in...				
Parent education intercepts	0.03736	0.03831	0.0219	0.02253
Parent occupation intercepts	0.02322	0.02284	0.01498	0.01496
Books intercepts	0.03698	0.03823	0.01607	0.01618
Level 3 - Residual variance between countries in...				
Parent education intercepts	0.05426	0.05362	0.03274	0.0325
Parent occupation intercepts	0.05227	0.0533	0.03748	0.0373
Books intercepts	0.1159	0.12149	0.05473	0.05587
Parent education cohort slopes	0.00004		0.00002	
Parent occupation cohort slopes	0.00003		0.00002	
Books cohort slopes	0.00007		0.00003	
Cohort slopes		0.00003		0.00002
N (Level 1 - gaps)	5541	5541	5541	5541
N (Level 2 - study-years)	1026	1026	1026	1026
N (Level 3 - countries)	100	100	100	100

+ $p<.1$, * $p<.05$, ** $p<.01$, *** $p < .001$. Note: All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper.

Appendix E. Changing distribution of SES

Educational attainment and the occupational structure have shifted dramatically in most countries since 1964. The trends reported in the main text of the paper refer to changes in the achievement gap between the 90th, 50th, and 10th percentiles of each SES variable in each country-year, even though the meaning of the 90th, 50th and 10th percentiles has changed over time. Table E1 reports the median category at which the 90th, 50th, and 10th percentiles of each SES variable fall across all participating countries in an older study and the same set of countries in PISA 2015. For parent education and occupation, the older study is FIMS 1964; for household books, it is the SISS 1984 8th grade sample, the first available high-quality measure of household books. It is clear that dramatic upgrading has occurred for both parent education and occupation during this period. However, the distribution of household books has remained more stable.

Table E1. Median Category at which 90th, 50th, and 10th SES Percentiles Fall, by Study

		FIMS 1964	PISA 2015
Parent education	90th percentile	13 years	ISCED 5A+
	50th percentile	9 years	ISCED 5B
	10th percentile	7 years	ISCED 3B, C
Parent occupation	90th percentile	Semi-Professional	Professionals
	50th percentile	Clerical & Sales	Technicians and Associate Professionals
	10th percentile	Manual Workers, Skilled & Semi-Skilled	Craft Etc Trades Workers
		SISS 1984	PISA 2015
Household books	90th percentile	251-500 books	201-500 books
	50th percentile	26-100 books	26-100 books
	10th percentile	11-25 books	0-10 books

Note: Countries included in FIMS-PISA sample (parent education and occupation) are Australia, Belgium-Flanders, Belgium-French, England, Finland, France, Germany, Israel, Japan, Netherlands, Scotland, and United States. Countries included in SISS-PISA sample (books) are Australia, Canada, England, Finland, Hong Kong, Hungary, Israel, Italy, Japan, Korea, Netherlands, Norway, Poland, Singapore, Sweden, Thailand, and United States.

Another way to describe the changing SES distribution is the changing share of the sample that falls into the same, consistently-defined set of high, medium, and low SES categories over time. Table E2 reports the average share of the sample falling into each of the three SES categories for a constant sample of countries participating in an older and more recent study. Each SES variable is coded into three categories for all studies, based on the ISCED 1997 scheme: parent education is coded into (1) less than secondary [less than ISCED 3], (2) secondary or non-degree vocational postsecondary [ISCED 3 or 4], and (3) an academically- or vocationally-oriented higher educational degree [ISCED 5A or 5B or more]. Parent occupation is coded into (1) working class [unskilled, semiskilled, or agricultural labor], (2) intermediate class [skilled trades, service, clerical, or small business], (3) salariat class [semi-professional, managerial, or professional]. Household books are coded into (1) 0-10 books, (2) 11-100 books, (3) 101 books or more. Once again, it is clear that the average level of parent education and occupation have increased dramatically, while the distribution of books has remained relatively constant or even *decreased* slightly. Comparing the SISS 1984 8th grade dataset to PISA 2015, the average share of students reporting 10 or fewer books at home has more than doubled (from 6% to nearly 14%), while the shares of students in the high and medium books categories have both declined by a few percentage points.

Table E2. Proportion of Sample in High, Medium, and Low Categories of SES Variables, by Study

		FIMS 1964	PISA 2015
Parent education	High (ISCED 5A, 5B)	0.106	0.644
	Medium (ISCED 3, 4)	0.168	0.288
	Low (<ISCED 3)	0.726	0.067
Parent occupation	High (Salariat)	0.202	0.460
	Medium (Intermediate)	0.313	0.343
	Low (Working Class)	0.485	0.196
		SISS 1984	PISA 2015
Household books	High (101+)	0.448	0.401
	Medium (11-100)	0.492	0.461
	Low (0-10)	0.060	0.138

Note: Countries included in FIMS-PISA sample (parent education and occupation) are Australia, Belgium-Flanders, Belgium-French, England, Finland, France, Germany, Israel, Japan, Netherlands, Scotland, and United States. Countries included in SISS-PISA sample (books) are Australia, Canada, England, Finland, Hong Kong, Hungary, Israel, Italy, Japan, Korea, Netherlands, Norway, Poland, Singapore, Sweden, Thailand, and United States.

The 90/10 percentile method (Reardon 2011b) was chosen to avoid changes in the selectivity of different SES categories as their frequencies changed over time. However, treating these historical and contemporary percentiles as equivalent also makes a theoretical assumption that these SES characteristics confer mainly positional advantages to children. Alternatively, it may be that, for example, having a parent with a university degree always confers the same absolute advantage, regardless of whether that parent was among the elite few who earned a degree in the mid-20th century or the larger share who earned a degree at the turn of the 21st century. One piece of evidence that the increasing SES achievement gaps reported here are not merely an artifact of the general upgrading of SES is that increases are found not only for parental education and occupation, whose levels have increased over time, but also for household books, whose levels have remained stable or even slightly declined, as seen above.

In addition to changing levels of the three SES variables, the dispersion of SES has also changed. The variance of parent education and occupation has declined somewhat in most countries over time, while the variance of household books has remained relatively constant. (The changing variances of these ordinal SES variables were computed after recoding into the same categories in every study—6 categories for parent education and occupation and 5 categories for books.) All else equal, if the variance of an independent variable—SES—decreases, then its unstandardized association with an outcome variable—achievement—will increase. The models in the main text of the paper avoid this problem, as converting the SES variables into percentiles is a form of standardization. However, these changes in variance should be kept in mind for the next set of models, where SES is unstandardized.

An additional piece of evidence that increasing SES achievement gaps are not an artifact of changing SES distributions comes from the models reported in Tables E3 and E4, which compute achievement gaps between three consistently-defined categories of each SES variable rather than percentiles. These analyses examine the robustness of the finding of increasing SES achievement gaps to treating SES as an absolute rather than a positional good. SES variables are coded as in Table E2 above. Table E3 reports trends in the gap between the top and bottom categories of each variable, while the top panel of Table E4 reports trends in the gap between the top and middle categories, and the bottom panel of Table E4 reports trends in the gap between the middle and bottom categories. Thus, these models are an alternative way to capture changing achievement gaps across the whole SES distribution, at the top, and at the bottom, rather than the 90/10, 90/50, and 50/10 gap trends reported in the main text.

In these models, unlike in the models predicting percentile-based gaps, the size of coefficients cannot be compared across different SES variables because gaps based on categories of different SES variables are not on equivalent scales. Therefore, we look only at whether cohort birth year coefficients are positive. The results in Table E3 show that gaps have increased between the top and bottom categories of all three variables. Table E4 shows that, for all three variables, increases between the top and middle categories are larger than increases between the middle and bottom categories. This is true both for parent education and occupation, where the top category has become a larger share of students, and for household books, where the top category has become a slightly smaller share of students. Thus, the achievement advantage of students with college-educated or professional parents or many books at home has increased, even as the share of students with college-educated or professional parents has increased (and the share with many books has declined).

That the gap between the top and middle SES categories has increased more than between the middle and bottom categories may at first appear inconsistent with the results from the main text of the paper showing that 50/10 gaps have increased more than 90/50 gaps in most countries. However, these two findings are simultaneously true. Comparing between descriptive Tables E1 and E2 shows why this is the case. In PISA 2015, the 50th and 10th percentiles of parent education and occupation in fact correspond to the “high” and “medium” categories of parent education and occupation. This demonstrates the difficulty of studying trends in achievement gaps between categories with drastically changing distributions. I choose to treat SES as a positional rather than an absolute good in this study, as it seems a more tenable assumption that the advantage conferred by a particular level of parental education or occupation changes over time with the changing distributions of these variables. However, the results of these models indicate that achievement gaps between high and low SES categories still appear to increase even when SES is treated as an absolute good.

Table E3. Estimated Trends in Achievement Gaps between High and Low SES Categories

	Education		Occupation		Books	
High-low						
Age 10 at testing	-0.038		0.099	***	-0.221	***
	(0.032)		(0.020)		(0.028)	
Age 15 at testing	-0.083	**			-0.001	
	(0.030)				(0.025)	
Math	0.03	**	0.002		-0.02	+
	(0.010)		(0.008)		(0.012)	
Science	0.032	***	0.001		0.044	***
	(0.009)		(0.007)		(0.013)	
Cohort birth year	0.007	***	0.002	*	0.003	+
	(0.001)		(0.001)		(0.002)	
Intercept	0.976	***	0.825	***	1.318	***
	(0.029)		(0.021)		(0.044)	
Residual variance (within countries)	0.03481		0.00851		0.02560	
Residual variance (country intercepts)	0.05793		0.02703		0.16224	
Residual variance (cohort slopes)	0.00009		0.00007		0.00015	
N (observations)	1889		1334		2086	
N (countries)	93		80		95	

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: See text for definitions of high, middle, and low for each SES variable. To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L). In models predicting achievement gaps based on parent occupation, the dummy variable for age 15 is omitted, as there are very few age-14 assessments that collected parent occupation. Therefore, in the parent occupation models, the reference category is assessments of students who are age 14 *or* 15.

Table E4. Estimated Trends in Achievement Gaps between High-Middle and Middle-Low SES Categories

	Education	Occupation	Books
High-middle			
Age 10 at testing	-0.029	-0.036 **	-0.169 ***
	(0.019)	(0.012)	(0.015)
Age 15 at testing	-0.088 ***		0.011
	(0.016)		(0.013)
Math	0.031 ***	0.012 *	0.001
	(0.006)	(0.005)	(0.008)
Science	0.019 ***	0.012 ***	0.0270 ***
	(0.005)	(0.004)	(0.007)
Cohort birth year	0.005 ***	0.002 +	0.002 *
	(0.001)	(0.001)	(0.001)
Intercept	0.496 ***	0.548 ***	0.62 ***
	(0.014)	(0.016)	(0.022)
Residual variance (within countries)	0.01243	0.00319	0.00907
Residual variance (country intercepts)	0.01691	0.01686	0.04203
Residual variance (cohort slopes)	0.00003	0.00005	0.00007
N (observations)	1889	1334	2086
N (countries)	93	80	95
Middle-low			
Age 10 at testing	-0.004	0.137 ***	-0.05 *
	(0.021)	(0.015)	(0.024)
Age 15 at testing	0.004		-0.004
	(0.023)		(0.018)
Math	-0.001	-0.009 +	-0.024 ***
	(0.007)	(0.005)	(0.007)
Science	0.012 +	-0.013 **	0.015 *
	(0.007)	(0.004)	(0.007)
Cohort birth year	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)
Intercept	0.484 ***	0.273 ***	0.698 ***
	(0.020)	(0.010)	(0.026)
Residual variance (within countries)	0.01781	0.00297	0.01407
Residual variance (country intercepts)	0.02717	0.00471	0.05037
Residual variance (cohort slopes)	0.00004	0.00004	0.00005
N (observations)	1889	1334	2086
N (countries)	93	80	95

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: See text for definitions of high, middle, and low for each SES variable. To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L). In models predicting achievement gaps based on parent occupation, the dummy variable for age 15 is omitted, as there are very few age-14 assessments that collected parent occupation. Therefore, in the parent occupation models, the reference category is assessments of students who are age 14 *or* 15.

Appendix F. Achievement gaps by mother's and father's SES characteristics

One possible explanation for increasing SES achievement gaps is the increasing educational attainment and occupational status of mothers. The dramatic global rise in women's status since the mid-20th century is easily visible in the international assessment data. In early international assessments, most students reported a higher level of educational attainment for their father than their mother. The mother's occupation was not collected at all in the earliest three datasets (FIMS 1964, FISS 1970, and FIRCS 1970). In the most recent assessments, mothers and fathers are about equally educated, and mothers have somewhat higher occupational status than fathers (reflecting women's greater likelihood of working in white collar jobs). Since the main analyses in study use only the *highest* of the two parents' education and occupation as a measure of the child's SES, the increasing education and occupational status of mothers means that the highest parent education and occupation are increasingly likely to come from the mother. If children's achievement tends to be more strongly associated with their mother's than with their father's education and occupation (e.g., because mothers perform the majority of childcare), then the increasing educational and occupational attainment of mothers could explain why SES achievement gaps are increasing. (Children and/or parents were asked to report the education and occupation of both parents, whether or not both were present in the home, and most international assessments did not collect data on whether each parent lived in the home. For parents not currently working, the most recent occupation was reported. For parents who had never worked—who were very likely to be mothers performing home duties, particularly in earlier years—this study treats that parent's occupation as missing and imputes an occupation as part of the multiple imputation model described in the Methods section of the main text of the paper.)[5]

To check the robustness of the main results to changes in the relative status of mothers and fathers, Table F1 reports trends in SES achievement gaps based on fathers' and mothers' education and occupation separately. The results show that trends in gaps based on all four characteristics are positive and significant. As expected, increases in gaps based on mothers' education and occupation are larger than those based on fathers' characteristics—a Wald joint hypothesis test shows that the increase in the mother's education gap is significantly larger than the increase in the father's education gap ($p < .001$). The trends in gaps based on mothers' and fathers' occupation differ only slightly, by less than 0.001 SD per year, a difference that is not significant ($p > 0.5$). Between the 1950 and the 2005 birth cohorts, the father's education achievement gap grew from about 0.72 SDs to 1.10 SDs (about a 54% increase), while the mother's education achievement gap grew from 0.68 SDs to 1.12 SDs (about a 65% increase). Between the 1966 and 2005 birth cohorts (1966 is the birth cohort corresponding to SIMS 1980, the first cohort for which both mothers' and fathers' occupation are available), the father's occupation achievement gap grew from 0.81 to 1.04 SDs (about a 29% increase), and the mother's occupation achievement gap grew from 0.83 to 1.07 SDs (about a 28% increase). That achievement gaps have increased not only for mothers' but also fathers' education and occupation suggests that the global increase in parent education and occupation achievement gaps is not fully explained by the increasing educational or occupational attainment of mothers.

[5] See Appendix L for models computing gap trends with listwise deletion of missing data rather than multiple imputation. In these models, for students with one missing and one nonmissing parent education or occupation, the nonmissing value was used as the "highest" parent education or occupation. Results are very similar to those from the models with imputed data.

Table F1. Trends in Achievement Gaps Based on Fathers' and Mothers' SES Characteristics

	Coef	(se)	
Father's education gaps intercept	0.991	(0.028)	***
Mother's education gaps intercept	0.993	(0.030)	***
Father's occupation gaps intercept	0.947	(0.027)	***
Mother's occupation gaps intercept	0.970	(0.029)	***
Level 1 - Gaps			
Subject (ref=Reading):			
Math	0.036	(0.007)	***
Science	0.028	(0.005)	***
SES variable quality measures			
Parent-reported × Parent education	0.091	(0.029)	**
Parent-reported × Parent occupation	-0.054	(0.025)	*
Number of categories (centered at 7)	0.001	(0.002)	
≥ 20% in bottom category	-0.044	(0.011)	***
≥ 20% in top category	-0.047	(0.009)	***
Level 2 - Study-years			
Age at testing (ref=14)			
Age 10 at testing	-0.091	(0.033)	**
Age 15 at testing	-0.116	(0.023)	***
Cohort birth year × Father's education	0.007	(0.001)	***
Cohort birth year × Mother's education	0.008	(0.001)	***
Cohort birth year × Father's occupation	0.006	(0.001)	***
Cohort birth year × Mother's occupation	0.006	(0.001)	***
Random effects			
Level 2 - Residual variance between studies in…			
Father's education intercepts	0.03050		
Mother's education intercepts	0.03658		
Father's occupation intercepts	0.01361		
Mother's occupation intercepts	0.01496		
Level 3 - Residual variance between countries in…			
Father's education intercepts	0.04857		
Mother's education intercepts	0.05634		
Father's occupation intercepts	0.03237		
Mother's occupation intercepts	0.04390		
Father's education cohort slopes	0.00006		
Mother's education cohort slopes	0.00005		
Father's occupation cohort slopes	0.00004		
Mother's occupation cohort slopes	0.00005		
N (Level 1 - gaps)	6502		
N (Level 2 - study-years)	866		
N (Level 3 - countries)	95		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper.

Another possible explanation for increasing SES achievement gaps is increasing homogamy among the parents of participating students. That is, the SES characteristics (education and occupation) of mothers and fathers are likely growing more correlated over time. Students with two highly-educated or high-occupational-status parents may be more advantaged than students with only one highly-educated or high-occupational-status parent. Table F2 reports hierarchical growth models estimating trends in the correlation between mothers' and fathers' education and occupation. The results show that the average correlation between mothers' and fathers' education has increased only slightly (from about 0.56 in the 1950 birth cohort to about 0.60 in the 2005 cohort), and the average correlation between mothers' and fathers' occupation has remained relatively constant (declining slightly from about 0.40 in the 1966 birth cohort to about 0.39 in the 2005 cohort). These results demonstrate a less pronounced increase in homogamy than expected, which may be because this analysis treats parent education and occupation as continuous positional goods, converted into percentiles within each country-year, as in the main text of this paper. While it is true that an increasing number of children have two parents with higher education degrees or professional occupations, the associations between the *relative* positions of mothers and fathers within their own gender distributions have not increased dramatically over time.

Table F2. Trends in Correlations between Mothers' and Fathers' Education and Occupation

	Education	Occupation
Intercept	0.589 ***	0.392 ***
	(0.006)	(0.009)
Level 1 - Study-years		
Age at testing (ref=14)		
Age 10 at testing	-0.025 ***	0.090 ***
	(0.007)	(0.008)
Age 15 at testing	-0.054 ***	
	(0.006)	
Cohort birth year	0.0006 *	-0.0003
	(0.0003)	(0.0005)
Level 1 residual variance	0.00234	0.00365
Level 2 residual variance in intercepts	0.00341	0.00523
Level 2 residual variance in cohort slopes	0.00000	0.00001
N (Level 1 - study-years)	866	550
N (Level 2 - countries)	95	82

+ p<.1, * p<.05, ** p<.01, *** p < .001

Appendix G. Achievement gaps conditional on other SES variables

Results in the main paper text estimate achievement gaps separately for each SES variable (parent education, parent occupation, and household books) rather than constructing an SES index to avoid loss of information because not all SES variables are available in every dataset. However, it could be the case that only one of the three SES variables is growing more strongly associated with achievement over time, while the other two SES variables only appear to be growing more strongly associated with achievement due to their correlations with this one most salient SES variable. Or it may even be the case that the independent associations between each SES variable and achievement have remained constant over time, but correlations between the three SES variables are growing stronger over time. This would create the appearance of increasing SES achievement gaps for all three variables because an increasing share of students would experience "double-" or "triple disadvantage." That is, students with university-educated parents would be more likely also to have parents with professional occupations and to have a large number of books at home. Conversely, students whose parents have not completed secondary education would be more likely to have parents with working-class occupations and very few books at home. Thus, there may be a pattern of increasing polarization of socioeconomic advantage and disadvantage among schoolchildren, which may completely explain away increasing SES achievement gaps for all three SES variables.

Table G1 reports hierarchical growth models estimating trends in the correlation between pairs of SES variables. The results show that the average correlation between parent education and parent occupation has increased quite substantially (from about 0.43 in the 1950 birth cohort to about 0.56 in the 2005 cohort). This finding is consistent with international research showing an increasing association between education and occupation across most countries (Kreidl, Ganzeboom and Treiman 2014).[6] In contrast, the average correlations between household books and each of parent education and occupation, respectively, have increased more modestly (the correlation between household books and parent education increased from about 0.31 in the 1956 birth cohort to about 0.35 in the 2005 cohort; the correlation between household books and parent occupation increased from about 0.31 in the 1956 birth cohort to about 0.39 in the 2005 birth cohort).

[6] Note the literature shows an increasing association between education and occupation only when both variables are treated as linear (as here); the true pattern of change may be more complex. For example, research on over-education shows a declining relationship between attainment of tertiary education and a professional occupation when treating both variables as categorical rather than linear—that is, assuming education and occupational status are absolute rather than positional goods (the opposite of the assumption made in this paper).

Table G1. Trends in Correlations between Three SES Variables

	Education & Occupation	Education & Books	Occupation & Books
Intercept	0.523 ***	0.338 ***	0.361 ***
	(0.042)	(0.007)	(0.014)
Level 1 - Study-years			
Age at testing (ref=14)			
Age 10 at testing	0.057	0.064 ***	-0.007
	(0.048)	(0.007)	(0.012)
Age 15 at testing	-0.048	-0.031 ***	-0.059 ***
	(0.044)	(0.005)	0.012
Cohort birth year	0.0025 **	0.0008 +	0.0016 ***
	(0.0008)	(0.0004)	(0.0004)
Level 1 residual variance	0.00529	0.00195	0.0018
Level 2 residual variance in intercepts	0.00437	0.00339	0.00541
Level 2 residual variance in cohort slopes	0.00001	0.00001	0.00001
N (Level 1 - study-years)	576	835	567
N (Level 2 - countries)	82	95	83

+ p<.1, * p<.05, ** p<.01, *** p < .001

But do these increasing correlations fully explain increasing SES achievement gaps for all three variables? And even if increasing correlations do not fully explain increasing gaps, could it be the case that only one or two SES variables are growing more strongly associated with achievement, while the other SES variable(s) only appear to be growing more strongly associated with achievement due to their correlation(s) with the most salient SES variable(s)? One way to address both of these questions is by computing SES achievement gaps for each variable conditional on one or both of the other SES variables. The results of these models are presented in Table G2. Conditional gaps can be estimated only from studies that collected more than one SES variable, meaning sample sizes are reduced. To obtain an accurate comparison to cohort trends in unconditional SES achievement gaps, trends in unconditional gaps are also estimated using the same reduced set of studies. In Table G2, each row reports two models estimating gaps based on a particular SES variable. The "Conditional gaps" columns report the intercept and cohort birth year coefficients for a model whose gaps are conditional on one or both of the other SES variables. The "Unconditional gaps" columns report the coefficients for a model based on unconditional gaps using the same sample of studies. In each model, the intercept is the estimated gap for the variable in question (conditional or unconditional) for the 1989 birth cohort, while the cohort trend is the estimated annual change in the gap. All models control for subject and age at testing. Since there is no established method to adjust conditional associations for attenuation due to measurement error, conditional gaps are not adjusted for test or SES reliability. For an accurate comparison, the unconditional gaps are also not adjusted for reliability.

Table G2. Intercept and Cohort Birth Year Coefficients from Models Predicting 90/10 SES Achievement Gaps Conditional on Other SES Variables

		Conditional gaps			Unconditional gaps		
Variable	Conditional on	Intercept	Cohort trend	(N)	Intercept	Cohort trend	(N)
Education	Occupation	0.444 ***	0.002	(81)	0.760 ***	0.005 **	(81)
Education	Books	0.389 ***	0.003 **	(95)	0.729 ***	0.006 ***	(95)
Education	Occupation & Books	0.302 ***	0.001	(81)	0.750 ***	0.003 *	(81)
Occupation	Education	0.361 ***	0.001	(81)	0.684 ***	0.004 ***	(81)
Occupation	Books	0.356 ***	0.001	(82)	0.672 ***	0.004 ***	(82)
Occupation	Education & Books	0.278 ***	-0.001	(81)	0.729 ***	0.003 *	(81)
Books	Education	0.483 ***	0.004 ***	(95)	0.806 ***	0.007 ***	(95)
Books	Occupation	0.502 ***	0.002 **	(82)	0.785 ***	0.005 ***	(82)
Books	Education & Occupation	0.475 ***	0.004 ***	(81)	0.797 ***	0.006 ***	(81)

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L).

By computing the predictions of the model with all three SES variables (rows 3, 6, and 9) for the 1956 birth cohort (the first cohort with all three variables available) and the 2005 cohort, it can be seen that in both years, books had the strongest independent association with achievement (0.54), followed by education (0.32) and then occupation (0.26). However, over this time period, this ranking became even more pronounced. The independent 90/10 books gap increased markedly from 0.34 to 0.54, while the independent 90/10 parent education gap increased more modestly from 0.27 to 0.32, and the independent parent occupation gap *decreased* from 0.31 to 0.26. (Note this does not mean parent occupation is not significantly related to student achievement after controlling for parent education and household books, only that the relationship between parent education and achievement is not growing, after accounting for the other two SES variables). Thus, it appears that the independent associations of each SES variable with achievement have changed at different rates over time.

That the conditional 90/10 parent education and household books gaps show increases suggests that the global increases in unconditional SES achievement gaps for these variables are not fully explained by their correlations with other SES characteristics. In contrast, the declining conditional 90/10 parent occupation gap suggests that the growing parent occupation gap may be fully explained by the correlations between parent occupation and other SES characteristics. However, it should be noted that the model with all three SES variables included is based on a substantially reduced sample of only those datasets that collected all three SES variables, and thus may not be representative.

Regardless of the relative importance of parent education, parent occupation, and books, the results in Table G2 clearly indicate that increasing correlations between SES variables do not fully explain increasing SES achievement gaps. Nearly all trends in conditional SES achievement gaps are positive, with the exception of parent occupation conditional on education and books, as noted above. These results suggest that, even after accounting for the growing number of children with “double-” or “triple-disadvantage” due to increasing correlations among SES variables, each SES variable (or at least parent education and books) has become more consequential for students’ academic achievement.

One further piece of evidence that increasing SES achievement gaps for each SES variable are not entirely due to increasing correlations between variables is that the R^2 of the models used to compute conditional gaps with all three variables has increased over time. Table G3 shows results from hierarchical growth models (country-subject-years within countries) predicting R^2 from cohort birth year (with controls for age and subject). (R^2 is adjusted for test reliability before running these models.) Results indicate that the R^2 of the model including all three SES variables nearly doubled from about 0.10 in the 1956 birth cohort to about 0.18 in the 2005 birth cohort. Thus, it appears that the overall predictive power of SES on achievement has grown substantially stronger over this 49-year time period.

Table G3. Trends in R^2 from Models with Two or Three SES Variables

	R^2 (Education & Occupation)	R^2 (Education & Books)	R^2 (Occupation & Books)	R^2 (Education, Occupation & Books)
Intercept	0.107 ***	0.138 ***	0.129 ***	0.154 ***
	(0.013)	(0.006)	(0.010)	(0.016)
Level 1 - R^2 measures				
Subject (ref=Reading):				
Math	0.006 ***	0.002	0.001	0.003 +
	(0.002)	(0.002)	(0.002)	(0.002)
Science	0.005 ***	0.009 ***	0.005 **	0.006 ***
	(0.001)	(0.001)	(0.002)	(0.001)
Level 2 - Study-years				
Age at testing (ref=14)				
Age 10 at testing	0.022	-0.015 **	-0.008	-0.010
	(0.013)	(0.006)	(0.008)	(0.015)
Age 15 at testing	0.002	-0.001	0.037 ***	0.02
	(0.014)	(0.005)	0.01	0.016
Cohort birth year	0.0010 ***	0.0021 ***	0.0014 ***	0.0017 ***
	(0.0003)	(0.0002)	(0.0003)	(0.0003)
Level 1 residual variance	0.000198	0.000322	0.000293	0.000323
Level 2 residual variance in intercepts	0.000806	0.001049	0.000728	0.000853
Level 3 residual variance in intercepts	0.001443	0.002646	0.002701	0.002623
Level 3 residual variance in cohort slopes	0.000002	0.000002	0.000002	0.000002
N (Level 1 - R^2 measures)	1382	1910	1389	1351
N (Level 2 - study-years)	576	836	567	545
N (Level 3 - countries)	82	95	83	82

+ p<.1, * p<.05, ** p<.01, *** p < .001

Appendix H. Changing measurement error of SES

Another reason why SES achievement gaps may artificially appear to be increasing over time is that the reliability with which SES is measured may be increasing over time. Lower reliability of SES in early years could cause the estimated association between SES and achievement in early years to be attenuated, creating the appearance of increasing SES achievement gaps over time. Reliability is defined as the ratio of the variance in true SES to the total variance in SES (including both true variance and the variance of errors of measurement). As reported in Appendix E, the total variance of SES has declined for two SES variables—parent education and occupation—while the total variance has remained relatively constant for household books. It is likely that much of the decline in total variance in parent education and occupation reflects a decline in true variance, due to the large increase in the average *levels* of both variables as more parents attain higher education and professional occupations.

However, it is also likely that some of the reduction in the total variance of parent education and occupation is due to declining measurement error in these variables. Measurement error could potentially decline, for example, because current students have more accurate knowledge of their families' SES characteristics than in the past, because of improvements in survey wording, or because more recent SES data are more likely to be reported by parents rather than students. As stated in the main paper text, some international assessments have added parent questionnaires in recent years (PIRLS 2001-2011, TIMSS 4th grade 2011-2015, and PISA 2006-2012). The main models use parent-reported SES variables when available and student-reported SES otherwise. However, it is expected that parents report SES variables more reliably than their children. Some other patterns reported in previous appendix sections are consistent with increasing reliability of SES due to declining measurement error, while other patterns are inconsistent with this story. The large increase in the correlation between parent education and occupation (reported in Appendix G) is consistent with increasing reliability in these variables. However, the increase in the correlations between household books and each of the other two SES variables have been more modest. An increasing correlation between mothers' and fathers' education and occupation (reported in Appendix F) could also be evidence for increasing reliability of these variables. However, the correlation between mothers' and fathers' education has increased only moderately, and the correlation between mothers' and fathers' occupations has *declined* slightly.

As described in the Methods section of the main paper text, all SES achievement gaps have been adjusted for estimated reporting reliability of SES variables. Following Reardon (2011a), this adjustment consists of by multiplying each gap estimate by $\frac{1}{\sqrt{r}}$, where r is the reliability of the SES measure. In order to estimate the reliability r of each student- or parent-reported SES measure, we can take advantage of having two measures of the same variable reported by different sources (i.e., students and parents) (Jerrim and Micklewright 2014). The reliability of students' and parents' SES reports can be computed from the following formulas (Reardon 2011a):

$$r_s = corr(s,p) \cdot \frac{corr(s,y)}{corr(p,y)}$$
$$r_p = corr(s,p) \cdot \frac{corr(p,y)}{corr(s,y)}$$

where r_s is the reliability of student-reported SES variable s, r_p is the reliability of parent-reported SES variable p, and y is a third variable that would have a particular correlation with true SES, were SES measured without error. In reliability calculations from PISA and TIMSS, math achievement is used for this third variable y; in reliability calculations from PIRLS, reading achievement is used for y. Parent education is reported by both 15-year-old students and parents in PISA 2006, 2009, and 2012. Parent occupation is reported by both 15-year-old students and parents in PISA 2006 and 2012. Household books are reported by both 4th grade students and parents in PIRLS 2001, 2006, and 2011 and in TIMSS 2011 and 2015. I estimate the average reliability of parents' reports of their own educational attainment (across PISA 2006-2012) at 0.84 and students' reports of their parents' education at 0.62. I estimate the average reliability of parents' reports of their own occupational category (across PISA 2006 and 2012) at 0.81 and students' reports of their parents' occupation at 0.79. I estimate the average reliability of parents' reports of the number of household books (across PIRLS 2001-2011 and TIMSS 2011-2015) at 0.52 and students' reports of household books at 0.46. The higher accuracy of parent occupation reports and low accuracy of household books reports is consistent with findings by Jerrim and Micklewright (2014) using some of the same international datasets. In order to estimate reliabilities for other age groups, I assume that 8th grade students report all SES variables with the same reliability as 15-year-old students, but 4th grade students report parent education 80% as reliably and parent occupation and household books 90% as reliably as 15-year-old students. Finally, in order to estimate reliabilities for other years where parent reports are unavailable, I use the average reliabilities for each of these age groups. Since the reliabilities applied to all years are derived from parent reports in recent years, this procedure adjusts only for differences in reliability between parents and children but cannot account for possible changes in reliability over time.

Table H1 reports estimated trends in 90/10 SES achievement gaps, using only student-reported data for all three SES variables (gaps are not adjusted for SES or test reliability). Sample sizes are reduced, mainly due to the omission of 4th grade assessments from recent years with parent-reported education and occupation. In these models, the estimated positive trends in 90/10 SES achievement gaps are reduced compared to those reported in the main paper text but are still positive and highly significant for all three SES variables. Using these results, it is possible to estimate the sensitivity of estimated gap trends to potential increases in the accuracy of students' reports of their parents' SES characteristics. For the increases in SES achievement gaps reported in Table H1 to be fully accounted for by measurement error alone, the reliability of students' reports of parental education, estimated at 0.62 for recent cohorts, would have to be only 0.46 for the 1950 cohort. The reliability of students' reports of parental occupation, estimated at 0.77 for recent cohorts, would have to be 0.50 for the 1950 cohort; and the reliability of students' reports of household books, estimated at 0.46 for recent cohorts, would have to be 0.34 for the 1956 cohort (the first cohort for which the household books variable was collected). Without parental reports, it is impossible to know from these data whether the reliability of students' reports could have increased by 35-55% over this 50-year period. However, a thorough literature search did not reveal published evidence that survey reporting of SES characteristics by either adults or children has become more accurate over time.

Table H1. Estimated trends in 90/10 SES achievement gaps, student-reported SES data only

	No adjustment					
	(H1)			(H2)		
	coef	(se)		coef	(se)	
Parent education gaps intercept	0.739	(0.022)	***	0.752	(0.021)	***
Parent occupation gaps intercept	0.772	(0.024)	***	0.782	(0.024)	***
Household books gaps intercept	0.829	(0.028)	***	0.838	(0.027)	***
Level 1 - Gaps						
Subject (ref=Reading):						
Math	0.018	(0.006)	***	0.019	(0.006)	***
Science	0.026	(0.004)	***	0.026	(0.004)	***
SES variable quality measures						
Number of categories (centered at 7)	-0.002	0.002		-0.001	0.002	
≥ 20% in bottom category	-0.025	0.014	+	-0.033	0.016	*
≥ 20% in top category	-0.105	0.01	***	-0.116	0.009	***
Level 2 - Study-years						
Age at testing (ref=14)						
Age 10 at testing	-0.186	(0.017)	***	-0.185	(0.017)	***
Age 15 at testing	0.014	(0.015)		0.007	(0.014)	
Cohort birth year × Parent education	0.004	(0.001)	***			
Cohort birth year × Parent occupation	0.006	(0.001)	***			
Cohort birth year × Books	0.005	(0.001)	***			
Cohort birth year				0.005	(0.001)	***
Random effects						
Level 2 - Residual variance between studies in…						
Parent education intercepts	0.01962			0.01991		
Parent occupation intercepts	0.01561			0.01553		
Books intercepts	0.01461			0.01514		
Level 3 - Residual variance between countries in…						
Parent education intercepts	0.02999			0.03059		
Parent occupation intercepts	0.03917			0.03960		
Books intercepts	0.06079			0.06187		
Parent education cohort slopes	0.00003					
Parent occupation cohort slopes	0.00002					
Books cohort slopes	0.00004					
Cohort slopes				0.00002		
N (Level 1 - gaps)	4980			4980		
N (Level 2 - studies)	1023			1023		
N (Level 3 - countries)	100			100		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper.

The estimates of the reliability of students' SES reports from recent years (for countries that collected SES from both students and parents) do vary substantially. Reliability estimates of students' reports of parent education from PISA 2012 range from 0.43 in Croatia to 0.91 in Portugal; of parent occupation range from 0.66 in Mexico to 0.92 for Croatia; and of household books from PIRLS 2011 range from 0.03 in Kuwait to 0.67 in Bulgaria. Yet for parent education and occupation, even the countries with the lowest estimated reliabilities barely reach what would need to be the *average* level of reliability for early cohorts in order to fully explain the global increase in SES achievement gaps. For household books, in contrast, there is more variability in the estimated accuracy of students' reports, and it is conceivable that average reliability could have increased by 37% in 49 years.

Estimated SES reliability could potentially also be compared across multiple waves of each study that collects a parent questionnaire (PIRLS 2001-2011 and PISA 2006-2012). Computed SES reliabilities for students and parents do appear to vary somewhat across waves, although it is not clear that these differences represent meaningful trends, given the small number of years and countries represented. It is also possible to compare trends in SES achievement gaps estimated from students' versus parents' reports across waves of PIRLS and PISA. These estimates are reported in Table H2. The estimated trends do not appear to differ systematically depending on whether they are estimated from parent or student reports (this comparison was done before adjusting for computed SES reliability, as gaps based on parent- and student-reported data will be nearly identical after adjustment by construction). Wald tests for joint null hypotheses that trends are equal for gaps based on parent- and student-reported SES cannot be rejected in any model except one. Correlations between country-specific random cohort slopes for gaps based on parent- and student-reported SES are also strongly positive all but one model (this excludes models predicting parent occupation gaps, where random cohort slopes could not be computed due to insufficient sample size). The exception in both cases is the model predicting parent education gaps in PISA science scores, which also showed negative trends in Appendix B. The results in Table H2 suggest this finding for science gaps based on student-reported parent education may be inaccurate, as it is inconsistent with the results for science gaps based on parent-reported education, which are likely to be more reliably estimated.

Table H2. Trends in Gaps by Student-Reported and Parent-Reported SES, by Test

Test instrument	PISA Math	PISA Reading	PISA Science	PISA Math	PISA Reading	PISA Science	PIRLS Reading
Test years	2006-2012	2006-2012	2006-2012	2006-2012	2006-2012	2006-2012	2001-2011
Cohort birth years	1991-1997	1991-1997	1991-1997	1991-1997	1991-1997	1991-1997	1991-2001
SES variable	Education	Education	Education	Occupation	Occupation	Occupation	Books
Parent education (parent report) intercept	0.828 ***	0.747 ***	0.810 ***				
	(0.066)	(0.069)	(0.070)				
Parent education (student report) intercept	0.747 ***	0.681 ***	0.744 ***				
	(0.064)	(0.064)	(0.069)				
Parent occupation (parent report) intercept				0.690 ***	0.649 ***	0.672 ***	
				(0.102)	(0.102)	(0.107)	
Parent occupation (student report) intercept				0.673 ***	0.629 ***	0.661 ***	
				(0.115)	(0.116)	(0.120)	
Household books (parent report) intercept							0.807 ***
							(0.029)
Household books (student report) intercept							0.756 ***
							(0.037)
Level 2 - Study-years							
Cohort birth year × Parent ed. (parent)	0.019 **	0.029 ***	0.014 *				
	(0.006)	(0.009)	(0.006)				
Cohort birth year × Parent ed. (student)	0.012 ***	0.015 *	-0.001				
	(0.003)	(0.006)	(0.007)				
Cohort birth year × Parent occ. (parent)				0.013 *	0.013 *	0.008	
				(0.005)	(0.005)	(0.005)	
Cohort birth year × Parent occ. (student)				0.017 **	0.014 *	0.010 +	
				(0.006)	(0.007)	(0.005)	
Cohort birth year × Books (parent)							0.005 *
							(0.003)
Cohort birth year × Books (student)							0.009 **
							(0.003)
N (Level 1 - gaps)	66	66	66	28	28	28	198
N (Level 2 - study-years)	33	33	33	14	14	14	99
N (Level 3 - countries)	13	13	13	7	7	7	40
p-value for H0: Cohort × SES (parent) = Cohort × SES (student)	0.233	0.025	0.071	0.272	>.500	>.500	0.118
Correlation between country random cohort slopes for parent- and student-reported SES	0.761	0.744	-0.066	a	a	a	0.951

+ p<.1, * p<.05, ** p<.01, *** p < .001 Note: To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L).

[a] Country-level random cohort slopes could not be included in the models predicting parent occupation gaps due to an insufficient sample size.

Increased accuracy of the SES variables could result not only from improvements in students' knowledge of their family SES characteristics but also from improved questionnaire wording over time. The wording of the background questionnaires differs across the older IEA studies, the new IEA studies (TIMSS and PIRLS), and PISA. However, the wording of background questionnaires has changed very little across multiple waves of the trend studies (TIMSS, PIRLS, and PISA). Yet, as demonstrated in Appendix B, trends in SES achievement gaps estimated from each of these studies individually are still mostly positive. Between 1995 and 2001, the IEA began including drawings of bookshelves in its TIMSS and PIRLS questionnaires for 4th grade students in order to assist them in estimating the number of books they have at home, which likely decreased measurement error and could bias estimates of books achievement gap trends upward. However, as we have seen, books achievement gaps also increased substantially for 8th grade and 15-year-old students, with no added drawings.

The one consistent exception to the increasing SES achievement gaps across all test instruments in Appendix B is trends estimated from PISA parent education, which are usually negative. This is especially surprising since 8th grade TIMSS tests a similar population in similar subjects and shows large increases in parent education achievement gaps. In the Appendix B results, it is difficult to discern whether the discrepancy between PISA and TIMSS 8th grade is the result of differences in samples, in test instruments, or in the measurement of SES. Table H3 estimates trends in parent education and household books achievement gaps for math and science in a constant sample of 42 countries that participated in at least two cycles each of PISA and TIMSS 8th grade. Trends in parent education gaps are close to 0 for PISA but are large and positive for TIMSS. Although trends in household books gaps are also larger for TIMSS than for PISA, they are nevertheless still large, positive, and significant for PISA. Wald tests for joint null hypotheses that trends are equal for PISA and TIMSS gaps can be rejected for both parent education and books gaps. Correlations between country-specific random cohort slopes for PISA and TIMSS gaps are weak for parent education but very strongly positive (over 0.90) for books gaps. These comparisons suggest that the difference in trends for PISA and TIMSS may be attributable not only to differences in test instruments and target populations but also likely due to differences in the measurement of SES, particularly parent education.

Table H3. Comparison of Trends in PISA and TIMSS 8th Grade 90/10 Parent Education and Household Books Gaps

	Math		Science	
Parent education intercept × PISA	0.720	***	0.713	***
	(0.033)		(0.031)	
Parent education intercept × TIMSS	0.765	***	0.738	***
	(0.032)		(0.034)	
Household books intercept × PISA	0.917	***	0.898	***
	(0.038)		(0.039)	
Household books intercept × TIMSS	0.799	***	0.800	***
	(0.037)		(0.044)	
Level 2 - Study-years				
Cohort birth year × Parent education × PISA	0.001		0.000	
	(0.002)		(0.002)	
Cohort birth year × Parent education × TIMSS	0.010	***	0.011	***
	(0.002)		(0.002)	
Cohort birth year × Books × PISA	0.005	**	0.007	***
	(0.002)		(0.002)	
Cohort birth year × Books × TIMSS	0.014	***	0.018	***
	(0.003)		(0.003)	
N (Level 1 - gaps)	711		711	
N (Level 2 - study-years)	358		358	
N (Level 3 - countries)	42		42	
p-value for H0: Cohort × Education × PISA = Cohort × Education × TIMSS	0.003		<0.001	
p-value for H0: Cohort × Books × PISA = Cohort × Books × TIMSS	0.001		<0.001	
Correlation between country random cohort slopes for PISA and TIMSS education gaps	-0.077		0.267	
Correlation between country random cohort slopes for PISA and TIMSS books gaps	0.941		0.924	

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L).

Table H4 compares the parent education item wording from the student questionnaire for PISA and the student/parent questionnaires for the IEA studies (TIMSS and PIRLS). TIMSS has a single education item for each parent; PISA has two items (schooling and higher education). TIMSS lists education levels in ascending order; PISA lists them in descending order. TIMSS includes an option for "I don't know"; PISA does not.[7] The highest two educational categories in TIMSS are "<ISCED Level 5A, first degree>" (i.e., BA) and "Beyond <ISCED Level 5A, first degree>" (i.e., MA, PhD, and professional degrees); the highest two categories in PISA are "<ISCED level 5A>" (i.e., BA, MA, and professional degrees) and "<ISCED level 6>" (i.e.,

[7] "I don't know" responses in TIMSS were treated as missing data and imputed as part of the multiple imputation procedure. TIMSS 8th grade still shows large and significant increases in parent education achievement gaps using unimputed data with listwise deletion of "I don't know" responses and other missing parent education data.

PhD).[8] Perhaps as a result of some or all of these differences, a substantially larger share of students select highest two categories (BA or above) in PISA than in TIMSS. Of the 29 countries participating in both PISA 2012 and TIMSS 2011 8th grade, 20 countries had a higher share of students reporting BA or above in PISA than TIMSS, by an average of 6 percentage points. For example, in Australia, 31% of TIMSS 8th grade 2011 students report BA or more, while 43% of PISA 2012 students report BA or more; in Finland, 42% of TIMSS 8th grade 2011 students report BA or more, while 55% of PISA 2012 students report BA or more. This pattern is in addition to the general increase in the share of students in the highest education categories seen across all datasets due to educational upgrading in the parents' generations. The larger share of students in the highest category in PISA means that the achievement at the 90th percentile of parent education is estimated less precisely. This means that parent education achievement gaps may be underestimated in more recent years, and consequently that gap trends may be underestimated.

Table H4. Comparison of PISA and TIMSS parent education questionnaire wording

PISA 2009-2015 student questionnaire	TIMSS 2003-2011 8th grade student questionnaire, TIMSS 2011 4th grade parent questionnaire, PIRLS 2006-2011 parent questionnaire
Q14. What is the <highest level of schooling> completed by your mother? • <ISCED level 3A> • <ISCED level 3B, 3C> • <ISCED level 2> • <ISCED level 1> • She did not complete <ISCED level 1> Q15. Does your mother have any of the following qualifications? • <ISCED level 6> • <ISCED level 5A> • <ISCED level 5B> • <ISCED level 4>	6A. What is the highest level of education completed by your mother <or stepmother or female guardian>? • Some <ISCED Level 1 or 2> or did not go to school • <ISCED Level 2> • <ISCED Level 3> • <ISCED Level 4> • <ISCED Level 5B> • <ISCED Level 5A, first degree> • Beyond <ISCED Level 5A, first degree> • I don't know

Note: All ISCED levels listed here refer to the ISCED 1997 scheme. Beginning in 2015, TIMSS updated its parent education item wordings to reflect the new ISCED 2011 scheme. PISA had not yet made any update in its 2015 cycle.

Models in the main text attempt to account for differences in variable quality across different studies or variables that may confound trend estimates. This is done by controlling for four SES variable quality measures (parent vs. student reporting, number of categories, 20% of

[8] All ISCED levels listed here refer to the ISCED 1997 scheme. Due to likely inaccuracies in students' reports of very high levels of parental education in PISA, in all analyses, I combine the highest two educational categories in PISA: ISCED 5A (BA, MA, and professional degrees) and ISCED 6 (PhD). Very large shares of students in PISA report at least one parent holding a PhD (over 10% in some countries), far outnumbering national statistics on PhD attainment. When the ISCED 5A and 6 categories are kept separate, results are broadly similar, but PISA parent education gaps are even smaller than when the categories are combined, due to the large number of low-achieving students reporting parents holding PhDs.

more students in the bottom category, and 20% or more students in the top category). Model H3 in Table H5 omits these controls (Model 1 is reproduced from the main text for comparison). Results for Model H3 are nearly identical to those for Model 1. The trend in parent education achievement gaps increases very slightly, from 0.007 SD of achievement per year to 0.008 SD of achievement per year. Thus, the trend reported in the main text controlling for SES variable quality measures is a slightly more conservative estimate of the trend. As in the main text, a Wald joint test of the hypothesis that the trends are equal across all three SES variables cannot be rejected (p = 0.11).

Table H5. Comparison of Trends in Gaps, Omitting SES Variable Quality Measures

	(1)			(H3)		
	with SES quality measures			no SES quality measures		
	coef	(se)		coef	(se)	
Parent education gaps intercept	1.032	(0.030)	***	0.989	(0.028)	***
Parent occupation gaps intercept	0.958	(0.030)	***	0.927	(0.028)	***
Household books gaps intercept	1.299	(0.041)	***	1.234	(0.039)	***
Level 1 - Gaps						
Subject (ref=Reading)						
Math	0.020	(0.007)	**	0.020	(0.007)	**
Science	0.034	(0.005)	***	0.034	(0.005)	***
SES variable quality measures						
Parent-reported × Parent education	0.132	(0.030)	***			
Parent-reported × Parent occupation	0.075	(0.025)	**			
Parent-reported × Books	-0.039	(0.029)				
Number of categories (centered at 7)	0.003	(0.003)				
≥ 20% in bottom category	-0.065	(0.021)	**			
≥ 20% in top category	-0.135	(0.013)	***			
Level 2 - Study-years						
Age at testing (ref=14)						
Age 10 at testing	-0.170	(0.024)	***	-0.136	(0.022)	***
Age 15 at testing	-0.024	(0.020)		-0.034	(0.020)	+
Cohort birth year × Parent education	0.007	(0.001)	***	0.008	(0.001)	***
Cohort birth year × Parent occupation	0.007	(0.001)	***	0.007	(0.001)	***
Cohort birth year × Books	0.008	(0.001)	***	0.008	(0.001)	***
Random effects						
Level 2 - Residual variance between studies in…						
Parent education intercepts	0.03736			0.04505		
Parent occupation intercepts	0.02322			0.02435		
Books intercepts	0.03698			0.04404		
Level 3 - Residual variance between countries in…						
Parent education intercepts	0.05426			0.05021		
Parent occupation intercepts	0.05227			0.04877		
Books intercepts	0.11590			0.11679		
Parent education cohort slopes	0.00004			0.00004		
Parent occupation cohort slopes	0.00003			0.00003		
Books cohort slopes	0.00007			0.00007		
N (Level 1 - gaps)	5541			5541		
N (Level 2 - study-years)	1026			1026		
N (Level 3 - countries)	100			100		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper.

Since the 90/10 percentile method may not perform as well when more than 20% of observations are in the bottom or top SES category (Reardon 2011a), the models in the main paper include two dummy variables indicating whether 20% or more of students fall into the bottom category (14.8% of country-study gaps) or the top category (38.7% of country-study gaps). Model H4 in Table H6 omits these gaps altogether (Model 1 is reproduced from the main text for comparison). Compared to Model 1, the sample size of Model H4 is substantially reduced—by more than half at level 1 due to the large shares of gaps with many students in the bottom or especially top categories, as reported above. While the estimates of the parent education and books gap trends remain relatively similar, the parent education gap trend is substantially reduced (from 0.007 SD of achievement per year to 0.004 SD per year). This is because parent education gaps are dropped for nearly all recent studies in high-income countries. I believe the estimate of the parent education gap trend in the main paper, while imprecise, is more accurate than the one in Model H4 omitting over half of the available data. Nevertheless, the estimate of the parent education gap trend is still positive and highly significant in Model H4.

Table H6. Comparison of Trends in Gaps, Dropping Observations with 20% or More of Students in the Top or Bottom SES Variable Category

	including all gaps			dropping if ≥ 20% in top or bottom category		
	(1)			(H4)		
	coef	(se)		coef	(se)	
Parent education gaps intercept	1.032	(0.030)	***	1.071	(0.036)	***
Parent occupation gaps intercept	0.958	(0.030)	***	0.998	(0.037)	***
Household books gaps intercept	1.299	(0.041)	***	1.366	(0.041)	***
Level 1 - Gaps						
Subject (ref=Reading)						
Math	0.020	(0.007)	**	0.004	(0.010)	
Science	0.034	(0.005)	***	0.034	(0.006)	***
SES variable quality measures						
Parent-reported × Parent education	0.132	(0.030)	***	0.129	(0.045)	**
Parent-reported × Parent occupation	0.075	(0.025)	**	0.049	(0.035)	
Parent-reported × Books	-0.039	(0.029)		-0.137	(0.037)	***
Number of categories (centered at 7)	0.003	(0.003)		0.000	(0.003)	
≥ 20% in bottom category	-0.065	(0.021)	**			
≥ 20% in top category	-0.135	(0.013)	***			
Level 2 - Study-years						
Age at testing (ref=14)						
Age 10 at testing	-0.170	(0.024)	***	-0.181	(0.035)	***
Age 15 at testing	-0.024	(0.020)		0.013	(0.029)	
Cohort birth year × Parent education	0.007	(0.001)	***	0.004	(0.002)	**
Cohort birth year × Parent occupation	0.007	(0.001)	***	0.006	(0.001)	***
Cohort birth year × Books	0.008	(0.001)	***	0.009	(0.002)	***
Random effects						
Level 2 - Residual variance between studies in…						
Parent education intercepts	0.03736			0.07398		
Parent occupation intercepts	0.02322			0.03327		
Books intercepts	0.03698			0.02828		
Level 3 - Residual variance between countries in…						
Parent education intercepts	0.05426			0.04230		
Parent occupation intercepts	0.05227			0.05367		
Books intercepts	0.11590			0.09283		
Parent education cohort slopes	0.00004			0.00002		
Parent occupation cohort slopes	0.00003			0.00003		
Books cohort slopes	0.00007			0.00005		
N (Level 1 - gaps)	5541			2577		
N (Level 2 - study-years)	1026			820		
N (Level 3 - countries)	100			91		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper.

In summary, it is likely that changing measurement error in SES confounds estimates of true global changes in SES achievement gaps. But it is not clear that measurement error in SES is uniformly declining over time, as error may be increasing for some variables in some recent tests (i.e., PISA parent education). This would lead gap increases to be *under*- rather than overestimated, making the main trend estimates more conservative. To the extent that the reliability of SES has improved, this increase would have to be very large to fully account for the increase in SES achievement gaps. There is no direct evidence for changes in measurement error of SES, either in the data used in this study or in published research. Future research should identify older datasets containing both children's and parents' reports of family SES characteristics in order to examine whether reliability in student reporting may have increased over time and the size of this possible increase.

Appendix J. Trend models run separately by SES variable

The models in the main text of the paper pool all gap types, regardless of which of the three SES variables was used to calculate gaps (parent education, parent occupation, or household books). Table J1 reports gap trends models run separately by the SES variable from which gaps were calculated. Results are similar to those for pooled models reported in the main text of the paper, though both the intercepts and the cohort birth year slopes are slightly smaller for all three SES variables in these models than in the main text. The discrepancy is mainly due to differences in coefficients for control variables (e.g., subject, SES quality measures, age), which are allowed to vary by SES variable in these models rather than constrained to take the same value for all gap types in the pooled models. However, it is not clear that the different results for control variables for different SES variables are meaningful, as each model in Table J1 includes a somewhat different set of study-years and countries. (Running the models only on country-study-years where all three SES variables are available would reduce the sample by 75%, eliminating a large amount of valuable information.)

In some cases, the differences in coefficients for control variables across models using different SES variables do appear meaningful. For example, the math coefficient in Table 1 in the main text of the paper is 0.02, meaning that, on average when pooling all gap types, math achievement gaps are larger than reading achievement gaps (the reference category) by about 0.02 SDs. In Table J1, the math coefficient when predicting parent education gaps is about 0.06, when predicting parent occupation gaps is about 0.03, and when predicting books gaps is about 0. It is easy to see why household books might truly have a stronger relationship with reading achievement compared to other SES variable-subject pairs. However, these coefficients may still overstate (or understate) the discrepancies across SES variables, as they are based on different samples. More importantly, though, the goal of the present study is not to describe the specific processes by which each type of achievement gap is generated, but instead to examine whether *trends* in gaps across cohorts are similar enough that they could plausibly be driven by the same underlying process of growing SES achievement gaps, broadly defined, and if so, to pool all gap types to obtain the most accurate estimate of that trend. Thus, the pooled models reported in the main text are preferred over those in Table J1, as they give the best summary of trends in SES achievement gaps, using the most complete data available.

The discrepancy between the results in Table J1 and those in the main text is also in small part due to the inability of the separate models to account for error covariances among different gap types, as the multivariate variance-known models do. However, this issue does not affect results substantially, as the results in Appendix L show that pooled models omitting error covariances produce very similar results to models incorporating error covariances.

Table J1. Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90^{th} and 10^{th} Percentiles of SES, Run Separately by SES Variable

	Parent education		Parent occupation		Books	
Intercept	1.016 ***	1.054 ***	0.921 ***	0.903 ***	1.211 ***	1.200 ***
	(0.031)	(0.033)	(0.023)	(0.037)	(0.040)	(0.061)
Level 1 - Gaps						
Subject (ref=Reading):						
Math	0.054 ***	0.055 ***	0.025 ***	0.025 ***	-0.004	-0.004
	(0.010)	(0.010)	(0.007)	(0.007)	(0.009)	(0.009)
Science	0.054 ***	0.054 ***	0.017 ***	0.017 ***	0.041 ***	0.041 ***
	(0.008)	(0.008)	(0.005)	(0.005)	(0.007)	(0.007)
SES variable quality measures						
Parent-reported		0.051		-0.051 +		0.096 ***
		(0.045)		(0.028)		(0.024)
Number of categories (centered at 7)		-0.002		0.040 **		-0.045
		(0.003)		(0.014)		(0.028)
≥ 20% in bottom category		-0.117 **		0.066		-0.120 **
		(0.044)		(0.087)		(0.037)
≥ 20% in top category		-0.101 ***		-0.096 **		-0.107 ***
		(0.020)		(0.030)		(0.023)
Level 2 - Study-years						
Age at testing (ref=14)						
Age 10 at testing	-0.003	-0.060	0.017	-0.003	-0.215 ***	-0.247 ***
	(0.030)	(0.054)	(0.019)	(0.027)	(0.028)	(0.028)
Age 15 at testing	-0.154 ***	-0.112 ***			0.105 ***	0.127 **
	(0.025)	(0.024)			(0.025)	(0.039)
Cohort birth year	0.006 ***	0.006 ***	0.004 ***	0.005 ***	0.008 ***	0.006 ***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Level 2 residual variance in intercepts	0.03390	0.03194	0.01852	0.01687	0.03395	0.03243
Level 3 residual variance in intercepts	0.04977	0.05164	0.03793	0.03616	0.11276	0.10128
Level 3 residual variance in cohort slopes	0.00005	0.00005	0.00004	0.00004	0.00008	0.00007
N (Level 1 - gaps)	1916	1916	1405	1405	2189	2189
N (Level 2 - study-years)	852	852	590	590	993	993
N (Level 3 - countries)	94	94	82	82	100	100

+ p<.1, * p<.05, ** p<.01, *** p < .001.

Note: In models predicting achievement gaps based on parent occupation, the dummy variable for age 15 is omitted, as there are very few age-14 assessments that collected parent occupation. Therefore, in the parent occupation models, the reference category is assessments of students who are age 14 *or* 15. All models in this table specify known level 1 error variances and covariances among gaps from different test subjects (but not different SES variables, as they are in separate models), estimated via bootstrapping.

Appendix K. SES achievement gap trends by age and subject

Table K1 reports gap trends models estimating different cohort birth year trends by level of schooling (primary versus secondary), rather than by SES variable as in the main text models. SES achievement gaps have increased substantially for both primary school students (0.007 SD per year) and secondary school students (0.009 SD per year). Both estimates are statistically significant ($p<0.001$). Gaps for secondary school students tend to be slightly larger than those for primary school students, and gaps for secondary students have also increased somewhat more than those for primary students. However, the primary and secondary school trend estimates are similar in size; a joint test of the null hypothesis that the two cohort slopes are equal cannot be rejected ($p=0.148$). Within countries, trends in gaps for primary and secondary school students tend to be similar: the correlation between country random cohort slopes for primary and secondary school gaps is positive and moderately strong ($r=0.472$). As differences in primary and secondary trends are small and comparing trends across age groups is not of substantive interest in the current project, I consider it justifiable to pool gaps from primary and secondary students in the main text models.

Table K2 reports gap trends models estimating different cohort birth year trends by test subject (math, science, or reading), rather than by SES variable as in the main text models. SES achievement gaps have increased substantially in reading tests (0.012 SD per year), math tests (0.010 SD per year), and science tests (0.012 SD per year). All three estimates are statistically significant ($p<0.001$). In the earliest birth cohorts, SES gaps in math achievement tended to be somewhat larger than those in reading or science achievement. However, SES gaps in reading and science achievement have increased at a slightly faster pace than those in math achievement, resulting in SES gaps that are very similar in size across the three subjects in the most recent cohorts. Yet the trend estimates for all three test subjects are similar in size; a joint test of the null hypothesis that the three cohort slopes are equal is only marginally significant ($p=0.074$). Within countries, trends in gaps based on different test subjects are very similar: pairwise correlations between country random slopes for different test subjects range from 0.785 to 0.854. As differences in trends by test subject are small and comparing trends across subjects is not of substantive interest in the current project, I consider it justifiable to pool gaps from all three test subjects in the main text models.

Table K1. Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th Percentiles of SES, Estimating Different Trends by Age (Primary or Secondary School)

	coef	(se)	
Primary school gaps intercept	0.988	(0.068)	***
Secondary school gaps intercept	1.079	(0.032)	***
Level 1 - Gaps			
Subject (ref=Reading):			
Math	0.021	(0.007)	**
Science	0.035	(0.005)	***
SES variable (ref=Parent education)			
Parent occupation	-0.005	(0.021)	
Household books	0.214	(0.035)	***
SES variable quality measures			
Parent-reported	0.021	(0.063)	
Parent-reported × Parent occupation	-0.072	(0.052)	
Parent-reported × Books	0.002	(0.071)	
Number of categories (centered at 7)	-0.013	(0.005)	**
≥ 20% in bottom category	-0.212	(0.038)	***
≥ 20% in top category	-0.035	(0.022)	
Level 2 - Study-years			
Age 15 at testing	-0.185	(0.029)	***
Primary × Parent occupation	-0.040	(0.046)	
Primary × Books	-0.112	(0.065)	+
Cohort birth year × Primary	0.007	(0.002)	***
Cohort birth year × Secondary	0.009	(0.001)	***
Random effects			
Residual variance between studies	0.02562		
Residual variance between countries in…			
Primary school intercepts	0.05685		
Secondary school intercepts	0.06327		
Primary school cohort slopes	0.00003		
Secondary school cohort slopes	0.00005		
N (Level 1 - gaps)	5541		
N (Level 2 - study-years)	1026		
N (Level 3 - countries)	100		
p-value for H0: Cohort × Primary = Cohort × Secondary	0.148		
Correlation between country random cohort slopes for primary and secondary gaps	0.472		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper.

Table K2. Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90^{th} and 10^{th} Percentiles of SES, Estimating Different Trends by Test Subject (Math, Science, or Reading)

	coef	(se)	
Reading gaps intercept	1.025	(0.032)	***
Math gaps intercept	1.037	(0.031)	***
Science gaps intercept	1.047	(0.032)	***
Level 1 - Gaps			
SES variable (ref=Parent education)			
Parent occupation	-0.073	(0.021)	***
Household books	0.362	(0.033)	***
SES variable quality measures			
Parent-reported	-0.034	(0.052)	
Parent-reported × Parent occupation	-0.053	(0.027)	+
Parent-reported × Books	-0.227	(0.036)	***
Number of categories (centered at 7)	-0.005	(0.005)	
≥ 20% in bottom category	-0.202	(0.037)	***
≥ 20% in top category	-0.102	(0.018)	***
Level 2 - Study-years			
Age at testing (ref=14)			
Age 10 at testing	-0.135	(0.031)	***
Age 15 at testing	-0.018	(0.019)	
Cohort birth year × Reading	0.012	(0.001)	***
Cohort birth year × Math	0.010	(0.001)	***
Cohort birth year × Science	0.012	(0.001)	***
Random effects			
Level 2 - Residual variance between studies in…			
Reading intecepts	0.03293		
Math intercepts	0.03216		
Science intercepts	0.03707		
Level 3 - Residual variance between countries in…			
Reading intercepts	0.06441		
Math intercepts	0.06027		
Science intercepts	0.07121		
Reading cohort slopes	0.00006		
Math cohort slopes	0.00003		
Science cohort slopes	0.00004		
N (Level 1 - gaps)	5541		
N (Level 2 - study-years)	1026		
N (Level 3 - countries)	100		
p-value for H0: Cohort × Reading = Cohort × Math = Cohort × Science	0.074		
Correlation between country random cohort slopes for reading and math gaps	0.808		
Correlation between country random cohort slopes for reading and science gaps	0.854		
Correlation between country random cohort slopes for math and science gaps	0.785		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper.

Appendix L. Specification of trend model

The coefficients for some control variables were omitted from the tables in the main text due to space constraints. Tables L1 and L2 report all coefficients from Table 2. Table L1 reports those from Model 3 (with country region interactions), and Table L2 reports all coefficients from Models 4-7.

Figure 3 in the main text showed estimated trends in SES achievement gaps for only 24 selected countries due to space constraints. Figures L1-L3 show estimated gap trends for a larger set of countries. Trend lines are derived from shrunken empirical Bayes estimates from Model 5 (Table 2 and Table L2). Countries with fewer than 7 study-years or whose data span fewer than 15 cohort birth years are excluded from the figures, as their trend estimates are imprecise. This leaves 55 countries displayed in the figures (of 100 countries in the full sample). For the remaining 45 countries with less available data, true country-specific gap trends are very uncertain. However, these countries still contribute to the estimation of the average global gap trend in the hierarchical growth curve models.

Table L1. All Coefficients from Model 3 Predicting Achievement Gaps between 90th and 10th Percentiles of SES with Interactions by Country Region

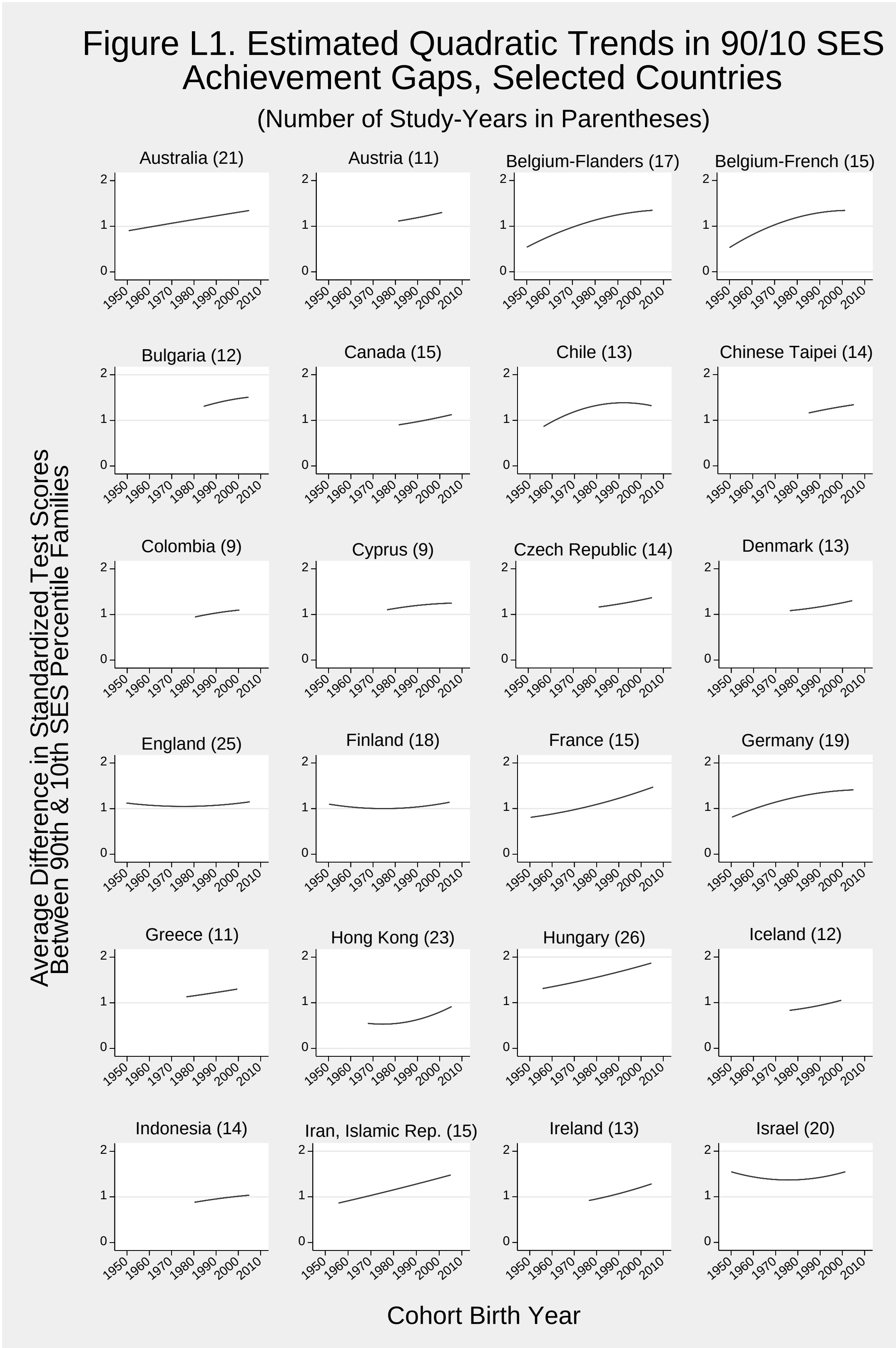
	(3) Region Interactions coef	(se)	
Parent education gaps intercept	1.126	(0.027)	***
Parent occupation gaps intercept	1.111	(0.029)	***
Household books gaps intercept	1.563	(0.038)	***
Level 1 - Gaps			
Subject (ref=Reading)			
Math	0.020	(0.007)	**
Science	0.034	(0.005)	***
SES variable quality measures			
Parent-reported × Parent education	0.111	(0.031)	***
Parent-reported × Parent occupation	0.069	(0.024)	**
Parent-reported × Books	-0.018	(0.026)	
Number of categories (centered at 7)	0.002	(0.003)	
≥ 20% in bottom category	-0.045	(0.020)	*
≥ 20% in top category	-0.147	(0.013)	***
Level 2 - Study-years			
Age at testing (ref=14)			
Age 10 at testing	-0.171	(0.024)	***
Age 15 at testing	-0.028	(0.020)	
Cohort birth year	0.008	(0.001)	***
Level 3 - Countries			
Region (ref=Western) × Intercept interactions			
Sub-Saharan Africa × Parent education	-0.307	(0.089)	***
Sub-Saharan Africa × Parent occupation	-0.373	(0.113)	**
Sub-Saharan Africa × Books	-0.707	(0.143)	***
East Asia & Pacific × Parent education	-0.160	(0.082)	+
East Asia & Pacific × Parent occupation	-0.316	(0.074)	***
East Asia & Pacific × Books	-0.402	(0.097)	***
Middle East & N. Africa × Parent education	-0.248	(0.078)	**
Middle East & N. Africa × Parent occupation	-0.317	(0.069)	***
Middle East & N. Africa × Books	-0.731	(0.082)	***
E. Europe & CIS × Parent education	-0.057	(0.062)	
E. Europe & CIS × Parent occupation	-0.091	(0.051)	+
E. Europe & CIS × Books	-0.162	(0.064)	*
Latin America & Caribbean × Parent education	0.077	(0.051)	
Latin America & Caribbean × Parent occupation	-0.022	(0.050)	
Latin America & Caribbean × Books	-0.220	(0.070)	**
Region (ref=Western) × Cohort interactions			
Sub-Saharan Africa × Cohort	0.004	(0.004)	
East Asia & Pacific × Cohort	-0.001	(0.002)	
Middle East & N. Africa × Cohort	-0.001	(0.004)	
E. Europe & CIS × Cohort	-0.001	(0.002)	
Latin America & Caribbean × Cohort	-0.009	(0.004)	*
Random Effects			
Level 2 - Residual variance between studies in…			
Parent education intercepts	0.03785		
Parent occupation intercepts	0.02247		
Books intercepts	0.03825		
Level 3 - Residual variance between countries in…			
Parent education intercepts	0.04199		
Parent occupation intercepts	0.03086		
Books intercepts	0.05514		
Cohort slopes	0.00003		
N (Level 1 - gaps)	5541		
N (Level 2 - study-years)	1026		
N (Level 3 - countries)	100		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: See Table A1 for country regions.

Table L2. All Coefficients from Models 4-7 Predicting Achievement Gaps between 90th, 50th, and 10th Percentiles of SES with Interactions by Country Income Level

	(4) Income Interaction			(5) Quadratic			(6) 90/50 Gap			(7) 50/10 Gap		
	coef	(se)		coef	(se)		coef	(se)		coef	(se)	
Parent education gaps intercept	1.074	(0.035)	***	1.066	(0.037)	***	0.548	(0.019)	***	0.514	(0.022)	***
Parent occupation gaps intercept	1.054	(0.038)	***	1.046	(0.042)	***	0.564	(0.020)	***	0.481	(0.022)	***
Household books gaps intercept	1.434	(0.047)	***	1.426	(0.049)	***	0.622	(0.022)	***	0.806	(0.029)	***
Level 1 - Gaps												
Subject (ref=Reading)												
Math	0.020	(0.007)	**	0.020	(0.007)	**	0.021	(0.004)	***	0.009	(0.004)	*
Science	0.034	(0.005)	***	0.034	(0.005)	***	0.022	(0.003)	***	0.013	(0.003)	***
SES variable quality measures												
Parent-reported × Parent education	0.113	(0.031)	***	0.112	(0.031)	***	0.105	(0.018)	***	0.016	(0.022)	
Parent-reported × Parent occupation	0.072	(0.024)	**	0.071	(0.024)	**	0.056	(0.017)	**	0.019	(0.017)	
Parent-reported × Books	-0.018	(0.027)		-0.018	(0.027)		0.140	(0.022)	***	-0.143	(0.022)	***
Number of categories (centered at 7)	0.002	(0.003)		0.002	(0.003)		0.002	(0.002)		0.001	(0.002)	
≥ 20% in bottom category	-0.061	(0.021)	**	-0.061	(0.021)	**	-0.055	(0.017)	**	-0.009	(0.020)	
≥ 20% in top category	-0.148	(0.013)	***	-0.147	(0.013)	***	-0.053	(0.011)	***	-0.086	(0.010)	***
Level 2 - Study-years												
Age at testing (ref=14)												
Age 10 at testing	-0.169	(0.024)	***	-0.163	(0.024)	***	-0.120	(0.016)	***	-0.041	(0.014)	**
Age 15 at testing	-0.023	(0.020)		-0.021	(0.020)		-0.044	(0.013)	**	0.020	(0.012)	+
Cohort birth year	0.007	(0.001)	***	0.007	(0.001)	***	0.001	(0.001)		0.006	(0.001)	***
Cohort birth year^2				0.00005	(0.00005)							
Level 3 - Countries												
Mid/low-income country × Intercept interactions												
Mid/low-income country × Parent education	-0.060	(0.048)		-0.040	(0.048)		0.007	(0.025)		-0.062	(0.029)	*
Mid/low-income country × Parent occupation	-0.166	(0.047)	***	-0.143	(0.050)	**	-0.007	(0.027)		-0.162	(0.028)	***
Mid/low-income country × Books	-0.251	(0.066)	***	-0.231	(0.066)	***	-0.054	(0.031)	+	-0.195	(0.042)	***
Mid/low-income country × Cohort	0.002	(0.002)		0.001	(0.002)		0.003	(0.001)	*	-0.001	(0.001)	
Mid/low-income country × Cohort^2				-0.00022	(0.00010)	*						
Random Effects												
Level 2 - Residual variance between studies in…												
Parent education intercepts	0.03822			0.03677			0.01671			0.01362		
Parent occupation intercepts	0.02302			0.02154			0.00877			0.00634		
Books intercepts	0.03839			0.03734			0.01715			0.01594		
Level 3 - Residual variance between countries in…												
Parent education intercepts	0.05256			0.05227			0.01294			0.01832		
Parent occupation intercepts	0.04810			0.05095			0.01448			0.01655		
Books intercepts	0.10677			0.10841			0.02128			0.03906		
Cohort slopes	0.00003			0.00003			0.00001			0.00001		
N (Level 1 - gaps)	5541			5541			5541			5541		
N (Level 2 - study-years)	1026			1026			1026			1026		
N (Level 3 - countries)	100			100			100			100		

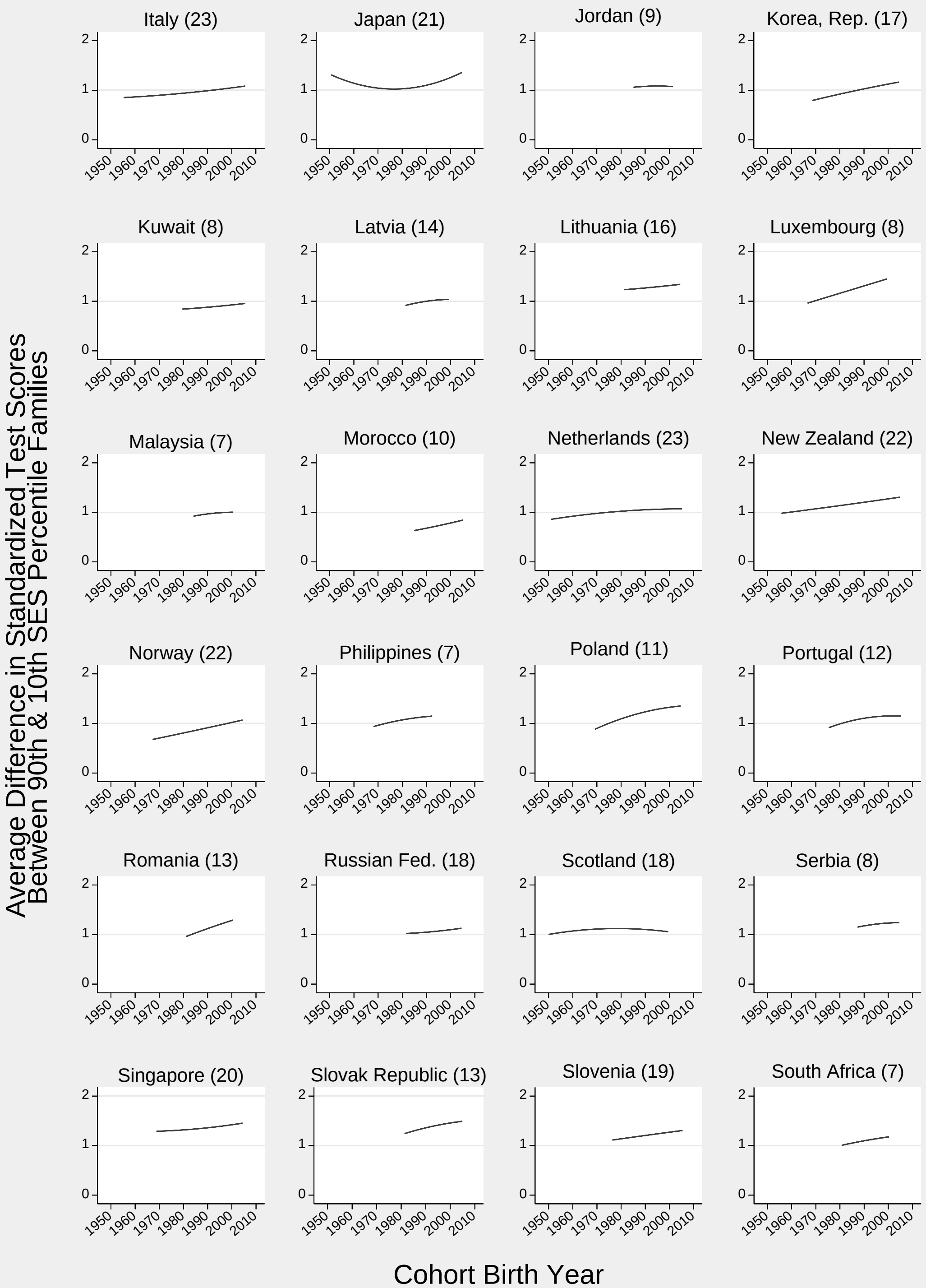
+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: “Middle/low income” countries had GDPs per capita of less than $6000 in 1980 (the reference category is high-income countries; see Table A1 for coding).



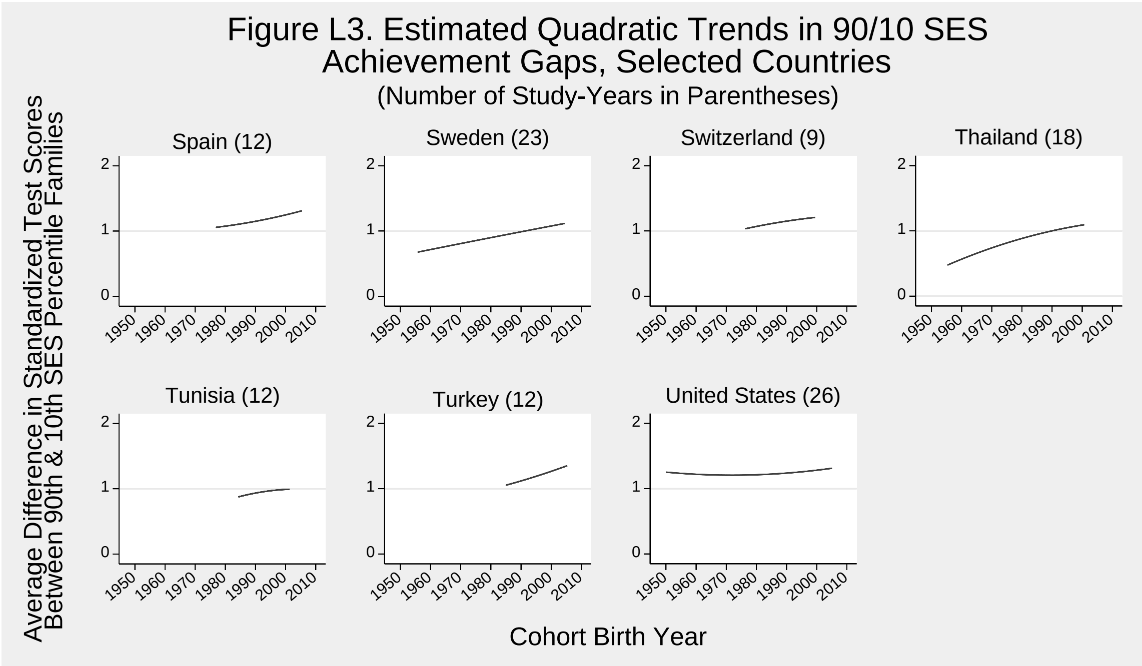
Notes: Countries sorted alphabetically. Trend lines are derived from shrunken empirical Bayes estimates from Model 5 (Table 2 and Table L2). Fixed values for control variables: SES=parent education, subject=math, all others=0 or reference category.

Figure L2. Estimated Quadratic Trends in 90/10 SES Achievement Gaps, Selected Countries

(Number of Study-Years in Parentheses)



Notes: Countries sorted alphabetically. Trend lines are derived from shrunken empirical Bayes estimates from Model 5 (Table 2 and Table L2). Fixed values for control variables: SES=parent education, subject=math, all others=0 or reference category.



Notes: Countries sorted alphabetically. Trend lines are derived from shrunken empirical Bayes estimates from Model 5 (Table 2 and Table L2). Fixed values for control variables: SES=parent education, subject=math, all others=0 or reference category.

Tables L3 and L4 report coefficients from six alternate specifications of the main gap trend model from the main text of the paper (Model 2 in Table 1). The first column of table L3 replicates Model 2 for comparison. All models in the main paper are multivariate variance-known models that specify the known the Level 1 (within-study) variance structure as equal to the estimated error variance-covariance matrix among different gap types in each country-study. Thus, the models take into account that the errors of different gap types within the same country-study are not independent. Error variances and covariances are estimated via bootstrapping. Due to the very long computation time of 1000 bootstrap estimates of 5541 SES achievement gaps, Models L1 and L2 examine the robustness of the results to excluding error covariances from level 1 while still specifying known error variances estimated via bootstrapping (Model L1) and to specifying known error variances estimated using the conventional normality-assumption-based formula (following Reardon (2011b)) rather than bootstrapping (Model L2). Results for both models are nearly identical to Model 2. Model L3 omits any known error variance from level 1 and freely estimates level 1 variance within the hierarchical model, meaning that the model no longer gives greater weight to more precisely-estimated gaps. Results again are nearly identical. As omitting error covariances and using conventional non-bootstrap error variances appears to produce very similar results, some subsequent models (L4, L5, and L6) and several others throughout the appendices use this alternate level 1 variance specification to avoid excessive computation time. They type of error variance computation used in each model is noted under the table where the model is reported.

Table L3. Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th Percentiles of Parent Education (Variety of Trend Model Specifications)

	(2)	(L1)	(L2)	(L3)
	Main text (with BS error covariance)	With BS, no error covariance	Non-BS, no error covariance	No weights
Parent education intercept	1.039 ***	1.034 ***	1.044 ***	1.035 ***
	(0.030)	(0.030)	(0.031)	(0.030)
Parent occupation intercept	0.964 ***	0.959 ***	0.962 ***	0.955 ***
	(0.030)	(0.030)	(0.031)	(0.029)
Household books intercept	1.294 ***	1.291 ***	1.301 ***	1.292 ***
	(0.041)	(0.041)	(0.040)	(0.040)
Level 1 - Gaps				
Subject (ref=Reading):				
Math	0.020 **	0.018 *	0.006	0.014 +
	(0.007)	(0.007)	(0.008)	(0.007)
Science	0.034 ***	0.031 ***	0.026 ***	0.033 ***
	(0.005)	(0.006)	(0.007)	(0.006)
SES variable quality measures				
Parent-reported × Parent education	0.112 ***	0.111 ***	0.080 **	0.117 ***
	(0.031)	(0.030)	(0.027)	(0.031)
Parent-reported × Parent occupation	0.073 **	0.069 **	0.051 *	0.074 **
	(0.024)	(0.023)	(0.021)	(0.025)
Parent-reported × Books	-0.017	-0.017	-0.067 **	-0.001
	(0.026)	(0.026)	(0.023)	(0.029)
Number of categories (centered at 7)	0.002	0.003	0.003	0.003
	(0.003)	(0.003)	(0.003)	(0.003)
≥ 20% in bottom category	-0.063 **	-0.062 **	-0.066 **	-0.072 ***
	(0.021)	(0.021)	(0.021)	(0.020)
≥ 20% in top category	-0.146 ***	-0.143 ***	-0.136 ***	-0.144 ***
	(0.013)	(0.013)	(0.012)	(0.013)
Level 2 - Study-years				
Age at testing (ref=14)				
Age 10 at testing	-0.168 ***	-0.163 ***	-0.147 ***	-0.166 ***
	(0.024)	(0.024)	(0.023)	(0.024)
Age 15 at testing	-0.023	-0.021	-0.027	-0.022
	(0.020)	(0.020)	(0.019)	(0.020)
Cohort birth year	0.007 ***	0.007 ***	0.008 ***	0.007 ***
	(0.001)	(0.001)	(0.001)	(0.001)
Random effects				
Level 1 - Residual variance between gaps				0.00774
Level 2 - Residual variance between studies in…				
Parent education gaps	0.03831	0.03972	0.02719	0.04135
Parent occupation gaps	0.02284	0.02309	0.01877	0.02297
Household books gaps	0.03823	0.03881	0.03187	0.04111
Level 3 - Residual variance between countries in…				
Parent education gaps	0.05362	0.05231	0.05517	0.05121
Parent occupation gaps	0.05330	0.05305	0.05490	0.05326
Household books gaps	0.12149	0.12167	0.12270	0.11918
Cohort slopes	0.00003	0.00003	0.00003	0.00003
N (Level 1 - gaps)	5541	5541	5541	5541
N (Level 2 - study-years)	1026	1026	1027	1027
N (Level 3 - countries)	100	100	100	100

+ p<.1, * p<.05, ** p<.01, *** p < .001

Model L4 uses gaps computed without multiple imputation of missing student-level data; listwise deletion of student observations with missing data on any variables is used instead. Level 1 (gaps) sample sizes are slightly smaller. Point estimates for coefficients are similar, though the standard errors on the intercepts and cohort trend increase substantially, and the cohort trend loses significance. These changes are due to increased heterogeneity across countries in gap trends estimated when dropping missing data (though note that imputation models were estimated separately for each country-study).

Model L5 uses gaps for which the 90th and 10th percentiles of each SES categorical variable were interpolated from linear rather than cubic weighted least squares models. Level 1 (gaps) and Level 2 (study-years) sample sizes are larger because a slightly larger number of gaps can be estimated reliably from these simpler models. The cohort birth year coefficient is slightly larger (i.e., the estimated increase in SES achievement gaps is larger) because the additional observations are very small gaps in early years. Cubic models are retained as the preferred models, as the trend estimates are more conservative, cubic gap functions allow more flexibility in the shape of the relationship between SES and achievement, and for comparability with Reardon (2011b).

Model L6 uses gaps between the 75th and 25th percentiles rather than the 90th and 10th percentiles of each SES variable. As expected, the intercepts, cohort birth year coefficient, and coefficients for most control variables are smaller than in models predicting 90/10 gaps, as 75/25 gaps are smaller in magnitude. However, the size of the cohort birth year coefficient is still substantial (0.005 SD per year) and highly significant. These results demonstrate that the finding of increasing SES achievement gaps is robust to using an alternate gap measure that may be more precisely estimated, as the 75th and 25th percentiles are less likely to be extrapolated outside the available SES data in high-income countries with large numbers of students in the top parent education and occupation categories. Additional models predicting 75/50 and 50/25 gaps (not shown) show that the estimated cohort coefficient for 75/50 gaps (0.002) is slightly smaller than the estimated cohort coefficient for 50/25 gaps (0.003). Both coefficients are highly significant ($p < 0.001$). Thus, the finding that gaps have increased more between the middle and bottom of the SES distribution than between the middle and top of the SES distribution—modeled in the main paper text using 50/10 and 90/50 gaps—is also robust to using alternative percentiles of the SES distribution. The 90/10, 90/50, and 50/10 gaps are still retained as the preferred results in the main paper text for comparability with Reardon (2011b).

Model L7 omits the random effect for the cohort birth year slope and instead estimates a fixed cohort slope for all countries. Results are very similar. The random cohort slope is retained as the preferred model, as a chi-square test shows the variance in the random cohort slopes to be significantly different from 0 in the main models and all other trend model specifications.

Table L4. Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th (or 75th and 25th) Percentiles of Parent Occupation (Variety of Trend Model Specifications)

	(L4) No imputation	(L5) Linear gaps	(L6) 75/25	(L7) Fixed slope
Parent education intercept	1.063 ***	0.992 ***	0.561 ***	1.036 ***
	(0.225)	(0.029)	(0.019)	(0.031)
Parent occupation intercept	1.000 ***	0.923 ***	0.555 ***	0.961 ***
	(0.178)	(0.029)	(0.021)	(0.030)
Household books intercept	1.311 ***	1.276 ***	0.794 ***	1.290 ***
	(0.174)	(0.040)	(0.022)	(0.042)
Level 1 - Gaps				
Subject (ref=Reading):				
Math	-0.007	0.013 +	0.002	0.020 **
	(0.010)	(0.007)	(0.006)	(0.007)
Science	0.023 **	0.020 ***	0.005	0.034 ***
	(0.009)	(0.005)	(0.006)	(0.005)
SES variable quality measures				
Parent-reported × Parent education	0.084 **	0.067 *	0.001	0.118 ***
	(0.031)	(0.026)	(0.002)	(0.031)
Parent-reported × Parent occupation	0.013	0.046 *	0.078 ***	0.080 **
	(0.025)	(0.021)	(0.023)	(0.025)
Parent-reported × Books	-0.075 **	-0.087 ***	-0.005	-0.012
	(0.027)	(0.022)	(0.016)	(0.027)
Number of categories (centered at 7)	0.002	0.004	-0.076 ***	0.002
	(0.006)	(0.002)	(0.017)	(0.003)
≥ 20% in bottom category	-0.049 *	-0.037 +	0.004	-0.065 **
	(0.023)	(0.019)	(0.015)	(0.021)
≥ 20% in top category	-0.129 ***	-0.104 ***	-0.069 ***	-0.148 ***
	(0.016)	(0.011)	(0.009)	(0.013)
Level 2 - Study-years				
Age at testing (ref=14)				
Age 10 at testing	-0.142 ***	-0.136 ***	-0.126 ***	-0.165 ***
	(0.024)	(0.022)	(0.015)	(0.024)
Age 15 at testing	-0.031	-0.020	-0.010	-0.021
	(0.022)	(0.019)	(0.012)	(0.020)
Cohort birth year	0.007	0.008 ***	0.005 ***	0.007 ***
	(0.006)	(0.001)	(0.001)	(0.001)
Random effects				
Level 2 - Residual variance between studies in…				
Parent education gaps	0.02784	0.02385	0.01715	0.04126
Parent occupation gaps	0.01356	0.01764	0.00615	0.02701
Household books gaps	0.03188	0.02749	0.01261	0.04209
Level 3 - Residual variance between countries in…				
Parent education gaps	0.05741	0.05079	0.01342	0.05533
Parent occupation gaps	0.04812	0.05529	0.01952	0.05526
Household books gaps	0.12037	0.12524	0.03298	0.12563
Cohort slopes	0.00003	0.00003	0.00002	
N (Level 1 - gaps)	5521	5584	5199	5541
N (Level 2 - study-years)	1026	1033	1015	1026
N (Level 3 - countries)	100	100	99	100

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: To avoid very long computation times, Models L4, L5, and L6 specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Table L3). Model L7 uses bootstrapped error variances and covariances.

Appendix M. Specification of multivariate model

The coefficients for some control variables were omitted from the tables in the main text due to space constraints. Tables M1 and M2 report all coefficients from Table 3 (with time-varying country covariates). Additionally, each table reports two further models predicting 90/50 and 50/10 SES achievement gaps. Table M1 reports all coefficients from models run on the analytic sample of countries with available covariate data; it shows that for not only 90/10 gaps but also 90/50 and 50/10 gaps, cohort trends are nearly identical in the analytic sample to in the full sample. In Table M2 (with time-varying country covariates), the coefficients for country covariates are similar in direction whether predicting 90/50 or 50/10 gaps, but the magnitude and significance of coefficients tends to be larger when predicting 90/50 gaps than 50/10 gaps. Additionally, by comparing the level 2 residual variances for these full models to reduced models that include study fixed effects and controls but no country covariates (not shown), we can compare how well the covariates account for changes in gaps at the top relative to the bottom of the SES distribution. Compared to a reduced model predicting 90/50 gaps, the country covariates in Model M3 explain an additional 5%, 10%, and 6% of the within-country variance in 90/50 gaps based on parent education, occupation and books. In contrast, the covariates in Model M4 explain only an additional 0.2%, 0%, and 5% of the within-country variance in 50/10 gaps based on parent education, occupation, and books. This is consistent with the fact that most of the covariate coefficients are larger and more significant when predicting 90/50 than 50/10 gaps and indicates that the country characteristics examined in this study do a better job of explaining changes in gaps at the top of the SES distribution than at the bottom. As 50/10 gaps have increased greatly in most countries (more than 90/50 gaps have increased), a large share of the within-country variance in 50/10 gaps is explained by the study-year fixed effects. Net of this large global secular increase, however, the country covariates do not tell us much about which countries experience larger or smaller increases in 50/10 gaps.

Table M1. Unstandardized Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th, 50th, and 10th Percentiles of SES

	(2B) 90/10 Gap			(M1) 90/50 Gap			(M2) 50/10 Gap		
	coef	(se)		coef	(se)		coef	(se)	
Parent education gaps intercept	1.066	(0.032)	***	0.569	(0.017)	***	0.491	(0.023)	***
Parent occupation gaps intercept	0.994	(0.031)	***	0.588	(0.019)	***	0.399	(0.022)	***
Household books gaps intercept	1.338	(0.045)	***	0.618	(0.021)	***	0.721	(0.031)	***
Level 1 - Gaps									
Subject (ref=Reading):									
Math	0.020	(0.007)	**	0.021	(0.004)	***	0.007	(0.004)	*
Science	0.033	(0.006)	***	0.021	(0.003)	***	0.012	(0.003)	***
SES variable quality measures									
Parent-reported × Parent education	0.002	(0.003)		0.002	(0.002)		0.002	(0.002)	
Parent-reported × Parent occupation	0.109	(0.037)	**	0.112	(0.020)	***	0.006	(0.026)	
Parent-reported × Books	0.055	(0.028)	*	0.041	(0.018)	*	0.014	(0.019)	
Number of categories (centered at 7)	-0.021	(0.030)		0.134	(0.018)	***	-0.136	(0.023)	***
≥ 20% in bottom category	-0.062	(0.024)	*	-0.071	(0.018)	***	0.003	(0.022)	
≥ 20% in top category	-0.155	(0.014)	***	-0.059	(0.011)	***	-0.091	(0.011)	***
Level 2 - Study-years									
Age at testing (ref=14)									
Age 10 at testing	-0.152	(0.026)	***	-0.117	(0.017)	***	-0.032	(0.015)	*
Age 15 at testing	-0.014	(0.023)		-0.048	(0.015)	**	0.032	(0.013)	*
Cohort birth year	0.007	(0.001)	***	0.002	(0.001)	**	0.005	(0.001)	***
Random effects									
Level 2 - Residual variance between studies in…									
Parent education intercepts	0.04072			0.01669			0.01475		
Parent occupation intercepts	0.02237			0.00882			0.00599		
Books intercepts	0.04049			0.01809			0.01581		
Level 3 - Residual variance between countries in…									
Parent education intercepts	0.04288			0.01000			0.01774		
Parent occupation intercepts	0.04244			0.01068			0.02182		
Books intercepts	0.10625			0.01935			0.04494		
Cohort slopes	0.00004			0.00001			0.00001		
N (Level 1 - gaps)	4604			4604			4604		
N (Level 2 - study-years)	855			855			855		
N (Level 3 - countries)	78			78			78		

+ p<.1, * p<.05, ** p<.01, *** p < .001.

Table M2. Unstandardized Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th, 50th, and 10th Percentiles of SES

	(8)			(M3)			(M4)		
	90/10 Gap			90/50 Gap			50/10 Gap		
	coef	(se)		coef	(se)		coef	(se)	
Parent education gaps intercept	0.970	(0.046)	***	0.552	(0.029)	***	0.401	(0.032)	***
Parent occupation gaps intercept	0.969	(0.049)	***	0.584	(0.030)	***	0.375	(0.032)	***
Household books gaps intercept	1.405	(0.057)	***	0.645	(0.033)	***	0.753	(0.034)	***
Level 1 - Gaps									
Subject controls (ref=Reading)									
Math	0.020	(0.007)	**	0.021	(0.004)	***	0.007	(0.004)	*
Science	0.033	(0.006)	***	0.022	(0.003)	***	0.012	(0.003)	***
SES variable quality measures									
Parent-reported × Parent education	0.023	(0.039)		0.071	(0.022)	**	-0.049	(0.028)	+
Parent-reported × Parent occupation	-0.026	(0.032)		0.002	(0.021)		-0.038	(0.022)	+
Parent-reported × Books	-0.128	(0.039)	**	0.084	(0.025)	***	-0.212	(0.026)	***
Number of categories (centered at 7)	0.000	(0.003)		0.001	(0.002)		0.001	(0.002)	
≥ 20% in bottom category	-0.055	(0.023)	*	-0.047	(0.019)	*	0.001	(0.021)	
≥ 20% in top category	-0.149	(0.015)	***	-0.061	(0.012)	***	-0.080	(0.011)	***
Level 2 - Study-years									
Study fixed effects (ref=TIMSS 2003 Grade 8)	yes			yes			yes		
School enrollment (proportion)	0.486	(0.107)	***	0.276	(0.067)	***	0.203	(0.068)	**
Immigrant background (proportion)	0.226	(0.250)		0.130	(0.145)		0.090	(0.134)	
GDP per capita (logged)	0.055	(0.059)		0.024	(0.049)		0.039	(0.039)	
Income inequality (Gini)	-1.913	(0.887)	*	-1.380	(0.476)	**	-0.554	(0.489)	
Mid/low-income country × Income inequality	2.539	(1.129)	*	1.687	(0.628)	**	0.753	(0.650)	
Age when tracking begins	-0.037	(0.016)	*	-0.023	(0.011)	*	-0.015	(0.008)	+
Private school enrollment (proportion)	0.240	(0.249)		0.147	(0.142)		0.070	(0.159)	
Expecting higher education (proportion)	-0.029	(0.094)		-0.043	(0.062)		0.017	(0.046)	
Level 3 - Countries									
Mid/low-income country × Intercept interactions									
Mid/low-income country × Parent education	0.192	(0.063)	**	0.071	(0.033)	*	0.138	(0.037)	***
Mid/low-income country × Parent occupation	0.060	(0.058)		0.056	(0.035)		0.017	(0.033)	
Mid/low-income country × Books	-0.082	(0.068)		-0.016	(0.040)		-0.059	(0.043)	
Mean school enrollment	0.640	(0.317)	*	0.139	(0.179)		0.531	(0.162)	**
Mean proportion immigrant background	0.134	(0.261)		-0.057	(0.146)		0.172	(0.140)	
Mean GDP per capita (logged)	0.142	(0.041)	***	0.078	(0.024)	**	0.070	(0.022)	**
Mean income inequality	0.605	(0.317)	+	0.594	(0.169)	***	-0.008	(0.174)	
Mean age when tracking begins	-0.035	(0.011)	**	-0.020	(0.005)	***	-0.018	(0.007)	*
Mean private school enrollment	0.043	(0.099)		0.006	(0.058)		0.029	(0.057)	
Mean proportion expecting higher education	-0.255	(0.169)		-0.113	(0.087)		-0.154	(0.101)	
Random effects									
Level 2 - Residual variance between studies in…									
Parent education gaps	0.038			0.017			0.015		
Parent occupation gaps	0.021			0.008			0.006		
Household books gaps	0.037			0.015			0.014		
Level 3 - Residual variance between countries in…									
Parent education gaps	0.028			0.007			0.010		
Parent occupation gaps	0.027			0.009			0.010		
Household books gaps	0.049			0.014			0.016		
N (Level 1 - gaps)	4604			4604			4604		
N (Level 2 - study-years)	855			855			855		
N (Level 3 - countries)	78			78			78		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Notes: "Middle/low income" countries had GDPs per capita of less than $6000 in 1980 (the reference category is high-income countries; see Appendix Table A1 for coding). All level 2 time-varying country covariates are mean-centered within countries, meaning results can be interpreted very similarly to a model with country fixed effects (as well as study-year fixed effects, included at level 2).

Tables M3 and M4 report intercorrelations among all country covariates, at the country-study-year level (n=855; Table M3) and at the country level (n=78; Table M4). None of the correlations is strong enough to cause concern about collinearity in the multivariate models.

Table M3. Correlations Among Time-Varying Country Covariates (n=855)

	School enrollment (proportion)	Immigrant background (proportion)	GDP per capita (logged)	Income inequality (Gini)	Age when tracking begins	Private school enrollment (proportion)	Expecting higher education (proportion)
School enrollment (proportion)	1						
Immigrant background (proportion)	0.258	1					
GDP per capita (logged)	0.587	0.441	1				
Income inequality (Gini)	-0.443	-0.281	-0.466	1			
Age when tracking begins	0.077	0.067	0.028	0.169	1		
Private school enrollment (proportion)	0.022	0.165	0.135	-0.052	-0.214	1	
Expecting higher education (proportion)	-0.063	-0.060	-0.086	0.432	0.307	0.017	1

Table M4. Correlations Among Country-Level Mean Variables (N=78)

	Mean school enrollment	Mean proportion immigrant background	Mean GDP per capita (logged)	Mean income inequality	Mean age when tracking begins	Mean private school enrollment	Mean proportion expecting higher education
Mean school enrollment	1.000						
Mean proportion immigrant background	0.320	1					
Mean GDP per capita (logged)	0.647	0.475	1				
Mean income inequality	-0.520	-0.388	-0.457	1			
Mean age when tracking begins	0.089	0.045	-0.015	0.082	1		
Mean private school enrollment	-0.009	0.089	0.151	-0.025	-0.262	1	
Mean proportion expecting higher education	-0.288	-0.251	-0.398	0.438	0.265	0.074	1

Table M5 further examines possible collinearity among time-varying country covariates by reporting seven models, each predicting 90/10 SES achievement gaps from only one time-varying country covariate at a time. The coefficients for all time-varying country covariates are very similar when each is entered individually to in models where all seven covariates are entered together. Thus, collinearity among country covariates does not appear to influence the results.

In the main text of the paper, I examined whether certain types of countries tend to have greater increases in SES achievement gaps by interacting two country-level variables (region and income level) with the cohort birth year slope. Table M6 further explores whether cohort birth year slopes systematically vary by other available country-level variables. As the country-level sample size with complete data is only 78, I run each interaction model separately. The coefficients of interest are the ‘Country variable × Cohort interactions’ near the bottom of the table. All coefficients are close to 0 and non-significant.

Table M5. Unstandardized Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th Percentiles of SES, Entering Covariates One at a Time

	Enrollment			Immigration			GDP per capita			Income inequality		
	coef	(se)		coef	(se)		coef	(se)		coef	(se)	
Parent education gaps intercept	1.065	(0.043)	***	1.082	(0.040)	***	0.969	(0.045)	***	1.016	(0.044)	***
Parent occupation gaps intercept	1.066	(0.047)	***	1.083	(0.046)	***	0.970	(0.049)	***	1.015	(0.046)	***
Household books gaps intercept	1.498	(0.057)	***	1.514	(0.055)	***	1.401	(0.055)	***	1.449	(0.056)	***
Level 1 - Gaps												
Subject controls (ref=Reading)												
Math	0.020	(0.007)	**	0.020	(0.007)	**	0.020	(0.007)	**	0.020	(0.007)	**
Science	0.033	(0.006)	***	0.033	(0.006)	***	0.033	(0.006)	***	0.033	(0.006)	***
SES variable quality measures												
Parent-reported × Parent education	0.027	(0.039)		0.027	(0.039)		0.025	(0.039)		0.020	(0.039)	
Parent-reported × Parent occupation	-0.022	(0.033)		-0.024	(0.032)		-0.025	(0.032)		-0.024	(0.032)	
Parent-reported × Books	-0.123	(0.039)	**	-0.124	(0.040)	**	-0.127	(0.039)	**	-0.136	(0.040)	***
Number of categories (centered at 7)	0.000	(0.003)		0.000	(0.003)		0.000	(0.003)		-0.002	(0.003)	
≥ 20% in bottom category	-0.059	(0.022)	**	-0.055	(0.022)	*	-0.056	(0.022)	*	-0.057	(0.022)	**
≥ 20% in top category	-0.154	(0.015)	***	-0.153	(0.015)	***	-0.152	(0.015)	***	-0.146	(0.015)	***
Level 2 - Study-years												
Study fixed effects (ref=TIMSS 2003 Grade 8)	yes			yes			yes			yes		
School enrollment (proportion)	0.505	(0.110)	***									
Immigrant background (proportion)				0.313	(0.301)							
GDP per capita (logged)							0.076	(0.068)				
Income inequality (Gini)										-1.623	(0.838)	+
Mid/low-income country × Income inequality										2.072	(1.078)	+
Level 3 - Countries												
Mid/low-income country × Intercept interactions												
Mid/low-income country × Parent education	0.021	(0.061)		-0.038	(0.050)		0.168	(0.062)	**	-0.070	(0.063)	
Mid/low-income country × Parent occupation	-0.110	(0.056)	*	-0.161	(0.051)	**	0.039	(0.057)		-0.195	(0.059)	***
Mid/low-income country × Books	-0.254	(0.072)	***	-0.314	(0.069)	***	-0.102	(0.066)		-0.347	(0.078)	***
Mean school enrollment	0.583	(0.291)	*									
Mean proportion immigrant background				0.358	(0.279)							
Mean GDP per capita (logged)							0.178	(0.036)	***			
Mean income inequality										0.140	(0.305)	
Random effects												
Level 2 - Residual variance between studies in...												
Parent education gaps	0.03892			0.03980			0.03957			0.04075		
Parent occupation gaps	0.02241			0.02219			0.02206			0.02300		
Household books gaps	0.03788			0.03949			0.04030			0.04011		
Level 3 - Residual variance between countries in...												
Parent education gaps	0.03905			0.04544			0.03646			0.04400		
Parent occupation gaps	0.03498			0.03694			0.02962			0.03736		
Household books gaps	0.07527			0.08838			0.06152			0.08712		
N (Level 1 - gaps)	4604			4604			4604			4604		
N (Level 2 - study-years)	855			855			855			855		
N (Level 3 - countries)	78			78			78			78		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Notes: “Middle/low income” countries had GDPs per capita of less than $6000 in 1980 (the reference category is high-income countries; see Appendix Table A1 for coding). All level 2 time-varying country covariates are mean-centered within countries, meaning results can be interpreted very similarly to a model with country fixed effects (as well as study-year fixed effects, included at level 2). All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper.

Model M5 (cont.) Unstandardized Coefficients from Hierarchical Growth Models Predicting Achievement Gaps between 90th and 10th Percentiles of SES, Entering Covariates One at a Time

	Tracking			Private			Higher ed. Expect.		
	coef	(se)		coef	(se)		coef	(se)	
Parent education gaps intercept	1.098	(0.039)	***	1.094	(0.040)	***	1.083	(0.043)	***
Parent occupation gaps intercept	1.098	(0.042)	***	1.095	(0.043)	***	1.084	(0.044)	***
Household books gaps intercept	1.531	(0.052)	***	1.527	(0.054)	***	1.515	(0.052)	***
Level 1 - Gaps									
Subject controls (ref=Reading)									
Math	0.020	(0.007)	**	0.020	(0.007)	**	0.020	(0.007)	**
Science	0.033	(0.006)	***	0.033	(0.006)	***	0.033	(0.006)	***
SES variable quality measures									
Parent-reported × Parent education	0.026	(0.039)		0.025	(0.038)		0.027	(0.038)	
Parent-reported × Parent occupation	-0.025	(0.032)		-0.026	(0.032)		-0.025	(0.032)	
Parent-reported × Books	-0.126	(0.040)	**	-0.126	(0.039)	**	-0.125	(0.039)	**
Number of categories (centered at 7)	0.000	(0.003)		0.000	(0.003)		0.000	(0.003)	
≥ 20% in bottom category	-0.052	(0.022)	*	-0.054	(0.022)	*	-0.056	(0.022)	*
≥ 20% in top category	-0.150	(0.015)	***	-0.152	(0.015)	***	-0.151	(0.015)	***
Level 2 - Study-years									
Study fixed effects (ref=TIMSS 2003 Grade 8)	yes			yes			yes		
Age when tracking begins	-0.032	(0.015)	*						
Private school enrollment (proportion)				0.206	(0.295)				
Expecting higher education (proportion)							-0.054	(0.101)	
Level 3 - Countries									
Mid/low-income country × Intercept interactions									
Mid/low-income country × Parent education	-0.068	(0.045)		-0.057	(0.046)		-0.037	(0.054)	
Mid/low-income country × Parent occupation	-0.194	(0.045)	***	-0.181	(0.045)	***	-0.160	(0.049)	***
Mid/low-income country × Books	-0.346	(0.063)	***	-0.332	(0.066)	***	-0.311	(0.067)	***
Mean age when tracking begins	-0.028	(0.013)	*						
Mean private school enrollment				0.213	(0.108)	*			
Mean proportion expecting higher education							-0.311	(0.176)	+
Random effects									
Level 2 - Residual variance between studies in...									
Parent education gaps	0.03950			0.03947			0.03939		
Parent occupation gaps	0.02165			0.02223			0.02209		
Household books gaps	0.03962			0.04009			0.04010		
Level 3 - Residual variance between countries in...									
Parent education gaps	0.04093			0.04483			0.04541		
Parent occupation gaps	0.03644			0.03729			0.03597		
Household books gaps	0.08150			0.08966			0.08300		
N (Level 1 - gaps)	4604			4604			4604		
N (Level 2 - study-years)	855			855			855		
N (Level 3 - countries)	78			78			78		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Notes: “Middle/low income” countries had GDPs per capita of less than $6000 in 1980 (the reference category is high-income countries; see Appendix Table A1 for coding). All level 2 time-varying country covariates are mean-centered within countries, meaning results can be interpreted very similarly to a model with country fixed effects (as well as study-year fixed effects, included at level 2). All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper.

Table M6. Unstandardized Coefficients from Hierarchical Growth Curve Models Predicting Achievement Gaps between 90th and 10th Percentiles, Adding Interactions by Country Variables

	Enrollment			Immigration			GDP per capita			Income inequality		
	coef	(se)		coef	(se)		coef	(se)		coef	(se)	
Parent education gaps intercept	1.064	(0.031)	***	1.067	(0.033)	***	1.063	(0.031)	***	1.066	(0.033)	***
Parent occupation gaps intercept	0.993	(0.029)	***	0.995	(0.030)	***	0.988	(0.028)	***	0.994	(0.031)	***
Household books gaps intercept	1.331	(0.039)	***	1.338	(0.044)	***	1.322	(0.036)	***	1.335	(0.043)	***
Level 1 - Gaps												
Subject controls (ref=Reading)												
Math	0.020	(0.007)	**	0.020	(0.007)	**	0.020	(0.007)	**	0.020	(0.007)	**
Science	0.033	(0.006)	***	0.033	(0.006)	***	0.033	(0.006)	***	0.033	(0.006)	***
SES variable quality measures												
Parent-reported × Parent education	0.109	(0.038)	**	0.108	(0.037)	**	0.107	(0.037)	**	0.109	(0.037)	**
Parent-reported × Parent occupation	0.055	(0.027)	*	0.055	(0.028)	*	0.053	(0.027)	+	0.055	(0.028)	*
Parent-reported × Books	-0.024	(0.030)		-0.020	(0.030)		-0.024	(0.030)		-0.021	(0.030)	
Number of categories (centered at 7)	0.003	(0.003)		0.002	(0.003)		0.002	(0.003)		0.002	(0.003)	
≥ 20% in bottom category	-0.049	(0.024)	*	-0.061	(0.025)	*	-0.047	(0.024)	+	-0.055	(0.024)	*
≥ 20% in top category	-0.155	(0.014)	***	-0.157	(0.013)	***	-0.156	(0.014)	***	-0.156	(0.014)	***
Level 2 - Study-years												
Age at testing (ref=14)												
Age 10 at testing	-0.153	(0.026)	***	-0.152	(0.026)	***	-0.151	(0.026)	***	-0.152	(0.026)	***
Age 15 at testing	-0.014	(0.023)		-0.013	(0.023)		-0.015	(0.023)		-0.014	(0.023)	
Cohort birth year	0.007	(0.001)	***	0.007	(0.001)	***	0.007	(0.001)	***	0.007	(0.001)	***
Level 3 - Countries												
Country variable × Intercept interactions												
Mean school enrollment × Parent education	0.597	(0.194)	**									
Mean school enrollment × Parent occupation	0.810	(0.252)	**									
Mean school enrollment × Books	1.635	(0.287)	***									
Mean proportion immigrant background × Parent education				0.358	(0.292)							
Mean proportion immigrant background × Parent occupation				0.934	(0.248)	***						
Mean proportion immigrant background × Books				1.288	(0.395)	**						
Mean GDP per capita (logged) × Parent education							0.091	(0.026)	***			
Mean GDP per capita (logged) × Parent occupation							0.152	(0.027)	***			
Mean GDP per capita (logged) × Books							0.291	(0.034)	***			
Mean income inequality × Parent education										-0.020	(0.241)	
Mean income inequality × Parent occupation										-0.510	(0.221)	*
Mean income inequality × Books										-1.326	(0.318)	***
Country variable × Cohort interactions												
Mean school enrollment × Cohort	0.003	(0.009)										
Mean proportion immigrant background × Cohort				0.013	(0.015)							
Mean GDP per capita (logged) × Cohort							0.001	(0.001)				
Mean income inequality × Cohort										-0.010	(0.012)	
Random Effects												
Level 2 - Residual variance between studies in…												
Parent education intercepts	0.04087			0.04074			0.04074			0.04071		
Parent occupation intercepts	0.02232			0.02233			0.02223			0.02237		
Books intercepts	0.04054			0.04042			0.04049			0.04051		
Level 3 - Residual variance between countries in…												
Parent education intercepts	0.03693			0.04246			0.0372			0.04387		
Parent occupation intercepts	0.03238			0.0371			0.02761			0.04022		
Books intercepts	0.06803			0.09794			0.05272			0.09185		
Cohort slopes	0.00004			0.00004			0.00003			0.00004		
N (Level 1 - gaps)	4604			4604			4604			4604		
N (Level 2 - study-years)	855			855			855			855		
N (Level 3 - countries)	78			78			78			78		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper.

Table M6 (cont). Unstandardized Coefficients from Hierarchical Growth Curve Models Predicting Achievement Gaps between 90th and 10th Percentiles, Adding Interactions by Country Variables

	Tracking			Private			Higher ed. Expect.		
	coef	(se)		coef	(se)		coef	(se)	
Parent education gaps intercept	1.068	(0.032)	***	1.065	(0.033)	***	1.065	(0.033)	***
Parent occupation gaps intercept	0.995	(0.031)	***	0.993	(0.031)	***	0.993	(0.031)	***
Household books gaps intercept	1.341	(0.044)	***	1.338	(0.045)	***	1.334	(0.044)	***
Level 1 - Gaps									
Subject controls (ref=Reading)									
Math	0.020	(0.007)	**	0.020	(0.007)	**	0.020	(0.007)	**
Science	0.033	(0.006)	***	0.033	(0.006)	***	0.033	(0.006)	***
SES variable quality measures									
Parent-reported × Parent education	0.108	(0.037)	**	0.109	(0.037)	**	0.108	(0.037)	**
Parent-reported × Parent occupation	0.055	(0.028)	*	0.055	(0.027)	*	0.054	(0.028)	*
Parent-reported × Books	-0.021	(0.030)		-0.021	(0.030)		-0.022	(0.030)	
Number of categories (centered at 7)	0.002	(0.003)		0.002	(0.003)		0.002	(0.003)	
≥ 20% in bottom category	-0.064	(0.025)	**	-0.063	(0.024)	*	-0.058	(0.025)	*
≥ 20% in top category	-0.156	(0.014)	***	-0.155	(0.014)	***	-0.155	(0.014)	***
Level 2 - Study-years									
Age at testing (ref=14)									
Age 10 at testing	-0.152	(0.026)	***	-0.153	(0.026)	***	-0.152	(0.026)	***
Age 15 at testing	-0.014	(0.023)		-0.014	(0.023)		-0.014	(0.023)	
Cohort birth year	0.007	(0.001)	***	0.007	(0.001)	***	0.007	(0.001)	***
Level 3 - Countries									
Country variable × Intercept interactions									
Mean age when tracking begins × Parent education	-0.044	(0.013)	***						
Mean age when tracking begins × Parent occupation	-0.027	(0.014)	*						
Mean age when tracking begins × Books	-0.058	(0.020)	**						
Mean private school enrollment × Parent education				0.232	(0.121)	+			
Mean private school enrollment × Parent occupation				0.278	(0.122)	*			
Mean private school enrollment × Books				0.165	(0.157)				
Mean proportion expecting higher education × Parent education							-0.269	(0.192)	
Mean proportion expecting higher education × Parent occupation							-0.615	(0.152)	***
Mean proportion expecting higher education × Books							-1.167	(0.279)	***
Mean age when tracking begins × Cohort	-0.001	(0.001)							
Mean private school enrollment × Cohort				-0.003	(0.007)				
Mean proportion expecting higher education × Cohort							-0.011	(0.007)	+
Random Effects									
Level 2 - Residual variance between studies in…									
Parent education intercepts	0.04075			0.0407			0.04072		
Parent occupation intercepts	0.0224			0.02249			0.02227		
Books intercepts	0.04058			0.04035			0.04044		
Level 3 - Residual variance between countries in…									
Parent education intercepts	0.03695			0.04189			0.04211		
Parent occupation intercepts	0.04095			0.0403			0.0366		
Books intercepts	0.0959			0.10571			0.08689		
Cohort slopes	0.00003			0.00004			0.00003		
N (Level 1 - gaps)	4604			4604			4604		
N (Level 2 - study-years)	855			855			855		
N (Level 3 - countries)	78			78			78		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper.

The country covariate models in the main text of the paper enter a series of study-year dummies in order to account for secular trends across study-years. An alternative way to account for secular trends in SES achievement gaps would be to include the cohort birth year variable in the model rather than study-year fixed effects. Table M7 reports such a model (Model M6), along with a model omitting country covariates but using the same analytic sample of 78 countries with complete covariate data (Model M5), and the original version of Model 8 with study-year fixed effects from the main paper text for comparison. Model M5 shows that the cohort birth year trend is identical in the reduced analytic sample of countries with full covariate data to in the full sample of 100 countries. Comparing across Models M6 and 8, coefficents for time-varying country covariates are broadly similar. Some coefficients change direction, but only those that are not significantly different from 0 in either model. The strongest predictor of increasing SES achievement gaps remains increasing school enrollment. Model M6 including both country covariates and cohort birth year also allows us to examine a different question: to what extent do the time-varying country covariates explain the average global increase in SES achievement gaps? Comparing across Model M5 and Model M6, the cohort birth year coefficient decreases from 0.007 to 0.001, an 86% reduction. Thus, the included country covariates do explain a great deal of the average global trend in SES achievement gaps. Model 8 is preferred over Model M6 because in addition to controlling for secular time trends in SES achievement gaps, it also accounts for possible differences across studies in the measurement error of SES achievement gaps (either due to unreliability or low quality of achievement or SES measures, as discussed in Appendices D and H). If we assume that all countries are likely to suffer from similar data reliability or quality issues in the same studies—a reasonable assumption because SES survey item wording is very similar across countries in a given year of a study—then these dummies control for study-specific biases in the estimation of SES achievement gaps. Thus, the coefficient estimates for covariates in Model 8 are less likely to be confounded by uneven quality of the outcome variable than those in Model M6.

Table M7. Comparison between Multivariate Models Predicting 90/10 Achievement Gaps from Country Covariates and Cohort Birth Year or Study-Year Dummies

	(M5) Analytic sample			(M6) Cohort year			(8) Study-year FE		
	coef	(se)		coef	(se)		coef	(se)	
Parent education gaps intercept	1.099	(0.030)	***	0.978	(0.039)	***	0.970	(0.046)	***
Parent occupation gaps intercept	1.097	(0.031)	***	0.975	(0.042)	***	0.969	(0.049)	***
Household books gaps intercept	1.522	(0.040)	***	1.391	(0.047)	***	1.405	(0.057)	***
Level 1 - Gaps									
Subject controls (ref=Reading)									
Math	0.020	(0.007)	**	0.020	(0.007)	**	0.020	(0.007)	**
Science	0.033	(0.006)	***	0.033	(0.006)	***	0.033	(0.006)	***
SES variable quality measures									
Parent-reported × Parent education	0.108	(0.037)	**	0.112	(0.037)	**	0.023	(0.039)	
Parent-reported × Parent occupation	0.054	(0.028)	*	0.061	(0.027)	*	-0.026	(0.032)	
Parent-reported × Books	-0.022	(0.030)		-0.017	(0.029)		-0.128	(0.039)	**
Number of categories (centered at 7)	0.002	(0.003)		0.002	(0.003)		0.000	(0.003)	
≥ 20% in bottom category	-0.055	(0.024)	*	-0.057	(0.025)	*	-0.055	(0.023)	*
≥ 20% in top category	-0.157	(0.014)	***	-0.156	(0.014)	***	-0.149	(0.015)	***
Level 2 - Study-years									
Age at testing (ref=14)									
Age 10 at testing	-0.153	(0.026)	***	-0.203	(0.031)	***			
Age 15 at testing	-0.014	(0.023)		-0.029	(0.022)				
Cohort birth year	0.007	(0.001)	***	0.001	(0.003)				
Study fixed effects (ref=TIMSS 2003 Grade 8)							yes		
School enrollment (proportion)				0.539	(0.145)	***	0.486	(0.107)	***
Immigrant background (proportion)				-0.215	(0.283)		0.226	(0.250)	
GDP per capita (logged)				0.127	(0.061)	*	0.055	(0.059)	
Income inequality (Gini)				-0.723	(0.997)		-1.913	(0.887)	*
Mid/low-income country × Income inequality				1.908	(1.138)	+	2.539	(1.129)	*
Age when tracking begins				-0.027	(0.016)	+	-0.037	(0.016)	*
Private school enrollment (proportion)				-0.042	(0.255)		0.240	(0.249)	
Expecting higher education (proportion)				0.050	(0.093)		-0.029	(0.094)	
Level 3 - Countries									
Mid/low-income country × Intercept interactions									
Mid/low-income country × Parent education	-0.059	(0.048)		0.206	(0.059)	***	0.192	(0.063)	**
Mid/low-income country × Parent occupation	-0.185	(0.045)	***	0.071	(0.057)		0.060	(0.058)	
Mid/low-income country × Books	-0.331	(0.063)	***	-0.044	(0.066)		-0.082	(0.068)	
Mean school enrollment				0.652	(0.282)	*	0.640	(0.317)	*
Mean proportion immigrant background				0.068	(0.239)		0.134	(0.261)	
Mean GDP per capita (logged)				0.143	(0.039)	***	0.142	(0.041)	***
Mean income inequality				0.653	(0.303)	*	0.605	(0.317)	+
Mean age when tracking begins				-0.040	(0.010)	***	-0.035	(0.011)	**
Mean private school enrollment				0.004	(0.089)		0.043	(0.099)	
Mean proportion expecting higher education				-0.155	(0.165)		-0.255	(0.169)	
Random effects									
Level 2 - Residual variance between studies in…									
Parent education gaps	0.04067			0.03993			0.03824		
Parent occupation gaps	0.02244			0.02222			0.02065		
Household books gaps	0.04051			0.03738			0.03683		
Level 3 - Residual variance between countries in…									
Parent education gaps	0.04228			0.02359			0.02822		
Parent occupation gaps	0.03519			0.02440			0.02732		
Household books gaps	0.08273			0.04353			0.04944		
Cohort slopes	0.00003			0.00004					
N (Level 1 - gaps)	4604			4604			4604		
N (Level 2 - study-years)	855			855			855		
N (Level 3 - countries)	78			78			78		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Notes: "Middle/low income" countries had GDPs per capita of less than $6000 in 1980 (the reference category is high-income countries; see Appendix Table A1 for coding). All level 2 time-varying country covariates are mean-centered within countries, meaning results can be interpreted very similarly to a model with country fixed effects (as well as study-year fixed effects, included at level 2). All models in this table specify known level 1 error variances and covariances, estimated via bootstrapping, consistent with models in the main text of the paper.

Finally, Table M8 reports multivariate results estimated using a fixed effects rather than a mixed effects (hierarchical growth curve) model. The model includes country fixed effects, meaning that country-level covariates are omitted due to collinearity. The model is estimated as follows:

$$\hat{G}_{ij} = \gamma_{00} + \gamma_{10}Y_{ij} + (\boldsymbol{X}_{ij})\mathbf{B} + \boldsymbol{\Gamma_j} + \boldsymbol{\Lambda}_{ij}, \quad \text{[M1]}$$

where $\hat{G}_{ij}$ is the estimated gap in country j in country-study-year i, γ_{10} is the coefficient for cohort birth year Y_{ij}, $\boldsymbol{X}_{ij}$ is a vector of time-varying country covariates in country-year i, $\boldsymbol{\Gamma}$ is a vector of country dummy variables, $\boldsymbol{\Lambda}_{ij}$ is a vector of dummy variables indicating age at testing and test subject, $\boldsymbol{B}$ is a vector of coefficients for the time-varying country covariates. The country fixed effects are estimated using weighted least squares (weighted by the inverse squared standard error associated with each gap estimate, estimated via bootstrapping) and are reported with robust Huber-White standard errors.

Results for the country fixed effects models are very similar to those in the main results for Models 2, 2B, and 8 of the main text, both in terms of the cohort trend estimates in Model M7 (and for the reduced analytic sample in Model M8) and the covariate results in Model M9. Recall that time-varying country covariates in the hierarchical growth curve models are mean-centered within countries, meaning that they have an interpretation very similar to a fixed effects model.

Table M8. Country Fixed Effects Models Predicting 90/10 Achievement Gaps from Country Covariates and Cohort Birth Year

	(M7)			(M8)			(M9)		
	Country FE - full sample			Country FE - analytic sample			Country & study FE, covariates		
	coef	(se)		coef	(se)		coef	(se)	
SES variable (ref=Parent education)									
Parent occupation	-0.083	(0.011)	***	-0.063	(0.012)	***	-0.009	(0.012)	
Household books	0.391	(0.012)	***	0.399	(0.013)	***	0.492	(0.014)	***
Subject (ref=Reading)									
Math	0.012	(0.008)		0.011	(0.008)		0.019	(0.008)	*
Science	0.025	(0.008)	***	0.023	(0.008)	**	0.031	(0.008)	***
SES variable quality measures									
Parent-reported × Parent education	0.083	(0.017)	***	0.071	(0.018)	***	-0.018	(0.019)	
Parent-reported × Parent occupation	-0.080	(0.018)	***	-0.080	(0.019)	***	-0.055	(0.018)	**
Parent-reported × Books	-0.206	(0.022)	***	-0.156	(0.022)	***	-0.187	(0.022)	***
Number of categories (centered at 7)	0.006	(0.004)		0.004	(0.004)		0.002	(0.004)	
≥ 20% in bottom category	-0.220	(0.014)	***	-0.213	(0.015)	***	-0.152	(0.015)	***
≥ 20% in top category	-0.128	(0.008)	***	-0.144	(0.009)	***	-0.124	(0.009)	***
Age at testing (ref=14)									
Age 10 at testing	-0.136	(0.017)	***	-0.145	(0.018)	***			
Age 15 at testing	-0.058	(0.012)	***	-0.058	(0.013)	***			
Cohort birth year	0.007	(0.001)	***	0.007	(0.001)	***			
Study fixed effects (ref=TIMSS 2003 Grade 8)							yes		
School enrollment (proportion)							0.397	(0.077)	***
Immigrant background (proportion)							0.362	(0.143)	*
GDP per capita (logged)							-0.051	(0.049)	
Income inequality (Gini)							-2.176	(0.449)	***
Mid/low-income country × Income inequality							3.429	(0.629)	***
Age when tracking begins							-0.035	(0.009)	***
Private school enrollment (proportion)							0.025	(0.144)	
Expecting higher education (proportion)							-0.095	(0.045)	*
Mid/low-income country × SES interactions									
Mid/low-income country × Parent occupation							-0.113	(0.015)	***
Mid/low-income country × Books							-0.212	(0.017)	***
Country fixed effects	yes			yes			yes		
Intercept	1.152	(0.013)	***	1.178	(0.014)	***	1.979	(0.451)	***
Adjusted R^2	0.677			0.675			0.711		
N (gaps)	5541			4604			4604		
N (study-years)	1026			855			855		
N (countries)	100			78			78		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: All models in this table are estimated via weighted least squares (weighted by the inverse squared standard error associated with each gap estimate, estimated via bootstrapping).

Appendix N. Trends in SES achievement gaps using a rank-based measure of achievement

The analyses in the main paper text standardize all achievement scores to a mean of 0 and standard deviation of 1 within each country-study-year-subject before calculating SES achievement gaps. Thus, trends in SES achievement gaps cannot be influenced by changes in the dispersion of achievement, as the dispersion is constrained to be constant over time. However, trends could still be influenced by changes in the *shape* of the achievement distribution. An examination of achievement distributions shows that, with very few exceptions, the distribution of achievement scores is close to normal in every country-study-year-subject. As an additional robustness check that trends in SES achievement gaps are not influenced by changes in the shape of the achievement distribution, I also estimate trends in gaps after converting achievement to a rank-based measure, which has a constant uniform distribution in every country-study-year-subject. Table N1 reports results of models estimating trends in SES achievement gaps computed from achievement rank. Within each country-study-year-subject, rather than standardizing achievement to a mean of 0 and standard deviation of 1, I instead convert achievement into percentiles scaled from 0 to 1. Using this transformed achievement measure, I then compute 90/10 gaps using the same method described in the main text of the paper.

The results for trends in SES gaps in achievement rank are very similar to those for trends in SES gaps in standardized achievement. The magnitude of all coefficients is smaller because gaps based on achievement rank are smaller than gaps based on standardized achievement. (A gap of 1.0 would indicate a gap of 100 percentile points of achievement between the 90^{th} and 10^{th} percentiles of SES, compared with a gap of 1.0 standard deviations of achievement between the 90^{th} and 10^{th} percentiles of SES.) Yet the overall story is very similar. 90/10 gaps based on all three SES measures have increased significantly ($p < .001$).

The trends in gaps in achievement rank also partially relax the assumption that achievement must be interval scaled. However, ranks are interval scaled only when interpreted *as ranks*; they are not necessarily interval scaled with respect to students' true achievement—the assumption that must hold in order for standardized achievement to be interval scaled in z-scores. Although it is not possible to test the interval scaling assumption directly, the consistency of results across the gaps using standardized achievement (in the main text), rank-based achievement in Table N1, and the original scales of PISA, TIMSS, and PIRLS (Appendix B) provides some indication that results are robust to the achievement scale used.

Table N1. Hierarchical Growth Models Predicting 90/10 SES Achievement Gaps Using a Rank-Based Measure of Achievement

	(1)			(2)		
	3 Cohort Slopes			1 Cohort Slope		
	coef	(se)		coef	(se)	
Parent education gaps intercept	0.300	(0.009)	***	0.303	(0.009)	***
Parent occupation gaps intercept	0.272	(0.009)	***	0.275	(0.009)	***
Household books gaps intercept	0.373	(0.012)	***	0.373	(0.012)	***
Level 1 - Gaps						
Subject (ref=Reading):						
Math	-0.001	(0.002)		-0.001	(0.002)	
Science	0.006	(0.002)	***	0.006	(0.002)	***
SES variable quality measures						
Parent-reported × Parent education	0.025	(0.007)	***	0.019	(0.008)	*
Parent-reported × Parent occupation	0.016	(0.006)	**	0.015	(0.006)	*
Parent-reported × Books	-0.029	(0.008)	***	-0.020	(0.007)	**
Number of categories (centered at 7)	0.001	(0.001)		0.001	(0.001)	
≥ 20% in bottom category	-0.014	(0.006)	*	-0.015	(0.006)	*
≥ 20% in top category	-0.029	(0.003)	***	-0.033	(0.003)	***
Level 2 - Study-years						
Age at testing (ref=14)						
Age 10 at testing	-0.042	(0.007)	***	-0.042	(0.007)	***
Age 15 at testing	-0.007	(0.006)		-0.007	(0.006)	
Cohort birth year × Parent education	0.0019	(0.0003)	***			
Cohort birth year × Parent occupation	0.0021	(0.0003)	***			
Cohort birth year × Books	0.0024	(0.0004)	***			
Cohort birth year				0.0022	(0.0003)	***
Random effects						
Level 2 - Residual variance between studies in…						
Parent education intercepts	0.002230			0.002270		
Parent occupation intercepts	0.001560			0.001580		
Books intercepts	0.002650			0.002750		
Level 3 - Residual variance between countries in…						
Parent education intercepts	0.004780			0.004730		
Parent occupation intercepts	0.004630			0.004750		
Books intercepts	0.010300			0.010660		
Parent education cohort slopes	0.000004					
Parent occupation cohort slopes	0.000003					
Books cohort slopes	0.000005					
Cohort slopes				0.000003		
N (Level 1 - gaps)	5541			5541		
N (Level 2 - study-years)	1027			1027		
N (Level 3 - countries)	100			100		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Note: To avoid very long computation times, all models in this table specify known level 1 error variances estimated using conventional non-bootstrap formulas and omit error covariances. This simplified specification appears to produce very similar results to models using bootstrapped error variances and covariances (see Appendix L).

References

Broer, Markus, Yifan Bai, and Frank Fonseca. forthcoming. *Socioeconomic Inequality and Educational Outcomes: Evidence from Twenty Years of TIMSS*: IEA Research for Education, Vol. 5.

Jerrim, John, and John Micklewright. 2014. "Socio-economic Gradients in Children’s Cognitive Skills: Are Cross-Country Comparisons Robust to Who Reports Family Background?" *European Sociological Review*:jcu072.

Kreidl, Martin, Harry BG Ganzeboom, and Donald J Treiman. 2014. "How did occupational returns to education change over time?" in *California Center for Population Research Online Working Paper Series PWP-CCPR-2014-017*: University of California - Los Angeles.

Mullis, Ina V. S., Michael O. Martin, Pierre Foy, and Kathleen T. Drucker. 2012. "PIRLS 2011 International Results in Reading." Boston, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, Ina V. S., Michael O. Martin, Pierre Foy, and Martin Hooper. 2016. "TIMSS 2015 International Results in Mathematics." Boston, MA: TIMSS & PIRLS International Study Center, Boston College.

OECD. 2016. "PISA 2015 Results: Excellence and Equity in Education (Volume I)." Paris: OECD Publishing.

Reardon, Sean F. 2011a. "Appendices for 'The Widening Academic Achievement Gap between the Rich and the Poor: New Evidence and Possible Explanations'." Pp. 1-49 in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, edited by Greg J. Duncan and Richard J. Murnane. New York: Russell Sage Foundation.

—. 2011b. "The Widening Academic-Achievement Gap between the Rich and the Poor: New Evidence and Possible Explanations." Pp. 91-115 in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, edited by Greg J. Duncan and Richard J. Murnane. New York: Russell Sage Foundation.