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

AUTHORS

Anna K. Chmielewski

University of Toronto

ABSTRACT

The existence of a "socioeconomic achievement gap"—a 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. What is unknown in most countries is whether the SES achievement gap might be 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 have 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 have also increased in many countries with consistently high enrollments, suggesting that cognitive skills are an increasingly important dimension of educational stratification worldwide.

Online appendix

VERSION

March 2019

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

Anna K. Chmielewski

March 18, 2019

ABSTRACT

The existence of a "socioeconomic achievement gap"—a 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. What is unknown in most countries is whether the SES achievement gap might be *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 have 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 have also increased in many countries with consistently high enrollments, suggesting that cognitive skills are an increasingly important dimension of educational stratification worldwide.

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

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 the sociology of education. 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 and/or out of school as their high-SES counterparts. Across many countries, SES achievement gaps impede upward mobility (Jackson 2013). This contradicts the traditional American view of education as a "great equalizer" (Downey and Condron 2016). However, it may be less surprising in many other 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 from three individual countries find increasing SES achievement gaps in the US (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)—find wide cross-national variation in trends in SES achievement gaps across the years of each assessment (Broer, Bai and Fonseca in press; OECD 2018).

This paper aims to provide the most comprehensive picture to date of cross-national trends in the SES achievement gap. It uses 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). It draws on 30 datasets across 100 countries representing some 5.8

million students and seeks to describe the global trend in SES achievement gaps, as well as to

describe cross-national variation in trends and 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, SES and achievement

measures, and covered different time periods. There is some early international evidence 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" (p. 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 that they

attribute to expanding school access and standardization of school quality.

More recently, single-country studies also suggest increasing SES achievement gaps but

produce some contradictory evidence. Using 19 nationally-representative US studies, Reardon

(2011b) shows that the gaps in reading and math achievement between students from families at

the 90th and 10th income percentiles grew by about 40% between children born in the 1970s and

the 1990s. However, the US 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 et al. (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 1985-1993), 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 1985-1997), Saw (2016) observes rapid growth in math and science achievement gaps between students whose parents attended post-secondary 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. For 60 counties participating in PISA 2015 and one earlier wave (corresponding to approximately birth years 1985-2000), associations between reading, math, and science achievement and an SES index (including parent education, parent occupation, and household possessions) have declined in a majority of countries (OECD 2018). In contrast, for 13 countries in 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) have increased in about half of countries (Broer, Bai and Fonseca in press). 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. Further, 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. At the same time, there is ample international comparative research on 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 the result of 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, Goesling and LeTendre 2002). Since 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 also 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 of the countries participating in international assessments, implying rising standards of living and capacity for public and private investment in education and child wellbeing. 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 et al. (2002) suggested 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 that 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 US and South Korea (Byun and Kim 2010; Reardon 2011b). Income inequality is increasing in many other countries as well, particularly in Europe and Asia (though 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. We may expect that countries with increasing income inequality 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). In the logic of this body of research, then, we would expect that countries that increase the rigidity of curricular differentiation or begin tracking at younger ages would experience increasing SES achievement gaps. However, it is not clear that such changes in tracking systems could explain increasing SES achievement gaps in many countries. Although Byun and Kim (2010) identify increasing 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 and/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.

While 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). In several different countries, research 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, others

have argued 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

growing more unequal 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 American 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 in press; Faircloth, Hoffman and Layne 2013; Gomez Espino 2013; Hays 1996; Karsten 2015; Katartzi 2017; Liu 2016; Quirke 2006; Tan 2017). In the US, 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 are expected to 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, although a small number of economics studies have combined 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 have used 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/or 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 (1995-2015), PIRLS (2001-2011) and PISA (2000-2015). All studies are conducted by the International Association for the Evaluation of Educational Achievement (IEA) except PISA, which is 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, though full population coverage was not achieved in every country-study-year. As 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 have participated in at least two different studies in different years, the final sample is 5541 country-study-gaps within 1026 country-

studies within 100 countries. The 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 Appendix A.

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. Although there are similarities among the different tests of the same subject, only the scores from multiple years of the TIMSS, PIRLS, and PISA studies are strictly comparable. Since 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 that 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% of observations) or science achievement (37.1%) versus reading achievement (reference category; 27.9%).

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 6-8 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, (8) Bachelor's

degree or more. Parent occupation was generally 9-10 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, (10)

Professional. Books in the household were usually reported in 5-6 categories, such as: (1) 0-10

books, (2) 11-25 books, (3) 26-100 books, (4) 101-200 books, (5) 201-500 books, (6) More than

500 books. In the final sample, 34.7% of country-study-gaps are based on parent education as the

SES measure, 25.8% are based on parent occupation, and 39.5% are based on household books.

Although the percentile method I use to calculate SES achievement gaps (described in the

Methods section below) 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 four variables to control for the quality of SES variables.

1. *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 fourth grade; and PISA 2006, 2009, and

    2012. As 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

    Methods), I also include a dummy variable indicating whether each gap was based on

student-reported SES (81.4% of country-study-gaps) or parent-reported SES (18.6%). I interact this variable with gap type (parent education, parent occupation, or household books), as students' and parents' relative accuracy depends on the SES variable they are reporting (Jerrim and Micklewright 2014).

2. *Number of categories.* The percentile method used to calculate SES achievement gaps (described in Methods) 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 larger 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 7 categories.

3. *20% or more students in the bottom SES category.* The percentile method may not perform as well when more than 20% of observations are in the bottom or top SES category (Reardon 2011a). I include a dummy variable indicating whether 20% or more of students fall into the bottom category (14.8% of country-study gaps) versus less than 20% in the bottom category (85.2%).

4. *20% or more students in the top SES category.* I also include a dummy variable indicating whether 20% or more of students fall into the top category (38.7% of country-study gaps) versus less than 20% in the top category (61.3%).[7]

Cohort birth year. I compute the mean birth year for each country-study from student reports either of birth year and month or of 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

13

ranges from 1949.86 in the England FIMS 1964 sample to 2005.78 in the New Zealand TIMSS 2015 4th grade sample. In 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 fourth grade/age 10 (FISS, FIRCS, SISS, RLS, TIMSS, and PIRLS), eighth 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% of observations) or age 15 (56.3%) versus age 14 (reference category; 23.0%).

      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 is obtained from the World Bank. For fourth grade testing cohorts, I use the proportion enrolled in primary school in the testing year; for eighth grade and 15-year-old cohorts, I use the proportion enrolled in secondary school.

Proportion Immigrant Background. Proportion of students reporting first- or second-generation immigrant status is computed from the microdata.

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

Income Inequality. Gini coefficient ranging from 0 (perfect equality) to 1 (perfect inequality) is obtained 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 US 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 is obtained 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 aged 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 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] Then 1000 bootstrap samples are drawn from each of the five imputed datasets. In each of these samples, 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 $\bar{Y}$ 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); a cubic function estimating the association between Y and θ is fit using weighted least squares (weighting by the inverse squared standard error of $\bar{Y}_k$).[13] This yields a fitted curve:

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

[1]

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

$$\hat{\delta}^{90/10} = \left[\hat{Y}|\theta = .9\right] - \left[\hat{Y}|\theta = .1\right] = 0.8\hat{b} + .8\hat{c} + .728\hat{d}$$

[2]

$$\hat{\delta}^{90/50} = \left[\hat{Y}|\theta = .9\right] - \left[\hat{Y}|\theta = .5\right] = 0.4\hat{b} + .56\hat{c} + .604\hat{d}$$

[3]

$$\hat{\delta}^{50/10} = \left[\hat{Y}|\theta = .5\right] - \left[\hat{Y}|\theta = .1\right] = 0.4\hat{b} + .24\hat{c} + .124\hat{d}$$

[4]

As mentioned in the Data section, gaps will tend to be attenuated in country-studies where SES is less reliably measured. Due to the standardization of achievement described above, 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, the 1000 bootstrap sample gaps are used 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 above, 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

17

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. Since there are up to nine different outcomes in each study-year (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, I implement this model using a multivariate variance-known model. The model was originally developed for use in meta-analysis with multiple outcomes, but 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 that 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:

$$\hat{G}_{pjk} = \boldsymbol{\delta}\left(\mathbf{T}_{pjk}\right) + \boldsymbol{\alpha}\left(\mathbf{S}_{pjk}\right) + \boldsymbol{\lambda}\left(\mathbf{A}_{jk}\right) + \boldsymbol{\gamma}\left(\mathbf{T}_{pjk}C_{jk}\right) + \boldsymbol{u}_k\left(\mathbf{T}_{pjk}\right) + \boldsymbol{w}_k\left(\mathbf{T}_{pjk}C_{jk}\right) + \boldsymbol{r}_{jk}\left(\mathbf{T}_{pjk}\right)$$
$$+ \epsilon_{pjk},$$

$$\begin{bmatrix} \epsilon_{1jk} \\ \vdots \\ \epsilon_{9jk} \end{bmatrix} \sim MVN[\mathbf{0}, \mathbf{V}_{jk}]; \quad \begin{bmatrix} r_{1jk} \\ r_{2jk} \\ r_{3jk} \end{bmatrix} \sim MVN[\mathbf{0}, \boldsymbol{\Sigma}]; \quad \begin{bmatrix} u_{1k} \\ u_{2k} \\ u_{3k} \\ w_{1k} \\ w_{2k} \\ w_{3k} \end{bmatrix} \sim MVN[\mathbf{0}, \boldsymbol{\tau}]$$

[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 $\boldsymbol{\gamma}$ 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 $6000). 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 very 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 US 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-old students 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 has declined slightly in the US 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 US 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 et al.'s (2019) finding of no change in US achievement gaps between the 90th and 10th percentiles of an

index of SES (including parent education and household possessions) and to Broer et al.'s (in press) finding of a small decline in US SES achievement gaps across recent waves of TIMSS. However, the slight decline in Figure 1 is less pronounced than the more marked decline in US SES achievement gaps reported for recent waves of PISA (OECD 2018). Possible methodological reasons for this discrepancy are discussed below. My estimates for US 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, though the parent occupation gap shows a slight decline like the parent education gap, while 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 US 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 US and several other high-income countries and appears to cause achievement gaps based on parent education and occupation but not books to be underestimated in later years (discussed further below).

(Figure 1 about here)

Table 1 puts the US 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, or household books) using interactions between cohort birth year and gap type indicators. As described in the Methods section, the multivariate variance-known models estimate a different intercept for each of the three gap types. Since 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.

(Table 1 about here)

Turning to the control variables, on average, math achievement gaps are significantly larger than reading gaps, which is consistent with prior US 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 noted 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, though the association is small and not significantly different from 0. More than 20% of students falling into the bottom or top category of the SES variable is associated with smaller estimated gaps. This suggests that 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 US and other wealthy countries with large proportions of students in the top categories in later years. Nevertheless, when SES variable quality measures are omitted, results are very similar.[22] SES achievement gaps tend to be smaller when they are estimated from students tested at age 10 than at age 14, while 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 have increased. Net of controls, 90/10 parent education and occupation gaps both increase at a rate of 0.007 SD of achievement per year, and 90/10 books gaps increase 0.008 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 0.4 SD of achievement for all three gap types. As mentioned above, 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 ($p > 0.5$). 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 0.58, for parent education and books gaps is 0.51, and for parent occupation and books gaps is 0.90. 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 very similar regardless of the SES variable used, suggesting that 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 0.007. The coefficient estimates for all control variables are very 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 above. 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 that 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% of countries' parent education cohort slopes fall within the range (-0.006, 0.019). The

95% plausible value ranges for parent occupation, books, and pooled cohort slopes are (-0.003,

0.017), (-0.009, 0.025), and (-0.004, 0.019), respectively. Also implied is that the share of

countries with trends greater than 0 is approximately 84% for parent education, 92% for parent

occupation, 82% for books, and 90% for pooled gaps. Thus, while a large majority of countries

experience increasing SES achievement gaps, the size of these increases varies widely, and gaps

decline in about 8-18% 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 Appendix Table A1 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 0.008 SDs. The gap increase for African

countries is larger though 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 very similar to in Western countries. The only region with a significantly different

gap trend from the West is Latin America and the Caribbean, where gaps have 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 $6000. 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 is expected to be

positive, as prior research suggests that countries at lower levels of economic development

experienced larger increases in SES achievement gaps between the 1970s and 1990s (Baker,

Goesling and LeTendre 2002). The coefficient is indeed positive but not significant.

(Table 2 about here)

However, Baker et al.'s (2002) findings pertain to cohorts born between approximately

1960 and 1980, a shorter time frame than the present study. Model 5 includes a squared cohort

birth year term and interaction with country income, in order to estimate curvilinear trends and

allow trends to differ by country income. The main effect for the squared term is positive, though

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 that the gap trend

for low-income countries curves downward. 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, while 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 high-income countries, consistent with earlier results from Baker et al. (2002). Yet in

recent years, this pattern reversed, and high-income countries experienced slightly greater

increases in gaps than 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.

(Figure 2 about here)

In order to examine cross-national variation in gap trends in more detail, Figure 3 plots estimated quadratic trends for 24 countries. The countries are selected from those with the most available data points over the longest time span, while also providing 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—though not all—experience 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, Scotland) tend to be high-income and, like the US, already had large gaps in early cohorts and have large proportions of students in the top categories of parent education and occupation in later years. However, it should be noted that there are several other high-income countries with large shares of students in the top education and occupation categories that nevertheless experience sizable increases in gaps (e.g., Australia, Canada, Norway and Sweden).

(Figure 3 about here)

Also visible in Figure 3 are the estimated trends for several countries that have been 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). It should be noted that both prior studies 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 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 et al.'s (in press) report on TIMSS trends, I find increasing SES achievement gaps for Hungary, Iran, New Zealand, and Singapore. Consistent 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 et al. (in press) 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 capture 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, while 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 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. While 90/50 gaps have increased very little in high-income countries, they have increased significantly more in low-income countries. 50/10 gaps have 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, or in other words, it is driven by the achievement of middle- and high-SES students pulling away from that of low-SES students.[27] However, it should be noted 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 0.56 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. 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, i.e. the intercept of the model (displayed in the lower "Level 3" portion of the table). 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. 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). It should be kept in mind that, with a country-level sample size of only 78, the level 3 portion of the model may be overfit.

(Table 3 about here)

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 those 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

0.04 SD ($p < .001$). Also as expected, an increasing share of immigrant students is associated with increasing achievement gaps, though this relationship is not significant.

The next two covariates pertain to economic changes. As expected based on previous research by Baker et al. (2002), countries with increasing GDPs per capita tend to experience increasing SES achievement gaps, though 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 0.1 in the Gini coefficient is associated with a decrease in the SES achievement gap of 0.19 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. However, joint hypothesis tests show 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 0.04 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, though 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, though 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 these full models to reduced models that include 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%, 7%, and 8% 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 15-35% 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]

A number of robustness checks were performed on these results, which are reported in the online appendices. 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 has 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 have increased by about 50%, gaps based on parent occupation by about 55%, and gaps based on household books by about 40%. 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 that 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 within-student 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 have increased slightly more than those based on parent education or occupation. 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.

While SES achievement gaps have increased in most countries, 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 have dramatically expanded enrollment. For example, Luxembourg and Ireland have both increased secondary school enrollment by over 20 percentage points over the years they have participated in international assessments, and Thailand has increased secondary enrollment by nearly 70 percentage points. The results for enrollment are consistent with Baker et al.'s (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 have 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, including the US, England, Finland, Israel, and Japan. These 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 US, 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, for example 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 US (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 that is associated with declining SES achievement gaps in the multivariate models, consistent with findings by van de Werfhorst (2018). However, changes in tracking policies cannot help to explain the secular global trend of increasing SES achievement gaps, as far more countries have moved the age of track selection later 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 US, 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. Although in such a society, all parents may equally recognize the importance of academic skills, higher-SES families have greater resources and information about how to foster their children's achievement (Ishizuka 2018; Lareau 2000). It should be noted 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 US.

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 US and UK (Bailey and Dynarski 2011; Belley and Lochner 2007; Galindo-Rueda and Vignoles 2005). However, in the US, 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 UK and Australia (Jerrim, Chmielewski and Parker 2015). Growing SES achievement gaps may also have political implications. Though belief in meritocracy is growing in many countries, this belief is strongly socioeconomically graded, particularly in the 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 US 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, results for trends across multiple waves of a single international assessment, such as PISA, are sometimes contradicted by results from other assessments, such as TIMSS or PIRLS. For example, the OECD (2016) finds declining SES achievement gaps across the last four waves of PISA, but this study finds increasing gaps when pooling these PISA

datasets together with 26 other assessments.[31] These differences are due not only to the broader

set of countries included when incorporating all international assessments, but also to occasional

differences in findings even for the same country. These discrepancies are likely due to

differences in the testing frameworks and SES measures of different international assessments,

although this paper has made a variety of efforts to harmonize measures where possible.

Nevertheless, the precise trends in the SES achievement gap for each individual country remains

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 US, France, Hong

Kong, and Russia (Broer, Bai and Fonseca in press; 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 educational

and social policies that may help to mitigate disparities in learning opportunities for high- and

low-SES children.

ENDNOTES

[1] See Appendix tables B13 and B14 for models based on TIMSS data, excluding country-study-years that IEA

considers not comparable for trend analysis.

[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 the within-country variance of achievement is declining across waves in PISA,

PIRLS, and TIMSS science (but not TIMSS math) (Appendix C). Yet SES gaps in unstandardized achievement are

increasing on average for all test instruments except PISA math and science (Appendix B).

[4] Supplemental analyses check the robustness of results by running models separately by subject (Appendix K); separately for TIMSS, PIRLS, and PISA (Appendix B); and computing gaps based on achievement rank (Appendix N). Results are similar.

[5] Additional detail on the treatment of mothers' and fathers' SES characteristics is reported in Appendix F.

[6] Models were also run with categories harmonized across datasets, and results were very similar (see Appendix B).

[7] The large number of gaps with more than 20% 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 and/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. In particular, these gaps usually appear to be underestimated, as they increase less than household books gaps 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, models were also run excluding gaps with 20% or more students falling into the top or bottom category, and results were very similar (see Appendix Table H6).

[8] Assessments of twelfth 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), while the World Bank uses household consumption in most countries (which I consider more comparable to disposable income) but uses gross income (pre-tax and transfer) in some countries (which I consider less comparable to disposable income). All three sources (the World Bank, LIS, and OECD) adjust by household size. Since this study is interested in comparing changes in time-varying covariates within countries over time, only one data source is used for each country. The validity of results, then, 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] Models were also run using listwise deletion rather than multiple imputation of missing data, and results were similar (see Appendix L).

[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 5 imputed datasets, the main analyses in this paper required computing nearly 28 million gap estimates (5541 SES achievement gaps*1000 bootstraps*5 imputed datasets), in addition to several million more gap estimates for robustness checks reported in the appendices. This required approximately 3240 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 5 new imputed datasets and rerun all analyses, the estimate for the trend in the SES achievement gap would likely remain very similar. However, the standard error would likely change (it might either increase or decrease). As 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 (2011). Quadratic or linear functions are used in country-years where there are insufficient SES categories. Linear functions are also used for country-years when more than 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. Models were also run with all linear gaps, and results are similar (see Appendix L).

[14] See Appendices D and H for more information on the reliability adjustment. Models were also run without adjusting for reliability, and results were similar (see Appendix Table D4).

[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, 10 imputed datasets were generated and combined with the 10 plausible values of achievement.

[17] This issue is discussed in more detail in Appendix E.

[18] Models run separately for each SES variable are reported in Appendix J. Additional analyses of gaps computed from models including all three SES variables are reported in Appendix G.

[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 $\epsilon^*_{pjk} \sim N(0, \boldsymbol{I}_P)$, where $\boldsymbol{I}_P$ 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 Appendix Table K2 for estimates separate gap trends by subject.

[21] See Appendix Table H2 for a comparison of gaps based on student- and parent-reported SES.

[22] See Appendix Table H5 for models omitting SES variable quality measures.

[23] See Appendix Table K1 for estimates of separate gap trends by age.

[24] See Appendix J for models run separately by SES variable.

[25] See Appendix B for a comparison of results between Broer et al. (in press) and the current study.

[26] See Appendix H 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 Appendix L).

[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 both a practical choice because of 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 US are "public," but academy schools in the UK 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 Appendices C

and E, and those pertaining to changing measurement error are available in Appendices D and H. Additional

supplementary analyses show that gap increases are positive when analyzing the different trend studies separately

(PISA reading, TIMSS 8th grade science, PIRLS, etc.), although the increases do not reach significance in every case (see Appendix B). Further examination of the trend studies shows that increasing gaps in these studies correspond to large increases in the achievement of high-SES students, coupled with stagnation or declines in the achievement of low-SES students (Appendix C). Increases in SES achievement gaps are robust to a variety of different specifications of the hierarchical growth model (Appendix L).

[30] See Appendix Table G1 for analyses of trends in within-student correlations between different SES variables.

[31] Trends are also positive when limiting only to recent years (see Appendix B).

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(1):1.

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*: Emerald Group Publishing.

Bailey, M.J., and S.M. Dynarski. 2011. "Inequality in Postsecondary Education." in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, edited by Richard 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 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. in press. *Socioeconomic Inequality and Educational Outcomes: Evidence from Twenty Years of TIMSS*: IEA Research for Education, Vol. 5.

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* 22(52):781-861.

Byun, Soo-yong, and Kyung-keun Kim. 2010. "Educational inequality in South Korea: The widening socioeconomic gap in student achievement." *Research in Sociology of Education* 17(155-182).

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." in *XVIII ISA World Congress of Sociology (July 13-19, 2014)*: Isaconf.

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. in press. "The many (subtle) ways parents game the system: Mixed-method evidence on the transition into secondary school tracks in Germany." *Sociology of Education*.

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*: Routledge.

Falch, Torberg, and Justina A.V. Fischer. 2012. "Public sector decentralization and school performance: International evidence." *Economics Letters* 114:276-79.

Galindo-Rueda, Fernando, and Anna Vignoles. 2005. "The declining relative importance of ability in predicting educational attainment." *Journal of Human Resources* 40(2):335.

Ganzeboom, Harry BG, 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 Woessmann. 2019. "The Unwavering SES Achievement Gap: Trends in U.S. Student Performance." NBER Working Paper No. 2548.

Hanushek, Eric A., and Ludger Wößmann. 2012. "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation." *Journal of Economic Growth* 17:267-321.

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*.

Jackson, Michelle. 2013. *Determined to Succeed?: Performance Versus Choice in Educational
    Attainment*: 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:jcu072.

Kalaian, Hripsime A, and Stephen W Raudenbush. 1996. "A multivariate mixed linear model for
    meta-analysis." American Psychological Association.

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: 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." in *2011 Annual meeting of the Population Association of America*. Washington, DC.

Lareau, Annette. 2000. *Home Advantage: Social Class and Parental Intervention in Elementary Education, Second Edition*. Lanham, MD: Rowman & Littlefield Publishers, Inc.

—. 2003. *Unequal Childhoods: Class, Race, and Family Life*: California.

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 JB. 2019. "The paradox of inequality: income inequality and belief in meritocracy go hand in hand." *Socio-Economic Review*.

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. 2015. "In It Together: Why Less Inequality Benefits All." Paris: OECD Publishing.

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

—. 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: 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." in *Brookings Papers on Economic Activity*.

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.

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):2332858416657343.

Roex, Karlijn LA, 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, M. 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.

Schomaker, Michael, and Christian Heumann. 2016. "Bootstrap inference when using multiple imputation." *arXiv preprint arXiv:1602.07933*.

Smith, Laura, Special Agent FBI. 2019. "College admissions bribery scheme affidavit." in *The Washington Post*.

Söderström, Martin, and Roope Uusitalo. 2010. "School choice and segregation: Evidence from an admission reform." *The Scandinavian Journal of Economics* 112(1):55-76.

Tan, Charlene. 2017. "Private supplementary tutoring and parentocracy in Singapore." *Interchange* 48(4):315-29.

UNDP. 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*.

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*: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 David P. Baker and Alexander W. Wiseman: Emerald Group Publishing Limited.

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

	(1) 3 Cohort Slopes			(2) 1 Cohort Slope		
	coef	(se)		coef	(se)	
Parent education gaps intercept	1.032	(0.030)	***	1.039	(0.030)	***
Parent occupation gaps intercept	0.958	(0.030)	***	0.964	(0.030)	***
Household books gaps intercept	1.299	(0.041)	***	1.294	(0.041)	***
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)	***	0.112	(0.031)	***
Parent-reported × Parent occupation	0.075	(0.025)	**	0.073	(0.024)	**
Parent-reported × Books	-0.039	(0.029)		-0.017	(0.026)	
Number of categories (centered at 7)	0.003	(0.003)		0.002	(0.003)	
≥ 20% in bottom category	-0.065	(0.021)	**	-0.063	(0.021)	**
≥ 20% in top category	-0.135	(0.013)	***	-0.146	(0.013)	***
Level 2 - Study-years						
Age at testing (ref=14)						
Age 10 at testing	-0.170	(0.024)	***	-0.168	(0.024)	***
Age 15 at testing	-0.024	(0.020)		-0.023	(0.020)	
Cohort birth year × Parent education	0.007	(0.001)	***			
Cohort birth year × Parent occupation	0.007	(0.001)	***			
Cohort birth year × Books	0.008	(0.001)	***			
Cohort birth year				0.007	(0.001)	***
Random effects						
Level 2 - Residual variance between studies in...						
Parent education intercepts	0.03736			0.03831		
Parent occupation intercepts	0.02322			0.02284		
Books intercepts	0.03698			0.03823		
Level 3 - Residual variance between countries in...						
Parent education intercepts	0.05426			0.05362		
Parent occupation intercepts	0.05227			0.05330		
Books intercepts	0.11590			0.12149		
Parent education cohort slopes	0.00004					
Parent occupation cohort slopes	0.00003					
Books cohort slopes	0.00007					
Cohort slopes				0.00003		
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. Two-tailed tests.

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

	(3) Region Interactions		(4) Income Interaction		(5) Quadratic		(6) 90/50 Gap		(7) 50/10 Gap	
	coef	(se)	coef	(se)	coef	(se)	coef	(se)	coef	(se)
Parent education gaps intercept	1.126	(0.027) ***	1.074	(0.035) ***	1.066	(0.037) ***	0.548	(0.019) ***	0.514	(0.022) ***
Parent occupation gaps intercept	1.111	(0.029) ***	1.054	(0.038) ***	1.046	(0.042) ***	0.564	(0.020) ***	0.481	(0.022) ***
Household books gaps intercept	1.563	(0.038) ***	1.434	(0.047) ***	1.426	(0.049) ***	0.622	(0.022) ***	0.806	(0.029) ***
Level 1 - Gaps										
Subject controls (ref=Reading)	x		x		x		x		x	
SES variable quality measures	x		x		x		x		x	
Level 2 - Study-years										
Age at testing controls (ref=14)	x		x		x		x		x	
Cohort birth year	0.008	(0.001) ***	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										
Region (ref=Western) × Intercept interactions	x									
Mid/low-income country × Intercept interactions			x		x		x		x	
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) *								
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) *				
N (Level 1 - gaps)	5541		5541		5541		5541		5541	
N (Level 2 - study-years)	1026		1026		1026		1026		1026	
N (Level 3 - countries)	100		100		100		100		100	

+ p<.1, * p<.05, ** p<.01, *** p < .001. Two-tailed tests.

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). 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 Appendix Table L2 for full results.

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

|	(8)		
	90/10 Gap		
	coef	(se)	
Parent education gaps intercept	0.970	(0.046)	***
Parent occupation gaps intercept	0.969	(0.049)	***
Household books gaps intercept	1.405	(0.057)	***
Level 1 - Gaps			
Subject controls (ref=Reading)	yes		
SES variable quality measures	yes		
Level 2 - Study-years			
Study fixed effects (ref=TIMSS 2003 Grade 8)	yes		
School enrollment (proportion)	0.486	(0.107)	***
Immigrant background (proportion)	0.226	(0.250)	
GDP per capita (logged)	0.055	(0.059)	
Income inequality (Gini)	-1.913	(0.887)	*
Mid/low-income country × Income inequality	2.539	(1.129)	*
Age when tracking begins	-0.037	(0.016)	*
Private school enrollment (proportion)	0.240	(0.249)	
Expecting higher education (proportion)	-0.029	(0.094)	
Level 3 - Countries			
Mid/low-income country × Intercept interactions	yes		
Mean school enrollment	0.640	(0.317)	*
Mean proportion immigrant background	0.134	(0.261)	
Mean GDP per capita (logged)	0.142	(0.041)	***
Mean income inequality	0.605	(0.317)	+
Mean age when tracking begins	-0.035	(0.011)	**
Mean private school enrollment	0.043	(0.099)	
Mean proportion expecting higher education	-0.255	(0.169)	
N (Level 1 - gaps)	4604		
N (Level 2 - study-years)	855		
N (Level 3 - countries)	78		

+ p<.1, * p<.05, ** p<.01, *** p < .001. Two-tailed tests.
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). Coefficents for control variables (subject, SES variable quality measures, and study fixed effects) are omitted due to space constraints; see Appendix Table M1 for full results.

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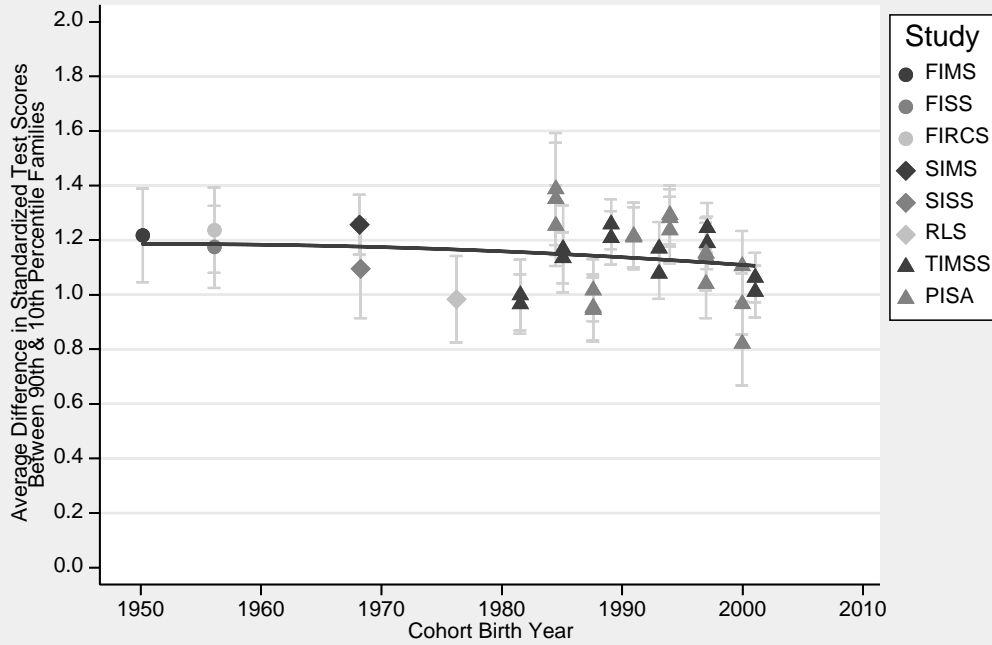
Notes: Gaps and quadratic fit line adjusted for age of testing and subject. Gray brackets are 95% confidence intervals.

Figure 2. Estimated Quadratic Trends in 90/10 SES Achievement Gaps, by Country Income Level, 1950-2005 Birth Cohorts
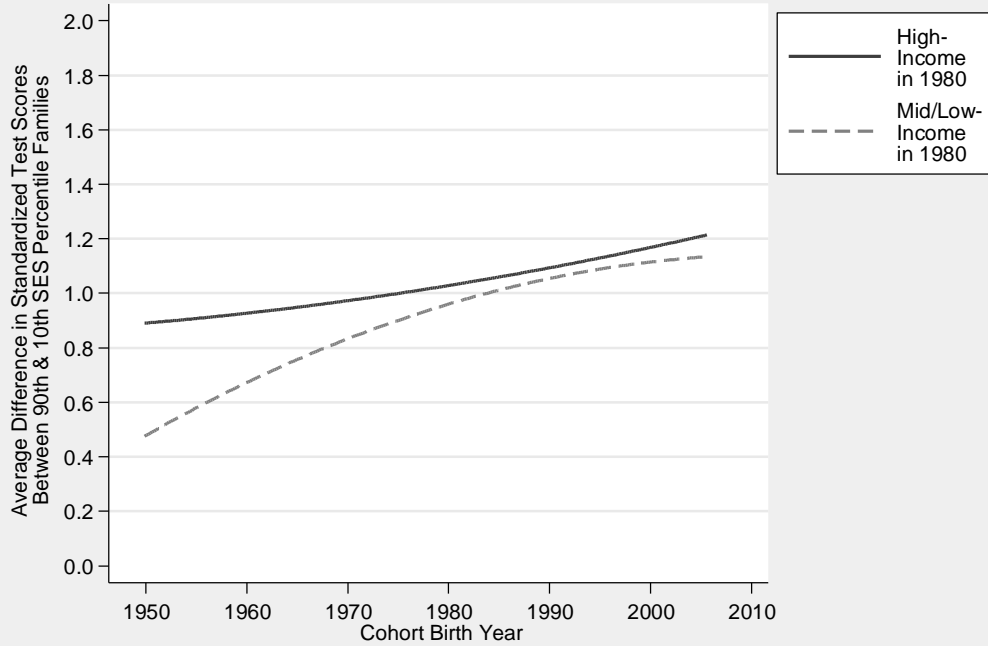
Notes: "High-income" countries had GDPs per capita of at least $6000 in 1980 (see Appendix Table A1 for coding). Trend lines are estimates from Model 5 (Table 2). Fixed values for control variables: SES=parent education, subject=math, all others=0 or reference category.

Figure 3. Estimated Quadratic Trends in 90/10 SES Achievement Gaps, Selected Countries
(Number of Study-Years in Parentheses)
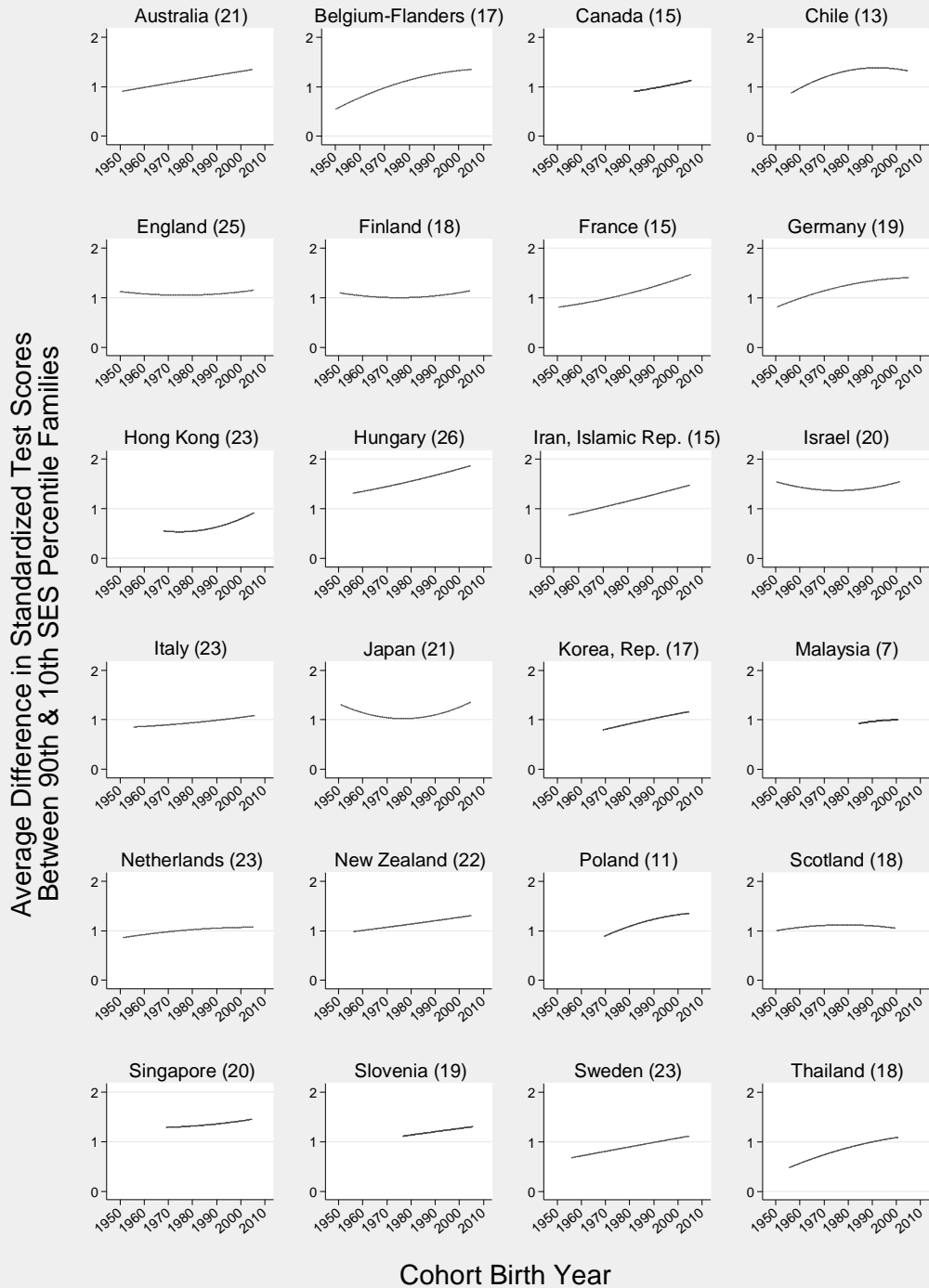
Notes: Countries sorted alphabetically. Trend lines are derived from shrunken empirical Bayes estimates from Model 5 (Table 2). Fixed values for control variables: SES=parent education, subject=math, all others=0 or reference category.