

Patterns of Cross-National Variation in the Association Between Income and Academic Achievement

Anna K. Chmielewski

University of Toronto

Sean F. Reardon

Stanford University

In a recent paper, Reardon found that the relationship between family income and children's academic achievement grew substantially stronger in the 1980s and 1990s in the United States. We provide an international context for these results by examining the income–achievement association in 19 other Organisation for Economic Co-operation and Development countries using data from the Progress in International Reading Literacy Study and the Programme for International Student Assessment. First, we calculate and compare the magnitude of “income achievement gaps” across this sample of countries. Second, we investigate the association between the size of a country's income achievement gap, its income inequality, and a variety of other country characteristics. We find considerable variation across countries in income achievement gaps. Moreover, the U.S. income achievement gap is quite large in comparison to this sample of countries. Our multivariate analyses show that the income achievement gap is positively associated with educational differentiation, modestly negatively associated with curricular standardization, and positively associated with national levels of poverty and inequality.

Keywords: *income achievement gap, cross-national comparisons, international large-scale assessments*

THE relationship between family income and children's academic achievement grew substantially stronger in the 1980s and 1990s in the United States (Reardon, 2011). This increase occurred at the same time that income inequality in the United States grew sharply, raising the question of whether and how income inequality relates to the size of the relationship between family income and student test performance. That family income/socioeconomic status (SES) is related to children's academic achievement is not surprising; that this relationship grew so rapidly in the last several decades, however, is rather surprising. The U.S. trends suggest that some of this growth may have been the result of rising income inequality. As one way of investigating the relationships between income inequality, school system characteristics, and the income achievement gap, we examine data from multiple countries with widely varying levels of income inequality and school institutional structures.

Specifically, we address two questions in the current paper. First, we investigate whether income is equally determinative of academic achievement across countries. We use data from the Progress in International Reading Literacy Study (PIRLS) and the Programme for International Student Assessment (PISA) to estimate the magnitude of the “income achievement gap” between children from high- and

low-income families in 20 Organisation for Economic Co-operation and Development (OECD) countries. Second, we investigate the association between the size of a country's income achievement gap, its poverty and inequality levels, and a variety of other country characteristics, including welfare policies, parental support policies, and the degree of differentiation and standardization of national school systems. Here we hypothesize (a) that poverty and inequality are positively associated with the strength of the relationship between income and achievement; (b) that social welfare policies are negatively related to the income achievement gap; (c) that differentiation within the school system—in the form of curricular tracking and size of the private sector—is positively associated with the strength of the relationship between income and achievement; and (d) that standardization of the school system—indicated by the use of centralized examinations—is negatively associated with the strength of the relationship between income and achievement.

To date, there has been little comparative cross-national research investigating the association between family income and student achievement in different countries, in large part because family income data are rarely available in cross-national studies of student achievement, such as PISA,



Trends in International Mathematics and Science Study (TIMSS), and PIRLS. In 2001, however, the PIRLS parent survey did ask respondents to report their household income. Likewise, in 2006, 2009, and 2012, the PISA study administered a parent survey that included a household income item in a limited set of countries. We make use of these data to compute income achievement gaps across the participating countries and to compare them to the magnitude of income achievement gaps in the United States.

The Income–Achievement Association in the United States

Reardon (2011), using data from 12 nationally representative samples of U.S. students tested in reading and math from 1960 through 2006, shows that the income achievement gap grew rapidly in the United States among cohorts of students born in the mid-1970s through 2000. Reardon measures the income achievement gap by estimating the average difference in test scores between students with family income at the 90th percentile of the income distribution and those with family income at the 10th percentile of the income distribution. This “90/10 income achievement gap” was roughly 0.85 to 0.90 standard deviations among cohorts of students born in the mid-1970s and roughly 1.20 to 1.25 standard deviations among those born in the mid-1990s, an increase of roughly 40% over 20 years. More recent data suggest that the gap declined by 15%, however, in the subsequent decade (Reardon & Portilla, 2015). A similar trend holds for the income achievement gap in math.

Reardon (2011) describes several other key features of the income achievement gap in the United States. First, the gap does not grow larger as children progress in school. In longitudinal studies, the income achievement gap is large when children enter kindergarten or elementary school and remains roughly the same size through elementary and secondary school. This finding suggests that schooling alone is not a primary cause of the gap (because the gap is present when children enter formal schooling). Second, most of the growth of the 90/10 income achievement gap in the United States resulted from the growth in the achievement gap between children with family incomes between the 90th and 50th percentiles of the income distribution; the achievement gap between children with family incomes between the 50th and 10th percentiles was unchanged in reading and grew only very slightly in math. Third, the trend in the income achievement gap is much steeper than the corresponding trend in the “parental-education achievement gap” (the achievement gap between children with highly and less educated parents. This difference in the income and education achievement gap trends suggests that—in the United States at least—family income is not identical to parental education in its relationship to children’s achievement over time. The explanation for this difference is not clear, however.

The reasons for the growth in the income achievement gap in the United States are also unclear. Certainly, the gap has grown at the same time that income inequality has grown, but the mechanisms linking this to the achievement gap are not obvious. If the ways that families use their income (and other income-related resources) to develop their children’s academic skills were constant over time (that is, if the cognitive returns to family income were constant), rising income inequality would lead, mechanically, to growing income achievement gaps. But achievement gaps might change as well because of a change in the returns to family income. If the ways that high- and low-income families use their economic, educational, and social resources have changed differently over time, then the association between income and academic performance might change differently than implied mechanically by changes in income inequality. The income achievement gap might also change because of changes in the correlation between family income and other resources that matter for children’s cognitive development. Reardon (2011) notes that family income has become more correlated with both parental education (because of the rising economic returns to education and because of increased marital homophily) and with the income of one’s neighbors (Bischoff & Reardon, 2014; Owens, 2015), both of which may affect children’s educational performance (Chetty, Hendren, & Katz, 2015; Reardon, 2011).

Reardon (2011) finds that the growth of the income achievement gap does not map well onto the growth of income inequality (income inequality among families with children in the United States grew most rapidly between families at the 50th and 10th percentiles of the income distribution in the 1970s and 1980s, but the achievement gap grew most rapidly between children born in the 1980s and 1990s in families at the 90th and 50th percentiles of the income distribution). In addition, the gap appears to have narrowed in recent years, even as income inequality continued to grow (Bassok & Latham, 2014; Magnuson & Duncan, 2014; Reardon & Portilla, 2015). These patterns suggest that the changes in the income achievement gap in the United States are not driven solely by mechanical effects of changing income inequality.

Nor is it clear that the growth of the gap can be attributed to the U.S. schooling system. The fact that the income achievement gap is very large as early as age 5 or 6 suggests that the gap has grown more because of early childhood experiences than because of elementary or secondary schooling. Likewise, the decline in the income achievement gap in recent years is driven by changes in gaps in school readiness when children enter kindergarten (Bassok & Latham, 2014; Magnuson & Duncan, 2014; Reardon & Portilla, 2015), a change that cannot be directly attributed to effects of schooling.

Nonetheless, the fact that the descriptive evidence in the United States does not clearly point to either rising income

inequality or features of the educational system as the source(s) of changes in the income achievement gap does not mean that inequality and education may not contribute to the achievement gap in more complex ways. Inequality may increase parents' status anxiety and lead to changes in how much time and resources they invest in their children's education. Those with more resources can invest more in their children than those with fewer resources, particularly if the social welfare system is too weak to support low-income families. The growing importance of standardized testing in U.S. schooling may lead parents and schools to focus increasingly on school readiness and test performance. And in a decentralized educational system like that of the United States, family resources may affect the quality of schooling available to children. Indeed, Owens (2015) suggests that parental focus on schooling, coupled with rising income inequality, has led to growing economic segregation both among neighborhoods and schools. Thus, income inequality, coupled with weak social welfare systems, high levels of income segregation, and a decentralized schooling system, will likely produce relatively large achievement disparities between children from rich and poor families. In the remainder of this paper, we use cross-national comparisons to investigate this possibility.

The Income–Achievement Association in Cross-National Comparison

In this paper, we focus on two main sets of national characteristics that may be associated with the strength of the income–achievement association: (a) broad social factors associated with educational (in)equality, including income inequality levels, poverty rates, school segregation, and equity-enhancing social welfare policies, and (b) features of the educational system that may affect educational inequality, including structural differentiation of the schooling system and standardization of the curriculum.

Social Factors Associated With Educational Inequality

As noted above, there is little cross-national research on the income–achievement association; there is, however, a great deal of research on the SES–achievement association. Much of this research is based on the student-reported socioeconomic background variables in the PISA studies, which included parents' educational attainment, parents' occupations, and an inventory of home possessions. The OECD combined these responses to create an index of economic, social, and cultural status (ESCS; OECD, 2009a, 2010a). Many secondary analyses of the PISA data sets use the ESCS index to capture students' SES, whereas some others use parent education, occupation, or the component possessions scales. Research using TIMSS or PIRLS often uses parent education or books in the home (which are

student reported in TIMSS and parent reported in PIRLS) as a measure of SES.

Previous cross-national research using PISA, TIMSS, and PIRLS has not found a strong relationship between income inequality and SES achievement gaps. Several authors, as well as the OECD's PISA reports, have found a very weak or no relationship between income inequality and SES achievement gaps (Dupriez & Dumay, 2006; Duru-Bellat & Suchaut, 2005; Marks, 2005; OECD, 2010a). Chudgar and Luschei (2009) found that there is more variation in test scores between schools in countries with high income inequality than in those with lower inequality. This suggests that income inequality may be associated with—and perhaps operate through—between-school segregation and/or resource differences between schools.

Indeed, increasing income inequality has been shown to be associated with increasing income segregation of neighborhoods, particularly, segregation of high-income families and families with children, at least in the United States (Owens, 2015; Reardon & Bischoff, 2011). If neighborhood segregation results in greater income segregation of schools, this could also drive socioeconomic disparities in achievement, especially those at the high end of the SES scale. Likewise, income inequality in wealthy countries is associated with high levels of poverty (conditional on per capita GDP); poverty rates may be predictive of income achievement gaps, particularly, achievement gaps between low- and middle-income students. In this paper, we measure income inequality directly, using the Gini index, but we also create a more general index of poverty and inequality. In developed countries, income inequality is strongly correlated with child poverty rates (because high poverty rates in a rich country implies substantial income inequality) as well as other sequelae of poverty and inequality, such as high rates of teen childbirth, high rates of low birthweight, and high levels of segregation by income among schools.

Finally, income inequality and poverty may matter less than social welfare policies and supports for families with children. Countries with a strong social safety net, generous social welfare policies, and publicly funded early childhood educational programs may have much lower educational inequality than countries with less effective social welfare regimes, even if both have the same level of income inequality and poverty.

Features of the Educational System Associated With Educational Inequality

Typologies of the institutional structures of educational systems focus on two key dimensions: differentiation and standardization (Allmendinger, 1989; Shavit & Müller, 1998; Van de Werfhorst & Mijs, 2010). Differentiation, when students are placed in different curricular tracks on the basis of ability or prior achievement, can reproduce social

class differences if lower-SES students are overwhelmingly assigned to lower tracks (either because of low prior performance or discrimination), where educational quality is much lower. On the other hand, standardization of resources and curricula may produce greater homogeneity of educational quality across schools.

Differentiation. Most international research on the topic of differentiation defines tracking in terms of overarching programs that determine course work in all subjects and are usually located in separate school buildings. Such tracking is widely practiced in Europe, in contrast with the within-school, subject-specific tracking common in the United States (Chmielewski, 2014). Although within- and between-school tracking have similar achievement gaps between tracks, between-school tracking is more predictive of SES achievement gaps because it is typically more socioeconomically segregated (Chmielewski, 2014). Among countries with some between-school tracking, the degree of differentiation varies according to how early tracking begins and the number of distinct tracks. A large body of literature has found that SES gaps are larger in countries with more differentiated school systems (Ammermüller, 2005; Bol, Witschge, Van de Werfhorst, & Dronkers, 2014; Brunello & Checchi, 2007; Causa & Chapuis, 2009; Dupriez & Dumay, 2006; Duru-Bellat & Suchaut, 2005; Field, Kuczera, & Pont, 2007; Horn, 2009; Marks, 2005; Marks, Cresswell, & Ainley, 2006; Schütz, Ursprung, & Wößmann, 2008; Schütz, West, & Wößmann, 2007) and in regions within countries with greater differentiation (Bauer & Riphahn, 2006; Mühlengeweg, 2007; Wößmann, 2007). However, two studies have cast some doubt on these findings by noting that the stronger relationship between SES and achievement in more differentiated countries was also evident in primary school math and reading tests before tracking had begun (Jakubowski, 2010; Waldinger, 2006). On the other hand, Schubert and Becker (2010) found that SES gaps did increase in Germany (which has a highly differentiated system) between the fourth grade and age 15. A further attribute of country school systems that characterizes the degree of differentiation is the prevalence of grade repetition. This could exacerbate socioeconomic inequalities in achievement if disadvantaged students tend to disproportionately repeat grades, thus progressing through school at a slower pace and being exposed to less advanced curricular content. Dupriez and Dumay (2006) found that, as part of index along with tracking, grade repetition was associated with larger SES gaps.

Another important characteristic related to differentiation is the size of the private school sector. The OECD's initial results using PISA found that private schooling overall is not associated with inequality in achievement (OECD, 2007), but other authors later pointed out that it was important to take into account that in some countries, privately managed schools are privately funded, whereas in others, they are

publicly funded (Alegre & Ferrer, 2010). When these two types are separated, private funding of schools is associated with greater SES gaps, whereas private management is associated with weaker SES gaps (Ammermüller, 2005; Schütz et al., 2007, 2008).

Tracking and private schooling are both theorized to increase socioeconomic achievement gaps because they segregate students of different social class backgrounds. Indeed, a number of comparative studies have found that tracking is associated with socioeconomic segregation between schools (Causa & Chapuis, 2009; Demeuse & Baye, 2008; Jenkins, Micklewright, & Schnepf, 2006; Willms, 2010). Private schooling may also be associated with greater segregation, although again this appears to depend on whether funding is public or private (Alegre & Ferrer, 2010; Jenkins et al., 2006). More generally, when schools are allowed to practice selective admissions, they tend to be more segregated (Alegre & Ferrer, 2010; Demeuse & Baye, 2008; Field et al., 2007; Jenkins et al., 2006; West, 2006), although evidence on the relationship between greater school choice by parents and segregation is mixed (Alegre & Ferrer, 2010; Demeuse & Baye, 2008; Field et al., 2007; Gorard & Smith, 2004). Another country factor that could contribute to school segregation is residential segregation, which is far higher in the United States than in Europe (Musterd, 2005). Perhaps the most general formulation of the idea of the effects of differentiation on inequality in achievement would be that greater socioeconomic segregation in a school system results in greater socioeconomic achievement gaps.

Standardization. The second dimension used to describe school systems is standardization. The hypothesis here is that standardization of national curriculum and exams should reduce SES achievement gaps, whereas decentralization of the school system should increase SES gaps. Furthermore, disparities in resources, such as those created by regional or local funding systems (Field et al., 2007) or variation in teacher quality (Chiu & Khoo, 2005), could disadvantage low-SES students. Most recent research on the impact of standardization on SES achievement gaps defines standardization either in terms of the presence of curriculum-based external exit exams and/or in terms of school autonomy over curricular and/or budgetary decision making. Curriculum-based external exit exams for secondary schools are defined by Bishop (1997) as exams that (a) have real consequences for students; (b) are defined relative to an external, centralized standard; (c) are tied to course curricula; (d) signal multiple levels of achievement, rather than merely pass/fail; and (e) are taken by almost all secondary school students. On the basis of these criteria, he classified countries into those with or without centralized exams in secondary school. Research on school autonomy often uses information from PISA principal questionnaires reporting school control

over curriculum, assessment, budgetary allocation, and hiring of teachers. Findings on standardization are mixed, depending on the country, time period, and exact measure of standardization used. Greater school autonomy appears to be associated with greater SES gaps, as expected (Horn, 2009), but this may be only for autonomy in hiring teachers and establishing salaries and not for curriculum or budget (Schütz et al., 2007). Regarding centralized exams, Park (2008) found that, as part of an index along with a standardized national curriculum and textbooks, exams were associated with smaller SES gaps, but other authors found no significant association between centralized exams and SES gaps (Horn, 2009; Wößmann, 2005), and some even found a positive association between centralized exams and greater SES gaps (Bol et al., 2014; Muller & Schiller, 2000; Schütz et al., 2007). Stevenson and Baker (1991) found that a standardized national curriculum was related to weaker effects of student and teacher characteristics on implemented curriculum, but Westbury and Hsu (1996) argued that this was actually conditional on between- and within-school tracking, which illustrates the complex, interdependent relationship between differentiation and standardization. Two sets of authors have found evidence for this proposed interaction effect between differentiation and standardization. Comparing the United States and Israel, Ayalon and Gamoran (2000) found that in the United States, greater within-school differentiation (math tracking) was associated with more inequality in achievement, whereas in Israel, which has a standardized national exam system, the opposite was true. For 36 countries participating in PISA 2009, Bol et al. (2014) found that the positive association between tracking and SES achievement gaps was weaker in systems with centralized exams.

Aside from differentiation and standardization, systems with a higher level of quality overall could also equalize opportunity by exposing even-lower-SES students to a high minimum standard of education. Educational expenditure is not associated with SES gaps (Schütz et al., 2008), nor are average length of school day and official school starting age (Schütz et al., 2008), but higher average instructional time is associated with smaller SES gaps (Ammermüller, 2005). Higher participation in childcare or preschool is associated with smaller SES gaps (Causa & Chapuis, 2009; Field et al., 2007; Schütz et al., 2008). Average educational quality could also be related to country level of wealth. Heyneman and Loxley (1983) found that the SES effect on achievement is stronger in wealthier countries, whereas the school quality effect is stronger in developing countries; but in a later study, Baker, Goesling, and LeTendre (2002) found that the SES effect has grown stronger even in developing countries as school resources have become more standardized worldwide. Schütz et al. (2008) found no association between gross national income per capita and SES gaps.

SES Gaps Versus Income Gaps

All of the literature described above focuses on the relationship between achievement and SES, as measured by parent education, parent occupation, or home educational possessions, such as books. Our focus in the present study, in contrast, is on a relationship that has received little attention in international education research: the relationship between academic achievement and family income. Given that income and parental education differ in their relationships to achievement in the United States (and these relationships exhibit different trends over time, as shown by Reardon, 2011), it is not clear that previously studied cross-national patterns of the SES (parent education, occupation, and possessions)–achievement gradient are identical to cross-national patterns of the income–achievement gradient. They may differ because the correlations between household income and other dimensions of SES differ across countries or because the independent effects of income relative to parental education differ among countries. Indeed, there is some evidence that income and wealth both operate differently in the United States than in many other countries. Compared to most European countries, the United States (and to a lesser extent the United Kingdom) have higher educational mobility but lower income mobility, most likely due to the strong associations between educational attainment and income in the United States and United Kingdom (Breen & Jonsson, 2005). In a sensitivity analysis of various constructions of the PISA ESCS index, Nonoyama-Tarumi (2008) found that home wealth-related possessions (dishwasher, cars, cellular phones) have additional predictive power for achievement over and above parent education, occupation, and home cultural possessions only in the United States and a very small set of other countries (all located in Latin America and Asia).

If the relationships between country characteristics and income achievement gaps differ from the relationships between those characteristics and SES achievement gaps, it points to the importance of material, and not only cultural, resources. Private schooling is perhaps the clearest example of the importance of material resources if private schools charge tuition that low-income families cannot afford. Material resources may also be particularly important in systems with decentralized funding, where schools in low-income regions or neighborhoods tend to have inadequate resources. Income segregation between schools and policies that facilitate segregation, such as tracking and school choice, can also allow for differentiation of educational resources between high- and low-income schools.

Data and Sample of Countries

Sample of Countries

We use data from the 2001 PIRLS and the 2006, 2009, and 2012 PISA studies. PIRLS tested fourth-grade students in reading and administered a background survey to their

parents. PISA tested 15-year-old students in reading, math, and science, and a small set of participating countries administered parent surveys. Not all countries administered the household income item in the parent surveys. Income responses are available for 22 countries for PIRLS 2001, 15 countries for PISA 2006, 14 countries for PISA 2009, and 10 countries for PISA 2012. We limit our sample to wealthy OECD member countries (defined as those with GDPs per capita of at least \$20,000).¹ This is consistent with the prior research on socioeconomic achievement gaps reviewed above, which has typically focused on developed countries. Additionally, data for country characteristics used in models were unavailable for many non-OECD members, as many of them were obtained from OECD publications. Furthermore, household income data were often low quality in these countries, leading to unreliable estimates of income achievement gaps. We also excluded two countries that met other sample criteria; we excluded the Czech Republic because its income achievement gaps were estimated with very large standard errors, and we excluded Chile because, as a recent addition to the OECD, it did not have available data for centralized exams.² About 4,000 students were sampled in each PIRLS country and about 4,500 in each PISA country, and after exclusions, this resulted in a total sample size of 48,074 students in 12 countries for PIRLS 2001,³ 38,434 students in eight countries for PISA 2006,⁴ 36,355 students in seven countries for PISA 2009,⁵ and 25,443 students in five countries for PISA 2012.⁶ Table 1 lists the countries in our sample.

In both PISA and PIRLS, parents reported annual household income before taxes in six categories based on country currencies. In PISA, the categories were equivalent to (a) less than 0.5 times median income, (b) 0.5 times median or more but less than 0.75 times median income, (c) 0.75 times median or more but less than median income, (d) median or more but less than 1.25 times median income, (e) 1.25 times median or more but less than 1.5 times median income, and (f) 1.5 times median income or more.⁷ In PIRLS, the categories were not explicitly defined in terms of national medians; category income ranges in national currency are reported in a supplement to the *PIRLS 2001 User Guide* (Gonzalez & Kennedy, 2003). In many countries, a moderately large amount of income data was missing because parents either did not complete the background survey or did not respond to the household income item. The percentage of missing income data for each country is reported in Table 1. We imputed missing data for each country using multiple imputation, using the iterated chained equations method. The results of all analyses reported below use imputed data, and all standard errors reflect uncertainty due to imputation.

We examine PIRLS reading scores and PISA reading, math, and science achievement scores. PIRLS reading score was based on performance on an International Association for the Evaluation of Educational Achievement–developed

assessment of reading literacy that requires students to retrieve information, make inferences, interpret, and evaluate texts that are literary or informational (Campbell, Kelly, Mullis, Martin, & Sainsbury, 2001). PISA reading score was based on performance on an OECD–developed assessment of reading literacy that requires students to retrieve information and interpret, reflect on, and evaluate texts that are literary, expository, and practical (e.g., charts, forms, and advertisements; OECD, 2006, 2009b, 2013). PISA 2009 and 2012 also included electronic texts (OECD, 2009b, 2013). PISA math score was based on performance on an OECD–developed assessment of mathematical literacy in the areas of quantity, space and shape, change and relationships, and uncertainty (OECD, 2006, 2009b, 2013). PISA science score was based on performance on an OECD–developed assessment of science literacy in the areas of physical systems, living systems, earth and space systems, and technology systems (OECD, 2006, 2009b, 2013).

Because the United States did not participate in the parent income survey in PISA, we estimate the income achievement gap using data from the Early Childhood Longitudinal Study–Kindergarten Cohort (ECLS-K). ECLS-K tested a nationally representative sample of U.S. third graders in reading and math in 2002 and again when they were eighth graders in 2007. The ECLS-K sample is therefore roughly the same birth cohort as the PIRLS and PISA 2006 and 2009 cohorts.

Measuring Income Achievement Gaps

We estimate the income achievement gap using the methods described in Reardon (2011; see Table A2 in the appendix). In brief, this method requires (1) standardizing test scores within each country, (2) estimating the mean standardized test score of students in each family income category within a country, (3) assigning each student an income percentile that corresponds to the middle percentile of his or her income category, (4) fitting a cubic function via weighted least squares (with an adjustment to account for nonlinearity) that describes the association between test scores and family income percentile, (5) using the fitted curve to estimate the difference in average test scores between students at the 90th and 10th percentiles of the income distribution, (6) using the estimated variance-covariance matrix of the estimates to compute a standard error for this difference, and (7) adjusting the estimated gaps to account for measurement error in the test scores and measurement error in family income.⁸ These methods allow us to compute the average difference in test scores between students at the 90th and 10th percentiles of the household income distribution within their respective country and cohort (the 90/10 income achievement gap) as well as the standard error of this estimated gap. We also compute the 90/50 and 50/10 income achievement gaps (by repeating Steps 5 through 7 above

TABLE 1
Income Data Availability in Sample Countries, PIRLS 2001, PISA 2006, 2009, and 2012

Country	Sample size	Missing (%)	Percentage in bottom income category	Percentage in top income category
PIRLS 2001				
Canada (Ontario)	4,295	25.24	6.99	46.05
Canada (Quebec)	3,958	26.31	9.37	36.37
Czech Republic ^a	3,022	40.32	45.89	1.36
England	3,156	51.20	15.06	10.10
Germany	7,633	26.25	17.99	10.06
Greece	2,494	19.13	17.02	13.29
Iceland	3,676	27.04	6.64	23.38
Netherlands	4,112	65.41	13.00	14.09
New Zealand	2,488	25.47	7.38	36.12
Norway	3,459	13.31	7.11	21.65
Slovak Republic	3,807	30.92	27.63	4.62
Slovenia	2,952	11.13	29.41	6.13
Sweden	6,044	12.31	9.62	15.37
PISA 2006				
Denmark	4,532	41.64	3.40	35.96
Germany	4,891	38.32	18.41	13.88
Iceland	3,789	45.56	5.15	29.55
Korea	5,176	2.63	17.53	16.82
Luxembourg	4,567	42.75	11.95	17.60
New Zealand	4,823	38.05	11.32	25.52
Poland	5,547	15.13	20.52	12.34
Portugal	5,109	24.19	15.88	10.95
PISA 2009				
Denmark	5,924	36.77	3.61	42.55
Germany	4,979	53.41	16.43	16.67
Hungary	4,605	19.36	23.84	8.98
Korea	4,989	2.69	20.63	18.70
New Zealand	4,643	33.10	17.83	24.62
Poland	4,917	16.38	9.98	10.06
Portugal	6,298	32.47	27.76	18.17
PISA 2012				
Belgium (Flanders)	4,877	43.88	12.50	20.89
Germany	5,001	57.97	13.21	21.84
Hungary	4,810	22.10	23.67	10.13
Korea	5,033	1.99	16.98	24.01
Portugal	5,722	25.48	32.27	16.98

Note. PIRLS = Progress in International Reading Literacy Study; PISA = Programme for International Student Assessment.

^aThe Czech Republic was excluded for unreliable income data. Its 90/10 income achievement gap was estimated with a very large standard error, resulting in a confidence interval for the gap ranging from 2 standard deviations to less than 0.

using the appropriate percentiles) to allow us to investigate whether income is more strongly related to achievement at the high or low end of the income distribution. We compute these gap estimates and standard errors for each test subject in each country and year in which household income data are available (19 countries, 32 country-by-year estimates, 72 country-by-year-by-subject estimates. Adding the U.S. data,

this becomes 20 countries, 34 country-by-year estimates, 75 country-by-year-by-subject estimates).

Measures of National Characteristics

We construct a set of four primary indices of national characteristics: (a) poverty/inequality, (b) social welfare

policies, (c) parental support, and (d) educational differentiation; and obtain a fifth measure, for curricular standardization, from the literature.

Poverty/inequality index. This index measures the level of socioeconomic inequality in a country. It is constructed as the first principal component of the Gini index of income inequality, the child poverty rate, the level of between-school income segregation, the low-birthweight rate, and the adolescent fertility rate (eigenvalue 2.90; explains 58% of the variance of the five items). We initially considered including the income segregation measure in the index of educational differentiation (described below), but it loaded much more strongly on the poverty/inequality dimension, likely because it is correlated with residential segregation, which is driven in part by income inequality (Reardon & Bischoff, 2011). The index is standardized to have a mean of 0 and a standard deviation of 1 among the 20 countries in our sample.

We obtain Gini coefficients for most countries from the Luxembourg Income Study (2009), supplemented by the World Bank Factbook (World Bank, 2012) and the OECD (2011a), and we interpolated missing years and then averaged over the years from birth to the test year. The percentage of children under 18 in poverty was measured in 1995 and 2008 and was obtained from the OECD Family Database (OECD, 2011b); we interpolated between these years and then averaged over the years up until each cohort was tested. The percentage of low-birthweight births and the adolescent fertility rate (births per 1,000 women ages 15 to 19) were obtained for the birth year of each cohort from the World Bank Factbook (World Bank, 2012). We compute income segregation as the segregation between students from families with above-median income (within their country) and students from families with below-median income. To compute this, we use methods similar to those described in Reardon and Bischoff (2011), estimating the segregation at each ordinal income category threshold and then interpolating (using a quadratic function fitted via weighted least squares) to estimate the segregation between above- and below-median income students.⁹ We compute U.S. income segregation for primary school from ECLS-K data from the kindergarten wave, as students may have changed schools by third grade, when we estimate income achievement gaps. We compute U.S. secondary school segregation from the Education Longitudinal Study, which has income data for a nationally representative sample of 10th-grade students in 2002.

Social welfare policy index. This index measures social welfare spending in a country. It is constructed as the first principal component of public health expenditures, public spending on family benefits in cash, public spending on family benefits in services, and preprimary school

enrollment rates (eigenvalue 2.06; explains 52% of the variance of the four items). Public spending on family benefits in cash and in services consisted of support exclusively for families (e.g., child payments and allowances, parental leave, and childcare) in 2007 as percentages of GDP, obtained from the OECD Family Database (OECD, 2011b). The percentage of the age cohort enrolled in preprimary education (gross) was averaged for the years from birth to age 5 for each cohort and was obtained from the World Bank Factbook (World Bank, 2012). Public health expenditure as a percentage of total health expenditure was averaged over the years from birth to testing for each cohort and was obtained from the World Bank Factbook (World Bank, 2012). We considered including per-pupil expenditures on primary and secondary school in the index, but this item was not correlated with the other measures or the constructed index ($r = -0.1$). The index is standardized to have a mean of 0 and a standard deviation of 1 among the 20 countries in our sample.

Parental support index. This index measures (albeit crudely) the extent to which a country requires paid parental leave following the birth of a child. We would like a more general measure of early childhood support—an index that would measure social policies that support families with young children and that provide educational opportunities for young children—reasoning that countries with stronger early childhood support policies would have smaller income achievement gaps because more children would have parents at home during infancy and more educational opportunities prior to enrolling in primary school. However, the only policy measures we were able to obtain were measures of parental leave policy. This index is constructed as the first principal component of the maximum number of weeks of leave for mothers (maternity leave plus parental or prolonged leave), the number of weeks of paternity leave for fathers (including parental or prolonged leave), the paid-leave full-rate equivalent pay for mothers (weeks of leave multiplied by rate of wage replacement), and the paid-leave full-rate equivalent for fathers (eigenvalue 2.27; explains 57% of the variance of the four items). All parental leave measures used in the parental support index are from 2007 or 2008 and were obtained from the OECD Family Database (OECD, 2011b). We considered including preprimary enrollment rates in this factor, but that item loaded much more heavily on the social welfare policy index and so was included in that factor. The index is standardized to have a mean of 0 and a standard deviation of 1 among the 20 countries in our sample.

Educational differentiation index. This index measures the extent to which the educational system is structured to provide highly differentiated learning environments for students. We reason that a more differentiated system—one with high levels of tracking and a large private school

TABLE 2

Correlation Matrix of Constructed Educational and Social Indexes

Variable	Educational differentiation index	Centralized exams	Parental support index	Social welfare index	Poverty/inequality index
Educational differentiation index	—				
Centralized exams	0.08	—			
Parental support index	0.10	0.21	—		
Social welfare policy index	0.36	0.46*	0.25	—	
Poverty/inequality index	-0.25	-0.38	-0.48 *	-0.58**	—

Note. Degrees of freedom for all correlations = 18.

* $p < .05$. ** $p < .01$. *** $p < .001$.

sector—may lead to greater stratification of students both between and within schools. If this stratification is associated with family socioeconomic background (as it generally is), more differentiation may contribute to widening academic achievement disparities.

The index is constructed as the first principal component of the proportion of students in private schools, the proportion of secondary school students in vocational tracks, the number of distinct tracks in secondary school, and the age at which students are first tracked (with lower ages implying greater differentiation; eigenvalue 2.50; explains 63% of the variance of the four items). Age of track selection and the number of tracks were taken from the PISA 2009 international report (OECD, 2010b). The percentage of secondary school students in the vocational track and the percentage of students enrolled in privately managed schools were obtained from the World Bank Factbook (World Bank, 2012). We used values for the year in which students were tested; for private schooling, we used primary school values for PIRLS and secondary school values for PISA. We considered including the between-school income segregation in this factor, but it loaded weakly on this factor and much more strongly with the poverty/inequality items above. The index is standardized to have a mean of 0 and a standard deviation of 1 among the 20 countries in our sample.

Curricular standardization. We use the existence of curriculum-based external exit exams in secondary school as a proxy for curricular standardization. This measure was first created by Bishop (1997) for the 39 countries that participated in TIMSS 1995 using comparative education studies and encyclopedias and interviews with embassy personnel to classify systems according to his five criteria that we summarize in the previous section. The measure is primarily dichotomous, as he classified most countries as either having (1) or not having (0) centralized exams, but a small number of countries with state or provincially based education systems were assigned decimal values reflecting the proportion of students in regions subject to centralized

exams. This measure has since been updated for 2003 data by Wößmann, Lüdemann, Schütz, and West (2009) and for 2009 data by Bol et al. (2014). We use the values from Wößmann et al. (2009), which pertain separately to secondary school math exams and science exams, and take the value for Slovenia (which was missing from Wößmann et al., 2009) from Bol et al. All other values are identical to those from Bol et al. for the countries in our sample. Following prior literature, to create a country-level measure of curricular standardization across all subjects and grade levels, we take the average of the math exam and science exam indicators and apply these values at both the primary and secondary levels. We standardize this measure to have a mean of 0 and a standard deviation of 1 among the 20 countries in our sample.

Table 2 reports the bivariate correlations among the four indices we constructed and the centralized exams measure. Note that the educational differentiation index and centralized exams are only weakly correlated with one another and with the parental support index but are moderately positively correlated with the social welfare policy index and moderately negatively correlated with the poverty/inequality index. The parental support index is moderately correlated with the social welfare policy index and with the poverty/inequality index. The social welfare policy and poverty/inequality indices are moderately negatively correlated with one another. The fact that most of these correlations are relatively low suggests that multicollinearity is not a concern when including them together in regression models.

Analysis

We are interested first in describing the magnitude and variation in the size of the income achievement gap across countries. Because the gaps are estimated with error, we use a random-effect model to estimate the mean and variance of the true gaps. Specifically, we fit the model

$$\hat{G}_{ci} = \mu + v_c + u_{ci} + \varepsilon_{ci}, \quad (1)$$

$$v_c \sim N(0, \tau); u_{ci} \sim N(0, \sigma^2); \varepsilon_{ci} \sim N(0, \omega_{ci}),$$

where \hat{G}_{ci} is the estimated gap in country c in study-year-subject i (we have multiple observations within a country because some countries are observed in multiple years and because multiple tests are given each year in the PISA countries), μ is the mean of the true gaps, τ is the between-country variance of the true gaps, σ^2 is the true within-country variance of the gaps (between studies, subjects, and years), and $\omega_{ci} = \text{var}(\hat{G}_{ci})$ is the sampling variance of \hat{G}_{ci} , which we compute using the squared standard error of \hat{G}_{ci} . Fitting this model via maximum likelihood produces estimates of μ and τ , the mean and variance of the income achievement gaps across countries.

We are next interested in the association between a set of country-level characteristics and the size of the income achievement gap. For this we fit models of the form

$$\begin{aligned} \hat{G}_{ci} &= \mu + \bar{X}_c \Gamma + (X_{ci} - \bar{X}_c) \mathbf{B} + \Lambda_i \\ &\quad + v_c + u_{ci} + \varepsilon_{ci}, \\ v_c &\sim N(0, \tau); u_{ci} \sim N(0, \sigma^2); \varepsilon_{ci} \sim N(0, \omega_{ci}), \end{aligned} \quad (2)$$

where X_{ci} is a vector of country-level covariates in study-year i , \bar{X}_c is the average of vector X_{ci} within country c , and Λ_i is a vector of dummy variables indicating study, year, and test subject. The key parameters of interest in these models are the vector of coefficients Γ , which describe the association between the country-level covariates and the average income achievement gap. In addition, we are interested in the value of τ , which describes the between-country variance of the income achievement gaps, conditional on the covariates included in the vector \bar{X} . We fit these models using the HLM software program (Raudenbush, Bryk, & Congdon, 2004).

The relatively small number of countries (20 countries; 75 total observations) limits the number of covariates that can be included in the vector \bar{X}_c ; as a result, we are parsimonious in our model specifications, generally limiting the models to fewer than five country-level covariates. We begin by fitting models using only the five national variables (educational differentiation, centralized exams, inequality/poverty, social welfare policy, and parental support). We then drop variables that are not significantly associated with achievement gaps in this model and refine the model to examine which elements of the indices are most consistently associated with achievement gaps. It is worth reiterating that this is strictly an exploratory analysis, designed to determine what national characteristics are most strongly associated with achievement gaps. Our estimates here should not be interpreted causally.

Results

Comparing Income Achievement Gaps Across Countries

Figure 1 shows the estimated income reading 90/10 achievement gaps for each of the countries in our sample, and Table A1 in the appendix reports estimated 90/10, 90/50, and 50/10 gaps for all subjects.

The 90/10 income achievement gaps for reading average around 1.0 standard deviation. Notably, the size of the gaps varies substantially across the countries in our sample, ranging from 1.39 standard deviations in Luxembourg in 2006 to 0.45 standard deviations in Iceland in 2006. With gaps of 1.28 standard deviations in elementary school and 1.36 standard deviations in high school, the United States ranks among the highest in the distribution of the size of income achievement gaps across countries. Gaps in the United States are comparable in size to those of Luxembourg, Portugal, Belgium (Flanders), and Hungary. The income achievement gaps also vary somewhat by test subject. Although the correlation between gaps in different test subjects is over 0.9 in every wave of PISA, average gaps are roughly 5% to 10% larger for the math tests than the reading and science tests, a finding consistent with the evidence on the size of the math and reading income achievement gaps in the United States (Reardon, 2011). The mean 90/50 income achievement gaps average around 0.58 standard deviations, while the mean 50/10 gaps average around 0.47 standard deviations. The slightly larger 90/50 gaps are also consistent with Reardon's (2011) finding that the 90/50 gaps in the United States are approximately 10% larger than the 50/10 gaps. Table A2 in the appendix reports the estimated means and standard deviations of the income achievement gaps across the countries in our sample.

The estimated gaps for PIRLS tend to be somewhat smaller than those for PISA for the 90/10, 90/50, and 50/10 gaps. However, this may be an artifact of our samples of countries for each study. There are very few overlapping countries across PIRLS and PISA; the only three countries in our sample for which we have data for both studies are Germany, Iceland, and New Zealand. In these three countries, the relative size of the PIRLS and PISA income gaps are inconsistent: In Germany, the PISA 2009 gap is larger than the gaps from PIRLS, PISA 2006, and PISA 2012; in Iceland, the gap is smaller in PISA 2006 than in PIRLS; and in New Zealand, the gaps are the roughly the same in PIRLS, PISA 2006, and PISA 2009. In general, there is no evidence of a strong trend in the size of the gaps across the three countries for which we have both PIRLS and PISA estimates.

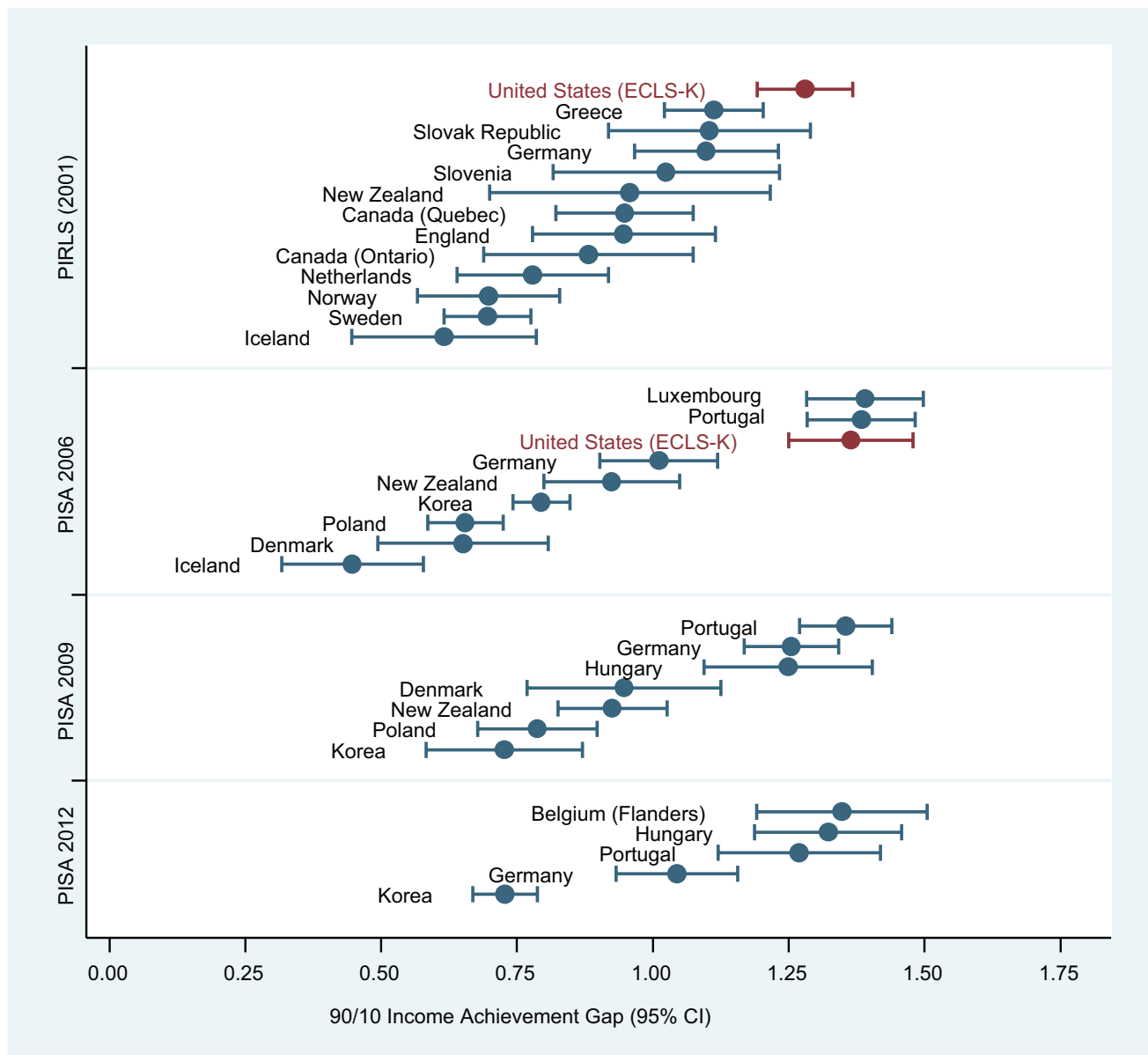


FIGURE 1. Estimated 90/10 income achievement gaps, reading, Progress in International Reading Literacy Study and Programme for International Student Assessment.

Bivariate Associations Between Income Gaps and Country Characteristics

Before turning to our model-based estimates, we present a set of descriptive figures showing the bivariate associations between income achievement gaps and country characteristics. Figures 2 through 6 plot the 90/10 income achievement gaps in each country against the four indices we constructed—poverty/inequality, social welfare policy, parental support, and educational differentiation—as well as our measure of curricular standardization based on centralized exams. To make each figure, we pooled the estimated 90/10 gaps in all available subjects from PIRLS and

PISA within each country to construct a single precision-weighted estimate of the gap in each country. For those indices that are time varying, we averaged the index across years to get an average index for each country. In each figure, the size of the circle indicates the precision of the 90/10 gap estimate, with the larger circles indicating the most precisely estimated gaps.

Each of Figures 2 to 6 includes three fitted lines. The thick solid line is the precision-weighted regression line through the 20 data points. Because the United States has extreme values of some of the indices (particularly, the poverty/inequality and social welfare policy indices) and because the achievement gap data for the United States come

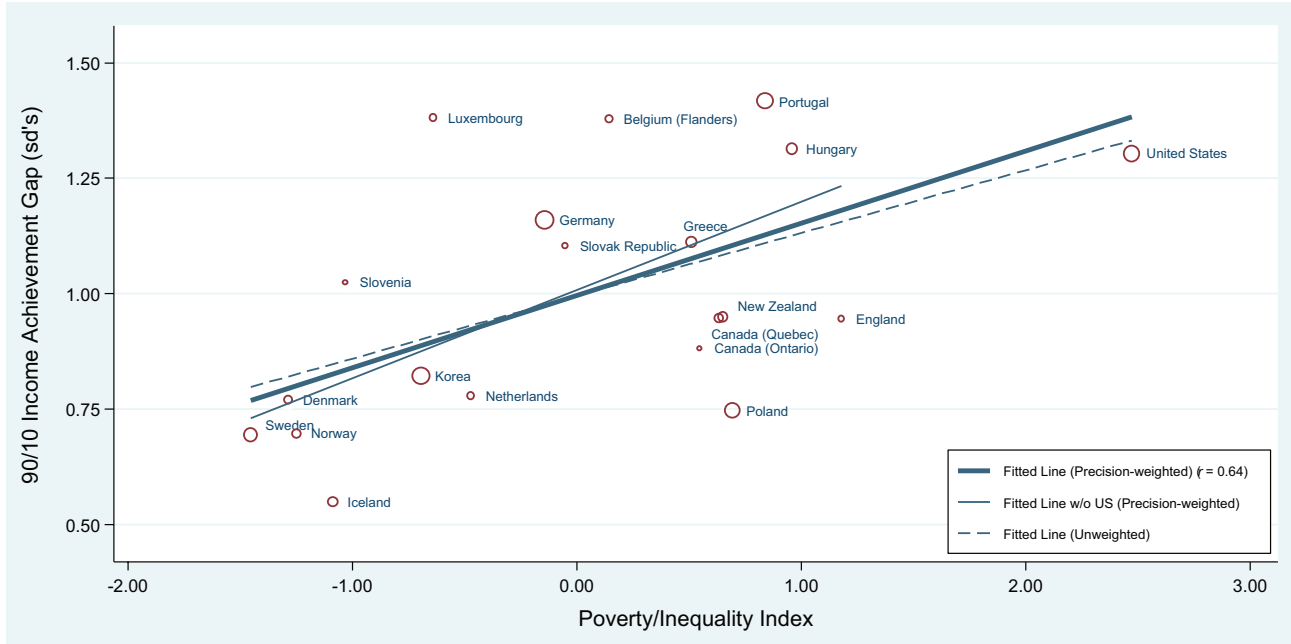


FIGURE 2. Association between income achievement gap and poverty/inequality index, wealthy Organisation for Economic Co-operation and Development countries, 2001 to 2012 (pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data).
 Note. The size of each circle indicates the precision of the 90/10 gap estimate, with the larger circles indicating the most precisely estimated gaps.

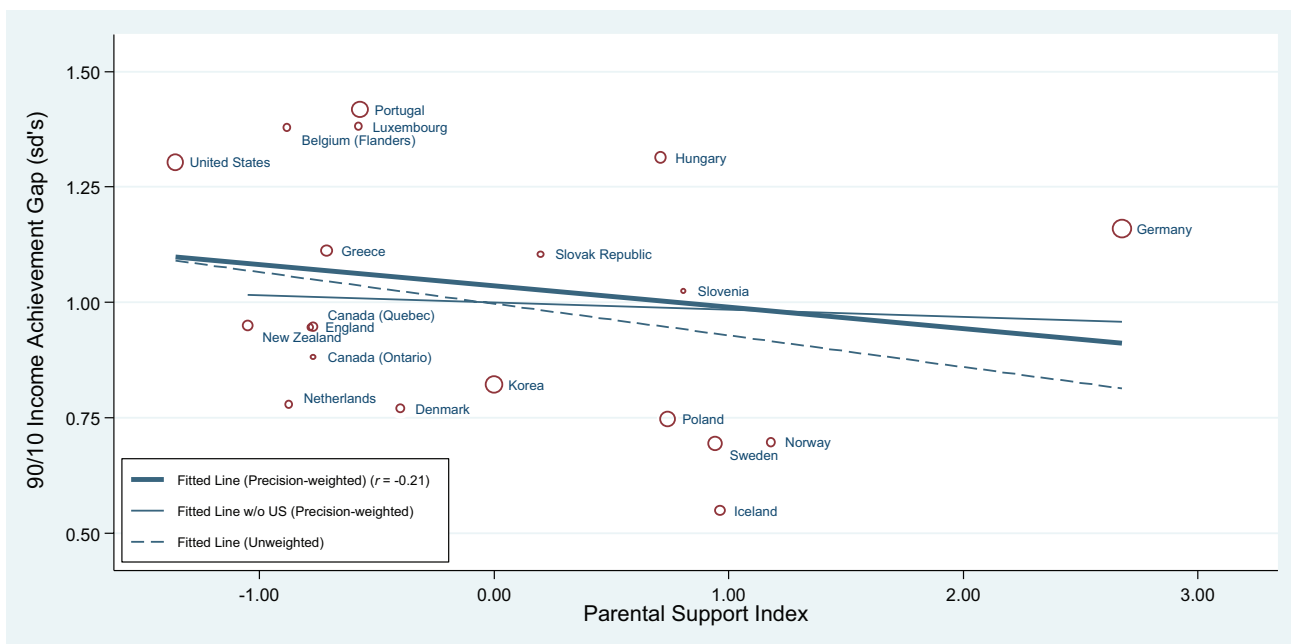


FIGURE 3. Association between income achievement gap and parental support index, wealthy Organisation for Economic Co-operation and Development countries, 2001 to 2012 (pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data).
 Note. The size of each circle indicates the precision of the 90/10 gap estimate, with the larger circles indicating the most precisely estimated gaps.

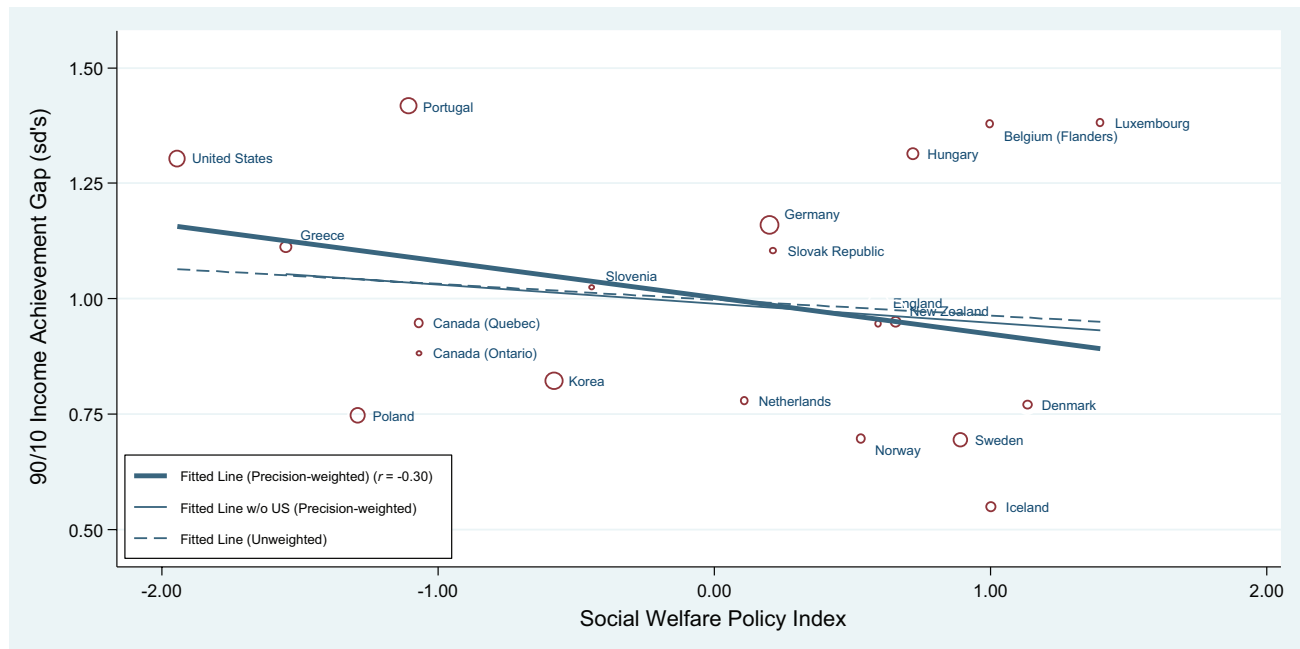


FIGURE 4. Association between income achievement gap and social welfare index, wealthy Organisation for Economic Co-operation and Development countries, 2001 to 2012 (pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data).
 Note. The size of each circle indicates the precision of the 90/10 gap estimate, with the larger circles indicating the most precisely estimated gaps.

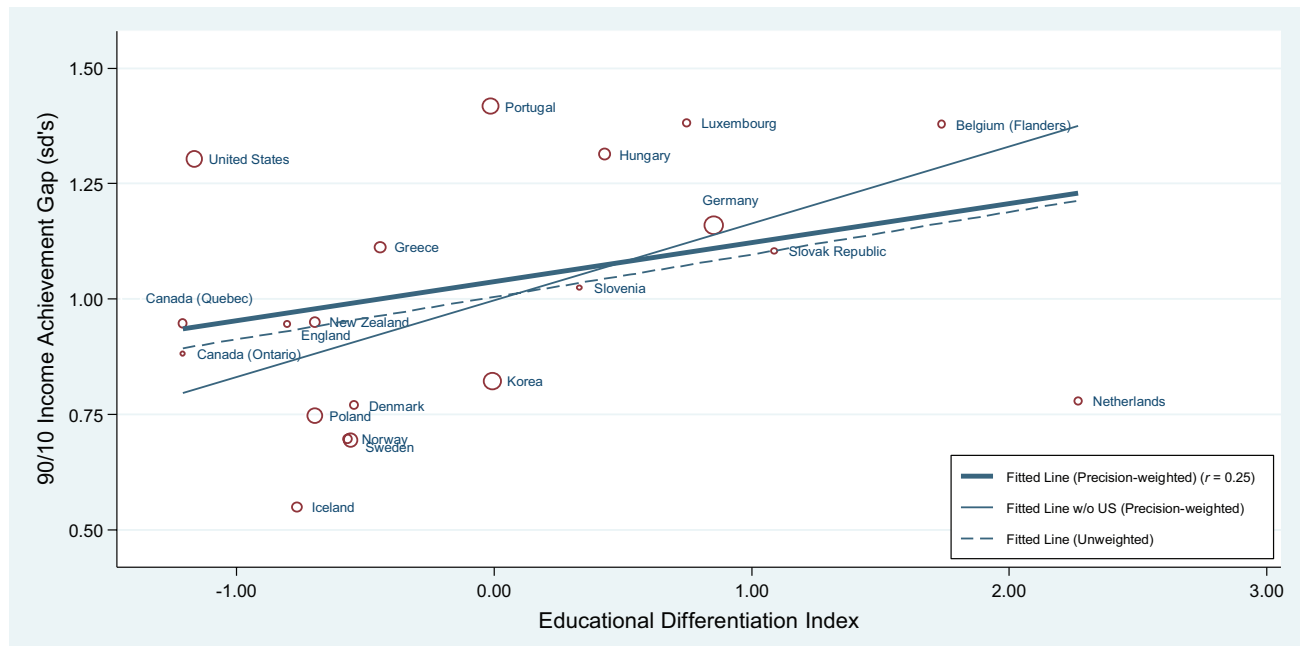


FIGURE 5. Association between income achievement gap and differentiation index, wealthy Organisation for Economic Co-operation and Development countries, 2001 to 2012 (pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data).
 Note. The size of each circle indicates the precision of the 90/10 gap estimate, with the larger circles indicating the most precisely estimated gaps.

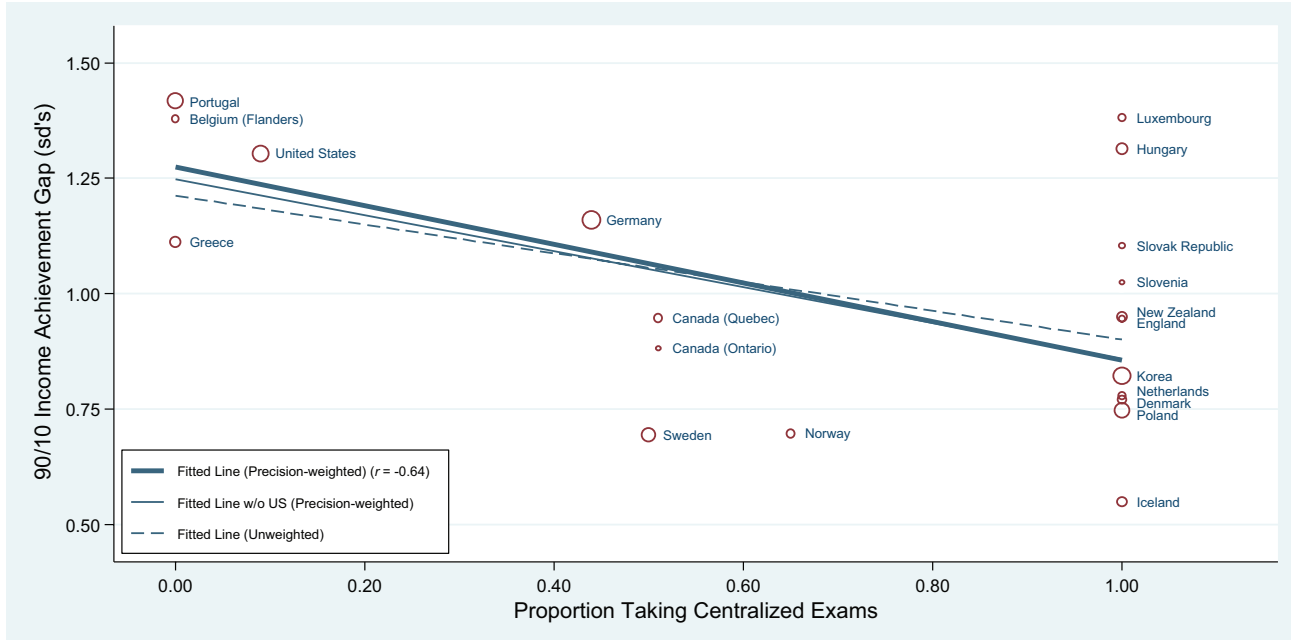


FIGURE 6. Association between income achievement gap and centralized exams, wealthy Organisation for Economic Co-operation and Development countries, 2001 to 2012 (pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data).

Note. The size of each circle indicates the precision of the 90/10 gap estimate, with the larger circles indicating the most precisely estimated gaps.

from the ECLS-K data rather than PISA and PIRLS, we also fit precision-weighted regression lines that exclude the United States in order to examine whether the fitted lines are heavily influenced by the presence of the United States in the sample. The estimates based on these regressions are shown as thinner solid lines in each figure. Finally, we also include a fitted line based on an unweighted regression to examine the sensitivity of our estimates to the precision weights. In most cases, the fitted lines are relatively insensitive to whether the United States is in the sample and to whether we use precision weights or not.

In general, all of the associations with the country characteristics are in the expected directions. The 90/10 gaps are most strongly associated with the poverty/inequality index, $r(18) = .64, p = .002$, and centralized exams, $r(18) = -.64, p = .002$, and are least strongly associated with the social welfare policy index, $r(18) = -.30, p = .19$; the educational differentiation index, $r(18) = .25, p = .28$; and the parental support index, $r(18) = -.21, p = .37$. Figures for the five separate components of the poverty/inequality index (income inequality, income segregation, child poverty rate, low-birthweight rate, and teenage birth rate) are presented in Appendix Tables A1 through A5.

Multivariate Associations Between Income Gaps and National Characteristics

Table 3 presents the results of fitting various versions of Equation (2) above to predict the 90/10 income achievement

gap. We begin with a null model (Model 0), which includes only a set of dummy variables indicating which study, year, and subject was tested.¹⁰ In this model—which, by construction, explains none of the between-country variance in gaps—the between-country variance of achievement gaps in our sample is .065 (corresponding to a standard deviation of 0.25). In Model 1, we include the four indices and centralized exams in the model. Recall that the indices are standardized, so the coefficients can be interpreted as the difference in the size of the 90/10 income achievement gap associated with a 1-standard-deviation difference in the index, conditional on the other variables in the model.

Income achievement gaps are larger in countries with higher degrees of educational differentiation and smaller in countries with centralized exams, as predicted. The poverty/inequality index is very strongly associated with the 90/10 achievement gap, whereas the social welfare policy and parental support indices are not significantly associated with achievement gaps, conditional on the poverty/inequality level. Based on this finding, we drop the social welfare policy and parental support indices from subsequent models. Model 2 then replaces the poverty/inequality index with three of its component items (Gini index, school income segregation, and the child poverty rate).¹¹ For ease of interpretation, we divide these variables by their standard deviation in the sample of 20 countries so that the coefficients can be compared to those of the four indices and centralized exams. With all three component variables in the model, none is significantly associated with achievement gaps. Because child poverty and the

TABLE 3

Estimated Multivariate Associations of Country Characteristics With 90/10 Income Achievement Gap

Variable	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Educational differentiation index		0.104** (0.040)	0.130** (0.045)	0.151** (0.048)	0.112** (0.041)	0.137** (0.043)
Centralized exams		-0.092* (0.039)	-0.086* (0.042)	-0.098* (0.045)	-0.093* (0.042)	-0.107* (0.042)
Poverty/inequality index		0.144** (0.047)				
Parental support index		-0.019 (0.038)				
Social welfare policy index		0.055 (0.047)				
Income inequality (Gini index)			0.010 (0.074)	0.105* (0.052)		
School income segregation			0.074 (0.048)		0.110* (0.043)	
Child poverty rate			0.061 (0.066)			0.107* (0.044)
Intercept	0.969*** (0.064)	0.919*** (0.042)	0.979*** (0.046)	0.970*** (0.048)	0.983*** (0.046)	0.970*** (0.047)
Within-country residual variance	0.002	0.001	0.002	0.002	0.002	0.002
Between-country residual variance	0.065	0.020	0.022	0.029	0.025	0.027
Variance explained (between countries)		0.698	0.663	0.562	0.622	0.586
<i>N</i> (observations)	75	75	75	75	75	75
<i>N</i> (countries)	20	20	20	20	20	20

Note. Estimates from random-effects models using pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data. United States included; U.S. data from Early Childhood Longitudinal Study–Kindergarten Cohort. Standard errors in parentheses.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Gini index of income inequality are highly correlated, $r(18) = .80$, $p < .001$, their associations with income achievement gaps are difficult to separate, and including both in the model leads to inflated standard errors on both coefficients. Income segregation is also quite highly correlated with both child poverty and income inequality, $r(18) = .59$, $p = .006$, and $r(18) = .56$, $p = .01$, respectively. In Models 3 through 5, then, we enter each of the three variables separately and find that each of them is significantly associated with the income achievement gap when the others are not in the model. Educational differentiation remains strongly associated with income achievement gaps in all models, and centralized exams remain strongly negatively associated with gaps in all models. Overall, Models 3 through 5 yield results consistent with our expectations. As we predicted, income achievement gaps are larger in countries with higher levels of income inequality and child poverty, in countries with higher levels of income segregation, and in countries with higher levels of educational differentiation and are smaller in countries with centralized exams.¹² Of the five models explored, Model 1 fits the data best, explaining 70% of the variance in achievement gaps across countries.

We fit the same set of models to predict both the 90/50 and 50/10 income achievement gaps (Model 1 for all three

types of gaps is presented in Table 4). In many ways, the results are quite consistent with the models predicting the 90/10 gaps, albeit with several interesting differences. First, educational differentiation is associated with significantly larger 50/10 gaps but not 90/50 gaps, which suggests that low-income students may be more likely than middle- and high-income students to be placed in the low track. Second, centralized exams are associated with significantly smaller 90/50 gaps but not 50/10 gaps, which suggests that standardization may be particularly beneficial to middle-income students (or detrimental to high-income students, although prior literature has mostly found that centralized exams are associated with higher average achievement [Fuchs & Woessmann, 2007], a finding that we do not test here). Third, the poverty/inequality index is associated with significantly larger 50/10 gaps but not 90/50 gaps. We also ran models entering income inequality, income segregation, and child poverty separately (not shown). Note that in these models, we use segregation of high-income students (above the 75th percentile) to predict 90/50 gaps and segregation of low-income students (below the 25th percentile) to predict 50/10 gaps. Both income inequality and child poverty significantly predict 50/10 gaps, whereas only income segregation significantly predicts 90/50 gaps. The differing

TABLE 4

Estimated Multivariate Associations of Country Characteristics With 90/10, 90/50, and 50/10 Income Achievement Gaps

Variable	90/10	90/50	50/10
	Model 1	Model 1	Model 1
Educational differentiation index	0.104** (0.040)	0.036 (0.021)	0.070* (0.032)
Centralized exams	-0.092* (0.039)	-0.080*** (0.019)	-0.014 (0.031)
Poverty/inequality index	0.144** (0.047)	0.039 (0.023)	0.110** (0.037)
Parental support index	-0.019 (0.038)	-0.017 (0.018)	-0.006 (0.029)
Social welfare policy index	0.055 (0.047)	0.019 (0.023)	0.044 (0.038)
Intercept	0.919*** (0.042)	0.525*** (0.024)	0.388*** (0.034)
Within-country residual variance	0.001	0.000	0.000
Between-country residual variance	0.020	0.004	0.012
Variance explained (between countries)	0.698	0.763	0.582
<i>N</i> (observations)	75	75	75
<i>N</i> (countries)	20	20	20

Note. Estimates from random-effects models using pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data. United States included; U.S. data from Early Childhood Longitudinal Study–Kindergarten Cohort. Standard errors in parentheses.

* $p < .05$. ** $p < .01$. *** $p < .001$.

results for child poverty are consistent with what we might expect, as it is plausible that child poverty most strongly affects gaps at the bottom of the income distribution (the 50/10 gap). However, the differing results for income segregation and income inequality were unexpected.

Finally, we ran two additional sets of analyses to confirm the robustness of our results. First, because we are concerned that our results may be influenced by the presence of the United States in the sample (recall that the test scores for the United States are based on ECLS-K rather than PISA and PIRLS and that some of the measures of U.S. characteristics, like school income segregation, are taken from different sources than in the other countries), we reran all the analyses in Tables 3 and 4 excluding the United States. The pattern of results (the magnitude and direction of signs as well as statistical significance) is largely unchanged with the United States excluded from the sample (results not shown). Last, because we are concerned that our results may be influenced by the restricted sample of 20 countries with available income data, we also computed gaps for the full population of OECD member countries. As household income data were unavailable in these additional countries, we computed gaps based on two other SES variables: highest parent educational attainment (available in both PISA and PIRLS) and ESCS (available only in PISA). Multivariate models predicting parent education gaps for the original restricted sample are presented in Appendix Table A3, and models predicting parent education

and ESCS gaps for the full OECD population are presented in Tables A4 and A5 (ESCS models cannot be run for the restricted sample, as ESCS is unavailable in PIRLS). Results differ very little between the restricted sample and the full OECD population, but compared to the models predicting income gaps, the associations with income inequality (Gini coefficient) and educational standardization (centralized exams) are weaker and generally nonsignificant. Thus, we tentatively conclude that the restricted sample does not bias our results extremely and that differences between our results for income and those previously published using other SES measures may be due to substantive differences in the relationship between family income versus other family socioeconomic characteristics and student achievement. These findings are discussed in further detail below.

Discussion

There is considerable variation across our sample of OECD countries in the strength of the association between family income and academic achievement, but the size of the gap is only modestly associated with national income inequality. The achievement gap between high- and low-income students is, on average, about 1.0 standard deviation, but this varies widely across countries. The income achievement gap in our sample of countries is largest in Portugal, Luxembourg, the United States, Belgium (Flanders), and

Hungary. Some of these countries, notably, the United States and Portugal, have very high levels of income inequality; others, such as Luxembourg and Belgium (Flanders), however, have moderately low levels of income inequality. The countries with the smallest income achievement gaps are Iceland, Norway, Sweden, Poland, and Denmark, most of which are Scandinavian countries with low levels of inequality. Poland, however, has both moderately high income inequality and low income achievement gaps. Our calculations of international income achievement gaps provide some valuable context for the findings on the U.S. income achievement gaps. The U.S. income gap is quite large, particularly at the primary school level, where it is larger than all of those that we estimated for PIRLS.

The second part of our analysis here focuses on investigating whether and how the income achievement gap is associated with a set of country characteristics, including poverty rates and inequality, social welfare policy, parental support policy, educational differentiation, and curricular standardization. Although evidence of cross-national associations between national characteristics and academic achievement gaps in cross-sectional data is certainly not sufficient to identify the extent to which the gaps are caused by the factors we investigate, evidence of associations is suggestive and may help guide theory and future research.

In exploring these associations, we hypothesized that countries with higher income inequality would have larger income achievement gaps. We found a modest positive association between the two in bivariate analyses, and this association remained evident when we included other covariates in a multivariate model. Prior studies have found no clear association between socioeconomic achievement gaps and measures of economic inequality, and our supplementary analyses for SES gaps (reported in Tables A3 through A5) also find no significant association with income inequality. That we find such an association when predicting income achievement gaps suggests that in countries with high income inequality, income may be particularly salient in shaping students' educational opportunities (compared to other measures of family SES, such as parental education and occupational status). This finding is consistent with a story by which income inequality leads to wider income achievement disparities largely through its effects on income segregation and child poverty rates. Our finding that both income inequality and child poverty are more strongly related to 50/10 than to 90/50 gaps points to the negative consequences of poverty for educational opportunity.

One consistent pattern in all of our analyses is that the extent to which students are segregated among schools by income is strongly related to the magnitude of achievement gaps. This finding is evident in our univariate analyses and in all of our multivariate models predicting 90/10 gaps. Moreover, this finding persists even after controlling for the

income inequality, child poverty rates, the differentiation of the schooling system (which includes measures of tracking and private school enrollment rates), and the standardization of the schooling system. Notably, the pattern persists even after the United States is excluded from the sample, an important finding, given that its income segregation measure is very high relative to the other OECD countries.

Although the strong association of income segregation and the income achievement gap is clear, the mechanisms that produce this association are not. Because residential segregation and school segregation are generally correlated, it is not clear how much to think of this association as reflecting school segregation (and inequalities in school quality associated with segregation) or residential segregation (and inequalities in environment and opportunities associated with residential segregation). The finding that, after controlling for other country characteristics, the segregation of high-income students significantly predicts 90/50 gaps, whereas the segregation of low-income students does not significantly predict 50/10 gaps, suggests that 50/10 gaps are better accounted for by poverty, inequality, and educational differentiation.

In addition to examining income inequality per se, we also explored the association of income achievement gaps with other aspects of social inequality, including social welfare policy, parental support policy, and other measures of childhood disadvantage, such as low-birthweight rates and teenage childbearing rates. In general, each of these was associated with achievement gaps in the predicted direction, but most were not strongly associated with achievement gaps after controlling for other factors. We are not confident, however, that the social welfare policy and parental support indices were sufficiently well defined to capture the true relationship of social policies to income achievement gaps. With only 20 countries, a poorly measured construct may not yield particularly reliable estimates of these associations.

Finally, we explored the association of income achievement gaps to two features of national education systems—structural differentiation and curricular standardization. Differentiation was positively associated with income achievement gaps, regardless of what other variables were in the model, suggesting that tracking regimes and private school enrollment may operate to exacerbate income achievement disparities. Our finding that differentiation is more strongly associated with 50/10 gaps than with 90/50 gaps is consistent with the possibility that tracking may be particularly harmful to low-income students. In all of the countries in our sample, the academic track comprises the majority of students rather than an elite few. Thus, if tracking results in a disproportionate share of low-income students being placed in lower-quality schooling, it may exacerbate achievement gaps between low-income students and their middle- and high-income peers.

Curricular standardization, indicated by the use of centralized examinations, was consistently negatively and significantly associated with income achievement gaps, in line with the hypothesis that standardization promotes greater homogeneity of quality across schools. Previous research has found inconsistent results for the association between socioeconomic achievement gaps and centralized exams, and our supplementary analyses for SES gaps (reported in Tables A3 through A5) also find no significant association with centralized exams. The strong and significant negative association for income gaps may indicate that in countries with low levels of curricular standardization, school quality is more strongly related to income than to SES more broadly. Moreover, our finding that standardization significantly predicts smaller 90/50 gaps but does not predict 50/10 gaps suggests that greater variation in school quality may provide relative benefits to high-income students over both middle- and low-income students.

In sum, then, we find that there is a great deal of variation among wealthy countries in the extent to which children from richer and poorer households do well on standardized tests. Both this wide variation and the fact that the income achievement gap is strongly related to features of the educational system (as well as to broader social inequality and segregation) suggest that the income achievement gap is affected by social conditions and public policy decisions.

Our additional analyses of gaps based on other family socioeconomic characteristics (reported in appendix Tables A3 through A5) suggest that income achievement gaps may be substantively different from SES gaps. Unlike income gaps, SES gaps are not significantly related to income inequality or to centralized examinations. The possible finding of larger income-achievement gaps (but not larger SES-achievement gaps) in countries with higher income inequality parallels prior research finding lower income mobility (but not lower educational mobility) in countries with higher income inequality (Breen & Jonsson, 2005; Corak, 2013) and merits further research. Although we did not investigate

patterns of social mobility here, our findings on cross-national differences in income achievement gaps may help to illuminate one mechanism underlying the association between income inequality and income mobility documented by Corak (2013). If school performance (as proxied by performance on standardized tests) is an important mechanism for upward mobility, then we might expect the income students ultimately attain as adults to be more strongly correlated with parental income in countries with large income achievement gaps. Clearly, more research investigating the role of educational inequality in patterns of social mobility is needed.

Because our findings are based on cross-sectional correlational patterns, they do not provide strong evidence regarding the causal processes underlying these associations. Yet they are consistent with the argument that broad societal inequality as well as features of the schooling system may play important roles in shaping patterns of educational inequality. To establish causal relationships, it would be preferable to observe how income achievement gaps change over time within countries as social and economic conditions and policies change.

Unfortunately, such a design is not currently possible with PISA and PIRLS data. Collection of household income data in international large-scale assessments is limited and declining. Income items were removed from the PIRLS parent survey after 2001, and the number of countries administering income items in the PISA survey has declined each year since 2006. This significantly limits the potential of countries to monitor trends in income achievement gaps and the potential of scholars to investigate their causes, consequences, and remedies. Because our results suggest that achievement gaps based on household income differ from gaps based on other family socioeconomic characteristics, the SES measures widely collected in studies like PIRLS and PISA may be insufficient to fully characterize socioeconomic disparities in school performance. More and better international data on household income are essential for cross-national research on educational inequality.

Appendix

TABLE A1
Income Achievement Gaps in Sample Countries, PIRLS 2001, PISA 2006, 2009, and 2012

Country	Reading			Math			Science		
	90/10 Gap (SE)	90/50 Gap (SE)	50/10 Gap (SE)	90/10 Gap (SE)	90/50 Gap (SE)	50/10 Gap (SE)	90/10 Gap (SE)	90/50 Gap (SE)	50/10 Gap (SE)
PIRLS 2001									
Canada (Ontario)	0.881 (0.098)	0.390 (0.114)	0.492 (0.159)						
Canada (Quebec)	0.948 (0.065)	0.445 (0.070)	0.503 (0.086)						
England	0.947 (0.086)	0.561 (0.081)	0.386 (0.071)						
Germany	1.098 (0.067)	0.564 (0.058)	0.534 (0.060)						
Greece	1.112 (0.046)	0.430 (0.062)	0.682 (0.068)						
Iceland	0.616 (0.087)	0.283 (0.085)	0.333 (0.084)						
Netherlands	0.779 (0.071)	0.295 (0.069)	0.484 (0.073)						
New Zealand	0.958 (0.132)	0.423 (0.143)	0.534 (0.175)						
Norway	0.698 (0.067)	0.199 (0.063)	0.499 (0.064)						
Slovak Republic	1.104 (0.095)	0.645 (0.098)	0.458 (0.060)						
Slovenia	1.025 (0.106)	0.619 (0.124)	0.406 (0.087)						
Sweden	0.695 (0.041)	0.235 (0.047)	0.460 (0.036)						
United States ^a	1.280 (0.045)	0.652 (0.045)	0.628 (0.036)	1.284 (0.038)	0.624 (0.037)	0.660 (0.034)			
PISA 2006									
Denmark	0.650 (0.080)	0.249 (0.082)	0.401 (0.115)	0.749 (0.068)	0.281 (0.080)	0.469 (0.111)	0.697 (0.065)	0.259 (0.081)	0.438 (0.102)
Germany	1.011 (0.055)	0.504 (0.051)	0.507 (0.047)	1.148 (0.058)	0.542 (0.052)	0.606 (0.058)	1.131 (0.069)	0.553 (0.063)	0.578 (0.061)
Iceland	0.447 (0.066)	0.154 (0.056)	0.293 (0.057)	0.557 (0.059)	0.202 (0.058)	0.355 (0.061)	0.539 (0.063)	0.190 (0.060)	0.349 (0.064)
Korea	0.795 (0.027)	0.399 (0.029)	0.396 (0.026)	1.042 (0.025)	0.525 (0.023)	0.517 (0.022)	0.850 (0.037)	0.432 (0.034)	0.418 (0.036)
Luxembourg	1.390 (0.055)	0.732 (0.046)	0.658 (0.048)	1.362 (0.085)	0.700 (0.077)	0.663 (0.078)	1.391 (0.078)	0.752 (0.070)	0.639 (0.071)
New Zealand	0.924 (0.063)	0.398 (0.079)	0.526 (0.085)	0.924 (0.070)	0.383 (0.074)	0.541 (0.083)	0.924 (0.068)	0.375 (0.074)	0.549 (0.084)
Poland	0.655 (0.036)	0.179 (0.037)	0.476 (0.037)	0.758 (0.051)	0.192 (0.050)	0.566 (0.049)	0.755 (0.039)	0.192 (0.044)	0.564 (0.048)
Portugal	1.383 (0.051)	0.588 (0.051)	0.795 (0.040)	1.471 (0.048)	0.606 (0.063)	0.866 (0.050)	1.441 (0.057)	0.594 (0.061)	0.846 (0.054)
United States ^a	1.364 (0.058)	0.715 (0.054)	0.649 (0.054)	1.303 (0.041)	0.607 (0.037)	0.696 (0.040)			
PISA 2009									
Denmark	0.947 (0.091)	0.266 (0.096)	0.681 (0.128)	0.846 (0.106)	0.304 (0.116)	0.543 (0.162)	0.906 (0.092)	0.320 (0.095)	0.586 (0.136)
Germany	1.255 (0.044)	0.611 (0.042)	0.643 (0.036)	1.361 (0.064)	0.651 (0.058)	0.711 (0.056)	1.310 (0.057)	0.660 (0.056)	0.650 (0.052)
Hungary	1.249 (0.079)	0.623 (0.083)	0.626 (0.054)	1.323 (0.065)	0.655 (0.064)	0.668 (0.046)	1.274 (0.059)	0.635 (0.066)	0.639 (0.044)
Korea	0.726 (0.073)	0.426 (0.075)	0.301 (0.063)	0.930 (0.099)	0.534 (0.090)	0.396 (0.083)	0.766 (0.088)	0.460 (0.086)	0.306 (0.075)
New Zealand	0.926 (0.051)	0.363 (0.052)	0.562 (0.051)	1.036 (0.037)	0.436 (0.039)	0.600 (0.034)	0.922 (0.049)	0.388 (0.053)	0.534 (0.045)
Poland	0.787 (0.056)	0.372 (0.055)	0.416 (0.057)	0.795 (0.063)	0.358 (0.056)	0.437 (0.061)	0.822 (0.069)	0.375 (0.071)	0.446 (0.071)
Portugal	1.354 (0.043)	0.480 (0.044)	0.874 (0.029)	1.478 (0.052)	0.551 (0.051)	0.927 (0.035)	1.414 (0.053)	0.501 (0.055)	0.913 (0.034)
PISA 2012									
Belgium (Flanders)	1.347 (0.080)	0.651 (0.082)	0.696 (0.069)	1.419 (0.072)	0.658 (0.068)	0.762 (0.067)	1.369 (0.072)	0.645 (0.072)	0.724 (0.060)
Germany	1.044 (0.057)	0.519 (0.064)	0.526 (0.049)	1.113 (0.043)	0.556 (0.055)	0.557 (0.056)	1.156 (0.063)	0.568 (0.066)	0.589 (0.059)
Hungary	1.323 (0.069)	0.678 (0.071)	0.645 (0.054)	1.385 (0.059)	0.652 (0.054)	0.733 (0.051)	1.319 (0.064)	0.636 (0.063)	0.684 (0.052)
Korea	0.728 (0.030)	0.372 (0.030)	0.356 (0.030)	0.879 (0.024)	0.426 (0.020)	0.453 (0.024)	0.701 (0.022)	0.553 (0.027)	0.348 (0.024)
Portugal	1.269 (0.076)	0.512 (0.080)	0.757 (0.033)	1.485 (0.082)	0.579 (0.091)	0.906 (0.046)	1.419 (0.068)	0.568 (0.075)	0.851 (0.037)

Note: PIRLS = Progress in International Reading Literacy Study; PISA = Programme for International Student Assessment.

^aU.S. data from Early Childhood Longitudinal Study—Kindergarten Cohort.

TABLE A2

Means and Standard Deviations of 90/10, 90/50, and 50/10 Income Achievement Gaps

Study	Year	Subject	90/10 Gap		90/50 Gap		50/10 Gap		Sample N	
			<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	Obs.	Countries
PIRLS	2001	Reading	0.934	(0.224)	0.498	(0.076)	0.438	(0.162)	13	13
PISA	2006	Reading	0.959	(0.320)	0.527	(0.158)	0.437	(0.208)	9	9
PISA	2006	Math	1.035	(0.272)	0.592	(0.132)	0.450	(0.158)	9	9
PISA	2006	Science	0.964	(0.288)	0.550	(0.154)	0.417	(0.181)	8	8
PISA	2009	Reading	1.037	(0.252)	0.587	(0.200)	0.455	(0.108)	7	7
PISA	2009	Math	1.116	(0.258)	0.622	(0.185)	0.507	(0.108)	7	7
PISA	2009	Science	1.064	(0.255)	0.587	(0.204)	0.484	(0.112)	7	7
PISA	2012	Reading	1.138	(0.318)	0.594	(0.196)	0.538	(0.136)	5	5
PISA	2012	Math	1.251	(0.281)	0.680	(0.212)	0.567	(0.119)	5	5
PISA	2012	Science	1.190	(0.402)	0.638	(0.256)	0.547	(0.154)	5	5
All	All	All	1.043	(0.267)	0.575	(0.169)	0.472	(0.139)	75	20

Note. U.S. data from Early Childhood Longitudinal Study–Kindergarten Cohort is included with PIRLS 2001 and PISA 2006. In each row, we can reject the null hypothesis that the true variance of the gaps is 0 ($p < .001$). PIRLS = Progress in International Reading Literacy Study; PISA = Programme for International Student Assessment; Obs. = observations.

TABLE A3

Estimated Multivariate Associations of Country Characteristics With 90/10 Parental Education Achievement Gap (Original Sample)

Variable	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Educational differentiation index		0.090*	0.064	0.103	0.054	0.099*
		(0.043)	(0.049)	(0.053)	(0.048)	(0.050)
Centralized exams		−0.062	−0.081	−0.068	−0.085 *	−0.057
		(0.040)	(0.042)	(0.049)	(0.041)	(0.045)
Poverty/inequality index		0.173***				
		(0.038)				
Parental support index		0.017				
		(0.034)				
Social welfare policy index		0.039				
		(0.045)				
Income inequality (Gini index)			−0.082	0.086		
			(0.077)	(0.052)		
School parent education segregation			0.086*		0.094*	
			(0.039)		(0.039)	
Child poverty rate			0.133*			0.114**
			(0.063)			(0.044)
Intercept	1.132***	1.123***	1.178***	1.159***	1.148***	1.136***
	(0.059)	(0.044)	(0.052)	(0.051)	(0.053)	(0.049)
Within-country residual variance	0.001	0.000	0.000	0.000	0.001	0.000
Between-country residual variance	0.042	0.011	0.016	0.025	0.022	0.023
Variance explained (between countries)		0.729	0.615	0.404	0.489	0.456
<i>N</i> (observations)	75	75	75	75	75	75
<i>N</i> (countries)	20	20	20	20	20	20

Note. Estimates from random-effects models using pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data. United States included; U.S. data from Early Childhood Longitudinal Study–Kindergarten Cohort. Standard errors in parentheses.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A4

Estimated Multivariate Associations of Country Characteristics With 90/10 Parental Education Achievement Gap (Full OECD Population; All Years of PIRLS and PISA)

Variable	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Educational differentiation index		0.137*** (0.029)	0.115** (0.035)	0.120** (0.040)	0.106*** (0.031)	0.115*** (0.035)
Centralized exams		0.010 (0.029)	0.011 (0.031)	0.008 (0.035)	-0.001 (0.030)	0.018 (0.034)
Poverty/inequality index		0.156*** (0.032)				
Parental support index		0.056* (0.028)				
Social welfare policy index		0.003 (0.033)				
Income inequality (Gini index)			-0.016 (0.058)	0.066 (0.040)		
School parent education segregation			0.080* (0.032)		0.097** (0.030)	
Child poverty rate			0.057 (0.054)			0.081* (0.035)
Intercept	1.015 *** (0.044)	1.019*** (0.035)	1.031*** (0.037)	1.015*** (0.040)	1.032*** (0.038)	1.014*** (0.039)
Within-country residual variance	0.010	0.010	0.009	0.010	0.010	0.009
Between-country residual variance	0.041	0.018	0.024	0.032	0.025	0.030
Variance explained (between countries)		0.560	0.428	0.234	0.399	0.280
<i>N</i> (observations)	474	474	474	474	474	474
<i>N</i> (countries)	33	33	33	33	33	33

Note. OECD = Organisation for Economic Co-operation and Development; PIRLS = Progress in International Reading Literacy Study; PISA = Programme for International Student Assessment. Estimates from random-effects models using pooled PIRLS (2001, 2006, 2011) and PISA (2000, 2003, 2006, 2009, 2012) data. United States included; U.S. data from PISA 2000 to 2012. Standard errors in parentheses.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A5

Estimated Multivariate Associations of Country Characteristics With 90/10 ESCS Achievement Gap (Full OECD Population; All Years of PISA)

Variable	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Educational differentiation index		0.094*** (0.022)	0.050 (0.030)	0.108*** (0.032)	0.044 (0.024)	0.102*** (0.028)
Centralized exams		-0.006 (0.023)	0.016 (0.024)	0.005 (0.030)	0.012 (0.023)	0.010 (0.028)
Poverty/inequality index		0.127*** (0.023)				
Parental support index		0.026 (0.022)				
Social welfare policy index		0.046 (0.025)				

(continued)

TABLE A5 (CONTINUED)

Variable	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Income inequality (Gini index)			-0.017 (0.045)	0.052 (0.033)		
School ESCS segregation			0.098*** (0.027)		0.107*** (0.025)	
Child poverty rate			0.034 (0.042)			0.062* (0.029)
Intercept	1.099*** (0.033)	1.103*** (0.022)	1.113*** (0.023)	1.100*** (0.028)	1.113*** (0.023)	1.099*** (0.028)
Within-country residual variance	0.004	0.004	0.003	0.004	0.003	0.004
Between-country residual variance	0.030	0.011	0.014	0.021	0.014	0.020
Variance explained (between countries)		0.623	0.539	0.290	0.529	0.333
<i>N</i> (observations)	437	437	437	437	437	437
<i>N</i> (countries)	30	30	30	30	30	30

Note. OECD = Organisation for Economic Co-operation and Development; PISA = Programme for International Student Assessment; ESCS = PISA index of economic, social, and cultural status. Estimates from random-effects models using pooled PISA (2000, 2003, 2006, 2009, 2012) data. United States included; U.S. data from PISA 2000 to 2012. Standard errors in parentheses.
 * $p < .05$. ** $p < .01$. *** $p < .001$.

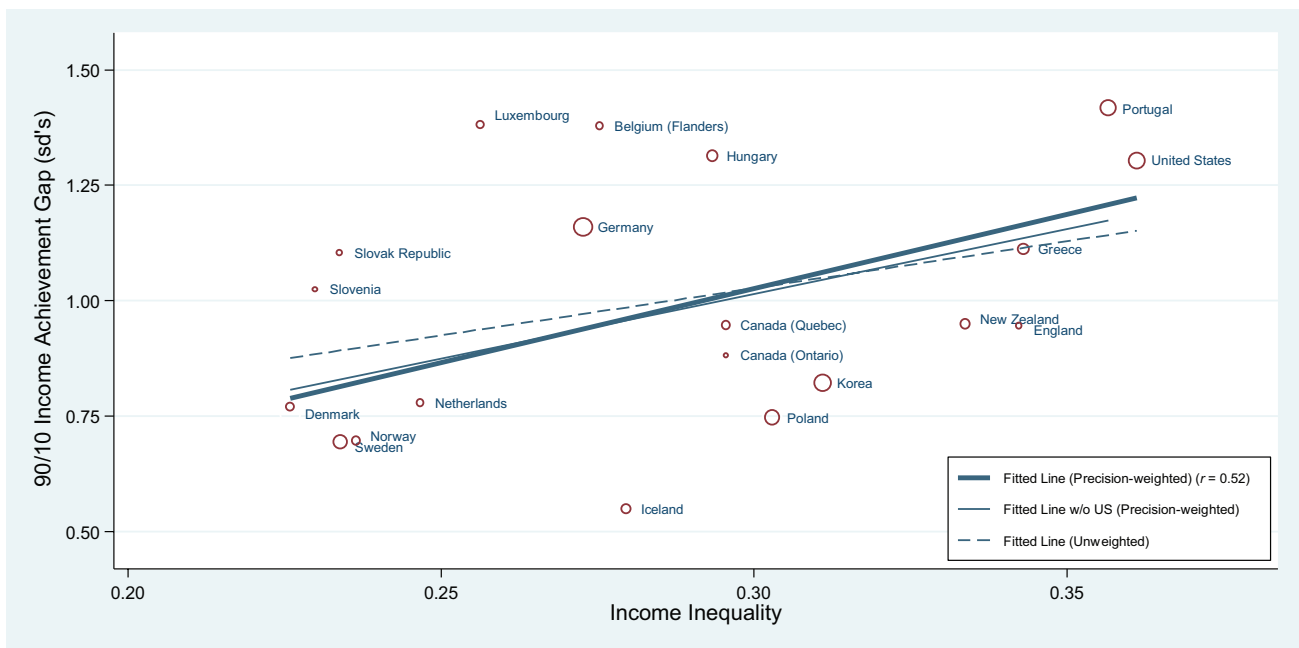


FIGURE A1. Association between income achievement gap and income inequality, wealthy Organisation for Economic Co-operation and Development countries, 2001 to 2012 (pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data).

Note. The size of each circle indicates the precision of the 90/10 gap estimate, with the larger circles indicating the most precisely estimated gaps.

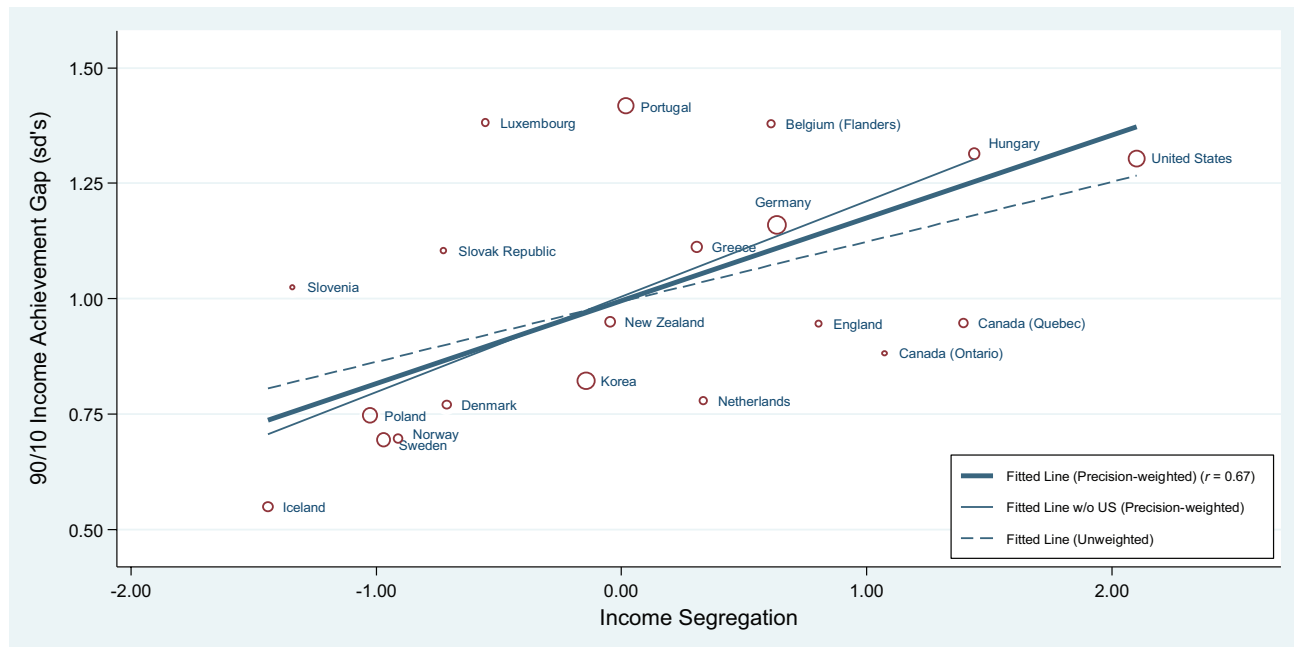


FIGURE A2. Association between income achievement gap and income segregation, wealthy Organisation for Economic Co-operation and Development countries, 2001 to 2012 (pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data).
 Note. The size of each circle indicates the precision of the 90/10 gap estimate, with the larger circles indicating the most precisely estimated gaps.

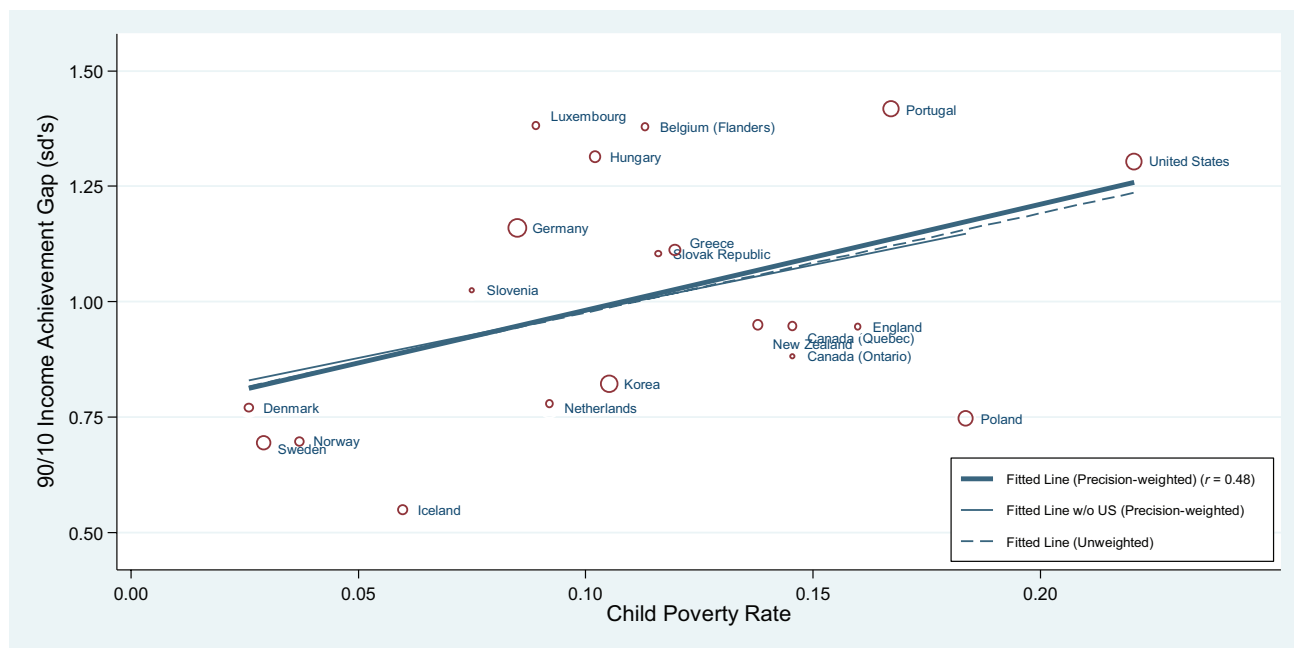


FIGURE A3. Association between income achievement gap and child poverty rate, wealthy Organisation for Economic Co-operation and Development countries, 2001 to 2012 (pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data).
 Note. The size of each circle indicates the precision of the 90/10 gap estimate, with the larger circles indicating the most precisely estimated gaps.

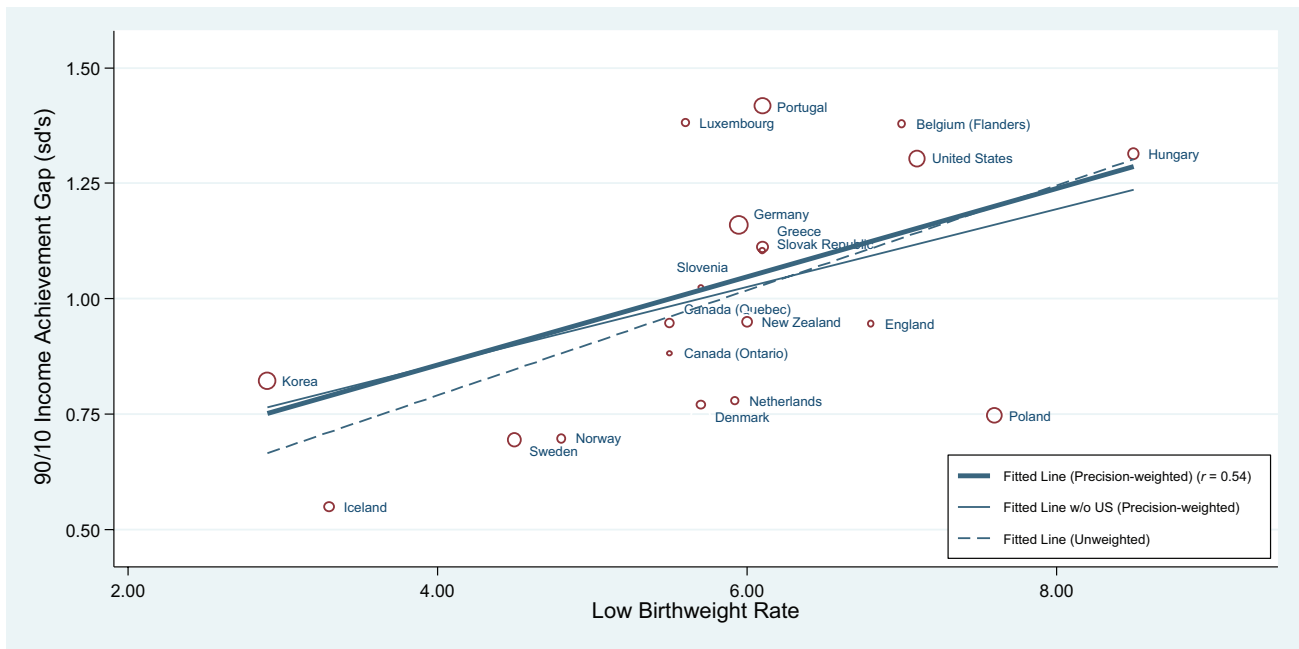


FIGURE A4. Association between income achievement gap and low-birthweight rate, wealthy Organisation for Economic Co-operation and Development countries, 2001 to 2012 (pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data).
 Note. The size of each circle indicates the precision of the 90/10 gap estimate, with the larger circles indicating the most precisely estimated gaps.

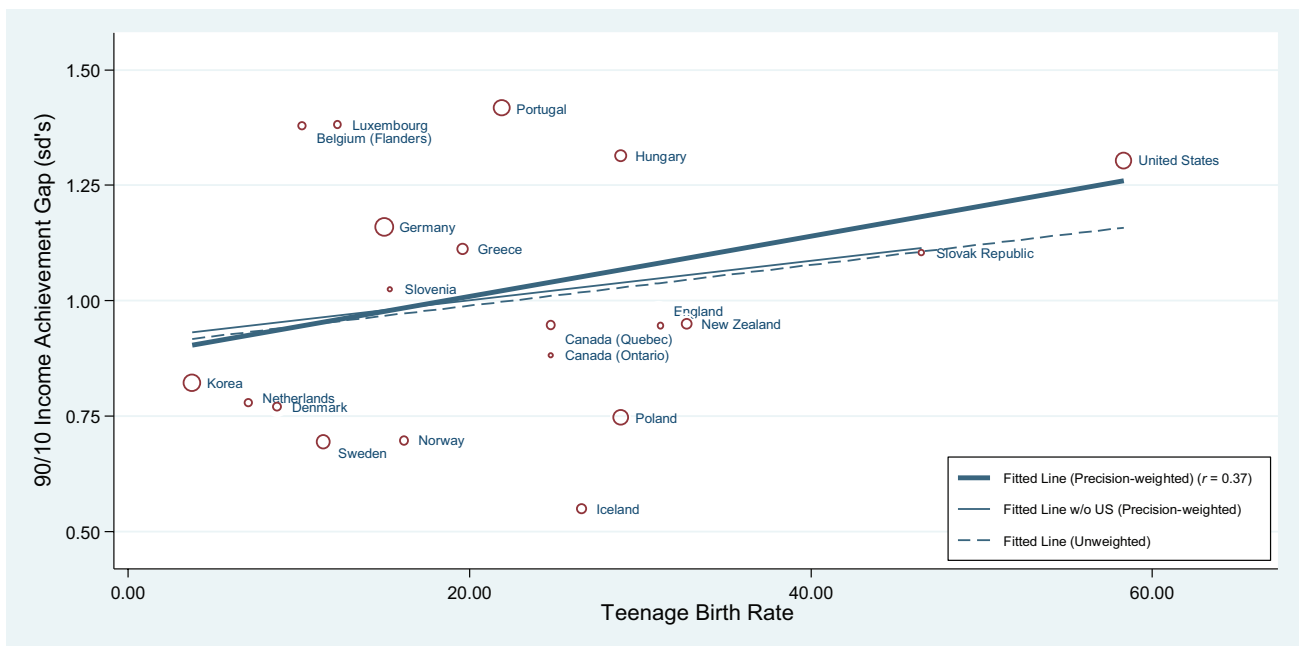


FIGURE A5. Association between income achievement gap and teenage birth rate, wealthy Organisation for Economic Co-operation and Development countries, 2001 to 2012 (pooled Progress in International Reading Literacy Study and Programme for International Student Assessment data).
 Note. The size of each circle indicates the precision of the 90/10 gap estimate, with the larger circles indicating the most precisely estimated gaps.

Authors' Note

An earlier version of this paper was prepared for the conference "Income, Inequality, and Educational Success: New Evidence about Socioeconomic Status and Educational Outcomes," supported by the American Educational Research Association, Stanford University, May 15 to 16, 2012.

Notes

1. We obtain GDP per capita from the World Bank Factbook (World Bank, 2012). We use the 2012 value converted to constant 2011 international dollars using purchasing-power parity rates.

2. We check the robustness of our results by loosening these exclusion criteria and rerunning multivariate models for as many countries as possible. Due to the large amount of missing data for country covariates, the only two countries with complete covariate data that we could include in our models are the Czech Republic and Mexico, both Organisation for Economic Co-operation and Development (OECD) member countries with low-quality income data, resulting in very large standard errors on their estimated income achievement gaps. Results for models including these two extra countries are nearly identical to our main results for the sample of 20 countries. This is unsurprising because gaps with larger standard errors contribute less in our precision-weighted models.

3. Canada (Ontario), Canada (Quebec), England, Germany, Greece, Iceland, the Netherlands, New Zealand, Norway, the Slovak Republic, Slovenia, and Sweden.

4. Denmark, Germany, Iceland, Korea, Luxembourg, New Zealand, Poland, and Portugal.

5. Denmark, Germany, Hungary, Korea, New Zealand, Poland, and Portugal.

6. Belgium (Flanders), Germany, Hungary, Korea, and Portugal.

7. In practice, it does not appear that the income categories were always defined this way in each country in the Programme for International Student Assessment (PISA) study. If they were defined this way, then the income categories in each country would span most of the income distribution (three categories above the median income and three below), but several countries have very large proportions (sometimes close to half) of the sample with incomes reported in the highest or lowest categories. The proportions of incomes reported in the highest and lowest categories are reported for each country in Table 1.

8. As described in Reardon (2011), the adjustment amounts to dividing each gap estimate (and the corresponding standard error) by $\sqrt{r_s r_i}$, where r_s and r_i are the reliabilities of the test and of parent-reported family income, respectively. For Progress in International Reading Literacy Study (PIRLS) 2001 and each year of PISA, we obtain country-specific estimates of r_s from the OECD's technical reports (in the case of PISA) and the International Association for the Evaluation of Educational Achievement's international report (in the case of PIRLS). For the United States, we use the published reliabilities of the Early Childhood Longitudinal Study-Kindergarten Cohort tests (Najarian et al., 2009). For r_i , we assign a value of 0.86. This is the average estimated reliability of parent-reported income reported in Marquis, Marquis, and Polich (1986).

9. We use the segregation between above- and below-median income students in the poverty/inequality index. In some of the

models described later, however, we also use measures of the segregation between students from households with incomes above and below the 75th percentile and of the segregation between students from households with incomes above and below the 25th percentile.

10. In a fully unconditional model (with no test or subject dummies), about 7% of the variance in gaps is within countries, and 93% is between countries.

11. Based on exploratory analyses, we did not include the low-birthweight rate or the teenage childbearing rate as they were not as strongly associated with the achievement gaps as these three items.

12. We also ran models examining the different variables composing the differentiation, parental support, and social welfare indices separately, controlling for poverty/inequality and all other indices. In the differentiation index, the variables related to tracking (especially the age of track selection and the number of tracks) are more strongly associated with gaps than was the share of private school enrollment. In the parental support index, none of the variables is strongly associated with gaps. In the social welfare policy index, the only variable strongly associated with gaps is public spending on family benefits in cash, but this association is positive rather than negative as hypothesized.

References

- Alegre, M. À., & Ferrer, G. (2010). School regimes and education equity: Some insights based on PISA 2006. *British Educational Research Journal*, 36(3), 433–461.
- Allmendinger, J. (1989). Educational systems and labor market outcomes. *European Sociological Review*, 5(3), 231.
- Ammermüller, A. (2005). *Educational opportunities and the role of institutions* (ZEW Discussion Paper No. 05-44). Mannheim, Germany: Center for European Economic Research.
- Ayalon, H., & Gamoran, A. (2000). Stratification in academic secondary programs and educational inequality in Israel and the United States. *Comparative Education Review*, 44(1), 54–80.
- Baker, D. P., Goesling, B., & LeTendre, G. K. (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.
- Bassok, D., & Latham, S. (2014, November). *Kids today: Changes in school readiness in an early childhood era*. Paper presented at the annual meeting of the Association for Public Policy Analysis and Management, Albuquerque, NM.
- Bauer, P., & Riphahn, R. T. (2006). Timing of school tracking as a determinant of intergenerational transmission of education. *Economics Letters*, 91(1), 90–97.
- Bischoff, K., & Reardon, S. F. (2014). Residential segregation by income, 1970–2009. In J. Logan (Ed.), *Diversity and disparities: America enters a new century* (pp. 208–233). New York, NY: Russell Sage Foundation.
- Bishop, J. H. (1997). The effect of national standards and curriculum-based exams on achievement. *American Economic Review*, 87(2), 260–264.
- Bol, T., Witschge, J., Van de Werfhorst, H. G., & Dronkers, J. (2014). Curricular tracking and central examinations: Counterbalancing the impact of social background on student achievement in 36 countries. *Social Forces*, 92(4), 1545–1572.

- Breen, R., & Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: recent research on educational attainment and social mobility. *Annual Review of Sociology, 31*, 223–243.
- Brunello, G., & Checchi, D. (2007). Does school tracking affect equality of opportunity? New international evidence. *Economic Policy, 22*(52), 781–861.
- Campbell, J. R., Kelly, D. L., Mullis, I. V. S., Martin, M. O., & Sainsbury, M. (2001). *Framework and specifications for PIRLS Assessment 2001*. Chestnut Hill, MA: Boston College.
- Causa, O., & Chapuis, C. (2009). *Equity in student achievement across OECD countries: an investigation of the role of policies*. Working paper, Organisation for Economic Co-operation and Development, Paris, France.
- Chetty, R., Hendren, N., & Katz, L. F. (2015). *The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment*. Cambridge, MA: Harvard University.
- Chiu, M. M., & Khoo, L. (2005). Effects of resources, inequality, and privilege bias on achievement: Country, school, and student level analyses. *American Educational Research Journal, 42*(4), 575–603.
- Chmielewski, A. K. (2014). An international comparison of achievement inequality in within-and between-school tracking systems. *American Journal of Education, 120*(3), 293–324.
- Chudgar, A., & Luschei, T. F. (2009). National income, income inequality, and the importance of schools: A hierarchical cross-national comparison. *American Educational Research Journal, 46*(3), 626.
- Corak, M. (2013). Inequality from generation to generation: The United States in comparison. In R. Rycroft (Ed.), *The economics of inequality, poverty, and discrimination in the 21st century* (pp. 107–126). Santa Barbara, CA: ABC-CLIO.
- Demeuse, M., & Baye, A. (2008). Measuring and comparing the equity of education systems in Europe. In N. C. Soguel & P. Jaccard (Eds.), *Governance and performance of education systems* (pp. 85–106). Dordrecht, Netherlands: Springer.
- Dupriez, V., & Dumay, X. (2006). Inequalities in school systems: Effect of school structure or of society structure? *Comparative Education, 42*(2), 243–260.
- Duru-Bellat, M., & Suchaut, B. (2005). Organisation and context, efficiency and equity of educational systems: What PISA tells us. *European Educational Research Journal, 4*(3), 181–194.
- Field, S., Kuczera, M., & Pont, B. (2007). *No more failures: Ten steps to equity in education*. Paris, France: Organisation for Economic Co-operation and Development.
- Fuchs, T., & Woessmann, L. (2007). What accounts for international differences in student performance? A re-examination using PISA data. *Empirical Economics, 32*(2), 433–464.
- Gonzalez, E. J., & Kennedy, A. M. (2003). *PIRLS 2001 user guide for the international database: Supplement two*. Boston, MA: International Association for the Evaluation of Educational Achievement.
- Gorard, S., & Smith, E. (2004). An international comparison of equity in education systems. *Comparative Education, 40*(1), 15–28.
- Heyneman, S. P., & Loxley, W. A. (1983). The effect of primary-school quality on academic achievement across twenty-nine high- and low-income countries. *American Journal of Sociology, 88*(6), 1162–1194.
- Horn, D. (2009). Age of selection counts: A cross-country analysis of educational institutions. *Educational Research and Evaluation, 15*(4), 343–366.
- Jakubowski, M. (2010). Institutional tracking and achievement growth: Exploring difference-in-differences approach to PIRLS, TIMSS, and PISA data. In J. Dronkers (Ed.), *Quality and inequality of education* (pp. 41–81). Dordrecht, Netherlands: Springer Science+Business Media.
- Jenkins, S. P., Micklewright, J., & Schnepf, S. V. (2006). *Social segregation in secondary schools: How does England compare with other countries?* (IZA Discussion Paper No. 1959). Bonn, Germany: Institute for the Study of Labor.
- Luxembourg Income Study. (2009). *Key figures*. Luxembourg: Author.
- Magnuson, K. A., & Duncan, G. J. (2014, October). *Can early childhood interventions decrease inequality of economic opportunity?* Paper presented at the Federal Reserve Bank of Boston Conference, Boston, MA.
- Marks, G. N. (2005). Cross-national differences and accounting for social class inequalities in education. *International Sociology, 20*(4), 483–505.
- Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socio-economic inequalities in student achievement: The role of home and school factors. *Educational Research and Evaluation, 12*(2), 105–128.
- Marquis, K. H., Marquis, M. S., & Polich, J. M. (1986). Response bias and reliability in sensitive topic surveys. *Journal of the American Statistical Association, 81*(394), 381–389.
- Mühlenweg, A. (2007). *Educational effects of early or later secondary school tracking in Germany* (ZEW Discussion Paper 07-079). Mannheim, Germany: Center for European Economic Research.
- Muller, C., & Schiller, K. S. (2000). Leveling the playing field? Students' educational attainment and states' performance testing. *Sociology of Education, 73*(3), 196–218.
- Musterd, S. (2005). Social and ethnic segregation in Europe: Levels, causes, and effects. *Journal of Urban Affairs, 27*(3), 331–348.
- Najarian, M., Pollack, J. M., Sogongon, A. G., & Hausken, E. G. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Psychometric Report for the Eighth Grade (NCES 2009-002)*. National Center for Educational Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- Nonoyama-Tarumi, Y. (2008). Cross-national estimates of the effects of family background on student achievement: A sensitivity analysis. *International Review of Education, 54*, 57–82.
- Organisation for Economic Co-operation and Development. (2006). *Assessing scientific, reading, and mathematical literacy: A framework for PISA 2006*. Paris, France: Author.
- Organisation for Economic Co-operation and Development. (2007). *PISA 2006: Science competencies for tomorrow's world*. Paris, France: Author.
- Organisation for Economic Co-operation and Development. (2009a). *PISA 2006 technical report*. Paris, France: Author.
- Organisation for Economic Co-operation and Development. (2009b). *PISA 2009 assessment framework: Key competencies in reading, mathematics and science*. Paris, France: Author.
- Organisation for Economic Co-operation and Development (2010a). *PISA 2009 results: Vol. 2. Overcoming social background:*

- Equity in learning opportunities and outcomes*. Paris, France: Author. <http://dx.doi.org/10.1787/9789264091504-en>
- Organisation for Economic Co-operation and Development. (2010b). *PISA 2009 results: Vol. 4. What makes a school successful? Resources, policies and practices*. Paris, France: Author. http://www.oecd-ilibrary.org/education/pisa-2009-results-what-makes-a-school-successful_9789264091559-en
- Organisation for Economic Co-operation and Development. (2011a). *Income distribution: Inequality*. Retrieved from <http://stats.oecd.org>
- Organisation for Economic Co-operation and Development. (2011b). *OECD family database*. Paris, France: Author.
- Organisation for Economic Co-operation and Development. (2013). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. Paris, France: Author.
- Owens, A. (2015). *Economic segregation between neighborhoods and schools*. Retrieved from http://scholar.harvard.edu/files/aowens/files/owens_inequality_childrens_contexts_feb15.pdf
- Park, H. (2008). The varied educational effects of parent-child communication: A comparative study of fourteen countries. *Comparative Education Review*, 52(2), 25.
- Raudenbush, S., Bryk, A., & Congdon, R. (2004). *HLM: Hierarchical linear and nonlinear modeling, v6.0* [Computer software].
- Reardon, S. F. (2011). The widening academic-achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–115). New York, NY: Russell Sage Foundation.
- Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology*, 116(4), 1092–1153.
- Reardon, S. F., & Portilla, X. (2015). *Recent trends in socioeconomic and racial school readiness gaps at kindergarten entry*. Working paper.
- Schubert, F., & Becker, R. (2010). Social inequality of reading literacy: A longitudinal analysis with cross-sectional data of PIRLS 2001 and PISA 2000 utilizing the pair wise matching procedure. *Research in Social Stratification and Mobility*, 28(1), 109–133.
- Schütz, G., Ursprung, H. W., & Wößmann, L. (2008). Education policy and equality of opportunity. *Kyklos*, 61(2), 279–308.
- Schütz, G., West, M. R., & Wößmann, L. (2007). *School accountability, autonomy, choice, and the equity of student achievement: international evidence from PISA 2003*. Paris, France: Organisation for Economic Co-operation and Development.
- Shavit, Y., & Müller, W. (1998). *From school to work: A comparative study of educational qualifications and occupational destinations*. New York, NY: Oxford University Press.
- Stevenson, D. L., & Baker, D. P. (1991). State control of the curriculum and classroom instruction. *Sociology of Education*, 64(1), 1–10.
- Van de Werfhorst, H. G., & Mijs, J. J. B. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology*, 36, 407–428.
- Waldinger, F. (2006). *Does Tracking affect the importance of family background on students' test scores?* Mimeo, London School of Economics, London, UK.
- West, A. (2006). School choice, equity and social justice: The case for more control. *British Journal of Educational Studies*, 54(1), 15–33.
- Westbury, I., & Hsu, C. S. (1996). Structures of curriculum governance and classroom practice in mathematics. *Educational Evaluation and Policy Analysis*, 18(2), 123.
- Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record*, 112(4), 3–4.
- World Bank. (2012). *Factbook*. Washington, DC: Author.
- Wößmann, L. (2005). The effect heterogeneity of central examinations: Evidence from TIMSS, TIMSS-Repeat and PISA. *Education Economics*, 13(2), 143–169.
- Wößmann, L. (2007). *Fundamental determinants of school efficiency and equity: German states as a microcosm for OECD countries* (IZA Discussion Paper No. 2880). Bonn, Germany: Institute for the Study of Labor.
- Wößmann, L., Lüdemann, E., Schütz, G., & West, M. R. (2009). *School accountability, autonomy, and choice around the world*. Cheltenham, UK: Edward Elgar.

Authors

ANNA K. CHMIELEWSKI is an assistant professor of educational leadership and policy at the Ontario Institute for Studies in Education (OISE) of the University of Toronto, 252 Bloor Street West, Toronto, ON M5S1V6, Canada; ak.chmielewski@utoronto.ca. Her research focuses on international comparisons of educational and social inequality; tracking and ability grouping; and international large-scale assessments.

SEAN F. REARDON is the endowed professor of poverty and inequality in education and is a professor (by courtesy) of sociology at Stanford University, 520 Galvez Mall, CERAS building, Stanford, CA 94305-3084; sean.reardon@stanford.edu. His research focuses on the causes, patterns, trends, and consequences of social and educational inequality; the effects of educational policy on educational and social inequality; and applied statistical methods for education research.