Autonomous Schools, Achievement, and Segregation*

Natalie Irmert† Jan Bietenbeck‡

Linn Mattisson§ Felix Weinhardt¶

October 8, 2024

Abstract

We study the impact of autonomous schools – publicly funded institutions that operate more independently than government-run schools – on student achievement and school segregation, using data from 15 countries over 16 years. Our triple-differences regressions exploit between-grade variation in the share of students attending autonomous schools within a given country and year. We find that autonomous schools do not raise overall achievement, and our estimates are precise enough to rule out even modest positive effects in math and small positive effects in science. However, these aggregate results mask important heterogeneity, with consistently positive effects for high-socioeconomic-status students and natives, and negative effects for low-socioeconomic-status students and immigrants. In line with these results, we also find that autonomous schools increase segregation by socioeconomic and immigrant status. We conclude that autonomous schools have not generated the anticipated system-wide benefits.

Keywords: autonomous schools, student achievement, school segregation
JEL codes: I21, I24, J15

---

*We thank Olmo Silva, Jonas Vlachos, Ludger Woessmann, and audiences at the spring 2023 Copenhagen Education Network Workshop, the 2023 AEDE Annual Meeting, the 2024 meeting of the VfS standing field committee of the economics of education, and the 2024 EALE Conference for helpful comments. We are grateful for funding from the Jan Wallanders och Tom Hedelius stiftelse samt Tore Browaldhs stiftelse. Jan Bietenbeck gratefully acknowledges the hospitality of the University of Duisburg-Essen while working on this paper.

†Lund University.

‡Lund University and CESifo, DIW Berlin, IZA.

§Socialstyrelsen.

¶European University Viadrina, Berlin School of Economics, DIW Berlin, CEP at the LSE, CESifo, IZA.

1 Introduction

In the past few decades, many countries have introduced new types of publicly funded schools that can operate more independently than traditional government-run schools. Examples of such "autonomous schools" include academies in England, church schools in Hungary, free schools in Sweden, and charter schools in the United States.[1] The central theoretical motivation for establishing these schools is the notion that there are competition-related welfare gains in the school sector: autonomous schools increase choice and thus competition for students who can vote with their feet (Tiebout 1956, Hoxby 2000). This creates an incentive for all schools to boost their productivity, which could lead to "a rising tide that lifts all boats" (Hoxby 2003). In practice, the rise of autonomous schools has sparked an intense academic and public debate, which primarily centers on their impacts on student achievement and school segregation.

In this paper, we provide a comprehensive assessment of the effects of autonomous schools on these outcomes. Based on data from 15 countries over 16 years, our results paint a rather sobering picture. We find no evidence that autonomous schools improve overall achievement on internationally comparable mathematics and science tests. Our estimates are precise and robust and, if anything, point towards small negative effects. However, these aggregate results mask considerable heterogeneity: we find consistently positive effects for high-socioeconomic status and native students, while low-socioeconomic status and immigrant students experience negative effects. Furthermore, we show that autonomous schools exacerbate school segregation by socioeconomic and immigrant status. We conclude that the anticipated competition-induced, system-wide benefits of autonomous schools have not materialized at the international level.

To reach this conclusion, we assemble data on the prevalence of autonomous schools from the OECD's annual *Education At a Glance* reports. The data capture the share of students in a country who attend autonomous schools, separately for primary and lower secondary education. To measure achievement and segregation, we use data from

[1] As we show in Section 2, what sets these schools apart from traditional government-run schools are very high levels of autonomy in budget and personnel decisions. We therefore refer to these schools as "autonomous schools" in the remainder of the paper.

five waves of the Trends in International Mathematics and Science Study (TIMSS), an international assessment of students' math and science knowledge. The assessment covers both fourth grade, which we link to the prevalence of autonomous schools in primary education, and eighth grade, which we link to the prevalence in lower secondary education. The linked data span the years 2003 to 2019 and contain individual-level information on achievement for 484,526 students. We add to this four measures of segregation by socioeconomic status (SES) and immigrant status at the country-year-grade level, which we compute using information gathered via the TIMSS student questionnaire and which we describe in more detail below.

Identifying the causal effects of autonomous schools in an international setting is challenging because national school systems differ in many unobserved ways. We address this challenge using a novel triple-differences research design, which exploits between-grade variation in the share of students attending autonomous schools within a given country and year. Our approach accounts for all country-year-specific shocks, grade-year-specific shocks, and time-invariant country-grade differences that could confound the effects of interest. Crucially, this means that we non-parametrically control for any system-wide changes in a country's education policy during our study period. A similar triple-differences approach has recently been used to estimate the effects of charter schools across regions in the United States (Monarrez et al. 2022). Our implementation extends this methodology to an international setting, yielding causal estimates of the country-wide impacts of autonomous schools.

The results show that autonomous schools do not affect overall math achievement and have a small negative effect on overall science achievement. Our estimates rule out even modest positive effects: the upper bounds of the 95 percent confidence intervals imply that a 10 percentage point (p.p.) increase in the share of students attending autonomous schools raises math achievement by only 0.013 standard deviations (SD) and has no effect on science achievement. These findings are robust to alternative sample restrictions and hold when exploiting only long-term variation in the prevalence of autonomous schools in a "long-differences" estimation. This suggests that the introduction and expansion of autonomous schools has not created the hoped-for "rising tide that lifts all boats."

In further analyses, we explore heterogeneity in these results by student background. We find consistently more positive effects for high-SES students and native students in both subjects. For example, we estimate that a 10 p.p. increase in the share of students attending autonomous schools raises math achievement by 0.022 SD among high-SES students, but reduces it by 0.010 SD among low-SES students. An implication of this heterogeneity is that existing achievement gaps by SES and immigrant status increase. Specifically, a 10 p.p. increase in the share of students attending autonomous schools widens the SES achievement gap by five percent and the native-immigrant achievement gap by 14 percent in math.

Building on the observed heterogeneity in the effects on achievement, we construct four measures of class-level segregation for high-SES and immigrant students. First, we calculate the variance ratio index, which captures the relative probability that a randomly selected classmate of a minority student belongs to the same minority group. Second, we compute the dissimilarity index, which measures the proportion of minority students who would need to change classes to achieve equal representation across classes. We construct these segregation measures separately by country, year, and grade, allowing us to apply the same triple-differences design used in the analysis of achievement effects. We find that autonomous schools increase segregation for high-SES and immigrant students. For example, we estimate that a 10 p.p. increase in the share of students attending autonomous schools raises segregation for high-SES students, as measured by the variance ratio index, by 0.14 SD.

Our paper contributes to the growing literature on the effects of autonomous schools on student achievement. These effects operate through two main channels: first, students attending autonomous schools may experience benefits, an effect referred to in the literature as the "direct effect" (Gilraine et al. 2021, Mumma 2022). Second, autonomous schools may exert competitive pressure on government-run schools, influencing achievement there; this is known as the "indirect effect." To accurately assess the overall impact of autonomous schools, both channels must be considered.

The largest portion of the related literature focuses exclusively on direct effects, with mixed findings (e.g. Abdulkadiroğlu et al. 2011, Angrist et al. 2013, Dobbie & Fryer 2015, Eyles et al. 2017, Eyles & Machin 2019, Bertoni et al. 2023). A much smaller portion examines indirect effects, also yielding mixed results (e.g. Winters 2012, Figlio et al. 2021, Mumma 2022).[2] The variation in settings and outcomes across these studies makes it difficult to draw firm conclusions about the overall effect of autonomous schools on achievement, which requires integrating both direct and indirect effects.

A handful of studies address this issue by exploiting regional variation in the prevalence of autonomous schools to estimate the "aggregate effect," which combines direct and indirect effects without separately identifying them. Gilraine et al. (2021) show that charter school entry in North Carolina improves math achievement at the market level. Using nationwide data, Chen & Harris (2023) find that charter schools improve overall market-level English Language Arts achievement, with no significant effect on math. At the municipality level, Böhlmark & Lindahl (2015) find that free schools boost aggregate student achievement in Sweden, whereas Hsieh & Urquiola (2006) report zero to negative effects of voucher schools in Chile. At the international level, Hanushek et al. (2013) use cross-country variation in decision-making autonomy, as reported by school principals, and student achievement in the Programme for International Student Assessment. They find that autonomy boosts achievement in richer countries, but that it is detrimental in poorer countries. In a related study, West & Woessmann (2010) find that private schools improve student achievement in an international setting. We contribute to this literature by providing international evidence on the aggregate effect of autonomous schools, adapting a novel estimation strategy that controls for unobserved country-by-year shocks.

Our paper also contributes to the much smaller literature on the effects of autonomous schools on segregation. It is most closely related to Monarrez et al. (2022), who find that charter schools modestly increase school segregation by race and ethnicity in the United States. Their conclusions are based on credible estimates from a triple-differences model that leverages variation in the share of students attending charter schools across locations,

[2] Cohodes & Parham (2021) provide an excellent overview of both strands of literature in the context of charter schools in the United States.

grades, and years. This approach inspired our empirical strategy, which exploits similar variation at the international level. Earlier research by Hsieh & Urquiola (2006) shows that the expansion of voucher schools in Chile during the 1980s increased student sorting, while Böhlmark et al. (2016) find that free school penetration in Sweden is associated with segregation by socioeconomic status and immigrant background. In contrast to this existing literature, our study is the first to examine the causal effect of autonomous schools on segregation at the international level.

In summary, by estimating both the aggregate effect of autonomous schools on achievement and their impact on segregation internationally, this study provides the most comprehensive assessment to date of the effects of autonomous schools. Additionally, ours is the first study to implement a triple-differences design to examine education outcomes at the international level. With grade-specific data on education policies becoming increasingly available, we expect this identification strategy to be applied in future cross-country studies of determinants of education.

2 Data

2.1 Data on the prevalence of autonomous schools

We collect data on the prevalence of autonomous schools from the OECD's annual *Education at a Glance* reports for the years 1999 to 2019. The reports provide internationally comparable indicators on education systems, including enrollment in different types of schools. The relevant indicator for our purposes is the share of students enrolled in government-dependent private schools, defined as schools that receive at least half of their funding from the government but are privately managed (OECD 2018). Crucially for our identification strategy, the share of students attending these schools is reported separately for primary and lower secondary education.

To further characterize government-dependent private schools, we use school-level data from the OECD's Programme for International Student Assessment (PISA), which asks principals to describe their school's funding sources and the extent of autonomy in various decision-making areas. Following Hanushek et al. (2013), we construct three indices measuring personnel, budget, and academic-content autonomy, with scores ranging from zero, indicating no autonomy, to 100, indicating full autonomy.[3] Figure 1 presents the means of these indices separately for government-run and government-dependent private schools. Principals of the latter report substantially greater autonomy across all three dimensions: the average budget autonomy score is 93, compared to 45 for government-run schools. Academic content autonomy is also higher, with an average score of 91 compared to 81. Finally, personnel autonomy scores average 74 for government-dependent private schools, compared to 40 for government-run schools. Figure 1 also shows that the share of government funding is similar for the two types of schools (80 percent and 87 percent). Taken together, the data confirm that government-dependent private schools are indeed publicly funded but enjoy significantly greater autonomy than government-run schools, justifying our labeling of them as autonomous schools.[4]

2.2 Data on student achievement and segregation

To measure achievement and segregation, we use individual-level data from TIMSS, an international student assessment conducted every four years since 1995 by the International Association for the Evaluation of Educational Achievement. TIMSS employs a two-stage clustered sampling approach to select nationally representative samples of fourth- and eighth-grade students. In the first stage, schools are selected, and in the second stage, classes within these schools are randomly chosen. All students in these selected classes take standardized math and science tests and provide comprehensive background information through questionnaires. We link fourth-grade students to the prevalence of autonomous schools in primary education and eighth-grade students to the prevalence of autonomous schools in lower secondary education.

Our first main outcome is student achievement, measured using TIMSS math and

[3] Details on the PISA data and the construction of the autonomy indices can be found in Online Appendix B.

[4] Note that while we can use the PISA data to characterize autonomous schools, we cannot use them for our main empirical analysis. The reason is that PISA samples 15-year-old students, which mostly attend the same grade. Our empirical strategy crucially depends on observing several grades per country and year, rendering an analysis with PISA data infeasible.

science scores. These scores are comparable across TIMSS waves and are reported as five plausible values, which are random draws from a posterior distribution.[5] To obtain unbiased coefficient estimates, we use the average of these five plausible values for each subject. For ease of interpretation, we standardize these averages to have a mean of zero and a standard deviation of one.

Our second main outcome is segregation by SES and immigrant status. We focus on these two dimensions due to our own findings on the differential impact of autonomous schools on achievement by SES and immigrant status, as well as documented differences in school choices along these lines in prior research (e.g. Böhlmark et al. 2016, Bertoni et al. 2020). We define SES based on the number of books in a student's home, with more than 100 books indicating high SES.[6] Immigrant status is determined using the student's country of birth, as reported in the student questionnaire.

We exploit TIMSS' sampling of entire classes within schools to measure segregation at the class level. This is particularly useful in education systems that track students within schools, as it captures both between-school and within-school segregation. We calculate two measures of segregation of high-SES students and immigrant students: the variance ratio index and the dissimilarity index. The variance ratio index captures the relative probability that a randomly selected classmate of a minority student belongs to the same minority group, while the dissimilarity index measures the proportion of minority students who would need to change classes to achieve equal representation. Both indices account for the overall student composition in a country, making them comparable across countries and time.[7] They range from 0 to 100, with higher values indicating greater segregation levels. We calculate the indices separately for each country, grade, and year.

[5] Plausible values are used in most international student assessments, see Jerrim et al. (2017).

[6] Books at home have been widely used as a proxy for SES in previous research, see for example Wößmann (2003). While parental education would have been an ideal additional measure of SES, this information is only available for eighth-grade students, and our regressions require a measure that is available for both fourth- and eighth-grade students. Notably, books at home are still highly correlated with parental education: among eighth-grade students in TIMSS 2019, 37 percent of those with more than 100 books at home have college-educated parents, compared to only 19 percent of those with fewer books.

[7] Online Appendix B provides additional details on the calculation of the segregation indices.

2.3 Sample selection, summary statistics, and trends in autonomous school attendance

Our identification strategy leverages within-country, between-grade variation in the share of students attending autonomous schools over time. Consequently, we select all countries that participated in at least two waves of TIMSS with both fourth- and eighth-grade students, and for which data on the share of students attending autonomous schools are available for these waves. Fifteen countries meet these criteria: Australia, Chile, Finland, Hungary, Ireland, Italy, Japan, South Korea, Lithuania, Norway, Russia, Slovenia, Sweden, Turkey, and the United Kingdom. The sample covers the years 2003 to 2019 and includes a total of 484,526 students across 106 country-grade-year cells.

Table 1 presents summary statistics for our sample. These statistics are computed using weights that adjust for individual sampling probabilities and give equal weight to all country-grade-year cells. In our sample, six percent of students are classified as immigrants, and roughly one-third have more than 100 books at home, indicating high SES. Across all years, about ten percent of students attend autonomous schools. The bottom of the table shows the means and standard deviations for the segregation indices, which are calculated at the country-grade-year level.[8] The means of these indices are broadly comparable to similar indices measuring school segregation of minority students at the metropolitan level in the United States (Monarrez et al. 2022).

Figure 2 illustrates the evolution of autonomous school attendance from 1999 to 2019, separately by country and grade level. Several countries experienced significant increases, most notably the United Kingdom, where lower-secondary autonomous school enrollment surged from zero to over 60 percent. Chile, Hungary, and Sweden also saw substantial growth, with smaller but notable increases in Australia, Finland, Norway, and Slovenia. Importantly, these trends differed by grade level: with the exception of Chile and Slovenia, the increase of autonomous school enrollment in lower secondary school outpaced the rise in primary school in all countries. Finally, six countries in our sample had no autonomous schools at either grade level throughout the entire sample period.

[8]There are only 80, rather than 106, observations for the segregation indices of immigrants because information on country of birth was not collected for fourth-grade students in TIMSS 2011.

Online Appendix Figures A.1 to A.6 present the corresponding trends in our outcome variables. For achievement, no clear overall trend emerges across the 15 countries; some show increasing trends, while others display decreasing trends. Similarly, trends in segregation remain relatively stable over this period for most countries, with a few exceptions. In the analysis below, we investigate whether these trends can be attributed to changes in autonomous school enrollment.

3 Empirical strategy

We estimate triple-differences specifications that exploit between-grade variation in the share of students attending autonomous schools within a given country and year. Intuitively, we examine whether an increase in autonomous-school enrollment at the secondary level, relative to the primary level, results in corresponding changes in the outcome variable between the two levels. When estimating effects on achievement, we use student-level data and the following specification:

$$A_{icgt} = \alpha^a + \beta^a ATS_{cgt} + X'_{icgt}\gamma^a + \delta^a_{cg} + \lambda^a_{ct} + \theta^a_{gt} + \varepsilon^a_{icgt}, \tag{1}$$

where i denotes students, c denotes countries, g denotes grades, and t denotes years. A_{icgt} is achievement in math or science and ATS_{cgt} is the share of students enrolled in autonomous schools. X_{icgt} is a vector containing the student characteristics listed in Table 1, which we interact with country dummies. δ_{cg} is a vector of country-grade fixed effects, λ_{ct} is a vector of country-year fixed effects, and θ_{gt} is a vector of grade-year fixed effects. When estimating this specification, we use weights that adjust for individual sampling probabilities and that give equal weight to all country-grade-year cells (unweighted estimates are qualitatively and quantitatively similar). We cluster standard errors at the country-grade level.

In contrast to achievement, segregation is a system-level outcome. Therefore, when estimating effects on segregation, we use data at the country-grade-year level and the following specification:

$$S_{cgt} = \alpha^s + \beta^s ATS_{cgt} + \delta^s_{cg} + \lambda^s_{ct} + \theta^s_{gt} + \varepsilon^s_{icgt}, \tag{2}$$

where S_{cgt} is the variance ratio index or the dissimilarity index measuring segregation of either high-SES students or immigrant students and the other variables and parameters are defined in a manner equivalent to those in Equation 1.

The identification of the effects of autonomous schools in Equations 1 and 2 relies on a comprehensive set of fixed effects. This approach is analogous to that used by Monarrez et al. (2022) in their analysis of how charter schools affect segregation in the United States and represents a novel application in an international setting. In what follows, we discuss the factors accounted for by this approach and the assumptions that remain.

First, and most importantly, by leveraging between-grade variation, we are able to include country-year fixed effects, λ^s_{ct}, which non-parametrically control for all grade-unspecific changes to a country's education system over time. This is an improvement over prior studies on international differences in achievement, which included country and year fixed effects separately but could not account for country-specific changes that may be correlated with both the treatment and the outcome (e.g. Brunello & Rocco 2013, Hanushek et al. 2013, Bergbauer et al. 2024).

Second, our approach addresses the potential concern that grade-specific education policies could influence both the prevalence of autonomous schools and student achievement. For instance, school assignment might involve catchment areas at the primary level but not at the secondary level in some countries. Such differences could affect the supply of autonomous schools and also influence achievement, for example through peer effects.[9] Similarly, student tracking differs among the countries in our sample: some countries track students after primary education, while others maintain comprehensive schools throughout lower secondary education. These differences have been shown to affect student outcomes (Hanushek & Wößmann 2006) and may also influence the prevalence of autonomous schools. Our regressions account for country-by-grade-specific institutional

[9] Although the peer effects literature (e.g. Carrell et al. 2013, Murphy & Weinhardt 2020) provides no definitive conclusion on how system-wide effects might operate through re-sorting, we use this as an illustrative example of how school assignment policies could relate to achievement.

features, including school assignment and tracking policies, by including country-by-grade fixed effects, δ^s_{cg}.

Finally, we account for any shocks that affect the prevalence of autonomous schools and student achievement in a specific grade and year by including grade-by-year fixed effects, θ^s_{gt}.

Conditional on the fixed effects mentioned above, the remaining variation occurs at the country-by-grade-by-year level. Our identification assumption is that there are no changes at this level that are correlated with both the prevalence of autonomous schools and the outcome. While it is difficult to think of potential confounders at this level, we conduct several robustness tests to validate our research design. Specifically, we present results from "long-difference" regressions, which use only the first and last observations for each country and therefore identify the effects from long-term trends rather than short-term (four-year) fluctuations. Additionally, we demonstrate that our results remain robust when excluding one country at a time, ensuring they are not driven by outliers.

4 Results

4.1 Treatment variation

One potential limitation of our identification strategy is that the many fixed effects may absorb much of the variation in the prevalence of autonomous schools and student achievement. This lack of conditional variation could result in imprecise estimates. To investigate this concern, we examine the residuals from a regression of our treatment variable, the share of students attending autonomous schools, on the fixed effects. Appendix Figure A.7 displays the distribution of residuals, revealing that changes in the prevalence of autonomous schools greater than 10 percentage points are still observed. As we will see below, this amount of conditional treatment variation is sufficient to estimate the effects of autonomous schools with meaningful statistical precision.

4.2 Effect on achievement: overall effect

Table 2 presents estimates of the aggregate effect of autonomous schools on math and science achievement. Columns 1 and 3 report results from regressions that omit student controls, revealing very small point estimates: a 10 percentage point (p.p.) increase in the share of students attending autonomous schools is estimated to raise math achievement by 0.007 standard deviations (SD) and decrease science achievement by 0.009 SD. Neither estimate is statistically significant at conventional levels. Columns 2 and 4 add student controls to these regressions, resulting in slightly more negative coefficients: a 10 p.p. increase in the share of students attending autonomous schools is now estimated to reduce math achievement by an insignificant 0.003 SD and decrease science achievement by 0.020 SD, which is statistically significant at the five percent level. Notably, these estimates rule out even modest positive effects: the upper bounds of the 95 percent confidence intervals suggest that a 10 p.p. increase in the share of students attending autonomous schools raises math achievement by only 0.013 SD and has no effect on science achievement. Overall, the key takeaway from these results is that the aggregate effect of autonomous schools on achievement appears to be very small and, if anything, negative, suggesting that autonomous schools have not created the anticipated ”rising tide that lifts all boats.”

One concern with the triple-differences design is that temporary and unobserved shocks at the country-by-grade-by-year level could affect our estimates. To ensure that our results are not driven by short-term fluctuations but instead reflect long-term developments within education systems, we estimate “long-differences” regressions as a robustness check (Online Appendix Table A.1). These regressions restrict the sample to only the first and last year in which each country is observed in the data.[10] Compared to our main results, the estimates from these regressions are even more negative, strengthening our conclusion that autonomous schools do not improve overall student achievement.

We also test the robustness of our findings with respect to sample selection choices. In Online Appendix Table A.2, we first exclude the six countries that never had autonomous

[10] Alternative specifications further restrict the sample to countries where the time span between the first and last observation is at least 12 years, see Online Appendix Table A.1.

schools from our sample. As expected, this leaves the estimates largely unchanged. We then proceed to exclude each of the remaining nine countries from the sample, one by one. This again does not alter the estimates significantly, with one exception: when we exclude England, the coefficients become more positive and much larger, but also much less precisely estimated. This is not entirely surprising. As Figure 2 shows, England contributes substantially to the between-grade variation in the share of students attending autonomous schools over time, which the triple-differences specification uses to estimate the effects. Moreover, the fact that the estimates become more positive when England is excluded is consistent with Eyles et al. (2017), who find that attending a primary school academy has a small negative effect on test scores in this country. We therefore conclude that the main findings remain robust.

How do these estimates compare to those in the existing literature? Chen & Harris (2023) use nationwide data from the United States and find that a 10 percentage point (p.p.) increase in charter market share raises English Language Arts achievement by 0.01 standard deviations (SD), with less robust results for math achievement. Gilraine et al. (2021) find that students exposed to charter school entry in North Carolina experience an average improvement in math achievement of 0.02 SD relative to untreated students. In the context of Chile, Hsieh & Urquiola (2006) find that voucher school enrollment, if anything, has a small negative effect on aggregate math achievement. Similarly, Böhlmark & Lindahl (2015) find that a 10 p.p. increase in the share of students attending free schools in Sweden is associated with a 1.7 percentile rank higher math and English achievement. In sum, the previous literature has found small effects of autonomous schools on achievement, ranging from negative to positive. Our results are fully consistent with these findings and provide a more general assessment using data from 15 countries.

4.3 Effect on achievement: heterogeneity

In Table 3, we examine whether the effect of autonomous schools on achievement varies by student background. Columns 1 and 2 split the sample by SES. A 10 p.p. increase in the share of autonomous schools is estimated to raise high-SES students' math achievement by a statistically significant 0.022 SD, with a smaller and insignificant 0.008 SD increase in science. In contrast, the effect for low-SES students is negative in both subjects, with a notable 0.027 SD decline in science scores. Columns 3 and 4 reveal similar differences by immigrant status, with consistently more positive estimates for natives than for immigrants across both subjects.

The bottom rows of Panels A and B show estimates of the implied change in the achievement gap between high- and low-SES students, as well as between native and immigrant students. In three out of four cases, these changes are both economically meaningful and statistically significant. For example, a 10 p.p. increase in the share of students attending autonomous schools increases the math achievement gap between high- and low-SES students by 0.033 SD, which corresponds to five percent of the sample mean. Similarly, the gap in math achievement between native and immigrant students increases by 0.050 SD, which corresponds to 14 percent of the sample mean.

We also test the sensitivity of our results to the definition of the SES variable, given that our classification of high SES as having more than 100 books at home is somewhat arbitrary. We explore heterogeneity using two alternative definitions of high SES: having more than 200 books at home, and having more books at home than the country median value. Online Appendix Table A.3 shows that our results are robust to these alternative definitions, with SES-related heterogeneity often becoming even more pronounced in these regressions.

4.4 Effects on segregation

Table 4 presents estimates of the effect of autonomous schools on segregation. Columns 1 and 2 focus on segregation by SES and reveal positive effects: a 10 p.p. increase in the share of students attending autonomous schools is estimated to raise the variance ratio index by 0.62 (approximately 0.14 SD) and the dissimilarity index by 0.89 (approximately 0.15 SD). Column 3 shows a similarly positive effect on segregation by immigrant status, with a coefficient of 0.043 (approximately 0.12 SD for a 10 p.p. increase) for the variance ratio index. In contrast, the estimated effect on the dissimilarity index in column 4 is close to zero and insignificant. Overall, these results indicate a non-negligible effect of autonomous schools on segregation.

We again test the robustness of these findings. Specifically, in Online Appendix Table A.4, we first exclude all countries that never had autonomous schools during our sample period, and then exclude the remaining countries from the sample one by one. Overall, the results hold. When excluding England, the variance radio index for immigrant students shows a larger positive effect but with lower precision. In contrast, excluding Hungary results in larger and statistically significant estimates for the variance ratio index and the dissimilarity index for high-SES students. Across all specifications, all estimates remain positive, confirming our main findings.

These results align well with the small existing literature on how autonomous schools affect segregation. Monarrez et al. (2022) find that charter schools in the United States modestly increase school segregation for Black, Hispanic, Asian, and White students. Their main finding is that a 10 p.p. increase in the share of students attending charter schools leads to about a 0.9 point increase in the variance ratio index for minority students. This corresponds to approximately 0.07 SD, which is slightly more than half the size of our estimates in SD terms. Notably, both Monarrez et al. (2022) and our study leverage variation across grades and regions in a triple-differences framework. In a more descriptive study, Böhlmark et al. (2016) show that free school penetration is associated with segregation by SES and immigrant status in Sweden. Overall, our results contribute to this small body of literature by estimating the effects of autonomous schools on segregation at an international level.

5 Conclusion

We examine the impact of autonomous schools – publicly funded institutions that operate more independently than government-run schools – on student achievement and school segregation across 15 countries over 16 years. Our analysis uses individual-level data on math and science achievement from TIMSS 2003 to 2019, merged with OECD data on the share of students attending autonomous schools by country, year, and grade. Additionally, we compute four measures of segregation by SES and immigrant status at the country-year-grade level, using detailed information on students' backgrounds and class composition.

To identify the causal effects of autonomous schools, we estimate triple-differences specifications, which exploit variation in the share of students attending autonomous schools between grades within a given country and year. Crucially, this strategy allows us to non-parametrically control for any system-wide changes in a country's education policy during our study period. This identification approach represents a significant improvement over previous studies on international differences in student achievement, which included country and year fixed effects separately but could not account for country-specific changes that may be correlated with both the treatment and the outcome.

Our findings show little evidence that autonomous schools improve overall student achievement. The estimates are precise enough to rule out even modest positive effects and, if anything, suggest negative effects in science. However, this null result masks important heterogeneity: high-SES students consistently benefit from autonomous schools, while low-SES and immigrant students experience negative effects. Rather than creating "a rising tide that lifts all boats," autonomous schools appear to widen existing achievement gaps by SES and immigrant status.

Consistent with these findings, we also observe that autonomous schools increase segregation by SES and immigrant status, as measured by both the variance ratio index and the dissimilarity index. This suggests that autonomous schools contribute to greater separation along socioeconomic and immigrant lines. Overall, our results indicate that, while autonomous schools do not improve overall achievement, they exacerbate existing inequalities in both academic outcomes and school composition.

References

Abdulkadiroğlu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J. & Pathak, P. A. (2011), 'Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots', *The Quarterly Journal of Economics* **126**(2), 699–748.

Angrist, J. D., Pathak, P. A. & Walters, C. R. (2013), 'Explaining Charter School Effectiveness', *American Economic Journal: Applied Economics* **5**(4), 1–27.

Bergbauer, A. B., Hanushek, E. A. & Woessmann, L. (2024), 'Testing', *Journal of Human Resources* **59**(2), 349–388.

Bertoni, M., Gibbons, S. & Silva, O. (2020), 'School Choice During a Period of Radical School Reform. Evidence from Academy Conversion in England', *Economic Policy* **35**(104), 739–795.

Bertoni, M., Heller-Sahlgren, G. & Silva, O. (2023), 'Free to improve? The Impact of Free School Attendance in England', CEP Discussion Paper 1946, Centre for Economic Performance, LSE.

Böhlmark, A., Holmlund, H. & Lindahl, M. (2016), 'Parental Choice, Neighbourhood Segregation or Cream Skimming? An Analysis of School Segregation after a Generalized Choice Reform', *Journal of Population Economics* **29**(4), 1155–1190.

Böhlmark, A. & Lindahl, M. (2015), 'Independent Schools and Long-Run Educational Outcomes: Evidence from Sweden's Large-Scale Voucher Reform', *Economica* **82**(327), 508–551.

Brunello, G. & Rocco, L. (2013), 'The Effect of Immigration on the School Performance of Natives: Cross Country Evidence Using PISA Test Scores', *Economics of Education Review* **32**, 234–246.

Carrell, S. E., Sacerdote, B. I. & West, J. E. (2013), 'From natural variation to optimal policy? the importance of endogenous peer group formation', *Econometrica* **81**(3), 855–882.

Chen, F. & Harris, D. N. (2023), 'The Market-Level Effects of Charter Schools on Student Outcomes: A National Analysis of School Districts', *Journal of Public Economics* **228**, 105015.

Cohodes, S. R. & Parham, K. S. (2021), Charter Schools' Effectiveness, Mechanisms, and Competitive Influence, *in* 'Oxford Research Encyclopedia of Economics and Finance', Oxford University Press.

Dobbie, W. & Fryer, R. G. (2015), 'The Medium-Term Impacts of High-Achieving Charter Schools', *Journal of Political Economy* **123**(5), 985–1037.

Duncan, O. D. & Duncan, B. (1955), 'A Methodological Analysis of Segregation Indexes', *American Sociological Review* **20**(2), 210–217.

Eyles, A. & Machin, S. (2019), 'The Introduction of Academy Schools to England's Education', *Journal of the European Economic Association* **17**(4), 1107–1146.

Eyles, A., Machin, S. & McNally, S. (2017), 'Unexpected School Reform: Academisation of Primary Schools in England', *Journal of Public Economics* **155**, 108–121.

Figlio, D. N., Hart, C. M. & Karbownik, K. (2021), Competitive Effects of Charter Schools, mimeo.

Gilraine, M., Petronijevic, U. & Singleton, J. D. (2021), 'Horizontal Differentiation and the Policy Effect of Charter Schools', *American Economic Journal: Economic Policy* **13**(3), 239–276.

Graham, B. S. (2018), 'Identifying and Estimating Neighborhood Effects', *Journal of Economic Literature* **56**(2), 450–500.

Hanushek, E. A., Link, S. & Woessmann, L. (2013), 'Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA', *Journal of Development Economics* **104**, 212–232.

Hanushek, E. A. & Wößmann, L. (2006), 'Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries', *The Economic Journal* **116**(510), C63–C76.

Hoxby, C. M. (2000), 'Does Competition among Public Schools Benefit Students and Taxpayers?', *American Economic Review* **90**(5), 1209–1238.

Hoxby, C. M. (2003), School Choice and school Productivity. Could School Choice be a Tide that Lifts all Boats?, *in* 'The economics of school choice', University of Chicago Press, pp. 287–342.

Hsieh, C.-T. & Urquiola, M. (2006), 'The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's Voucher Program', *Journal of Public Economics* **90**(8-9), 1477–1503.

Jerrim, J., Lopez-Agudo, L. A., Marcenaro-Gutierrez, O. D. & Shure, N. (2017), 'What Happens When Econometrics and Psychometrics Collide? An Example Using the PISA Data', *Economics of Education Review* **61**, 51–58.

Massey, D. S. & Denton, N. A. (1988), 'The Dimensions of Residential Segregation', *Social Forces* **67**(2), 281–315.

Monarrez, T., Kisida, B. & Chingos, M. (2022), 'The Effect of Charter Schools on School Segregation', *American Economic Journal: Economic Policy* **14**(1), 301–40.

Mumma, K. S. (2022), 'The Effect of Charter School Openings on Traditional Public Schools in Massachusetts and North Carolina', *American Economic Journal: Economic Policy* **14**(2), 445–474.

Murphy, R. & Weinhardt, F. (2020), 'Top of the Class: The Importance of Ordinal Rank', *The Review of Economic Studies* **87**(6), 2777–2826.

OECD (2018), 'OECD Handbook for Internationally Comparative Education Statistics 2018 - Concepts, Standards, Definitions and Classifications', *OECD Publishing* .

Schneeweis, N. (2011), 'Educational Institutions and the Integration of Migrants', *Journal of Population Economics* **24**(4), 1281–1308.

Tiebout, C. M. (1956), 'A Pure Theory of Local Expenditures', *Journal of Political Economy* **64**, 416–416.

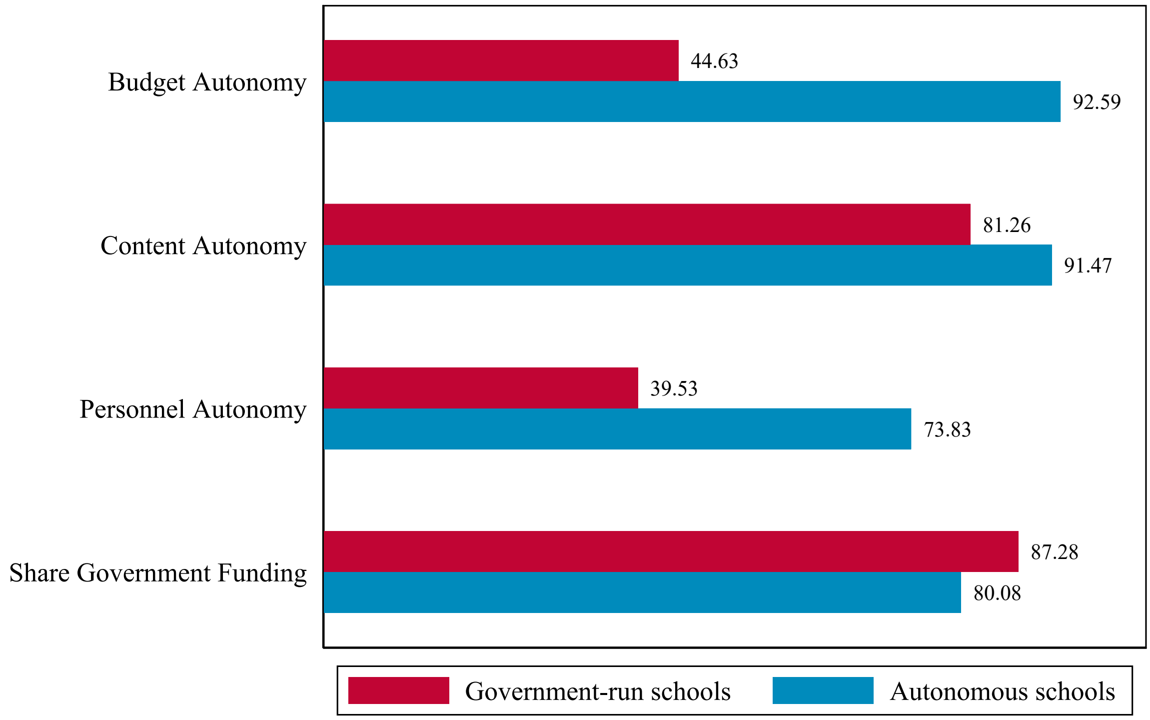
West, M. R. & Woessmann, L. (2010), 'Every Catholic Child in a Catholic School: Historical Resistance to State Schooling, Contemporary Private Competition and Student Achievement Across Countries', *The Economic Journal* **120**(546), F229–F255.

Winters, M. A. (2012), 'Measuring the Effect of Charter Schools on Public School Student Achievement in an Urban Environment: Evidence from New York City', *Economics of Education Review* **31**(2), 293–301.

Wößmann, L. (2003), 'Schooling Resources, Educational Institutions and Student Performance: the International Evidence', *Oxford Bulletin of Economics and Statistics* **65**(2), 117–170.

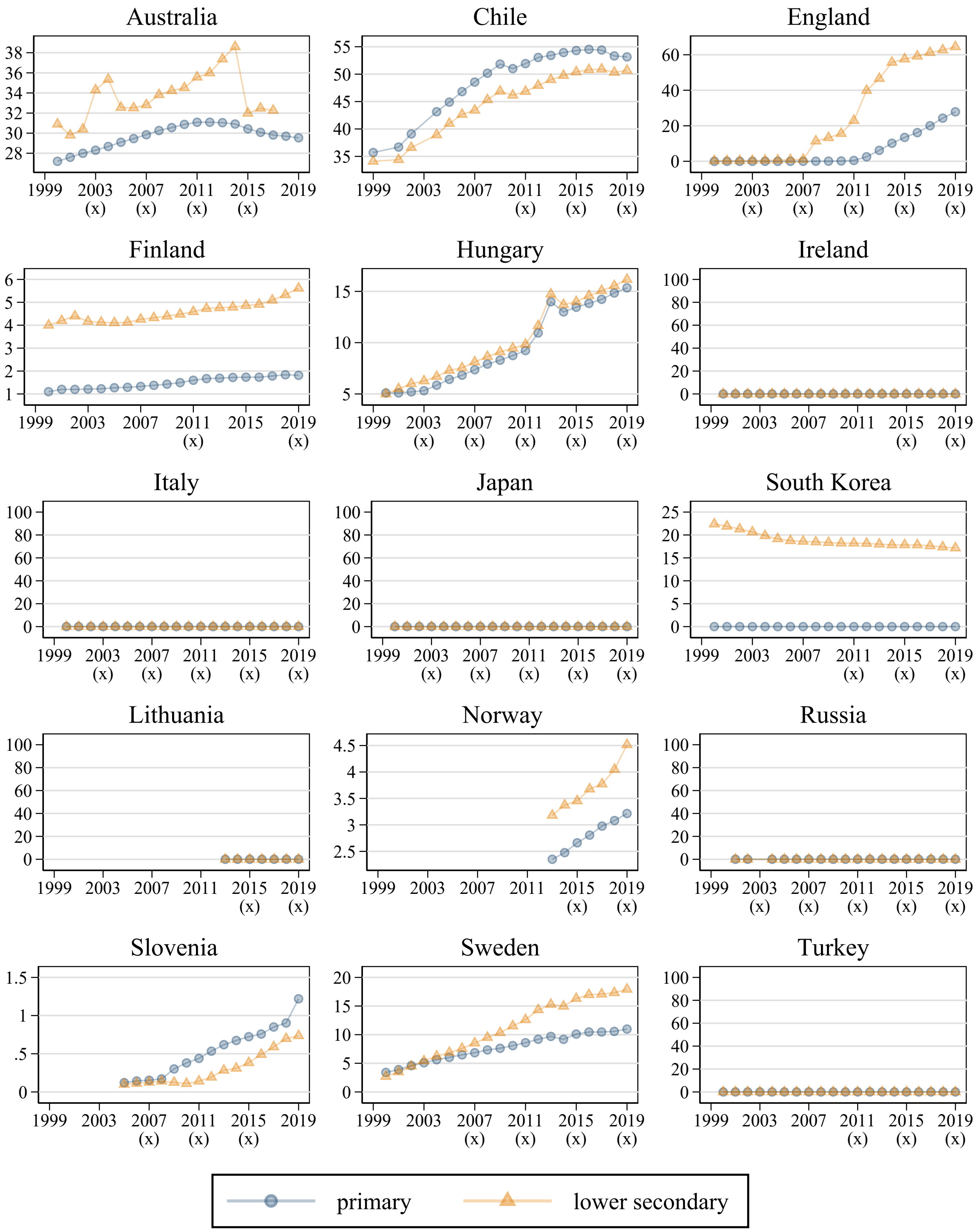
Figures and Tables

Figure 1: Comparing government-run and autonomous schools in PISA 2015



Notes: The figure shows means of autonomy indices and the share of government funding, separately for autonomous and government-run schools. The three indices measuring personnel, budget, and academic-content autonomy are constructed following Hanushek et al. (2013). We classify schools as autonomous schools if principals report the school to be privately managed but receiving at least 50 percent of funding from the government, following the definition by the OECD (2018). Details on the PISA data and the construction of the indices can be found in Online Appendix B.

Figure 2: Trends in the share of students attending autonomous schools, by country and grade level



Notes: The figure shows trends in the share of students attending autonomous schools, separately by country and for primary and lower secondary education. (x) indicates years in which the country participated in TIMSS with both fourth and eighth grade; these are the years included in the analysis sample.

Table 1: Summary statistics

	Mean	SD	N
Student characteristics			
Female	0.49	0.50	484,423
Age	12.35	2.06	483,424
Immigrant	0.06	0.24	418,694
Books at home:			
11–25	0.24	0.43	476,495
26–100	0.31	0.46	476,495
101–200	0.17	0.37	476,495
more than 200	0.16	0.36	476,495
Autonomous schools			
Share autonomous	9.70	15.75	484,526
Student achievement			
Math score	0.00	1.00	484,526
Science score	0.00	1.00	484,526
Segregation (country-grade-year level)			
High SES: variance ratio index	14.49	4.46	106
High SES: dissimilarity index	33.92	5.93	106
Immigrant status: variance ratio index	9.48	3.71	80
Immigrant status: dissimilarity index	56.68	14.89	80

Notes: The table shows means and standard deviations and the number of students observed with each variable for the 484,526 students included in the analysis sample. Statistics are computed using weights that adjust for individual sampling probabilities and that give equal weight to all country-grade-year cells. Information on country of birth was not collected in fourth grade in TIMSS 2011, which explains the larger number of missing values for the immigrant variable. The omitted category for books at home is 0–10 books. The bottom four rows show statistics for segregation indices, which are measured at the level of 106 country-grade-year cells; the number of observations for the indices related to immigrant status is reduced because of the missing information on country of birth in 2011.

Table 2: Effect of autonomous schools on student achievement

	Math		Science	
	(1)	(2)	(3)	(4)
Share autonomous	0.0007	–0.0003	–0.0009	–0.0020**
	(0.0008)	(0.0008)	(0.0011)	(0.0009)
Observations	484,526	484,526	484,526	484,526
Student controls	No	Yes	No	Yes
Country-grade fixed effects	Yes	Yes	Yes	Yes
Country-year fixed effects	Yes	Yes	Yes	Yes
Grade-year fixed effects	Yes	Yes	Yes	Yes

Notes: The table shows estimates of the effect of autonomous schools on student achievement in math and science. Student controls include age and dummies for female, books at home, and immigrant, all of which are interacted with country dummies. Regressions use weights that adjust for individual sampling probabilities and that give equal weight to all country-grade-year cells. Standard errors in parentheses are clustered at the country-grade level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 3: Effect of autonomous schools on student achievement, heterogeneity

	By SES		By immigrant status	
	High	Low	Native	Immigrant
	(1)	(2)	(3)	(4)
Panel A: math achievement				
Share autonomous	0.0022***	–0.0010	0.0011*	–0.0039*
	(0.0006)	(0.0009)	(0.0006)	(0.0019)
Observations	159,262	317,233	325,955	21,705
Change in gap	0.0033***		0.0050***	
	(0.0007)		(0.0018)	
Panel B: science achievement				
Share autonomous	0.0008	–0.0027***	–0.0012	–0.0020
	(0.0010)	(0.0009)	(0.0009)	(0.0016)
Observations	159,262	317,233	325,955	21,705
Change in gap	0.0034***		0.0008	
	(0.0005)		(0.0012)	
Student controls	Yes	Yes	Yes	Yes
Country-grade fixed effects	Yes	Yes	Yes	Yes
Country-year fixed effects	Yes	Yes	Yes	Yes
Grade-year fixed effects	Yes	Yes	Yes	Yes

Notes: The table shows estimates of the effect of autonomous schools on student achievement in math and science, separately by socioeconomic status and immigrant status. High (low) SES is defined as having more than 100 (at most 100) books at home. The combined sample sizes in columns 1-2 and columns 3-4 differ from those in Table 2 because of missing values on the SES and immigrant variables. Student controls include age and dummies for female, books at home (not in columns 1-2), and immigrant (not in columns 3-4), all of which are interacted with country dummies. The bottom rows in Panels A and B show the change in the achievement gap between groups, which is estimated using seemingly unrelated regression. Sample means of these gaps are: high-SES-low-SES math (science) gap: 0.5954 (0.6363); native-immigrant math (science) gap: 0.3590 (0.4526). Regressions use weights that adjust for individual sampling probabilities and that give equal weight to all country-grade-year cells. Standard errors in parentheses are clustered at the country-grade level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 4: Effect of autonomous schools on segregation

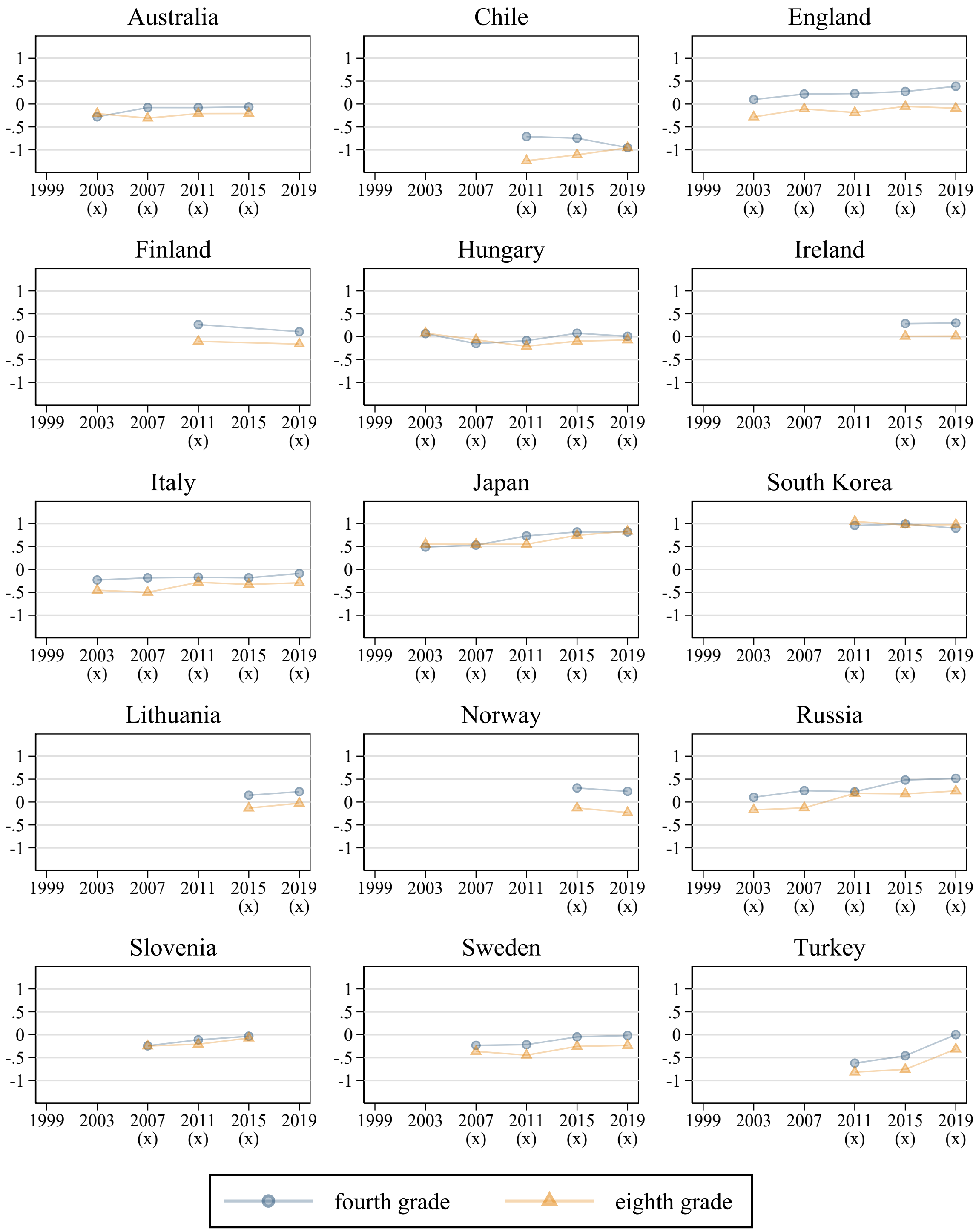
	By SES		By immigrant status	
	Variance ratio index (1)	Dissimilarity index (2)	Variance ratio index (3)	Dissimilarity index (4)
Share autonomous	0.062**	0.089*	0.043*	0.004
	(0.030)	(0.046)	(0.025)	(0.062)
Observations	106	106	80	80
Country-grade fixed effects	Yes	Yes	Yes	Yes
Country-year fixed effects	Yes	Yes	Yes	Yes
Grade-year fixed effects	Yes	Yes	Yes	Yes

Notes: The table shows estimates of the effect of autonomous schools on segregation. For this analysis, the data are collapsed at the country-grade-year level. The sample size is smaller in columns 3-4 because information on country of birth was not collected in fourth grade in TIMSS 2011. See text for details on how the four outcome measures of segregation are computed. Standard errors in parentheses are clustered at the country-grade level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

– ONLINE APPENDIX –

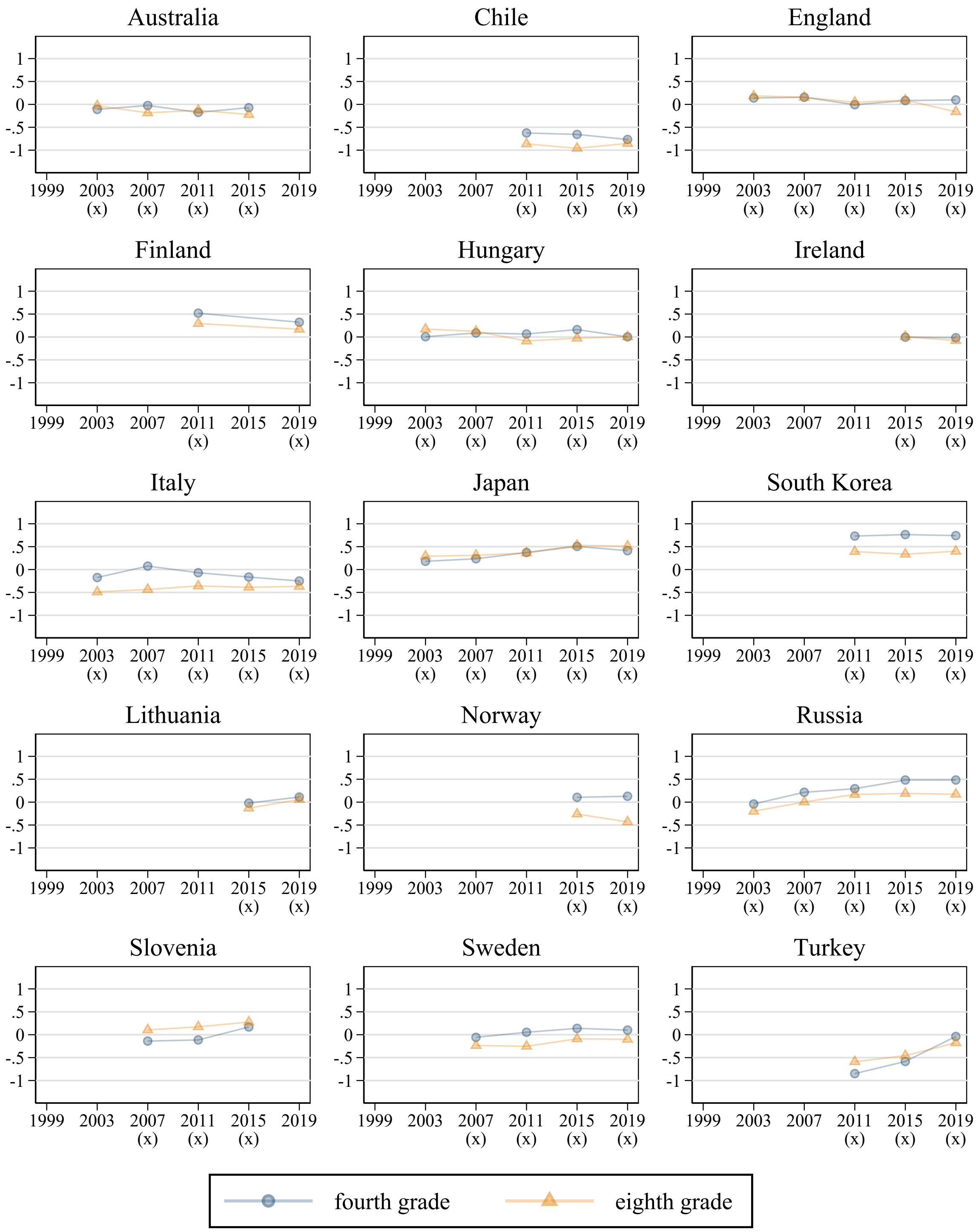
A Additional figures and tables

Online Appendix Figure A.1: Trends in math achievement, by country and grade level



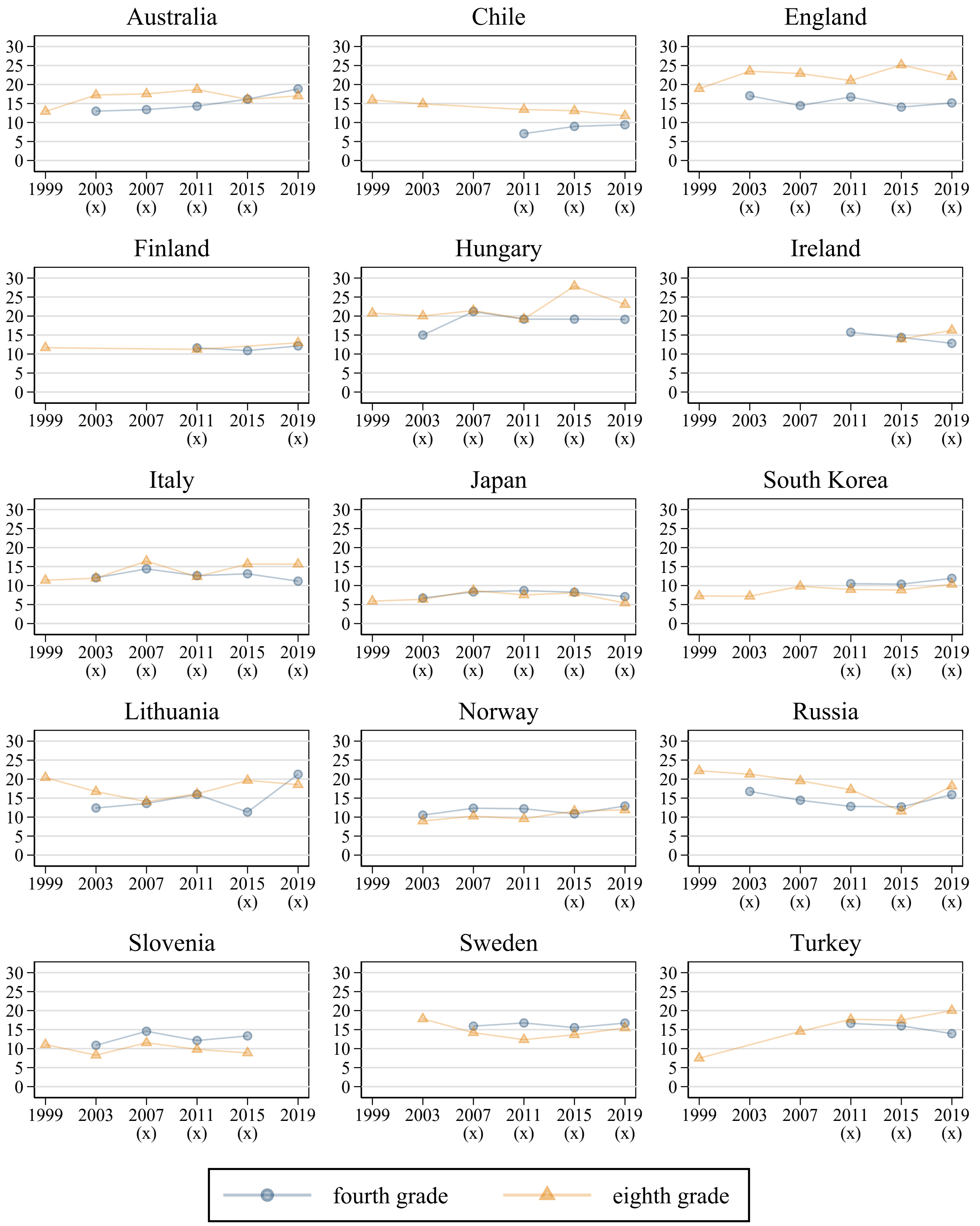
Notes: The figure shows trends in math achievement, separately by country and grade level. (x) indicates years included in the analysis sample.

Online Appendix Figure A.2: Trends in science achievement, by country and grade level



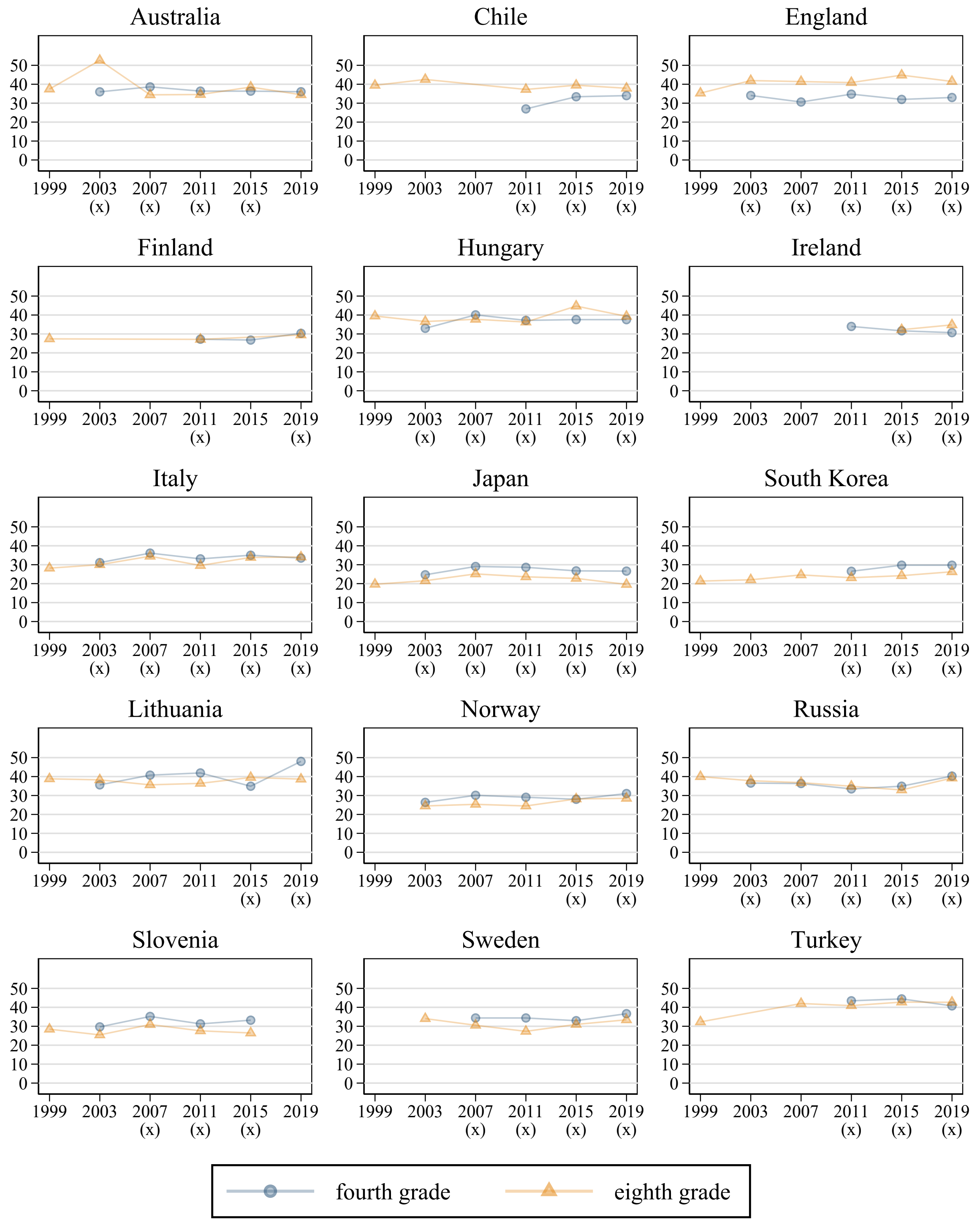
Notes: The figure shows trends in science achievement, separately by country and grade level. (x) indicates years included in the analysis sample.

Online Appendix Figure A.3: Trends in variance ratio index for high-SES students, by country and grade level



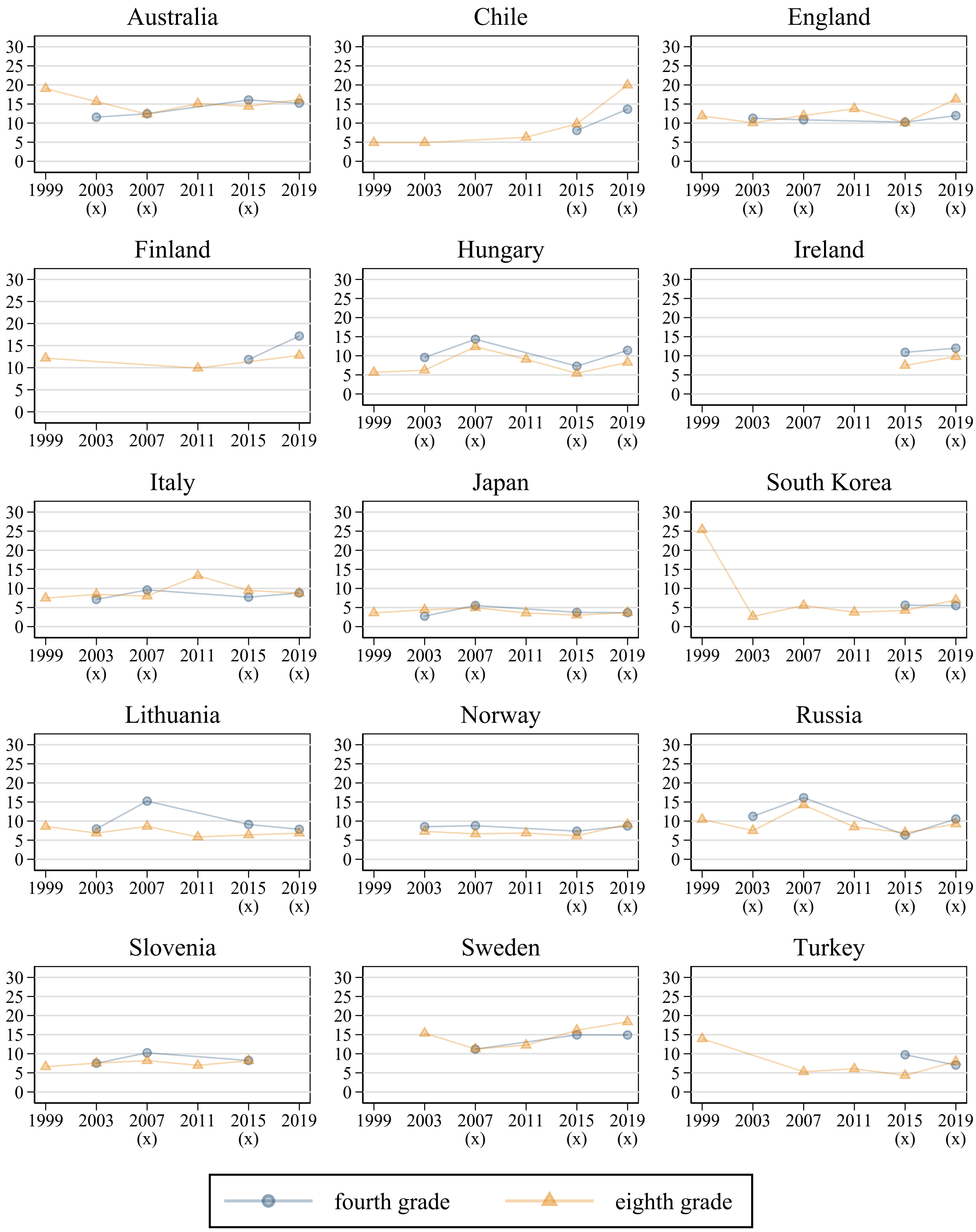
Notes: The figure shows trends in segregation of high-SES students, separately by country and grade level. Segregation is measured using the variance ratio index. (x) indicates years included in the analysis sample.

Online Appendix Figure A.4: Trends in dissimilarity index for high-SES students, by country and grade level



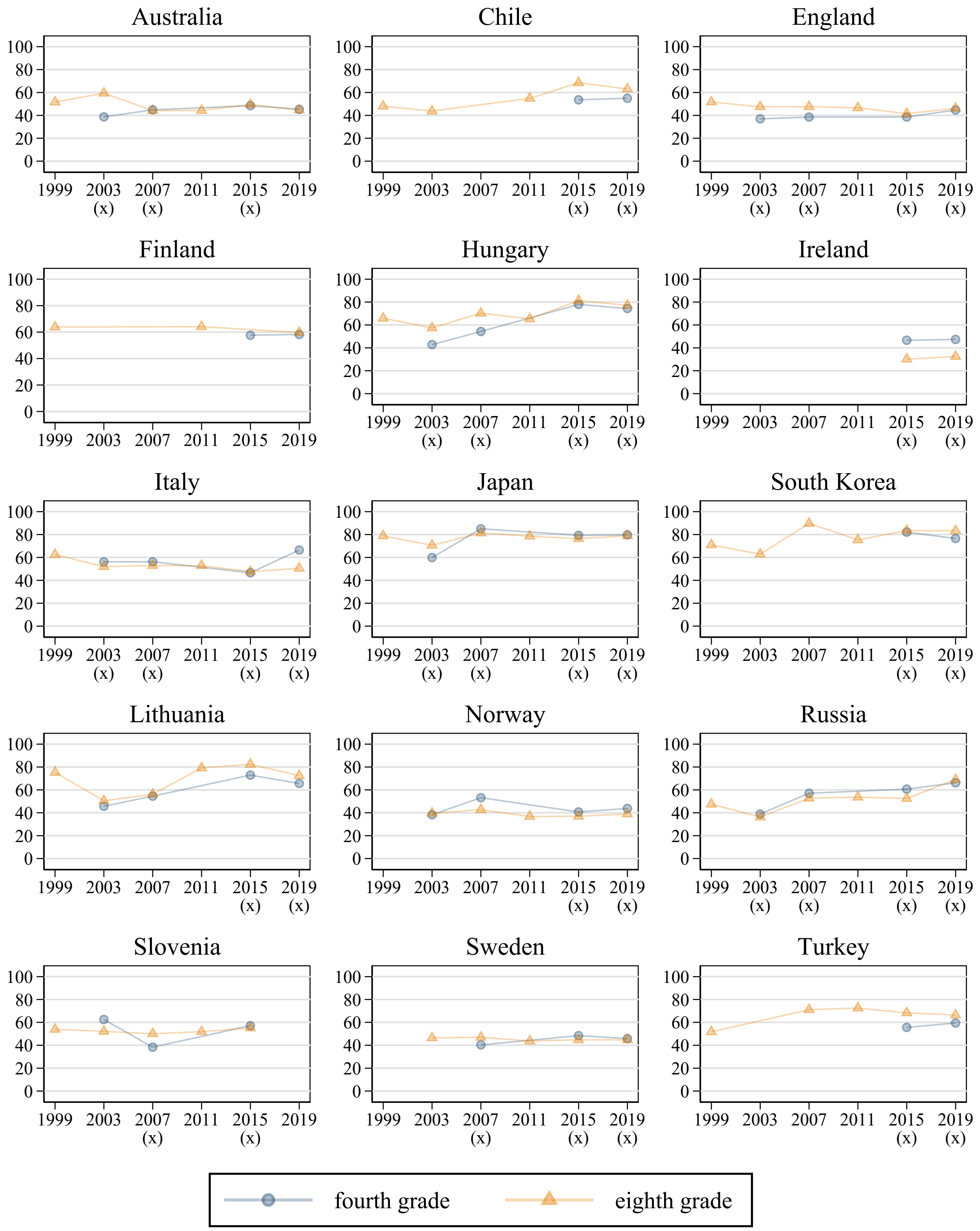
Notes: The figure shows trends in segregation of high-SES students, separately by country and grade level. Segregation is measured using the dissimilarity index. (x) indicates years included in the analysis sample.

Online Appendix Figure A.5: Trends in variance ratio index for immigrant students, by country and grade level



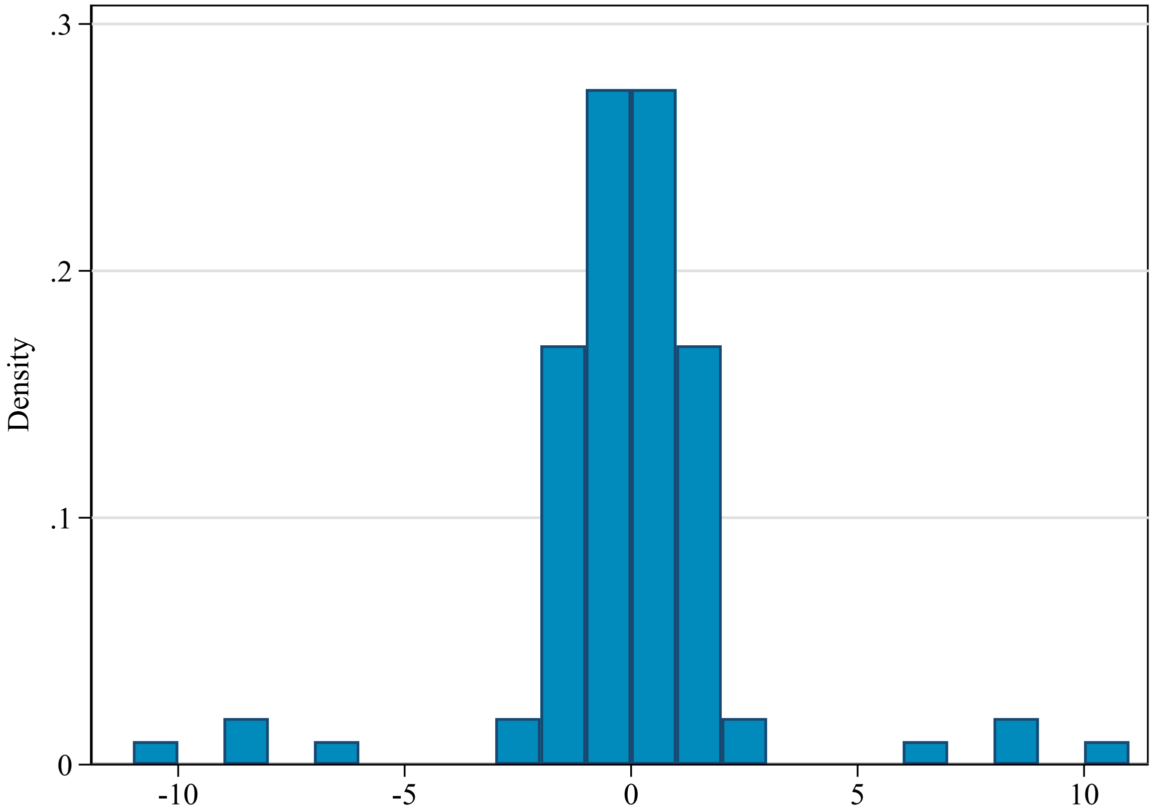
Notes: The figure shows trends in segregation of immigrant students, separately by country and grade level. Segregation is measured using the variance ratio index. (x) indicates years included in the analysis sample. The year 2011 is not included in the analysis sample for any country because information on country of birth was not collected in fourth grade in that wave. Finland is dropped from the sample because it only has one year of data.

Online Appendix Figure A.6: Trends in dissimilarity index for immigrant students, by country and grade level



Notes: The figure shows trends in segregation of immigrant students, separately by country and grade level. Segregation is measured using the dissimilarity index. (x) indicates years included in the analysis sample. The year 2011 is not included in the analysis sample for any country because information on country of birth was not collected in fourth grade in that wave. Finland is dropped from the sample because it only has one year of data.

Online Appendix Figure A.7: Residual Variation in share autonomous after taking out fixed effects



Notes: This figure shows the distribution of residuals from a regression of the share of students attending autonomous schools on country-by-year, country-by-grade, and grade-by-year fixed effects in our estimation sample.

Online Appendix Table A.1: Effect of autonomous schools on student achievement, long-differences estimates

	Math		Science	
	(1)	(2)	(3)	(4)
Share autonomous	–0.0000	–0.0015	–0.0040***	–0.0061***
	(0.0011)	(0.0011)	(0.0013)	(0.0011)
Observations	269,624	122,802	269,624	122,802
Student controls	Yes	Yes	Yes	Yes
Country-grade fixed effects	Yes	Yes	Yes	Yes
Country-year fixed effects	Yes	Yes	Yes	Yes
Grade-year fixed effects	Yes	Yes	Yes	Yes

Notes: The table shows estimates of the effect of autonomous schools on student achievement in math and science. In columns 1 and 3, the sample is restricted to the first and last available TIMSS wave for each country. In columns 2 and 4, the sample is further restricted to those countries for which the time between the first and last available TIMSS wave is at least 12 years. Student controls include age and dummies for female, books at home, and immigrant, all of which are interacted with country dummies. Regressions use weights that adjust for individual sampling probabilities and that give equal weight to all country-grade-year cells. Standard errors in parentheses are clustered at the country-grade level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Online Appendix Table A.2: Effect of autonomous schools on student achievement, samples excluding specific countries

	All with	Countries excluded from the sample: all without variation and additionally								
	variation	Australia	Chile	England	Finland	Hungary	S. Korea	Norway	Slovenia	Sweden
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: math achievement										
Share autonomous	0.0003	–0.0005	0.0003	0.0073	0.0003	0.0008*	0.0003	0.0003	0.0006*	0.0002
	(0.0004)	(0.0004)	(0.0005)	(0.0103)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0004)
Observations	284,122	236,736	254,808	246,546	265,614	239,891	256,890	266,570	258,119	247,802
Panel B: science achievement										
Share autonomous	–0.0005	–0.0008*	–0.0005	0.0086*	–0.0005	–0.0001	–0.0005	–0.0005*	–0.0007**	–0.0002
	(0.0003)	(0.0004)	(0.0004)	(0.0040)	(0.0003)	(0.0005)	(0.0004)	(0.0003)	(0.0003)	(0.0004)
Observations	284,122	236,736	254,808	246,546	265,614	239,891	256,890	266,570	258,119	247,802
Student controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-grade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table shows estimates of the effect of autonomous schools on student achievement in math and science for different subsamples of countries. Column 1 drops the following six countries with no variation in share autonomous from the sample: Ireland, Italy, Japan, Lithuania, Russia, and Turkey. Columns 2-10 additionally drop individual remaining countries as indicated in the column headers. Student controls include age and dummies for female, books at home, and immigrant, all of which are interacted with country dummies. Regressions use weights that adjust for individual sampling probabilities and that give equal weight to all country-grade-year cells. Standard errors in parentheses are clustered at the country-grade level. * p<0.10, ** p<0.05, *** p<0.01.

Online Appendix Table A.3: Effect of autonomous schools on student achievement, heterogeneity using alternative definitions of SES

	Highest category		Above country median	
	High	Low	High	Low
	(1)	(2)	(3)	(4)
Panel A: math achievement				
Share autonomous	0.0048***	–0.0008	0.0023***	–0.0011
	(0.0009)	(0.0008)	(0.0006)	(0.0009)
Observations	77,908	398,587	169,451	307,044
Change in gap	0.0055***		0.0034***	
	(0.0008)		(0.0007)	
Panel B: science achievement				
Share autonomous	0.0022	–0.0023**	0.0009	–0.0027***
	(0.0014)	(0.0009)	(0.0010)	(0.0009)
Observations	77,908	398,587	169,451	307,044
Change in gap	0.0044***		0.0035***	
	(0.0008)		(0.0005)	
Student controls	Yes	Yes	Yes	Yes
Country-grade fixed effects	Yes	Yes	Yes	Yes
Country-year fixed effects	Yes	Yes	Yes	Yes
Grade-year fixed effects	Yes	Yes	Yes	Yes

Notes: The table shows estimates of the effect of autonomous schools on student achievement in math and science, separately by socioeconomic status. In columns 1-2, high (low) SES is defined as having more than 200 (at most 200) books at home. In columns 3-4, high (low) SES is defined as having more than (at most) the country-specific median number of books at home. The bottom rows in Panels A and B show the change in the achievement gap between groups, which is estimated using seemingly unrelated regression. Sample means of these gaps are 0.56 (0.61) in math (science) when using the highest category classification, and 0.45 (0.53) in math (science) when using the above country median classification. Student controls include age and dummies for female and immigrant, all of which are interacted with country dummies. Regressions use weights that adjust for individual sampling probabilities and that give equal weight to all country-grade-year cells. Standard errors in parentheses are clustered at the country-grade level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Online Appendix Table A.4: Effect of autonomous schools on segregation, samples excluding specific countries

	All with	Countries excluded from the sample: all without variation and additionally								
	variation	Australia	Chile	England	Finland	Hungary	S. Korea	Norway	Slovenia	Sweden
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: variance ratio index for high-SES students										
Share autonomous	0.063	0.034	0.060	0.125	0.065	0.103***	0.062	0.063	0.058	0.063
	(0.041)	(0.037)	(0.041)	(0.498)	(0.043)	(0.020)	(0.045)	(0.041)	(0.051)	(0.047)
Observations	62	54	56	52	58	52	56	58	56	54
Panel B: dissimilarity index for high-SES students										
Share autonomous	0.117	0.033	0.111	0.135	0.116	0.207***	0.115	0.117	0.101	0.115
	(0.085)	(0.042)	(0.086)	(0.428)	(0.086)	(0.045)	(0.090)	(0.085)	(0.090)	(0.089)
Observations	62	54	56	52	58	52	56	58	56	54
Panel C: variance ratio index for immigrant students										
Share autonomous	0.048	0.005	0.048	0.656*	0.048	0.046	0.048	0.048	0.062*	0.053
	(0.037)	(0.033)	(0.038)	(0.360)	(0.037)	(0.053)	(0.038)	(0.037)	(0.033)	(0.036)
Observations	44	38	40	36	44	36	40	40	40	38
Panel D: dissimilarity index for immigrant students										
Share autonomous	0.129***	0.110***	0.129***	0.232	0.129***	0.151***	0.135***	0.128***	0.105***	0.129***
	(0.026)	(0.020)	(0.024)	(0.893)	(0.026)	(0.049)	(0.031)	(0.026)	(0.011)	(0.031)
Observations	44	38	40	36	44	36	40	40	40	38
Student controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-grade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table shows estimates of the effect of autonomous schools on segregation for different subsamples of countries. Column 1 drops the following six countries with no variation in share autonomous from the sample: Ireland, Italy, Japan, Lithuania, Russia, and Turkey. Columns 2-10 additionally drop individual remaining countries as indicated in the column headers. The sample size is smaller in panels C and D because information on country of birth was not collected in fourth grade in TIMSS 2011. Standard errors in parentheses are clustered at the country-grade level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

B Data Appendix

This appendix further describes our data sources, variable definitions, and the construction of our estimation sample.

B.1 Data sources

B.1.1 Share of students in autonomous schools

We obtain data on the share of students enrolled in autonomous schools from the OECD Education at a Glance (EAG) database, which provides internationally comparable indicators related to education outcomes, investments, and policy contexts. Importantly for our analysis, one of these indicators reports the enrollment of students across different types of educational institutions. In particular, the data include the distribution of students across three types of institutions: public, government-dependent private, and independent-private, separately for three levels of education (primary, lower-secondary, upper-secondary). The relevant category for our analysis are government-dependent private institutions, which are defined as educational institutions for which a private organization has ultimate control, but at least half of the funding is received from the government (OECD, 2018).

For the years 1999-2013, the EAG data are published in the annual EAG reports. The data are reported with a two year lag, which means that data for the time period 1999-2013 are published in the EAG reports from the years 2001-2015. The relevant tables are: Table C1.4. in the EAG reports 2001, 2011, 2012, and 2013; Table C2.4. in 2002, 2003, 2004, 2008; Table D5.1. in 2005; Table C2.9. in 2007; Table C1.5. in 2009, 2010, 2014; and Table C1.4a. in 2015. Data for the year 2014 and onward are no longer included in the EAG reports, but are instead available online (on https://stats.oecd.org/; table 'Enrolment by gender, programme orientation, mode of study and type of institution' with code EAG ENRL SHARE CATEGORY).

B.2 PISA data

We use school-level data from the 2015 wave of the OECD's Programme for International Student Assessment (PISA) to characterize government-dependent private schools. We follow the OECD (2018) and classify schools as government-dependent private if the school is a private school (variable SC013Q01TA) and receives at least 50 percent of its funding from the government (variable SC016Q01TA). We construct three autonomy indices following Hanushek et al. (2013): personnel autonomy, budget autonomy, and academic-content autonomy. The three indices are based on six domains that capture different areas of decision-making authority in the school: 1. Selecting teachers for hire (variables SC010Q01TA - SC010Q01TE), 2. Establishing teachers' starting salary (variables SC010Q03TA - SC010Q03TE), 3. Formulating the school budget (variables SC010Q05TA - SC010Q05TE), 4. Deciding which courses are offered (variables SC010Q12TA - SC010Q12TE), 5. Determining course content (variables SC010Q11TA - SC010Q11TE), 6. Choosing which textbooks are used (variables SC010Q10TA - SC010Q10TE).

For each domain, the principal reports whether decisions are made by the principal, teachers, the school governing board, the regional or local education authority, or the national education authority. Principals could tick as many options as appropriate. We classify schools as having autonomy in the given domain if within-school authorities (principal, teachers, school governing board) hold the sole responsibility for this domain. If any outside education authority (the regional or local education authority or the national education authority) shares responsibility, the school is not considered as having autonomy in this domain. We then aggregate the domain autonomy to construct the three indices: we use domains 1. and 2. for the personnel autonomy index, domain 3. for the budget autonomy index, and domain 4.-6. for the academic-content autonomy index.

B.2.1 Student achievement and segregation

To measure achievement and segregation, we use individual-level data from TIMSS. TIMSS is an international student assessment administrated by the International Association for the Evaluation of Educational Achievement every four years since 1995. The study draws nationally representative samples of fourth- and eighth-grade students through a two-stage clustered sampling design, where schools are selected in the first stage and classes within these schools are randomly sampled in the second stage. All students in the selected classes participate in the assessment and provide detailed background information via questionnaires. Test scores from the assessments in both math and science are reported as five plausible values. This reflects the fact that TIMSS tests each student only on a subset of questions and uses these answers and Item Response Theory to compute a distribution of test scores for each student. The plausible values are random draws from this posterior distribution. We use the mean of the five plausible values in math and science as measures of achievement and use the rich information on the social composition of the sampled classes to construct indices of segregation at the class level.

B.3 Measuring segregation

The variance ratio index is commonly used by economists to study school or neighborhood segregation (Graham 2018, Monarrez et al. 2022) and is defined as follows. Let the number of minority students in class $i = 1, ..., N$ be denoted by m_i and the total class size p_i. M is the total number of minority students in the country, and Q the overall share of minority students in the country. Then the index is calculated by:

$$VR = \frac{Isolation - Q}{1 - Q} \quad \text{with} \quad Isolation = \sum_{i=1}^{N} \frac{m_i}{p_i} \cdot \frac{m_i}{M} \tag{3}$$

As evident from the equation, the variance ratio index builds on an index of isolation, which describes the probability that a randomly selected classmate of a minority student is from this minority group, too. In a perfectly segregated school system, the isolation index will be equal to one, meaning that all classmates of a minority student are also members of this group. In contrast, in absence of segregation, the isolation index equals the overall share of minority students in the country. The variance ratio index is obtained by scaling the isolation index by the overall share of minority students in the country. Ultimately, the index measures the excess isolation of minority students compared to a perfectly integrated school system relative to the excess isolation of a perfectly segregated school system. The index is bounded between zero and one, with higher values indicating higher degrees of segregation. To facilitate interpretation of our results, in our analysis we scale the index to be between zero and 100.

The dissimilarity index is the most widely used measure of segregation in the sociology literature (Duncan & Duncan 1955, Massey & Denton 1988) but has also been applied by economists (Graham 2018, Monarrez et al. 2022). It is defined as follows:

$$DI = \frac{1}{2} \sum_{i=1}^{N} \left| \frac{m_i}{M} - \frac{r_i}{R} \right| \tag{4}$$

Here, m_i denote again the number of minority students in class i and M the total number of minority students in the country. r_i and R are the number of students who belong to the majority group in class i and the country as whole, respectively. Intuitively, the index measures the proportion of minority students who need to change classes to ensure an even distribution across classrooms, relative to the share of students that would need to change when there is perfect segregation (Graham 2018). If there is no segregation, the relative share of minority and other students will be the same in all classes (which means $\frac{m_i}{M} = \frac{r_i}{R}$ for all i), so the dissimilarity index will be equal to zero. If, on the other hand, there exists perfect segregation, minority students are in classes without any students from the majority group (which means in a given class either m_i or r_i is equal to zero), so that the index is equal to one. Like with the variance ratio index, in our analysis we scale the dissimilarity index to lie between zero and 100.

We compute our segregation measures at the class level for both fourth- and eighth-grade students. In doing so, we use the appropriate TIMSS sampling weights in order to ensure that our segregation measures are nationally representative.[11]

B.4 Sample construction

In our empirical analysis, we employ a triple-differences model to estimate the effects of autonomous schools on achievement and segregation. This identification strategy relies on within-country, between-grade variation over time. Therefore, we need to restrict our sample to countries for which we have data on both the primary and lower-secondary level for at least two time periods. This means that we restrict the sample to countries that participated in at least two waves of the TIMSS assessment with both fourth- and eighth grade students, and for which data on the share of students in autonomous schools is available for the years of their TIMSS participation.

In addition, we restrict the sample as follows:

- We exclude observations from Norway before 2013 because the share of students in autonomous schools cannot be clearly identified: before 2013, Norway did not separately report enrollment shares for independent-private and government-dependent private institutions but instead had these two types of institutions pooled in one category.

- We exclude the United States because the National Center for Education Statistics (NCES), the agency that reports education statistics to the OECD, does not consider charter schools as government-dependent private schools, but instead includes them in the category of public institutions. We received this information from the NCES after requesting information about charter school categorization in summer 2022.

- We exclude New Zealand because autonomous schools were first introduced, but shortly after disestablished within our study period. When they were disestablished, schools previously operating as autonomous schools were not closed but instead reorganized as public schools. This implies that while the share of students in

[11] TIMSS samples representative samples of students, not classes. We follow Schneeweis (2011) and weight a class with the sum of all student weights within this class.

autonomous schools in our data would be zero for the time period after the abolition, both the segregation of students and their achievement could still be influenced by the fact that some schools operated as autonomous schools before, which could bias our estimates.

Moreover, we apply the following imputations:

- We impute the share of students in autonomous schools to zero if data on autonomous schools in the EAG data set is coded as missing with either one of the following two flags: a. category does not apply or n. magnitude is negligible or zero; and if at the same time the share of students enrolled in the two other categories (public institutions and independent private institutions) sums up to 100 percent.

- We impute the share of students attending autonomous schools in the Russian Federation in the year 2003 for both primary and lower-secondary institutions. As the data are available for two preceding years and all subsequent years until 2019, we take the average of the enrollment share in 2004 and 2002 to impute the data for the year 2003. Since the share of students in autonomous schools in the Russian Federation is zero throughout our study period, we impute a share of zero for the year 2003.

- The EAG data base does not report statistics separately for countries within the United Kingdom. However, the reported share of students in autonomous schools across the United Kingdom mimics closely the share of students enrolled in academies in England.[12] We thus impute the share of students in autonomous schools in England as the corresponding share in the United Kingdom.

To obtain our final sample, we merge the TIMSS data with the EAG data on autonomous schools. We merge TIMSS fourth-grade data with the data on autonomous schools at the primary level, and TIMSS eighth-grade data with data on autonomous schools at the lower-secondary level. Our final analysis sample consists of 15 countries over the time period from 2003 to 2019, which implies that it covers five TIMSS waves. Note that not every country is participating with both grade levels in each TIMSS wave, which implies that the number of country-grade-wave cells in our sample is 106 instead of 150 (which would be the case with 15 countries $\times$ 5 waves $\times$ 2 grades).

[12] See https://www.gov.uk/government/statistics/schools-pupils-and-their-characteristics-january2019.