The Effect of School Starting Age on Student's Achievement: Evidence From a Large Sample of OECD Countries

Team 30

Abstract

We study the effect of school starting age on educational achievement. Therefore, we use three waves of the PISA study and a large sample of OECD countries to document the development of test scores over time and between countries. We are able to identify school entry age rules at the country level from the data and exploit these in a fuzzy regression discontinuity design to estimate causal effects of school starting age on students' performance. We conclude that students who started school being one year older compared to their peers perform on average 0.16 standard deviations better in reading and 0.15 standard deviations better in mathematics tests.

Key Words: School starting age, Fuzzy regression discontinuity, OECD

Econometric Game 2024
April 18, 2024

1 Introduction

Nearly every educational system in the world enforces a formal rule which influences a child's relative age compared to her peers in the classroom. Typically, there is a specific cutoff date that determines eligibility for school enrollment: children must reach a certain age by a designated day of the year before starting school. This rule can introduce age differences of up to one year within a single classroom. Such differences may significantly impact students' academic achievement, their educational career and even labor market outcomes.

Consequently, a large body of economic literature on education explores whether school starting age is a crucial determinant of students' achievements as several studies indicate that older children do outperform their younger peers in standardised tests or probability of on-time promotion (Black et al., 2011; Cook & Kang, 2020). Using data from England, Crawford et al. (2013) found that those children who enroll in school at an older age are even more likely to enroll in a university and outperform their peers there. However, it is widely recognized that using school starting age as an explanatory variable might suffer from endogeneity issues due to the following reasons. First, the primary variation in school entrance age arises from differences in birthdates, which could correlate with unobserved parental factors, such as education, earnings or family background. Second, some parents might disregard the enrollment rule based on certain characteristics of their child. For example, if the child shows development delays, her parents might postpone her school entry. Lastly, in some countries, repeating a grade after failing an academic year is not uncommon. Additionally, aside from the school starting age, there appear to be two more factors which might drive this result: age-at-test and the relative age in class (Attar & Cohen-Zada, 2018). However, those effects tend to be co-linear and hence, are not trivial to disentangle (Black et al., 2011; Cascio & Schanzenbach, 2016; Peña, 2017).

To address potential endogeneity issues related to parental choice of school entrance dates, many studies employ a quasi-experimental approach. These studies use the legal age set by the official assignment rule as an instrument for the actual enrolment age, finding rather mixed evidence (see Bedard and Dhuey (2006), Black et al. (2011), and Datar (2006) among others). However, these studies do not fully tackle the arising endogeneity concerns as birth dates might still correlate with student outcomes, which could affect the validity of the results.

Another strand of studies exploits the fact that some children enroll at school at an older age because they are born after a specific cutoff date using a regression discontinuity design. Focusing mostly on the United States and OECD countries, the results suggest that a one-year increase in the enrollment age of a child leads to a decreased probability of grade retention and a substantial increase in overall test scores. However, there is little consensus, whether this effect is vanished by entering the middle school and about the long-run effects in general (Bedard & Dhuey, 2006; Black et al., 2011; Datar, 2006; Dobkin & Ferreira, 2010; Fredriksson & Öckert, 2014; Robertson, 2011).

As most of the existing studies investigate the effect of school starting age on a student's outcomes on a national level for a single country, we contribute to the current literature by running a multinational analysis for selected OECD countries. For our empirical analysis we draw data from PISA (Programme for International Student Assessment) considering the last three waves (2015, 2018, 2022) and hence, focusing on the short-run outcomes. This database contains reading, mathematics and science test scores for 15-year-olds conducted in an international assessment survey. Applying a regression discontinuity design, we focus on the test outcomes for reading and mathematics and limit our analysis to ten OECD countries (Finland, Luxembourg, Chile, Switzerland, Costa Rica, Germany, South Korea, Austria, the Netherlands and Israel). The rational behind this step is to only include countries with a clear school starting rule which we estimate from our data. Our results align with the current literature in economics of education and confirm that within a classroom, older students indeed outperform their younger peers. Considering a pooled sample of the ten selected OECD countries, our results indicate that children whose age is just above the school starting cutoff have a 0.16 of a standard deviation higher performance in reading and 0.15 in math, respectively. Our results appear to be robust to different bandwidth specifications. Furthermore, they do not seem to be driven by other mechanisms, such as grade retention or early childhood development.

This paper is organized as follows. Section 2 describes the data used in our empirical analysis as well as details on our sample selection and existing trends within and across the OECD countries. Section 3 presents our empirical strategy and Section 4 outlines our main results as well as heterogeneity analyses. Section 5 discusses the possible mechanisms through which the school starting age can affect test scores of teenage students and a robustness check. Finally, Section 6 contains the conclusions.

2 Data and Descriptive Results

2.1 PISA Database

We obtain our data on students' performance in mathematics and reading from the Programme for International Student Assessment (PISA) and explicitly focus on the waves 2015, 2018 and 2022. PISA was first conducted in 2000 by the OECD and the data collection has since been repeated every three years, with the exception of 2021, due to the COVID-19 pandemic. The number of participating countries in every sampling period is more than 50 developing and developed countries. In each wave, PISA randomly selects nationally representative samples of 15-year-old students and assesses them using standardised multiple-choice tests in science, mathematics and reading. The main goal of these tests is to measure students' knowledge of the subject to obtain internationally comparable data on their educational achievement. This can be leveraged to improve educational policies, human capital accumulation and to resolve inequalities (Bietenbeck & Collins, 2023; Hanushek et al., 2013).

PISA employs a two-stage sampling design in most of the countries. First, a random sample of schools in which 15-year-old students are enrolled, is drawn. Hereby, the probability for a school being selected is proportional to its size i.e., the estimated number of 15-year-old students attending. In the second stage, PISA samples 35 students of the 15-year-students from the eligible schools where each student has the same sampling probability (Hanushek et al., 2013).

The test scores are standardised to have a mean of 500 and a standard deviation of 100 across all OECD countries participating in PISA 2000. PISA uses a method of plausible values to account for the uncertainty of reported statistics resulting from differences in the distributed tasks. Hereby, ten plausible values are provided for each subject. For the sake of simplicity, we use the average test score per student for each subject as our main outcome variable[1].

The test score results are accompanied by a rich set of background information on each student and school. Students are asked to fill out a questionnaire on their demographic characteristics but also their family background. Furthermore, school principals are asked to provide information on a school's institutional settings and its endowments, such as admittance policies, school's location or its digital resources.

2.2 Descriptive Statistics

Table 6 and Table 7 presents the mean, median, and standard deviation for the test scores in mathematics and reading, respectively for all OECD countries in our sample, i.e. 38 in total. In our main analysis in Section 4, we will only focus on ten selected OECD countries (Finland, Luxembourg, Chile, Switzerland, Costa Rica, Germany, South Korea, Austria, Netherlands and Israel) as only those indicate a clear school starting rule described in Section 3.2.

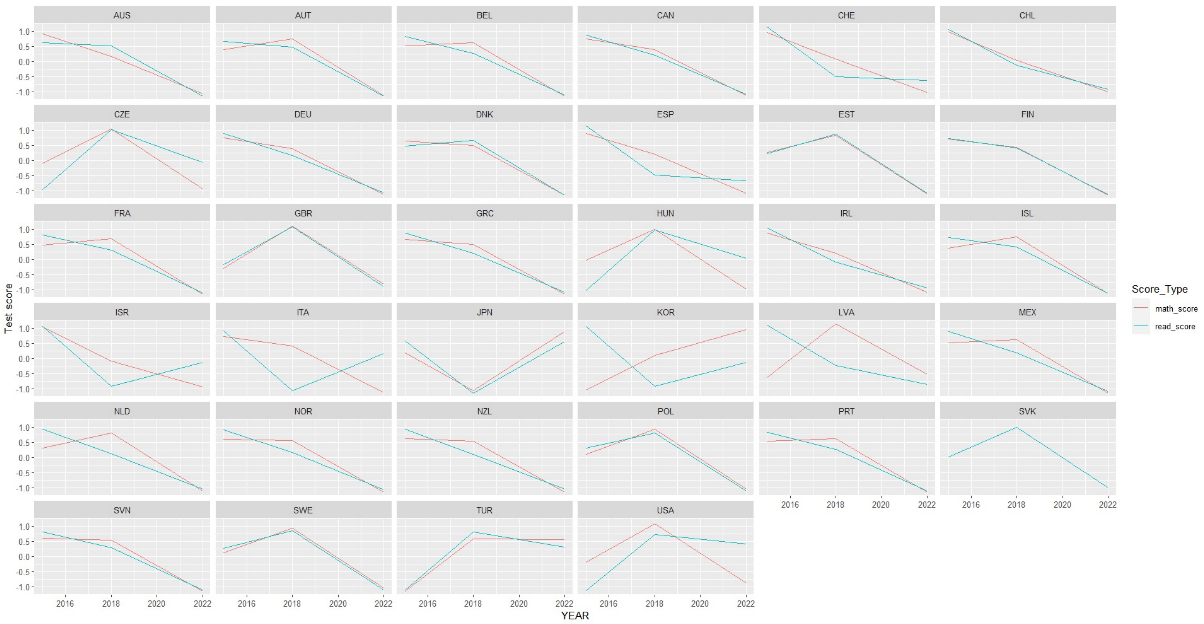

Figure 1: Standardised test scores by country over time for the PISA waves 2015, 2018 and 2022.

Figure 1 presents the development of the standardised test scores in mathematics (orange) and reading (blue) over time. Here, we limit our investigation to 34 countries as only those have test scores results for each wave. Test scores are standardised by centring each country's scores around zero and dividing it by its standard deviation at the country level. Overall, there seems to be a declining trend in both, mathematics and reading since 2018, with only single countries being able to improve their test scores over the years, such as Japan, South Korea, Hungary or the United States. As the test scores are already falling since 2015, the overall decline can only partially be attributed to the COVID-19 pandemic.

There persists a relatively huge variation in test scores across countries prior to the pandemic (2015 - 2018). Several countries, for example Czech Republic, Hungary, and Great Britain indicate an improvement in their test scores, while others experience a decline, e.g. Germany,

[1] An alternative approach would be to use jackknife repeated replication technique as a resampling method

Canada and Spain. In contrast to different developments from 2015 to 2018, we observe a decrease of test scores in 2022 compared to the previous wave for most of the countries in both reading and mathematics. Exceptions are Japan and Korea which were able to increase student performance in the last period in both reading and math, while Israel and Italy improved their average reading scores. The fact that some countries are able to increase their performance while others experience a decrease evokes the question whether some of these countries implemented educational policies that are more successful than those of other countries. Moreover, it motivates our investigation of whether changing school starting age is an effective policy to increase student performance.

Figure 4 and Figure 5 show respectively the demeaned mathematics and reading test scores of students from different countries. It allows us to consider the differences between children coming from households with at least one highly educated parent (red) or none (blue). For this visualisation, we exclude pupils who have missing information about their parents' educational level, as there is only a small number of them, less than two percent of the OECD countries' dataset. Parents with more than twelve years of schooling, which in most countries constitutes as the end of secondary education, are considered high-educated while those with less time spent in academia are low-educated. Within the sample, a small portion of children, 2.7 percent have parents with less than 9 years of schooling, and 34.3 percent with 9 to 12 years. The remaining 63 percent have parents with tertiary education. The observed downward trend in performance is consistent with our previous findings. This movement seems to be independent of a given student's familial background, the performance lines move perfectly together in some countries, suggesting that the cause of the decline is independent of households, rather affecting a whole generation within a country. This cause should be explored on a national level, as improving countries manage to increase the performance of both student groups, suggesting the root of the problem is not the potential difference in student composition, but rather regional effects. One noteworthy exception is Italy, where the reading capabilities of students from high-educated parents seem to have increased while the abilities of pupils from low-educated parents are on a steady decline.

3 Empirical Approach

3.1 Identification Strategy

To identify a causal effect of starting school one year earlier compared to peers, we exploit potentially exogenous variation created by the cutoff date for school starting age rules similar to the sharp regression discontinuity design of (Oosterbeek et al., 2021). The basic idea is that children born just before the cutoff date will start school in a given year while those born after the cutoff start school one year later, such that they are respectively the youngest and oldest students in their class.

Not all the countries that we use in our sample have strictly enforced school starting age cutoffs. As a consequence, we use a fuzzy RDD, where the effect of the school starting rule on actual school starting age is used as a first stage of an instrumental variable regression. The first-stage equation that we estimate for each has the following form:

$$A_{ict} = \alpha_c + \lambda_t + \pi \mathbf{1}\{\text{birthmonth}_i > \text{cutoff}\} + f(\text{birthmonth}_i) + \xi_{ict} \tag{1}$$

where A_{ict} is the school starting age of student i belonging to school grade c that is interviewed at wave t. The reduced form equation for each country is:

$$y_{ict} = \alpha_c + \lambda_t + \delta \mathbf{1}\{\text{birthmonth}_i > \text{cutoff}\} + f(\text{birthmonth}_i) + \epsilon_{ict} \tag{2}$$

The variable, $\mathbf{1}\{\text{birthmonth}_i > \text{cutoff}\}$, classifies whether a student is born before or after the cutoff and is used as an instrument for A_{ict}. We include up to a third-degree polynomial, $f(\text{birthmonth}_i)$, of the running variable to make sure that we do not interpret a nonlinear relationship between birth month and school starting age as a discontinuity. Furthermore, we allow the coefficients of the polynomials to differ on each side of the cutoff. We include grade fixed effects, α_c, as students born close around the cutoff have approximately the same age, however, those born just before the cutoff date have spend a year more in school and are thus expected to have on average higher test scores compared to those who were born just after the cutoff. We also estimate pooled versions of these equations using data of all the countries with a clear cutoff in which we include country level fixed effects. Including those is necessary as tracking systems on country level might be affected by performance differences that are due to the different school starting ages assigned by the cutoff (Oosterbeek et al., 2021).

An instrument variable estimate is obtained by scaling the reduced form estimate of δ by the first-stage estimate of π (see e.g. Angrist & Pischke, 2009). It should be noted that the estimand obtained from a fuzzy RDD is not an average treatment effect of the treated but a local average treatment effect (LATE). It summarises the effect of the treatment on those who complied to the cutoff of the school starting rule. Standard errors for our results are obtained using the sample weights provided by PISA.

3.2 Identifying Cutoff Dates

To implement our identification strategy, it is necessary to know which countries of our PISA sample have clearly enforced school-starting rules and to identify the exact cutoff dates for those. Therefore, we plot the average school grade for each birth month and year for every country. Then, we mark the month during which there was the largest drop. This is our tentative cutoff month. We use only those countries for our analysis for which we can identify a clear jump at some month of birth. Furthermore, to make our choice of countries more robust we only use those countries in our analysis that have a large F-statistic of the instrument when estimating the first-stage equation for each country.

Figure 2 shows the average age of students at school start relative to the cutoff month. We can use it to visually identify which countries enforce school-starting age laws strictly, as the age of students will produce a sudden jump after the cut-off in these instances. Children born right before the date their government determined as the school starting threshold will be the youngest school starters, while the ones born right afterwards will be the oldest. This can be clearly observed in the figures of Finland, Luxembourg, Chile, Switzerland, Costa Rica,

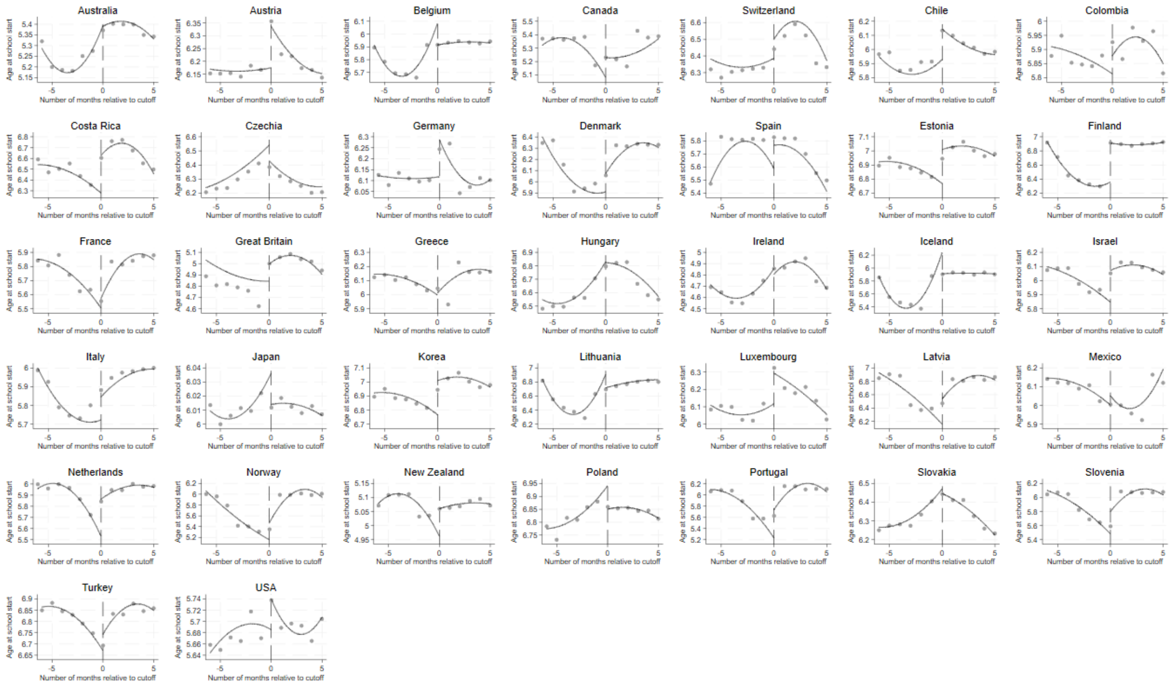

Figure 2: Average age of students at school start relative to the cutoff month

Germany, South Korea, Austria, the Netherlands, and Israel. From Table 8 we can obtain the corresponding F-statistics of the aforementioned ten countries and conclude that they fulfil our eligibility criteria of an F-score close to ten.

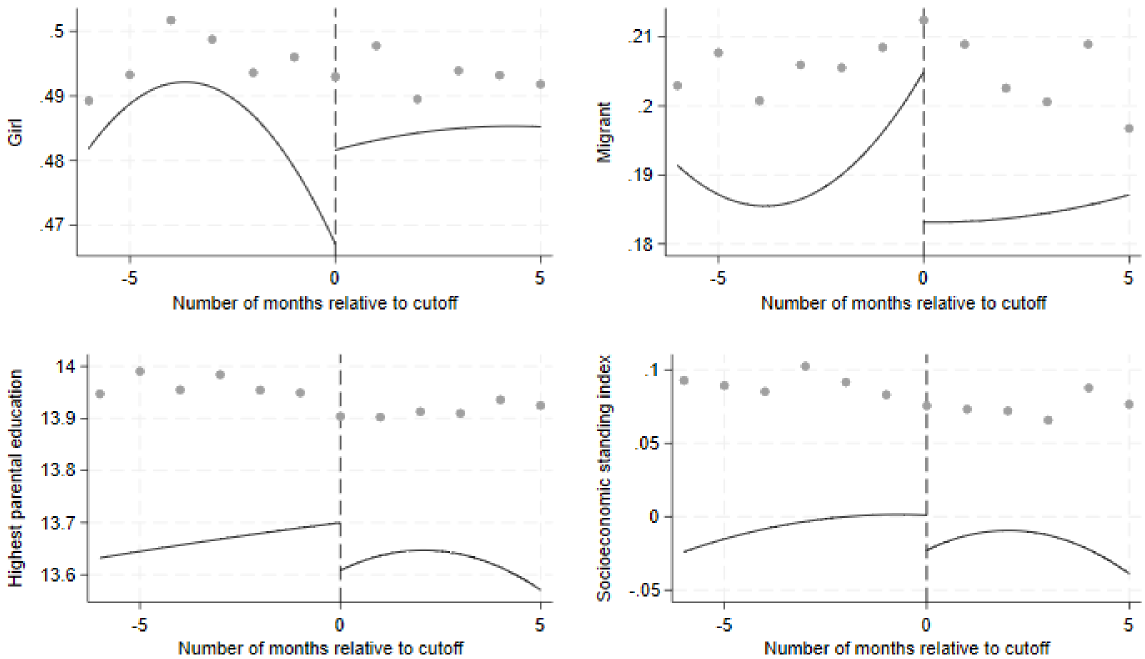

Figure 3: Balancing of predetermined individual characteristics.

3.3 Validity of Identification Strategy

The key identifying assumption of the empirical strategy described above is that the student being assigned just above and below the cutoff is independent of their potential outcomes when being assigned just above or below the cutoff conditioned on the included fixed effect. We are by construction not able to test this assumption directly, however, we investigate whether being born just below or above the cutoff has an effect on predetermined characteristics such as gender, immigration status, and parents' socioeconomic status. We plot the average of each predetermined outcome binned at month of birth relative to the cutoff in fig. 3. We observe no significant jumps at the cutoffs which lends credibility to our empirical strategy as it provides evidence that timing of birth around the cutoff is not related to other important characteristics that determine educational attainment but that observations are similar around the cutoff.

4 Results

4.1 Preliminary Analysis

As a preliminary analysis we investigate possible correlations between the mean school starting age of primary school and the test scores in mathematics and reading across all 38 OECD countries. Table 1 presents the estimated coefficients of interest from standardised test scores regressions using an OLS specification. Age at school start is a discrete variable measured in years and the test scores are stadardised with a mean zero and a standard deviation of one. While columns (1) and (3) do not consider any controls for student characteristics, such as gender or the socieconomic status, columns (2) and (4) include the controls.

The estimated effect of the age at school start on the test scores is negative and highly significant (1%-level) for reading and mathematics irrespective of the controls included. Without considering any controls, this translates to a negative effect of 0.099 of a standard deviation in the overall student test score for reading and 0.074 for math. The result is still negative and highly significant after including the controls but its absolute value decreases by roughly 50% for reading and by roughly 75% for math.

Importantly, these results only present correlations between the age at school start and the test scores. Therefore, they need to be interpreted with caution as they are prone to endogeneity issues, such as selection bias.

Table 1: Individual level partial correlations of school starting age and test scores.

	(1) Reading Score	(2) Reading Score	(3) Math Score	(4) Math Score
Age at School Start	-0.0988***	-0.0474***	-0.0744***	-0.0191***
	(0.0011)	(0.0011)	(0.0011)	(0.0011)
Controls	No	Yes	No	Yes
N	785,101	747,170	785,101	747,170

Notes: Controls include gender, first or second-generation immigrant status, parents' highest educational attainment, age, and socioeconomic status. Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).

4.2 First Stage Results

As described in Section 3, we use only those countries for which we can identify a clear jump at a cutoff date for school starting age. Furthermore, it is necessary to have a strong first-stage regression, when estimating the effect of the cutoff date on school starting age. To test the strength of the cutoff variable as an instrument we use F-statistics and R-squared as recommended by Bound et al. (1995). We use only those countries with an F-statistics larger than 10 as a rule of thumb. Our first-stage estimate of the pooled sample suggests that the jump at the threshold is a 15 percentage point increase in school starting age which is close to two months. The estimate is statistical significantly different from zero at the 1%-level and has an F-statistic of the excluded instrument of 616.29. As we have established a strong and economically significant first-stage, we continue with the reduced form results.

4.3 Reduced Form Results

We begin our discussion of the effect of school starting age on student performance with the estimates using the full sample of countries that have a clear school entry rule. Those results are reported in Table 2. For our main specification, we focus on a rather conservative approach and consider a bandwidth of two (columns (1) and (3)). Our treatment indicates a positive and highly significant effect on the standardised reading and math test scores. Being highly significant (1%-level), our estimates show that children whose age is just above the school starting cutoff have a 0.16 of a standard deviation higher performance in reading and 0.15 in math, respectively. Therefore, our results align with the current literature in economics of education and confirm that within a classroom, older students indeed outperform their younger peers (see Oosterbeek et al. (2021) for high school tracking test scores). Scaling these results by the pooled first-stage, one obtains an effect as large as one standard deviation for those students who were assigned to enter school later by being marginally above the cutoff. Furthermore, our results emphasize the important difference between correlation and causality. Comparing these results to Table 1 shows that using a naive OLS approach might lead to an endogeneity issue from unobservables.

We extend our analysis and consider country-specific estimates. Table 9 presents the effect of school starting age on math and reading scores by country. Here, our evidence is rather mixed. We still find a positive effect for seven countries but only for Austria, Finland, South Korea and Luxembourg this effect is statistically significant (mostly at the 1%-level). Notably, we find the largest effect in Finland where students whose age is just above the school starting cutoff have a 0.743 of a standard deviation higher performance in math and 0.911 in reading, respectively. The estimates for Chile, Costa Rica and Germany is negative but not statistically significant.

4.4 Heterogeneity Analysis

The heterogeneity produced by the student composition could come from factors external to national educational jurisdiction. One such example is immigration. Since different countries have different school starting ages and more importantly, different educational systems, migrant 15-year-olds' results could come from differences not present in the native population. Research from Hermansen (2017) suggests that the age at immigration of children influences later job market outcomes. As most immigrants move to an economically more developed country than

Table 2: Reduced form results for the pooled sample of countries.

	(1) Reading Score	(2) Reading Score	(3) Math Score	(4) Math Score
Treatment	0.160***	0.125***	0.150***	0.119***
	(0.0143)	(0.0207)	(0.0142)	(0.0204)
Country fixed-effects	Yes	Yes	Yes	Yes
Grade fixed-effects	Yes	Yes	Yes	Yes
Wave fixed-effects	Yes	Yes	Yes	Yes
Bandwidth in month	2	3	2	3
N	81,015	53,427	81,015	53,427

Notes: Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).

their home nation, the earlier they arrive, the longer they will be exposed to a presumably better education system. This is even more relevant in our collection of selected countries, as well as OECD countries as a whole since they are well-known destinations of immigration.

We hypothesise that students arriving before the age of 6 will behave the same way as non-immigrants, while those arriving later will not have any advantage from being older. In order to test this we divided first-generation immigrants into two groups, according to their age at arrival to the country. The first group arrived before the age of 6 presumably before the start of their journey in formal education, while the second group arrived later. In Table 3 we can see that, contrary to the whole group of students, we cannot conclude that first-generation immigrants as a whole would benefit from being older when they start school. This is also true for those who arrived before the school starting age and for those who arrived afterwards. The only exception is a decrease in reading scores for immigrants arriving at an older age. The lack of significance might be due to the small number of immigrant students in the dataset though.

Another dimension along which outcomes may vary is parents' background. Socioeconomic status may be measured in many different ways, but here we have chosen the simple division highly educated (14 or more years of education) and lowly educated (12 or less years of schooling). The results for the two groups of children are shown in columns (2) and (3) in Table 4. Surprisingly, the effect of being enrolled in school one year later than your peers is very similar for math scores in both groups. However, smaller divergences can be seen with respect to reading scores, where those with low-education parents benefit more from waiting a year. This may be because parents with low education have less resources to compensate for children's education. However, differences are small and should not be interpreted too literally. It may also be the case that specific countries with larger treatment effects are overrepresented in the sample of children with low-education parents.

5 Potential Mechanisms and Robustness Checks

5.1 Potential Mechanisms

As mentioned in Section 1, the effect of school starting age summarises many underlying mechanisms such as absolute age at school start and relative age to peers. Therefore, we explore

Table 3: Heterogeneous effects on standardised test scores in mathematics and reading by migration status

	(2) 1. Gen	(3) Age 0–5	(4) 6–16
Effect on math scores	0.107	-0.0411	-0.0464
	(0.0821)	(0.107)	(0.1016)
Effect on reading scores	0.0586	0.0432	-0.203*
	(0.0879)	(0.111)	(0.110)
N	3,226	2,041	2,302

Notes: Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).

Table 4: Heterogeneous effects on standardised test scores in mathematics and reading by parents' level of education

	(1) Full sample	(2) High ed. (14+ years)	(3) Low ed. (12- years)
Effect on math scores	0.119***	0.111***	0.122***
	(0.0204)	(0.0250)	(0.0343)
Effect on reading scores	0.125***	0.108***	0.139***
	(0.0207)	(0.0254)	(0.0347)
N	53,427	36,330	17,097

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).

two different mechanisms which might potentially drive our results. First, in many countries, grade repetition is common if students fail the academic year. As younger students tend to under-perform in overall tests, there might be a concerning correlation which we cannot observe in the data directly. Hence, we address this issue by introducing an alternative estimation where we exclude students who repeated a grade at least once from our analysis. Second, we exploit potential effects of early childhood development as this also might affect the standardised test scores. Children who have received a certain kind of early childhood development (e.g., attended a pre-school) might perform differently on tests compared to those who did not. Therefore, we control for this fact by including a variable for the duration in early childhood education and care.

Table 5 presents the results for the pooled sample of all ten OECD countries and a bandwidth of two. Columns (1) and (3) show the results for standardised reading and math scores, by additionally excluding students who have repeated the grade at least once. The coefficients have remained positive and statistically significant at the 1%-level compared to our main results in Columns (1) and (3) in Table 9, although being half the size. Columns (2) and (4) consider all students but additionally control for early childhood development. The results differ only little from our main specification.

Table 5: Evidence on potential mechanisms: Grade retention and early childhood development.

	(1) Reading Score	(2) Reading Score	(3) Math Score	(4) Math Score
Treatment	0.0971***	0.172***	0.0812***	0.154***
	(0.0227)	(0.0239)	(0.0228)	(0.0239)
N	43,825	39,390	43,825	39,390

Notes: Columns (1) and (3) use samples that exclude students that have repeated a grade. Columns (2) and (4) control for early childhood education. Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).

5.2 Robustness Check

To assess the robustness of our findings using our RDD approach, we test different bandwidths around the cutoff. We decide against the use of optimal bandwidth tests, see e.g. Imbens and Kalyanaraman (2012), as we do not observe birthday but only months such that we have only a small number of bins. Results for the pooled sample reported in table 2 include estimates for two-months and three-months bandwidths. Estimates using both bandwidths are highly significant for both reading and math scores. However, they are slightly lower when a three months bandwidth is used. This observation seems highly plausible as one would expect a decreasing effect of relative age as those students further away from the cutoff are closer to the mean age of their peers.

6 Conclusion

In this paper, we study the effect of school starting age on educational achievement. Hereby, we use three waves of the PISA study and a large sample of OECD countries to document the development of test scores over time and between countries.

Using individual level student data, we identify school entry age rules at the country level from the data and exploit these in a fuzzy regression discontinuity design to estimate causal effects of school starting age on students' performance. We find that students who started school being one year older compared to their peers perform on average 0.16 standard deviations better in reading and 0.15 standard deviations better in math tests using a pooled sample of the countries with a clear cutoff rule. Our results appear to be robust to different bandwidth specifications. Furthermore, they do not seem to be driven by other mechanisms, such as grade retention or early childhood development.

From a policy perspective, our paper underlines the importance of the school starting age. Nevertheless, our results should be considered with caution as we are only focusing on the short-term effects. This is due to the fact that our data does not allow us to exploit long-term dynamics which opens a gap for further research.

References

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Attar, I., & Cohen-Zada, D. (2018). The effect of school entrance age on educational outcomes: Evidence using multiple cutoff dates and exact date of birth. *Journal of Economic Behavior & Organization*, *153*, 38–57.

Bedard, K., & Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *The Quarterly Journal of Economics*, *121*(4), 1437–1472.

Bietenbeck, J., & Collins, M. (2023). New evidence on the importance of instruction time for student achievement on international assessments. *Journal of Applied Econometrics*, *38*(3), 423–431.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2011). Too Young to Leave the Nest? The Effects of School Starting Age. *The Review of Economics and Statistics*, *93*(2), 455–467.

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak. *Journal of the American Statistical Association*, *90*(430), 443–450.

Cascio, E. U., & Schanzenbach, D. W. (2016). First in the class? age and the education production function. *Education Finance and Policy*, *11*(3), 225–250.

Cook, P. J., & Kang, S. (2020). Girls to the front: How redshirting and test-score gaps are affected by a change in the school-entry cut date. *Economics of Education Review*, *76*, 101968.

Crawford, C., Dearden, L., & Greaves, E. (2013). *When you are born matters: Evidence for england*. IFS Report.

Datar, A. (2006). Does delaying kindergarten entrance give children a head start? *Economics of Education review*, *25*(1), 43–62.

Dobkin, C., & Ferreira, F. (2010). Do school entry laws affect educational attainment and labor market outcomes? *Economics of education review*, *29*(1), 40–54.

Fredriksson, P., & Öckert, B. (2014). Life-cycle effects of age at school start. *The Economic Journal*, *124*(579), 977–1004.

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.

Hermansen, A. S. (2017). Age at arrival and life chances among childhood immigrants. *Demography*, *54*(1), 201–229.

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, *79*(3), 933–959.

Oosterbeek, H., ter Meulen, S., & van der Klaauw, B. (2021). Long-term effects of school-starting-age rules. *Economics of Education Review*, *84*, 102144.

Peña, P. A. (2017). Creating winners and losers: Date of birth, relative age in school, and outcomes in childhood and adulthood. *Economics of Education Review*, *56*, 152–176.

Robertson, E. (2011). The effects of quarter of birth on academic outcomes at the elementary school level. *Economics of Education Review*, *30*(2), 300–311.

A Appendix

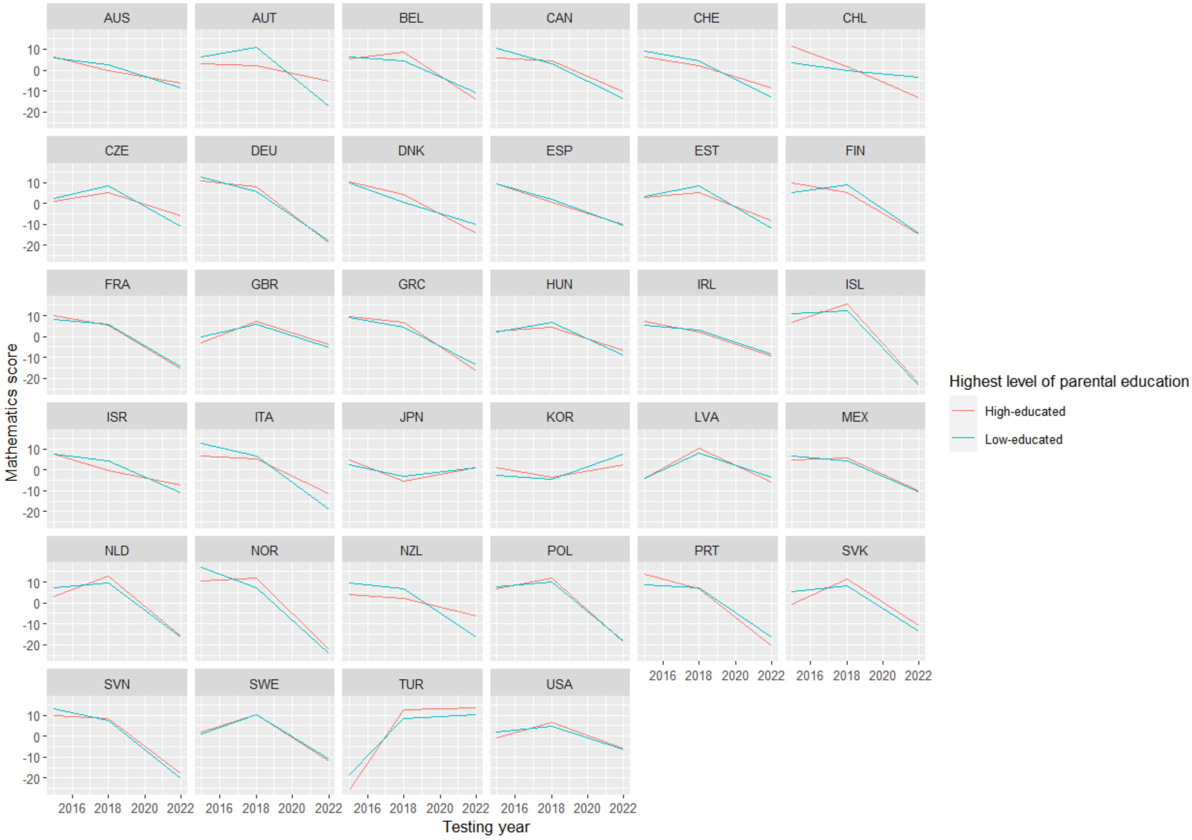

Figure 4: Demeaned mathematics test scores grouped by the highest level of at least one parent's education by country over time for the PISA waves 2015, 2018 and 2022.

Table 6: Mean, median, and standard deviation of mathematics scores by country in each year of testing

		Wave 2015			Wave 2018			Wave 2022		
	Country	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
1	AUS	493.90	483.09	87.82	491.36	492.07	85.17	487.08	485.07	94.62
2	AUT	496.74	503.17	87.81	498.94	507.15	86.43	487.27	490.83	88.18
3	BEL	506.98	517.40	91.29	508.07	518.48	89.42	489.49	496.67	92.32
4	CAN	515.65	505.30	78.61	512.02	505.15	82.82	496.95	482.89	88.22
5	CHE	521.25	523.72	88.50	515.31	519.06	86.39	507.99	507.06	92.08
6	CHL	422.67	443.94	83.53	417.41	433.51	80.89	411.70	425.15	76.05
7	COL				390.93	395.28	75.45	382.70	385.07	69.70
8	CRI							384.58	379.99	62.02
9	CZE	492.33	505.86	87.22	499.47	519.92	89.72	487.00	500.97	92.82
10	DEU	505.97	510.84	83.40	500.04	507.84	90.13	474.83	477.97	90.20
11	DNK	511.09	499.13	78.53	509.40	498.86	77.83	489.27	478.78	79.91
12	ESP	485.84	493.66	77.62	481.39	496.10	79.10	473.14	484.65	80.39
13	EST	519.53	522.15	74.45	523.41	524.18	74.62	509.95	512.40	80.30
14	FIN	511.08	514.46	75.73	507.30	512.06	75.70	484.14	477.29	88.44
15	FRA	492.92	504.26	89.28	495.41	495.61	91.86	473.94	470.75	90.89
16	GBR	492.48	493.29	81.07	501.77	496.23	81.91	488.98	480.74	91.52
17	GRC	453.63	462.73	81.39	451.37	455.89	80.36	430.15	429.24	78.54
18	HUN	476.83	487.33	86.10	481.08	490.82	84.18	472.78	481.66	88.52
19	IRL	503.72	505.88	74.36	499.63	502.35	71.96	491.65	492.82	76.40
20	ISL	488.03	488.68	85.72	495.19	500.17	82.76	458.91	458.71	83.54
21	ISR	469.67	476.40	96.40	463.03	471.14	100.00	457.90	457.93	101.57
22	ITA	489.73	501.61	83.94	486.59	498.65	84.42	471.26	472.24	83.84
23	JPN	532.44	537.33	82.17	526.97	530.47	80.43	535.58	540.01	90.14
24	KOR	524.11	528.50	92.69	525.93	531.62	93.17	527.30	535.65	101.82
25	LTU				481.19	479.64	83.54	475.15	468.26	83.11
26	LUX	485.77	487.73	87.97	483.42	484.77	92.13			
27	LVA	482.31	484.74	70.36	496.13	493.94	71.66	483.16	480.07	75.66
28	MEX	408.02	411.69	67.23	408.80	413.29	68.87	395.03	390.31	64.78
29	NLD	512.25	519.58	86.49	519.23	518.28	90.24	492.68	494.13	103.10
30	NOR	501.73	503.04	78.66	500.96	502.04	84.32	468.45	470.29	89.37
31	NZL	495.22	496.66	85.77	494.49	497.70	86.03	479.07	477.78	94.84
32	POL	504.47	506.23	81.08	515.65	517.65	81.54	488.96	495.73	84.25
33	PRT	491.63	481.32	90.69	492.49	499.15	89.62	471.91	475.23	84.29
34	SVK	475.23	479.62	88.96	486.16	494.50	93.12	463.99	472.43	96.84
35	SVN	509.92	496.47	82.23	508.90	496.32	81.20	484.53	467.31	83.44
36	SWE	493.92	495.47	83.10	502.39	505.31	84.35	481.77	483.55	91.63
37	TUR	420.45	409.04	73.63	453.51	447.96	81.22	453.15	444.57	86.16
38	USA	469.63	469.34	82.61	478.24	473.91	86.47	464.89	458.27	91.13

Notes: Statistics are obtained using sample weights provided by PISA. The sample is based on non-imputed data.

Table 7: Mean, median, and standard deviation of reading scores by country in each year of testing

		Wave 2015			Wave 2018			Wave 2022		
	Country	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
1	AUS	502.90	497.03	98.12	502.63	508.68	105.44	498.05	503.78	102.24
2	AUT	484.87	493.04	94.18	484.39	492.06	95.69	480.41	489.45	96.94
3	BEL	498.52	512.03	93.83	492.86	501.18	98.72	478.85	488.52	96.18
4	CAN	526.67	519.22	84.91	520.09	513.01	98.22	507.13	493.25	99.96
5	CHE	492.20	495.18	90.74	483.93	488.65	97.76	483.33	484.55	99.25
6	CHL	458.57	479.95	83.52	452.27	472.04	90.18	447.98	468.14	87.61
7	COL				412.30	418.04	86.72	408.67	414.84	87.13
8	CRI							415.23	413.28	79.98
9	CZE	487.25	505.85	97.02	490.22	513.11	97.25	488.60	505.06	94.68
10	DEU	509.10	520.93	92.88	498.28	507.55	102.82	479.79	484.93	98.93
11	DNK	499.81	492.67	85.32	501.13	493.13	91.18	488.80	481.61	88.61
12	ESP	495.58	507.42	80.14	476.54	486.97	88.09	474.31	485.59	86.25
13	EST	519.14	524.43	81.80	523.02	525.59	89.66	511.03	517.51	84.27
14	FIN	526.42	536.86	87.53	520.08	528.27	96.13	490.22	485.15	101.83
15	FRA	499.31	517.37	105.22	492.61	490.18	102.59	473.85	474.43	103.34
16	GBR	497.97	497.12	87.08	503.93	502.74	94.62	494.40	492.52	97.04
17	GRC	467.04	484.55	88.44	457.41	463.94	93.32	438.44	444.53	86.71
18	HUN	469.52	486.19	89.80	475.99	487.62	93.85	472.97	489.12	93.41
19	IRL	520.81	525.95	81.07	518.08	520.82	87.19	516.01	522.19	82.95
20	ISL	481.53	484.80	92.66	473.97	477.09	101.11	435.92	438.57	95.36
21	ISR	478.96	490.45	106.46	470.42	482.03	120.50	473.83	481.84	113.51
22	ITA	484.76	498.02	83.93	476.28	486.85	91.07	481.60	485.50	84.75
23	JPN	515.96	526.51	85.76	503.86	508.07	93.98	515.85	522.31	90.84
24	KOR	517.44	524.99	90.06	514.05	525.05	97.91	515.42	527.29	93.19
25	LTU				475.87	474.84	91.31	471.83	468.31	87.71
26	LUX	481.44	488.17	101.01	469.99	473.11	105.60			
27	LVA	487.76	492.39	77.37	478.70	478.61	85.73	474.57	476.23	82.65
28	MEX	423.28	430.88	71.22	420.47	427.06	78.68	415.36	413.66	78.42
29	NLD	502.96	512.45	95.82	484.78	480.19	102.89	459.24	459.44	108.64
30	NOR	513.19	518.99	92.23	499.45	506.35	102.64	476.52	482.74	104.97
31	NZL	509.27	515.72	98.45	505.73	514.20	102.48	500.85	504.71	102.46
32	POL	505.70	512.51	83.33	511.86	515.63	93.17	488.71	503.03	95.92
33	PRT	498.13	491.49	87.64	491.80	497.19	92.96	476.59	485.11	84.84
34	SVK	452.51	460.37	98.03	457.98	461.73	97.97	446.86	456.05	97.60
35	SVN	505.22	490.59	87.87	495.35	482.22	90.71	468.54	458.30	89.83
36	SWE	500.16	507.16	94.94	505.79	512.42	103.70	486.98	496.07	104.10
37	TUR	428.34	424.93	74.53	465.63	465.48	84.36	456.08	457.03	82.34
38	USA	496.94	501.69	93.56	505.35	504.84	104.98	503.94	504.03	106.28

Notes: Statistics are obtained using sample weights provided by PISA. The sample is based on non-imputed data.

Table 8: Selection of countries with a clear school starting rule

	Country	First stage	SE	Clear jump in graph	F-statistic	F large
1	FIN	0.46	0.02	1.00	903.86	1.00
2	LUX	0.28	0.04	1.00	57.16	1.00
3	CHL	0.25	0.06	1.00	20.06	1.00
4	CHE	0.23	0.07	1.00	9.78	1.00
5	CRI	0.21	0.05	1.00	18.14	1.00
6	DEU	0.17	0.04	1.00	22.44	1.00
7	KOR	0.13	0.02	1.00	31.74	1.00
8	AUT	0.12	0.02	1.00	56.89	1.00
9	NLD	0.11	0.02	1.00	23.04	1.00
10	ISR	0.09	0.02	1.00	19.17	1.00
11	GBR	0.30	0.21	1.00	2.09	0.00
12	EST	0.03	0.02	1.00	2.81	0.00
13	ITA	0.26	0.01	0.00		
14	IRL	0.21	0.04	0.00		
15	LTU	0.18	0.02	0.00		
16	AUS	0.15	0.04	0.00		
17	USA	0.07	0.06	0.00		
18	SVK	0.07	0.02	0.00		
19	ISL	0.07	0.03	0.00		
20	COL	0.06	0.09	0.00		
21	ESP	0.05	0.03	0.00		
22	BEL	0.03	0.02	0.00		
23	DNK	0.00	0.02	0.00		
24	HUN	-0.00	0.04	0.00		
25	NZL	-0.01	0.02	0.00		
26	MEX	-0.01	0.02	0.00		
27	JPN	-0.01	0.01	0.00		
28	GRC	-0.01	0.02	0.00		
29	CZE	-0.01	0.02	0.00		
30	CAN	-0.02	0.02	0.00		
31	POL	-0.08	0.06	0.00		
32	LVA	-0.08	0.02	0.00		
33	SVN	-0.09	0.02	0.00		
34	TUR	-0.10	0.02	0.00		
35	PRT	-0.13	0.02	0.00		
36	NOR	-0.21	0.03	0.00		
37	FRA	-0.22	0.03	0.00		

Notes: The first ten countries (Finland, Luxembourg, Chile, Switzerland, Costa Rica, Germany, South Korea, Austria, the Netherlands and Israel) are selected to have a clear school starting rule.

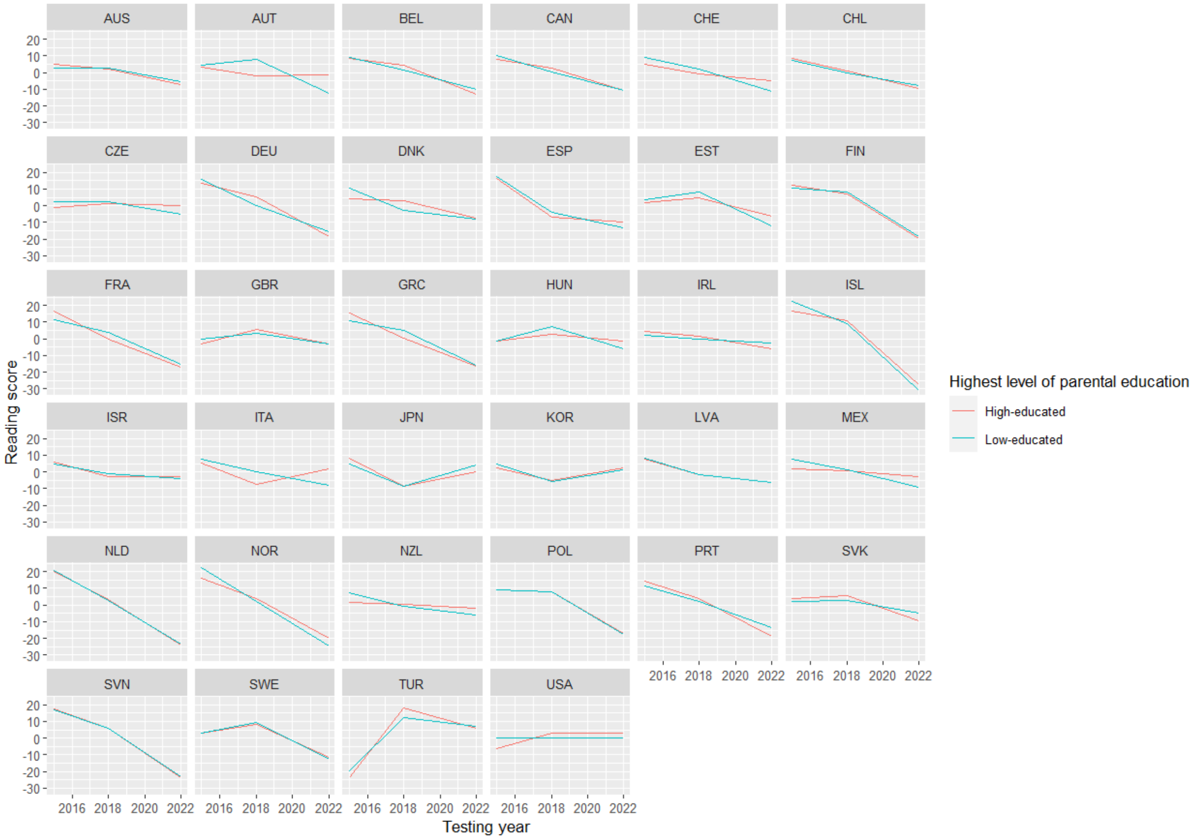

Figure 5: Demeaned reading test scores grouped by the highest level of at least one parent's education by country over time for the PISA waves 2015, 2018 and 2022.

Table 9: Reduced form results for the pooled sample of countries.

	(1) AUT	(2) CHE	(3) CHL	(4) CRI	(5) DEU	(6) FIN	(7) ISR	(8) KOR	(9) LUX	(10) NLD
Effect on Math Scores	0.211**	0.0496	-0.0850	-0.0130	-0.0431	0.743***	0.0155	0.265***	0.423***	0.00134
	(0.0768)	(0.0532)	(0.0527)	(0.0617)	(0.0499)	(0.0591)	(0.0677)	(0.0773)	(0.0636)	(0.0571)
Effect on Reading Scores	0.227**	0.0986	-0.0916	-0.00985	-0.0354	0.911***	-0.0827	0.286***	0.473***	-0.0210
	(0.0782)	(0.0528)	(0.0541)	(0.0744)	(0.0516)	(0.0649)	(0.0714)	(0.0724)	(0.0692)	(0.0595)
Country fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth in month	2	2	2	2	2	2	2	2	2	2
N	4,713	5,945	6,818	2,237	6,237	6,876	6,585	5,567	3,557	4,892

Notes: Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).