



NOVA SCHOOL OF BUSINESS AND ECONOMICS

UNIVERSIDADE NOVA DE LISBOA

Dissertation, presented as part of the requirements for the degree of Doctor of Philosophy in
Economics

Essays in Economics of Education

Afonso Saraiva Câmara Leme

Student nº 25798

A dissertation carried out on the PhD in Economics, under the supervision of Professor Ana
Balcão Reis and Professor Luís Catela Nunes

October 31st, 2022

Abstract

This dissertation consists of three independent essays in economics of education.

In the first chapter, through a cross-county analysis, we study the effect of tracking students in specialization areas *versus* keeping the system more general in academic upper-secondary school, on students' achievement and inequality. The second chapter provides a comprehensive analysis of the Danish school well-being indicators and their determinants. Finally, in the third chapter, I study how a large increase in the compulsory schooling leaving age from 15 to 18 years old in Portugal can contribute to reducing early school leaving and change students' educational paths.

Keywords: tracking; school well-being; compulsory schooling; inequality

Acknowledgments

This dissertation marks the end of a long journey that would have not been possible without the help of many people that accompanied me along this way.

First of all, I want to thank my supervisors, Ana Balcão Reis and Luís Catela Nunes, for continued advice and support. Your honesty and feedback has been an invaluable source of motivation to improve my work and has stopped me from getting stuck on dead end roads many times. Thanks you for your kindness and patience.

The first chapter of this dissertation started with the help of Josep-Oriol Escardíbul at the University of Barcelona. Thank you for your guidance in the early stages of my career as a researcher and for sparking my excitement about the field of Economics of Education.

The second chapter was developed during my research visit in Denmark. I want to thank Helene Bie Lilleør and Anna Folke Larsen for having received me so well at Rockwool Foundation. It was wonderful to have worked in such a vivid, bright, and inclusive workplace. A special thanks goes to Mikkel Stahlschmidt for outstanding research assistance in this project. Thanks to Marianne Simonsen for hosting me at the Department of Economics and Business Economics at Aarhus University during this research stay. Thank you for giving me the opportunity to experience an outstanding research environment and for your essential feedback on my other research projects. This experience really set out the course for the rest of my PhD. The warmest thanks goes to Frederik Jensen for your friendship and help during these months: *tak og vær stærk!*

Thank you to all my colleagues and friends at Nova SBE. Your help and support during the first year courses and the never-ending studying for the Comprehensive exams was precious. A special thanks goes to João Vasco Gama, Matilde Grácio, Mattia Fracchia and Sara Almeida. Thank you to my colleagues at the Social Sciences DataLab. Another special mention goes to Pedro Freitas and João Firmino for your endless help in my research projects involving the Portuguese Administrative Education dataset.

The most immeasurable contribution in my education production function goes to my friends and family: *Aos meus queridos amigos Vasco, Salvador e António — que já viajaram tantas canções comigo, e ainda há tantas por viajar*; to Margherita, for your love and for always being there for me; to my sister Sara, for your inspiration and attitude — *continua rebelde!* Above all, thanks to my parents for believing in me and in my future — it should go without saying that none of this would have been possible without your support and love.

I gratefully acknowledge the support of the grant SFRH/BD/118966/2016 and PTDC/EGE-ECO/4764/2021 from FCT (Fundação para a Ciência e a Tecnologia). The work in this dissertation used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA010145FEDER007722) and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

Table of Contents

<i>Abstract</i>	2
<i>Acknowledgments</i>	3
<i>Introduction</i>	7
<i>Chapter 1</i>	9
<i>The effect of a Specialized versus a General Upper-Secondary school curriculum on students' performance and inequality. A difference-in-differences cross country analysis....</i>	9
1.1 Introduction.....	9
1.2 Empirical Strategy	14
1.2.1 Difference-in-differences model.....	16
1.3 Data	19
1.3.1 Data sources	19
1.3.2 Differences between Specialized and General systems	24
1.4 Results	30
1.4.1 Results from the difference-in-differences model	30
1.4.2 Heterogeneous effects	31
1.4.3 Robustness checks.....	33
1.5 Conclusion	38
<i>Chapter 2</i>	41
<i>Pupil Well-being in Danish Primary and Lower Secondary Schools</i>	41
2.1 Introduction.....	41
2.2 Measuring well-being in elementary and lower secondary school in Denmark	44
2.3 Data, key measures, and samples	45
2.4 Correlates of social well-being.....	52
2.4.A Individual characteristics and family background	52
2.4.B The role of teacher and school characteristics.....	63
2.4.C The role of classroom composition	64

2.4.D Persistence over time	65
2.4.E Survey response.....	67
2.5 Associations with other relevant outcomes.....	68
2.6 Conclusion	69
<i>Chapter 3</i>	<i>70</i>
<i>Too young to quit school? Increasing the Compulsory Schooling Leaving age and educational outcomes.....</i>	<i>70</i>
3.1 Introduction.....	70
3.2 Institutional Setting	78
3.3 Identification Strategy.....	82
3.4 Data	89
3.4.1 Data description	89
3.4.2 Analysis samples	90
3.4.3 Retained and Promoted students.....	92
3.5 Results	97
3.5.1 Who are main compliers of the new CSL?	97
3.5.2 Main Difference-in-Differences Results	103
3.5.3 Heterogeneity analysis.....	107
3.5.4 Placebo test.....	112
3.6 Conclusion	115
<i>References.....</i>	<i>117</i>
<i>A1 Appendix Chapter 1</i>	<i>126</i>
<i>A2 Appendix Chapter 2</i>	<i>128</i>
<i>A3 Appendix Chapter 3</i>	<i>145</i>
A3.1 - Control variables used in Section 3.5.....	145
A3.2 - Machine Learning Prediction of students at Retention Margin in 2006/07.....	146

Introduction

This dissertation is composed of three chapters in the Economics of Education field. In these chapters I have analysed the effects of educational policies; school conditions and systems; as well as socio-economic and family environments on both cognitive and non-cognitive skills, and educational attainment.

In the first chapter, we explore the fact that countries differ in their academic upper-secondary school systems whereby some require their students to choose a specialization track from a set of areas — typically natural sciences, economic sciences, or humanities — and follow that specialization for the course of their upper secondary education years whereas, by contrast, others follow a general curriculum in which students are not required to follow a single specialization and thus, may receive a more general education. Because countries follow only one system or the other, a cross-country analysis is required to estimate the possible effects on students' achievement of these institutional differences. We measure achievement with the scores in reading and mathematics in international assessment programmes. An international difference-in-differences approach is chosen to account for country heterogeneity and unobserved factors influencing student outcomes, by using data from three different large- scale international student assessments. Results suggest that there is a negative effect from specialization tracking and that this may be greater for specific sub-groups of the student population — in particular, for students with a low socioeconomic status.

In the second chapter, we exploit that since 2014, the Danish Ministry of Education has conducted yearly national well-being surveys for children of all ages in public school. The ministry introduced the survey as a tool for schools to monitor well-being of their pupils, to make

informed adjustments of their own related practices, and to inform education policy at the municipal level. This paper studies the characteristics of the social well-being segment of the survey. We document that low school social well-being correlates meaningfully with standard measures of disadvantage at the pupil and parental level, just as teacher characteristics and classroom composition are additional important predictors of well-being. We also show that school social well-being exhibits high degrees of persistence over time, regardless of whether or not we control for a wide range of background characteristics. We finally show that high school social well-being is positively associated with academic performance and negatively associated with absence from school, though estimates are not large in size.

Finally, in the third chapter, I study how a large increase in the Compulsory Schooling Leaving age (CSL) from 15 to 18 years old can contribute to reducing early school leaving and change students' educational paths. The Portuguese reform determined exposure to the new CSL based on enrolment in grades up to the 7th grade in the academic year of 2009/10, while students in grades 8 or above remained under the old policy. This allows me to use grade-retention in the 7th grade in the previous academic year as quasi-natural variation in exposure to the new CSL. As grade- retention only lead to different treatment exposure in 2008/09, I apply a local Difference-in-Differences that uses data from students at the retention margin from previous grade-cohorts to isolate the causal effect of the policy from the effect of retention on outcomes. I find substantial heterogeneity in both the compliance with the new CSL and graduation probabilities, with lower-achieving students experiencing stronger positive effects. Furthermore, results suggest that the effectiveness of CSL increases may depend on other policies taken to prevent early school leaving.

Chapter 1

The effect of a Specialized versus a General Upper-Secondary school curriculum on students' performance and inequality. A difference-in-differences cross country analysis¹

Joint work with Josep-Oriol Escardibul, Luís Catela Nunes, Ana Balcão Reis, and Carmo Seabra

1.1 Introduction

The efficiency and equity of a country's educational system are topics of major concern and interest for policymakers, researchers, parents, and students. An important characteristic of any educational system is whether it applies educational tracking and the timing in which it does so. Tracking can be generally described as separating students into different school classes, and Betts (2011) identifies two types of tracking, as this separation can be based on academic ability or curriculum focus. In the ability-grouping tracking, students whose overall achievement is above average are assigned to the same class, as are students whose achievement is at an average level or below it. Usually, this separation by ability goes hand in hand with the division of students by curriculum focus, as it typically also determines whether a student follows an academic-based track that prepares them for entrance in a university, or a vocational-based track that prepares students for entrance to the labor market straight after their school years. Furthermore, tracking can occur within schools, as is more common in the US — a form of ability-grouping also called streaming — or it can lead to a division of students between different

¹Published in *Applied Economics* 52, no. 39 (2020): 4317-4331

<https://doi.org/10.1080/00036846.2020.1734183>

It includes Supplemental Material classifying Educational systems, discussed in Appendix A1.

schools, as is the norm in countries such as Germany, the Netherlands, and several other European countries (Maaz et al., 2008). Some countries such as Germany and Austria track students by ability as early as age 10, while most OECD countries do so at age 15 or 16 (Woessmann, 2009). While on a theoretical level different arguments for and against tracking can be made², most recent empirical research has concluded that tracking increases inequality (Hanushek and Woessmann, 2006) and reinforces the differences of family background on educational outcomes (Brunello and Checchi, 2007).

However, there is a type of tracking that has not been investigated in the literature before. This type of tracking is present at the upper secondary school level for students who are following an academic path. Similarly to the other types of tracking, it is applied by some countries and absent in others. Countries differ in their upper secondary school systems such that some require their students to choose a specialization track from a set of areas — typically natural sciences, economic sciences, humanities or arts — and follow that specialization for the course of their upper secondary education years (e.g., Portugal, Spain, Sweden). By contrast, others including Finland, Germany, and the UK, follow a general curriculum in which students, although able to choose between different classes in distinct areas, are not required to follow a single specialization, and may thus receive a more general education. The defining difference between the two systems analyzed is therefore that in a general upper-secondary school system, the academic track is undifferentiated and students are allowed to take courses in several different academic fields, such as advanced mathematics and natural sciences as well as social sciences

² From a theoretical point of view, the effects of tracking are widely controversial, suggesting a substantial uncertainty about its impact on both the level and distribution of students' achievement - see Betts (2011); Meier & Schutz (2007) for a review on the theoretical considerations on the impact of tracking.

and humanities. On the other hand, in a specialized system, students are sorted into distinct academic tracks from the beginning until the end of upper-secondary schooling in which each specialization track is typically composed of a set of classes with pre-defined hours, even if the student may also have the opportunity to choose some optional classes beyond the ones offered in that track. Therefore, we classify an educational system as specialized if there exists a clear set of specialization tracks from which students have to choose upon starting academic upper-secondary school. While comparing the academic upper-secondary school curriculum of different countries with a specialized system, a clear difference that we observe is that between specialization areas, the natural sciences and economics tracks consistently have a greater focus on mathematics than the others, while to some lesser extent the humanities tracks have a greater focus on languages, writing and reading skills. The label “specialized” academic upper-secondary school system may be confusing, as most or all upper-secondary school systems offer students some degree of choice to specialize in certain subjects or areas. Our classification of a specialized system is defined by the existence of a clear set of specialization tracks, such as the ones mentioned above. Thus, for the sake of simplicity, throughout this article we use the expressions “specialized” and “specialization tracking” interchangeably.

One difference from other types of tracking, in which schools make strong recommendations for what type of track the student should follow based on his/her previous achievement, is that the choice of track placement — in this case the specialization — is solely the student’s and parents’ choice, at times aided by psychological advice or tests. This decision is usually based on the academic interests of the student and his/her plans for his/her area of study at the tertiary education level, as some university programs often ask for the completion of some courses specific to a certain area (e.g., it is unlikely for a student following an arts specialization to study

medicine at the tertiary level). Hence, this type of tracking is not designed to divide students by ability but rather by area of interest and specialization, and for this reason the theoretical framework concerning the effects of having classrooms with homogenous or heterogeneous levels of student ability does not directly apply. However, we may expect that students self-select into the specialization track that best fits their type of ability — for instance, we suppose that students with higher relative ability in mathematics and natural sciences courses to be more likely to choose a natural sciences or mathematics intensive specialization track (Arcidiacono, 2010 and Bartolj & Polanec, 2012 show that perceived and relative ability have a great influence on choice of college major). Because students self-select into specialization tracks, it is plausible that the distribution of students across specialization tracks differs in terms of students' absolute or relative ability, parental background, and other individual characteristics of students. In fact, De Groote (2019) shows that students in the classical languages and mathematics track of the Flemish system have higher average language and math ability at the start of secondary education as well as higher parental background than students of other tracks. Gender also matters for track choice, as there is a greater share of boys in mathematics-intensive tracks in the Flemish system — a result also confirmed by Rapoport & Thibout (2018), who show that in the French system, boys more often choose the Sciences track while girls more often choose the Humanities track. Although the different specialization areas are not designed to have different levels of difficulty *per se*, as Bishop (2010) points out, the maths-science tracks have a reputation for being the most difficult and prestigious, giving better chances of being accepted in varied tertiary education programs. Rapoport & Thibout (2018) also confirm that the majority of students enrolled in the most elite colleges in France studied in the Natural Sciences track, and that almost all post-

secondary pathways remain open for students from this track, while future options are much narrower for students in the Humanities track.

These patterns of distribution of students across specialization tracks may be relevant for predicting the effects of sorting students into different academic tracks, instead of keeping the system more general and comprehensive. In the tracking literature, the arguments for comprehensive schools, as opposed to tracking students by ability or curricular focus, claim that lower-level students will be systematically disadvantaged by slower learning environments if they are separated into different tracks, leaving them further behind the level of the upper groups, thereby increasing inequality — both in their educational achievement and later in their life, through lower earnings in their adult years (Pekkarinen et al., 2009). This possibility that tracking aggravates economic inequality is also related to the fear that it leads to a distribution of students through socio-economic background, perpetuating a bias against more disadvantaged students (Brunello & Chechi, 2007). It is possible that some of these effects are also present in a specialization tracking system.

In this article we investigate whether international tests show an effect of this type of tracking, either through average effects or heterogeneous effects for different sub-groups of the student population on reading and mathematics achievement. We explore if there are different effects from specialization by the socioeconomic background and gender of students. Since countries only follow one system or the other, this type of tracking requires a cross-country analysis to estimate the possible effects of these institutional differences. For this reason, an international difference-in-differences analysis is used to account for country heterogeneity and unobserved factors influencing student outcomes. The impact of having a specialized system in upper-

secondary school, as opposed to a general one, is identified by comparing the change in outcomes between a point just before students have started this stage of schooling (measured using PISA data) and a later point, after they have completed it (using data from the recent PIAAC study) across countries with a specialized and general systems. We therefore follow the student micro-level data difference-in-differences methodology of Waldinger (2007), Jakubowski (2010), and Ruhose & Schwerdt (2016) with Repeated Cross Sections of PISA and PIAAC data combinations to study the effect from specialization tracking on students' achievement.

1.2 Empirical Strategy

The common method to estimate the determinants of a student's achievement is to consider an education production function in which the achievement of a student — usually measured by test scores — depends on various factors: personal characteristics of the student such as gender and innate ability; family characteristics such as parental education and income; school characteristics and even specific characteristics of a country's educational system.

One could consider a model like the one shown in equation (1) to estimate the impact of a student following a specialized system, as opposed to a general one, on his/her achievement:

$$Y_{igs} = X'_{igcs}\beta + \gamma SPEC_{gs} + \varepsilon_{igs} \quad (1.1)$$

where the achievement of student i in grade g and school s (Y_{igs}) is regressed on several characteristics of the student, her family and school (vector X), and the existence of a specialized system in his/her upper-secondary school ($SPEC$). The main issue with this methodology that is focused on the achievement of students within a single country is that the existence of a

specialized system in upper-secondary school is determined at the national level, and therefore, there is no variation in the variable *SPEC*. Even if there are schools that follow a different type of curriculum than the one defined at the national level, these schools are usually also private or international schools and in this way, vary significantly in terms of school resources and parental background of their students. Thus, an unbiased estimate of the effect of a specialized system, γ , could not be obtained.

Instead, an alternative would be to focus on a cross-country analysis, in which the achievement of students in a country following a general system is compared to that of students in a specialized system. Following this model, we would obtain unbiased estimates of the average effect of having a specialized system in upper-secondary school as opposed to a general one, only if the conditional independence assumption would hold — that is, conditional on X , the potential outcomes of each student are independent of the treatment, *SPEC*. However, it is very unlikely that this assumption holds, as there may be several unmeasurable factors at the country level that may be correlated with the existence of each of the two systems. For example, the quality of the teaching force could be systematically correlated with the existence of a specialized system.

If longitudinal data were available that would allow us to follow the achievement of students through the course of their school education in several different countries with each of the two different systems, this could be considered a reliable approach to eliminate the existence of the unmeasurable confounding factors mentioned above. Unfortunately, these types of data are not available at a cross-country level.

1.2.1 Difference-in-differences model

The lack of longitudinal data at a cross-country level along with the issues just described are also an issue for research dealing with the impact of early tracking by ability on academic achievement and inequality. However, as first noted by Hanushek & Woessmann (2006), the existence of large-scale assessment programs, such as PISA, TIMSS, and PIRLS, that assess cohorts of students at different educational stages, allows for a difference-in-differences methodology. To solve the issues described above, Hanushek & Woessmann (2006) effectively apply an international difference-in-differences approach at the country-level, by comparing the average achievement gain in countries with early ability tracking to that of countries without it. As no country applies tracking in the early primary grades, it is possible to compare the level and distribution of student performance in these grades by using international assessment programs, to these levels in secondary school where some countries have separated students into differing-ability schools and others have not, by looking at the results for secondary school. This difference-in-differences methodology (DID) was later expanded to a student micro-level data approach by Waldinger (2007), Jakubowski (2010), and Ruhose & Schwerdt (2016), who studied whether early tracking affects different groups in a distinct manner, namely students with a low socioeconomic status and migrant students.

Given the potential for a difference-in-differences methodology to account for time-invariant unobservables affecting outcomes, we choose to implement this estimation strategy to identify the impact of a specialized system on students' achievement in academic upper-secondary school, as opposed to maintaining a general system. The simplest DID estimator of this effect, $\hat{\gamma}$, is presented below:

$$\hat{\gamma} = (\bar{Y}_{SPEC,2} - \bar{Y}_{SPEC,1}) - (\bar{Y}_{GEN,2} - \bar{Y}_{GEN,1}) \quad (1.2)$$

where $\bar{Y}_{g,t}$ represents the average achievement of students in group g , specialized (SPEC), or general (GEN), in a period t , 2 or 1 — that is, after and before students have been through a specialized or general system in upper-secondary school, respectively. The DID estimator, $\hat{\gamma}$, is therefore obtained by comparing the difference in average achievement growth between a point just before upper-secondary school and after this educational stage, across countries with a specialized or general system. This parameter produces unbiased estimates of the effect of having a specialized system on achievement in academic upper-secondary schooling, as opposed to having a general system, if the identifying assumption of DID models, the parallel-trends assumption, holds. However, there may be time-varying factors affecting the achievement of students in specialized and general systems differently, implying a violation of the parallel-trends assumption. If, for instance, parental background characteristics of the two samples analyzed in $t=2$ — that is, after students have been through a specialized or general system in upper-secondary school — and $t=1$ (before students have been through this two upper-secondary school systems) were very different and correlated with the existence of a specialized system, this would imply a violation of the parallel-trends assumption. Therefore, including control variables that account for these characteristics minimizes the possibilities of a violation of our identifying assumption.

To allow for the addition of such control variables at the student level, this estimator is in fact implemented in a regression framework with country-fixed effects³ that follows the student micro-level DID approach of Waldinger (2007), Jakubowski (2010), and Ruhose & Schwerdt (2016) in Repeated Cross-Sections, shown below:

$$Y_{ict} = \alpha_c + \lambda UPSEC_t + \gamma SPEC_c * UPSEC_t + X'_{ict}\beta + \varepsilon_{ict} \quad (1.3)$$

where Y_{ict} is the standardized test score of student i , in country c and time t ; α_c are country fixed-effects; $UPSEC_t$ is a dummy variable that takes the value 1 if an observation is taken from a student who has finished upper secondary school and the value 0 if from a student in an educational stage just before upper-secondary schooling; $SPEC_c$ is a dummy variable that identifies whether an observation comes from a country with a specialized system in upper-secondary school; X_{ict} is a vector of student background variables that includes gender, the number of books at home, the highest level of parental education, and whether the student speaks the language of the test at home; and ε_{ict} is the error term. The key parameter of interest is therefore γ , which under the identifying assumptions of this difference-in-differences methodology, estimates the effect of a specialized system, in contrast with a general system, on a student's achievement. It is worthwhile noting that since we do not observe the same individuals in the two time periods (before and after upper-secondary schooling), a further identifying assumption of this Repeated Cross-Section approach is that the stratified sampling

³ To account for country heterogeneity, we include country fixed effects in all of our regressions. This specification excludes $SPEC_c$ from the most basic DID specification $Y_{ict} = \alpha_1 SPEC_c + \alpha_2 UPSEC_t + \gamma SPEC_c * UPSEC_t + \varepsilon_{ict}$ to avoid a perfect multicollinearity problem.

nature of the surveys used in this analysis guarantees that a representative sample of students in academic tracks is tested.

Besides studying whether there are average effects from specialization tracking for the complete sample, we also explore if specialization tracking has heterogeneous effects on different sub-groups of the student population, which could give us some indication of the possible effect of this type of tracking on students' inequality. The sub-groups that we analyze are students with a low and high socioeconomic status, and a division by gender.

1.3 Data

1.3.1 Data sources

Our measures of achievement in reading and mathematics come from PISA and PIAAC. To complement the description of the differences in educational systems, we also use TIMSS data, as described in subsection 1.3.2. To assess students in the different systems before specialization tracking has occurred we use data from PISA. Students who take part in this exam are 15 years old (or have turned 16 very recently), just before the entrance to upper secondary schooling⁴.

To assess whether specialization in upper-secondary school has an impact on students' performance, we use a test similar to PISA, which evaluates students just after they have concluded secondary education. In 2012 the OECD released the results of the first edition of PIAAC, a study that evaluates adults from 16 to 65 years old in the areas of literacy, numeracy, and problem-solving. Around 155,000 individuals were tested, over 24 countries in this first

⁴ While some 15 year old students assessed in PISA may have already been in upper-secondary school for a short while, and therefore exposed to the effects of a general or specialized system, we assume these effects are still null or, at least, small.

round of the study. By restricting the sample of this study to individuals who are 18 or 19 years old and have recently finished academic upper-secondary schooling, it is then possible to obtain a sample of interest for the analysis proposed. We therefore include observations from countries that were present in both the PISA and PIAAC studies and use test-scores for subjects that are common to both surveys, that is, the reading and mathematics parts of the tests⁵.

As the methodology of this chapter proposes to compare the performance of students at a point before and after they have been through an academically-oriented education in upper-secondary school, several observations had to be dropped from the original PIAAC database and to a lesser extent, from the PISA database. Starting with PIAAC, this database has information on the highest level of education obtained by each examinee according to the ISCED 97 classification (UNESCO, 2003). This tells us not only whether an individual has completed at least upper secondary education — the basic condition to be eligible for this analysis — but also if this education had an academic or vocational orientation. As we are interested in analyzing only students that followed an academic education in upper secondary school (since they are the only ones subject to either a specialized or general system), students with a vocational education at this level were removed from the study. Some individuals did not report their highest level of education and for those from Estonia, this information is not available at all — therefore, they were also removed. Furthermore, PIAAC has information on the area of study for the highest level of education obtained — with this, individuals who followed an area which was clearly not academic were also removed. Concerning the age group of individuals in the PIAAC database, we chose to include only individuals with a maximum age of 19 years old in order to minimize

⁵ For more information regarding the PISA and PIAAC frameworks for numeracy and mathematical literacy and their comparability, refer to Gal & Tout (2014).

the chances of having other confounding factors such as the quality of post-secondary and tertiary education influencing a student's achievement. As for the PISA database, removals were made mostly in countries that have tracking between academic and vocational schools before or at the age when students take this exam.

Since our methodology proposes to compare individuals that have been through an academic upper secondary education following either a specialized or general system with their equivalents at a point before this stage of education, students who are already following a vocational track at the age of 15 are not eligible for this analysis. In this way, in countries with no early tracking between vocational and academic tracks such as Norway or Finland, no removals needed to be made, whereas in countries with early tracking such as Austria or Germany, several removals were done. These removals were then made for students in a vocational or pre-vocational program (ISCED 3 or ISCED 2) and in the PIAAC database, for students with missing information on the type of program they were attending. Finally, we perform our main regression analysis with native students only as we can better assure that these students have spent all of their educational years in their home country – and therefore, have been exposed to the effects of a general or specialized upper-secondary school system. In Section 1.4.3 we perform a robustness check in which we also include migrant students in the analysis (1st and 2nd generation migrants).

Regarding the PISA cohorts used to run the difference-in-differences estimation, two possibilities exist: we can follow the PISA 2009 cohort, as 15 years old students were approximately 18 or 19 years old in PIAAC 2012, or focus on contemporaneous measures of performance at the two different educational stages by using the PISA 2012 and PIAAC 2012. According to Hanushek & Woessman (2006), the former possibility has the advantage of

minimizing the possibilities of having confounding factors from any inherent variation in family background and peer characteristics, while the latter has the benefit of assessing the two different cohorts in a stable educational system. In order to provide a robustness check on our results, we test with the two combinations in our DID regressions. The control variables used in these micro-level regressions have to be available in both surveys. We chose to include gender, the number of books at home when the student was 16 years old, the highest level of parental education, and a variable indicating whether the individual speaks the language of the test at home.

Concerning the information about whether a country follows a specialized or general system in upper-secondary school, these data were gathered from different sources, the most important being the Eurydice Network (Eurydice, 2017). Other sources range from OECD documents, information available on the websites of the Ministry of Education of each country, and other articles about the curriculum and educational system of a country, such as the country reports of the TIMSS and PISA studies. A document describing and justifying the classification of each of the countries used in our sample is available as a support to this article.

Table 1.1 shows the division of systems by country in each of the two survey combinations used in our analysis. Because we are more confident in comparing the achievement of students in more similar educational systems and because our main and most reliable source for the classification of each country's upper-secondary school system, Eurydice, provides information only for European countries, we choose to restrict the sample to include students from European countries only.

Table 1.1
Division of systems by country

Specialized	General
Austria	Czech Republic
Belgium (Flanders only)	Finland
Denmark	Germany
France	Ireland
Italy	Poland
Netherlands	Slovak Republic
Norway	United Kingdom
Spain	(England and Northern Ireland only)
Sweden	

Number of countries: 16; Number of countries with a Specialized system: 9;
Number of countries with a General system: 7

Test scores for each individual in PISA and PIAAC are estimated with plausible values based on multiple imputation techniques. As PISA and PIAAC have different scales (PISA from 200- 800 and PIAAC from 0-500), the test scores analyzed are standardized to have a cross-country mean of zero and a standard deviation of 1, within each survey. Thus, results presented in Section 1.4 should be interpreted in terms of changes in standard deviations. We take into account the complex survey designs of PISA and PIAAC by including all plausible values available in these surveys and their sampling weights. To adjust standard errors for stratification, we take into account the replicate weights provided in PISA and PIAAC. In PISA and PIAAC these weights are named Balanced-Repeated-Replication (BRR) weights. As argued by Jerrim et al. (2017), regression analysis with international large-scale assessment programs should take these weights into account for valid standard error estimation.

1.3.2 Differences between Specialized and General systems

Table 1.2 presents some statistics of demographic and socioeconomic background variables for the two assessment periods, before and at the end of upper-secondary school, for the Repeated Cross Sections of PISA-PIAAC. In terms of these demographic and socioeconomic background variables, it is possible to see that the differences between Specialized and General systems tend to be small and are mostly maintained from the PISA 2009 and 2012 cohorts to the PIAAC 2012 cohort.

Table 1.2
Summary statistics - demographics and socioeconomic background in PISA and PIAAC cohorts

	PISA 2009			PISA 2012			PIAAC 2012		
	Specialized	General	Difference	Specialized	General	Difference	Specialized	General	Difference
Male	0.478	0.501	- 0.022 *** (0.003)	0.473	0.497	- 0.024 *** (0.003)	0.410	0.401	0.009 (0.028)
0-10 books	0.098	0.114	-0.016 *** (0.002)	0.117	0.116	0.001 (0.002)	0.028	0.042	- 0.013 (0.010)
11-25 books	0.140	0.151	-0.012 *** (0.002)	0.152	0.157	- 0.005 ** (0.002)	0.066	0.114	- 0.047 *** (0.016)
26-100 books	0.305	0.309	- 0.004 (0.003)	0.296	0.301	- 0.004 (0.002)	0.313	0.300	0.016 (0.026)
101-200 books	0.198	0.188	0.009 (0.003)	0.192	0.189	0.003 (0.002)	0.207	0.221	- 0.013 (0.023)
201-500 books	0.161	0.152	0.009 *** (0.002)	0.152	0.152	0.000 (0.002)	0.246	0.197	0.048 ** (0.023)
> 500 books	0.097	0.084	0.013*** (0.002)	0.089	0.084	0.004 ** (0.002)	0.139	0.129	0.009 (0.019)
Highest Parental Education: Primary or Lower-Secondary	0.142	0.068	0.073 *** (0.002)	0.121	0.099	0.021 *** (0.002)	0.141	0.032	0.110 *** (0.015)
Highest Parental Education: Upper-Secondary	0.332	0.486	- 0.154 *** (0.003)	0.319	0.416	- 0.097 *** (0.003)	0.363	0.448	- 0.085 *** (0.028)
Highest Parental Education: Tertiary	0.525	0.444	0.081 *** (0.003)	0.559	0.484	0.075 *** (0.003)	0.495	0.518	- 0.023 (0.028)
1st generation immigrant	0.045	0.039	0.005 *** (0.001)	0.055	0.040	0.017 *** (0.001)	0.061	0.019	0.042 *** (0.010)
2nd generation immigrant	0.059	0.062	-0.003 ** (0.001)	0.062	0.058	0.004 *** (0.001)	0.048	0.069	- 0.020 (0.013)
Different language	0.101	0.063	0.038 *** (0.001)	0.108	0.055	0.053 *** (0.002)	0.040	0.024	0.016 (0.009)
Late immigrant	0.021	0.014	0.006 *** (0.001)	0.021	0.015	0.005 *** (0.001)	0.014	0.007	0.006 (0.005)
Observations	68,148	35,528		69,360	37,978		555	721	

Summary statistics show means of the variables for the PISA 2009, PISA 2012, and PIAAC 2012 samples in Specialized and General upper-secondary school systems. A late immigrant is defined as an individual who migrated to the country where the test was taken at an age older than 10. Students weights are used for weighting. Significance levels for testing whether the difference is equal to zero or not: *** p < 0.01, ** p < 0.05, * p < 0.1.

Because PIAAC does not provide information on students' choice of courses (besides knowing whether the education was academic or vocational), we use data from TIMSS 1995 to compare

the type of upper-secondary school courses taken by students in Specialized and General systems, and between specialization tracks within Specialized systems. In its 1995 edition, TIMSS had the unique characteristic of assessing students in the last year of a country's upper-secondary school cycle in mathematics and science literacy. This version was called "TIMSS Advanced" and has several items of information regarding students' choice of courses: we know if they were enrolled in advanced mathematics and physics courses; the amount of hours of mathematics instruction they had per week; how many science courses they took; and furthermore, for a subgroup of countries, we know whether a student from a country with a specialized system was in a Natural Sciences specialization track or not. We believe extending our analysis to the TIMSS 1995 study has the advantage of giving us the opportunity of comparing students' family background and their set of courses in the different systems, and between specialization tracks in specialized systems. However, one drawback is that the data are older and that the educational systems of these countries have changed considerably since 1995. Also, in our TIMSS sample only 4 of the 14 countries analyzed have a general system⁶, meaning that we do not have as much system variability at the country level as we do with PISA and PIAAC. Therefore, we only use TIMSS data to complement the description of educational systems and do not use these data in our regression analysis.

Taking advantage of the information that TIMSS Advanced 1995 provides regarding students' choice of courses and their parental background, in Table 1.3 we explore whether there are

⁶ In the TIMSS Advanced 1995 sample the countries following a Specialized system are Austria, Cyprus, Czech Republic, Denmark, France, Iceland, Lithuania, Netherlands, Sweden, and Switzerland. The countries following a General System are Germany, Hungary, Norway, and Slovenia.

differences between Specialized and General systems in the last year of upper secondary school. The last column of Table 1.3 also shows that, in fact, there seem to be quite large differences between students' choice of courses in Specialized and General systems. Namely, there is a much higher proportion of students taking 5 hours or more per week of Mathematics in Specialized systems. Regarding the number of Natural Science courses taken in this school year, we also observe a significantly higher proportion of students taking 3 or more Natural Science courses in Specialized systems. Finally, TIMSS Advanced also provides information regarding whether students take at least one of the two following subjects in the last year of upper secondary school: Advanced Mathematics and Physics (the TIMSS Advanced 1995 survey names these students "specialists"). We observe that mostly all students in General systems are "specialists" while almost half of the students in Specialized systems do not take either of these two subjects. It appears therefore, that there is a higher degree of heterogeneity in terms of students' course enrollment in Specialized systems, with students in General systems taking intermediate amounts of weekly hours of Mathematics more commonly, as well as for the number of Natural Sciences courses (around half of the students in General systems take 3 to 4 hours of Mathematics per week and 2 Natural Science courses).

Table 1.3

Summary statistics - demographics, socioeconomic background and course of studies in TIMSS Advanced 1995

Last year of upper-secondary school			
	Specialized	General	Difference
Male	0.449	0.426	0.022 *** (0.008)
0-10 books	0.012	0.004	0.008 *** (0.002)
11-25 books	0.048	0.030	0.018 *** (0.004)
26-100 books	0.282	0.137	0.145 *** (0.007)
101-200 books	0.264	0.200	0.064 *** (0.007)
> 200 books	0.394	0.628	- 0.235 *** (0.009)
Highest Parental Education: Primary or Lower-Secondary	0.228	0.018	0.210 *** (0.007)
Highest Parental Education: Upper-Secondary	0.483	0.478	0.005 (0.009)
Highest Parental Education: Tertiary	0.289	0.503	- 0.214 *** (0.008)
Migrant	0.053	0.042	0.011 *** (0.004)
Different language	0.034	0.041	- 0.007 *** (0.003)
Not taking Mathematics	0.042	0.132	- 0.089 *** (0.005)
< 3 hours of Mathematics per week	0.225	0.343	- 0.118 *** (0.009)
3 to <4 hours of Mathematics per week	0.225	0.486	- 0.262 *** (0.010)
4 to <5 hours of Mathematics per week	0.101	0.025	0.076 *** (0.006)
> 5 hours of Mathematics per week	0.408	0.014	0.394 *** (0.010)
Not taking any Natural Sciences courses	0.230	0.046	0.183 *** (0.007)
Taking 1 Natural Science course	0.082	0.307	- 0.225 *** (0.006)
Taking 2 Natural Science courses	0.127	0.447	- 0.319 *** (0.007)
Taking ≥3 Natural Science courses	0.562	0.200	0.362 *** (0.009)
Specialist	0.490	0.938	- 0.448 *** (0.010)
Observations	11,336	4,660	

Summary statistics show means of the variables for the TIMSS Advanced 1995 samples in Specialized and General upper-secondary school systems. In this sample the countries following a Specialized system are Austria, Cyprus, Czech Republic, Denmark, France, Iceland, Lithuania, Netherlands, Sweden, and Switzerland. The countries following a General System are Germany, Hungary, Norway, and Slovenia. A "Specialist" is defined by TIMSS 1995 as a student that takes at least one of the two following subjects in the last year of upper secondary school, Advanced Mathematics or Physics. Students weights are used for weighting. Significance levels for testing whether the difference is equal to zero or not: *** p < 0.01, ** p < 0.05, * p < 0.1.

Given that our definition of a Specialized system is related with the existence of a clear set of specialization tracks in upper secondary school, we also explore the extent to which demographic, socioeconomic background characteristics, and students' choice of courses differ between specialization tracks within Specialized systems in Table 1.4. Namely, for 5 of the 10 countries with a Specialized system (Cyprus, Czech Republic, France, Sweden, and Switzerland) we are able to identify if students were enrolled in a Natural Sciences specialization track or not. With this information we can also study whether the heterogeneity in students' course enrollment in Specialized systems, seen in Table 1.3, is related with differences that exist between specialization tracks. Regarding demographic and socioeconomic background characteristics, we observe that these tend to be mostly similar between Natural Science and non-Natural Science tracks, except for the facts that there is a significantly higher proportion of male students in Natural Science tracks and that there are differences between parental education of these students, with students in Natural Science tracks having higher levels of more highly educated parents. This result is in line with those reported in the literature and discussed in the introduction of this article, that students in Natural Sciences tracks typically have higher levels of parental background and have a higher proportion of male students. Moreover, we observe large differences regarding students' choice of courses in these different tracks. Namely, there is a much higher proportion of students in Natural Science tracks that take high levels of weekly hours of Mathematics instruction (5 or more hours per week) and 3 or more Natural Science courses. Furthermore, roughly all students in Natural Science tracks are "specialists" while only 8% of students not in Natural Science tracks take at least one of the two subjects, Advanced Mathematics or Physics.

Table 1.4

Summary statistics - demographics, socioeconomic background, and course of studies for students in Specialized systems enrolled in Natural Science tracks or not (TIMSS 1995)

	Natural Science Track	Not in a Natural Science track	Difference
Male	0.603	0.313	0.290 *** (0.013)
0-10 books	0.002	0.017	- 0.015 *** (0.003)
11-25 books	0.038	0.051	- 0.013 ** (0.006)
26-100 books	0.285	0.297	- 0.012 (0.013)
101-200 books	0.285	0.261	0.024 * (0.012)
> 200 books	0.390	0.373	0.016 (0.013)
Highest Parental Education: Primary or Lower-Secondary	0.228	0.317	- 0.087 *** (0.013)
Highest Parental Education: Upper-Secondary	0.423	0.447	- 0.024 *** (0.014)
Highest Parental Education: Tertiary	0.348	0.237	0.112 *** (0.013)
Migrant	0.049	0.051	- 0.002 (0.007)
Different language	0.021	0.044	- 0.023 *** (0.005)
Not taking Mathematics	0	0.041	- 0.041 *** (0.005)
< 3 hours of Mathematics per week	0.074	0.405	- 0.331 *** (0.013)
3 to <4 hours of Mathematics per week	0.058	0.189	- 0.131 *** (0.010)
4 to <5 hours of Mathematics per week	0.050	0.157	- 0.111 *** (0.009)
≥ 5 hours of Mathematics per week	0.822	0.208	0.614 *** (0.014)
Not taking any Natural Sciences courses	0	0.470	- 0.470 *** (0.013)
Taking 1 Natural Science course	0.034	0.050	- 0.016 *** (0.006)
Taking 2 Natural Science courses	0.093	0.125	- 0.032 *** (0.009)
Taking ≥ 3 Natural Science courses	0.873	0.354	0.519 *** (0.014)
Specialist	0.984	0.082	0.903 *** (0.014)
Observations	2.063	3.614	

Summary statistics show means of the variables for TIMSS Advanced 1995 sample in Specialized upper-secondary school systems, separately for students in Natural Science tracks and not in these tracks. The countries included are Cyprus, Czech Republic, France, Sweden, and Switzerland. A "Specialist" is defined by TIMSS 1995 as a student that takes at least one of the two following subjects in the last year of upper secondary school, Advanced Mathematics or Physics. Students weights are used for weighting. Significance levels for testing whether the difference is equal to zero or not: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

1.4 Results

1.4.1 Results from the difference-in-differences model

Table 1.5 presents the difference-in-differences estimates of equation 4, using Repeated Cross Sections of PISA 2009 with PIAAC 2012 (columns 1, 2, 5 and 6) and PISA 2012 with PIAAC 2012 (columns 3, 4, 7, and 8). The first four columns report the regression results for reading test scores, while the last four columns present results for mathematics. Odd numbered columns show results without further controls, whereas in even numbered columns the control variables mentioned in Section 1.2 are added.

All estimates for the coefficients of the control variables have their expected sign and magnitude. Regarding our variable of interest, the interaction of the Upper Secondary with the Specialized system dummy variables, estimates without further control variables suggest that specialized systems in upper-secondary school are associated with lower scores of around 23 percent of a standard deviation in both mathematics and reading. Estimates decrease slightly to around 20 percent of a standard deviation when adding further control variables. The fact that estimates of the interaction variable with and without the addition of controls variables are very similar and not statistically significantly different from each other at a 1% significance level can be considered as a sign that our identifying parallel-trends assumptions may in fact hold. We therefore estimate that specialization tracking has a negative average effect on both mathematics and reading achievement.

Table 1.5

Difference-in-differences estimation with PISA and PIAAC

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Reading				Mathematics			
	PISA 09 - PIAAC 12	PISA 12 - PIAAC 12	PISA 12 - PIAAC 12	PISA 12 - PIAAC 12	PISA 09 - PIAAC 12	PISA 12 - PIAAC 12	PISA 12 - PIAAC 12	PISA 12 - PIAAC 12
Specialized system x Upper Secondary	-0.250*** (0.050)	-0.227*** (0.046)	-0.246*** (0.051)	-0.207*** (0.045)	-0.239*** (0.032)	-0.220*** (0.028)	-0.202*** (0.032)	-0.169*** (0.027)
Upper-Secondary	0.112*** (0.027)	-0.088*** (0.025)	0.113*** (0.031)	-0.116*** (0.027)	0.121*** (0.024)	-0.036* (0.020)	0.104*** (0.027)	-0.084*** (0.022)
<i>Further Controls</i>								
Male		-0.089*** (0.034)		-0.072** (0.034)		0.293*** (0.019)		0.287*** (0.022)
11-25 books		0.264*** (0.053)		0.292*** (0.054)		0.244*** (0.060)		0.271*** (0.055)
26-100 books		0.536*** (0.042)		0.542*** (0.041)		0.546*** (0.045)		0.559*** (0.041)
101-200 books		0.844*** (0.052)		0.831*** (0.056)		0.826*** (0.038)		0.838*** (0.035)
201-500 books		1.087*** (0.033)		1.090*** (0.030)		1.080*** (0.046)		1.111*** (0.042)
>500 books		1.088*** (0.029)		1.076*** (0.031)		1.083*** (0.040)		1.112*** (0.036)
Highest Parental Education: Upper-Secondary		0.268*** (0.038)		0.222*** (0.038)		0.222*** (0.025)		0.208*** (0.025)
Highest Parental Education: Tertiary		0.408*** (0.045)		0.351*** (0.042)		0.385*** (0.035)		0.367*** (0.036)
Different language		-0.220*** (0.080)		-0.203** (0.082)		-0.164* (0.088)		-0.166* (0.088)
Constant	0.434*** (0.083)	-0.542*** (0.073)	0.446*** (0.076)	-0.481*** (0.066)	0.562*** (0.053)	-0.558*** (0.064)	0.535*** (0.056)	-0.578*** (0.063)
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.045	0.211	0.001	0.184	0.066	0.247	0.057	0.241
Individual Observations	96,458	90,604	97,869	91,320	96,458	90,604	97,869	91,320
Number of countries	16	16	16	16	16	16	16	16

Dependent variables are the original test scores (plausible values), which are standardized to have an international mean of 0 and a standard deviation of 1 within each survey. BRR (Balanced Repeated Replication) standard errors in parentheses. All regressions include Country Fixed effects and a dummy variable indicating that an observation is taken from the upper-secondary school sample. The reference categories for the books and Highest Parental Education dummy variables are 0-10 books and lower than Upper-Secondary, respectively. Number of countries with a Specialized system: 9. Significance levels: *** 1%, ** 5%, * 10%.

1.4.2 Heterogeneous effects

Results from the previous section suggest that, on average, specialized systems in upper-secondary school are associated with lower scores in both mathematics and reading. It is also interesting to investigate whether particular sub-groups of the student population lose more from specialization tracking than others or whether some sub-groups may even gain from it. Following the structure of Table 1.5, Table 1.6 presents estimates of the coefficient for our variable of interest in our difference-in-differences specification, with the addition of further control variables, for specific sub-groups of the student population. The sub-groups that we analyze are

the following: students with a low Socioeconomic status (low SES), defined as having fewer than 100 books at home or no parent with tertiary education, and students with a high Socioeconomic status (high SES); female and male students. We follow the choice of Jakubowski (2010) of defining “low SES” as having fewer than 100 books at home or no parent with tertiary education. Changing the definition for the different SES groups does not change our conclusions.

Table 1.6

Heterogenous effects with PISA-PIAAC

	(1)	(2)	(3)	(4)
	Reading		Mathematics	
	PISA 09 - PIAAC 12	PISA 12 - PIAAC 12	PISA 09 - PIAAC 12	PISA 12 - PIAAC 12
Low SES	-0.306*** (0.048)	-0.274*** (0.052)	-0.287*** (0.047)	-0.224*** (0.043)
High SES	-0.054 (0.131)	-0.078 (0.133)	-0.082 (0.057)	-0.062 (0.064)
Boys	-0.211** (0.083)	-0.187** (0.081)	-0.307*** (0.037)	-0.249*** (0.039)
Girls	-0.230*** (0.063)	-0.214*** (0.062)	-0.153*** (0.037)	-0.108*** (0.040)

All entries show the coefficient of the interaction between Upper Secondary x Specialized System. BRR Standard errors in parentheses. All regressions include Country Fixed effects and a dummy variable indicating that an observation is taken from the upper-secondary school sample. Further controls added are the same as used before. Low SES is defined as having fewer than 100 books at home or no parent with tertiary education. Number of Countries: 16. Number of countries with a Specialized system: 9. Significance levels: *** 1%, ** 5%, * 10%.

Negative effects from specialization tracking in upper-secondary school seem to be concentrated on students with a low socioeconomic status (low SES), as there is no statistically significant effect for high SES students. When dividing our sample by gender we find that specialization tracking has a negative effect on reading achievement for both boys and girls. We also estimate

a negative effect on mathematics achievement for both groups, but this effect appears to be stronger for boys.

It is worth noting that whenever statistically significant results are found for the effect of specialization tracking they are estimated to be negative. While we have no way of assessing the mechanism of these effects with PISA and PIAAC, we believe they may be related with the specialization track choice of these different sub-groups of students in specialized systems – more specifically, the pattern of track choices in specialized systems of different SES groups and genders could be related with the pattern of results we observe in Table 1.6.

1.4.3 Robustness checks

In this section we perform five sets of robustness checks on the results found in the previous sections. First, we perform a placebo test, also known as a falsification test, by running the same difference-in-differences regressions with a different sample that should not be influenced by the treatment we are analyzing. We restrict the PISA samples to include only students in vocational or pre-vocational tracks and the PIAAC sample to include only students that completed upper-secondary school in a vocational track, following the same age criterion as before of including only students with a maximum age of 19 years old. As the treatment that we are analyzing, the specialization tracking, only affects students in academic tracks, we would expect to find a statistically insignificant estimate of the interaction coefficient. Table 1.7 presents results for reading achievement in the first four columns, while results for mathematics are presented in the last four columns. Odd numbered columns present results without further controls, while in even numbered columns these control variables are added. Each entry of Table 1.7 shows only the

coefficient of the interaction term, following the full specification of our difference-in-differences model.

Table 1.7

Placebo test - regressions with Vocational students in PISA and PIAAC

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Reading				Mathematics			
	PISA 09 - PIAAC 12	PISA 12 - PIAAC 12	PISA 12 - PIAAC 12	PISA 12 - PIAAC 12	PISA 09 - PIAAC 12	PISA 12 - PIAAC 12	PISA 12 - PIAAC 12	PISA 12 - PIAAC 12
Specialized system x Upper Secondary	0.058 (0.103)	0.051 (0.094)	-0.053 (0.154)	-0.032 (0.145)	0.094 (0.077)	0.078 (0.065)	0.035 (0.125)	0.042 (0.136)
Further Controls	-	Yes	-	Yes	-	Yes	-	Yes
R-squared	0.056	0.122	0.083	0.145	0.063	0.136	0.079	0.158
Individual Observations	26,042	22,353	16,761	14,483	26,042	22,353	16,761	14,483
Number of countries	16	16	16	16	16	16	16	16

Dependent variables are the original test scores (plausible values), which are standardized to have an international mean of 0 and a standard deviation of 1 within each survey. BRR (Balanced Repeated Replication) standard errors in parentheses. All regressions include Country Fixed effects and a dummy variable indicating that an observation is taken from the upper-secondary school sample. Further controls added are the same as used before. Number of countries with a Specialized system: 9. Significance levels: *** 1%, ** 5%, * 10%.

We indeed find a statistically insignificant impact of the interaction term in all cases, which reassures us that differences in academic upper-secondary test scores between the two systems do not appear to be a simple by-product of differences that exist in the educational systems of the two systems as a whole. The estimates we find for the placebo sample of vocational students are also statistically significantly different from the main effects presented in Table 1.5, at a 1% significance level.

In the previous section we argued that a sample of native students is more appropriate to study educational system effects. Still, it is relevant to analyze whether our results change when including migrant students (1st and 2nd generation). In this case our sample comprises 90% natives and we estimate a highly statistically significant negative effect (p-value < 0.01) from specialization of around 14 percent of a standard deviation. When considering a sample of migrant students only we estimate statistically insignificant effects for our variable of interest.

There is a possibility that our estimated negative effect of specialization tracking is driven by how restricted the sample of academic students is in each system. Because in our main analysis we focus only on students in academic or general tracks, as those are the only ones that will likely be subject to a specialized or general academic upper-secondary school system, we would expect that for countries with early tracking by ability we have a more restricted sample that is associated with higher test scores. Even though the tetrachoric correlation of the early tracking status of country with the existence of a specialized system in academic upper-secondary school of 0.017 is quite negligible, we test whether including this variable in our regressions changes our main results and in this way, make sure that differences in test scores in upper-secondary school between the two systems are related in fact with the tracking of students by specialization areas and not with how restricted our sample in each system is. We follow the commonly used definition in the literature of defining an early tracking country as a country that tracks students before the age of 15, that is, before the time of assessment of the PISA survey. Data on the timing of each country's tracking were obtained from Ruhose & Schwerdt (2016) and 7 of the 16 countries used in our sample have early tracking (Austria, Belgium, Czech Republic, Germany, Italy, Netherlands, and Slovak Republic). Table 1.8 presents regression results for reading and mathematics achievement with the two Repeated Cross Sections analyzed, following the same specification as our main results shown in Table 1.5, but adding a dummy variable indicating the early tracking status of a country. Adding this variable does not change our results for the estimated effect of specialization tracking.

Table 1.8

Robustness check - Accounting for early ability tracking

	(1)	(2)	(3)	(4)
	Reading		Mathematics	
	PISA 09 - PIAAC 12	PISA 12 - PIAAC 12	PISA 09 - PIAAC 12	PISA 12 - PIAAC 12
Specialized system x Upper Secondary	-0.233*** (0.046)	-0.207*** (0.046)	-0.226*** (0.028)	-0.169*** (0.027)
Early trackers	0.073 (0.066)	0.005 (0.067)	0.077 (0.049)	-0.002 (0.047)
Further Controls	Yes	Yes	Yes	Yes
R-squared	0.211	0.202	0.247	0.241
Individual Observations	90,604	91,320	90,604	91,320
Number of countries	16	16	16	16

Dependent variables are the original test scores (plausible values), which are standardized to have an international mean of 0 and a standard deviation of 1 within each survey. BRR (Balanced Repeated Replication) standard errors in parentheses. All regressions include Country Fixed effects and a dummy variable indicating that an observation is taken from the upper-secondary school sample. Further controls added are the same as used before. The "Early trackers" variable takes the value 1 if the observation comes the upper-secondary school sample and from a country that tracks students by ability before the age of 15, and 0 otherwise. Number of countries with a Specialized system: 9. Significance levels: *** 1%, ** 5%, * 10%.

We also investigate whether these results could be driven by a heterogeneous treatment of specialization tracking. As previously discussed, all upper-secondary school systems offer students some degree of choice to specialize in certain subjects or areas, regardless of the existence of specialization tracking areas or not. As such, we believe that the fact that mathematics is not compulsory at any point of academic upper-secondary school could be considered a reasonable proxy for the degree of specialization of a country. More specifically, in 5 of the 16 countries analysed with PISA and PIAAC, it is possible for students to go through academic upper-secondary schooling without enrolling in any mathematics course. These countries are France, Ireland, Netherlands, Spain, and the United Kingdom. Information on whether mathematics is compulsory for academic upper-secondary school is described in the supplementary document to this article that describes and justifies the upper-secondary school system classification of each of the countries used in our sample. In Table 1.9 we repeat our difference-in-differences regressions, but add instead a dummy variable indicating that

mathematics is not compulsory in any of the grades in upper-secondary schooling. That is, this variable takes the value one for students from the upper-secondary school sample that come from a country where mathematics is not compulsory in any of upper-secondary schooling. Adding this variable does not change our conclusions regarding the estimated impact of specialization tracking on achievement.

Table 1.9

Robustness check - Accounting for whether Mathematics is compulsory in upper-secondary school

	(1)	(2)	(3)	(4)
	Reading		Mathematics	
	PISA 09 - PIAAC 12 PISA 12 - PIAAC 12		PISA 09 - PIAAC 12 PISA 12 - PIAAC 12	
Specialized system x Upper Secondary	-0.263*** (0.048)	-0.230*** (0.049)	-0.221*** (0.032)	-0.173*** (0.030)
Maths not compulsory	0.111* (0.056)	0.080 (0.058)	0.003 (0.048)	0.014 (0.048)
Further Controls	Yes	Yes	Yes	Yes
R-squared	0.211	0.202	0.247	0.241
Individual Observations	90,604	91,320	90,604	91,320
Number of countries	16	16	16	16

Dependent variables are the original test scores (plausible values), which are standardized to have an international mean of 0 and a standard deviation of 1 within each survey. BRR (Balanced Repeated Replication) standard errors in parentheses. All regressions include Country Fixed effects and a dummy variable indicating that an observation is taken from the upper-secondary school sample. Further controls added are the same as used before. The "Maths not compulsory" variable takes the value 1 if the observation comes from the upper-secondary school sample and from a country where Mathematics is not compulsory at any point of academic upper-secondary school, and 0 otherwise. Number of countries with a Specialized system: 9. Significance levels: *** 1%, ** 5%, * 10%.

Finally, we perform a robustness check inspired by Ruhose & Schwerdt (2016) of estimating our difference-in-differences specification with native students only, several times and excluding each time a different country. Results of this piecewise country deletion robustness check show that in fact results are not driven by the inclusion of a single country as all estimates of the interaction term remain statistically significant and around the previously calculated magnitude.

1.5 Conclusion

The analysis carried out in this article provides results about the effects of a specialized versus a general upper-secondary school curriculum on students' performance and inequality, which to the best of our knowledge is the first analysis regarding this topic.

Because countries only follow one system or the other, a cross-country analysis is required to estimate the possible effects of these institutional differences. We choose an international difference-in-differences analysis to account for country heterogeneity and unobserved factors influencing student outcomes. The impact of having a specialized system in upper-secondary school, as opposed to a general one, is identified by comparing the change in outcomes between a point just before students have started this stage of schooling (measured using PISA data) and a later point after they have completed it (using data from the PIAAC study) across countries with specialized and general systems. Results suggest that there is a negative effect from specialization tracking in both mathematics and reading achievement, and that these may be greater for specific sub-groups of the student population – in particular, students with a low socioeconomic status. While we have no way of assessing the mechanism of these effects with PISA and PIAAC data, we believe that they may be related with the specialization track choice of these different sub-groups of students in specialized systems. In fact, when analyzing data from TIMSS Advanced 1995 study, we see that there are differences between the socioeconomic background and other individual characteristics of students of different specialization tracks, within specialized systems. Moreover, we find large differences in the type of courses that students of different specialization tracks take during upper-secondary school. Thus, the pattern of track choices in specialized systems of different SES groups and genders could be related with the pattern of results we estimate for different sub-groups of the student population. In fact, the

descriptive analysis with TIMSS data is in line with the findings in the literature, described in Section 1.1, that students from higher socioeconomic backgrounds are more likely to choose more prestigious and academically demanding specialization tracks, such as the math and science tracks.

A limitation of our study however, is that, while the classification of educational systems is based on official information, some scholars have shown that systems might work quite differently than stated in these official descriptions. For instance, while we assume that lower-secondary programs all have a general curriculum, in fact some preparation for the later specialization tracks might already be occurring. Boone et al. (2018) describe how in Flanders, the A-stream in lower-secondary school, which is classified as a general program, actually already makes students choose some optional courses which are related with the track chosen in upper-secondary school. Tarabini et al. (2018) also describes that in Catalonia, some ability-grouping is informally done which determines track choice in upper-secondary school. Unfortunately, with our data we are not able to check whether students have been part of these programs in lower secondary school that prepare for or determine later specialization track choices. If we had student level data on lower-secondary course choices, we could try to disentangle the effects of specialization tracking by groups of students with different lower-secondary school preparation. Also, the datasets used in our study do not allow us to follow the same student over time. If it was possible to link the students across different assessments in time, the impact of specialized systems could be more precisely estimated using DID regressions. Such an approach would however require using administrative datasets for each country. However, that approach would be difficult to implement as different countries use different assessments systems.

Taken together, these results suggest that some specific groups of the student population lose from a system that tracks students by specialization areas, while some may even gain from it. Because specialization track choices of students are also directly related with enrollment choices in tertiary education, and consequently, labor market outcomes, we believe that future research could also address its effect on these two issues. Educational policy should also consider the possible negative effects of specialization tracking. In the case that specialized systems are detrimental, and more so for students from low socioeconomic status, it might be relevant to consider and analyze the implementation of policies that delay the specialization track choice, reduce the curricular differences between tracks, or provide more guidance in track choice.

Chapter 2

Pupil Well-being in Danish Primary and Lower Secondary Schools⁷

Joint work with Anna Folke Larsen and Marianne Simonsen

2.1 Introduction

Children spend a large fraction of their time in school, with the primary purpose of accumulating academic skills. At the same time, pupil mental health and well-being more broadly have received considerable interest (e.g. Tsang, Wong and Lo, 2012, WHO, 2016) and one of the three national objectives of the Danish public school is, in fact, to increase pupil well-being. Since 2014, the Ministry of Education has conducted yearly national well-being surveys (Nationale trivselsmålinger) for children of all ages in public school. The survey was introduced as a tool for schools to monitor the subjective well-being of their pupils, to make informed adjustments of their own related practices, and to inform education policy at the municipal level. These data hold an untapped potential to expand our knowledge of pupil well-being in Danish classrooms, to the benefit of parents, teachers, and policy-makers. But well-being measures could also be used to inform and enrich analyses of school-related interventions and maybe even social policy. This chapter explores three specific research questions: 1) who in the classroom experiences low and high levels of well-being and, implicit in this, what is the role of peers and schools? 2) Is there a tendency for the same children to experience low levels of well-being over time? And finally, 3) what is the relationship between pupil well-being, academic performance, and absence?

⁷ Working paper published in the [Aarhus BSS Economics Working Papers](#), Aarhus University.

The starting point for our analyses is the school social well-being segment of the nationally administered well-being survey, which we combine with register-based, population-wide data available for the children in question as well as for their families and teachers. In the interest of simplicity, we will subsequently refer to school social well-being as “social well-being”.

Our approach is pragmatic and positive in that we describe the existing well-being measure as is and study its relationship to other variables that decision-makers usually care for; our purpose is *not* to validate the measure formally. As such, our analysis complements a recent paper that speaks more directly to the content of the well-being scales: Niclasen et al. (2018) study the psychometric properties of the national well-being surveys for grades 4-9. They generally support the idea of a factor structure and their measure of “school connectedness” is closely related to the social well-being measure.⁸ Our paper is also related to Andersen et al. (2020) who use the national well-being survey to study the relationship between personality and academic performance over time. Their measure of emotional stability, in particular, is constructed from three questions that also enter the social well-being measure and is positively correlated with reading scores.

We first confirm previous findings that show that pupils on average report high social well-being (Knoop, et al., 2017). Three quarters of children across grades 1-9 report high levels of social well-being, while only slightly more than one pupil per class reports average social well-being below the midpoint of the scale. Though many have high levels of well-being, the share of children reporting low well-being is still non-negligible. We continue to document that lower social well-being correlates meaningfully with standard measures of disadvantage at the pupil

⁸ Social well-being, as studied in our paper, corresponds closely to what they label school connectedness.

and parental level. Although we stress that we are not making causal claims, we also note that teacher characteristics and classroom composition are additional important predictors of well-being. A particular concern is that well-being is a latent characteristic that is notoriously difficult to measure, and a pupil's answers to the test may well be sensitive to temporary conflicts on the day of the test (Tsang, Wong and Lo, 2011). Yet, we document high degrees of persistence in social well-being over time; the (conditional) correlation between current social well-being and social well-being one year prior is as high as .45, whereas the correlation between current social well-being and social well-being two years back in time is about .15. These results hold whether or not we control for a wide range of background characteristics. We finally show that social well-being is positively associated with academic performance and negatively associated with absence from school; both outcomes that are straightforward in their interpretation. An underlying assumption that permeates work on pupil well-being is, in fact, the inherent expectation of a positive link with academic performance (Adler, 2017; Conti and Heckman, 2014). Taken together and in the light of the previous results of Andersen et al. (2020) and Niclasen et al. (2018), these results suggest that the current measure could reasonably be used as a first, convenient indicator of well-being among Danish pupils.

The remainder of the chapter is structured as follows: Section 2.2 provides a brief description of elementary and lower secondary school in Denmark, and Section 2.3 describes the data and our sampling strategies. Section 2.4 shows how social well-being correlates with pupil, parent, teacher, and classroom characteristics, and Section 2.5 documents the association between social well-being and other important outcomes. Finally, Section 2.6 concludes.

2.2 Measuring well-being in elementary and lower secondary school in Denmark

In December 2013, The Danish Ministry of Education established an expert group that were to guide the development of the national well-being survey. Their work was accompanied by a pilot study (Keilow et al., 2014) and the final version was implemented in Danish public school starting in the spring of 2015.

In Denmark, compulsory education comprises primary and lower secondary education (ISCED 1 and 2) and lasts 10 years, from grade 0 to grade 9, with the possibility of attending an optional 11th year (grade 10). Children are supposed to enter school in the year in which they turn six years old. Among children in primary and lower secondary education, 76 percent attend the municipal public school, Folkeskolen (as shown in Table 2.3). In public schools, children are divided into classes of maximum 28 pupils during grade 0. In practice, however, the average class size is much smaller amounting to 21.3 pupils on average (see Table A2.2). Typically, children stay together in these classes until they leave school. A class receives education in all subjects together, and is headed by a “class teacher”, who follows the class for several years. This teacher, who is usually also the Danish or Math teacher of the class, coordinates the activities of the group of subject teachers associated with the class, and is the primary point person in cases of academic, behavioral or social problems. While a teacher is usually only class teacher for one class, subject teachers teach their subjects to several classes.

2.3 Data, key measures, and samples

We make use of a series of data sources with individual level information about children, their families, and their teachers. Key to our analyses are the nationally administered well-being surveys developed by the Danish Ministry of Education, but we also exploit nationally administered IT-based tests of Danish reading skills and Math (see Beuchert and Nandrup 2018), absence data from the Danish Ministry of Education, administrative data linking teachers to classes, and register-based data maintained by Statistics Denmark. In addition to socio-economic background information, the latter includes information about 9th grade GPA. We merge this variety of data sources by use of the central personal register number, a unique individual identifier available for all Danish residents.

Our primary variable of interest is a measure of attitudes towards school and emotional well-being in the classroom. We base the measure on the recently implemented national well-being indicators that are collected during the first quarter of each calendar year (Andersen et al., 2015). The well-being survey varies across primary and lower secondary school, with children in primary school (grades 0-3) answering 20 questions and children in lower secondary school (grades 4-9) answering 40 questions. Among the full list of questions in the national well-being survey, we use only the ten questions that enter into the social well-being subscale for children in lower secondary school (The Danish Ministry of Education, 2016).⁹ The Ministry only constructs a social well-being subscale for grade 4-9. Still, our overarching goal was to

⁹ The subscales are constructed based on factor analyses that reveal high levels of reliability (Ministry of Education, 2016). The social well-being scale is closely related to the “school connectedness” scale of Niclasen et al. (2018). In fact, all of the seven questions entering into “school connectedness” are included in the social well-being subscale. The latter includes three additional questions: “Are you afraid of being ridiculed at school?”, “Since the start of the school year, did anyone bully you?”, and “I like the breaks at school”.

investigate well-being across lower secondary *as well as* primary school. Among the ten questions, therefore, we have found eight questions in the questionnaire for grade 0-3 that correspond well to those included in the grade 4-9 social well-being subscale.¹⁰ We use these to construct a similar social well-being measure for primary school children. The answers to all questions for children in primary school are coded to range from one to three, with three being the most positive. For children in lower secondary school, the answers are coded to range from one to five, with five being the most positive.¹¹ We list the questions in Table 2.1. In line with the Danish Ministry of Education (2019), we subsequently calculate social well-being as the within-individual average of the answers provided.

Table 2.1
Questions that enter the social well-being indicator

Grades 0-3	Grades 4-9
Do you like your school?	Do you like your school?
Do you like the other children in your classroom?	Do you like the other children in your classroom?
Do you feel lonely at school?	Do you feel lonely?
Are you afraid that the other children at school will laugh at you?	Are you afraid of being ridiculed at school?
Does anyone at school tease you and make you sad?	How often do you feel safe at school?
Do you like the breaks at school?	Since the start of the school year, did anyone bully you?
Are the children in your classroom good at helping each other?	I feel I belong at my school.
Do you think the other children in your classroom like you?	I like the breaks at school.
	Most of the pupils in my classroom are kind and helpful.
	Other pupils accept me as I am.

¹⁰ As will be clear below, we do exclude children in grade 0 because of concerns about their ability to independently fill out the survey.

¹¹ For positive questions like “Do you feel safe at school?” the value five (three) corresponds to “very often” (“yes, very”). For negative questions like “Do you feel lonely?” five (three) corresponds to “never” (“no”). In this sense, a higher number is always a better outcome.

Table 2.2 and Figure 2.1 together show aspects of the distribution of social well-being for children in primary school and lower secondary school. The first and very positive message is that the vast majority of children report high levels of social well-being. On the scale from 1 to 5 we consider values above (or equal to) 4 as high levels of social well-being in line with Knoop et al. (2017) and values below (or equal to) 3 as low levels of social well-being in line with Nielsen and Rangvid (2016). For children in grade 1-3 we use 2 (the midpoint on the scale) as the natural cut-off. More than 60% of children across grades 4-9 report average levels of social well-being above 4, and on the smaller scale for children in grades 1-3, 90% report above 2.¹² Still, non-negligible shares of children report lower levels of well-being. Roughly 6% of children in grades 4-9, for example, have average social well-being below 3 corresponding to a little more than one pupil per class; and as many as 30% of children in 4-9 have at least one report below 3. Persistency in low social well-being initially appears low – 0.6% of the children exposed to at least three survey rounds (grades 6-9) score on average 3 or lower across all of the three most recent years. Yet, in a relative sense, 0.6% is still non-negligible: it amounts to 10% of the children who have low levels of well-being in the most recent year.

Here, it is also important to note that 15% of the relevant children do not answer the well-being survey in a given year, and an even larger fraction misses at least one of the well-being surveys in the past three years (see Table 2.3). Our analyses suggest that these children are more likely to have low levels of well-being that go unreported, which indicates that we are likely to

¹² For children in grade 4-9 the mean social well-being is 4.09 (standard deviation 0.63) and for children in grade 1-3 the mean is 2.56 (standard deviation 0.34).

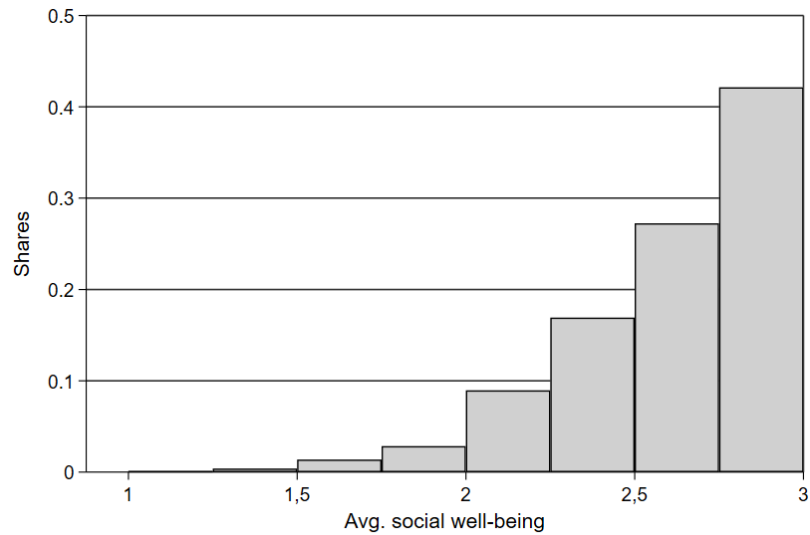
underestimate the prevalence and persistence of low levels of well-being. We return to the role of survey response below.

Table 2.2
Low and high social well-being

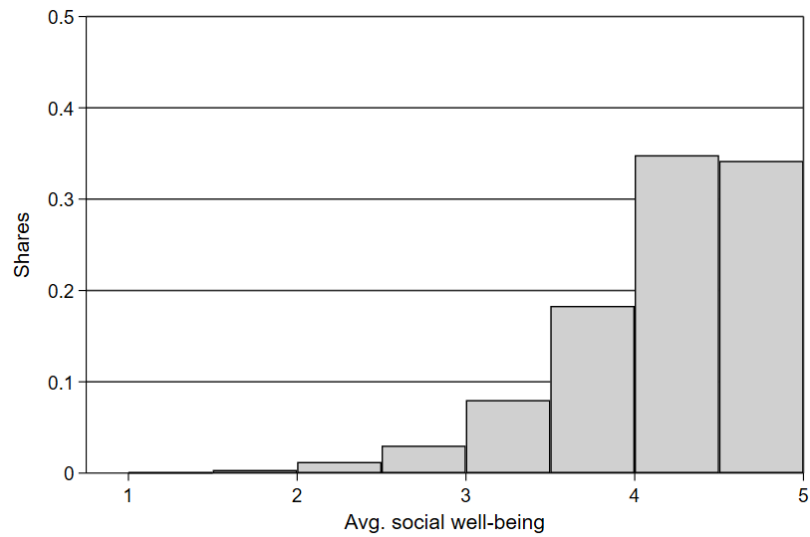
	Grades 1-9	Grades 1-3	Grades 4-9	Grades 6-9
<i>Low well-being:</i>				
Average well-being < 3 (2)	0.056	0.052	0.058	0.054
Average well-being ≤ 3 (2)	0.078	0.092	0.071	0.066
Any answers < 3 (2)	0.280	0.224	0.312	0.300
Any answers ≤ 3 (2)	0.744	0.870	0.674	0.667
Average well-being < 3 across last three years	N/A	N/A	N/A	0.005
Average well-being ≤ 3 across last three years	N/A	N/A	N/A	0.006
<i>High well-being:</i>				
Average well-being > 4 (2)	0.711	0.908	0.601	0.600
Average well-being ≥ 4 (2)	0.764	0.948	0.662	0.663
Any answers > 4 (2)	0.934	0.966	0.916	0.922
Any answers ≥ 4 (2)	0.993	0.999	0.989	0.991
Average well-being > 4 across last three years	N/A	N/A	N/A	0.350
Average well-being ≥ 4 across last three years	N/A	N/A	N/A	0.413
# Observations answered survey, last year	380,680	136,003	244,677	150,750
# Observations answered survey, last three years	N/A	N/A	N/A	127,090

Notes: This table shows shares of pupils across various levels of social well-being. Social well-being is on a scale from 1 to 3 for 1st to 3rd grade pupils and on a scale from 1 to 5 for 4th to 9th grade pupils, hence threshold values depend on the grade: We first present thresholds for 4th to 9th grade pupils, while thresholds in parentheses are for 1st to 3rd grade.

Figure 2.1
Distribution of social well-being



(A) Grades 1-3



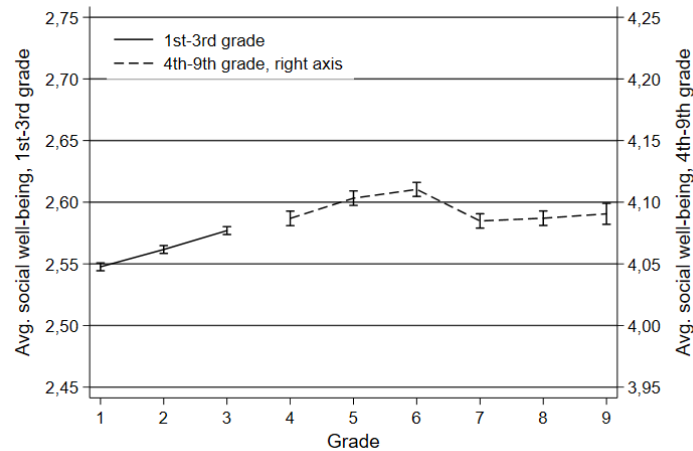
(B) Grades 4-9

Notes: This figure shows the distribution of average (within-individual) social well-being for the *cross-sectional* sample (see Table 2.2), by grades 1-3 (primary school) and grades 4-9 (lower secondary school).

Figure 2.2 shows the average social well-being by grade for both the primary school measure (left axis) and secondary school measure (right axis). Well-being increases during the primary school years. In secondary school, average social well-being increases until age 11-12 followed by a decrease until the end of lower secondary school. However, while this pattern is clear, the differences themselves are in fact very small. In practice, in order to study social well-being across grades using one common specification, we standardize the measure at each grade level to have a mean of zero and a variance of one.

Figure 2.2

Social well-being across grades



Notes: This figure shows average social well-being by grades for the *cross-sectional* sample (see Table 2.2).

Equipped with a measure of social well-being, we next select our analysis samples. Our starting point is the group of children enrolled in publicly provided primary and lower secondary school in the 2018/2019 school year. Among these, we select children enrolled in grades 1-9. Table 2.3 shows our sample selection criteria and their implications for our sample size. Critically, not all children complete the well-being surveys; our analyses will explore this selection in detail. In practice, we will consider a child to have completed the social well-being segment of the survey if he/she provides an answer to more than 50% of the questions related to social well-being. Most of our analyses will make use of the *cross-sectional* sample, consisting of the 380,680 children who all answered more than 50% of the questions related to social well-being. In addition to speak about persistency in social well-being, some analyses will exploit a *longitudinal* sample.

This sample consists of children who not only completed the social well-being questions in the 2018/2019 school year, but who also answered the social well-being survey in the previous two years. Naturally, this sample is smaller by about 130,000 children. This is primarily because we, by construction, exclude children in grades 1-2.

Table 2.3
Sample selection

Sample selection criteria	Sample size	% of original sample	Sample reduction
Pupils in primary and lower secondary school 2018/2019	696,543	100.00	-
... enrolled in public schools	530,497	76.16	-166,046
... in grades 1-9	472,250	67.80	-58,247
... in regular classrooms separated by grade	462,671	66.42	-9,579
... didn't immigrate after 6th birthday	453,206	65.07	-9,465
... in classrooms of 5 to 40 students	449,971	64.60	-3,235
... who participated in the well-being survey	382,921	54.97	-67,050
... and who answered more than 50% of the questions related to social well-being	380,680	54.65	-2,241
... And who has answered the survey the last three consecutive years	247,987	35.60	-132,693

Notes: This table shows the sample selection criteria along with the consequences for sample size.

2.4 Correlates of social well-being

2.4.A Individual characteristics and family background

Through a straightforward multiple regression analysis based on the cross-sectional sample, we first consider the (conditional) correlations between social well-being and pupil as well as family background characteristics. We primarily focus on groups of variables that signify prior disadvantage or disruptions at the individual as well as at the family level. This is both because these variables are of interest to decision makers and because a wide range of studies suggest that this type of variables is associated with – and even has a causal effect on – future child human capital and well-being more broadly. See for example excellent surveys by Almond and Currie (2010), Almond and Currie (2011), and Almond, Currie and Duque (2018). Some

examples include birth weight (Black et al., 2005) and childhood health more generally (Currie et al., 2010). Age at school start (Black et al., 2008, Landersø et al., 2017) and family income (Dahl and Lochner, 2012; Aizer et al., 2016) are other prominent, well-documented examples. To the extent that the social well-being scale is unassociated with this range of variables, there is reason to worry about the nature and information content of the measure.

In practice, for the pupils themselves, in addition to gender, we consider measures of disadvantage at early life (birth weight); school (age at school start, grade repetitions, school switches, and special needs); social disadvantage (municipal preventive measures and out-of-home placement); and health disadvantage (hospital use and psychiatric contacts).¹³ At the family level, we consider measures of parental income, education, receipt of unemployment insurance (UI) or cash benefits, health, and incidents of crime as well as a range of indicators of family stability (birth order, family size, separations, and single parenthood). Table A2.1 presents the full list of variables, their definitions, and sources.

We present our main results in Table 2.4. Column 1 is our baseline model that controls for observable characteristics at the pupil and family level.¹⁴ Column 2 presents the full model adding teacher, school and classroom characteristics to the list of controls. Column 3 adds classroom random effects and Column 4 adds classroom fixed effects to account for unobservable factors that correlate with both observable characteristics and social well-being. Importantly, all our main conclusions are qualitatively robust to these variations in model

¹³ We also condition on migrant status but refrain from interpreting the coefficients because migrant status correlate highly with indicators for missing information on other variables such as birthweight and parental educational attainment.

¹⁴ We do control for a series of variables describing immigrant status but because these variables correlate highly with our indicators for missing family level information, we refrain from showing and commenting on these results.

specification. Table A2.2 complements the regression exercise by instead showing how means of the included observable characteristics vary across the distribution (quartiles) of social well-being. Overall, we find that social well-being is vastly higher among boys (about 25% of a standard deviation or 0.16 points on the scale from 1 to 5), which is in contrast to earlier studies linking gender and measures of subjective school well-being; see, for example, Bradshaw et al. (2011).¹⁵ This result is particularly interesting given the public debate on how the organization of (Danish) schools tend to favor girls.¹⁶ Contrary to expectations, social well-being also increases slightly with low birthweight.¹⁷ We note that social well-being correlates negatively – and significantly – with a pupil’s own prior experiences of disadvantage. These include late school start, school switches, contacts with psychiatric hospitals, and special needs more generally. Some estimates are particularly large: indicators of receiving preventive measures, having contact to a psychiatric hospital, and being a special needs child in an ordinary classroom all predict lower well-being and some estimates are as large as 25% of a standard deviation. Meanwhile, family background also correlates with well-being: lower parental income, low education levels, receipt of UI or cash benefits, psychiatric hospitalization and incidents of crime are all strong predictors of lower social well-being too. None of the parental variables appears as important as the pupil’s own profile, but the occurrence of family separations, measured as an indicator for whether parents have lived apart for two consecutive years, stand out with a

¹⁵ Bradshaw et al. (2011) do find that boys have higher personal and family well-being than girls.

¹⁶ See for example “We need to change the culture among boys in the classroom” (“Vi skal ændre drengekulturen i klasseværelset”), April 12 2018 (<https://videnskab.dk/kultur-samfund/forsker-vi-skal-aendre-drengekulturen-i-klassevaerelset>) and “Head of teacher’ union: We need to accommodate boys’ needs in schools” (“Lærerformand: Skolegang skal indrettes mere til drenge”), September 20, 2015 (<https://www.berlingske.dk/samfund/laererformand-skolegang-skal-indrettes-mere-til-drenge>).

¹⁷ While many studies argue that birth weight correlates negatively with later academic performance, a recent Danish study (Maruyama and Heinesen, 2020) shows, in contrast, no strong evidence for non-health long-run outcomes, such as test scores.

negative coefficient of 9% of a standard deviation or 0.06 points on the scale from 1 to 5. If we run two separate regressions with child characteristic and parent/household characteristics separately, we find that child characteristics can explain almost twice as much of the variation in social well-being as parent/household characteristics.¹⁸

Given this, the social well-being scale seems to relate more closely to the pupil's own experiences at school as well as his or her level of human capital. Parental and family background does correlate with well-being but to a lesser extent. Of course, many other factors contribute to well-being; our baseline model explains just above 4% of the observed variation.

¹⁸ Results available upon request.

Table 2.4
Correlates of social well-being

	Baseline model	Full model	Random effects model	Fixed effects model
<i>Child characteristics</i>				
Male	0.24669 *** (0.00366)	0.24551 *** (0.00369)	0.24517 *** (0.00368)	0.25433 *** (0.00363)
Birthweight is below 2500 grams	0.02806 *** (0.00774)	0.02767 *** (0.00765)	0.02783 *** (0.00762)	0.02037 *** (0.00755)
Age relative to cohort (Ref.cat.: Adequate for grade)				
Young for grade	-0.02224 (0.01424)	-0.01880 (0.01403)	-0.01673 (0.01401)	-0.01679 (0.01385)
Late school start	-0.04072 *** (0.00654)	-0.04776 *** (0.00638)	-0.04740 *** (0.00638)	-0.05400 *** (0.00645)
Repeated a grade	-0.05889 *** (0.00932)	-0.06191 *** (0.00895)	-0.06306 *** (0.00895)	-0.06994 *** (0.00916)
Switched address, since school start	0.00806 (0.00458)	0.00023 (0.00438)	-0.00164 (0.00436)	-0.00316 (0.00444)
Switched school, since school start	-0.08179 *** (0.01011)	-0.08826 *** (0.01004)	-0.08748 *** (0.01002)	-0.08180 *** (0.01001)
Switched school but not address , since school start	-0.01260 (0.01079)	-0.02222 * (0.01072)	-0.02374 * (0.01071)	-0.03274 *** (0.01074)
Placed outside home, from 2010	0.09120 *** (0.02077)	0.08329 *** (0.02054)	0.08674 *** (0.02040)	0.07572 *** (0.02068)
Received social preventive measures by municipality, since birth	-0.18007 *** (0.00827)	-0.17872 *** (0.00813)	-0.18159 *** (0.00811)	-0.18161 *** (0.00814)
Avg. no. of contacts per year to hospital, last 5 years	-0.02971 *** (0.00305)	-0.02796 *** (0.00298)	-0.02716 *** (0.00297)	-0.02696 *** (0.00299)
Been in contact with psychiatric hospital, since birth	-0.23508 *** (0.00933)	-0.24065 *** (0.00912)	-0.24292 *** (0.00909)	-0.24799 *** (0.00931)
Special needs student in ordinary classroom	-0.19383 *** (0.03673)	-0.18112 *** (0.03630)	-0.18062 *** (0.03609)	-0.16995 *** (0.03716)
Special needs student in special needs classroom	-0.03055 (0.01752)	0.00217 (0.03185)	0.00709 (0.03194)	0.02267 (0.01511)

Notes: Dependent variable is social well-being. Column 1 shows results including child, parental, and household characteristics. Column 2 adds teacher, school, and classroom characteristics. Column 3 adds classroom random effects, while column 4 adds classroom fixed effects. Indicators for missing values are included in all regressions. Standard errors clustered at the classroom level and shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. No. observations 380,680; no. classrooms 21,135. R^2 -adjusted of 0.039 (column 1), 0.044 (column 2), 0.045 (column 3), and 0.037 (column 4). Fraction of variance due to variation between classrooms (rho) 0.104 (column 3, random effects), 0.167 (column 4, fixed effects).

Table 2.4 continued
Correlates of social well-being

	Baseline model		Full model		Random effects model		Fixed effects model	
<i>Maternal characteristics in 2018</i>								
Age at child birth	0.00317 *** (0.00046)		0.00253 *** (0.00046)		0.00242 *** (0.00045)		0.00231 *** (0.00045)	
Education (Ref.cat.: Vocational)								
Unskilled	-0.01877 ** (0.00627)		-0.01113 (0.00625)		-0.00825 (0.00605)		-0.00788 (0.00612)	
Highschool	0.00295 (0.00799)		0.00091 (0.00798)		0.00105 (0.00776)		-0.00022 (0.00786)	
Short further	0.00846 (0.00733)		0.00393 (0.00731)		0.00087 (0.00707)		-0.00189 (0.00715)	
Medium further	0.05408 *** (0.00445)		0.04937 *** (0.00439)		0.04632 *** (0.00429)		0.03890 *** (0.00432)	
Long further or PhD	0.04805 *** (0.00618)		0.03929 *** (0.00610)		0.04148 *** (0.00593)		0.03700 *** (0.00600)	
Taxable income in 2nd quartile, last year	0.01165 * (0.00518)		0.01121 * (0.00514)		0.01294 * (0.00511)		0.01294 ** (0.00507)	
Taxable income in 3rd quartile, last year	0.01784 ** (0.00555)		0.01641 ** (0.00548)		0.01922 *** (0.00543)		0.02302 *** (0.00539)	
Taxable income in 4th quartile, last year	0.03008 *** (0.00586)		0.02487 *** (0.00575)		0.02616 *** (0.00570)		0.03051 *** (0.00567)	
Received unemployment assistance, last year	-0.01372 ** (0.00448)		-0.01285 ** (0.00445)		-0.01242 ** (0.00435)		-0.01102 ** (0.00436)	
Received unemployment insurance, last year	-0.04471 *** (0.00667)		-0.04222 *** (0.00665)		-0.03863 *** (0.00649)		-0.03583 *** (0.00655)	
Avg. no. of contacts per year to hospital, last 5 years	-0.00435 * (0.00184)		-0.00376 * (0.00183)		-0.00548 ** (0.00177)		-0.00631 *** (0.00179)	
Been in contact with psychiatric hospital, since birth of pupil	-0.02676 *** (0.00633)		-0.02677 *** (0.00628)		-0.02766 *** (0.00620)		-0.02818 *** (0.00622)	
Committed criminal offense, since birth of pupil	-0.00983 (0.01134)		-0.00599 (0.01134)		-0.00527 (0.01113)		-0.00162 (0.01130)	

Notes: Dependent variable is social well-being. Column 1 shows results including child, parental, and household characteristics. Column 2 adds teacher, school, and classroom characteristics. Column 3 adds classroom random effects, while column 4 adds classroom fixed effects. Indicators for missing values are included in all regressions. Standard errors clustered at the classroom level and shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. No. observations 380,680; no. classrooms 21,135. R^2 -adjusted of 0.039 (column 1), 0.044 (column 2), 0.045 (column 3), and 0.037 (column 4). Fraction of variance due to variation between classrooms (rho) 0.104 (column 3, random effects), 0.167 (column 4, fixed effects).

Table 2.4 continued
Correlates of social well-being

	Baseline model	Full model	Random effects model	Fixed effects model
<i>Paternal characteristics in 2018</i>				
Age at child birth	-0.00170 *** (0.00035)	-0.00180 *** (0.00035)	-0.00169 *** (0.00034)	-0.00167 *** (0.00034)
Education (Ref.cat.: Vocational)				
Unskilled	-0.03359 *** (0.00543)	-0.02766 *** (0.00541)	-0.02300 *** (0.00524)	-0.01969 *** (0.00530)
Highschool	0.00022 (0.00769)	-0.00129 (0.00766)	0.00349 (0.00742)	0.00609 (0.00751)
Short further	0.01232 * (0.00627)	0.00862 (0.00626)	0.00902 (0.00605)	0.00944 (0.00611)
Medium further	0.02006 *** (0.00520)	0.01615 ** (0.00516)	0.01638 ** (0.00500)	0.01366 *** (0.00506)
Long further or PhD	0.02892 *** (0.00581)	0.01966 *** (0.00577)	0.02465 *** (0.00556)	0.02441 *** (0.00564)
Taxable income in 2nd quartile, last year	-0.01190 * (0.00515)	-0.01056 * (0.00511)	-0.00957 (0.00508)	-0.01007 ** (0.00504)
Taxable income in 3rd quartile, last year	0.01407 ** (0.00532)	0.01234 * (0.00526)	0.01337 * (0.00523)	0.01419 *** (0.00521)
Taxable income in 4th quartile, last year	0.03622 *** (0.00558)	0.02735 *** (0.00546)	0.02802 *** (0.00542)	0.02435 *** (0.00541)
Received unemployment assistance, last year	-0.01341 ** (0.00516)	-0.01127 * (0.00514)	-0.01252 * (0.00501)	-0.01135 ** (0.00505)
Received unemployment insurance, last year	-0.03694 *** (0.00773)	-0.03346 *** (0.00772)	-0.02992 *** (0.00753)	-0.02635 *** (0.00761)
Avg. no. of contacts per year to hospital, last 5 years	-0.00540 * (0.00221)	-0.00520 * (0.00220)	-0.00659 ** (0.00215)	-0.00732 *** (0.00218)
Been in contact with psychiatric hospital, since birth of pupil	-0.01829 * (0.00733)	-0.01782 * (0.00731)	-0.01745 * (0.00719)	-0.01602 ** (0.00723)
Committed criminal offense, since birth of pupil	-0.03019 *** (0.00718)	-0.02608 *** (0.00716)	-0.02841 *** (0.00700)	-0.02343 *** (0.00705)

Notes: Dependent variable is social well-being. Column 1 shows results including child, parental, and household characteristics. Column 2 adds teacher, school, and classroom characteristics. Column 3 adds classroom random effects, while column 4 adds classroom fixed effects. Indicators for missing values are included in all regressions. Standard errors clustered at the classroom level and shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. No. observations 380,680; no. classrooms 21,135. R^2 -adjusted of 0.039 (column 1), 0.044 (column 2), 0.045 (column 3), and 0.037 (column 4). Fraction of variance due to variation between classrooms (ρ) 0.104 (column 3, random effects), 0.167 (column 4, fixed effects).

Table 2.4 continued
Correlates of social well-being

	Baseline model		Full model		Random effects model		Fixed effects model
<i>Household characteristics in 2018</i>							
Number of siblings	0.00493 *		0.00471 *		0.00371 *		0.00331 *
	(0.00196)		(0.00195)		(0.00189)		(0.00191)
Birth order on mother's side	-0.02210 ***		-0.02045 ***		-0.02076 ***		-0.02105 ***
	(0.00244)		(0.00243)		(0.00235)		(0.00237)
Parents have lived apart for two consecutive years	-0.09424 ***		-0.09103 ***		-0.08827 ***		-0.08656 ***
	(0.00583)		(0.00582)		(0.00567)		(0.00574)
Single provider household	-0.02704 ***		-0.02338 ***		-0.01920 ***		-0.01644 ***
	(0.00593)		(0.00592)		(0.00576)		(0.00583)
<i>Teacher characteristics</i>							
Male, Danish			-0.02558 ***		-0.02677 ***		
			(0.00743)		(0.00751)		
Age, Danish			0.00017		0.00013		
			(0.00036)		(0.00036)		
Educated in teaching Danish or comparable			0.01969		0.02460		
			(0.01693)		(0.01697)		
Years since degree, 0-10 years, Danish			0.00047		0.00057		
			(0.00212)		(0.00214)		
10 years or above since degree, Danish			0.01633		0.01689		
			(0.01472)		(0.01487)		
Tenure of teacher, Danish(Ref.cat.: 0-1 year)							
Tenure is 2 to 3 years, Danish			0.04307 ***		0.04233 ***		
			(0.00936)		(0.00942)		
Tenure is 4 to 5 years, Danish			0.07464 ***		0.07503 ***		
			(0.00935)		(0.00942)		
Tenure is 6 years or above, Danish			0.08751 ***		0.08662 ***		
			(0.00929)		(0.00934)		
Share of days absent, Danish			-0.09571 **		-0.10360 **		
			(0.03358)		(0.03397)		

Notes: Dependent variable is social well-being. Column 1 shows results including child, parental, and household characteristics. Column 2 adds teacher, school, and classroom characteristics. Column 3 adds classroom random effects, while column 4 adds classroom fixed effects. Indicators for missing values are included in all regressions. Standard errors clustered at the classroom level and shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. No. observations 380,680; no. classrooms 21,135. R^2 -adjusted of 0.039 (column 1), 0.044 (column 2), 0.045 (column 3), and 0.037 (column 4). Fraction of variance due to variation between classrooms (ρ) 0.104 (column 3, random effects), 0.167 (column 4, fixed effects).

Table 2.4 continued
Correlates of social well-being

	Baseline model	Full model	Random effects model	Fixed effects model
Age, Math		-0.00028 (0.00035)	-0.00023 (0.00036)	
Male, Math		-0.00132 (0.00573)	-0.00070 (0.00575)	
Educated in teaching math or comparable		0.02103 (0.01194)	0.02319 (0.01204)	
Years since degree, 0-10 years, Math		0.00020 (0.00214)	-0.00029 (0.00217)	
10 years or above since degree, Math		0.01159 (0.01433)	0.01092 (0.01443)	
Tenure of teacher, Math(Ref.cat.: 0-1 year)				
Tenure is 2 to 3 years, Math		0.02250 * (0.00906)	0.02394 ** (0.00915)	
Tenure is 4 to 5 years, Math		0.03028 ** (0.00924)	0.03239 *** (0.00928)	
Tenure is 6 years or above, Math		0.04094 *** (0.00916)	0.03950 *** (0.00923)	
Share of days absent, Math		-0.03356 (0.03284)	-0.02680 (0.03276)	
<i>School characteristics</i>				
Turnover rate of Danish and math teachers		-0.00035 (0.00059)	-0.00039 (0.00060)	
Turnover rate of all teachers		-0.00013 (0.00060)	-0.00008 (0.00060)	
Size		-0.00003 (0.00001)	* -0.00004 (0.00001)	**

Notes: Dependent variable is social well-being. Column 1 shows results including child, parental, and household characteristics. Column 2 adds teacher, school, and classroom characteristics. Column 3 adds classroom random effects, while column 4 adds classroom fixed effects. Indicators for missing values are included in all regressions. Standard errors clustered at the classroom level and shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. No. observations 380,680; no. classrooms 21,135. R^2 -adjusted of 0.039 (column 1), 0.044 (column 2), 0.045 (column 3), and 0.037 (column 4). Fraction of variance due to variation between classrooms (ρ) 0.104 (column 3, random effects), 0.167 (column 4, fixed effects).

Table 2.4 continued
Correlates of social well-being

	Baseline model	Full model	Random effects model	Fixed effects model
<i>Classroom characteristics</i>				
Size		0.00104 (0.00077)	0.00072 (0.00077)	
Share of males		-0.17980 *** (0.02412)	-0.17843 *** (0.02445)	
Share with birthweight below 2500 grams		0.11592 * (0.04575)	0.12143 ** (0.04630)	
Share of who are young for grade		-0.04935 (0.09009)	0.00225 (0.09222)	
Share of late starters		0.09067 ** (0.03241)	0.10756 ** (0.03319)	
Share who repeated a grade		0.09999 ** (0.03795)	0.08552 * (0.03903)	
Share who switched address, since school start		0.02527 (0.02088)	0.01234 (0.02103)	
Share who switched school, since school start		-0.09736 (0.05868)	-0.09593 (0.05969)	
Share who switched school but not address, since school start		0.17048 ** (0.06053)	0.15825 * (0.06156)	
Share who have been placed outside home, from 2010		0.21594 * (0.09155)	0.21928 * (0.09202)	
Share who received social preventive measures by municipality, since birth		-0.02262 (0.03922)	-0.03029 (0.04014)	
Avg. no. of contacts per year to hospital, last 5 years		-0.01619 (0.01542)	-0.00630 (0.01555)	
Share who has been in contact with psychiatric hospital, since birth		0.07794 * (0.03884)	0.06127 (0.03937)	
Share with special needs in ordinary classrooms		-0.14418 (0.14598)	-0.14170 (0.13906)	

Notes: Dependent variable is social well-being. Column 1 shows results including child, parental, and household characteristics. Column 2 adds teacher, school, and classroom characteristics. Column 3 adds classroom random effects, while column 4 adds classroom fixed effects. Indicators for missing values are included in all regressions. Standard errors clustered at the classroom level and shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. No. observations 380,680; no. classrooms 21,135. R^2 -adjusted of 0.039 (column 1), 0.044 (column 2), 0.045 (column 3), and 0.037 (column 4). Fraction of variance due to variation between classrooms (rho) 0.104 (column 3, random effects), 0.167 (column 4, fixed effects).

Table 2.4 continued
Correlates of social well-being

	Baseline model	Full model	Random effects model	Fixed effects model
Share with a criminal parent		-0.21287 *** (0.02881)	-0.16151 *** (0.02853)	
Share with parents have med./long education		0.13709 *** (0.02057)	0.14098 *** (0.02059)	
Share with parents who received unemployment assistance/insurance, last year		-0.06251 * (0.02521)	-0.04194 (0.02519)	
Share with a parent who has been in contact with psychiatric hospital		0.01609 (0.03159)	-0.00382 (0.03182)	
Share with mother's income in 2nd quartile, last year		0.00918 (0.03108)	0.01914 (0.03140)	
Share with mother's income in 3rd quartile, last year		-0.05718 (0.03223)	-0.04666 (0.03244)	
Share with mother's income in 4th quartile, last year		-0.08829 ** (0.03372)	-0.07615 * (0.03388)	
Share with father's income in 2nd quartile, last year		0.02033 (0.03011)	0.02834 (0.03042)	
Share with father's income in 3rd quartile, last year		-0.00871 (0.03062)	0.00218 (0.03095)	
Share with father's income in 4th quartile, last year		0.07386 * (0.03110)	0.08515 ** (0.03136)	

Notes: Dependent variable is social well-being. Column 1 shows results including child, parental, and household characteristics. Column 2 adds teacher, school, and classroom characteristics. Column 3 adds classroom random effects, while column 4 adds classroom fixed effects. Indicators for missing values are included in all regressions. Standard errors clustered at the classroom level and shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. No. observations 380,680; no. classrooms 21,135. R^2 -adjusted of 0.039 (column 1), 0.044 (column 2), 0.045 (column 3), and 0.037 (column 4). Fraction of variance due to variation between classrooms (rho) 0.104 (column 3, random effects), 0.167 (column 4, fixed effects).

2.4.B The role of teacher and school characteristics

Our next set of explanatory variables concern teacher and school characteristics. Whether teachers and schools affect child human capital is a question that has attracted overwhelming interest from scholars. See e.g. Rivkin et al. (2005) and Chetty et al. (2011) for studies on teacher quality. Among teacher characteristics, gender has received particular attention (e.g. Nixon and Robinson, 1999; Bettinger and Long, 2005; Dee, 2007; Antecol et al., 2015).

Included variables describe Danish and math teachers in terms of their gender, age, training, experience, and tenure, as well as their absence from work. School characteristics include measures of teacher turnover and size. Results are shown in panel C of Table 2.4. For obvious reasons, we cannot study teacher characteristics in models that include classroom fixed effects. Interestingly, we find that social well-being is higher in classrooms with a female Danish teacher, just as higher tenure correlate positively with well-being. Teacher absence, in contrast, is negatively associated with well-being. Teacher turnover at the school level and school size are not strongly related to well-being. Presumably, and in line with existing literature, teachers are instrumental in establishing a well-functioning classroom culture, but many confounding factors could contribute to this too and causality may run in both directions. Teachers in problematic classrooms may experience burnout, for example, and substitute teacher may often be younger and less experienced. Accordingly, as above, we explicitly refrain from making causal interpretations and leave it to other researchers to study any effects of teacher profile on pupil well-being.

2.4.C The role of classroom composition

Our final set of variables describe classroom composition. Here, we include variables measured at the pupil and family level but now calculated as means among the other pupils in the classroom, while explicitly excluding the pupil him or herself. Previous studies suggest that classroom composition explain pupil achievement; see for example Carrell and Hoekstra (2010) on criminal background, Cho (2012) and Diette et al. (2014) on English language learners and immigrants, and Figlio (2007) and Kristoffersen et al. (2015) on disruptive students.

Our estimates indicate that a higher share of males and a higher share of classmates with parents with a criminal record correlate negatively with social well-being. The share of late starters and children who repeat a grade, on the other hand, correlate positively with social well-being. This is presumably because these factors imply that the children are more mature. School switchers, children placed outside of the home, and to a lesser extent the share of children with contacts to psychiatric hospitals are similarly associated with higher levels of well-being. The latter two cases could, however, be because of additional resources allocated to the classroom. The share of parents with higher levels of education correlate positively with higher well-being but the association with income levels is less clear and results vary across mothers and fathers; to the extent that there is assortative matching, these results are slightly more difficult to interpret.¹⁹

Overall, we find compelling evidence that the social well-being measure is associated with individual, family, and teacher variables indicative of disadvantage. Hence, given the results from our analysis, there is no reason to think that the social well-being scale is not an informative proxy for actual well-being. Of course, the explanatory power of our models is low and the R^2

¹⁹ Of course, these results could link to survey response tendencies as well, which we explore in Section 2.4.E.

lies in the neighborhood of 0.04 depending on the exact model; indicating that much of the variation in children's social well-being remains unexplained. This is not uncommon, even for well-validated measures; see also Gutman and Feinstein (2008) and Bradshaw et al. (2011). It is noteworthy, however, that our conditioning set is considerably richer than that of prior studies. Moreover, much of the explained variation seems to be driven by individual level factors: estimated intraclass correlations lie between 10 and 15% indicating that between 85 and 90% of the explained variation in social well-being is due to individual level factors, i.e. factors beyond the grouping structure (the classroom); see also Verkuyten and Thijs (2002) who find that similar results.

2.4.D Persistence over time

One concern is that well-being is particularly sensitive to day-to-day variation, potentially prohibiting a meaningful interpretation. We saw indications in Table 2.2, however, that well-being was fairly persistent over time. To address the issue further, we investigate the extent to which this conclusion holds once we control for background characteristics as well. Keep in mind that the outcome is standardized relative to the cohort, which means that we implicitly explore persistency in the relative position in the distribution, rather than the absolute value, over time. In practice, we extend the full model from Table 2.4 to include the first and second lag of social well-being and base our analysis on the longitudinal sample. Results are shown in Table 2.5. Of course, as in the descriptive analyses, this limits the sample to consist of pupils for whom we have at least three measures of well-being and de facto excludes the youngest pupils in the sample. Results are very clear: there is a high association between current and lagged measures

of social well-being, regardless of the empirical specification. The correlation is higher ($>.47$) between the current and the first lag of social well-being but is still substantial (about .15) two years back in time. Moreover, lagged social well-being is key in explaining current social well-being; adding control variables in column 2 only increases the explained variation slightly. We similarly investigate persistence in the likelihood of lying in the lowest quartile of social well-being. These results are presented in Table A2.3. All of these findings imply to us that the measure of social well-being carries relevant information.²⁰

Table 2.5
Persistence in social well-being measure

	Lagged well-being measures only		Add background variables	
Social well-being, 1 year lag	0.45852 (0.00239)	***	0.44306 (0.00240)	***
Social well-being, 2 year lag	0.17718 (0.00226)	***	0.16699 (0.00225)	***

Notes: Table shows regressions of social well-being on lags of social well-being. First column only includes lags of social well-being; second column also include child, parental, teacher and school, and classroom characteristics in addition to indicators for missing values as in Table 2.4, column 2. Robust standard errors shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. No. observations 247,987. R^2 -adjusted of .318 (column 1), 0.327 (column 2).

²⁰ Many of the same background variables that correlate with social well-being in levels also correlate with consistently being in the 1st quartile in terms of social well-being. Results are available upon request.

2.4.E Survey response

Although obligatory at Danish public schools, some children do not complete the national well-being survey, with absence from school being the most likely explanation. Out of the relevant 449,971 children, 380,680 pupils, or 85%, participated in the survey (see Table 2.3). If disadvantaged children with low well-being are less likely to respond to the survey, the real link between well-being and background characteristics could be even stronger than what we observe in the data. We saw above, for example, that special needs children in ordinary classrooms report lower levels of well-being. If it is the case that special needs children with low levels of well-being are more likely to be absent on the day of survey, we will tend to *underestimate* the relationship between social well-being and this particular variable. To investigate this, we regress an indicator for survey response on the same set of characteristics as above.

Table A2.4 shows clearly that individual and family level proxies for prior disadvantage correlate with survey response in the way we expect. This is comforting because it ensures that the signs of the estimated relationships from above are, in fact, reliable. Teacher characteristics are to some extent predictive: exposure to an older and more experienced teacher or a male teacher generally lowers response rates. Classroom composition, on the other hand, seems predictive of survey response but the direction is not always in line with what one would expect. For example, we detect higher response rates in classrooms where the share of children with special needs is high, but lower response rates if the share of high-income mothers is high. This calls for caution when interpreting the results.

2.5 Associations with other relevant outcomes

Table 2.6 finally explores the degree to which other relevant outcomes such as absence rates during the school year, standardized national test scores, and 9th grade GPA correlate with previous measures of social well-being. This is important, because the social well-being measures could then reasonably be used as early indicators for later academic outcomes that may have consequences for children's entire life path. It also speaks to the information content of the social well-being measure. The conditional correlations between absence rates and prior measures of social well-being are negative but small. Test scores and 9th grade GPA, on the other hand, correlate positively with social well-being, both one and two years prior to the measurement of the outcome. An increase in social well-being in 2018 of one standard deviation is associated with a 0.07 standard deviations higher national test score in reading and a 0.07 higher level of 9th grade GPA, for example. These positive correlations are in line with international studies; see B  cker et al. (2018) for an overview.

Table 2.6

Associations between social well-being, school absence, and academic outcomes

	Share Absent		Danish (reading), national test		Math, national test		9th grade GPA	
Social well-being, 1 year lag	-0.00145	***	0.06838	***	0.06176	***	0.07569	***
	(0.00010)		(0.00253)		(0.00300)		(0.01892)	
Social well-being, 2 year lag	-0.00057	***	0.04625	***	0.04636	***	0.12044	***
	(0.00011)		(0.00284)		(0.00297)		(0.01853)	
# Observations	373,203		176,668		129,732		20,538	

Notes: Table shows regressions of outcomes on lags of social well-being. Share absent measures the share of the school year a child was absent; national test scores are standardized to have mean 0 and variance of 1; 9th grade GPA measures average exam grades in obligatory courses. Regressions also include child, parental, teacher and school, and classroom characteristics in addition to indicators for missing values as in Table 2.4. Robust standard errors shown in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. R²-adjusted of 0.102 (column 1), 0.180 (column 2), 0.186 (column 3), and 0.278 (column 4).

2.6 Conclusion

This chapter studies the characteristics of the social well-being segment of the Danish national well-being survey, distributed to children of all ages in public school. We document that low social well-being correlates meaningfully with standard measures of disadvantage at the pupil and parental level, just as teacher characteristics and classroom composition are additional important predictors of well-being. We also show that social well-being exhibits high degrees of persistence over time, regardless of whether or not we control for a wide range of background characteristics. We finally show that higher social well-being is positively associated with academic performance and negatively but only slightly associated with absence from school. All of these findings are important for a range of decision-makers. Seen in combination with the findings of Andersen et al. (2020) and Niclasen et al. (2018), we also propose that the school social well-being indicator from the national well-being survey could reasonably be used as a first, convenient indicator of well-being among Danish pupils.

Chapter 3

Too young to quit school? Increasing the Compulsory Schooling Leaving age and educational outcomes

3.1 Introduction

Reducing early school leaving, defined as leaving education with at most lower-secondary schooling²¹, is generally considered to be one of the top priorities of educational policy. By the year 2000, the early school leaving rate in the European Union averaged about 20%. Because of the high social and economic costs of early school leaving, both to individuals and societies (Brunello and Paola, 2014), the European Union set as a Europe 2020 goal to bring this rate to under 10%, a target that has been met, with 2020 rates at 9.9%. While these numbers have been improving, the general consensus is that there is still substantial room for improvement.

A policy that has been applied with the objective of reducing early school leaving and dropout is to increase the compulsory schooling leaving age. By mandating that students stay longer in schools, we would expect that more of them eventually obtain an Upper-Secondary school diploma — specifically, the *marginal* students that change their educational path because of the policy. However, it is still unclear how effective raising the compulsory schooling leaving age is as a policy instrument. While it could increase graduation probabilities and reduce early school

²¹ “Early leaver from education and training, previously named early school leaver, refers to a person aged 18 to 24 who has completed at most lower secondary education and is not involved in further education or training; the indicator ‘*early leavers from education and training*’ is expressed as a percentage of the people aged 18 to 24 with such criteria out of the total population aged 18 to 24.” [Source: Eurostat](#)

leaving rates (Cabus and De Witte, 2011), it could also make students stay longer in school but have no effect on graduation probabilities (Adamecz-Völgyi, 2021; Landis and Reschly, 2011).

I analyze the Portuguese reform which provides a noteworthy opportunity to study the effects of increasing the Compulsory Schooling Leaving (CSL) age and have more recent evidence on the effectiveness of such a policy. Portugal had one of the EU's highest early school leaving rates in 2008 at 35%, increased its CSL age from 15 to 18 years old in 2009, and recently achieved an early school leaving rate of 8.9% in 2020, attaining the EU 2020 target of a rate under 10%. The law determined that all students enrolled in grades 1 to 7 from the academic year of 2009/10 onwards need to stay in school until they have finished the 12th grade of schooling or have turned 18 years old, while students enrolled in the 8th grade or above in 2009/10 are still under the old CSL and can leave school when they have finished the 9th grade or have turned 15 years old. All else constant, were students under the new CSL more likely to get an Upper-Secondary school diploma?

A straightforward, yet naïve, answer would simply compare graduation rates of the older cohorts (in the 7th grade in 2008/09) that were under the old CSL, with those of the younger cohorts (in the 7th grade 2009/10) which were under the higher CSL. However, this approach cannot account for unobservable cohort effects that may drive educational outcomes. Nonetheless, there is a group of students from the same grade-cohort for whom exposure to the new CSL was determined by their academic achievement in that year. Students enrolled in the 7th grade in 2008/09 that were retained had to repeat this grade in the subsequent academic year and were, therefore, exposed to the new CSL. On the other hand, students in the 7th grade in 2008/09 that were promoted to the 8th grade still fell under the old CSL.

As such, I compare 7th grade students from the same grade-cohort in 2008/09 for whom their academic performance determined grade retention and hence, exposure to the higher CSL age. Yet, this would assume that retained *vs* promoted students are not different, on average, in terms of unobservables that may also affect the outcomes in question. To deal with this issue, I apply a Difference-in-Differences methodology based on data from students from previous 7th grade cohorts for whom retention *vs* promotion did not determine exposure to different CSL ages, to isolate the effect of the reform from the effect of retention on outcomes. Furthermore, to add credibility to the parallel-trends assumption that is needed for Difference-in-Differences estimates, I compare students at the grade-retention margin only, identified by school subject scores and retention rules. This methodology rests on the assumption that, during the period of analysis, the difference on educational outcomes between retention *vs* promotion at the margin remained constant.

A threat to the aforementioned identifying assumption would be if schools and teachers purposely changed retention practices in the academic year before the policy was implemented in the 7th grade (2008/09), or if these students changed their effort levels in response to an anticipation of different exposures to the policy. Balance tests comparing students' characteristics at the retention margin in the 2008/09 academic year with students at this margin in prior and subsequent years show that these differences were in fact small. Moreover, retention rates in the 7th grade also remained stable around the period of analysis. It is important to mention that retention levels in Portugal are quite high and a common practice — the retention rate in the 7th grade between 2007 and 2011 was around 15% and about 30% students in this grade were at least one year above the reference age, a strong proxy for having had previous grade retentions.

Furthermore, having at least one retention is a key predictor of early school dropout — 7th grade students in the analysis period with at least one retention were around 30% less likely to graduate, which implies this group of students is of special interest for studying increases in the CSL.

I find that as a result of the higher CSL age, students stayed longer in school although it is less clear if graduation probabilities were affected. While the effects are estimated for a policy-relevant group — those at the grade retention margin, which are at higher risk of dropout — there is still substantial heterogeneity within this margin, with positive treatment effects concentrated on lower-achieving students. I find that the positive effect on the compliance outcome (defined as complying with the new CSL, i.e., staying in school until 18 years old, or graduating Upper-Secondary school) is driven by students above the 7th grade reference age — a strong proxy for having had previous retentions and thus, lower prior achievement — while students at reference age experience no effect. For students above the reference age, I estimate that this group was 3.2 percentage points (p.p.) more likely to graduate from Upper-Secondary school, although this effect is not statistically significant. Exploring further heterogeneity within this sub-group, I find positive effects for students that ever enrolled in Upper-Secondary School (6.1 p.p.), and Boys (5.7 p.p.), that are statistically significant at the 10% level. A law that was passed in August of 2012, right before the first treated cohort could have reached Upper-Secondary school, determining measures to prevent early school leaving, such as better guidance and support to at-risk students, could have been of particular relevance in improving the graduation probabilities of some sub-groups. This suggests that the effectiveness of CSL increases may depend on other policies taken to prevent early school leaving, and their success may be compromised otherwise.

I complement the main Difference-in-Differences results with those coming from Event- Study estimates, that compare the first treated cohorts with the last untreated ones. While this latter analysis is only able to provide truly credible causal estimates under stronger assumptions regarding unobservable cohort effects, it has the advantage of allowing us to understand who the likely *compliers* of the policy were — i.e., those changing educational choices because they were required to stay in school for longer, and would have dropped out earlier in the absence of the reform. This is not entirely possible with the main Difference- in-Differences identification strategy due to the fact that by construction, it captures effects for students at the grade retention margin only and thus, are lower achieving by definition. I look at dimensions which the literature identifies as being key predictors of school dropout: namely socio-economic status, prior academic achievement, and reference age in the 7th grade. In line with the literature, I find that the main *compliers* from the policy were students from both lower socio-economic status and with prior lower-achievement, whereas no effects are found for higher-achieving and higher socio-economic status students.

I contribute to the literature in two ways. First, I provide more up-to-date and comprehensive evidence to the literature that examines the effectiveness of policies that raise the CSL age in improving educational outcomes. This is of particular policy-relevance for governments considering raising their CSL. Such changes have not been frequent in Europe as of late and increases in the CSL age are usually lower than three years. Moreover, only a minority of European countries set the compulsory education requirement at the Upper-Secondary level and large changes in this age limit have not been thoroughly studied. In fact, Eurydice (2019) shows that in the 43 European education systems analyzed, only 8 had a duration of full-time

compulsory education of at least 12 years (Belgium, Luxembourg, Germany, Hungary, Netherlands, Northern Ireland, Portugal, and Turkey).

While the effects of Compulsory Schooling Laws on various social and economic outcomes have been widely documented in the literature — e.g., on wages²², crime, voting behavior, fertility decisions, or mortality²³ — there is a smaller body of literature studying the effects of CSL on school outcomes. Cabus and De Witte (2011) show that a one year increase in the CSL age in 2007, from 17 to 18, decreased dropout rates in the Netherlands by around 2.5 p.p. However, as noted by the authors, the observed effect comes entirely from the group non-labile to the policy reform — the control group more often left school in the immediate period before the policy reform, which given the economic revival at that time, may reflect anticipation of labor market opportunities. On the other hand, the treatment group's dropout rates did not significantly change. Landis and Reschly (2011) find that higher CSL ages in the US had an effect on the timing of dropping out, but not on high school completion rates, by comparing states that changed the legal requirement from 2002 to 2006 with states that did not. An interesting case to evaluate the effects of changes in the CSL age comes from Hungary, which raised its CSL age from 16 to 18 in 1996 but brought it back to 16 in 2012. Adamecz-Völgyi (2021) finds that while the first reform increased the length of schooling on average by half a year, it did not improve graduation probabilities nor did it have any effect on the probability of employment at age 20 and 25. Moreover, she finds that student composition in vocational schools changed substantially after the first treated cohort reached high-school. In particular, the share of disadvantaged

²² Oreopoulos (2007) & Devereux and Hart (2010) find that increasing the CSL age had positive wage returns while in other cases, no wage returns were found (Oosterbeek and Webbink, 2007; Pischke and Von Wachter, 2008).

²³ See Anderson (2014) and Bell et al. (2016) for studies on CSL and crime; voting behaviour (Milligan et al., 2004); fertility decisions (Black et al., 2011); and mortality (Lleras-Muney, 2005).

students in these schools increased dramatically and the reform might have even increased the probability of dropping out in vocational schools, an effect the author attributes to potential negative peer effects and a crowding out of resources in these schools. Due to these challenges and several other political considerations, the Hungarian CSL age was reduced back to 16. Adamecz-Völgyi et al. (2021) find that the reform increased the probabilities of dropping out by 4 p.p., and of being neither in education nor in employment (NEET) by the same effect size of 4 p.p., with higher effects for the most at-risk students. No effects are found for the probability of being employed. Finally, Erten and Keskin (2019) analyze the policy change in Turkey that increased compulsory schooling from 8 to 12 years in 2012 and focus on identifying the main compliers of the policy. They find that among boys, those who would have participated in paid employment prior to the change in law were now more likely to attend high school. On the other hand, female compliers would have likely been NEET, or they could have been employed in unpaid work.

The literature therefore presents mixed evidence on the effectiveness of compulsory schooling laws in decreasing dropout. It seems clear that CSL age increases keep students in school for longer but only a limited number of studies show that graduation probabilities were also increased. However, a common finding across this literature is that the effects of these policies are heterogeneous, with positive effects, both in educational and labor market outcomes, typically concentrated in more marginalized and at-risk groups (Harmon, 2017). In line with the literature, the findings of this chapter also suggest substantial heterogeneity on the effectiveness of the CSL reform, with lower-achieving students experiencing stronger positive effects. By using a detailed administrative dataset covering the Portuguese student population, I'm able to give a clearer picture of who the student sub-groups that can potentially benefit from the reform

are. Moreover, my results indicate that lower-achieving boys and lower-achieving students that enrolled in Upper-Secondary schooling could have increased their graduation probabilities, providing further evidence that in some cases, CSL age increases can help reducing early school leaving for some students.

My second contribution is methodological. In some cases within educational systems, exposure to a policy is dependent, directly or indirectly, on grade retention *vs* promotion in some years, but not in others. Examples could include a scholarship, or another type of benefit, that is given to promoted students but not to those that were grade-retained. If this policy differentiates retained from promoted students only from some year onward, an identification strategy similar to the one used in this chapter could be applied. In fact, in a Placebo test in Section 3.5.4, I give credibility to this identification strategy by showing that outcomes for 8th grade-cohorts evolved parallelly between retained and promoted students, as expected by the absence of any policy differentially treating the two groups. To the best of my knowledge, no other study has applied a similar identification strategy.

This chapter is organized as follows. Section 3.2 presents the Portuguese institutional setting. Next, in Section 3.3 the identification strategy is discussed, and Section 3.4 describes the data used. Finally, Section 3.5 presents results, while Section 3.6 concludes.

3.2 Institutional Setting

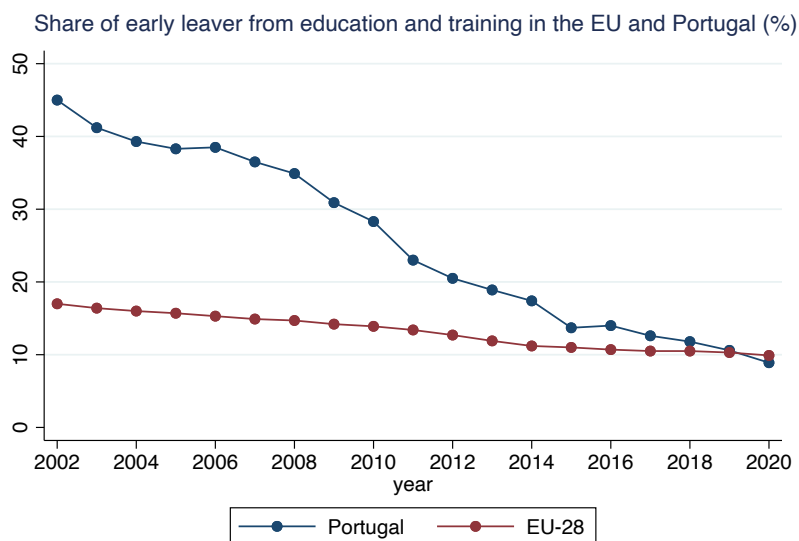
Portugal increased its Compulsory School Leaving (CSL) age from 15 to 18 years old in 2009. The fact that the country had one of the European Union's highest early school leaving rates in 2008 at 35%, significantly above the EU rate of 14.4%, was one of the main drivers of this policy change. 11 years later, in 2020, Portugal was able to achieve an early school leaving rate of 8.9%, and attained the EU 2020 target of a rate under 10% (Figure 3.1). While this was a very substantial improvement, it remains to understand what contribution the CSL age policy exactly had, as the country's rate was already on a clear downwards trajectory.

The government first announced in 2007 their plans to increase the CSL age from 15 to 18 years old two years later; in April of 2009 the parliament approved the law proposal; and in August of that year the law was finally approved²⁴. Support for the policy was generally widespread and political discussions focused mostly on the details of its implementation and resources needed. The law determined that students enrolled in grades 1 to 7 from the academic year of 2009/10 onward would be required to stay in school until they finished the 12th grade of schooling (the last grade of schooling in either the academic or vocational tracks) or turned 18 years old²⁵, while students enrolled in the 8th grade or above in 2009/10 would still be under the old CSL and could leave school when they finished the 9th grade or turned 15 years old.

²⁴ Law [“Lei n.º 85/2009”](#)

²⁵ Note that students in Portugal can graduate Upper-Secondary School at 17 years old if their birthday is after the end of the school year, in July, and they complied with the regular school starting age rule of starting school the year they turn 6 years old.

Figure 3.1



Share of early leavers from education and training in the EU-28 and Portugal. An “Early leaver from education and training, previously named early school leaver, refers to a person aged 18 to 24 who has completed at most lower secondary education and is not involved in further education or training; the indicator ‘*early leavers from education and training*’ is expressed as a percentage of the people aged 18 to 24 with such criteria out of the total population aged 18 to 24.”. Source: Eurostat

The academic year starts in September and finishes at the end of August of the following calendar year, with the summer holiday months going from July to September. The first cohort of students affected by the new CSL age reached 10th grade, the first year of Upper-Secondary Schooling, in the academic year of 2012/13, if they did not have any grade retentions since the implementation of the policy. In August of 2012, two laws were approved in light of the increase of the CSL age and the arrival of the first exposed cohort in Upper- Secondary school. The first²⁶ altered the Labour code to allow youngsters under 16 years old to work only if they had

²⁶ Law [“Decreto-Lei n.º 176/2012”](#)

completed compulsory schooling (i.e., for students still under the old CSL), or if they are enrolled in Upper-Secondary Schooling — serving therefore, as a way to enforce the new policy. The second²⁷ altered the Labour code to allow youngsters under 16 years old to work only if they had completed compulsory schooling (i.e., for students still under the old CSL), or if they are enrolled in Upper-Secondary Schooling — serving therefore, as a way to enforce the new policy.

In Portugal, children start school the year they turn 6 years old and private and public schools coexist in all levels of education²⁸. Elementary Schooling in Portugal goes from the 1st to the 6th grade and Lower-Secondary Schooling goes from the 7th to the 9th grade. During these first two educational stages there is no ability-tracking and the general track has the vast majority of students enrolled, around 90%. Upper-Secondary Schooling Portugal starts in the 10th grade. Students enrolling in this grade need to make the choice of whether to enroll in the Vocational track, from which students can choose one of several different fields directed towards earlier integration in the labor market, or the Academic track, targeting students who want to pursue a university degree, which is composed of 4 different specialization sub-tracks: Sciences & Technologies; Socio-Economic Sciences; Humanities & Languages; and Arts. Although students are not tracked by ability into their Upper-Secondary education paths, student and parental selection make it more likely for students coming from higher socio-economic status, or with higher prior achievement, to pursue the academic track, and within the academic track, to choose the Sciences and Technologies track.

²⁷ Law “Lei n.º 47/2012”

²⁸ The percentage of students in private schools is around 13% at all cycles of basic education and of around 20% in secondary education.

Students in the Portuguese school system are evaluated through teacher assessment and national exams. National exams in Portuguese Language and Mathematics are performed by every student in the system, by the end of the 4th and 6th grades (until 2015), and 9th grade. Children take these exams at exactly the same time, facing the same questions. Exams are then evaluated by a randomly allocated evaluator teacher, from schools other than the school in which the student is enrolled, in an anonymous fashion. In order to complete the general academic track of Upper-Secondary education, students must also sit through national exams — typically completing two track-specific exams in the 11th grade and another two in the 12th grade, in most cases, Portuguese language and Mathematics exams. Students can only gain admission to tertiary education if they have a passing grade in both 12th grade exams. Teacher-assigned school grades are based on several coursework elements, that include in- class tests but also homework, oral presentations, class participation and student behavior. In the 1st to 9th grade, grades are given on 1–5 scale, where 3–5 are passing grades, while in Upper-Secondary education the 1–20 scale is used, where the passing grades range from 10–20. Until the 9th grade, grades assigned by teachers at the end of each academic year, along with national exam scores in the grades they are taken, are used to determine whether a student is promoted to the next grade or retained, in which case the student has to repeat the same school-grade in the following academic year. Retention is decided by the student's teachers and a class committee but, as determined by national law, is only considered when a student has at least two or three subjects (depending on the school-grade and subjects) with failing scores. Retention levels in Portugal are high and a common practice (Eurydice, 2011; OECD, 2014). Although this practice has been steadily falling since around 2013, the retention rate in the 7th grade between 2007 and 2011 was around 15% and about 30%

of students in this grade were at least one year above the reference age, a strong proxy for having had previous grade retentions.

3.3 Identification Strategy

As described in the previous section, exposure to the new CSL was determined by the school-grade in which the student enrolled in the academic year of 2009/2010: students enrolled in grades 1 to 7 from the academic year of 2009/10 were required to stay in school until they finished the 12th grade of schooling or turned 18 years old, while students enrolled in the 8th grade or above in 2009/10 were still under the old CSL and could leave school when they completed the 9th grade or turned 15 years old.

A straightforward, yet naïve, approach to analyze the effect of the policy change would simply compare educational outcomes of older cohorts (in the 7th grade in 2008/09, or before) that were under the old CSL, with those of the younger cohorts under the new CSL (in the 7th grade in 2009/10, or later):

$$Y_{it} = \gamma \cdot CSL_t + \mathbf{X}'_{it}\beta + \varepsilon_{it} \quad (3.1)$$

where Y_{it} is an educational outcome (e.g., graduation) of student i , enrolled in the 7th grade in the academic year t ; CSL_t is a binary variable indicating exposure to the new CSL, i.e., it takes the value 1 if the student enrolled in the 7th grade in the academic year 2009/10 or later²⁹, and 0 otherwise; and $\mathbf{X}_{it}$ is a vector of student background characteristics, including gender, parental

²⁹ If a student was ever retained in the 7th grade and repeated it, I consider the last academic year in which the student enrolled in this grade.

education, baseline 6th grade national exam scores, as well as indicators for migrant status, socio-economic support, previous retentions, and resources at home. Despite its limitations in providing a credible causal estimate of the effect of the reform, the event-study approach of equation 3.1 is useful to understand the temporal evolution of the educational outcomes of the different cohorts, before and after the CSL change, and to understand which groups of students, identified by the background characteristics described above, changed their educational paths the most, or if at all, after the reform. This analysis is developed in Section 3.5.1 and focuses on understanding and identifying who the main compliers with the CSL reform were.

Nonetheless, there is a group of students from the same grade-cohort for whom exposure to the new CSL was determined by their academic achievement in that year. Students enrolled in the 7th grade in 2008/09 that were retained had to repeat this grade in the subsequent academic year and were therefore exposed to the new CSL, having to stay in school until they turned 18 years old or finished the 12th grade. On the other hand, students in the 7th grade in 2008/09 that were promoted to the 8th grade still fell under the old CSL and could leave school when they turned 15 years old or finished the 9th grade.

As such, I compare 7th grade students from the same grade-cohort in 2008/09 for whom their academic performance determined grade retention and hence, exposure to the higher CSL age. Yet, this assumes that retained *vs* promoted students are not different, on average, in terms of unobservables that may also affect the outcomes in question. To deal with this issue, I apply a Difference-in-Differences methodology based on data from students from previous 7th grade cohorts for whom retention *vs* promotion did not determine exposure to different CSL ages, to isolate the effect of the CSL from the effect of retention on outcomes. Furthermore, to add

credibility to the parallel-trends assumption that is needed for Difference-in-Differences estimates, I only compare students at the grade-retention margin, identified by school subject scores and retention rules. The Difference-in-Differences estimates are therefore estimated locally, as they compare students just above the retention threshold with those just below it. If the identifying assumptions discussed below hold, it is possible to identify the effect of the CSL reform for a specific group of students — those in the grade retention margin. However, being at this margin also serves as a proxy for being a lower-achieving student — a group that CSL policies specifically target, and for whom the literature identifies potentially larger effects.

Figure 3.2, panel A, illustrates the timing of exposure to the new CSL of different 7th grade cohorts and the identification strategy. For the *Pre-Intervention* cohorts, students always fell under the old CSL, irrespective of retention status. However, for the *Mid-Intervention* cohort, retained students became exposed to the new CSL due to enrollment in the 7th grade in 2009/10, while promoted students still fell under the old CSL as they will have enrolled in the 8th grade in 2009/10. Finally, in the *Post-Intervention* cohorts, all students were exposed to the new CSL, regardless of retention status. For all cohorts, I consider the first time a student was enrolled in the 7th grade³⁰.

The identification of the effect of the CSL policy for a group of students from the first treated cohort comes from comparing the evolution of outcomes of the retained group from 07/08 to 08/09, for whom their treatment status changed in the latter academic year, with that of the

³⁰ To guarantee that retained 7th grade students in the 2006/07 and 2007/08 cohorts do not end up exposed to the new CSL because of additional retentions, I focus only on students with one 7th grade retention at most, excluding around 1% of the sample.

promoted group, who remained under the old CSL in both 7th grade academic year cohorts. The evolution of outcomes of the promoted group in the academic years of 07/08 to 08/09 constitutes a valid counterfactual to that of the retained group, as long as the parallel- trends assumption holds. This assumption implies that the trend in outcome variables of the retained group from one cohort to the next would not be different from that experienced by the promoted group, had the retained group remained unexposed to the higher CSL in 08/09. This implies that, during the period of analysis, the difference on educational outcomes between retention *vs* promotion at the margin remained constant, allowing us to isolate the causal effect of the CSL from the effect of retention on outcomes.

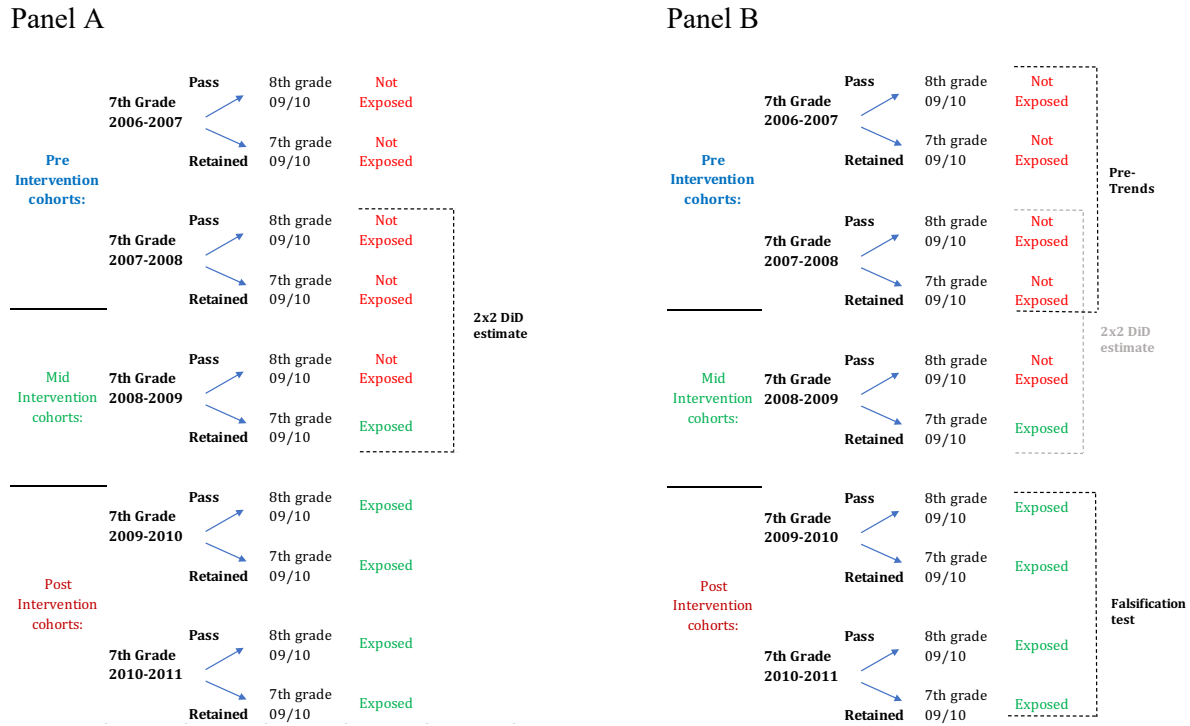
If, for instance, schools and teachers purposely changed retention practices in the academic year before the policy was implemented in the 7th grade, or if these students changed their effort levels in response to an anticipation of different exposures to the policy, this would imply changes on relevant unobservables, constituting a violation of the identifying assumptions. The comparison of retained and promoted students across years, in Section 3.4.3, as well as the Placebo tests discussed below and presented in Sections 3.5.2 and 3.5.4, provide good indication that this is not the case.

The identification of the CSL treatment effect is, in practice, implemented through the use of repeated cross-sections of 7th grade students at the retention margin, in a 2x2 Difference- in- Differences specification (i.e., the canonical DiD with 2 time periods and 2 groups), shown in equation 3.2:

$$Y_{it} = \alpha \cdot Retain_i + \lambda \cdot 2008/09_t + \gamma \cdot Retain_i * 2008/09_{it} + \mathbf{X}'_{it}\beta + \varepsilon_{it} \quad (3.2)$$

where Y_{it} is an educational outcome (e.g., graduation) of student i , enrolled in the 7th grade in the academic year t ; $Retain_i$ indicates retention in the 7th grade; $2008/09_t$ indicates that the observation comes from the 08/09 7th grade cohort instead of the 07/08 cohort; and X_{it} is a vector of student background characteristics listed above. The parameter of interest is therefore γ , the parameter on the interaction term of $Retain_i$ with $2008/09_t$, which under the identifying assumptions previously discussed, identifies the effect of the new CSL on these students' educational paths.

Figure 3.2



Timing of exposure to the new CSL and Difference-in-Differences Identification Strategy. Panel A illustrates where the identification of the new CSL effect comes from, while Panel B shows two possible tests of the validity of the identification strategy

The setting illustrated in Figure 3.2 indicates variation in treatment timing of the different groups, retained and promoted. While the retained group change their treatment status from 07/08 to 08/09 and the promoted group remain unexposed, the promoted group change their treatment status from 08/09 to 09/10 while the retained group remain exposed. In principle, we could then estimate a second set of treatment effects, comparing the evolution of outcomes of the promoted group from 08/09 to 09/10 with that of the retained group — i.e., comparing the *late* switchers with the *early* switchers. However, as discussed by the recent literature on Difference-in-Differences with variation in treatment timing (also known as staggered-adoption), comparisons of *late* with *early* switchers are biased in the presence of treatment effect heterogeneity across time (Goodman-Bacon, 2021). In this setting, this would mean that the treatment effect experienced by the retained group differs across 7th grade cohorts from 08/09 onward — this could happen if, for instance, Upper-Secondary schools became better prepared to receive later cohorts of retained students, positively affecting their outcomes. Under that scenario, comparing what happens when the promoted group change their treatment status with what happened with the retained group leads to bias. One solution proposed by Callaway and Sant’Anna (2021) is to consider a “*Never-Treated*” group that remains unexposed to the policy in the period of analysis and therefore, can be used to estimate the second set of treatment effects for when the promoted group becomes treated. However, because there is no such group in this setting, I focus only on the treatment effect arising from the change of the treatment status of the retained group from 07/08 to 08/09.

To test the validity of this identification strategy, I consider two other sets of effects illustrated in panel B of Figure 3.2, as well as a Placebo test. The first inspects if pre-treatment trends are parallel — that is, if from 06/07 to 07/08 the outcomes of the retained group and the promoted

group evolved in a parallel way, when both groups remained untreated. While verifying this does not guarantee that the parallel-trends assumption is satisfied in the periods for which the treatment effect is estimated (as this is related with a potential and unobservable outcome for the retained group in 08/09), it makes it more plausible. However, because data on school subject scores, used to define the analysis sample at the retention margin, are not available in the first year of the dataset, 2006/07, this is not directly possible. Nevertheless, it is possible to use Machine Learning classification methods to identify who the students that were likely to be at the retention margin in that academic year were, based on all observed covariates and subsequent academic path. Specifically, I train and test a model for 7th grade cohort between 2007/08 to 2010/11 (the last cohort used in the main analysis), for whom all variables are observed, and use it to make predictions for being at the retention margin in 2006/07, year in which all variables, except for school subject scores, are observed. Performance metrics such as Sensitivity, Specificity and Accuracy are around 80 to 85%, and Area Under the Curve (AUC-ROC), a commonly used measure to evaluate the performance of im- balanced classification tasks, is 91%. This indicates that the Machine Learning classification (using an *Extreme Gradient Boosting model*) has good performance. The procedure to get these predictions is further detailed in Appendix A3.2.

The second validity test lies in inspecting the evolution of outcomes for the retained and promoted groups after 2009/10, i.e., when both groups are exposed to the new CSL. This test is valid under the assumption that treatment effect dynamics of both groups are homogeneous from 2009/10 onward. Thus, while not verifying this homogeneity does not constitute a violation of the parallel-trends assumption, it still provides an important description of treatment effect dynamics in any case.

Finally, in a Placebo test in Section 3.5.4, I give credibility to this identification strategy by showing that outcomes for 8th grade-cohorts evolved parallelly between retained and promoted students in all years, as expected by the absence of any policy differentially treating the two groups.

I analyze two educational outcomes: compliance with the new CSL — i.e., staying in school until reaching at least 18 years old or graduating Upper-Secondary School; and Graduation from Upper-Secondary School. Results for each set of outcomes are presented in Section 3.5.2 both through estimates of equation 3.2, and graphical representations of the evolution of these outcomes for the retained and promoted groups.

3.4 Data

3.4.1 Data description

To analyze the question at hand, I use an administrative dataset³¹ containing detailed information on every student enrolled in public and private schools in mainland Portugal from the academic year 2006/07 to 2017/18. It contains relevant data on personal and socioeconomic characteristics of each student — such as gender, birthday, home neighborhood and school attended, country of origin, parents' education, parents' employment status, eligibility for social support, access to computer or Internet at home – with minimum measurement error or missing information. This

³¹ The administrative data is collected and maintained by the Directorate General of Education and Science Statistics (*Direcção-Geral de Estatísticas de Educação e Ciência*), a department under the administration of the Ministry of Education in Portugal

information is used to construct the control variables used in Section 3.5. The description of covariates used is included in Appendix A3.1 and besides these, measures of the mean level of these covariates at the school level are also included to control for school composition effects. A unique student identifier allows us to track students throughout grades and gather additional information about their educational pathway. We thus have a panel dataset of students since they are first observed in the Portuguese education system. A student's track is lost when he/she moves abroad, drops from the education system altogether, or dies. We may also lose track of students if these move to a different public or private school and the matching algorithm is unable to correctly assign the unique identifier to new instances of the same student in the system. The data is merged with two other datasets (ENEB and ENES) containing comprehensive information on student achievement in the standardized national exams of the 4th, 6th, 9th, 11th, and 12th grades.

3.4.2 Analysis samples

Our dataset also includes data on students' subject scores. As the identification strategy proposes to compare students near the 7th grade retention margin, determined by the number of failing scores, I use these data to identify these students in each cohort. Students in the 7th grade across all Portuguese schools take 10 subjects — Portuguese Language, Mathematics, Natural Sciences, Physics & Chemistry, History, Geography, English Language, a second Foreign Language (usually French), Visual Arts, and Physical Education — and grade retention is considered when a student has 3 or more failing scores (i.e., Scores of 1 or 2 in a scale of 1 to 5). Data on these subject scores is only available for Public Schools and from the academic year of 2007/08 on.

For this sample we have subject scores data for around 80% of 7th grade students. Because these data on school subjects are central to the Difference- in-Differences identification strategy, I restrict the main analysis to students who attended a Public School in the 7th grade³².

Furthermore, in all analysis presented in Section 3.5, only students aged at most 15 years old in their last 7th grade are considered. Under the old CSL, students reaching 15 years old could drop out of school at the end of the academic year. As such, to make sure that retention in the 7th grade leads to an exposure of the new CSL for the 2008/09 cohort, I have to make sure that students were at most 14 years old in their first enrollment in the 7th grade (and hence, at most 15 years old in their last 7th grade). The reference age at the end of the 7th grade — i.e., under a regular school starting age and with no accumulated retentions — is 13 years old if a student's birthday is between the 1st of January and the 31st of August, or 12 years old if the birthday is after that date. This means that the main sample of analysis includes students who are at most one year above the reference age in the 7th grade.

Finally, to make sure that more recent 7th grade cohorts are comparable with older cohorts, the educational outcomes analyzed are measured with the same time horizons. Specifically, because Upper-Secondary school graduation and compliance with the new CSL need time to be observed, the 7th grade cohorts from the academic years 2006/07 to 2010/11 are included, and measured 7 years later for all cohorts.

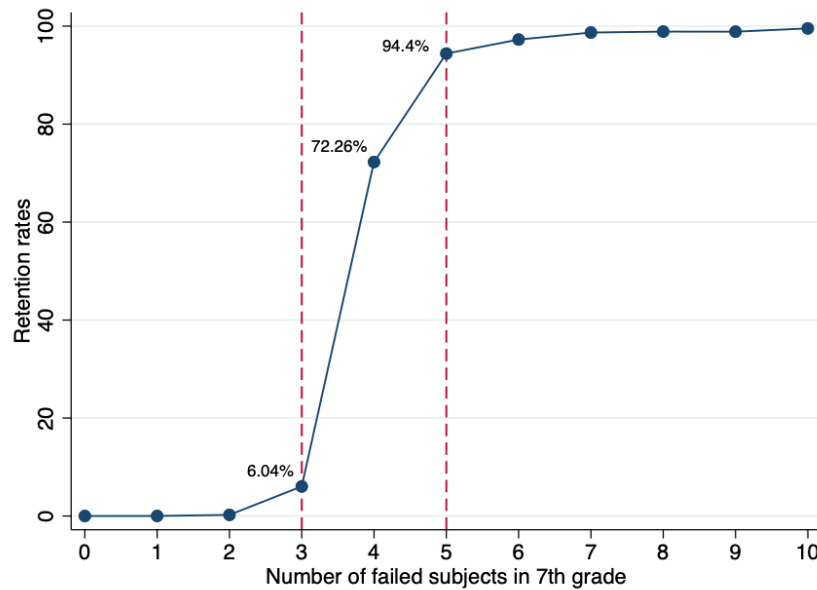
³² Students included in the main sample of analysis for the DiD estimates may have enrolled in a Private School before or after their first enrollment in 7th grade. The only condition of inclusion is therefore, that the enrollment in 7th grade was in Public School.

3.4.3 Retained and Promoted students

As discussed in Section 3.3, the local Difference-in-Differences methodology applied in this chapter compares retained with promoted students in the 7th grade at the retention margin, as retention in the 2008/09 academic year lead to an exposure of the new CSL, while promoted students remained under the old CSL. To capture the effect of the new CSL and separate it from the effect of retention vs promotion at the margin on the outcomes analyzed, I use data from other 7th grade cohorts for whom the retention or promotion status did not lead to different CSL exposure. This identification strategy therefore, rests on the assumption that, during the period of analysis, the effect of grade retention at the margin of promotion vs retention remained constant. I identify this margin by looking at retention rules and school subject scores. Retention is decided by the student's teachers and the class committee but, as determined by national law, is only considered when a student has 3 or more failing scores in the 10 subjects taken in the 7th grade. However, in practice, grade retention is much more commonly applied when a student has 4 failing scores — only 6% of students with 3 failing scores were retained, while 72.3% with 4 failing scores were retained, and 94.4% of students with 5 failing scores were retained, as seen in Figure 3.3.

Figure 3.3

Retention rates and number of failing scores



Retention rates and number of failing scores out of the 10 school subjects taken in the 7th grade. Raw rates are calculated based on all 7th grade cohorts used in the analysis, i.e., from 2006/07 to 2010/11.

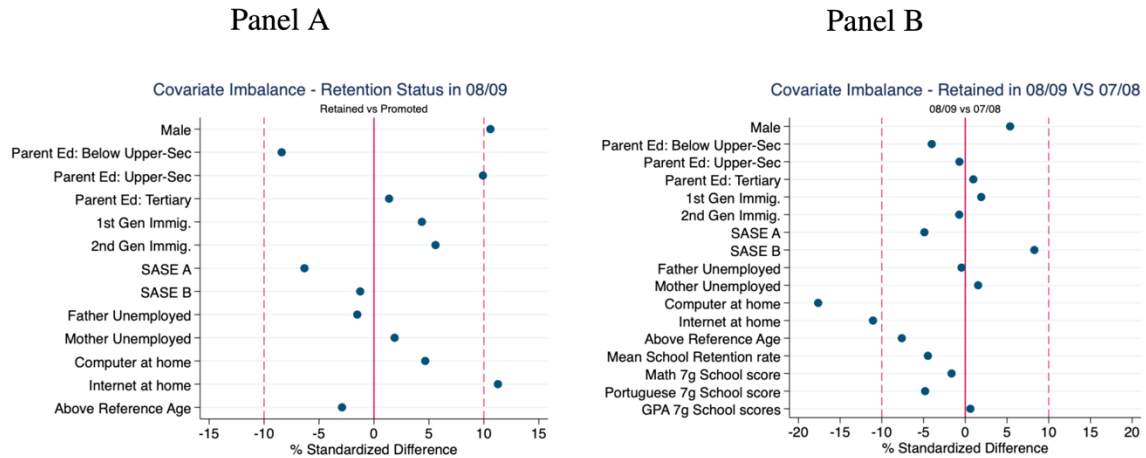
I focus on the group of students that had between 3 and 5 failing scores. First of all, we would like that the retained and promoted groups identified in this retention margin are as comparable as possible, and consequently, differences in covariates observed before the 7th grade should be small. Figure 3.4, panel A, plots standardized differences of pre-treatment covariates between retained and promoted students in the academic year of 2008/09. Standardized differences provide a scale and sample size free way of assessing covariate balance between groups (Imbens, 2015). The literature typically considers that having standardized differences below $|10\%|$ is a good balance. Panel A of Figure 3.4 shows that these differences are all under this threshold, except for the indicators of male gender and having access to an internet connection at home that are only slightly above the 10% threshold. However, regarding variables that directly determine

retention vs promotion, we observe large differences between the two groups in the academic year of 08/09. Table 3.1, panel A, shows that differences in the Portuguese Language and Mathematics subjects in the 7th grade, the GPA of all 10 subjects taken in the 7th grade, as well as mean retention rates at the school level, are very different between the two groups. Nonetheless, this would only be problematic if retained students in 08/09 were very different from those retained in 07/08. Panels B of Table 3.1 and Figure 3.4 show that this was not, in fact, the case. Differences in school subject scores, as well as on other variables, between retained students in 08/09 and retained students in 07/08 are small and below the |10%| threshold. The only exceptions are in the indicators for having computer and internet access at home. The same pattern is present for promoted students in 08/09 vs promoted in 07/08, as well as in differences at the retention margin across other subsequent 7th grade cohorts³³.

As we want to assume that the effect of retention at the margin remained constant across years, the fact that students at the retention margin are homogeneous in terms of observable covariates across cohorts is reassuring. A threat to the aforementioned identifying assumption would be if schools and teachers purposely changed retention practices in the academic year before the policy was implemented in the 7th grade, or if these students changed their effort levels in response to an anticipation of different exposures to the policy. Figure 3.5 shows that retention rates in the 7th grade for the cohorts of analysis, both for the overall population and for students at the retention margin, remained stable around 15% and 40%, respectively.

³³ Results available upon request.

Figure 3.4



Standardized differences in students' characteristics between retained and promoted students in the academic year of 2008/09 (Panel A), and retained students in 2008/09 compared with retained students in 2007/08. The literature typically identifies standardized differences below $|10\%|$ as a good balance.

Table 3.1

Panel A**Differences in Covariates - Retained VS Promoted in 2008/09**

	Retained	Promoted	% Std Diff
Male	0.60	0.55	10.6
Parent Ed: Below Upper-Sec	0.69	0.72	-8.4
Parent Ed: Upper-Sec	0.17	0.13	9.9
Parent Ed: Tertiary	0.04	0.03	1.4
1st Gen Immigrant	0.05	0.04	4.4
2nd Gen Immigrant	0.04	0.03	5.6
SASE A	0.29	0.32	-6.3
SASE B	0.12	0.13	-1.3
Father Unemployed	0.05	0.06	-1.5
Mother Unemployed	0.11	0.11	1.9
Computer at home	0.52	0.49	4.7
Internet at home	0.33	0.27	11.3
Above Reference Age	0.37	0.39	-2.9
Mean School Retention rate	0.20	0.17	36.4
Math 7g School Course	2.17	2.34	-37.4
Portuguese 7g School Course	2.34	2.64	-62.7
GPA 7g School scores	2.62	2.78	-150.7
Observations	2,684	4,253	6,937

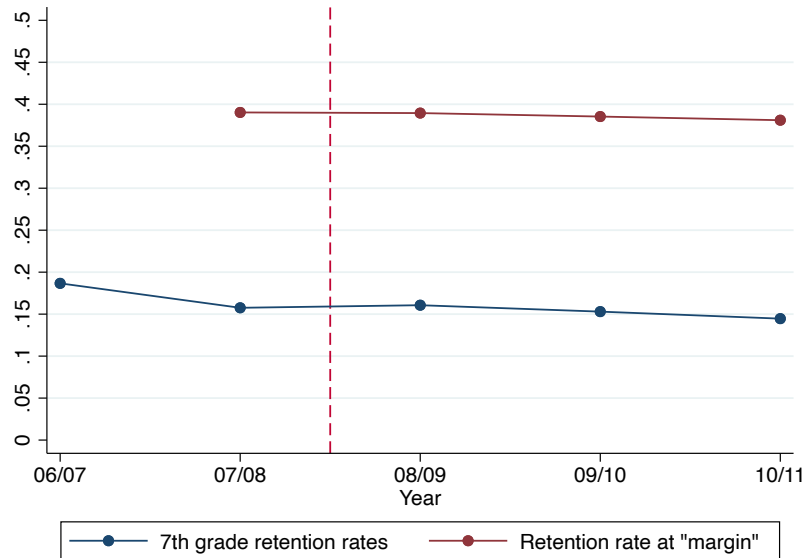
Panel B**Differences in Covariates - Retained in 08/09 VS in 07/08**

	08/09	07/08	% Std Diff
Male	0.60	0.58	5.4
Parent Ed: Below Upper-Sec	0.69	0.70	-4.0
Parent Ed: Upper-Sec	0.17	0.17	-0.7
Parent Ed: Tertiary	0.04	0.03	1.0
1st Gen Immigrant	0.05	0.05	1.9
2nd Gen Immigrant	0.04	0.04	-0.7
SASE A	0.29	0.31	-4.9
SASE B	0.12	0.10	8.3
Father Unemployed	0.05	0.05	-0.4
Mother Unemployed	0.11	0.11	1.5
Computer at home	0.52	0.60	-17.6
Internet at home	0.33	0.38	-11.1
Above Reference Age	0.37	0.41	-7.6
Mean School Retention rate	0.20	0.21	-4.5
Math 7g School Course	2.17	2.18	-1.6
Portuguese 7g School Course	2.34	2.36	-4.8
GPA 7g School scores	2.62	2.62	0.6
Observations	2,684	2,561	5,245

Means and Standardized differences in students' characteristics between retained and promoted students in the academic year of 2008/09 (Panel A), and retained students in 2008/09 compared with retained students in 2007/08.

Figure 3.5

Retention rates in 7th grade



Retention rates in the 7th grade student population in blue, and in the 7th grade retention margin, defined as having between 3 and 5 failing scores out of the 10 school subjects taken in that grade.

3.5 Results

3.5.1 Who are main compliers of the new CSL?

The literature on increases in the Compulsory Schooling Leaving age suggests that the effects from these policies tend to be concentrated on specific groups of the student population, and are far from being homogeneous. It is therefore most important to identify who the *compliers* from the new CSL policy are — i.e., the students who would have left school without a high-school diploma, or without reaching 18 years old, in the absence of the policy. I look at dimensions which the literature identifies as being key predictors of school dropout — namely socio-

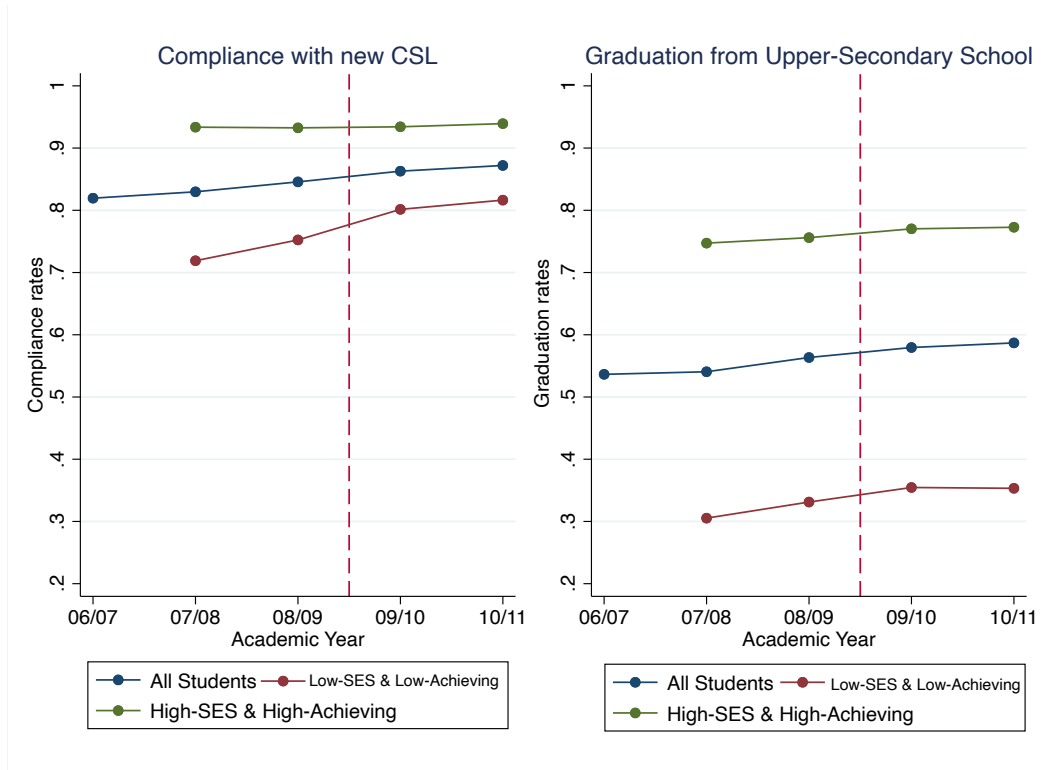
economic status, prior academic achievement, and reference age in the 7th grade (Cratty, 2012) — and find that dropout rates across time, and following the reform, changed much more significantly for disadvantaged students, while these same rates changed very little, or not at all, for students from more advantaged backgrounds.

Figure 3.6 shows compliance with the new CSL and Upper-Secondary school graduation rates for different sub-groups of the student population, across 7th grade cohorts. Compliance with the new CSL is defined as staying in school until 18 years old, or graduating Upper- Secondary School. While we observe a steady increase in these rates for the overall student population, represented in blue, we also notice great heterogeneity across sub-groups. Specifically, for students that come from lower socio-economic status (*Low-SES*) and also have lower levels of prior achievement (*Low-Achieving*), we observe a 5 percentage points (p.p.) increase in their compliance rates from the last unexposed cohort to the first cohort exposed to the higher CSL. On the other hand, for students that come from both higher socio-economic status (*High-SES*) and also have higher levels of prior achievement (*High-Achieving*), we observe almost no change in their compliance or graduation rates across 7th grade cohorts. A *Low-SES* student is defined as being a beneficiary of socioeconomic support (SASE), or with neither parent having completed Upper-Secondary schooling or higher; while a *Low-Achieving* student is defined as having failed at least one of the 6th grade exams of Portuguese Language and Mathematics^{34,35}.

³⁴ A High-SES student is thus defined as having at least one parent that has completed Upper-Secondary education or higher; while a High-Achieving student did not fail any of the two 6th grade national exams.

³⁵ Slightly changing the definitions of Low-SES or Low-Achieving does not alter results meaningfully. Moreover, presenting results for the Low-SES and Low-Achieving groups individually shows the same positive trend in outcomes but are stronger for the group that is simultaneously Low-SES & Low-Achieving

Figure 3.6
Compliance and Graduation rates for different student groups

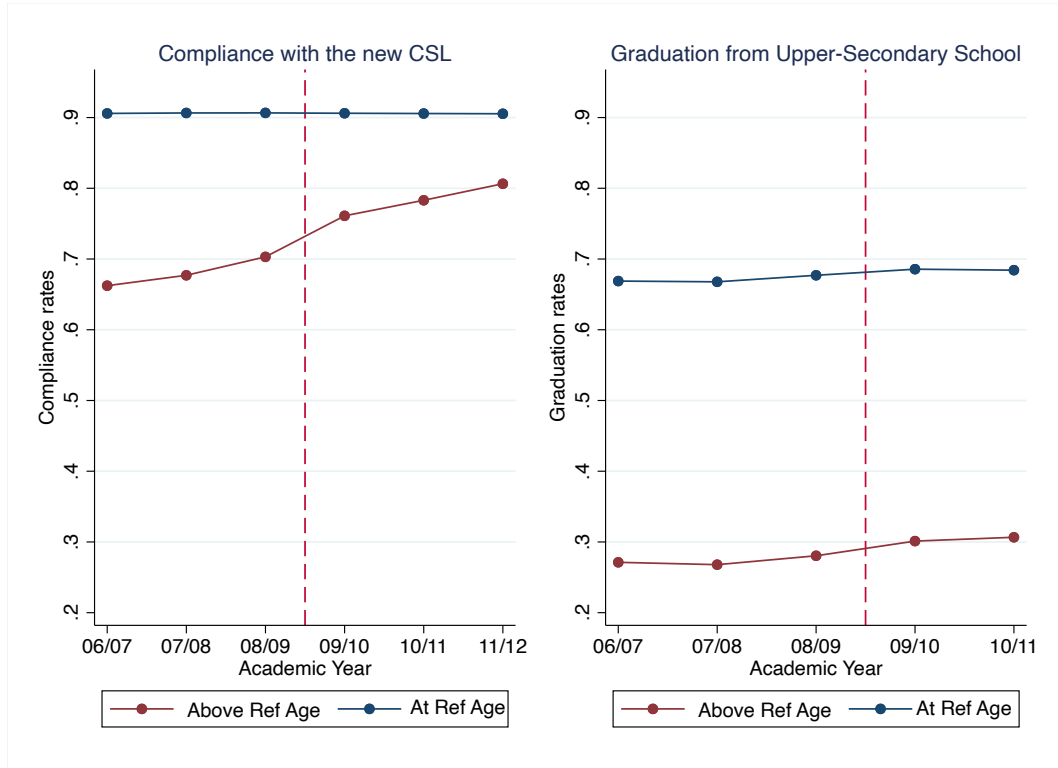


Compliance with the new CSL (defined as staying in school until 18 years old or graduating Upper-Secondary School) and Graduation from Upper-Secondary School rates in the population in blue, for *Low-SES & Low-Achieving* in red, and for *High-SES & High-Achieving* in green. A *Low-SES* student is defined as being a beneficiary of socioeconomic support (SASE), or with neither parent having completed Upper-Secondary schooling or higher; while a *Low-Achieving* student is defined as having failed at least one of the 6th grade exams of Portuguese Language and Mathematics.

In the same line, if we analyze differences in the trends of these outcomes for students that are on the 7th grade reference and above this reference age (a strong proxy for previous grade retentions and thus, low prior achievement), shown in Figure 3.7, we also notice that the positive evolution is only observed for the lower achieving group.

Figure 3.7

Compliance and Graduation rates for different student age groups



Compliance with the new CSL (defined as staying in school until 18 years old or graduating Upper-Secondary School) and Graduation from Upper-Secondary School rates for students at the 7th grade reference age in blue, and for students above the reference age in red. The reference age at the end of the 7th grade — i.e., under a regular school starting age and with no accumulated retentions — is 13 years old, if a student's birthday is between the 1st of January and the 31st of August, or 12 years old if the birthday is after that date.

By estimating Event-Study effects with the specification presented in equation 3.1, we can check how these outcome trends are influenced by changes in observable composition of 7th grade cohorts across time³⁶. Figures 3.8 and 3.9 show Event-Study plots for different SES & Achievement groups, for the compliance and graduation outcomes, respectively. Figures 3.10 and 3.11 show them for the different Reference age groups. These results are estimated by

³⁶ While equation 1 presents a specification with a simple dummy variable for post-reform periods, I choose to show results in event-study plots, with dummy variables for each 7th grade cohort, to get a better representation of treatment effect dynamics.

comparing different 7th grade cohorts to the comparison group of the 2008/09 7th grade cohort, the last cohort not affected by the new CSL. I take into account all covariates described in Appendix A3.1, as well as municipality fixed effects, and standard errors are clustered at the municipality level.

Event-Study results suggest that compliance rates increased by around 4 to 5 percentage points (p.p.) for the *Low-SES & Low-Achieving*, as well as the Above Reference Age group, when comparing the first cohort affected by the new CSL (09/10) with the last cohort under the old CSL (08/09). Regarding graduation probabilities, these effects are about half in terms of magnitude, i.e., around 2 p.p. for both sub-groups. Meanwhile, for more advantaged students (*High-SES & High-Achieving*, and *At Reference Age*) I do not find meaningful differences across cohorts — these are either non-statistically different or at most, different in terms of 1 percentage point. Thus, controlling for background covariates does not alter the substantial heterogeneity in the evolution of compliance and graduation rates shown in Figures 3.6 and 3.7.

While, as argued in Section 3.3, these Event-Study estimates should not be interpreted as causal estimates of the effect of the CSL reform, their main purpose lies in identifying who the main *compliers* of the policy may be. Moreover, because the DiD estimates of section 3.5.2 that aim at recovering causality are restricted to low-achieving students as a consequence of the identification strategy, it is relevant that Event-Study estimates identify this group as a main *complier* from the policy.

Event-Study estimates for different student groups

Figure 3.8

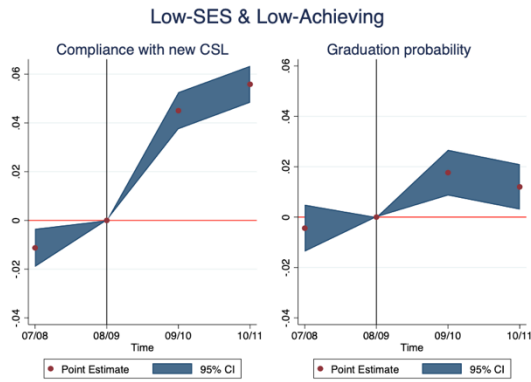


Figure 3.9

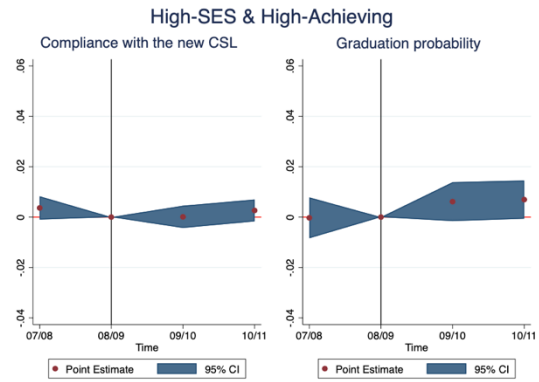


Figure 3.10

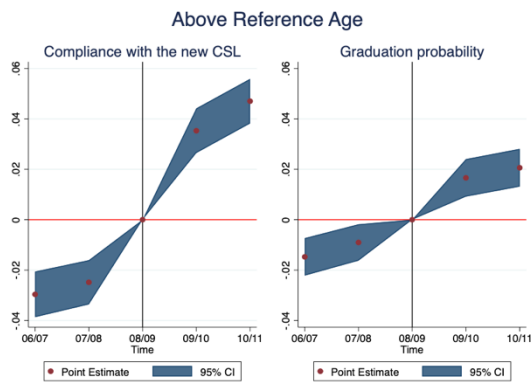
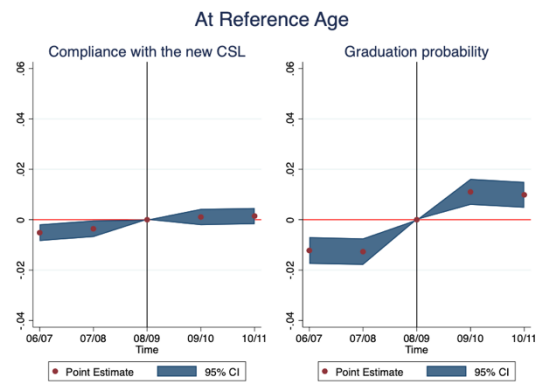


Figure 3.11



Event-Study estimates for the outcomes of compliance with the new CSL (defined as staying in school until 18 years old or graduating Upper-Secondary School) and Graduation from Upper-Secondary School, for different sub-groups of the students population. Results are estimated by comparing different 7th grade cohorts to the comparison group of the 2008/09 7th grade cohort, the last cohort not affected by the new CSL. All covariates described in Appendix A, as well as municipality fixed effects, are accounted for. Standard errors are clustered at the municipality level.

3.5.2 Main Difference-in-Differences Results

This section presents the main results of the Difference-in-Differences (DiD) identification strategy discussed in Section 3.3. Results are presented for the two following outcomes: Compliance with the new CSL (i.e., staying in the school system until 18 years old or graduating from Upper-Secondary school); and Graduation from Upper-Secondary School, in either the Academic or Vocational tracks. Results are presented both through estimates of equation 3.2 and graphical representations of the evolution of these outcomes for the retained and promoted groups across 7th grade cohorts. The graphical representations are presented both with raw means and treatment effect estimates, that take into account the covariates described in Appendix A.3.1.

Figure 3.12 shows the evolution of compliance for retained and promoted groups across 7th grade cohorts in the left-side panel. As seen in the right-side panel, the hypothesis that this outcome evolved parallelly between retained and promoted students before the policy change is not rejected, giving credibility to the identification strategy. We observe a positive treatment effect when the retained group changes their treatment status from 07/08 to 08/09 of 2.5 percentage points (p.p.), when no covariates are considered (column 1 of Table 3.2), and of 2.8 p.p. when covariates are considered (column 2 of Table 3.2). This indicates that students were more likely to stay in school for longer as a result of the higher CSL, an effect which is statistically significant at the 5% level.

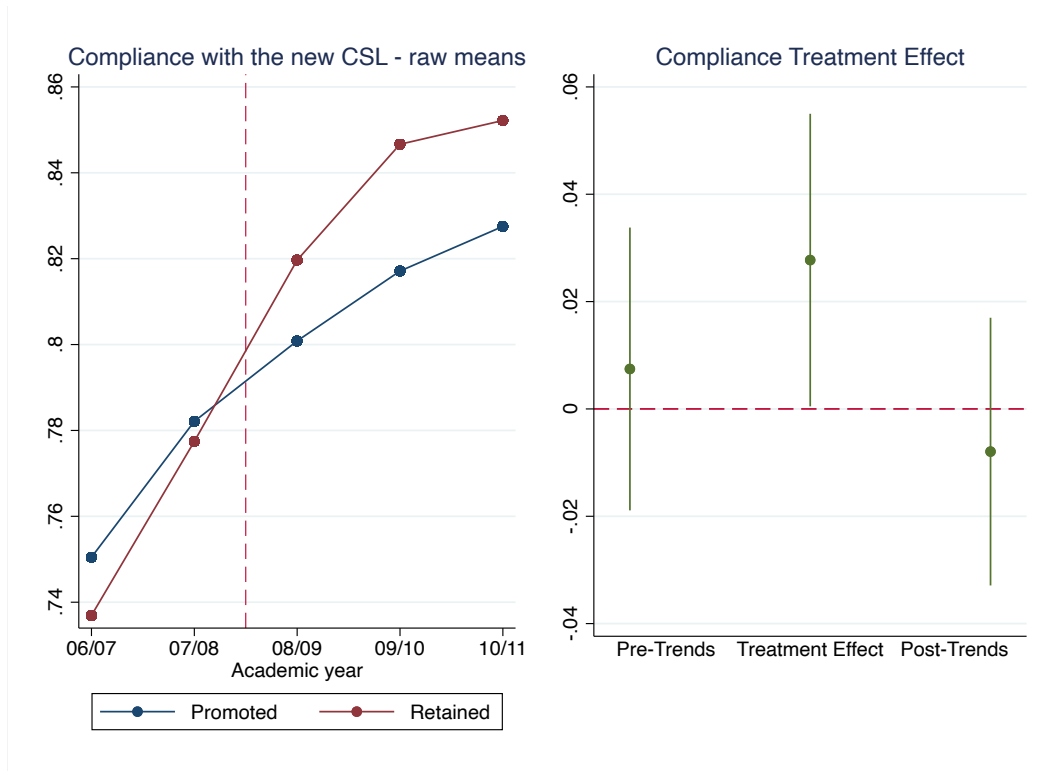
Regarding graduation probabilities, seen in Figure 3.13, the evolution of this outcome across 7th grade-cohorts also appears parallel between retained and promoted groups but in this case, no statistically significant treatment effect is found. Columns 3 and 4 of Table 3.2 show that the estimated treatment effect is very close to zero, irrespective of the inclusion of covariates.

It is worth noting that both the compliance and graduation probabilities of the retained group increase by at least as much from 08/09 to 09/10 (the second treated cohort) as they did from 07/08 to 08/09 (the first treated cohort). This could indicate positive treatment effects for the second treated cohort of retained students, perhaps because Upper-Secondary schools became better prepared to receive later cohorts of students as time went by. However, it is not possible to identify this time-group treatment effect due to the inexistence of a *Never-Treated* group in 7th grade cohorts.

Furthermore, it is possible to see in Figures 3.12 and 3.13 that from the 09/10 to the 10/11 cohorts, compliance and graduation outcomes evolved parallelly for the retained and promoted groups, suggesting homogeneity of treatment effect dynamics from 09/10 onwards. Difference-in-Differences estimates considering the 09/10 and 10/11 cohorts — i.e., when retained and promoted groups are both treated — show no statistically significant differences for the interaction term, thus suggesting that Post-Trends estimates are a valid falsification test.

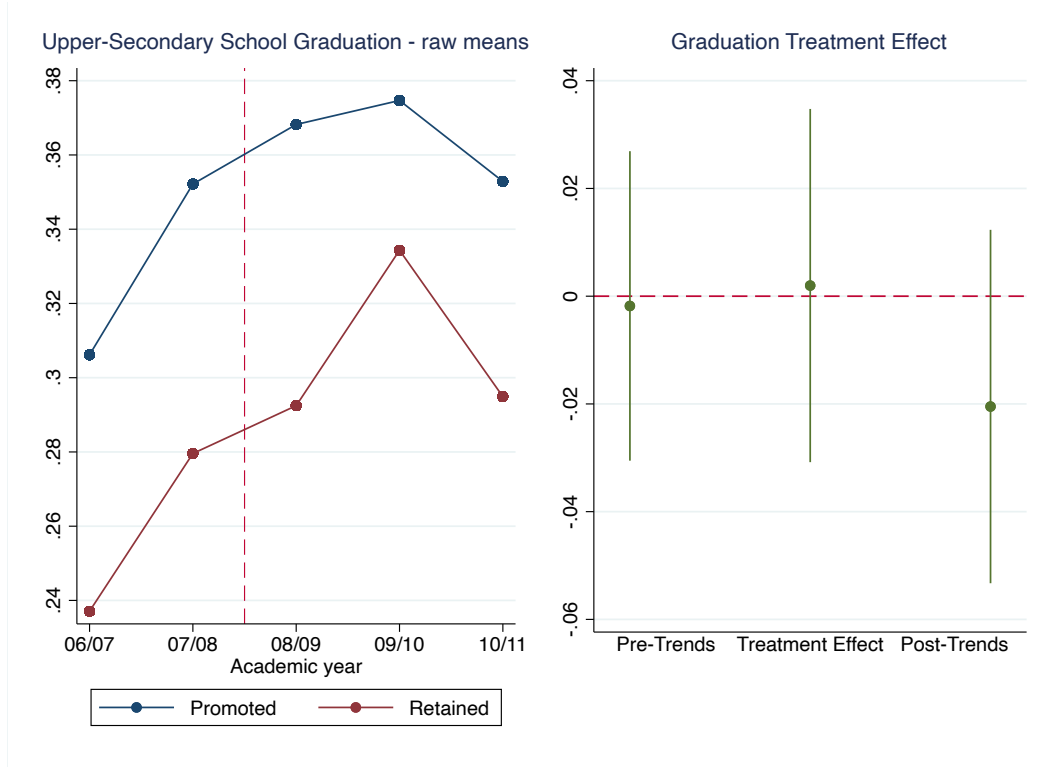
Figure 3.12

Compliance treatment effects



The left-side plot shows Compliance raw means for promoted students in blue and retained students in red. Compliance with the new CSL is defined as staying in school until 18 years old or graduating Upper-Secondary School. The right-hand plot shows Treatment effect estimates in 95% confidence interval plots for the Compliance outcome. The “Pre-Trends”, “Treatment Effect”, and “Post-Trends” represent the interaction coefficient of Equation 3.2, the 2x2 DiD estimate that compares retained with promoted students in the academic years 07/08 vs 06/07, 08/09 vs 07/08, and 10/11 vs 09/10, respectively. The sample of analysis considers 7th grade cohorts in Public Schools, in their first 7th grade, and in the grade-retention margin. Treatment effect estimates take into account all covariates described in Appendix A3.1, as well as means of these covariates at the school level. Standard Errors are clustered at the school level. Data on school subject scores, used to define the analysis sample at the retention margin, are not available in the first year of the dataset, 2006/07. The missing outcomes are imputed according to the procedure detailed in Appendix A3.2.

Figure 3.13
Graduation treatment effects



The left-side plot shows Graduation raw means for promoted students in blue and retained students in red. Graduation is defined as completing the 12th grade, the last grade of either the Academic or Vocational tracks. The right-side plot shows Treatment effect estimates in 95% confidence interval plots for the Graduation outcome. The “Pre-Trends”, “Treatment Effect”, and “Post-Trends” represent the interaction coefficient of Equation 3.2, the 2x2 DiD estimate that compares retained with promoted students in the academic years 07/08 vs 06/07, 08/09 vs 07/08, and 10/11 vs 09/10, respectively. The sample of analysis considers 7th grade cohorts in Public Schools, in their first 7th grade, and in the grade-retention margin. Treatment effect estimates take into account all covariates described in Appendix A3.1, as well as means of these covariates at the school level. Standard Errors are clustered at the school level. Data on school subject scores, used to define the analysis sample at the retention margin, are not available in the first year of the dataset, 2006/07. The missing outcomes are imputed according to the procedure detailed in Appendix A3.2.

Table 3.2
Main Difference-in-Differences estimates for Compliance and Graduation outcomes

	(1)	(2)	(3)	(4)
	Compliance	Compliance	Graduation	Graduation
Post_period x Retained	0.024* (0.014)	0.028** (0.014)	-0.003 (0.017)	0.002 (0.017)
Post_period	0.019** (0.009)	0.006 (0.015)	0.016 (0.012)	-0.028 (0.018)
Retained	-0.005 (0.012)	0.010 (0.015)	-0.073*** (0.013)	-0.011 (0.018)
Constant	0.782*** (0.008)	0.583*** (0.186)	0.352*** (0.009)	0.009 (0.201)
Controls		X		X
N	13,445	13,445	13,445	13,445
R-squared	0.001	0.059	0.006	0.054

Each column shows estimates of the Difference-in-Differences specification of equation 3.2. Columns 1 and 2 show results for the Compliance with the new Compulsory Schooling Law outcome, defined as staying in school until 18 years old or graduating Upper-Secondary School. Columns 3 and 4 show results for the Graduation outcome, defined as completing the 12th grade, the last grade of either the Academic or Vocational tracks. Columns 2 and 4 take into the covariates described in Appendix A3.1, as well as means of these covariates at the school level. Standard Errors are clustered at the school level. Significance levels: ***p < 0.01, **p < 0.05, *p < 0.1.

3.5.3 Heterogeneity analysis

The main results of the previous sub-section suggest positive effects of around 2.8 p.p. for compliance with the new CSL but no effects for graduation probabilities. Thus, taken together, they suggest that the higher CSL induced the first treated cohort of students at the retention margin to stay longer in school, but did not alter significantly their probability to graduate from Upper-Secondary school. Still, even within the retention margin there exists substantial heterogeneity across student groups that could lead to different treatment effects for some. I

analyze treatment effect heterogeneity in the lines of Socio-Economic status, Gender, and Reference age in the 7th grade and find meaningful differences in the latter dimension.

As shown in sub-section 3.5.1, students above the reference age (a strong proxy for having had previous grade retentions and thus, having lower levels of prior achievement) show a positive evolution in the compliance and graduation outcomes over time, while students at the 7th grade reference age show very steady rates across cohorts. Table 3.3, Panel A, shows positive and statistically significant treatment effects for the compliance outcome of around 5 p.p. for the group of students above the reference age, while for the graduation outcome a 3.2 p.p. effect is estimated, but it is not statistically different from zero. For the group of students at reference age, no statistically significant effects are found for either outcome, as seen in Panel B of Table 3.3. These results could indicate that, even within the retention margin — a group of students that is already low-achieving by definition — students with even lower levels of prior achievement could benefit more from CSL policies.

As positive treatment effects are driven by students above the reference age, I further explore whether some sub-groups of these lower-achieving students could have benefited in terms of Graduation probabilities. Once again, I explore treatment effect heterogeneity in the lines of Socio-Economic status, Gender, and additionally, whether students ever enrolled in Upper-Secondary School. This last condition is of particular relevance for several reasons. First, students that enrolled in Upper-Secondary at some point (70.9% of this sample, as seen in column 2 of Table 3.4) might have been more motivated to complete this stage of education, knowing that they were required by law to stay in school until 18 years old. Moreover, Upper-Secondary schools may have also been better prepared to receive and engage students from 2012/13

onwards, the first academic year when students under the new CSL could have started this stage of education. As described in Section 3.2, in August of 2012, a law was approved determining measures to prevent early school leaving and failure, such as better guidance and support to at-risk students, giving further emphasis to the enrollment in vocational and alternative tracks as a preventive measure.

Table 3.4 shows positive and larger Graduation treatment effects for lower-achieving students (i.e., above the 7th grade reference age) that enrolled in Upper-Secondary School (6.1 p.p., p-value: 0.055), and boys (5.7 p.p., p-value: 0.058). When combining these two conditions in column 5, I estimate that lower-achieving boys that enrolled in Upper-Secondary school increased their graduation probabilities by 9.1 p.p (p-value: 0.025). These results could give support to the hypothesis that Upper-Secondary schools, and policies targeting at-risk students, may have meaningful impacts in reducing early school leaving.

The fact that the estimated treatment effect for girls is not statistically significant could indicate, on the other hand, that these policies were not successful in engaging girls, or that there might have been other factors pushing their earlier dropout. Differences across Socio- Economic status (SES) are not meaningful, as the *High-SES* group constitutes only 10% of the sample above reference age.

Table 3.3
Treatment effect heterogeneity

Panel A - Students above reference age

	(1)	(2)	(3)	(4)
	Compliance	Compliance	Graduation	Graduation
Post_period x Retained	0.054** (0.027)	0.051* (0.026)	0.030 (0.024)	0.032 (0.024)
Post_period	0.023 (0.015)	0.023 (0.024)	0.004 (0.015)	-0.029 (0.022)
Retained	-0.033* (0.019)	-0.020 (0.023)	-0.069*** (0.017)	-0.019 (0.023)
Constant	0.726*** (0.013)	0.283 (0.316)	0.284*** (0.012)	0.066 (0.268)
Controls		X		X
N	5,369	5,369	5,369	5,369
R-squared	0.003	0.046	0.004	0.036

Panel B - Students at reference age

	(1)	(2)	(3)	(4)
	Compliance	Compliance	Graduation	Graduation
Post_period x Retained	0.002 (0.017)	0.010 (0.017)	-0.024 (0.022)	-0.018 (0.022)
Post_period	0.011 (0.011)	-0.004 (0.015)	0.017 (0.014)	-0.027 (0.020)
Retained	0.013 (0.012)	0.033** (0.017)	-0.077*** (0.016)	-0.001 (0.022)
Constant	0.823*** (0.008)	0.721*** (0.188)	0.402*** (0.010)	-0.078 (0.246)
Controls		X		X
N	8,076	8,076	8,076	8,076
R-squared	0.001	0.047	0.008	0.039

Each column shows estimates of the Difference-in-Differences specification of equation 3.2 but considering only students above the 7th grade reference age, in Panel A, and students at the 7th grade reference age, in Panel B. The reference age at the end of the 7th grade — i.e., under a regular school starting age and with no accumulated retentions — is 13 years old, if a student's birthday is between the 1st of January and the 31st of August, or 12 years old if the birthday is after that date. Columns 1 and 2 show results for the Compliance with the new Compulsory Schooling Law outcome, defined as staying in school until 18 years old or graduating Upper-Secondary School. Columns 3 and 4 show results for the Graduation outcome, defined as completing the 12th grade, the last grade of either the Academic or Vocational tracks. Columns 2 and 4 take into the covariates described in Appendix A3.1, as well as means of these covariates at the school level. Standard Errors are clustered at the school level. Significance levels: ***p < 0.01, **p < 0.05, *p < 0.1.

Table 3.4

Heterogeneous Graduation Effects within Students above reference age						
	(1)	(2)	(3)	(4)	(5)	(6)
	All Above Ref Age	Enroll Upper-Sec School	Boys	Girls	Enroll Upp-Sec Boys	Enroll Upp-Sec Girls
Post_period x Retained	0.032 (0.024)	0.061* (0.032)	0.057* (0.030)	-0.002 (0.041)	0.091** (0.040)	0.024 (0.052)
Controls	X	X	X	X	X	X
N	5,369	3,807	3,191	2,178	2,244	1,563
Percent of Sample	100%	70.9%	59.4%	40.6%	41.8%	29.1%
R-squared	0.044	0.048	0.036	0.057	0.046	0.054

Each column shows estimates of the Difference-in-Differences specification and the interaction term of interest of equation 3.2 but considering only students above the 7th grade reference age — i.e., that were 14 years old at the end of the year. Column 1 shows the estimate for all students above the reference age, while columns 2, 3, and 4, show estimates for these lower-achieving students that ever enrolled in Upper-Secondary School, for boys, and for girls, respectively. Column 5 shows estimates for lower-achieving boys that enrolled in Upper-Secondary School, while column 6 shows these for girls. All estimates take into the covariates described in Appendix A3.1, as well as means of these covariates at the school level. Standard Errors are clustered at the school level. Significance levels: ***p < 0.01, **p < 0.05, *p < 0.1.

3.5.4 Placebo test

The identification strategy applied in this chapter assumes that if there was no policy differentially treating retained and promoted students in the 7th grade in 2008/09, outcomes for the two groups would have evolved parallelly. Therefore, a check that gives credibility to the assumptions of this identification strategy is to compare what happens to the evolution of outcomes for 8th grade students at the margin of retention, where retention rules are the same as those applied in the 7th grade. As defined by the CSL reform, all students enrolled in the 8th grade or above in 2009/10 would be under the old CSL.

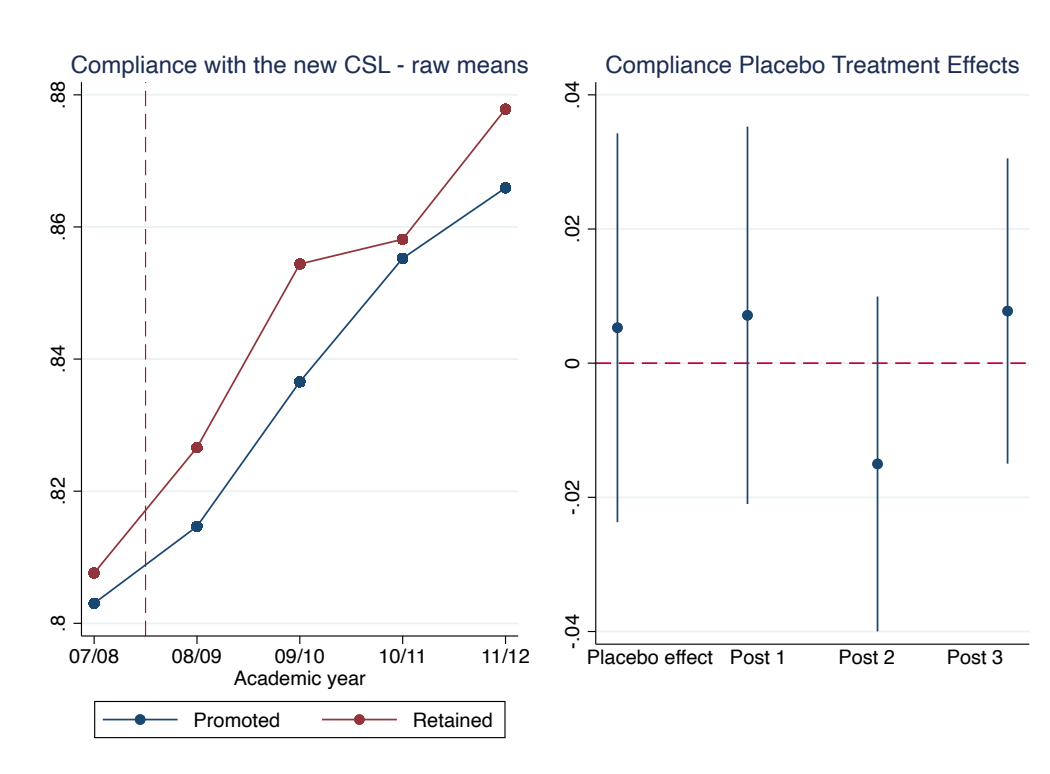
In a similar fashion to the main results presented in Figures 3.12 and 3.13, Figure 3.14 and 3.15 respectively show compliance and graduation outcomes for 8th grade students at the retention margin. We verify that both outcomes evolved parallelly between retained and promoted students, as expected by the absence of any policy differentially treating the two groups in the 8th grade. The Placebo treatment effect is not statistically different from zero for both outcomes and moreover, all Post-Trends placebo treatment effects are statistically insignificant.

Besides corroborating the plausibility of the identifying assumptions, these results also provide further external validity to this methodology, in general. In some cases within educational systems, exposure to a policy is dependent on grade retention vs promotion in some years, but not in others. Examples could include a scholarship, or another type of benefit, that is given to promoted students but not to those that were grade-retained. Another example could be that participation in a high-stakes exam at the end of the academic year is conditional

on grade-promotion. If such policy differentiates retained from promoted students only from some year onward, an identification strategy similar to the one used in this chapter could be applied.

Figure 3.14

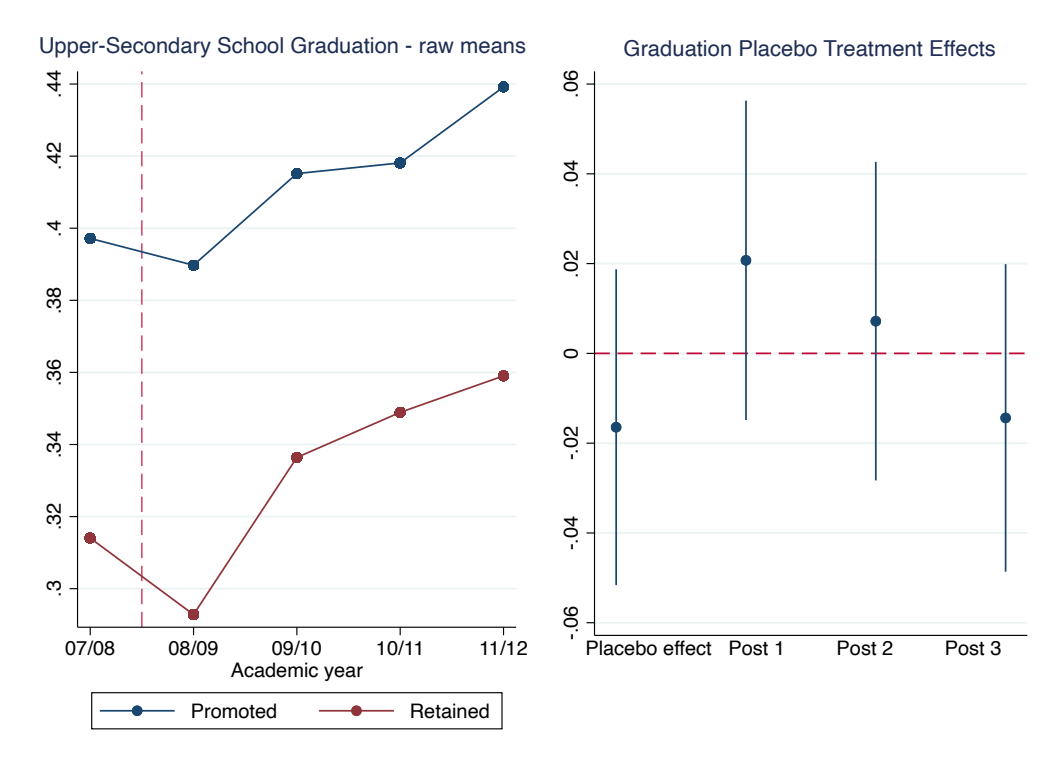
Compliance placebo treatment effects



The left-side plot shows Compliance raw means for promoted 8th grade students in blue and retained students in red. Compliance with the new CSL is defined as staying in school until 18 years old or graduating Upper- Secondary School. The right-hand plot shows Placebo Treatment effect estimates in 95% confidence interval plots for the Compliance outcome. The “Placebo effect” and all Placebo Post trends (“Post 1” to “Post 3”) represent the interaction coefficient of Equation 3.2, the 2x2 DiD estimate that compares retained with promoted students in the academic years 08/09 vs 07/08, and 10/11 vs 09/10, through 11/12 vs 10/11, respectively. The sample of analysis considers 8th grade cohorts in Public Schools, in their first 8th grade, and in the grade- retention margin, defined in an equal manner to the 7th grade. Treatment effect estimates take into account all covariates described in Appendix A3.1.

Figure 3.15

Graduation placebo treatment effects



The left-side plot shows Graduation raw means for promoted 8th grade students in blue and retained students in red. Graduation is defined as completing the 12th grade, the last grade of either the Academic or Vocational tracks. The right-hand plot shows Placebo Treatment effect estimates in 95% confidence interval plots for the Graduation outcome. The “Placebo effect” and all Placebo Post trends (“Post 1” to “Post 3”) represent the interaction coefficient of Equation 3.2, the 2x2 DiD estimate that compares retained with promoted students in the academic years 08/09 vs 07/08, and 10/11 vs 09/10, through 11/12 vs 10/11, respectively. The sample of analysis considers 8th grade cohorts in Public Schools, in their first 8th grade, and in the grade-retention margin, defined in an equal manner to the 7th grade. Treatment effect estimates take into account all covariates described in Appendix A3.1.

3.6 Conclusion

This chapter studies the effect of a Compulsory Schooling Leaving age (CSL) increase, from 15 to 18 years old, on students' educational paths. The Portuguese reform provides a noteworthy opportunity to study the effects of increasing the CSL age as the country had one of the EU's highest early school leaving rates in 2008 at 35%, increased its CSL age from 15 to 18 years old in 2009, and achieved a rate of 8.9% in 2020, attaining the EU 2020 target of a rate under 10%. Moreover, CSL age changes have not been frequent in Europe as of late and these increases are usually lower than three years.

Students enrolled in grades 1 to 7 from the academic year of 2009/10 onward were required by law to stay in school until they have finished the 12th grade of schooling or have turned 18 years old, while students enrolled in the 8th grade or above in 2009/10 still fell under the old CSL and could leave school when they had finished the 9th grade or turned 15 years old. As such, grade-retention in 7th grade in the 2008/09 academic year determined exposure to the new CSL, while for other 7th grade cohorts, retention did not lead to different treatment exposure. The Difference-in-Differences identification strategy used in this chapter takes advantage of the way the legislation was set, using grade-retention as quasi-experimental variation to estimate causal effects of the reform on students' educational outcomes. To add credibility to the parallel-trends assumption that is needed for Difference-in-Differences estimates, only students at the margin of retention are considered. At the same time, these results are complemented with Event-Study estimates that compare the first treated cohorts with the last untreated cohorts, helping us identify who the main *compliers* with the CSL reform were — that is, those who would have dropped out of the school system earlier, in the absence of the new

policy. In line with the literature, Event-Study estimates suggest that the main *compliers* are both lower-achieving and lower socio-economic status students.

The main Difference-in-Differences estimates suggest that as a result of the higher CSL age students stayed longer in school, although it is less clear if graduation probabilities were affected. I estimate that the reform increased compliance with the new CSL — i.e., staying in school until 18 years old, or graduating Upper-Secondary schooling — in the first affected cohort of students at the retention margin by 2.8 percentage points, but no statistically significant effects are found for graduation probabilities. However, treatment effect heterogeneity analysis indicate that, even within the retention margin, groups of students that had even lower levels of prior achievement, such as students that have likely had previous grade retentions before the 7th grade, could benefit more from CSL policies. For these students, a positive effect of 3.2 p.p. for graduation probabilities is found but it is not statistically significant. Exploring further heterogeneity within this sub-group of students with lower prior achievement, I find positive graduation treatment effects for students that ever enrolled in Upper-Secondary School (6.1 p.p.), and Boys (5.7 p.p.), that are statistically significant at the 10% level. These results could be related with the fact that right before 2012/13 — the first academic year when students under the new CSL could have started Upper-Secondary school — a law determining measures to prevent early school leaving, such as better guidance and support to at-risk students, giving further emphasis to the enrollment in vocational and alternative tracks as a preventive measure, was approved. These results give support to the hypothesis that alongside CSL age increases, Upper-Secondary schools, and policies targeting at-risk students, may have meaningful impacts in reducing early school leaving.

References

- Adamecz-Völgyi, Anna. 2021. "Is Raising the School Leaving Age Enough to Decrease Dropping Out?" Technical report, GLO Discussion Paper.
- Adamecz-Völgyi, Anna, Dániel Prinz, Ágnes Szabó-Morvai, and Sunčica Vujić. 2021. "The Labor Market and Fertility Impacts of Decreasing the Compulsory Schooling Age." Technical report, KRTK-KTI Working Papers.
- Adler, Alejandro. 2017. "Well-Being and Academic Achievement: Towards a New Evidence-Based Educational Paradigm." In *Future Directions in Well-Being*, 203–8. Springer.
- Aizer, Anna, Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney. 2016. "The Long-Run Impact of Cash Transfers to Poor Families." *American Economic Review* 106 (4): 935–71.
- Almond, Douglas, and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives* 25 (3): 153–72.
- Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature* 56 (4): 1360–1446.
- Andersen, S C, M Gensowski, S Ludeke, J H Pedersen, and others. 2018. "Evaluering Af Den Nationale Trivselsmåling for Folkeskoler--Og Forslag Til Justeringer.[Evaluation of the National Social Well-Being Measurement--and Suggestions for Adjustments]". Aarhus; Trygfondens Børneforskningcenter, Aarhus Universitet
- Andersen, Simon Calmar, Miriam Gensowski, Steven G Ludeke, and Oliver P John. 2020. "A Stable Relationship between Personality and Academic Performance from Childhood through Adolescence. An Original Study and Replication in Hundred-Thousand-Person Samples." *Journal of Personality* 88 (5): 925–39.

- Anderson, D Mark. 2014. "In School and out of Trouble? The Minimum Dropout Age and Juvenile Crime." *Review of Economics and Statistics* 96 (2): 318–31.
- Antecol, Heather, Ozkan Eren, and Serkan Ozbeklik. 2015. "The Effect of Teacher Gender on Student Achievement in Primary School." *Journal of Labor Economics* 33 (1): 63–89.
- Arcidiacono, Peter, V Joseph Hotz, and Songman Kang. 2012. "Modeling College Major Choices Using Elicited Measures of Expectations and Counterfactuals." *Journal of Econometrics* 166 (1): 3–16.
- Bartolj, Tjaša, and Sašo Polanec. 2012. "College Major Choice and Ability: Why Is General Ability Not Enough?" *Economics of Education Review* 31 (6): 996–1016.
- Belfi, Barbara, Mieke Goos, Bieke De Fraine, and Jan Van Damme. 2012. "The Effect of Class Composition by Gender and Ability on Secondary School Students' School Well-Being and Academic Self-Concept: A Literature Review." *Educational Research Review*. Elsevier. <https://doi.org/10.1016/j.edurev.2011.09.002>.
- Bell, Brian, Rui Costa, and Stephen Machin. 2016. "Crime, Compulsory Schooling Laws and Education." *Economics of Education Review* 54: 214–26.
- Bettinger, Eric P, and Bridget Terry Long. 2005. "Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students." *American Economic Review* 95 (2): 152–57.
- Betts, Julian R. 2011. "The Economics of Tracking in Education." In *Handbook of the Economics of Education*, 3:341–81. Elsevier.
- Beuchert, Louise, and Anne Nandrup. 2017. "The Danish National Tests at a Glance." *Available at SSRN 2996431*.
- Bishop, John H. 2010. "Which Secondary Education Systems Work Best? The United States or Northern Europe." Cornell University Working papers. www.digitalcommons.ilr.cornell.edu/workingpapers/

- Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *The Quarterly Journal of Economics* 122 (1): 409–39.
- Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2011. "Too Young to Leave the Nest? The Effects of School Starting Age." *The Review of Economics and Statistics* 93 (2): 455–67.
- Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2011. "Too Young to Leave the Nest? The Effects of School Starting Age." *The Review of Economics and Statistics* 93 (2): 455–67.
- Boone, Simon, Marie Seghers, and Mieke Van Houtte. 2018. "Transition from Primary to Secondary Education in a Rigidly Tracked System: The Case of Flanders." In *Educational Choices, Transitions and Aspirations in Europe: Systemic, Institutional and Subjective Challenges*, 53–70. Routledge.
- Bradshaw, Jonathan, Antonia Keung, Gwyther Rees, and Haridhan Goswami. 2011. "Children's Subjective Well-Being: International Comparative Perspectives." *Children and Youth Services Review* 33 (4): 548–56. <https://doi.org/10.1016/j.childyouth.2010.05.010>.
- Brunello, Giorgio, and Daniele Checchi. 2007. "Does School Tracking Affect Equality of Opportunity? New International Evidence." *Economic Policy* 22 (52): 782–861.
- Brunello, Giorgio, and Maria de Paola. 2014. "The Costs of Early School Leaving in Europe." *IZA Journal of Labor Policy* 3 (1): 1–31.
- Bücker, Susanne, Sevim Nuraydin, Bianca A Simonsmeier, Michael Schneider, and Maike Luhmann. 2018. "Subjective Well-Being and Academic Achievement: A Meta-Analysis." *Journal of Research in Personality* 74: 83–94.
- Cabus, Sofie J., and Kristof De Witte. 2011. "Does School Time Matter? On the Impact of Compulsory Education Age on School Dropout." *Economics of Education Review*. <https://doi.org/10.1016/j.econedurev.2011.07.003>.

- Callaway, Brantly, and Pedro H C Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225 (2): 200–230.
- Carrell, Scott E, and Mark L Hoekstra. 2010. "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids." *American Economic Journal: Applied Economics* 2 (1): 211–28.
- Chetty, Raj, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR." *The Quarterly Journal of Economics* 126 (4): 1593–1660.
- Cho, Rosa Minhyo. 2012. "Are There Peer Effects Associated with Having English Language Learner (ELL) Classmates? Evidence from the Early Childhood Longitudinal Study Kindergarten Cohort (ECLS-K)." *Economics of Education Review* 31 (5): 629–43.
- Cratty, Dorothyjean. 2012. "Potential for Significant Reductions in Dropout Rates: Analysis of an Entire 3rd Grade State Cohort." *Economics of Education Review* 31 (5): 644–62.
- Conti, G. and J. J. Heckman (2014). "Economics of child well-being." In: Handbook of child well- being, eds. A. Ben-Arieh, F. Casas, I. Frønes and J. Korbin (Springer, Dordrecht), 363-401.
- Currie, Janet, and Douglas Almond. 2011. "Human Capital Development before Age Five." In *Handbook of Labor Economics*, 4:1315–1486. Elsevier.
- Currie, Janet, Mark Stabile, Phongsack Manivong, and Leslie L Roos. 2010. "Child Health and Young Adult Outcomes." *Journal of Human Resources* 45 (3): 517–48.
- Dahl, Gordon B, and Lance Lochner. 2012. "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." *American Economic Review* 102 (5): 1927–56.

- Danish Ministry of Education (2016), “Methods Brief: Calculating the Well-being Indicators” (Metodenotat om beregning af trivselsindikatorer, 2014/2015), <https://www.uvm.dk/-/media/filer/uvm/stat/pdf16/160316-metodenotat-vedr-indikatorberegningerne-24022016.pdf.%20Accessed%20August%202020>
- Danish Ministry of Education (2019), “Brief on Pupil Well-being in Primary and Lower Secondary School” (Tabelnotat om trivsel hos eleverne i folkeskolen, 2019) <https://www.uvm.dk/-/media/filer/uvm/stat/pdf19/191112-figurnotat-trivsel.pdf>. Accessed August 2020.
- Groote, Olivier De. 2019. “A Dynamic Model of Effort Choice in High School.” Toulouse, France: TSE Working Papers 19- 1002. Toulouse School of Economics (TSE).
- Dee, Thomas S. 2007. “Teachers and the Gender Gaps in Student Achievement.” *Journal of Human Resources* 42 (3): 528–54.
- Devereux, Paul J, and Robert A Hart. 2010. “Forced to Be Rich? Returns to Compulsory Schooling in Britain.” *The Economic Journal* 120 (549): 1345–64.
- Diette, Timothy M, and Ruth Uwaifo Oyelere. 2014. “Gender and Race Heterogeneity: The Impact of Students with Limited English on Native Students’ Performance.” *American Economic Review* 104 (5): 412–17.
- Erten, Bilge, and Pinar Keskin. 2019. “Compulsory Schooling for Whom? The Role of Gender, Poverty, and Religiosity.” *Economics of Education Review*. <https://doi.org/10.1016/j.econedurev.2019.06.001>.
- Eurydice. 2011. *Grade Retention during Compulsory Education in Europe: Regulations and Statistics*. EU: European Union.
- Eurydice. 2017. <https://eurydice.eacea.ec.europa.eu/>
- Eurydice. 2019. *Compulsory Education in Europe : 2018/19*. European Commission. <https://doi.org/doi/10.2797/20126>.

- Figlio, David N. 2007. "Boys Named Sue: Disruptive Children and Their Peers." *Education Finance and Policy* 2 (4): 376–94.
- Gal, Iddo, and Dave Tout. 2014. "Comparison of PIAAC and PISA Frameworks for Numeracy and Mathematical Literacy."
- Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225 (2): 254–77.
- Gutman, Leslie, and Leon Feinstein. 2008. *Children's Well-Being in Primary School: Pupil and School Effects [Wider Benefits of Learning Research Report No. 25]*. Centre for Research on the Wider Benefits of Learning.
- Hanushek, Eric A, and Ludger W ößmann. 2006. "Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries." *The Economic Journal* 116 (510): C63--C76.
- Harmon, Colm P. 2017. "How Effective Is Compulsory Schooling as a Policy Instrument?" *IZA World of Labor*.
- Imbens, Guido W. 2015. "Matching Methods in Practice: Three Examples." *Journal of Human Resources* 50 (2). <https://doi.org/10.3368/jhr.50.2.373>.
- Jakubowski, Maciej. 2010. "Institutional Tracking and Achievement Growth: Exploring Difference-in-Differences Approach to PIRLS, TIMSS, and PISA Data." In *Quality and Inequality of Education*, 41–81. Springer.
- Jerrim, John, Luis Alejandro Lopez-Agudo, Oscar D Marcenaro-Gutierrez, and Nikki Shure. 2017. "What Happens When Econometrics and Psychometrics Collide? An Example Using the PISA Data." *Economics of Education Review* 61: 51–58.
- Keilow, Maria, Anders Holm, Stine Bagger, and Sofie Henze-Pedersen. 2014. *Udvikling Af Trivselsmålinger i Folkeskolen: En Pilotundersøgelse*. SFI.

- Knoop, Hans Henrik, Bjørn Evald Holstein, Hanne Viskum, and Jannie Moon Lindskov. 2017. "Elevernes Fællesskab Og Trivsel i Skolen.: Analyser Af Den Nationale Trivselsmåling."
- Kristoffersen, Jannie Helene Grøne, Morten Visby Krægpøth, Helena Skyt Nielsen, and Marianne Simonsen. 2015. "Disruptive School Peers and Student Outcomes." *Economics of Education Review* 45: 1–13.
- Landersø, Rasmus, Helena Skyt Nielsen, and Marianne Simonsen. 2017. "School Starting Age and the Crime-Age Profile." *The Economic Journal* 127 (602): 1096–1118.
- Landis, Rebecca N., and Amy L. Reschly. 2011. "An Examination of Compulsory School Attendance Ages and High School Dropout and Completion." *Educational Policy*. <https://doi.org/10.1177/0895904810374851>.
- Lleras-Muney, Adriana. 2005. "The Relationship between Education and Adult Mortality in the United States." *The Review of Economic Studies* 72 (1): 189–221.
- Maaz, Kai, Ulrich Trautwein, Oliver Lüdtke, and Jürgen Baumert. 2008. "Educational Transitions and Differential Learning Environments: How Explicit between-School Tracking Contributes to Social Inequality in Educational Outcomes." *Child Development Perspectives* 2 (2): 99–106.
- Maruyama, Shiko, and Eskil Heinesen. 2020. "Another Look at Returns to Birthweight." *Journal of Health Economics* 70: 102269.
- Meier, Volker, and Gabriela Schütz. 2007. "The Economics of Tracking and Non-Tracking."
- Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. 2004. "Does Education Improve Citizenship? Evidence from the United States and the United Kingdom." *Journal of Public Economics* 88 (9–10): 1667–95.
- Niclasen, Janni, Maria Keilow, and Carsten Obel. 2018. "Psychometric Properties of the Danish Student Well-Being Questionnaire Assessed in >250,000 Student

Responders.” *Scandinavian Journal of Public Health* 46 (8): 877–85.
<https://doi.org/10.1177/1403494818772645>.

Nielsen, Chantal Pohl, Beatrice Schindler Rangvid, C P Christensen, C B Dyssegaard, N Egelund, and S Henze-Pedersen. 2016. *Inklusion i Folkeskolen: Sammenfatning Af Resultaterne Fra Inklusionspanelet*. SFI-Det Nationale Forskningscenter for Velfærd.

Nixon, Lucia A, and Michael D Robinson. 1999. “The Educational Attainment of Young Women: Role Model Effects of Female High School Faculty.” *Demography* 36 (2): 185–94.

OECD. 2014. “Are Disadvantaged Students More Likely to Repeat Grades?” *PISA in Focus*, no. 09.

Oosterbeek, Hessel, and Dinand Webbink. 2007. “Wage Effects of an Extra Year of Basic Vocational Education.” *Economics of Education Review* 26 (4): 408–19.

Oreopoulos, Philip. 2007. “Do Dropouts Drop out Too Soon? Wealth, Health and Happiness from Compulsory Schooling.” *Journal of Public Economics* 91 (11–12): 2213–29.

Organization, World Health, and others. 2016. *Growing up Unequal: Gender and Socioeconomic Differences in Young People’s Health and Well-Being*. World Health Organization. Regional Office for Europe.

Pekkarinen, Tuomas, Roope Uusitalo, and Sari Kerr. 2009. “School Tracking and Intergenerational Income Mobility: Evidence from the Finnish Comprehensive School Reform.” *Journal of Public Economics* 93 (7–8): 965–73.

Pischke, Jörn-Steffen, and Till Von Wachter. 2008. “Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation.” *The Review of Economics and Statistics* 90 (3): 592–98.

- Rapoport, Benôit, and Claire Thibout. 2018. "Why Do Boys and Girls Make Different Educational Choices? The Influence of Expected Earnings and Test Scores." *Economics of Education Review* 62: 205–29.
- Rivkin, Steven G, Eric A Hanushek, and John F Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73 (2): 417–58.
- Ruhose, Jens, and Guido Schwerdt. 2016. "Does Early Educational Tracking Increase Migrant-Native Achievement Gaps? Differences-in-Differences Evidence across Countries." *Economics of Education Review* 52: 134–54.
- Tarabini, Aina, Marta Curran, Alba Castejón, and Alejandro Montes. 2018. "Framing Youth Educational Choices at the End of Compulsory Schooling: The Catalan Case." In *Educational Choices, Transitions and Aspirations in Europe*, 95–109. Routledge.
- Tsang, Kwan Lan Vicky, Paul Y H Wong, and Sing Kai Lo. 2012. "Assessing Psychosocial Well-Being of Adolescents: A Systematic Review of Measuring Instruments." *Child: Care, Health and Development* 38 (5): 629–46.
- UNESCO United Nations Educational, Scientific, and Cultural Organization. 2003. *International Standard Classification of Education, ISCED 1997*. Springer.
- Waldinger, Fabian. 2007. "Does Ability Tracking Exacerbate the Role of Family Background for Student's Test Scores?" Unpublished manuscript. London, UK: London School of Economics.
- Woessmann, Ludger. 2009. "International Evidence on School Tracking: A Review." *CESifo DICE Report* 7 (1): 26–34.

A1 Appendix Chapter 1

The present document has the purpose of justifying the classification of each country's academic Upper-Secondary school system used in the first chapter of this thesis, previously published as "The effect of a Specialized versus a General Upper-Secondary school curriculum on students' performance and inequality. A difference-in-differences cross country analysis" (Câmara Leme et al., 2020). The 49 pages document is available in <https://www.tandfonline.com/doi/suppl/10.1080/00036846.2020.1734183>

Countries differ in their academic upper-secondary school systems in a way that some require their students to choose a specialization track from a set of areas — typically natural sciences, economic sciences, humanities or arts - and follow that specialization for the course of their upper secondary education years (e.g. Portugal, Spain, Sweden). By contrast, others including Finland, Germany, or England, follow a general curriculum where students, although able to choose between different classes in distinct areas, are not required to follow a single specialization and thus, may receive a more general education. A country is therefore classified as having a specialized system in academic upper secondary school if there is a clear set of specialization tracks that a student has to choose from in this stage of education. The references used for classifying countries' upper secondary school systems in the 2009 to 2012 period come mostly from Eurydice Network (Eurydice, 2017) — a network of the European Commission that compiles a vast set of information about the structure of education in Europe at all levels, besides detailed descriptions and overviews of national education systems. Other sources range from OECD documents, information available on the websites of the Ministry of

Education of each country and other articles about the curriculum and educational system of a country. Because several of these references come directly from webpages, the weblink to these pages were added in the main text of the article and citations from these pages were also included, whenever we considered appropriate. This document also provides information on whether Mathematics is compulsory for all students in academic tracks at any point of academic upper secondary school — a variable used as a robustness check in the main article.

A2 Appendix Chapter 2

Table A2.1
Variable definitions

Background Variables	Definition	Source
<i>Child characteristics</i>		
Male	Indicator if male	BEF
Birthweight is below 2500 grams	Indicator if birthweight is below 2500 grams	MFR
Young for grade	Indicator if pupils age is young for grade ("ahead of schedule")	UDSP, BEF
Late school start	Indicator if pupil started in the school later than year the child turned 6	KOTRE, BEF
Repeated a grade	Indicator if the pupil ever has repeated a grade	KOTRE, BEF
Switched address, since school start	Indicator if the pupil has switched address since school start	KOTRE, BEF
Switched school, since school start	Indicator if the pupil has switched schools since school start	KOTRE
Switched school but not address , since school start	Indicator if the pupil has switched school but not address since school start	KOTRE, BEF
Placed outside home, 2010 to 2018	Indicator if placed outside home	BUAS
Received social preventive measures by municipality, since birth	Indicator if received social preventive measures	BUFO
Avg. no. of contacts per year to hospital, 2014 to 2018	Avg. no. of daily contacts per year to hospital (Max 1 contact per day)	LPR_ADM
Been in contact with psychiatric hospital, since birth	Indicator if pupil has been in contact with psychiatric hospital	LPSYADM
Special needs student in ordinary classroom	Indicator if student received special education	UDSP
Special needs student in special needs classroom	Indicator if student received special education	UDSP
<i>Maternal characteristics in 2016</i>		
Age at child birth	Age at birth of child in years (birth date of parent - birth date of child)/365.25	BEF
Unskilled	Indicator if parent's highest obtained education is unskilled	UDDA
Highschool	Indicator if parent's highest obtained education is highschool	UDDA
Short further	Indicator if parent's highest obtained education is short further	UDDA
Medium further	Indicator if parent's highest obtained education is medium further	UDDA
Long further or PhD.	Indicator if parent's highest obtained education is long further or ph.d.	UDDA
Taxable income in 2nd quartile, last year	Indicator if parent's disposable income is in 2nd quartile	IND
Taxable income in 3rd quartile, last year	Indicator if parent's disposable income is in 3rd quartile	IND
Taxable income in 4th quartile, last year	Indicator if parent's disposable income is in 4th quartile	IND
Received unemployment assistance, last year	Indicator if received unemployment assistance	DREAM
Received unemployment insurance, last year	Indicator if received unemployment insurance	DREAM
Avg. no. of contacts per year to hospital, 2014 to 2018	Avg. no. of daily contacts per year to hospital (Max 1 contact per day)	LPR_ADM
Been in contact with psychiatric hospital, since birth of pupil	Indicator if parent has been in contact with psychiatric hospital, since birth of pupil	LPSYADM
Committed criminal offense, since birth of pupil	Indicator if ever committed a criminal offense (Note: Not traffic related)	KRAF

<i>Paternal characteristics in 2016</i>		
Age at child birth	Age at birth of child in years (birth date of parent - birth date of child)/365.25	BEF
Unskilled	Indicator if parent's highest obtained education is unskilled	UDDA
Highschool	Indicator if parent's highest obtained education is high school	UDDA
Short further	Indicator if parent's highest obtained education is short further	UDDA
Medium further	Indicator if parent's highest obtained education is medium further	UDDA
Long further or PhD.	Indicator if parent's highest obtained education is long further or ph.d.	UDDA
Taxable income in 2nd quartile, last year	Indicator if parent's disposable income is in 2nd quartile	IND
Taxable income in 3rd quartile, last year	Indicator if parent's disposable income is in 3rd quartile	IND
Taxable income in 4th quartile, last year	Indicator if parent's disposable income is in 4th quartile	IND
Received unemployment assistance, last year	Indicator if received unemployment assistance	DREAM
Received unemployment insurance, last year	Indicator if received unemployment insurance	DREAM
Avg. no. of contacts per year to hospital, 2014 to 2018	Avg. no. of daily contacts per year to hospital (Max 1 contact per day)	LPR_ADM
Been in contact with psychiatric hospital, since birth of pupil	Indicator if parent has been in contact with psychiatric hospital, since birth of pupil	LPSYADM
Committed criminal offense, since birth of pupil	Indicator if ever committed a criminal offense (Note: Not traffic related)	KRAF
<i>Household characteristics in 2016</i>		
Number of siblings	Number of siblings (both half and full)	BEF
Birth order on mother's side	Birth order on mother's side	BEF
Parents have lived apart for two consecutive years	Indicator if parents had different adress last two years consecutively	BEF
Single provider household	Indicator if pupil is living in a single provider household	BEF
<i>Teacher characteristics</i>		
Note for teachers variables: They are calculated as weighted averages across the 4 teachers, which spent most time with students last year, weighted by their time spent with the pupil		
Male, Danish	Indicator if teacher is male	BEF
Age, Danish	Age of teacher	BEF
Educated in teaching Danish or comparable	Indicator if teacher had Danish specialization or have skills comparable to it	Teacher Competencies
Years since degree, 0-10 years, Danish	Spline for years since degree from 0 to 10 years	UDDA
10 years or above since degree, Danish	Indicator if years since degree is 10 or over	UDDA
Tenure is 2 to 3 years, Danish	Indicator if teachers tenure is 2 to 3 years at current school	Teacher Competencies
Tenure is 4 to 5 years, Danish	Indicator if teachers tenure is 4 to 5 years at current school	Teacher Competencies
Tenure is 6 years or above, Danish	Indicator if teachers tenure is 6 years or above at current school	Teacher Competencies
Share of days absent, Danish	Share of days absent	FRPE
Age, Math	Indicator if teacher is male	BEF
Male, Math	Age of teacher	BEF
Educated in teaching math or comparable	Indicator if teacher had math specialization or have skills comparable to it	Teacher Competencies
Years since degree, 0-10 years, Math	Spline for years since degree from 0 to 10 years	UDDA
10 years or above since degree, Math	Indicator if years since degree is 10 or over	UDDA
Tenure is 2 to 3 years, Math	Indicator if teachers tenure is 2 to 3 years at current school	Teacher Competencies
Tenure is 4 to 5 years, Math	Indicator if teachers tenure is 4 to 5 years at current school	Teacher Competencies
Tenure is 6 years or above, Math	Indicator if teachers tenure is 6 years or above at current school	Teacher Competencies
Share of days absent, Math	Share of days absent	FRPE

<i>Classroom characteristics</i>		
Size	Size of classroom	UDSP
Share of males	Share of males in classroom	BEF
Share with birthweight below 2500 grams	Share of pupils with a birthweight below 2500 grams	MFR
Share of who are young for grade	Share of pupils who are young for grade ("Ahead of schedule")	UDSP, BEF
Share of late starters	Share that started school after the year they turned 6	KOTRE
Share of who repeated a grade	Share of pupils who have ever repeated a grade	KOTRE, BEF
Share of who switched address, since school start	Share of pupils who have switched address since school start	KOTRE, BEF
Share of who switched school, since school start	Share of pupils who have switched schools since school start	KOTRE, BEF
Share of who switched school but not address, since school start	Share of pupils who have switched school but not address since school start	KOTRE, BEF
Share of who have been placed outside home, from 2010	Share of pupils who have been placed outside home	BUAS
Share of who received social preventive measures by municipality, since birth	Share of pupils who have received social preventive measures	BUFO
Avg. no. of contacts per year to hospital, last 5 years	Avg. no. Of daily contacts per year to hospital (Max 1 contact per day)	LPR_ADM
Share of who has been in contact with psychiatric hospital, since birth	Share of who has been in contact with psychiatric hospital, since birth	LPSYADM
Share with special needs in ordinary classrooms	Share with special education needs in regular classrooms	UDSP
Share with a criminal parent	Share with a criminal parent	KRAF
Share with parents have med./long education	Share with parents that have medium or long further education	UDDA
Share with parents who received unemployment assistance/insurance, last year	Share of parents which received unemployment assistance or insurance, last year	DREAM
Share with a parent who has been in contact with psychiatric hospital, since birth	Share with a parent who has been in contact with psychiatric hospital, since birth	LPSYADM
Share with mother's income in 2nd quartile, last year	Share of mothers which belonged to 2nd income quartile	IND
Share with mother's income in 3rd quartile, last year	Share of mothers which belonged to 3rd income quartile	IND
Share with mother's income in 4th quartile, last year	Share of mothers which belonged to 4th income quartile	IND
Share with father's income in 2nd quartile, last year	Share of fathers which belonged to 2nd income quartile	IND
Share with father's income in 3rd quartile, last year	Share of fathers which belonged to 3rd income quartile	IND
Share with father's income in 4th quartile, last year	Share of fathers which belonged to 4th income quartile	IND
<i>School characteristics</i>		
Turnover rate of Danish and math teachers	Turnover rate of Danish and math teachers at school	Teachers
Turnover rate of all teachers	Turnover rate of all teachers at school	Competencies
Size	Size of school	Teachers
<i>HIDDEN CONTROLS</i>		
Immigrant	Indicator if immigrant	BEF
Immigrant, nonwestern	Indicator if immigrant, nonwestern	BEF
Descendant	Indicator if descendant	BEF
Descendant, nonwestern	Indicator if descendant, nonwestern	BEF
Class share: Immigrant, nonwestern	Share in class that is nonwestern immigrant	BEF
Class share: Immigrant	Share in class that is immigrant	BEF
Class share: Descendant, nonwestern	Share in class that is nonwestern descendant	BEF
Class share: Descendant	Share in class that is descendant	BEF
More than 4 teachers, Danish	Indicator if pupil has more than 4 danish teachers connected	Teachers
More than 4 teachers, Math	Indicator if pupil has more than 4 math teachers connected	Competencies

MISSING CONTROLS

M: Danish teacher	Indicator if no Danish teacher is traceable
M: Share of days absent, Danish	Indicator if Danish teachers absence data is missing
M: Danish teacher characteristics	Indicator if one or multiple Danish teacher characteristics is missing
M: Math teacher	Indicator if no math teacher is traceable
M: Share of days absent, Math	Indicator if math teachers absence data is missing
M: Math teacher characteristics	Indicator if one or multiple math teacher characteristics is missing
M: Late school start	Indicator if started school late is missing (This can be missing if they e.g. go directly into 1st grade and perhaps had kindergarten in another country). Can be changed.
M: Turnover rate of teachers at school	Indicator if turnover rate of teachers is missing (Likely due to schools not reporting)
M: Birth information	Indicator if birth information is missing (This is not perfectly correlated with being immigrant)
M: Birth order on mothers side	Indicator if birth order is missing (If e.g. Mor_id is missing)
M: Fathers taxable income	indicator if fathers income is missing
M: Mothers taxable income	indicator if mothers income is missing
M: father receives unemployment assistance	Indicator if missing if father received any unemployment assistance
M: mother receives unemployment assistance	Indicator if missing if mother received any unemployment assistance
M: Father educ.	Indicator if missing fathers education
M: Mother educ.	Indicator if missing mothers education.

Table A2.2

Background characteristics by social well-being quartiles

Background Variables	Full Sample	1st quartile	2nd quartile	3rd quartile	4th quartile
<i>Child characteristics</i>					
Male	0.51373 (0.49981)	0.42523 (0.49438)	0.49193 (0.49994)	*** 0.54272 (0.49817)	*** 0.57973 (0.49361)
Birthweight is below 2500 grams	0.05160 (0.22121)	0.05278 (0.22360)	0.05103 (0.22006)	* 0.04951 (0.21693)	* 0.05172 (0.22146)
Age relative to cohort (Ref.cat.: Adequate for grade)					
Age adequate for grade	0.86819 (0.33828)	0.86101 (0.34594)	0.87496 (0.33077)	*** 0.88325 (0.32112)	*** 0.88814 (0.31520)
Young for grade	0.01365 (0.11605)	0.01489 (0.12113)	0.01357 (0.11570)	** 0.01325 (0.11434)	** 0.01196 (0.10869)
Late school start	0.08324 (0.27624)	0.08438 (0.27796)	0.07872 (0.26931)	*** 0.07452 (0.26261)	*** 0.07312 (0.26033)
Repeated a grade	0.04548 (0.20835)	0.05096 (0.21991)	0.04235 (0.20138)	*** 0.03754 (0.19008)	*** 0.03406 (0.18140)
Switched address, since school start	0.25322 (0.43486)	0.26316 (0.44035)	0.24990 (0.43296)	*** 0.22642 (0.41851)	*** 0.21161 (0.40845)
Switched school, since school start	0.19327 (0.39486)	0.20004 (0.40003)	0.18322 (0.38685)	*** 0.15667 (0.36349)	*** 0.13974 (0.34672)
Switched school but not address, since school start	0.15620 (0.36304)	0.15757 (0.36434)	0.14765 (0.35475)	*** 0.12574 (0.33156)	*** 0.11151 (0.31477)
Placed outside home, from 2010	0.01097 (0.10418)	0.01447 (0.11942)	0.00921 (0.09554)	*** 0.00790 (0.08851)	*** 0.00723 (0.08474)
Received social preventive measures by municipality, since birth	0.07994 (0.27120)	0.10339 (0.30447)	0.07006 (0.25525)	*** 0.05541 (0.22877)	*** 0.04824 (0.21428)
Avg. no. of contacts per year to hospital, last 5 years	0.38499 (0.60267)	0.39204 (0.58994)	0.36691 (0.55954)	*** 0.36176 (0.56106)	*** 0.35629 (0.54720)
Been in contact with psychiatric hospital, since birth	0.06198 (0.24111)	0.07320 (0.26047)	0.05094 (0.21988)	*** 0.04269 (0.20215)	*** 0.03651 (0.18756)
Special needs student in ordinary classroom	0.00337 (0.05798)	0.00449 (0.06687)	0.00256 (0.05054)	*** 0.00244 (0.04932)	*** 0.00200 (0.04466)
Special needs student in special needs classroom	0.02907 (0.16800)	0.03245 (0.17719)	0.02324 (0.15067)	*** 0.01906 (0.13674)	*** 0.01810 (0.13332)

<i>Maternal characteristics in 2018</i>								
Age at child birth	30.37012 (5.49963)	30.11042 (5.71404)	30.42961 (5.47734)	***	30.53152 (5.33170)	***	30.61715 (5.24650)	***
Education (Ref.cat.: Vocational)								
Unskilled	0.12831 (0.33443)	0.15144 (0.35848)	0.12232 (0.32766)	***	0.11092 (0.31403)	***	0.10107 (0.30143)	***
Highschool	0.04971 (0.21735)	0.05248 (0.22299)	0.04928 (0.21644)	***	0.04846 (0.21473)	***	0.04669 (0.21098)	***
Vocational	0.31979 (0.46640)	0.33735 (0.47281)	0.32269 (0.46751)	***	0.31219 (0.46339)	***	0.30004 (0.45828)	***
Short further	0.05585 (0.22963)	0.05492 (0.22783)	0.05803 (0.23380)	***	0.05758 (0.23294)	***	0.05783 (0.23342)	***
Medium further	0.28583 (0.45181)	0.26260 (0.44005)	0.28666 (0.45220)	***	0.29919 (0.45790)	***	0.31530 (0.46464)	***
Long further or PhD.	0.14623 (0.35334)	0.12480 (0.33049)	0.14683 (0.35394)	***	0.15884 (0.36553)	***	0.16763 (0.37354)	***
Taxable income quartile (Ref.cat.: 1st)								
Taxable income in 1st quartile, last year	0.24868 (0.43225)	0.28177 (0.44986)	0.23873 (0.42631)	***	0.22467 (0.41737)	***	0.21188 (0.40864)	***
Taxable income in 2nd quartile, last year	0.24868 (0.43225)	0.25778 (0.43741)	0.25351 (0.43502)	**	0.24839 (0.43208)	**	0.24597 (0.43067)	**
Taxable income in 3rd quartile, last year	0.24868 (0.43225)	0.23748 (0.42554)	0.25181 (0.43405)	***	0.25684 (0.43689)	***	0.26296 (0.44025)	***
Taxable income in 4th quartile, last year	0.24867 (0.43224)	0.21717 (0.41232)	0.25064 (0.43338)	***	0.26542 (0.44156)	***	0.27485 (0.44644)	***
Received unemployment assistance, last year	0.26424 (0.44093)	0.29291 (0.45510)	0.25451 (0.43559)	***	0.23922 (0.42661)	***	0.22754 (0.41925)	***
Received unemployment insurance, last year	0.06721 (0.25038)	0.07480 (0.26307)	0.06779 (0.25139)	***	0.06535 (0.24714)	***	0.06096 (0.23925)	***
Avg. no. of contacts per year to hospital, last 5 years	0.82333 (1.00021)	0.87431 (1.04085)	0.80522 (0.96931)	***	0.78971 (0.96916)	***	0.77310 (0.94475)	***
Been in contact with psychiatric hospital, since birth of pupil	0.10139 (0.30185)	0.11817 (0.32282)	0.09726 (0.29632)	***	0.08714 (0.28204)	***	0.08001 (0.27131)	***
Committed criminal offense, since birth of pupil	0.03043 (0.17177)	0.03687 (0.18844)	0.02768 (0.16406)	***	0.02425 (0.15382)	***	0.02197 (0.14658)	***

Paternal characteristics in 2018								
Age at child birth	32.13440 (8.07182)	31.93417 (8.45100)	32.22563 (8.01749)	***	32.31674 (7.73720)	***	32.40648 (7.44471)	***
Education (Ref.cat.: Vocational)								
Unskilled	0.15529 (0.36218)	0.17906 (0.38340)	0.15046 (0.35752)	***	0.13966 (0.34663)	***	0.13000 (0.33630)	***
Highschool	0.05063 (0.21923)	0.04949 (0.21688)	0.05147 (0.22097)	**	0.05047 (0.21891)	**	0.04994 (0.21783)	**
Vocational	0.38799 (0.48729)	0.39901 (0.48970)	0.39022 (0.48780)	***	0.38801 (0.48730)	***	0.38172 (0.48581)	***
Short further	0.07782 (0.26789)	0.07395 (0.26168)	0.07931 (0.27022)	***	0.08018 (0.27158)	***	0.08351 (0.27665)	***
Medium further	0.14485 (0.35195)	0.13063 (0.33699)	0.14673 (0.35384)	***	0.15229 (0.35930)	***	0.15905 (0.36572)	***
Long further or ph.d.	0.14061 (0.34762)	0.11934 (0.32420)	0.14041 (0.34741)	***	0.15203 (0.35905)	***	0.16261 (0.36901)	***
Taxable income quartile (Ref.cat.: 1st)								
Taxable income in 1st quartile, last year	0.24378 (0.42936)	0.26851 (0.44319)	0.23597 (0.42461)	***	0.22518 (0.41770)	***	0.21522 (0.41098)	***
Taxable income in 2nd quartile, last year	0.24378 (0.42936)	0.25744 (0.43723)	0.24642 (0.43093)	***	0.24240 (0.42854)	***	0.23475 (0.42385)	***
Taxable income in 3rd quartile, last year	0.24378 (0.42936)	0.23332 (0.42294)	0.24804 (0.43188)	***	0.25098 (0.43358)	***	0.25636 (0.43663)	***
Taxable income in 4th quartile, last year	0.24378 (0.42936)	0.21258 (0.40914)	0.24496 (0.43007)	***	0.25973 (0.43849)	***	0.27454 (0.44628)	***
Recieved unemployment assistance, last year	0.18568 (0.38885)	0.20923 (0.40676)	0.18116 (0.38515)	***	0.16682 (0.37282)	***	0.15739 (0.36417)	***
Recieved unemployment insurance, last year	0.05145 (0.22091)	0.05813 (0.23400)	0.05135 (0.22071)	***	0.04856 (0.21496)	***	0.04612 (0.20975)	***
Avg. no. of contacts per year to hospital, last 5 years	0.55012 (0.83680)	0.57830 (0.84939)	0.53996 (0.78801)	***	0.52926 (0.86299)	***	0.51517 (0.79049)	***
Been in contact with psychiatric hospital, since birth of pupil	0.07350 (0.26096)	0.08507 (0.27899)	0.07004 (0.25521)	***	0.06339 (0.24366)	***	0.05833 (0.23437)	***
Committed criminal offense, since birth of pupil	0.08188 (0.27419)	0.09651 (0.29529)	0.07691 (0.26646)	***	0.06793 (0.25163)	***	0.06128 (0.23984)	***
Household characteristics in 2018								
Number of siblings	1.68814 (1.13885)	1.71789 (1.18139)	1.67402 (1.12369)	***	1.65270 (1.08790)	***	1.62227 (1.06387)	***
Birth order on mother's side	1.82359 (0.93618)	1.84877 (0.94742)	1.82316 (0.92841)	***	1.80999 (0.92045)	***	1.78759 (0.91309)	***
Parents have lived apart for two consecutive years	0.26767 (0.44275)	0.30690 (0.46121)	0.26272 (0.44011)	***	0.23871 (0.42629)	***	0.21766 (0.41266)	***
Single provider household	0.22184 (0.41548)	0.25071 (0.43342)	0.21757 (0.41259)	***	0.19842 (0.39881)	***	0.18141 (0.38536)	***

<i>Teacher characteristics</i>							
Male, Danish	0.17513 (0.37017)	0.16918 (0.36398)	0.17511 *** (0.37016)	0.16768 *** (0.36418)	0.15688 *** (0.35489)		
Age, Danish	43.98348 (9.87307)	43.80572 (9.90664)	43.87648 (9.89354)	44.00977 (9.86853)	44.27215 (9.78593)		
Educated in teaching danish or comparable	0.96342 (0.16587)	0.96040 (0.17160)	0.96384 *** (0.16486)	0.96356 *** (0.16608)	0.96663 *** (0.15803)		
Years since degree, 0-10 years, Danish	1.39579 (2.67607)	1.43440 (2.69829)	1.41494 (2.68332)	1.40473 (2.68656)	1.34032 (2.64765)		
10 years or above since degree, Danish	0.74080 (0.43819)	0.73188 (0.44298)	0.73589 ** (0.44086)	0.73961 ** (0.43885)	0.75365 ** (0.43089)		
Tenure of teacher, Danish(Ref.cat.: 0-1 year)							
Tenure is 0 to 1 years, Danish	0.18620 (0.37594)	0.20591 (0.39016)	0.19042 *** (0.37929)	0.18653 *** (0.37646)	0.17028 *** (0.36350)		
Tenure is 2 to 3 years, Danish	0.19676 (0.38751)	0.20239 (0.39089)	0.20032 (0.39000)	0.19297 (0.38503)	0.18654 (0.38045)		
Tenure is 4 to 5 years, Danish	0.24818 (0.42325)	0.24182 (0.41862)	0.24611 ** (0.42199)	0.24816 ** (0.42336)	0.25109 ** (0.42578)		
Tenure is 6 years or above, Danish	0.36887 (0.47489)	0.34988 (0.46891)	0.36314 *** (0.47335)	0.37234 *** (0.47593)	0.39209 *** (0.48088)		
Share of days absent, Danish	0.04900 (0.09155)	0.05033 (0.09174)	0.04901 *** (0.09072)	0.04857 *** (0.09057)	0.04835 *** (0.09280)		
Age, Math	43.63583 (10.26406)	43.59684 (10.31617)	43.65714 (10.31509)	43.58799 (10.23548)	43.73325 (10.20142)		
Male, Math	0.44771 (0.48802)	0.43081 (0.48512)	0.44629 *** (0.48823)	0.44044 *** (0.48732)	0.43154 *** (0.48694)		
Educated in teaching math or comparable	0.93162 (0.23742)	0.92341 (0.24964)	0.93255 *** (0.23596)	0.93111 *** (0.23822)	0.93308 *** (0.23583)		
Years since degree, 0-10 years, Math	1.33409 (2.57252)	1.35125 (2.57926)	1.34011 (2.56550)	1.34089 (2.57462)	1.31521 (2.56031)		
10 years or above since degree, Math	0.73569 (0.44096)	0.73080 (0.44355)	0.73256 (0.44263)	0.73370 (0.44202)	0.74082 (0.43819)		
Tenure of teacher, Math(Ref.cat.: 0-1 year)							
Tenure is 0 to 1 years, Math	0.20114 (0.38990)	0.21556 (0.40006)	0.20415 *** (0.39220)	0.20383 *** (0.39185)	0.19194 *** (0.38347)		
Tenure is 2 to 3 years, Math	0.19981 (0.39209)	0.20431 (0.39531)	0.20051 ** (0.39260)	0.19585 ** (0.38923)	0.19442 ** (0.38837)		
Tenure is 4 to 5 years, Math	0.23423 (0.41719)	0.22974 (0.41387)	0.23246 (0.41613)	0.23423 (0.41711)	0.23662 (0.41900)		
Tenure is 6 years or above, Math	0.36483 (0.47503)	0.35039 (0.47045)	0.36288 *** (0.47466)	0.36609 *** (0.47534)	0.37702 *** (0.47868)		
Share of days absent, Math	0.04584 (0.08765)	0.04648 (0.08637)	0.04591 (0.08816)	0.04514 (0.08623)	0.04563 (0.08881)		

Classroom characteristics								
Size	21.32845 (4.23776)	21.17975 (4.30615)	21.41371 (4.11619)	***	21.56581 (4.06161)	***	21.63471 (4.03603)	***
Share of males	0.51373 (0.11510)	0.51696 (0.11711)	0.51277 (0.11250)	***	0.51040 (0.10991)	***	0.50785 (0.10821)	***
Share with birthweight below 2500 grams	0.05158 (0.05708)	0.05137 (0.05685)	0.05121 (0.05596)		0.05117 (0.05599)		0.05094 (0.05575)	
Share of who are young for grade	0.01365 (0.02907)	0.01332 (0.02874)	0.01370 (0.02837)	***	0.01338 (0.02837)	***	0.01304 (0.02785)	***
Share of late starters	0.08375 (0.09070)	0.08109 (0.08899)	0.08132 (0.08634)		0.07831 (0.08306)		0.07757 (0.08357)	
Share of who repeated a grade	0.04548 (0.07572)	0.04572 (0.07442)	0.04435 (0.07213)	***	0.04244 (0.06963)	***	0.04112 (0.06947)	***
Share of who switched address, since school start	0.25322 (0.16516)	0.24210 (0.16504)	0.25331 (0.16251)	***	0.24101 (0.16196)	***	0.23263 (0.15812)	***
Share of who switched school, since school start	0.19323 (0.23200)	0.17836 (0.22299)	0.18804 (0.22483)	***	0.17332 (0.21519)	***	0.16040 (0.20673)	***
Share of who switched school but not address, since school start	0.15620 (0.21402)	0.14199 (0.20378)	0.15075 (0.20738)	***	0.13807 (0.19738)	***	0.12686 (0.18926)	***
Share of who have been placed outside home, from 2010	0.01097 (0.03442)	0.01134 (0.03429)	0.01031 (0.03094)	***	0.00977 (0.02958)	***	0.00967 (0.03003)	***
Share of who received social preventive measures by municipality, since birth	0.07994 (0.09970)	0.08360 (0.10320)	0.07829 (0.09332)	***	0.07428 (0.08883)	***	0.07156 (0.08594)	***
Avg. no. of contacts per year to hospital, last 5 years	0.38499 (0.18226)	0.38157 (0.17676)	0.38088 (0.16960)		0.37825 (0.16684)		0.37371 (0.16710)	
Share of who has been in contact with psychiatric hospital, since birth	0.06198 (0.11573)	0.06197 (0.12133)	0.05847 (0.10660)	***	0.05521 (0.09878)	***	0.05277 (0.09489)	***
Share with special needs in ordinary classrooms	0.00347 (0.01706)	0.00380 (0.01835)	0.00343 (0.01652)	***	0.00346 (0.01708)	***	0.00332 (0.01609)	***

Share with a criminal parent	0.18939 (0.12184)	0.19753 (0.12504)	0.18876 *** (0.11964)	0.18092 *** (0.11662)	0.17538 *** (0.11445)
Share with parents have med./long education	0.50437 (0.19949)	0.48879 (0.19840)	0.50315 *** (0.19747)	0.51549 *** (0.19710)	0.52599 *** (0.19891)
Share with parents who received unemployment assistance/insurance, last year	0.42162 (0.15890)	0.43363 (0.16009)	0.42089 *** (0.15666)	0.41433 *** (0.15543)	0.40700 *** (0.15444)
Share with a parent who has been in contact with psychiatric hospital, since bir	0.15694 (0.10625)	0.16066 (0.10868)	0.15597 *** (0.10418)	0.15144 *** (0.10265)	0.14773 *** (0.10079)
Taxable income quartile (Ref.cat.: 1st)					
Share with mother's income in 1st quartile, last year	0.24868 (0.14268)	0.25882 (0.14715)	0.24712 *** (0.14010)	0.24345 *** (0.13771)	0.23809 *** (0.13577)
Share with mother's income in 2nd quartile, last year	0.24868 (0.12242)	0.25482 (0.12302)	0.25018 *** (0.12201)	0.24887 *** (0.12186)	0.24914 *** (0.12241)
Share with mother's income in 3rd quartile, last year	0.24868 (0.11217)	0.24620 (0.11233)	0.24966 *** (0.11196)	0.25029 *** (0.11155)	0.25059 *** (0.11126)
Share with mother's income in 4th quartile, last year	0.24867 (0.16192)	0.23476 (0.15830)	0.24781 *** (0.16014)	0.25239 *** (0.16249)	0.25737 *** (0.16505)
Taxable income quartile (Ref.cat.: 1st)					
Share with father's income in 1st quartile, last year	0.24378 (0.13402)	0.25086 (0.13733)	0.24248 *** (0.13192)	0.23791 *** (0.12989)	0.23136 *** (0.12737)
Share with father's income in 2nd quartile, last year	0.24378 (0.12524)	0.25072 (0.12579)	0.24552 *** (0.12460)	0.24311 *** (0.12544)	0.24156 *** (0.12625)
Share with father's income in 3rd quartile, last year	0.24378 (0.11063)	0.24330 (0.11111)	0.24472 *** (0.10962)	0.24583 *** (0.10982)	0.24787 *** (0.10977)
Share with father's income in 4th quartile, last year	0.24378 (0.16103)	0.22946 (0.15665)	0.24246 *** (0.15971)	0.24938 *** (0.16244)	0.25622 *** (0.16587)
<i>School characteristics</i>					
Turnover rate of Danish and math teachers	17.89773 (20.30256)	18.17307 (20.37762)	17.80744 *** (19.46363)	17.42887 *** (18.45994)	17.20917 *** (18.46575)
Turnover rate of all teachers	17.26428 (20.09400)	17.50427 (20.15308)	17.15226 *** (19.23596)	16.82285 *** (18.21994)	16.62136 *** (18.22736)
Size	514.40450 (210.44609)	505.90195 (211.07069)	512.27578 *** (211.34562)	515.40906 *** (212.74823)	513.35672 *** (214.51812)
# Observations	380.680	96.611	94.104	96.378	93.587

Notes: This table presents means and std. dev. across standardized quartiles of social well-being.

Tests represent two-sided t-tests compared to 1st quartile. * p < 0.1, ** p < 0.05, *** p < 0.01

Table A2.3

Persistence in social well-being: propensity to belong to 1st quartile

	1st quartile in 2019	
1st quartile in social well-being 2018	0.31813	***
	(0.00236)	
1st quartile in social well-being 2017	0.14764	***
	(0.00226)	

Notes: Table shows regression of indicator for belonging to 1st quartile of social well-being on lags of similar measure. Regressions also include child, parental, teacher and school, and classroom characteristics in addition to indicators for missing values as in Table 3. Robust standard errors shown in parentheses. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.01$. No. observations 247,987. R^2 -adjusted is 0.184.

Table A2.4
Selection into answering well-being survey

	Answering well-being survey
<i>Child characteristics</i>	
Male	-0.00680 *** (0.00111)
Birthweight is below 2500 grams	0.00639 *** (0.00241)
Age relative to cohort (Ref.cat.: Adequate for grade)	
Young for grade	-0.00440 (0.00457)
Late school start	-0.01649 *** (0.00207)
Repeated a grade	-0.01653 *** (0.00285)
Switched address, since school start	-0.00992 *** (0.00137)
Switched school, since school start	-0.00631 * (0.00336)
Switched school but not address , since school start	-0.02275 *** (0.00355)
Placed outside home, from 2010	0.03023 *** (0.00633)
Received social preventive measures by municipality, since birth	-0.05216 *** (0.00257)
Avg. no. of contacts per year to hospital, last 5 years	-0.02020 *** (0.00099)
Been in contact with psychiatric hospital, since birth	-0.08844 *** (0.00293)
Special needs student in ordinary classroom	-0.05689 *** (0.01168)
Special needs student in special needs classroom	0.10837 *** (0.01804)

<i>Maternal characteristics in 2018</i>		
Age at child birth	-0.00016 (0.00015)	
Education (Ref.cat.: Vocational)		
Unskilled	-0.00717 (0.00208)	***
Highschool	0.00126 (0.00266)	
Short further	0.00730 (0.00239)	***
Medium further	0.00323 (0.00140)	**
Long further or PhD	0.00035 (0.00207)	
Taxable income in 2nd quartile, last year	0.00839 (0.00169)	***
Taxable income in 3rd quartile, last year	0.00793 (0.00178)	***
Taxable income in 4th quartile, last year	0.00786 (0.00189)	***
Received unemployment assistance, last year	-0.01247 (0.00147)	***
Received unemployment insurance, last year	-0.00073 (0.00221)	
Avg. no. of contacts per year to hospital, last 5 years	0.00063 (0.00062)	
Been in contact with psychiatric hospital, since birth of pupil	-0.00019 (0.00205)	
Committed criminal offense, since birth of pupil	-0.01119 (0.00370)	***

<i>Paternal characteristics in 2018</i>		
Age at child birth	0.00012 (0.00012)	
Education (Ref.cat.: Vocational)		
Unskilled	-0.00445 (0.00177)	**
Highschool	-0.00320 (0.00260)	
Short further	0.00164 (0.00204)	
Medium further	-0.00226 (0.00167)	
Long further or PhD	-0.00131 (0.00199)	
Taxable income in 2nd quartile, last year	0.00423 (0.00165)	**
Taxable income in 3rd quartile, last year	0.00473 (0.00172)	***
Taxable income in 4th quartile, last year	0.00907 (0.00177)	***
Received unemployment assistance, last year	-0.00387 (0.00169)	**
Received unemployment insurance, last year	0.00267 (0.00253)	
Avg. no. of contacts per year to hospital, last 5 years	-0.00027 (0.00071)	
Been in contact with psychiatric hospital, since birth of pupil	-0.00295 (0.00239)	
Committed criminal offense, since birth of pupil	-0.01375 (0.00235)	***
<i>Household characteristics in 2018</i>		
Number of siblings	-0.00267 (0.00067)	***
Birth order on mother's side	-0.00112 (0.00082)	
Parents have lived apart for two consecutive years	0.00102 (0.00193)	
Single provider household	-0.01396 (0.00194)	***

<i>Teacher characteristics</i>		
Male, Danish	-0.02306 (0.00509)	***
Age, Danish	0.00003 (0.00021)	
Educated in teaching Danish or comparable	0.00718 (0.00923)	
Years since degree, 0-10 years, Danish	0.00058 (0.00125)	
10 years or above since degree, Danish	0.00277 (0.00863)	
Tenure of teacher, Danish(Ref.cat.: 0-1 year)		
Tenure is 2 to 3 years, Danish	-0.00609 (0.00550)	
Tenure is 4 to 5 years, Danish	-0.00868 (0.00540)	
Tenure is 6 years or above, Danish	-0.00778 (0.00523)	
Share of days absent, Danish	-0.01928 (0.01962)	
Age, Math	0.00055 (0.00022)	**
Male, Math	-0.01369 (0.00348)	***
Educated in teaching math or comparable	-0.00600 (0.00569)	
Years since degree, 0-10 years, Math	-0.00175 (0.00130)	
10 years or above since degree, Math	-0.01377 (0.00870)	
Tenure of teacher, Math(Ref.cat.: 0-1 year)		
Tenure is 2 to 3 years, Math	-0.00871 (0.00531)	
Tenure is 4 to 5 years, Math	-0.01302 (0.00541)	**
Tenure is 6 years or above, Math	-0.01512 (0.00530)	***
Share of days absent, Math	-0.04692 (0.01945)	**

<i>Classroom characteristics</i>		
Size	0.00277 (0.00046)	***
Share of males	-0.01028 (0.01442)	
Share with birthweight below 2500 grams	-0.01649 (0.02637)	
Share of who are young for grade	-0.08195 (0.05442)	
Share of late starters	-0.20064 (0.02138)	***
Share who repeated a grade	-0.00846 (0.02284)	
Share who switched address, since school start	-0.14280 (0.01284)	***
Share who switched school, since school start	0.16316 (0.03828)	***
Share who switched school but not address, since school start	-0.32187 (0.04020)	***
Share who have been placed outside home, from 2010	-0.14482 (0.04935)	***
Share who received social preventive measures by municipality, since birth	0.06252 (0.02143)	***
Avg. no. of contacts per year to hospital, last 5 years	-0.05412 (0.00945)	***
Share who has been in contact with psychiatric hospital, since birth	-0.08378 (0.02295)	***
Share with special needs in ordinary classrooms	0.24491 (0.08174)	***

Share with a criminal parent	-0.01765 (0.01690)	
Share with parents have med./long education	0.02488 (0.01194)	**
Share with parents who received unemployment assistance/insurance, last year	0.01762 (0.01470)	
Share with a parent who has been in contact with psychiatric hospital, since bir	0.00023 (0.01808)	
Share with mother's income in 2nd quartile, last year	-0.01113 (0.01819)	
Share with mother's income in 3rd quartile, last year	-0.07759 (0.01883)	***
Share with mother's income in 4th quartile, last year	-0.11253 (0.02052)	***
Share with father's income in 2nd quartile, last year	0.04846 (0.01779)	***
Share with father's income in 3rd quartile, last year	0.03660 (0.01855)	**
Share with father's income in 4th quartile, last year	0.02667 (0.01828)	
<i>School characteristics</i>		
Turnover rate of Danish and math teachers	0.00021 (0.00035)	
Turnover rate of all teachers	-0.00049 (0.00035)	
Size	-0.00003 (0.00001)	***

Notes: Table shows regression of indicator for participating in the social well-being survey on child, parental, teacher and school, and classroom characteristics in addition to indicators for missing values as in Table 3. Robust standard errors shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. No. observations 449,971. R^2 -adjusted is 0.065.

A3 Appendix Chapter 3

A3.1 - Control variables used in Section 3.5

Scores in the 6th grade National Exams of Portuguese and Mathematics:	discrete variables on a scale of 1-5, where 3-5 are passing scores.
Male:	binary variable taking the value 1 if the student's sex is male, 0 if female.
Parental Education:	dummy variable categories for the highest level of education by either of the student's parents. The categories are: lower than Upper-Secondary School (used as the reference group), at most Upper-Secondary School, Tertiary education.
Migrant status:	dummy variable categories for the migrant status of a students. The categories are: native students (used as the reference group), 1 st generation immigrant, and 2 nd generation immigrant.
SASE support:	dummy variable categories indicating the type of social support (SASE – <i>Serviço de Ação Social Escolar</i>), given in case of low family income. The categories are: no support (used as the reference group), SASE B, and SASE A (the category for the lowest family income).
Parental employment status:	indicators for the employment status of a student's father and mother, taking the value 1 if unemployed.
Resources at home:	indicator for whether the student has a computer at home, and another indicator for whether the student has internet connection at home.
Years above reference age:	dummy variable indicating whether the student is above the reference age in the 7 th grade. The reference age at the end of the 7th grade — i.e., under a regular school starting age and with no accumulated retentions — is 13 years old, if a student's birthday is between the 1st of January and the 31st of August, or 12 years old if the birthday is after that date.

A3.2 - Machine Learning Prediction of students at Retention Margin in 2006/07

The identification strategy used in this paper compares retained with promoted students at the 7th grade retention margin across years, as retention in this grade in 2008/09 lead to exposure to the new Compulsory Schooling Leaving age. To add credibility to the parallel-trends assumption that is needed for these Difference-in-Differences estimates, I only compare students at the grade-retention margin, identified by having between 3 and 5 failing scores out of the 10 school subjects taken in this grade. To test the plausibility of this identification strategy we would like to verify that pre-treatment trends are parallel — that is, if from 2006/07 to 2007/08 the outcomes of the retained group and the promoted group evolved in a parallel way, when both groups remained untreated.

However, because data on school subject scores, used to define the analysis sample at the retention margin, are not available in the first year of the dataset, 2006/07, this is not directly possible. Nevertheless, it is possible to use Machine Learning classification methods to identify who the students that were likely to be at the retention margin in that academic year were, based on all observed covariates and subsequent academic path.

The procedure used to identify who the students that were likely to be at the retention margin in 2006/07 is the following:

1. Construct classifications models using all available variables in Appendix A3.1, as well as variables for students' educational paths: e.g., school track choices, dropout and its timing, being at the grade retention margin in $t+1$ and $t+2$, and grade-level enrollment in $t + 1$ and $t + 2$.

2. Train the different classification models with data from the 2008/09 to 2010/11 7th grade cohorts. The classification task consists in identifying from the full population of 7th grade students, who the students at the grade-retention margin were. This is an imbalanced classification task as the *at the grade-retention margin* group is around 9% of the population.
3. An *Extreme Gradient Boosting (XGBoost)* model is chosen for its overall better performance with standard parameters against other models such as a *Random Forest*.
4. A standard Grid Search is performed to tune the model's hyper-parameters and optimize its performance. Performance metrics such as Sensitivity, Specificity and Accuracy are around 80 to 85%, and Area Under the Curve (AUC-ROC), a commonly used measure to evaluate the performance of imbalanced classification tasks, is 91%. This indicates that the *XGBoost* model has good performance.
5. In step 2, the 2007/08 data is purposely left-out of the training step to test the predictions of the model chosen in the previous step. I choose the probability cutoff $p > c$ that most closely matches the Compliance and Graduation outcomes of retained and promoted students observed in this year, seen in Figures 3.12 and 3.13. The cutoff c is then set at 0.76. This means that students with a probability of being at the grade-retention margin greater than 0.76 are classified as *at the grade-retention margin*, whereas those with probability lower than that threshold are classified as *not at the grade-retention margin*.
6. I then construct a new *XGBoost* model with data from all academic years, i.e., from 2007/08 to 2010/11, with the same hyper-parameters found in step 4.

7. Finally, I get predictions on the 2006/07 data using the model of the previous step, and use the same probability cutoff c as in step 5 to classify students as being *at the grade- retention margin*. The missing outcomes for 2006/07 are then imputed in Figures 3.12 and 3.13.