

Stockholm School of Economics
Department of Economics
5350 Master's Thesis in Economics
Spring 2025

Remote Instruction and Failure Rates in Swedish Schools during the COVID-19 Pandemic

Andres Cruz (25199) & Klara Holmer (25037)

This thesis examines whether remote instruction during the COVID-19 pandemic affected students' academic achievements in Sweden. Through a Difference-in-Differences design, we study the effects of remote instruction on failure rates in Swedish, English and Mathematics by comparing grade 10 students, who were largely subject to remote instruction during the pandemic, to grade 9 students, whose education largely remained in-person. To improve the comparability of grade 9 and grade 10 students, we match students based on observed shared characteristics through Propensity Score Matching. The Difference-in-Differences results for English are inconclusive due to indicated violations of the parallel trends assumption. However, for subjects Swedish and Mathematics, the use of remote instruction did not affect failure rates, which contrasts much of the previous research on this topic. This thesis thereby adds to the existing literature by providing evidence that remote instruction during the pandemic may not always have a negative effect on students' learning. Since this result differs from previous research on this topic, this thesis also highlights the need for continued research in this area.

Keywords: Difference-in-Differences, Propensity Score Matching, COVID-19 Pandemic, Education, Remote Instruction

JEL: I21, I28, I18

Supervisor:	Ulrika Ahrsjö
Date submitted:	May 13th, 2025
Date examined:	May 26th, 2025
Discussants:	Clara Falkenek & Herman Johansson
Examiner:	Elena Paltseva

Acknowledgements

We would like to thank our supervisor Ulrika Ahrsjö for her guidance, support, and engagement throughout this entire process. She has provided us with valuable, concrete feedback, which has positively contributed to this thesis, and we are deeply grateful for all her help. We would also like to extend our thanks to Jaakko Meriläinen, Wilma Liljequist, Zain Awad, Oscar Nilsson Hallgren, Zoltan Racz, Clara Falkenek, Herman Johansson and Elena Paltseva for their time and valuable comments, which has improved the quality of this thesis.

Contents

1	Introduction	1
2	Literature Review	3
2.1	The Pandemic's Effects on Swedish Students	3
2.2	The Pandemic's Effects on Learning	3
2.3	Remote Instruction's Effects on Learning	4
2.4	Remote Instruction's Effects on Learning in the Pandemic	5
2.5	Research Gap	7
3	Background	9
3.1	The Swedish School System	9
3.2	Swedish Schools During the Pandemic	10
4	Data	13
5	Methodology	16
5.1	Empirical Strategy	16
5.2	Propensity Score Matching	17
5.3	Model Specification	18
5.4	Identifying Assumptions	19
6	Results	21
6.1	Subject Swedish	21
6.2	Subject English	23
6.3	Subject Mathematics	24
7	Sensitivity Analyses	26
7.1	Heterogeneity Analyses: Subjects Swedish and Mathematics	26
7.2	Bounding Relative Magnitudes: Subject Mathematics	28
7.3	Robustness Check: Subjects Swedish, English and Mathematics	31
8	Discussion	33
8.1	Interpretation of Results	33

8.1.1	Pre-Pandemic Trends	33
8.1.2	Results for Subject Swedish	33
8.1.3	Potential Causes	34
8.1.4	Results for Subjects English and Mathematics	35
8.2	Implications for Academia and Policy	36
8.3	Limitations	37
8.3.1	Generalizability of Results	37
8.3.2	Failure Rates as a Measure of Learning	37
8.3.3	Attendance During the Pandemic	39
8.3.4	Infrequent Data Collection and Short Exposure to Treatment	39
9	Conclusion	41
	References	42
	Appendix	50
A	Data Sources and Sample	50
B	Covariate Balance After Propensity Score Matching	53
C	List of Tables for the Results Section	54
D	Summary Statistics and Results Prior to Propensity Score Matching	57
E	Heterogeneity Analysis: Subject English	59
	Description of AI Usage	60

List of Tables

1	Remote Instruction During the Pandemic	12
2	Summary Statistics by Educational Level: Pre- and Post-Pandemic	15
3	Heterogeneous Effects, Subject Swedish	27
4	Heterogeneous Effects, Subject Mathematics	28
5	Dynamic Difference-in-Differences Models by Subject	32
6	Required Data Files	50
7	Number of Observations and Share of Missing Values by School Size Group . .	51
8	Number of Unique Schools Across Sample Restrictions	52
9	Covariate Imbalance Before and After Matching (Excluding Exact-Match Variables)	53
10	Dynamic Difference-in-Differences Models, Subject Swedish	54
11	Dynamic Difference-in-Differences Models, Subject English	55
12	Dynamic Difference-in-Differences Models, Subject Mathematics	56
13	Summary Statistics by Educational Level: Pre- and Post-Pandemic	57
14	Heterogeneous Effects, Subject English	59

List of Figures

1	Time Trend of Average Grade Points	10
2	Timeline of Recommendations During the Pandemic	11
3	Trends in Failure Rates in Subject Swedish	22
4	Dynamic Difference-in-Differences of Failure Rates in Subject Swedish	22
5	Trends of Student Failure Rates in Subject English	23
6	Dynamic Difference-in-Differences of Failure Rates in Subject English.	24
7	Trends of Student Failure Rates in Subject Mathematics.	25
8	Dynamic Difference-in-Differences of Failure Rates in Subject Mathematics.	25
9	Bounded Relative Magnitudes	31
10	Propensity Score Distributions Before and After Matching	53
11	Results for Subject Swedish Prior to PSM	57
12	Results for Subject English Prior to PSM	58
13	Results for Subject Mathematics Prior to PSM	58

1 Introduction

The COVID-19 pandemic was an unprecedented health, economic and social crisis (United Nations, 2024). In 2020, the pandemic claimed the lives of 4.1 million people globally, causing it to be among the top three leading causes of death that year, and the following year, the death count more than doubled. In two years, over a decade of progress in life expectancy and healthy life expectancy was undone (World Health Organisation, 2024).

The United Nations' Sustainable Development Goals Report warned of a global reading crisis prior to the pandemic (United Nations, 2019). Progress in education was not on track to reach Goal 4, which aims to achieve “inclusive and equitable quality education and promote lifelong learning opportunities for all” (United Nations, n.d.). In 2020, progress in all Sustainable Development Goals had been affected by the pandemic. For education, progress had not only slowed, but regressed, removing around 20 years of progress in reading proficiency and Mathematics (United Nations, 2020; United Nations, n.d.). At its peak, 1.6 billion students were subject to school closures and many countries had implemented remote instruction (UNESCO, 2021; World Bank Group, 2025). While many countries closed all schools, this was not the case in Sweden (Our World in Data, 2020; Folkhälsomyndigheten, 2024).

During the pandemic, many countries used lockdowns or other regulatory interventions, but Sweden used an alternative approach, which was based on recommendations and personal responsibility (Regeringen, 2021). Similarly, Sweden differed in their approach to education during the pandemic. The Public Health Agency in Sweden wanted to avoid school closures and minimise the use of remote education to avoid negative effects on students' health. Still, during periods with increased infection rates, the agency recommended certain levels of education to move to online instruction (Folkhälsomyndigheten, 2023).

Considering the large number of students that were affected by physical school closures and a move to remote instruction during the pandemic, it is highly important to study the impacts of this change on students' learning outcomes. This would also be informative for policy decisions on potential support actions for affected students and for education decisions in future crises. Examining specifically failure rates is important, as previous literature indicates that lower achieving students may be especially affected by this change (Bettinger et al. 2017; Heppen et al., 2016; Werner and Woessmann, 2023). In a Swedish context, it is particularly valuable to study failure rates in Swedish, English and Mathematics, as they influence the eligibility for continued studies. Thus, this thesis aims to answer the following research question:

Did remote instruction during the COVID-19 pandemic impact failure rates in Swedish, English, and/or Mathematics in Sweden?

We construct a yearly school-level panel dataset for upper secondary grade 10 students enrolled in academic programs and grade 9 students in lower secondary school for the academic years 2014/15 to 2020/21, which includes the failure rates in each subject as well as school characteristics.

To estimate the effect on failure rates of the transition to remote instruction we use a Difference-in-Differences design, comparing grade 10 students, who to a large extent transitioned to remote instruction during the academic year 2019/20 and 2020/21, with grade 9 students, whose education to a higher degree remained in-person. Specifically, we specify a dynamic Difference-in-Differences model to check the pre-trend dynamics and see how the treatment effect changes in the academic years following the recommendations of remote instruction. Prior to the estimation, we employ Propensity Score Matching to account for compositional differences between untreated lower secondary schools and treated upper secondary schools.

The results suggest that remote instruction did not impact failures rates in the subjects Swedish and Mathematics, while for the subject English the results are inconclusive given that the parallel trends assumption appears to be violated. Our results indicate that in a crisis, remote instruction may be a suitable alternative to in-person instruction in Sweden. Moreover, this thesis sheds light on a research gap, since there appears to be differences between countries in how students' learning is affected by remote instruction.

The remainder of this thesis is organized as follows. Section 2 reviews the relevant literature on the pandemic's effects on Swedish students, the pandemic's effects on learning and remote instruction's effects on learning prior to and during the pandemic. Section 3 provides background on the Swedish school system and the Swedish approach to remote instruction during the COVID-19 pandemic. Section 4 describes the data sources, sample construction, and variable definitions used in the analysis. Section 5 details the empirical strategy, including the Difference-in-Differences approach, the Propensity Score Matching methodology, and our model specification. Section 6 presents the empirical results. Section 7 presents a set of heterogeneity analyses and sensitivity analyses. Section 8 discusses the results and broader implications of the findings, addresses key limitations, and suggests areas for future research. Finally, Section 9 concludes.

2 Literature Review

This literature review begins by describing research on how Swedish students have been affected by the pandemic in ways other than learning. Thereafter, we review the existing literature on the effects of the pandemic on learning. This is followed by research on the effects of remote instruction on learning in a non-pandemic setting. Finally, we cover research on how the use of remote instruction during the pandemic has affected learning. This literature review focuses on studies from high and middle income countries as they are likely to provide the most relevant insights for a Swedish context.

2.1 The Pandemic's Effects on Swedish Students

Through a Difference-in-Differences approach, Björkegren, Svaleryd and Vlachos (2024) isolate the effect of remote instruction on Swedish students' mental health, by exploiting an exogenous variation in exposure to remote instruction, where upper secondary students received remote instruction while lower secondary schools stayed open. They found that moving to online learning decreased use of mental health care, especially for depression and anxiety. While the underlying causes remain unclear, their findings indicate that the school environment negatively affects students' mental health. Potential reasons behind this that they suggest are reduced stress, increased flexibility, reduced social pressures or a perception of more lenient academic standards. Utilising the same exogenous variation in remote instruction, Vlachos, Hertegård and Svaleryd (2021) found that school closures in Sweden did not appear to have a large effect on the spread of COVID-19 in general, but that the infection rates of teachers doubled. Aalto, Müller and Tilley (2023) study the effect of the pandemic on choices between upper secondary programs in Sweden. Their results indicate that labor market conditions affect students' choices of school programs. Specifically, the pandemic made students less prone to apply to service-related vocational programs, especially programs in the accommodation and food services industry. However, there was no change in the interest in academic programs.

2.2 The Pandemic's Effects on Learning

A meta-analysis by Betthäuser, Bach-Mortensen and Engzell (2023) on the effect of the COVID-19 pandemic on learning outcomes found "a substantial overall learning deficit [...], which arose

early in the pandemic and persists over time” (p. 375). The meta-analysis was published in January 2023, so the long term effects were measured at most three years from the beginning of the pandemic. The effect on learning was larger for students from lower socio-economic backgrounds, and results in Mathematics were more negatively affected by the pandemic compared to reading.

A working paper by Hall, Hardoy and Lundin (2022) describes that the Nordic countries were relatively well prepared for remote instruction compared to other countries. They also describe that the pandemic likely had long term negative effects on Nordic students’ learning and that certain groups were more negatively affected than others, such as students that were younger, had special needs or had poor skills in the country’s main language.

In a Swedish setting, one study examines whether the COVID-19 pandemic affected reading skills of primary school students in grades 1-3 (Hallin et al., 2022). Contrary to much of the previous literature, they found no effect on students’ reading skills. Furthermore, students from low socioeconomic backgrounds were not found to be more affected by the pandemic than others. One crucial difference between this study and other papers on children’s learning during the pandemic is that since it studied Swedish students in grades 1-3, it studied students whose schools remained open. They therefore examine the effect of the pandemic excluding physical school closures and a switch to remote instruction on students’ academic results. Since they do not find a negative effect on learning, while other studies have, this could indicate that the pandemic mainly affected students through the removal of in-person instruction, or that students were less affected by the pandemic due to fewer restrictive measures such as lockdowns. However, the lack of studies on how Swedish students were affected by the pandemic makes it unclear whether the contrasting results in the study by Hallin et al. (2022) were due to schools remaining open or due to the Swedish setting.

Research is currently being conducted on relationships between mental health, school grades and school support in Sweden (Giota et al., 2024). Part of the research aim of Giota et al. is to examine the effects of the COVID-19 pandemic on students’ mental health and academic achievement. Since their research paper is not finalised, it does not yet provide useful findings for this literature review, but it further illustrates a research gap concerning the effects of the pandemic on Swedish students. They describe that this area requires immediate attention, both to provide support for the affected students and to produce research that can aid policy.

2.3 Remote Instruction’s Effects on Learning

Existing studies on the effectiveness of online courses indicate that substituting in-person courses with courses that are exclusively online have a negative effect on learning (Escueta et al., 2020). The effects also appear to differ between student groups, with lower achieving students being more negatively affected by online instruction (Figlio, Rush and Yin, 2013; Bettinger et al., 2017; Heppen et al., 2016). In a study by Bettinger et al. (2017), students that

attended class online did not only receive lower grades in said class, but also in subsequent courses where teaching was in-person. However, Heppen et al. (2016) found no significant difference in results between lower and average achieving students one year later.

In their review article, Escueta et al. (2020) describe a lack of rigorous studies on the effectiveness of online courses, despite their increased popularity. Additionally, most studies on this topic focuses on post-secondary students, and the existing studies on primary and secondary school students lack a causal interpretation, as few of them have used an experimental or quasi-experimental approach. A majority of the studies were also conducted in the US (Escueta et al., 2020).

2.4 Remote Instruction's Effects on Learning in the Pandemic

Research on school closures in the pandemic generally indicates a negative effect on learning. A meta-analysis from the Joint Research Centre of the European Commission (2023) examines OECD countries and finds that on average, school closures during the pandemic had a negative effect on learning. Contrary to previous meta-analyses, they find that the negative effect increased at later stages of the pandemic. This was due to the inclusion of newer studies that the previous analyses had not had access to (European Commission: Joint Research Centre and Di Pietro, 2023).

Agostinelli et al. (2022) investigated how high school students in the US were affected by school closures during the COVID-19 pandemic through the use of a structural model of skill acquisition. They analyze the effect of school closures through three different channels: schools, peers and parents, as these three channels were connected to previous empirical evidence. The results of the study indicates that school closures during the pandemic caused a higher educational inequality through each of the three channels, with the largest effect coming from the peer effects channel. Fuchs-Schündeln et al. (2022) have calculated long term effects of school closures during the pandemic using a structural life cycle model. On average, they find a 2.1% loss in lifetime earnings for the affected children and a 1.2% loss in welfare. The effect of school closures is also found to be larger for children whose parents have low educational levels and low assets. However, in their main analysis, they equate school closures to a complete loss in schooling. At a global level, learning losses estimated using PISA test scores were associated with 17 trillion dollars of economic loss (Jakubowski, Gajderowicz and Patrinos, 2025). Countries with longer closures were associated with larger losses than those that had shorter closures, and the effects were worse for boys and students with immigrant backgrounds.

Through a Difference-in-Differences approach, Halloran et al., (2021) found that US districts with less in-person instruction had a larger decline in passing rates compared to schools with more in-person instruction. The difference in effect was larger for Mathematics compared to English. Engzell, Frey and Verhagen (2021) use the exogenous change of an 8 week lock-

down in the Netherlands to estimate the effect on students' learning, by creating a benchmark from previous years' results and comparing those to the results in 2020 in a Difference-in-Differences model. The lockdown caused a learning loss equal to what the learning loss would be from no schooling during the same time period. The authors thereby note that "the findings imply that students made little or no progress whilst learning from home". Students whose parents had lower education levels were also more negatively affected by the lockdown, with an up to 60% higher learning loss.

In Switzerland, researchers found that remote instruction caused a substantial reduction in time spent studying; from 35 hours per week to 23 hours per week. However, they found no difference in relative reduction between genders or between students whose parents had different education levels (Grätz and Lipps, 2020). Werner and Woessmann (2023) used panel data from time use surveys to analyse how children in Germany were affected by school closures. They found that the students' studying time approximately halved during the first school closures and had not returned to previous levels during the second wave of school closures. Time spent on activities like reading, exercising and being creative only increased with 6 minutes per day during the first closures, while time spent on activities such as using social media, TV or video games increased with 1.3 hours per day. This indicates that some of the time previously spent studying was substituted for activities that are less suitable for learning. Low achieving students replaced studying time with these less suitable activities for learning to a larger extent than other students. Parents also viewed the time their children spent studying as being less efficient during the physical school closures, especially parents with a lower level of education and parents to lower achieving students.

In a Nordic setting, Skar, Graham and Huebner (2021) used school closures in Norway as a natural experiment and compared writing skills of cohorts of first grade students that were taught during the pandemic with those taught one year before the pandemic. They found negative effects of the school closures on students' writing ability and attitude toward writing. However, in a follow-up replication study (Skar, Graham and Huebner, 2023), these negative effects were no longer present one year later. A Danish study (Birkelund and Karlson, 2021) on the effect of school closures on learning outcomes found a negative effect for older students, but a positive effect for younger ones. These results contrast those of other studies, which indicate that the younger students should not only have had a negative effect from remote instruction, but a larger negative effect than the older students (Hall, Hardoy and Lundin, 2022). They also found little evidence of an increased difference in learning outcomes based on family background, which also differs from previous studies (e.g. Betthäuser, Bach-Mortensen and Engzell, 2023; Engzell, Frey and Verhagen, 2021). In a Nordic setting, the existing research is scarce and presents results that contrast those from other countries. It is therefore unclear in a Nordic context whether transitioning to remote instruction during the pandemic had a negative effect on students, whether it was lasting and whether it increased the learning gap based on family background.

A report by Svaleryd and Vlachos (2021) found that average grades in Sweden during the first year of the pandemic followed similar trends as previous years for both upper and lower secondary students. Since the upper secondary students had received online instruction, this could indicate that remote instruction did not negatively affect Swedish students' learning, but since they did not test this causally, we can not draw strong conclusions. Qualitative research on 153 teachers in Sweden found that the majority of teachers considered the transition to online education to have been a large change, but that it worked surprisingly well (Bergdahl and Nouri, 2020). Teachers were surprised by students' motivation and engagement, and also described that the students often focused better during remote instruction, because of fewer distractions from their classmates. They also noted that some at-risk students had higher participation rates during remote instruction. Teachers described that this new mode of instruction appeared to fit these students better, while certain other students that had previously performed well in school showed difficulties adapting to remote instruction.

2.5 Research Gap

The three research areas on learning that have been examined in this literature review – research on remote instruction, the pandemic and remote instruction during the pandemic – all describe negative effects on learning. These effects are also estimated to have large consequences for society as a whole. In Sweden however, the existing research has not found that the pandemic had a negative effect on learning, although it is unclear whether this could be due to researchers examining a setting where schools remained open. Qualitative research on the school closures in Sweden found that teachers considered remote instruction to work surprisingly well and that students focused better due to fewer distractions from classmates. This could indicate that school closures may not have had a negative effect in Sweden, and potentially even had a positive effect, but strong conclusions should be cautioned against due to the study's small sample size. Moreover, research in the Nordics add to this ambiguity, as research in Norway found negative effects of remote instruction, but that these disappeared shortly after, and research in Denmark found negative effects for some students, but positive for others.

Furthermore, each of the three research areas on learning indicate that some subgroups of students appear to have been more affected by the pandemic than others. Remote instruction, both in and outside a pandemic, appears to affect lower achieving students more negatively than others. Research describes a negative effect on lower achieving students' academic results, a larger increase in their time spent on activities that are less suitable for learning and, in their parents' view, a less efficient use of studying time. In Sweden, the effects of school closures during the pandemic on lower achieving students' learning has not yet been researched through quantitative analysis. The existing qualitative research indicates that lower achieving students may fare better under remote instruction, but also that previously higher achieving students may instead reduce their learning. As this conclusion differs from other papers on this topic, and

current research in a Swedish setting is scarce, this indicates a research gap worth examining.

Our approach utilises methodologies from previous research papers studying related topics in Sweden, while adding components from other areas of empirical research. Following Björkegren, Svaleryd, and Vlachos (2024), we conduct a Difference-in-Differences design where we compare Swedish upper secondary students, most of whom were exposed to remote instruction during the pandemic, to lower secondary students, who instead largely had in-person instruction. To improve comparability of the two groups, we focus exclusively on students in the last year of lower secondary school and students in the first year of upper secondary school, in line with Vlachos, Hertegård and Svaleryd (2021). Moreover, to account for remaining differences between these groups of students, we implement Propensity Score Matching, a method commonly used in empirical research (Caliendo and Kopeinig, 2005).

3 Background

3.1 The Swedish School System

The Swedish school system consists of ten years of compulsory schooling, which are divided into preschool, primary school and lower secondary school. After lower secondary school, students may choose to attend upper secondary school, which usually lasts three years. Out of the national programs for Swedish upper secondary school, 67% of the students attend academic programs and 33% attend vocational programs. Following upper secondary school, students may choose to continue with higher education (University Admissions, 2022; Skolverket and Sveriges Officiella Statistik, 2024; Skolverket, 2025a). In this paper, the last year of lower secondary school and the first year of upper secondary school will be referred to as grade 9 and grade 10, respectively.

For both lower and upper secondary students in Sweden, the grade scale goes from A to F, with A being the highest, E the lowest passing grade and F a failing grade. In lower secondary school, students receive grades from year 6, or in some cases from year 4, but only receive their final grade in year 9. In upper secondary school, students instead receive final grades in all courses attended over the three years, rather than only receiving final grades in the final year (Skolverket, n.d.; Skolverket, 2025b; Skolverket, 2024b). To qualify for upper secondary school, lower secondary students must have a passing grade in Swedish or Swedish as a Second Language, Mathematics, English and five other courses. Similarly, to qualify for colleges or universities, upper secondary students must have a passing grade in Swedish or Swedish as a Second Language, Mathematics, English and in a certain number of additional courses (Skolverket, 2025a; Antagning, 2022; Skolverket, 2025c). Since 1992, students in Sweden have been choosing which school to attend. Both lower and upper secondary schools can be privately or publicly owned, and are open to all students (Edmark, Frölich and Wondratschek, 2014; Skolverket, n.d.; Skolverket, 2025b). If demand exceeds supply for an upper secondary school, then the grades set in grade 9 determine which students will be admitted to the school. For colleges and universities, admission is often based on grades set in grades 10 to 12 in upper secondary school (Skolverket, 2025a; Universitets- och högskolerådet, 2025).

Average grade points in grade 9 and grade 10 have remained relatively stable over the past ten years, although there have been a gradual increase in grades up until 2020/21, followed by a gradual decrease, see Figure 1. In international tests such as TIMSS, Swedish lower

secondary students have improved their results since 2011 in Mathematics and Natural Sciences, but differences in results between students have gradually increased during the same time period (Skolverket, 2024c). Similarly, when average grades of upper secondary students gradually increased over time, the share of students that meet the requirements of an upper secondary degree saw a slight decrease (Skolverket, 2022a). This indicates that learning outcomes of the average upper and lower secondary student may have improved over time, but that those of lower achieving students may not have. A report by Holmlund, Sjögren and Öckert (2020) describes that differences in results between Swedish schools has been increasing and that the reason behind it is school segregation; students in Sweden are increasingly attending schools where students share a similar background. Swedish schools have become more socially and ethnically segregated over the past thirty years, largely due to increased housing segregation, but this is also affected by the ability for students to choose which school to attend.

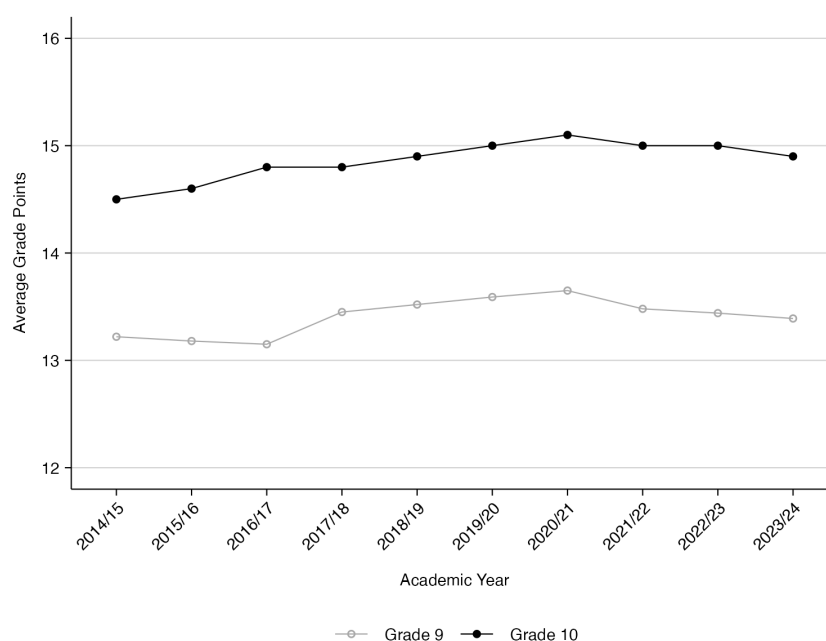


Figure 1 – Time Trend of Average Grade Points

Notes: The figure shows the time trends in average grade points by level of education for the academic years 2014/15 to 2023/2024. For grade 10, the values correspond to academic programs. Read the last paragraph in Appendix A for a walk-through on how find the data used for this graph.

3.2 Swedish Schools During the Pandemic

The Swedish approach to the pandemic was largely based on recommendations (Regeringen, 2021). For schools, the recommendations differed between different educational stages, see Figure 2. This difference in recommendations between upper and lower secondary schools appears to have been due to the age difference between the two groups. Swedish authorities aimed to keep children in focus and preferred interventions that affected adults rather than children

(European Centre for Disease Prevention and Control and Folkhälsomyndigheten, 2023). They also reasoned that children are less affected by the virus than adults and may be less contagious due to milder symptoms (Folkhälsomyndigheten and Institutet för hälsa och välfärd, 2020).

On March 17 in 2020, the Public Health Agency of Sweden announced the first recommendation that upper secondary education should be conducted through online instruction. This recommendation was withdrawn from June 15 in 2020, and upper secondary schools could return to in-person instruction at the start of the fall semester of 2020. However, on December 3 in 2020, the agency once again recommended upper secondary schools to switch to online instruction. On January 20 in 2021, upper secondary schools were recommended to gradually return to in-person instruction, with between 20-80% of remote instruction. From April 1 of 2021, the recommendation of remote instruction was lifted. Lower secondary schools received no recommendation of remote instruction for the spring or fall of 2020. On January 8 in 2021, the Swedish government decided that owners of lower secondary schools would be allowed to implement remote instruction in their schools if they considered it to be necessary for reducing the spread of COVID-19. In August of 2021, lower secondary schools were to return in-person instruction (Folkhälsomyndigheten, 2024).

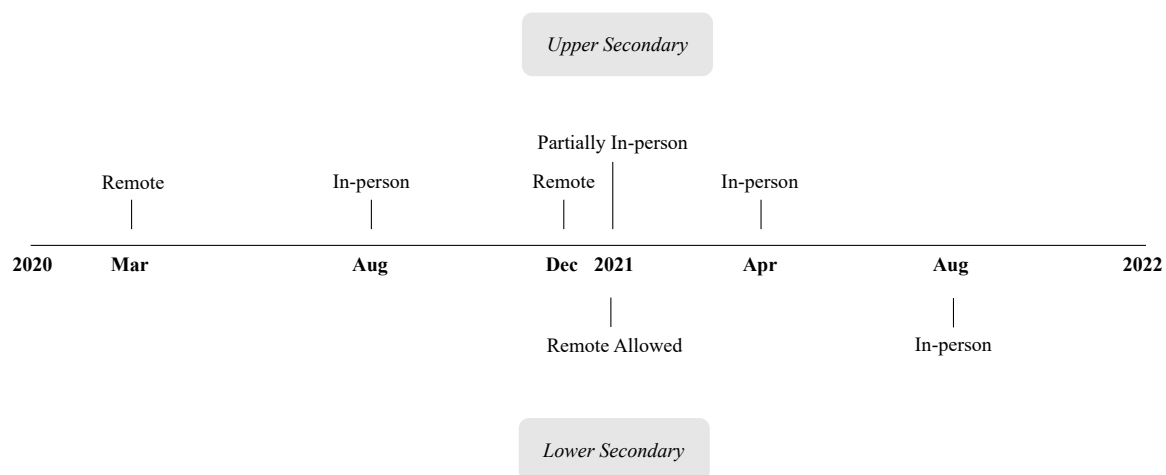


Figure 2 – Timeline of Recommendations During the Pandemic

Notes: This figure shows the timeline of recommendations for Swedish upper and lower secondary schools during the pandemic. Text above the line refers to recommendations for upper secondary schools and text below the line refers to recommendations for lower secondary schools. The information depicted in the timeline is from Folkhälsomyndigheten (2024).

To study whether schools followed these recommendations, we use principal survey data from Statistics Sweden (Statistikmyndigheten SCB, 2022) on the use of remote instruction, see Table 1. The results show that Swedish schools largely followed the recommendations

given by the Public Health Agency of Sweden. 84% of upper secondary students received remote instruction during the spring of 2020 and the median of 12 weeks approximately equals the remaining weeks of the school semester from the date of the recommendation announcement. Meanwhile, lower secondary education generally remained in-person during the spring of 2020, with 7% of students receiving remote instruction for a median of 2.5 weeks. In the fall of 2020, most upper secondary students received remote instruction, but only for a median of 3 weeks, which again aligns with the time that remained of the school semester at the time of the announcement to reintroduce remote instruction. Similarly to spring of 2020, lower secondary students generally received in-person teaching, with one third of them receiving remote instruction, but for a median of only one week. During the spring of 2021, the exposure to remote instruction was similar for lower and upper secondary students. Most of these students received some remote instruction and the median number of weeks was similar for both groups. This shows that upper secondary schools followed the recommendation of 20-80% of remote instruction and that many owners of lower secondary schools decided to use some remote instruction after being presented with the option.

Table 1 – Remote Instruction During the Pandemic

	Spring 2020	Fall 2020	Spring 2021
Lower Secondary Students	7 (2.5)	33 (1.0)	79 (5.0)
Grade 7 Students	8 (2.5)	33 (1.0)	80 (5.4)
Grade 8 Students	7 (3.0)	34 (1.0)	81 (5.4)
Grade 9 Students	6 (2.0)	34 (1.0)	76 (3.5)
Upper Secondary Students	84 (12.0)	75 (3.0)	69 (7.0)
Grade 10 Students	89 (12.0)	81 (3.0)	78 (8.0)
Grade 11 Students	88 (12.0)	81 (3.4)	75 (8.0)
Grade 12 Students	84 (11.0)	77 (3.0)	69 (6.0)
Academic Programs	94 (12.0)	84 (3.0)	81 (8.0)
Vocational Programs	74 (11.0)	70 (3.0)	60 (6.0)
Preparatory Programs	59 (11.0)	38 (2.0)	33 (5.0)

Notes: This table is inspired by Björkegren, Svaleryd, and Vlachos (2024) and uses data from Statistics Sweden (Statistikmyndigheten SCB, 2022). The left number in each column shows the percentage of students in remote instruction and the number in parentheses shows the median number of weeks of remote instruction.

These results also show differences in exposure to remote instruction for different subsets of students. Among upper secondary students, grade 10 students were those with the most exposure to remote instruction, and among lower secondary students, grade 9 students were those with the least. Furthermore, academic programs received more remote instruction than the other upper secondary programs.

4 Data

Similarly to previous studies on the effect of the COVID-19 pandemic on Swedish students, such as the paper by Aalto, Müller, and Tilley (2023), this thesis uses yearly school-level data, which was collected by Statistics Sweden (SCB) and publicly published by the Swedish National Agency for Education (Skolverket, 2025d).

We constructed a panel dataset for grade 10 students enrolled in academic programs in upper secondary schools and grade 9 students in lower secondary schools. The panel dataset includes statistics on failure rates in subjects Swedish, Mathematics, and English for graduating students, as well as school characteristics for the academic years 2014/15 to 2020/21. Furthermore, we excluded schools with fewer than 50 students due to confidentiality-related missing data, and then removed any remaining schools with missing values and kept only those that reported data in all years to form a balanced panel. Lastly, we applied Propensity Score Matching to improve comparability between groups. The main analysis will be on the sample after Propensity Score Matching. For more details on the data sources and the construction of the sample, see Appendix A. Generative AI has aided the production of code for appending the data sets and for the formatting of tables, graphs and equations, see *Description of AI Usage* in the final section of this thesis for more details.

In our data, the pre-treatment period covers the academic years 2014/15 to 2018/19, and the treatment period includes the academic years 2019/20 and 2020/21. For the pre-treatment period there are no hard-and-fast rules regarding the number of periods to include in the pre-period (Miller, 2023). For transparency, we chose to include the maximum number of available years in our dataset, which are five academic years, as opposed to Björkegren, Svaleryd, and Vlachos (2024), who included only four academic years. However, it is important to note that the periods furthest away may not be suitable for our counterfactual (Miller, 2023). For the post-period, we included two academic years since these were the years for which remote instruction was a recommendation.

For this thesis, we assume that all upper secondary schools transitioned to remote instruction, while lower secondary schools did not, in accordance with the official recommendations. Table 1 presents the percentage of students in remote instruction and the median number of weeks of remote instruction during the pandemic, indicating that this assumption broadly holds. However, for both academic years in the post-period, some upper secondary schools did not transition to remote instruction, while some lower secondary schools did. Based on Table 1,

it is evident that our assumption holds particularly well for the academic year 2019/20, but is less clear-cut for 2020/21. In addition, the intensity of treatment (i.e., the length of remote instruction) varied across schools. The figures in Table 1 are based on national aggregate data. School-level data on treatment intensity and on which schools actually transitioned to remote instruction exist, but were unavailable to us.

The dependent variables are *Failure Rates by Subject* for students in grade 9 and grade 10, and they refer to the proportion of students who received a grade of F and are reported separately for Swedish, English, and Mathematics (expressed as percentages). We denote our treatment indicator as *Grade10*, which is a binary variable equal to 1 if the failure rates belong to grade 10 students in upper secondary school, and 0 for grade 9 students in lower secondary school. The independent variables consist of school-level characteristics. *Foreign Background* denotes the proportion of students with a foreign background. A student is classified as having a foreign background if they were born abroad or if both parents were born abroad. Students lacking a personal identity number are also included. *Active Certified Teachers* is the proportion of teachers holding a professional license and subject qualification, actively engaged in teaching roles. *Post-secondary Parents* refers to the proportion of students with at least one parent having completed post-secondary education (minimum 30 ECTS). *Female Students* indicates the proportion of female students enrolled at the school level. *Private Ownership* is a binary variable equal to 1 if the school is privately owned and 0 otherwise. All variables except for *Private Ownership* are expressed as percentages.

In Table 2 we present summary statistics on the sample after Propensity Score Matching. It can be observed that there is less variation for upper secondary schools compared to lower secondary schools in the failure rates for each subject, as well as lower overall average failure rates. Furthermore, the means for Foreign Background, Active Certified Teachers, and Female Students area are all higher for upper secondary schools than for lower secondary schools in both the pre- and post-pandemic periods. This is not the case for Post-secondary Parents, where the means are similar. An interesting observation is that Female Students varies considerably more in upper secondary schools than in lower secondary schools. Lastly, public schools are more common than schools with private ownership.

Table 2 – Summary Statistics by Educational Level: Pre- and Post-Pandemic

	Lower Secondary		Upper Secondary	
	Pre-Pandemic	Post-Pandemic	Pre-Pandemic	Post-Pandemic
Failure Rates, Subject Swedish	3.49 (4.14)	4.01 (4.16)	0.26 (0.98)	0.29 (0.90)
Failure Rates, Subject English	6.69 (6.95)	6.63 (6.61)	0.32 (0.94)	0.38 (1.11)
Failure Rates, Subject Mathematics	9.05 (8.35)	8.71 (7.74)	1.09 (2.17)	1.32 (2.67)
Foreign Background	23.75 (16.65)	26.08 (18.13)	25.02 (17.38)	28.83 (20.05)
Active Certified Teachers	73.29 (13.91)	76.08 (13.22)	81.11 (13.43)	85.27 (11.89)
Post-secondary Parents	62.00 (16.86)	64.80 (16.95)	61.93 (12.51)	64.78 (13.70)
Female Students	49.09 (5.52)	49.29 (5.01)	52.67 (14.77)	53.34 (14.24)
Public Schools	750	300	750	300
Private Schools	455	182	455	182
All Schools	1205	482	1205	482

Notes: This table is inspired by Björkegren, Svaleryd, and Vlachos (2024). It presents averages and standard deviations (in parentheses) of school-level characteristics (expressed as percentages) before and after the Public Health Agency of Sweden’s first remote instruction recommendation. The pre-pandemic period covers academic years 2014/15 to 2018/19, and the post-pandemic period ranges from 2019/20 to 2020/21. The number of observations includes repeated unique schools, i.e., five years in the pre-pandemic period and two years in the post-pandemic period. For a unique school count see "After PSM" in Table 8. All variables are reported after Propensity Score Matching.

5 Methodology

5.1 Empirical Strategy

In line with Hallin et al. (2022), we aim to analyse the effect of the COVID-19 pandemic on Swedish students' learning outcomes. Specifically, we examine the impact of remote instruction during the pandemic on failure rates in Sweden, similarly to Halloran et al. (2021), who study this effect in the United States. Our empirical strategy is inspired by Björkegren, Svaleryd, and Vlachos (2024), who apply a Difference-in-Differences approach to compare Swedish upper secondary students, whose schools generally switched to remote instruction, with lower secondary students, whose schools largely remained open. Building on these studies, this thesis examines whether the transition to remote instruction during the COVID-19 pandemic affected student failure rates in the subjects Swedish, English, and Mathematics. To isolate this effect, a Difference-in-Differences strategy is used, comparing grade 9 students in lower secondary schools with grade 10 students in upper secondary schools, who transitioned to remote instruction to a greater extent. It would not have been enough to compare upper secondary students' failure rates before and after the implementation of remote instruction, as this identification strategy rather would have captured the effects of the pandemic. Comparing the post-pandemic failure rates of upper secondary students with those of lower secondary students also would have been insufficient, due to potential selection bias. The students who transitioned to remote instruction could be systematically different to the students who remained in-person in ways that affect failure rates. For these reasons, we follow the identification strategy by Björkegren, Svaleryd and Vlachos (2024), which addresses these potential biases.

In line with Vlachos, Hertegård and Svaleryd (2021), we decided to focus exclusively on grade 9 and grade 10 students, as these groups are likely to share more similarities. These students are only one year apart in age and they all receive final grades, compared to for instance grade 8 students, who do not receive a final grade until year 9. Passing the subjects Swedish, English and Mathematics also holds similar importance for grade 9 and grade 10 students, as it is a degree requirement for both groups. Additionally, the age difference is not sufficiently large to cause differences in how these student groups were affected by the virus (Turke, 2020). As highlighted by Table 1, grade 9 and grade 10 students are also especially suitable for our analysis, since grade 10 students received more remote instruction than other upper secondary students, while grade 9 students received less remote instruction than other lower secondary

students¹. Our treatment group is further narrowed by focusing exclusively on academic upper secondary programs, similarly to Björkegren, Svaleryd and Vlachos (2024). The reason we chose to focus on academic programs rather than all national upper secondary programs is due to the variation in use of remote instruction between the programs, as previously highlighted by Björkegren, Svaleryd and Vlachos (2024), where students in academic programs received more remote instruction than students in other national upper secondary programs, see Table 1².

While grade 9 and grade 10 students are similar in many ways, there are also important differences between these two groups. The main difference is that they attend different levels of education. Some of the differences associated with different educational levels do not pose a problem for identification, such as differences in the courses' content and difficulty level, since level differences in failure rates between grade 9 and grade 10 students is accounted for by the Difference-in-Differences methodology. Rather, there is cause for concern if there are differences between grade 9 and grade 10 students that influence these students' failure rates over time, as this would cause selection bias (Stuart et al., 2014). Lower secondary school is mandatory in Sweden, while upper secondary school is not, which means that grade 10 students are the subgroup of previous grade 9 students who chose to continue their education. Which students choose to study at the upper secondary level is likely influenced by a number of factors, such as their academic abilities and job aspirations. Consequently, there is likely a larger variation of academic results, including failure rates, for grade 9 students than for grade 10 students. This larger variation is confirmed by Table 2 in Section 4. Comparing these students may therefore cause biased estimates that include the effect of confounders. To address this issue, we employ Propensity Score Matching.

5.2 Propensity Score Matching

To account for compositional differences between untreated lower secondary schools and treated upper secondary schools, we implement Propensity Score Matching, which is widely used in empirical research (Caliendo and Kopeinig, 2005). The propensity score is defined as the conditional probability of assignment to a treatment given a set of observed characteristic. Given this, the distribution of the observed covariates between treated and untreated schools should be comparatively close after matching (Austin, 2011; Abadie and Imbens, 2016). In our context, a move to remote instruction is based on the level of education, which is either upper secondary

¹For an additional analysis, it could have been interesting to also compare grade 9 and grade 12 students, as they were both in their final year of their respective level of education, but this was not possible since the Swedish National Agency for Education only presents data on failure rates for grade 9 and grade 10 students.

²If not accounting for this difference in the use of remote instruction, students in vocational upper secondary programs would in one way have been more similar to grade 9 students, as they have more similar average grades (Skolverket, 2022a; Figure 1). However, they represent a smaller group of students, since 33% of those in national programs attend vocational programs while 67% attend academic programs. There are also greater gender differences in vocational programs overall and within specific programs, with for instance 96% of students in the electricity and energy program consisting of men (Skolverket and Sveriges Officiella Statistik, 2024). These issues indicate that students in vocational programs are not a suitable comparison to grade 9 students.

grade 10 or lower secondary grade 9. Since upper secondary school moved to remote instruction, the estimated propensity score becomes the probability that a school is an upper secondary school³. Thus, rather than selecting all lower secondary schools as the untreated group, we select those that are similar to the upper secondary schools available in our sample, in terms of observable characteristics.

We estimate a logistic regression since it is most commonly used in practice (Austin, 2003). We choose a one-to-one nearest neighbour matching without replacement, which means that each treated upper secondary school is matched with the lower secondary school that has the closest propensity score. According to Austin (2011) one-to-one is most commonly used, so therefore we opted for one-to-one matching. Without replacement implies that once a lower secondary school has been chosen as a match then it is not allowed to be matched again. Furthermore, matching is done in descending order from the highest to lowest propensity score. This makes sure that the schools that have difficulties in finding a match are prioritized first (Rubin, 1973). We also imposed a high tolerance level, i.e. caliper, on the maximum allowed distance in propensity score between matches to avoid extremely bad treated-untreated matches (Caliendo and Kopeinig, 2005). Lastly, to further increase the quality of the matches we only allow schools that are from the same municipality and have the same ownership structure to be matched.

The choice of which covariates to include was based on two aspects: i) covariates included should affect both the outcome variable and the treatment decision, and ii) matches should only be done on the past covariate values that have not been affected by the treatment or by anticipation of treatment (Caliendo and Kopeinig, 2005). Following these criteria, the covariates included in the propensity score estimation are Active Certified Teachers, Post-secondary parents and Foreign Background. As illustrated by the literature review, these variables often affect learning outcomes. More importantly, they also affect learning outcomes in Sweden (Andersson, Johansson and Waldenström, 2010; Von Kügelgen, 2023; Bohlmark and Lindahl, 2008; Edmark and Persson, 2021). Additionally, the estimation is conducted using school characteristics from the academic year 2014/15, which ensures that the covariates in the Propensity Score Matching are not affected by the treatment. In Appendix B, the covariate balance prior to and after propensity score matching is depicted.

5.3 Model Specification

To capture the pre-trend dynamics and estimate the effects in the years following the onset of the pandemic, we specify the following dynamic Difference-in-Differences model:

³With school-level data on which schools transitioned to remote instruction, the probability of moving to remote instruction could have been estimated, which would have yielded a more accurate propensity score.

$$Y_{ist} = \alpha + \sum_{t=2014/15, t \neq 2018/19}^{2020/21} \beta_t (Grade10_s \cdot Year_t) + \mathbf{X}_{it} + \phi_t + \eta_i + \varepsilon_{ist}, \quad (1)$$

where Y_{ist} is the share of students with an F grade in school i of the educational stage s at time t . α is the intercept and $Grade10_s$ is an indicator for whether a school i is an upper secondary school. $Year_t$ denotes the academic year corresponding to time period t . The β_t coefficients capture the interaction effects between $Grade10_s$ and $Year_t$, and thus represent the Difference-in-Differences estimates for each year from 2014/15 to 2020/21, relative to the reference year 2018/19. $\mathbf{X}_{it}$ is a vector of controls for the share of students with foreign background, active certified teachers, postsecondary parents, and the share of female students⁴. ϕ_t are academic year (time) fixed effects, which absorbs the $Year_t$ effects. η_i are school (individual) fixed effects, which absorbs the $Grade10_s$ effects. ε_{ist} is the error term. Cluster-robust standard errors are used, with the clustering being at the school level to account for serial correlation in errors across years within schools (Cameron and Miller, 2015). As a last note, three distinct Difference-in-Difference models will be estimated, one for each subject.

5.4 Identifying Assumptions

Following Callaway and Sant’Anna (2020) and Roth et al. (2023), we set the following identifying assumptions:

Assumption I (Irreversibility of Treatment). *No unit is treated in the first period, and once a unit becomes treated, it remains treated in all subsequent periods.*

In the academic year 2014/15, the subject of remote instruction was not yet included in the Swedish school law and the rule was to use in-person instruction (Regeringen, 2014). Therefore, it is likely to assume that grade 9 and grade 10 students did not receive remote instruction in the academic year 2014/15. Moreover, Table 1 depicts how grade 10 students in academic programs overall receive remote instruction the academic years 2019/20 to 2020/21. Assumption I is therefore likely to hold.

Assumption II (No Treatment Anticipation). *The treatment has no causal effect on outcomes prior to its implementation.*

Assumption II is likely to hold as well. The first documented case of COVID-19 was in December 2019 (U.S. Centers for Disease Control and Prevention, 2024), making it unlikely that the sudden implementation of remote instruction in the spring of 2020 would have been predicted in previous academic years. Therefore, the use of remote instruction during the pandemic should not have affected failure rates of students in the years prior to academic year 2019/20.

⁴Recall the variable definitions in Section 4.

Assumption III (Conditional Parallel Trends). *In the absence of treatment, the difference between the treatment and control group is constant over time.*

Assumption III is the key identifying assumption for the Difference-in-Differences approach. It means that in the absence of remote instruction, the failure rates of the matched grade 9 and grade 10 students would have evolved similarly during the pandemic. In our case, the parallel trends assumption is conditional, since we use control variables in our specifications. If the control and treatment groups are similar to each other in observed characteristics, it is more likely that they evolve similarly. As previously discussed, grade 9 and grade 10 students share many similarities and the use of propensity score matching further removes observed differences between the groups. While it is not possible to directly test the parallel trends assumption, since a counterfactual can not be measured (Cunningham, 2021), it is possible to examine whether the trend of the treatment group followed the trend of the control group prior to treatment, which provides an indication of whether the parallel trends assumption is more or less likely to hold. We visually inspect the average failure rates over time for grade 9 and grade 10 students and test the difference in pre-trends using our dynamic Difference-in-Differences specification. If this shows that there were no significant Difference-in-Differences estimates prior to treatment, this increases the likelihood that the parallel trends assumption holds (Roth et al., 2023). If our dynamic Difference-in-Differences specification provides significant estimates for the post-treatment periods, we will conduct a sensitivity analysis based on recommendations presented by Rambachan and Roth (2023). The rationale for estimating Difference-in-Differences in the pre-pandemic periods is the assumption that pre-trends are informative about what the trends following treatment would have been in the absence of treatment. The sensitivity analysis presented by Rambachan and Roth (2023) uses a similar rationale, but describes how Difference-in-Differences estimates can be interpreted based on how large the violations to the parallel trends assumption are.

Assumption IV (Overlap). *For every treated unit with a given set of covariates, there exists a positive probability of observing untreated units in the population with the sufficient similar covariate values.*

There is overlap in observed characteristics between the grade 9 and grade 10 students in our sample prior to matching and this overlap increases from matching, see Table 10 in Appendix B. Therefore, Assumption IV is expected to hold.

6 Results

The following sections present our main results for Swedish, English and Mathematics. All results are based on the matched sample, see Appendix D for results based on an unmatched sample.

6.1 Subject Swedish

Figure 3 plots average failure rates over time for grade 9 and grade 10 students separately. Like the following graphs in this section, the time period in this figure is from 2014/15 to 2020/21 and a vertical dotted line indicates the transition to remote instruction. Figure 3 shows a level difference between the two groups, with failure rates being consistently higher for grade 9 students compared to grade 10 students across all academic years. For both groups, failure rates appear relatively stable throughout the period, although grade 9 students exhibit a modest upward trend in the pretreatment period. Failure rates do not appear to have changed much during the pandemic for either group.

Figure 4 presents the estimated difference in failure rates between grade 9 and grade 10 students relative to the differences in failure rates in the reference year of 2018/19. For instance, if an estimate would have been equal to 0, then the difference in failure rates between grade 9 and grade 10 would have been the same in that year as it was in the reference year. Each estimate includes 95% confidence intervals, visually represented by the vertical lines. There is no confidence interval for the reference year since it is not an estimate. Figure 4 shows that the difference in failure rates between grade 9 and grade 10 students is relatively even over time, although the estimates indicate a slight trend in the difference in failure rates. The estimates in the pre-pandemic period are statistically insignificant, as indicated by the confidence intervals' inclusion of 0 in Figure 4, suggesting that the parallel trend assumption holds.

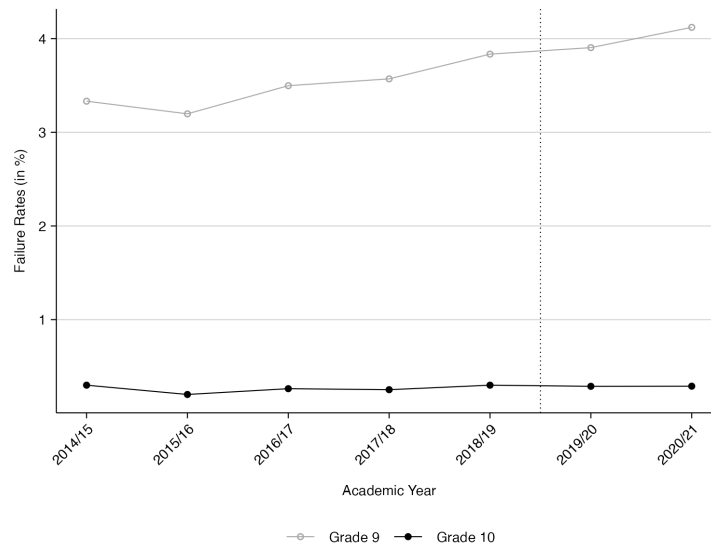


Figure 3 – Trends in Failure Rates in Subject Swedish

Notes: The figure shows the time trends in failure rates in subject Swedish by level of education for the academic years 2014/15 to 2020/21. The vertical reference line indicates the transition to remote instruction.

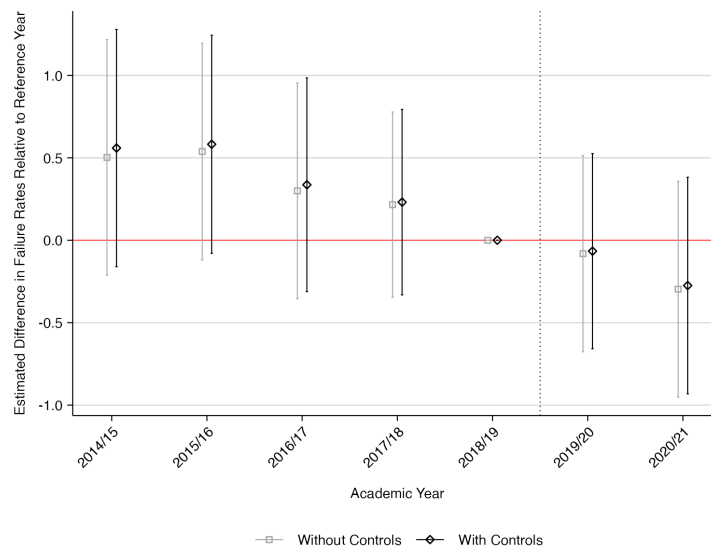


Figure 4 – Dynamic Difference-in-Differences of Failure Rates in Subject Swedish

Notes: The figure shows the estimated difference in failure rates in subject Swedish between grade 9 and grade 10, relative to reference academic year 2019/20. Both models include school-level and time fixed effects. Controls are Foreign Background, Active Certified Teachers, Post-secondary Parents, and Female Students. 95% confidence interval are included. Standard errors are clustered at the school level.

The Difference-in-Differences estimates for the two academic years in the pandemic are relatively small, negative estimates of -0.07 and -0.28, see Table 10 in Appendix C. However, neither of these estimates are statistically significant. We therefore find no statistically significant effect of remote instruction on failure rates in Swedish in the academic years 2019/20 and 2020/21. This result holds in both the unconditional and conditional model. Adding covariates

one by one does not change the direction or the size of the interaction terms, see Table 10 in Appendix C.

6.2 Subject English

Similarly to failure rates in Swedish, failure rates in English are constantly higher for grade 9 students than for grade 10 students, see Figure 5. The pre-pandemic trend in failure rates for grade 10 students is stable over time, but this is not the case for grade 9 students. Figure 5 shows a sharp increase in failure rates in academic year 2016/17, followed by a gradual decline. Thus, visually, there appears to be a difference in trends prior to treatment.

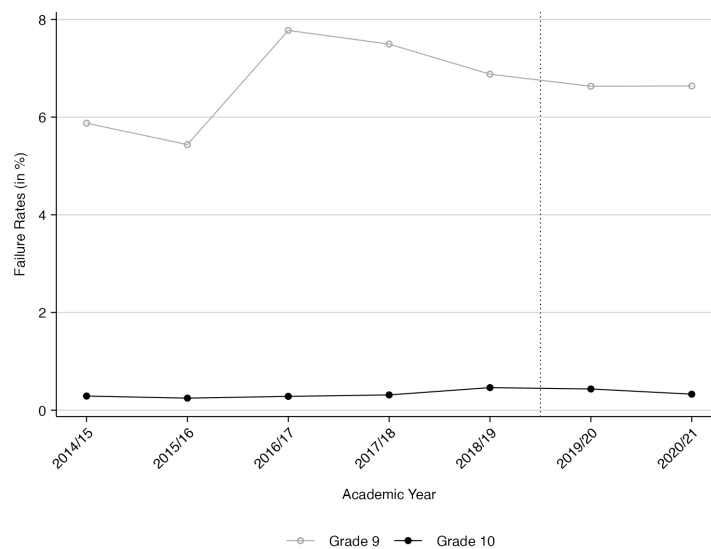


Figure 5 – Trends of Student Failure Rates in Subject English

Notes: The figure shows the time trends in failure rates in subject English by level of education for the academic years 2014/15 to 2020/21. The vertical reference line indicates the transition to remote instruction.

The difference in trends prior to treatment is also evident in Figure 6, where each pre-pandemic estimate produced by the conditional model is statistically significant. For the identification strategy, this implies that the Difference-in-Differences estimates will not capture the causal impact of remote instruction on failure rates. The Difference-in-Differences regressions produced relatively small, positive, but insignificant estimates, with or without control variables, see Table 11 in Appendix C. These results would indicate that remote instruction had no effect on failure rates in English, but due to clear indications that the parallel trends assumption is violated, the results for English are inconclusive.

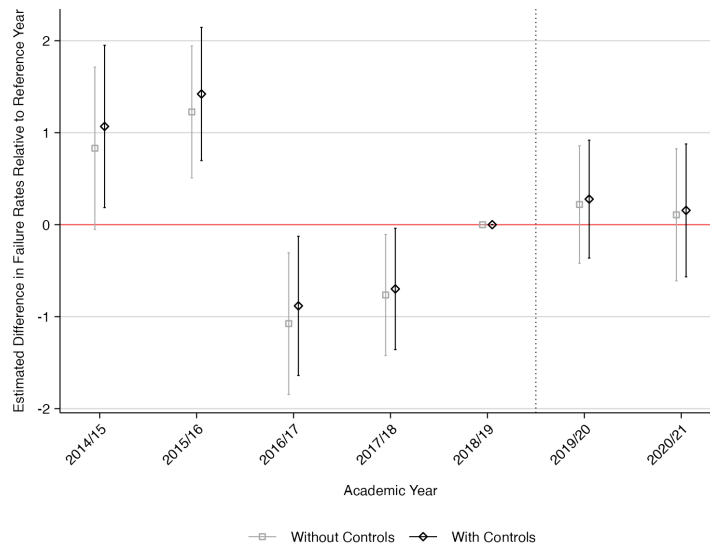


Figure 6 – Dynamic Difference-in-Differences of Failure Rates in Subject English.

Notes: The figure shows the estimated difference in failure rates in subject English between grade 9 and grade 10, relative to reference academic year 2019/20. Both models include school-level and time fixed effects. Controls are Foreign Background, Active Certified Teachers, Postsecondary Parents, and Female Students. 95% confidence interval are included. Standard errors are clustered at the school level.

6.3 Subject Mathematics

Figure 7 presents the time trends of failure rates in Mathematics for grade 9 and grade 10 students from the academic year 2014/15 to 2020/21. Similarly to Swedish and English, failure rates are consistently higher for grade 9 students in comparison to grade 10 students. There is a clear increase in grade 9 students' failure rates in academic year 2016/17, similar to the time trend of failure rates in English, see Figure 5. We also see a sharp decrease in grade 9 students' failure rates in the previous academic year.

Figure 8 shows that the Difference-in-Differences estimates in the first two pre-pandemic academic years are statistically significant, indicating violations to the parallel trends assumption. However, if the analysis includes fewer pre-pandemic years, the pre-pandemic trends for grade 9 and grade 10 students appear similar. For the pandemic years, Figure 8 shows positive estimates that are statistically significant at the 95% confidence level. In 2019/20 and 2020/21, remote instruction is associated with a 1.28 and 1.79 percentage point increase in failure rates in Mathematics, see Table 12 in Appendix C. These estimates are relatively large in magnitude in relation to the initial average estimates of 9.05% and 1.09% in the pre-pandemic periods for grade 9 and grade 10 respectively, see Table 2. These results appear to be largely driven by grade 9 students, see Figure 7.

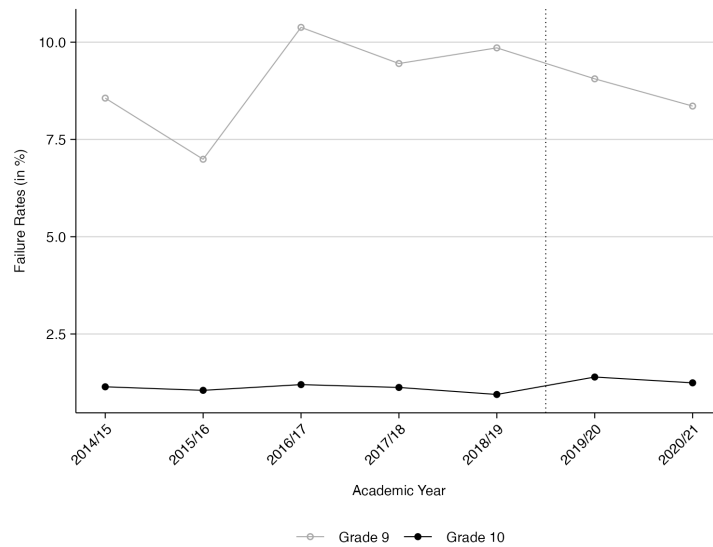


Figure 7 – Trends of Student Failure Rates in Subject Mathematics.

Notes: The figure shows the time trends in failure rates in subject Mathematics by level of education for the academic years 2014/15 to 2020/21. The vertical reference line indicates the transition to remote instruction.

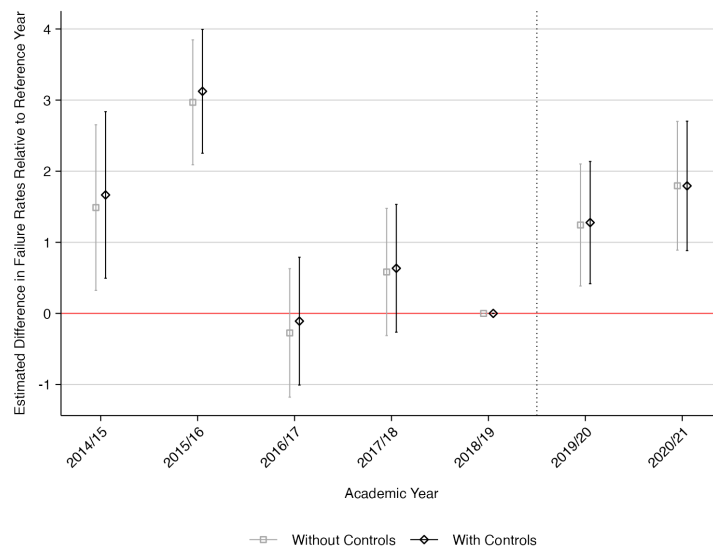


Figure 8 – Dynamic Difference-in-Differences of Failure Rates in Subject Mathematics.

Notes: The figure shows the estimated difference in failure rates in subject Mathematics between grade 9 and grade 10, relative to reference academic year 2019/20. Both models include school-level and time fixed effects. Controls are Foreign Background, Active Certified Teachers, Postsecondary Parents, and Female Students. 95% confidence interval are included. Standard errors are clustered at the school level.

7 Sensitivity Analyses

7.1 Heterogeneity Analyses: Subjects Swedish and Mathematics

Previous research has indicated that different subgroups may be especially affected by remote instruction (Betthäuser, Bach-Mortensen and Engzell, 2023; Hall, Hardoy and Lundin, 2022; Jakubowski, Gajderowicz and Patrinos, 2025; Engzell, Frey and Verhagen, 2021). To study this, we conduct heterogeneity analyses for the two ownership types and for high and low shares in Foreign background, Postsecondary parents, Female students, and Active Certified Teachers. The heterogeneity analyses are conducted for Swedish and Mathematics; for Swedish this is due to pre-pandemic trends appearing parallel, and for Mathematics it is due to the ambiguity surrounding whether post-pandemic failure rates would be expected to have evolved similarly in the absence of remote instruction or not. For the subject English, given the indication that the parallel trends assumption is violated regardless of the pre-treatment window chosen we decided to exclude the heterogeneity analysis from the main text. For those interested, the corresponding heterogeneity analysis for the subject English is presented in Appendix E.

Regarding failure rates in Swedish, Table 3 shows statistically insignificant Difference-in-Differences estimates for all pre- and post-pandemic years for schools of both ownership types, for different shares of postsecondary parents, and for different shares of active certified teachers (based on a 95% confidence level). For these subgroups, remote instruction does not appear to have had an effect on failure rates in Swedish. The same holds for schools with a high share of students with a foreign background (based on a 95% confidence level). For schools with a low share of female students, there is a significant negative effect of remote instruction on failure rates in the academic year 2020/21, indicating that the use of remote instruction compared to in-person instruction reduced failure rates. However, since the result for academic year 2019/20 is statistically insignificant, it is unclear whether or not this subgroup generally benefits from the use of remote instruction. For the remaining subgroups, the results are inconclusive due to indications of violations of the parallel trends assumption.

Table 3 – Heterogeneous Effects, Subject Swedish

	Failure Rates in Subject Swedish									
	Private (1)	Public (2)	Foreign Background High	Low	Postsecondary Parents High	Low	Female Students High	Low	Active Certified Teachers High	Low
2014/15 × Grade 10	0.281 (0.563)	0.759 (0.468)	−2.448* (1.391)	0.785** (0.367)	0.495 (0.358)	0.188 (1.036)	0.697 (0.706)	0.387 (0.513)	0.592 (0.388)	1.640 (2.578)
2015/16 × Grade 10	0.487 (0.550)	0.658 (0.417)	−1.164 (1.686)	0.700** (0.333)	0.416 (0.312)	0.452 (0.999)	1.666*** (0.637)	−0.030 (0.471)	0.635* (0.348)	−0.786 (2.482)
2016/17 × Grade 10	0.206 (0.577)	0.419 (0.394)	−0.644 (1.191)	0.424 (0.342)	0.110 (0.318)	0.269 (1.008)	0.924 (0.657)	−0.046 (0.446)	0.393 (0.318)	−0.223 (2.015)
2017/18 × Grade 10	0.246 (0.455)	0.233 (0.369)	−0.177 (1.286)	0.302 (0.288)	0.268 (0.301)	0.065 (0.830)	0.708 (0.473)	−0.024 (0.400)	0.197 (0.294)	−0.445 (2.462)
(Reference Year)	0	0	0	0	0	0	0	0	0	0
2019/20 × Grade 10	0.020 (0.560)	−0.120 (0.346)	0.030 (1.412)	−0.044 (0.300)	−0.003 (0.276)	−0.691 (0.995)	0.713 (0.596)	−0.685 (0.420)	−0.117 (0.288)	0.270 (1.512)
2020/21 × Grade 10	0.266 (0.563)	−0.591 (0.413)	−0.463 (1.387)	−0.241 (0.340)	−0.183 (0.331)	−1.062 (1.091)	0.988 (0.602)	−1.118** (0.447)	−0.319 (0.328)	−0.096 (1.650)
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Observations	1,274	2,100	322	3,052	2,621	753	1,611	1,763	3,208	166

Notes: This table presents heterogeneous treatment effects by school ownership and composition characteristics. For Foreign Background, Post-secondary Parents, Female Students and Active Certified Teachers, the sample was split as follows: "High" refers to schools with shares above 50%, and "Low" to those with shares at or below 50%. It was not possible to add controls to Low Postsecondary Parents sample, and to our best knowledge this was due to high collinearity. Therefore, controls are excluded in both subgroup analyses of Postsecondary Parents. Standard errors are clustered at the school level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

For failure rates in Mathematics, all subgroups – except the low share group in active certified teachers – have statistically significant estimates in the pre-pandemic period, but these are generally restricted to academic year 2014/15 and 2015/16, see Table 4. In academic year 2015/16, estimates were considerably larger compared to other years for all subgroups. Estimates in the pandemic are positive in both academic years for all subgroups – except the low share group in active certified teachers yet again – but both estimates are only significant for publicly owned schools, schools with low share of students with foreign background, schools with a high share of parents with post-secondary education, schools with high share of female students, and high share of active certified teachers. For schools with a low share of female students, the Difference-in-Differences estimate for academic year 2020/21 is statistically significant, but not for academic year 2019/20.

Table 4 – Heterogeneous Effects, Subject Mathematics

	Failure Rates in Subject Mathematics									
	Private (1)	Public (2)	Foreign Background High	Low	Postsecondary Parents High	Low	Female Students High	Low	Active Certified Teachers High	Low
2014/15 × Grade 10	0.935 (0.867)	1.953** (0.803)	−0.019 (3.566)	1.562*** (0.548)	1.740*** (0.541)	1.651 (1.622)	1.254 (1.055)	1.987** (0.826)	1.823*** (0.633)	−1.500 (5.261)
2015/16 × Grade 10	1.883** (0.757)	3.779*** (0.526)	4.557** (1.975)	2.787*** (0.443)	2.152*** (0.453)	5.599*** (1.277)	3.277*** (0.879)	3.155*** (0.625)	3.315*** (0.446)	−2.237 (4.276)
2016/17 × Grade 10	0.579 (0.713)	−0.556 (0.598)	−0.657 (1.763)	−0.092 (0.478)	−0.050 (0.458)	−0.301 (1.328)	0.537 (0.854)	−0.137 (0.629)	−0.068 (0.459)	−0.331 (5.168)
2017/18 × Grade 10	0.668 (0.748)	0.615 (0.576)	−0.183 (2.508)	0.667 (0.443)	0.232 (0.464)	2.187* (1.259)	1.455** (0.666)	0.315 (0.671)	0.636 (0.464)	−2.065 (2.691)
(Reference Year)	0	0	0	0	0	0	0	0	0	0
2019/20 × Grade 10	0.693 (0.706)	1.665*** (0.562)	1.509 (2.277)	1.257*** (0.432)	1.209*** (0.439)	0.979 (1.204)	1.495** (0.734)	1.137* (0.610)	1.341*** (0.432)	−0.923 (4.554)
2020/21 × Grade 10	1.171* (0.655)	2.246*** (0.638)	0.909 (2.296)	1.909*** (0.470)	1.783*** (0.466)	0.880 (1.393)	2.071*** (0.740)	1.507** (0.677)	1.753*** (0.463)	−0.444 (4.633)
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Observations	1,274	2,100	322	3,052	2,621	753	1,611	1,763	3,208	166

Notes: This table presents heterogeneous treatment effects by school ownership and composition characteristics. For Foreign Background, Post-secondary Parents, Female Students and Active Certified Teachers, the sample was split as follows: "High" refers to schools with shares above 50%, and "Low" to those with shares at or below 50%. The reference academic year is 2018/19 and is excluded. It was not possible to add controls to Low Postsecondary Parents sample, and to our best knowledge this was due to high collinearity. Therefore, controls are excluded in both subgroup analyses of Postsecondary Parents. Standard errors are clustered at the school level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

7.2 Bounding Relative Magnitudes: Subject Mathematics

Significant post-treatment estimates were observed for failure rates in the subject of Mathematics, but our dynamic Difference-in-Differences graphs indicate that the parallel trends assumption is unlikely to hold if all four pre-pandemic academic years are taken into account. Most Difference-in-Differences applications require the parallel trends assumption to hold, and therefore a violation of this assumption would mean an end to our analysis of subject Mathematics. However, recent advances by Rambachan and Roth (2023) have highlighted alternative approaches to Difference-in-Differences estimation that relaxes the parallel trends assumption. The parallel trends assumption builds on the intuition that the difference in trends between upper and lower secondary schools prior to the pandemic – when there was largely only in-person instruction – is informative about what the difference in trends would have been in 2019/20 and 2020/21, had there been no switch to remote instruction. The approaches presented by Rambachan and Roth (2023) build on the same intuition, but present the parallel trends assumption as a special case of the types of possible restrictions, rather than the rule. This allows for further investigations of potential treatment effects even when the parallel trends assumption appears to be violated. They present two main approaches, *Smoothness Restrictions* denoted $\Delta^{SD}(M)$ and *Bounding Relative Magnitudes* denoted $\Delta^{RM}(\bar{M})$, where the choice between them depends

on the research setting.

Let us first consider *Smoothness Restrictions*. If there is reason to suspect that there may exist trend differences between the treatment and control groups that evolve smoothly over time, then using the same rationale as for the parallel trends assumption, it would be reasonable to expect that this difference in trends would have evolved similarly pre- and post-treatment, had there been no treatment. This is the underlying idea used for the *Smoothness Restrictions* where the slope is allowed to change across succeeding periods.

Formally, following Rambachan and Roth (2023), let Δ be a set of possible differences in trends, where δ denotes the difference in trends between the treated and untreated groups. In *Smoothness Restrictions* let $\delta \in \Delta^M(M)$ for $M \geq 0$ where

$$\Delta^{SD}(M) := \{\delta : |(\delta_{t+1} - \delta_t) - (\delta_t - \delta_{t-1})| \leq M, \forall t\}.$$

The extent to which the slope δ can change between subsequent periods is set by $M \geq 0$. When $M = 0$, then the difference in trends is required to be linear, and when $M > 0$, the slope is allowed to be increasingly non-linear. In our context, there seems to exist certain confounders such as for instance parents' educational level, but these should already be accounted for through the use of Propensity Score Matching. Therefore, *Smoothness Restrictions* are not applicable in our context.

Now to *Bounding Relative Magnitudes*, which is applicable when there are concerns about shocks that may have influenced the control and treatment group differently. In our setting, there is a concern that the sudden influx of refugees in 2015 changed the composition of students, which may have influenced failure rates. Therefore, the use of *Bounding Relative Magnitudes* is useful in our context. This also aligns with our results, which indicate sudden deviations in the difference in pre-pandemic trends, rather than a smoothly evolving difference in trends over time.

Similarly to the *Smoothness Restrictions*, in *Bounding Relative Magnitudes* the pre-treatment difference in trends does not have to be statistically insignificant. Instead, jumps between consecutive periods in the pre-treatment difference in trends are allowed, but this has the implication that jumps in the post-treatment period are only considered to be signs of a treatment effect if they are of a certain size relative to the maximum jump in the pre-treatment period.

Formally, following Rambachan and Roth (2023), let $\delta \in \Delta^{RM}(\bar{M})$ for $\bar{M} \geq 0$ where

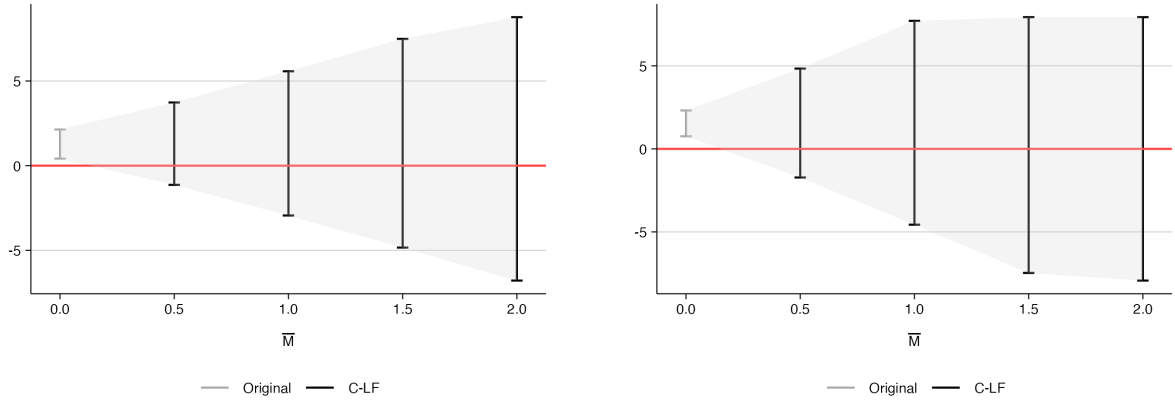
$$\Delta^{RM}(\bar{M}) = \left\{ \delta : \forall t \geq 0, |\delta_{t+1} - \delta_t| \leq \bar{M} \cdot \max_{s < 0} |\delta_{s+1} - \delta_s| \right\}.$$

Here, $t \geq 0$ indexes post-treatment periods and $s < 0$ indexes pre-treatment periods. This equation says that the largest post-treatment violation of parallel trends between subsequent periods is limited to be less than or equal to $(\bar{M})$ times the absolute value of the maximum pre-treatment violation of parallel trends.

In our setting, the rationale is that if there is a large enough change in the difference in trends

in academic years 2019/20 and 2020/21, then this would indicate that remote instruction had an effect on failure rates even if there are jumps in the difference in trends in the periods 2014/15 - 2018/19. Based on this approach, we can give a range for where the Difference-in-Differences estimate should be, rather than a precise estimate. This range, determined by $\Delta^{RM}(\bar{M})$, is based on how large we require the jumps in 2019/20 and 2020/21 to be for us to be able to reject no effect. Therefore, the significance of the estimate depends on the range. For instance, if $M = 0.5$, then jumps in 2019/20 and 2020/21 have to be greater than half of the largest jump before 2019/20, while if $M = 2$, then jumps in 2019/20 and 2020/21 have to be twice as large as the largest jump before 2019/20. The size of the range where the Difference-in-Differences estimate is no longer significant is called the breakdown value of M .

Using *Bounding Relative Magnitudes*, we estimate the immediate effect of remote instruction on failure rates in Mathematics for the first year of the pandemic, and the average effect in the years of the pandemic, see Figure 9. These two methods provide similar results, where almost regardless of the value of M , we can not reject the null hypothesis of no effect; the breakdown value of M is approximately 0. For instance, even if the jump in the difference in trends in 2019/20 only needs to be half as large as the maximum jump in the difference in trends before the pandemic ($M=0.5$), then the Difference-in-Differences estimate for 2019/20 is still not significantly different from 0. This is reasonable, considering that Figure 8 in Section 6 shows that the Difference-in-Differences estimate in 2019/20 and 2020/21 are relatively small in comparison to the Difference-in-Differences estimate in academic year 2015/16. Based on this analysis using *Bounding Relative Magnitudes*, our results therefore do not support that remote instruction had an effect on failure rates in Mathematics, at least when considering all the pre-pandemic academic years from 2014/15 in the analysis.



(a) First Year After Treatment

(b) Average Effects

Figure 9 – Bounded Relative Magnitudes

Notes: The graphs evaluate the sensitivity of the estimated treatment effect in the subject Mathematics for (a) the academic year 2019/20, which is the first period of treatment and (b) the treatment effect over the two post-treatment periods – average effect – 2019/20 and 2020/21. $\bar{M}$ is the Relative Magnitudes parameter. For each value of $\bar{M}$, it gives a robust confidence interval that allows for violations of parallel trends in the post-treatment period to be no more than $\bar{M}$ times the worst pre-treatment violation of parallel trends. We consider a range of values for $\bar{M}$, taking values of 0.5, 1.0, 1.5, and 2.0. C-LF refers to Conditional-Least Favourable Hybrid Confidence sets (Rambachan and Roth, 2022).

7.3 Robustness Check: Subjects Swedish, English and Mathematics

During the academic years 2019/20 and 2020/21, factors such as local COVID-19 incidence, municipal support from the government, and school-level investments in IT infrastructure varied across schools and municipalities (Finansdepartementet, 2020; Florida and Mellander, 2021). These factors could influence failure rates and may not be fully accounted for by school fixed effects. To better isolate the effect of remote instruction from these simultaneous shocks, one would ideally include municipality-times-year controls. To assess whether these factors alter our results, we provide a robustness analysis in which municipality-times-year controls are included.

In Table 5, the results for the specification that includes municipality-times-year controls are presented. These results are similar those shown in Table 10 for the subject Swedish, Table 11 for the subject English, and Table 12 for the subject Mathematics, all found in Appendix C. This suggests that including municipality-times-year controls does not significantly alter the results.

Table 5 – Dynamic Difference-in-Differences Models by Subject

	Failure Rates in Subject Swedish		Failure Rates in Subject English		Failure Rates in Subject Mathematics	
	(1)	(2)	(3)	(4)	(5)	(6)
2014/15 × Grade 10	0.502 (0.357)	0.574 (0.356)	0.831* (0.437)	1.060** (0.438)	1.489** (0.585)	1.702*** (0.586)
2015/16 × Grade 10	0.538* (0.325)	0.587* (0.324)	1.226*** (0.350)	1.414*** (0.356)	2.969*** (0.432)	3.148*** (0.431)
2016/17 × Grade 10	0.300 (0.319)	0.335 (0.314)	−1.076*** (0.362)	−0.896** (0.358)	−0.274 (0.445)	−0.101 (0.441)
2017/18 × Grade 10	0.216 (0.277)	0.234 (0.276)	−0.764** (0.325)	−0.700** (0.327)	0.583 (0.442)	0.646 (0.443)
(Reference Year)	0	0	0	0	0	0
2019/20 × Grade 10	−0.081 (0.292)	−0.069 (0.291)	0.220 (0.309)	0.276 (0.311)	1.244*** (0.425)	1.269*** (0.426)
2020/21 × Grade 10	−0.297 (0.323)	−0.272 (0.326)	0.107 (0.338)	0.161 (0.344)	1.795*** (0.442)	1.790*** (0.445)
Foreign Background		0.006 (0.015)		0.051** (0.025)		0.074*** (0.029)
Active Certified Teachers		0.007 (0.007)		−0.002 (0.008)		−0.006 (0.010)
Postsecondary Parents		−0.030 (0.018)		−0.092*** (0.028)		−0.036 (0.031)
Female Students		−0.0001 (0.013)		0.018 (0.017)		0.033 (0.024)
School Fixed Effects	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Time Fixed Effects	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Municipality × Year	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Observations	3,374	3,374	3,374	3,374	3,374	3,374

Notes: This table reports estimates of the effect of remote instruction on failure rates in the subjects Swedish, English, and Mathematics, based on two specifications of the Dynamic Difference-in-Differences model for each subject: one excluding all controls and one including the full set of controls. All models include school and time fixed effects, as well as Municipality × Year controls. The reference academic year is 2018/19. Standard errors are clustered at the school level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

8 Discussion

8.1 Interpretation of Results

8.1.1 Pre-Pandemic Trends

Prior to the pandemic, failure rates in Swedish appear to have evolved similarly for the matched grade 9 and grade 10 students. For Mathematics, we see a dip in failure rates for grade 9 students in the academic year 2015/16. Svaleryd and Vlachos (2021) describe that the grade 9 national exam in Mathematics appears to have been easier in academic year 2015/16 compared to other years, which may explain this result. Our heterogeneity analysis of failure rates in Mathematics supports this conclusion, since each subgroup had statistically significant estimates in the academic year 2015/16 that were considerably large relative to the estimates in other years, indicating that for each subgroup, grade 9 students saw a decrease in failure rates relative to grade 10 students.

In the academic year 2016/17, there appears to be a sudden increase in failure rates for grade 9 students in Mathematics and English, but not in Swedish. One potential explanation could be that students that attend Swedish as a Second Language, and therefore are not included in the failure rates for Swedish, are driving these results. Here, one potential reason behind the difference in trends between grade 9 and grade 10 students could be the large influx of refugees to Sweden around that time. Our heterogeneity analysis of failure rates in Mathematics does not indicate that belonging to grade 10 had an effect on failure rates for schools with a high share of students with a foreign background in that academic year. While the measure of foreign background includes people of heterogeneous backgrounds, where only a portion of the people are refugees, this provides some indication that the change in failure rates in 2016/17 was not driven by compositional changes in the student population due to an influx of refugees. Determining the cause of this change would therefore require further analysis.

8.1.2 Results for Subject Swedish

Our results for failure rates for subject Swedish indicate that remote instruction did not have an effect on lower achieving students' learning, which contrasts much of the previous literature on the topic. However, it aligns with previous research on learning in Sweden during the pandemic (Svaleryd and Vlachos, 2021; Hallin et al., 2022). From research on a related topic, Björkegren,

Svaleryd and Vlachos (2024) have also found that remote instruction had a positive effect on students' mental health, which also contrasts previous literature in that area (Xu and Wang, 2023). Our results, along with these previous papers, therefore indicate that Swedish students may have been less affected by remote instruction during the pandemic than students in other countries.

Moreover, our result of no effect from the remote instruction on learning in subject Swedish also holds for most subgroups that we analysed, see Table 3. This is also a result that contrasts the previous research on the topic. The only subgroup where we saw an indication that there may have been an effect was for schools with a low share of female students, but since this effect was only apparent in one out of two pandemic years, it is unclear whether there generally was an effect for this group. The significant effect on was also negative, which would go against the previous literature on the topic.

8.1.3 Potential Causes

Overall, for subject Swedish, remote instruction does not appear to have had a negative effect on lower achieving students generally, or on certain subgroups of lower achieving students. To discuss potential causes of this, we return to the three channels presented by Agostinelli et al. (2022): schools, peers and parents. In Sweden, the direct effect from school closures may have been smaller than in other countries. Shorter school closures have been associated with a smaller effect on learning and Swedish schools were closed for a shorter time than many other countries (Jakubowski, Gajderowicz and Patrinos, 2025). The Nordic countries were also relatively well prepared for remote instruction compared to other countries (Hall, Hardoy and Lundin, 2022) and teachers have described that it worked surprisingly well (Bergdahl and Nouri, 2020). There may also be existing issues with the Swedish school environment that remote instruction mitigated. Findings from Björkegren, Svaleryd and Vlachos (2024) indicate that the school environment negatively affects students' mental health and teachers have described that students often focused better in remote instruction due to fewer distractions from their classmates (Bergdahl and Nouri, 2020).

Regarding peer effects, Agostinelli et al. (2022) describe that school closures decrease interactions with peers at school and interactions become more restricted to peers who live in the same neighborhood. Thereby, school closures reduce interactions between students of different backgrounds. As previously mentioned, school segregation has increased in Sweden, largely due to increased housing segregation (Holmlund, Böhlmark and Lindahl, 2015). However, the peer effects channel may still be influential in Sweden, since Sweden is still considered a low-to-medium segregated school system in an international context. Agostinelli et al. (2022) also describe that less socialising with friends also can negatively affect learning. Since Sweden used less restrictive measures to limit the spread of the virus (Regeringen, 2021), children had more opportunities to socialise during the pandemic compared to many other countries.

Therefore, it is possible that this aspect of the peer channel had a smaller effect in Sweden.

The third channel presented by Agostinelli et al. (2022) is that remote instruction puts more pressure on parents to help their children with their studies. The ability of parents to help their children in this way varies, as it for instance is restricted by whether the parents can work from home during the pandemic and how much time they are able to spend with their children. In Sweden, this effect may have been larger due to the lack of a lockdown, so some parents were at home more than others. Among Swedish upper secondary students, lower parental education is also associated with a more crowded living situation, which may influence studies during remote instruction (Statistikmyndigheten SCB, 2021). This indicates that while remote instruction did not appear to generally have an effect on failure rates in Swedish, it may have affected failure rates of students whose parents did not have the same opportunities to help their children with their studies. However, our heterogeneity analysis showed that schools with students whose parents have lower educational levels do not appear to be adversely affected by remote instruction. We therefore find no indications of a parent effect in this setting. This result is also interesting since it differs from the large differences based on parental education that have been observed in previous research (Engzell, Frey and Verhagen, 2021).

In sum, all three channels presented by Agostinelli et al. (2022) may have been less influential in Sweden, which may explain our results.

8.1.4 Results for Subjects English and Mathematics

For English and Mathematics, the main results are inconclusive due to indications that the parallel trends assumption does not hold, see Figure 6 and Figure 8. However, our sensitivity analysis using Bounding Relative Magnitudes indicates that for subject Mathematics, similarly to subject Swedish, there was no effect of the use of remote instruction on lower achieving students' learning. This result could thereby be interpreted similarly to the result for Swedish, where the direct effect of school closures, peer effects and parental effects may have been less influential in Sweden, resulting in no effect on learning.

As there are no hard-and-fast rules in the literature on how to select the number of years to include in a Difference-in-Differences analysis (Miller, 2023), it is also worth considering how this choice may affect the results. In our main analysis, we kept all years included in the data, but if we instead were to only consider time periods from 2016/17 to be informative about the evolution of failure rates in 2019/20 and 2020/21, had there been no implementation of remote instruction, then the conclusions would have changed. In this case, the parallel trends assumption would appear to hold, see Figure 8. For both academic years of the pandemic, the Difference-in-Differences estimates are positive and significant, which at a first glance would indicate that remote instruction negatively affected students' learning in Mathematics. However, Figure 7 indicates that this change is largely driven by grade 9 students. The implication would then rather be that grade 9 students' learning in Mathematics was positively affected by

the pandemic.

One potential explanation for this could be that higher student absence reduced class size, which may have positively affected lower achieving students' learning. There is considerable evidence that smaller class sizes improve learning outcomes and these effects tend to be driven by lower achieving students (Sjögren et al., 2021). Although, student absence is also likely to have a negative effect on learning through reduced instruction. If the results are explained through this channel, then this would either indicate that lower achieving grade 9 students were less absent than their peers during the pandemic, or that the benefits of a smaller class size outweighed the negative effects of reduced instruction for this subject.

Consequently, our results for Mathematics indicate no effect of remote instruction on failure rates, but there are some indications that in terms of learning, lower achieving students with in-person instruction benefitted from the pandemic, potentially due to reduced class size. Whether this channel exists and whether lower achieving students benefitted from the pandemic would depend on future research. The heterogeneity analyses also illustrate that the significant effect appears to be driven by specific subgroups. It would therefore also be valuable to further investigate why this may be.

8.2 Implications for Academia and Policy

We add to the existing research by providing evidence that remote instruction during the pandemic may not always have a negative effect on students' learning. This is especially interesting considering that we study lower achieving students, for whom previous research describes the effects to be larger. The results also hold for specific subgroups that previous research has found to also be especially affected by remote instruction. Our findings indicate that using remote instruction in Sweden in the same way it was conducted during the pandemic may not have a negative effect on students' learning. This could present remote instruction as a reasonable alternative to in-person instruction in Sweden, especially in a crisis.

Another main contribution of this thesis is that it highlights the need for further research on this topic. Our results contrast much of the previous research on this topic. Considering the large effects that remote instruction during the pandemic has had on learning in other countries, it is worth examining what the underlying mechanisms for this difference in effects may be. Determining the mechanisms would be instructive for policy decisions in different countries on the uses of remote instruction. Moreover, it would be valuable to continue studying the effects of remote instruction on learning in Sweden and to expand the analysis to cover more subjects. One potential avenue for future research could be to evaluate the effects of a newly introduced measure, where certain lower secondary students with high absence are allowed to receive remote instruction instead of in-person learning (SVT Nyheter, 2024). Future research in Sweden could potentially also use data collected through a recently proposed method, where Statistics Sweden (SCB) would be able to connect directly to individual schools' IT-systems,

enabling access to more data (Regeringen och Regeringskansliet, 2024). When a longer time period has passed since the pandemic, it would also be valuable to examine the effect of the school closures on long term measures, such as students' future incomes.

Another more subtle contribution of this paper is the presentation of pre-pandemic Difference-in-Differences estimates where the parallel trends assumption does not appear to hold. This is rarely done in Difference-in-Differences papers, which is a problem, since only publishing results where the pre-trends are not statistically significant may cause selection bias (Roth, 2021). While inconclusive results may not appear important for an individual study, it is valuable for the field as a whole.

8.3 Limitations

8.3.1 Generalizability of Results

The decisions made regarding our data, which resulted in the final sample, have restricted the interpretation of our results. Firstly, we excluded schools with fewer than 50 students due to a large amount of missing data among small schools. Secondly, the results are mostly applicable to academic programs. Third, and lastly, our findings are based on a sample of 94 out of 290 Swedish municipalities, primarily those containing larger cities. This reduced sample is partly due to the exclusion of municipalities that only have lower secondary schools, which stems from the propensity score matching, which required that upper secondary schools be matched with lower secondary schools within the same municipality. It is also due to the general exclusion of smaller municipalities because of missing data. However, the municipalities span the entirety of Sweden, with Gotlands län being the only county not included in the sample. As a result, our findings are most representative of schools that are not especially small and of municipalities that include medium-sized to large cities, and of students enrolled in academic programs.

8.3.2 Failure Rates as a Measure of Learning

One issue with studying the effects of the pandemic is that different data collection programmes have been negatively affected by the pandemic (United Nations, 2020). This has limited the data available for this thesis through the delay of PISA tests (Skolverket, 2023) and cancellation of national exams (Skolverket, 2022a). We also did not use average grade points as an outcome variable due to data limitations, although these limitations were not a result of the pandemic. For grade 9 students, the Swedish National Agency for Education reports average grade points based on the grades received in all courses attended in grade 9 (Skolverket, 2024a). However, for upper secondary students, the Swedish National Agency for Education instead reports a measure of average grade points based on the grades received in all courses attended in all three years of upper secondary school (Skolverket, 2024b). Therefore, while students in grade

10 receive final grades, we are unable to access data on these grades in isolation. Consequently, using average grade points as the outcome would make it difficult to isolate the effect of remote instruction through a Difference-in-Differences strategy and would also increase the age difference between the control and the treatment group. As a result of these data limitations, this thesis has exclusively studied one measure of academic achievement, when multiple measures would have been more instructive.

Failure rates – the learning measure used in this thesis – also has a number of limitations. In Sweden, teachers can set grades without any calibration or standardization (Bergling, 2024). Failure rates are thereby not objective measures, but rather a measure that may vary depending on which teacher is grading. This may be especially inappropriate for a country where education has been marketized (Bergling, 2024). Consequently, there is a risk that failure rates do not only measure learning but also differences in grading practices between different teachers. There is a risk that the pandemic may have changed how teachers grade, which could bias our estimates. Survey data shows that many owners of lower and upper secondary schools believe that the conditions for grading were affected by the pandemic (Skolverket, 2021a). Remote education and cancellation of national exams were presented as main reasons, along with high student and teacher absence for lower secondary schools.

In order for a change in grading practices to influence our results, they would need to change differently for grade 9 and grade 10 students. Cancellations of national exams is likely to influence both groups similarly and therefore poses less of a concern. Remote instruction, along with student and teacher absence, may affect grades in either positive or negative ways. Teachers may be more lenient due to students facing different circumstances than previous students, or they may be stricter due to a perceived increased risk of cheating, or simply due to having less material to grade students on. Another potential concern would be an increased prevalence of cheating in remote education, but while cheating appears to have increased among Swedish university students (Sveriges Radio, 2021), principals of Swedish upper secondary schools have described that remote instruction does not appear to have increased cheating and that many measures have been taken to reduce the risk of cheating (Skolinspektionen, 2020).

In our setting, the main cause for concern surrounding this topic is any effect that would have a positive effect on grades of grade 10 students in the pandemic, as this could conceal the negative effects of remote instruction on learning that are expected based on the previous literature. Therefore, we aim to understand whether remote instruction during the pandemic increased leniency in grading. Previous research indicates that privately owned schools in Sweden grade more generously than publically owned schools (Hinnerich and Vlachos, 2017; Edmark and Persson, 2021). Accordingly, if remote instruction had impacted failure rates in privately owned schools but not publicly owned schools, then this could have provided some cause for concern that the results were driven by changes in grading rather than changes in learning. Our heterogeneity analysis shows that remote instruction had no effect on failure rates in Swedish for privately or publicly owned schools, reducing the risk that changes in

leniency of grading are driving the results. Still, there may remain a degree of leniency in the failure rates measure, which would cause a downward bias of the estimate.

8.3.3 Attendance During the Pandemic

Swedish principals and teachers have stated that both upper and lower secondary students have suffered learning losses in the pandemic, but for different reasons. For upper secondary students this is due to remote instruction, for lower secondary students it is due to student and teacher absence (Skolverket, 2022b).

Since in-person instruction appears to have doubled the COVID-19 infection rates of teachers (Vlachos, Hertegård and Svaleryd, 2021), it is likely that teacher absence became a larger issue for lower secondary schools during the pandemic than for upper secondary schools. Sweden has no national data on student attendance (Sjögren et al., 2021), but for lower secondary schools there is survey data on school owners' accounts of absence and for upper secondary schools there is data from the Swedish Board of Student Finance (CSN) on absence without a given cause (Skolverket, 2021b). The survey data compares how the share of students with at least 15% absence differs in the fall semesters 2019 and 2020, but may provide some indication of how attendance changed in the spring of 2020. The results show a large increase in student absence in primary school in the fall 2020. Grade 9 has the smallest increase compared to the other grades, both in percent and percentage points, but still had a large increase of around 10 percentage points. A report from the Swedish National Agency for Education also notes that there are indications that the numbers provided by the school owners may be an underestimation (Skolverket, 2021b). For upper secondary schools, CSN has data on students with at least four hours of absence without a given cause in one month. In academic year 2019/20, there were fewer students that had this number of hours of absence without a given cause compared to academic year 2018/19 (Skolverket, 2021b). It is hard to draw strong conclusions based on this, because lower secondary schools also saw similar levels of absence without a given cause in 2019 and 2020, even though the overall attendance was negatively affected. Therefore, the pandemic appears to have affected attendance of grade 9 students and their teachers negatively, but it is less clear whether grade 10 students' attendance was affected. The existing data on attendance thereby indicates that the choice of instruction may have influenced quantity of education as well as quality.

8.3.4 Infrequent Data Collection and Short Exposure to Treatment

Another data limitation is that the frequency of data collection is relatively low, with failure rates being reported annually. Meanwhile, remote instruction was only implemented during parts of the academic years 2019/20 and 2020/21. It is therefore possible that remote instruction had an effect on failure rates, but that this effect was too small to identify through this specification. This is especially relevant for Mathematics, because some upper secondary pro-

grams have this course in the fall of grade 10 rather than the entire year, which reduces the share of students that were exposed to online instruction for this subject. To address this issue, it would have been valuable to have data on which specific schools used remote instruction at which times, as this would have enabled a comparison between the specific schools that used remote instruction with those that did not. Instead, there was only data available on the use of remote instruction at an aggregate level based on grade, level of education or type of program, as presented in Table 1, resulting in a comparison between groups that overall used more or less remote instruction.

9 Conclusion

This thesis has examined how remote instruction during the COVID-19 pandemic has affected lower achieving students' learning, through an analysis of failure rates in Swedish, English and Mathematics. To isolate the effect of remote instruction, we exploit differences in recommendations for lower and upper secondary schools in Sweden. We use Propensity Score Matching to match grade 9 and grade 10 students with similar observable characteristics and we compare the evolution of failure rates for these groups with a dynamic Difference-in-Differences design. The results for English are inconclusive. For subjects Swedish and Mathematics, our results indicate that remote instruction did not have a negative effect on lower achieving students' learning, which contrasts much of the previous literature.

We discuss that physical school closures may have affected Swedish students less because Sweden was relatively well prepared for remote instruction, physical school closures were relatively short and that there may have been existing issues in the Swedish school environment that were mitigated through remote instruction. Continued peer interactions may also have been more feasible in Sweden due to less restrictive pandemic measures and parents' different opportunities to help their children with school work may have had a smaller effect in Sweden. One main limitation for identification is that failure rates are not an objective measure of learning and that grading may have changed in the pandemic.

This thesis contributes to the existing literature by highlighting that remote instruction may not always reduce learning. We therefore recommend future research to examine possible mechanisms behind our results and expand on research on the use of remote instruction in Sweden.

References

- Aalto, Müller, and Tilley (2023). *From epidemic to pandemic: Effects of the COVID-19 outbreak on high school program choices in Sweden*, p. 102346. DOI: 10.1016/j.labeco.2023.102346.
- Abadie and Imbens (2016). *Matching on the estimated propensity score*. 2, pp. 781–807. DOI: 10.3982/ecta11293.
- Agostinelli et al. (2020). *It Takes a Village: The Economics of Parenting with Neighborhood and Peer Effects*. DOI: 10.3386/w27050.
- (2022). *When the great equalizer shuts down: Schools, peers, and parents in pandemic times*, p. 104574. DOI: 10.1016/j.jpubeco.2021.104574.
- Andersson, Johansson, and Waldenström (2010). *Do you want your child to have a certified teacher?* 1, pp. 65–78. DOI: 10.1016/j.econedurev.2010.07.003.
- Antagning (2022). *Grundläggande behörighet*. URL: <https://www.antagning.se/sv/betyg-och-behorighet/gymnasieskolan/gymnasieexamen-2014-och-framat/grundlaggande-behorighet/> (visited on 03/18/2025).
- Austin (2011). *An introduction to propensity score methods for reducing the effects of confounding in observational studies*. 3, pp. 399–424. DOI: 10.1080/00273171.2011.568786.
- Bergdahl and Nouri (2020). *Covid-19 and Crisis-Prompted Distance Education in Sweden*. 3, pp. 443–459. DOI: 10.1007/s10758-020-09470-6.
- Bergling (2024). *Avslöjar: F-betyg skrotas – inför nytt betygssystem*. URL: <https://www.vilarare.se/nyheter/betyg/avslojar-skrotar-f--nytt-betygssystem/> (visited on 06/12/2025).
- Bergström and Andersson (2018). *Om sambandet mellan skolstorlek och utbildningsresultat*. URL: <https://skr.se/skr/tjanster/rapporterochskrifter/publikationer/omsambandetmellanskolstorlekoctutbildningsresultat.65328.html>.
- Bethäuser, Bach-Mortensen, and Engzell (2023). *A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic*. 3, pp. 375–385. DOI: 10.1038/s41562-022-01506-4.
- Bettinger et al. (2017). *Virtual Classrooms: How Online College Courses Affect Student Success*. 9, pp. 2855–2875. DOI: 10.1257/aer.20151193.

- Birkelund and Karlson (2022). *No evidence of a major learning slide 14 months into the COVID-19 pandemic in Denmark*. 3, pp. 468–488. DOI: 10.1080/14616696.2022.2129085.
- Björkegren, Svaleryd, and Vlachos (2024). *Remote Instruction and Student Mental Health: Swedish Evidence from the Pandemic*. 4, pp. 491–517. DOI: 10.1257/pol.20220699.
- Bohlmark and Lindahl (2008). *Does School Privatization Improve Educational Achievement? Evidence from Sweden's Voucher Reform*. DOI: 10.2139/ssrn.1267832.
- Caliendo and Kopeinig (2008). *Some Practical Guidance for the Implementation of Propensity Score Matching*. 1, pp. 31–72. DOI: 10.1111/j.1467-6419.2007.00527.x.
- Callaway and Sant'Anna (2020). *Difference-in-Differences with multiple time periods*. 2, pp. 200–230. DOI: 10.1016/j.jeconom.2020.12.001.
- Cameron and Miller (2015). *A Practitioner's Guide to Cluster-Robust Inference*. 2, pp. 317–372. DOI: 10.3368/jhr.50.2.317.
- Cunningham (2021). *Causal Inference: The Mixtape*. URL: https://mixtape.scunning.com/09-difference_in_differences (visited on 06/12/2025).
- Edmark, Frölich, and Wondratschek (2014). *Sweden's school choice reform and equality of opportunity*, pp. 129–142. DOI: 10.1016/j.labeco.2014.04.008.
- Edmark and Persson (2021). *The impact of attending an independent upper secondary school: Evidence from Sweden using school ranking data*, p. 102148. DOI: 10.1016/j.econedurev.2021.102148.
- Engzell, Frey, and Verhagen (2021). *Learning loss due to school closures during the COVID-19 pandemic*. 17. DOI: 10.1073/pnas.2022376118.
- Escueta et al. (2020). *Upgrading Education with Technology: Insights from Experimental Research*. 4, pp. 897–996. DOI: 10.1257/jel.20191507.
- European Centre for Disease Prevention and Control and Folkhälsomyndigheten (2023). *The Swedish advice-making process for distance learning in schools, November 2020–April 2021*. URL: <https://www.ecdc.europa.eu/en/publications-data/covid-19-pandemic-after-action-review-schools-sweden> (visited on 05/02/2025).
- European Commission: Joint Research Centre and Di Pietro (2023). *The impact of COVID-19 physical school closure on student performance in OECD Countries – A meta-analysis*. URL: <https://data.europa.eu/doi/10.2760/197242>.
- Figlio, Rush, and Yin (2013). *Is it live or is it Internet? Experimental estimates of the effects of online instruction on student learning*. 4, pp. 763–784. DOI: 10.1086/669930.
- Finansdepartementet (2020). *Tillskott till kommunsektorn med anledning av Covid-19*. URL: <https://www.regeringen.se/artiklar/2020/04/om-krispaketet-till-kommunsektorn/> (visited on 06/04/2025).
- Florida and Mellander (2021). *The geography of COVID-19 in Sweden*. 1, pp. 125–150. DOI: 10.1007/s00168-021-01071-0.

- Folkhälsomyndigheten (2023). *Folkhälsa och barns bästa vägledande i skolbeslut under pandemin*. URL: <https://www.folkhalsomyndigheten.se/nyheter-och-press/nyhetsarkiv/2023/oktober/folkhalsa-och-barns-basta-vagledande-i-skolbeslut-under-pandemin/> (visited on 03/22/2025).
- (2024). *När hände vad under pandemin?* URL: <https://www.folkhalsomyndigheten.se/smittskydd-beredskap/utbrott/utbrottsarkiv/covid-19-pandemin-2019-2023/nar-hande-vad-under-pandemin/> (visited on 03/22/2025).
- Folkhälsomyndigheten and Institutet för hälsa och välfärd (2020). *Covid-19 hos förskole- och skolbarn – En jämförelse mellan Finland och Sverige*. URL: <https://www.folkhalsomyndigheten.se/contentassets/1249ec40eb2345d4a4b2416d8c0239de/covid-19-forskole-och-skolbarn.pdf>.
- Fuchs-Schündeln et al. (2022). *The Long-Term Distributional and welfare Effects of COVID-19 school closures*. 645, pp. 1647–1683. DOI: 10.1093/ej/ueac028.
- Giota et al. (2024). *Psykisk hälsa, skolprestationer och Covid-19 – långsiktiga konsekvenser och skyddsfaktorer*. URL: <https://www.gu.se/forskning/psykisk-halsa-skolprestationer-och-covid-19-langsiktiga-konsekvenser-och-skyddsfaktorer> (visited on 02/24/2025).
- Goldhaber et al. (2023). *The Educational Consequences of Remote and Hybrid Instruction during the Pandemic*. 3, pp. 377–392. DOI: 10.1257/aeri.20220180.
- Grätz and Lipps (2020). *Large loss in studying time during the closure of schools in Switzerland in 2020*, p. 100554. DOI: 10.1016/j.rssm.2020.100554.
- Hall, Hardoy, and Lundin (2022). *Schooling in the Nordic countries during the COVID-19 pandemic*. URL: <https://www.ifau.se/en/Research/Publications/Working-papers/2022/schooling-in-the-nordic-countries-during-the-covid-19-pandemic/>.
- Hall and Lundin (2024). *Technology in the classroom: Personal computers and learning outcomes in primary school*, p. 102536. DOI: 10.1016/j.econedurev.2024.102536.
- Hall, Lundin, and Sibbmark (2021). *A laptop for every child? The impact of technology on human capital formation*, p. 101957. DOI: 10.1016/j.labeco.2020.101957.
- Hallin et al. (2022). *No learning loss in Sweden during the pandemic: Evidence from primary school reading assessments*, p. 102011. DOI: 10.1016/j.ijer.2022.102011.
- Halloran et al. (2021). *Pandemic Schooling Mode and Student Test Scores: Evidence from US States*. URL: <https://ideas.repec.org/p/nbr/nberwo/29497.html>.
- Heppen, Sorensen, Allensworth, Walters, Rickles, et al. (2016). *The Struggle to Pass Algebra: Online vs. Face-to-Face Credit Recovery for At-Risk Urban Students*. 2, pp. 272–296. DOI: 10.1080/19345747.2016.1168500.
- Heppen, Sorensen, Allensworth, Walters, Stachel, et al. (2012). “Efficacy of Online Algebra I for Credit Recovery for At-Risk Ninth Graders: Consistency of Results from Two Cohorts.”

- In: Society for Research on Educational Effectiveness (SREE) Spring 2012 Conference, Washington, DC. URL: <https://eric.ed.gov/?id=ED562703>.
- Hinnerich and Vlachos (2017). *The impact of upper-secondary voucher school attendance on student achievement. Swedish evidence using external and internal evaluations*, pp. 1–14. DOI: 10.1016/j.labeco.2017.03.009.
- Holmlund, Böhlmark, and Lindahl (2015). *Skolsegregation och skolval*. URL: <https://www.ifau.se/Forskning/Publikationer/Rapporter/2015/Skolsegregation-och-skolval/>.
- Holmlund, Sjögren, and Öckert (2020). *Jämlikhet i möjligheter och utfall i den svenska skolan*. URL: <https://www.ifau.se/Forskning/Publikationer/Rapporter/2020/jamlikhet-i-mojligheter-och-utfall-i-den-svenska-skolan/>.
- Jakubowski, Gajderowicz, and Patrinos (2025). *COVID-19, school closures, and student learning outcomes. New global evidence from PISA*. 1. DOI: 10.1038/s41539-025-00297-3.
- Mack, Su, and Westreich (2018). *Types of missing data*. URL: <https://www.ncbi.nlm.nih.gov/books/NBK493614> (visited on 04/01/2025).
- Miller (2023). *An Introductory guide to event study models*. 2, pp. 203–230. DOI: 10.1257/jep.37.2.203.
- Our World in Data (2020). *School closures during the COVID-19 pandemic*. URL: <https://ourworldindata.org/grapher/school-closures-covid?time=2020-03-22> (visited on 04/06/2025).
- Rambachan and Roth (2023). *A more credible approach to parallel trends*. 5, pp. 2555–2591. DOI: 10.1093/restud/rdad018.
- Regeringen (2014). *Möjligheter till fjärrundervisning*. URL: <https://www.regeringen.se/rattsliga-dokument/proposition/2014/12/prop.-20141544> (visited on 05/12/2025).
- (2021). *Summary Sweden in the Pandemic*. URL: <https://www.regeringen.se/rattsliga-dokument/statens-offentliga-utredningar/2022/02/sou-202210/> (visited on 03/08/2025).
- Regeringen and Regeringskansliet (2020). *En mer likvärdig skola – minskad skolsegregation och förbättrad resurstilldelning*. URL: <https://www.regeringen.se/rattsliga-dokument/statens-offentliga-utredningar/2020/04/sou-202028/> (visited on 02/18/2025).
- (2024). *SCB ska samla in registerdata från skolornas IT-system*. URL: <https://www.regeringen.se/pressmeddelanden/2024/12/scb-ska-samla-in-registerdata-fran-skolornas-it-system/> (visited on 02/18/2025).
- Ringer and Kreitz-Sandberg (2022). *Swedish Pupils’ Perspectives on Emergency Remote Teaching during COVID-19 - A Qualitative Study*, p. 100167. DOI: 10.1016/j.ijedro.2022.100167.

- Rosenbaum and Rubin (1983). *The central role of the propensity score in observational studies for causal effects*. 1, pp. 41–55. DOI: 10.1093/biomet/70.1.41.
- Roth (2021). *Jonathan Roth "Testing and Sensitivity Analysis for Parallel Trends"*. URL: <https://www.youtube.com/watch?v=F8C1xaPoRvM>.
- Roth et al. (2023). *What's trending in difference-in-differences? A synthesis of the recent econometrics literature*. 2, pp. 2218–2244. DOI: 10.1016/j.jeconom.2023.03.008.
- Rubin (1973). *Matching to remove bias in observational studies*. 1, pp. 159–183. DOI: 10.2307/2529684.
- Sjögren et al. (2021). *Swedish children and youth during the COVID-19 pandemic*. URL: <https://www.ifau.se/en/Research/Publications/Working-papers/2021/swedish-children-and-youth-during-the-covid-19-pandemic/>.
- Skar, Graham, and Huebner (2021). *Learning loss during the COVID-19 pandemic and the impact of emergency remote instruction on first grade students' writing: A natural experiment*. 7, pp. 1553–1566. DOI: 10.1037/edu0000701.
- (2023). *The Long-Term Effects of the COVID-19 Pandemic on Children's Writing: a Follow-up Replication Study*. 1. DOI: 10.1007/s10648-023-09729-1.
- Skolinspektionen (2020). *Gymnasieskolors distansundervisning under covid-19 pandemin*. URL: <https://www.skolinspektionen.se/globalassets/02-beslut-rapporter-stat/granskningsrapporter/ovriga-publikationer/2020/covid-19/uppfoljning-av-gymnasieskolors-distansundervisning---slutrapport.pdf>.
- Skolverket (2020). *Skolverkets lägesbedömning 2020*. URL: <https://www.skolverket.se/sok-publikationer/publikationsserier/rapporter/2020/skolverkets-lagesbedomning-2020?id=6436>.
- (2021a). *Covid-19-pandemins påverkan på skolväsendet - Delredovisning 4*. URL: <https://www.skolverket.se/publikationer?id=9007>.
- (2021b). *Nationell kartläggning av elevfrånvaro*. URL: <https://www.skolverket.se/publikationer?id=8573>.
- (2022a). *Betyg- och studieresultat i gymnasieskolan år 2021*. URL: <https://www.skolverket.se/publikationer?id=8892>.
- (2022b). *Covid-19-pandemins konsekvenser för skolväsendet – slutredovisning*. URL: <https://www.skolverket.se/publikationer?id=10589>.
- (2022c). *Covid-19-pandemins påverkan på skolväsendet - Delredovisning 5*. URL: <https://www.skolverket.se/sok-publikationer/publikationsserier/regeringsuppdrag/2022/covid-19-pandemins-paverkan-pa-skolvasendet-delredovisning-5?id=9732>.
- (2023). *PISA 2022*. URL: <https://www.skolverket.se/publikationer?id=12177>.
- (2024a). *Sätta betyg i grundskolan*. URL: <https://www.skolverket.se/undervisning/grundskolan/betyg-i-grundskolan/satta-betyg-i-grundskolan> (visited on 05/07/2025).

- Skolverket (2024b). *Sätta betyg i gymnasieskolan*. URL: <https://www.skolverket.se/undervisning/gymnasieskolan/betyg-i-gymnasieskolan/satta-betyg-i-gymnasieskolan> (visited on 05/06/2025).
- (2024c). *TIMSS 2023*. URL: <https://www.skolverket.se/publikationer?id=13178>.
- (2025a). *Behörighetsregler och meritvärde - Utbildningsguiden*. URL: <https://utbildningsguiden.skolverket.se/gymnasieskolan/behorighet-och-meritvarde/behorighetsregler-och-meritvarde> (visited on 03/06/2025).
- (2025b). *Om gymnasieskolan - Utbildningsguiden*. URL: <https://utbildningsguiden.skolverket.se/gymnasieskolan/om-gymnasieskolan> (visited on 03/06/2025).
- (2025c). *Om högskoleförberedande program - Utbildningsguiden*. URL: <https://utbildningsguiden.skolverket.se/gymnasieskolan/gymnasieskolans-program/om-hogskoleforberedande-program> (visited on 03/18/2025).
- (2025d). *Om Skolverkets statistik*. URL: <https://www.skolverket.se/skolutveckling/statistik/om-skolverkets-statistik> (visited on 05/12/2025).
- (2025e). *Grundskolan - utbildningsguiden*. URL: <https://utbildningsguiden.skolverket.se/languages/svenska/grundskolan> (visited on 03/06/2025).
- Skolverket and Sveriges Officiella Statistik (2024). *Elever i gymnasieskolan Läsåret 2023/24*. URL: <https://www.skolverket.se/download/18.3531f8d618df22bffb7c/1709217798857/pdf12565.pdf>.
- Statistikmyndigheten SCB (2020). *Fler grundskollärare än någonsin – men lägre andel med examen*. URL: <https://www.scb.se/hitta-statistik/artiklar/2020/fler-grundskollarare-an-nagonsin--men-farre-med-examen/> (visited on 03/05/2025).
- (2021). *Boendesituation bland gymnasieelever*. URL: https://www.scb.se/contentassets/4cac5bbb01fd44e691c79da85569fd9c/uf0549_2020a01_br_a40br2101.pdf.
- (2022). *Analyser och statistik om befolkningens utbildning*. URL: <https://www.scb.se/hitta-statistik/statistik-efter-amne/utbildning-samt-forskning-inom-hogskolan/befolkningens-utbildning-och-studiedeltagande/analyser-och-statistik-om-befolkningens-utbildning/> (visited on 05/06/2025).
- Stockholms Universitet (2020). *Ny studie: Så påverkades lärandet av distansundervisningen under pandemin - Institutionen för pedagogik och didaktik*. URL: <https://www.su.se/institutionen-for-pedagogik-och-didaktik/nyheter/ny-studie-s%C3%A5-p%C3%A5verkades-l%C3%A4randet-av-distansundervisningen-under-pandemin-1.631430> (visited on 02/13/2025).
- Stuart et al. (2014). *Using propensity scores in difference-in-differences models to estimate the effects of a policy change*. 4, pp. 166–182. DOI: 10.1007/s10742-014-0123-z.
- Sveriges Radio (2021). *Fusk bland studenter ökade under pandemin*. URL: <https://www.sverigesradio.se/artikel/fusk-bland-studenter-okade-under-pandemin> (visited on 04/10/2025).

- Sveriges Riksdag (2010). *Skollag (2010:800)*. URL: https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/skollag-2010800_sfs-2010-800/#K2 (visited on 04/26/2025).
- SVT Nyheter (2024). *Nu kan de första hemmasittarna i landet få distansundervisning*. URL: <https://www.svt.se/nyheter/lokalt/orebro/nu-kan-de-forsta-hemmasittarna-i-landet-fa-distansundervisning> (visited on 04/10/2025).
- Turke (2020). *Five reasons COVID-19 is less severe in younger age-groups*. 1, pp. 113–117. DOI: 10.1093/emph/eoaa050.
- U.S. Centers for Disease Control and Prevention (2024). *CDC Museum COVID-19 timeline*. URL: <https://www.cdc.gov/museum/timeline/covid19.html> (visited on 05/03/2025).
- UNESCO (2020). *1.37 billion students now home as COVID-19 school closures expand, ministers scale up multimedia approaches to ensure learning continuity*. URL: <https://www.unesco.org/en/articles/137-billion-students-now-home-covid-19-school-closures-expand-ministers-scale-multimedia-approaches> (visited on 04/12/2025).
- (2021). *One year into COVID-19 Education disruption: Where do we stand?* URL: <https://www.unesco.org/en/articles/one-year-covid-19-education-disruption-where-do-we-stand> (visited on 04/14/2025).
- United Nations (2019). *The Sustainable Development Goals Report 2019*. URL: <https://unstats.un.org/sdgs/report/2019/The-Sustainable-Development-Goals-Report-2019.pdf>.
- (2020). *The Sustainable Development Goals Report 2020*. URL: <https://unstats.un.org/sdgs/report/2020/The-Sustainable-Development-Goals-Report-2020.pdf>.
- (2024). *The Sustainable Development Goals Report 2024*. URL: <https://unstats.un.org/sdgs/report/2024/>.
- (2025). *Sustainable Development Goals Indicators*. URL: <https://unstats.un.org/sdgs/report/2021/goal-04/> (visited on 04/21/2025).
- Universitets- och högskolerådet (2025). *Meriter och meritvärde*. URL: <https://www.studera.nu/sa-kommer-du-in/anmalan-och-antagning/meritvarde/> (visited on 03/22/2025).
- University Admissions (2022). *The Swedish higher education system*. URL: <https://www.universityadmissions.se/en/find-out-more/the-swedish-higher-education-system/> (visited on 03/18/2025).
- Vlachos, Hertegård, and Svaleryd (2021). *The effects of school closures on SARS-CoV-2 among parents and teachers*. 9. DOI: 10.1073/pnas.2020834118.
- von Kügelgen (2023). *Höga ambitioner och goda resultat – många unga med utländsk bakgrund klarar sig bra i den svenska skolan | Integration Norden*. URL: <https://www.integrationnorden.se/>

//nordicwelfare.org/integration-norden/en/nyheter/hoga-ambitioner-och-goda-resultat-manga-unga-med-utlandsk-bakgrund-klarar-sig-bra-i-den-svenska-skolan/ (visited on 05/01/2025).

Werner and Woessmann (2023). *The legacy of Covid-19 in education*. 115, pp. 609–668. DOI: 10.1093/epolic/eiad016.

Wing, Simon, and Bello-Gomez (2018). *Designing Difference in Difference Studies: Best Practices for Public Health Policy Research*. 1, pp. 453–469. DOI: 10.1146/annurev-publhealth-040617-013507.

World Bank Group (2025). *Remote Learning During COVID-19: Lessons from Today, Principles for Tomorrow*. URL: <http://documents.worldbank.org/curated/en/160271637074230077>.

World Health Organisation (2024). *World Health Statistics 2024*. URL: <https://iris.who.int/bitstream/handle/10665/376869/9789240094703-eng.pdf?sequence=1>.

Xu and Wang (2023). *High prevalence of anxiety, depression, and stress among remote learning students during the COVID-19 pandemic: Evidence from a meta-analysis*. DOI: 10.3389/fpsyg.2022.1103925. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC9871576/>.

Zeldow and Hatfield (2021). *Confounding and regression adjustment in difference-in-differences studies*. 5, pp. 932–941. DOI: 10.1111/1475-6773.13666.

Appendix

A Data Sources and Sample

Table 6 lists all data files used in constructing the dataset. Each export had to be repeated for every academic year from 2014/15 to 2020/21. Then, for each educational level and statistical area, the data had to be downloaded for each academic year and then appended across years to construct separate panel datasets for grades, students, and staff. These panel datasets were then merged to create a panel datasets for upper- and lower secondary schools. Finally, the two educational levels were appended to end up with a combined panel dataset. This sample is referred to "Full Sample" in Table 8.

Table 6 – Required Data Files

Level of Education	Level of Aggregation	Statistical Domain	Export File Name
Upper Secondary	School Level	Grades	Graduating students, all programs
		Students	Number of students who enrolled, by program
		Teachers	Teachers with licensing and qualified certification
Lower Secondary		Grades	Final grades by subject Final grades, all students
		Students	Student enrollment by grade
		Teachers	Teachers with licensing and qualified certification

Notes: This table lists the relevant datasets required to construct our panel dataset. Each entry corresponds to a specific combination of level of education level, level of aggregation, statistical domain, and the name of the file to be exported. Additionally, Municipality ID was added using data from OpenDataSoft. Complete source code is available on GitHub upon request.

In the data provided by the Swedish National Agency for Education, when outcome statistics are entirely missing, a single dot (“.”) is used. Additionally, outcome statistics based on fewer than 10 students are marked with a double dot (“..”) to protect the confidentiality of student identities. This confidentiality-based suppression is determined by student group size and the composition of these students. As a result, the missingness mechanism can be classified as missing not at random (MNAR), since statistics on schools who have fewer students are more likely to be suppressed. By looking at Table 7, it is evident that there is a significant amount of missing values within schools with less than 50 students. For schools with 50 or more students, the issue is not as widespread.

Table 7 – Number of Observations and Share of Missing Values by School Size Group

School Size	Observations	Female Students	Foreign Background	Postsecondary Parents	Certified Teachers	Swedish Failure	English Failure	Math Failure
1–49	1139	0.60	0.84	0.52	0.10	0.78	0.73	0.74
50–99	1296	0.09	0.32	0.09	0.09	0.29	0.22	0.23
100–199	3341	0.03	0.13	0.04	0.09	0.07	0.04	0.04
200–299	4007	0.02	0.05	0.03	0.04	0.03	0.01	0.01
300–399	2833	0.01	0.02	0.01	0.02	0.03	0.01	0.01
400–499	1819	0.01	0.02	0.01	0.04	0.03	0.00	0.00
500–599	1024	0.02	0.02	0.02	0.08	0.02	0.00	0.00
600–699	509	0.01	0.01	0.01	0.07	0.01	0.00	0.00
700–799	319	0.01	0.01	0.01	0.14	0.01	0.00	0.00
800–899	203	0.00	0.00	0.00	0.15	0.00	0.00	0.00
900+	267	0.02	0.02	0.02	0.20	0.00	0.00	0.00

Notes: This table presents the number of observations across all academic years and the proportion of missing values for the selected variables, grouped by school size. The share of missing values is computed within each school size category, which is defined as total student enrollment per school. The size categorization follows the grouping used in Bergström and Andersson (2018).

MNAR might or might not pose a challenge for causal identification. Whether or not it does depends on whether the missingness is independent of the outcomes variable or not (Mack, Su and Westreich, 2018). It is not possible to determine whether smaller schools systematically have either higher or lower failure rates based on the schools with fewer students that are left in the sample. The small schools that are left in the in the full sample are likely to be schools with a homogenous composition of students, e.g. failure rates of zero or no one with foreign background, because of the confidentiality-based suppression threshold. In other words, small schools that have a heterogenous composition are more likely to have missing values, since there is a higher likelihood that an outcome statistic is based on fewer than 10 students. Thus, the resulting bias from the confidentiality-based suppression makes it not possible for us to determine whether the missingness in schools with fewer students than 50 is independent of the outcomes or not. Therefore, for the main analysis we exclude all schools that have fewer than 49 students, and focus on the effect on schools with 50 or more students.

Even after excluding all schools that have fewer than 50 students, there are still schools that have missing values. Consequently, the schools that have at least one missing value are excluded from the sample. As a last step, we only keep schools that show up in all years, which implies that new schools that opened up or schools that closed down in the any academic year are excluded, as well as schools that fail to report at least one entry in any given area in any given academic year.

It is important to mention that the balanced panel dataset consists of schools that are diligent in reporting to Statistics Sweden. This, in itself, might be introducing bias, since schools that are more consistent in reporting could be systematically different from those schools that report less consistently. Nonetheless, there is little we are able to do to guard against this.

For our sample used in the main analysis, we employed Propensity Score Matching to account for compositional differences between untreated lower secondary schools and treated upper secondary schools. "After PSM" refers to the sample after Propensity Score Matching and it is discussed in Section 5.

In Table 8 we present how our data management decisions impacted the number of unique schools in each step. Also, following Propensity Score Matching – in the "After PSM" Sample – 94 of the 290 Swedish municipalities remain. The 94 municipalities listed represent 20 out of the 21 counties in Sweden, with Gotlands län being the only one not included. Here is the list of municipalities represented in the "After PSM" Sample: Alingsås, Älvsbyn, Arvika, Bollnäs, Borås, Borlänge, Botkyrka, Danderyd, Degerfors, Enköping, Eskilstuna, Eslöv, Falkenberg, Falköping, Falun, Gävle, Gislaved, Gnosjö, Göteborg, Håbo, Hagfors, Halmstad, Haninge, Haparanda, Härryda, Hässleholm, Hedemora, Helsingborg, Hudinge, Hudiksvall, Järfälla, Jönköping, Kalmar, Karlshamn, Karlskrona, Kristianstad, Kungälv, Leksand, Lerum, Lidingö, Lidköping, Linköping, Ljungby, Ljusdal, Ludvika, Luleå, Lund, Lysekil, Malmö, Markaryd, Mjölby, Motala, Nacka, Nässjö, Norrköping, Nyköping, Nynäshamn, Örebro, Örensköldsvik, Österåker, Östersund, Övertorneå, Piteå, Ronneby, Salem, Sigtuna, Skövde, Söderhamn, Södertälje, Solna, Stockholm, Strängnäs, Sundsvall, Täby, Tierp, Timrå, Tingsryd, Torsås, Torsby, Tranås, Trollhättan, Uddevalla, Ulricehamn, Umeå, Upplands Väsby, Upplands-Bro, Uppsala, Vallentuna, Vänersborg, Varberg, Värnamo, Västerås, Växjö, and Vellinge.

Table 8 – Number of Unique Schools Across Sample Restrictions

Level of Education	Type of Ownership	Sample Restrictions				
		Full Sample	Size ≥ 50	No Missing	Balanced Panel	After PSM
Grade 9	Public	1576	1505	1413	659	150
	Private	578	505	373	179	91
Grade 10	Public	668	637	527	223	150
	Private	331	289	257	142	91
Total		3153	2936	2570	1203	482

Notes: This table presents the number of unique schools across sample restrictions. "Full Sample" refers to the entire dataset before any exclusions. "Size ≥ 50 " excludes schools with fewer than 50 students. "No Missing" excludes observations with missing values. "Balanced Panel" keeps schools only observed in all academic years from 2014/15 to 2020/21. "After PSM" refers to the matched sample after propensity score matching.

To construct Figure 1, in the statistics section of the Swedish National Agency for Education choose upper secondary schools and lower secondary schools, respectively. The level of aggregation should be set to the national level. The statistical domain is grades and select comparison of average grade points over time. When in the files, for upper secondary schools the values corresponds to the average grade points of graduating students in academic programs only. For lower secondary, the values corresponds to average grade points based on 17 subjects, which are then divided by 17 to approximate the average grade points in upper secondary. Please be aware that the instructions are translations from Swedish to English.

B Covariate Balance After Propensity Score Matching

Table 9 – Covariate Imbalance Before and After Matching (Excluding Exact-Match Variables)

Covariate	Before Matching					After Matching				
	Mean Treat	Mean Control	Std. Diff	Std Dev. Treat	Std Dev. Control	Mean Treat	Mean Control	Std. Diff	Std Dev. Treat	Std Dev. Control
Distance	0.378	0.271	0.699	0.153	0.135	0.369	0.328	0.270	0.140	0.139
Active Certified Teachers	72.041	65.522	0.430	15.165	13.540	71.239	65.724	0.364	13.349	15.532
Postsecondary Parents	60.852	54.517	0.539	11.760	16.013	60.938	60.539	0.034	12.089	16.904
Foreign Background	22.137	19.393	0.187	14.679	14.683	22.166	22.083	0.006	15.569	16.000
Private Ownership	0.378	0.378	0.000	0.486	0.486	0.378	0.378	0.000	0.486	0.486

Notes: This table presents covariate means and standard deviations for treated and control groups before and after matching, excluding variables exact-matched on. "Std. Diff" refers to standardized mean differences.

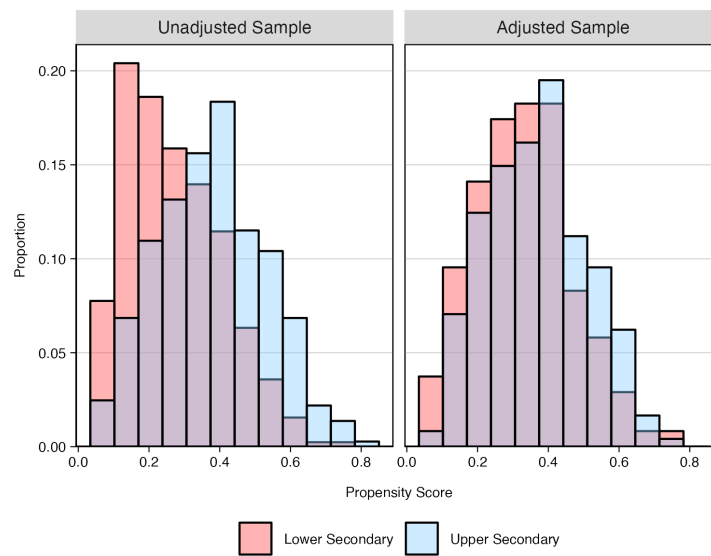


Figure 10 – Propensity Score Distributions Before and After Matching

Notes: Figure 10 shows the distribution of estimated propensity scores for upper secondary schools and lower secondary schools in the unadjusted sample before matching and the adjusted sample after matching.

C List of Tables for the Results Section

Table 10 – Dynamic Difference-in-Differences Models, Subject Swedish

	Failure Rates in Subject Swedish				
	(1)	(2)	(3)	(4)	(5)
2014/15 × Grade 10	0.502 (0.365)	0.546 (0.364)	0.556 (0.368)	0.559 (0.367)	0.559 (0.367)
2015/16 × Grade 10	0.538 (0.336)	0.571* (0.334)	0.575* (0.337)	0.581* (0.337)	0.582* (0.337)
2016/17 × Grade 10	0.300 (0.334)	0.331 (0.331)	0.331 (0.331)	0.336 (0.331)	0.337 (0.331)
2017/18 × Grade 10	0.216 (0.287)	0.228 (0.286)	0.230 (0.287)	0.231 (0.287)	0.231 (0.287)
(Reference Year)	0	0	0	0	0
2019/20 × Grade 10	−0.081 (0.304)	−0.079 (0.304)	−0.081 (0.302)	−0.065 (0.302)	−0.066 (0.302)
2020/21 × Grade 10	−0.297 (0.334)	−0.300 (0.335)	−0.300 (0.335)	−0.273 (0.335)	−0.275 (0.335)
Foreign Background		0.017 (0.012)	0.017 (0.011)	0.006 (0.013)	0.006 (0.013)
Active Certified Teachers			0.003 (0.006)	0.002 (0.006)	0.002 (0.006)
Postsecondary Parents				−0.030** (0.015)	−0.030** (0.015)
Female Students					0.004 (0.010)
School Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	3,374	3,374	3,374	3,374	3,374

Notes: This table reports estimates of the effect of remote instruction on failure rates in the subject Swedish from the Dynamic Difference-in-Differences models. The reference academic year is 2018/19. Standard errors are clustered at the school level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 11 – Dynamic Difference-in-Differences Models, Subject English

	Failure Rates in Subject English				
	(1)	(2)	(3)	(4)	(5)
2014/15 × Grade 10	0.831* (0.450)	1.076** (0.454)	1.058** (0.454)	1.068** (0.450)	1.068** (0.450)
2015/16 × Grade 10	1.226*** (0.366)	1.409*** (0.368)	1.401*** (0.369)	1.418*** (0.369)	1.421*** (0.370)
2016/17 × Grade 10	−1.076*** (0.393)	−0.899** (0.387)	−0.899** (0.387)	−0.884** (0.386)	−0.883** (0.386)
2017/18 × Grade 10	−0.764** (0.336)	−0.697** (0.335)	−0.701** (0.335)	−0.700** (0.336)	−0.698** (0.337)
(Reference Year)	0	0	0	0	0
2019/20 × Grade 10	0.220 (0.326)	0.232 (0.327)	0.236 (0.327)	0.281 (0.327)	0.278 (0.327)
2020/21 × Grade 10	0.107 (0.366)	0.088 (0.367)	0.087 (0.367)	0.161 (0.369)	0.155 (0.368)
Foreign Background		0.093*** (0.018)	0.092*** (0.018)	0.061*** (0.020)	0.062*** (0.020)
Active Certified Teachers			−0.005 (0.007)	−0.005 (0.007)	−0.005 (0.007)
Postsecondary Parents				−0.084*** (0.022)	−0.084*** (0.022)
Female Students					0.011 (0.013)
School Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	3,374	3,374	3,374	3,374	3,374

Notes: This table reports estimates of the effect of remote instruction on failure rates in the subject English from the Dynamic Difference-in-Differences models. The reference academic year is 2018/19. Standard errors are clustered at the school level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 12 – Dynamic Difference-in-Differences Models, Subject Mathematics

	Failure Rates in Subject Mathematics				
	(1)	(2)	(3)	(4)	(5)
2014/15 × Grade 10	1.489** (0.594)	1.708*** (0.597)	1.663*** (0.598)	1.666*** (0.597)	1.666*** (0.598)
2015/16 × Grade 10	2.969*** (0.448)	3.133*** (0.447)	3.113*** (0.444)	3.119*** (0.445)	3.123*** (0.444)
2016/17 × Grade 10	−0.274 (0.461)	−0.116 (0.458)	−0.115 (0.458)	−0.110 (0.459)	−0.109 (0.458)
2017/18 × Grade 10	0.583 (0.456)	0.643 (0.457)	0.632 (0.458)	0.633 (0.458)	0.635 (0.458)
(Reference Year)	0	0	0	0	0
2019/20 × Grade 10	1.244*** (0.438)	1.255*** (0.438)	1.265*** (0.439)	1.282*** (0.439)	1.277*** (0.438)
2020/21 × Grade 10	1.795*** (0.462)	1.778*** (0.462)	1.775*** (0.463)	1.802*** (0.464)	1.794*** (0.464)
Foreign Background		0.083*** (0.022)	0.082*** (0.022)	0.070*** (0.025)	0.071*** (0.025)
Active Certified Teachers			−0.012 (0.009)	−0.012 (0.009)	−0.013 (0.009)
Postsecondary Parents				−0.031 (0.025)	−0.030 (0.025)
Female Students					0.015 (0.019)
School Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	3,374	3,374	3,374	3,374	3,374

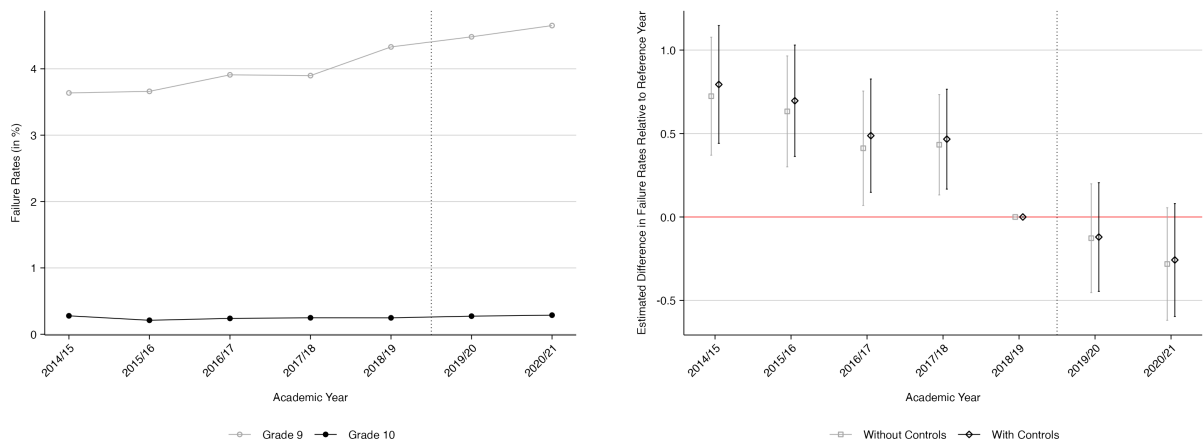
Notes: This table reports estimates of the effect of remote instruction on failure rates in the subject Mathematics from the Dynamic Difference-in-Differences models. The reference academic year is 2018/19. Standard errors are clustered at the school level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D Summary Statistics and Results Prior to Propensity Score Matching

Table 13 – Summary Statistics by Educational Level: Pre- and Post-Pandemic

	Lower Secondary		Upper Secondary	
	Pre-Pandemic	Post-Pandemic	Pre-Pandemic	Post-Pandemic
Failure Rates, Subject Swedish	3.89 (3.97)	4.57 (4.33)	0.24 (0.91)	0.28 (0.93)
Failure Rates, Subject English	7.64 (6.89)	7.62 (6.58)	0.28 (0.87)	0.38 (1.12)
Failure Rates, Subject Mathematics	9.84 (8.15)	9.56 (7.69)	0.96 (2.03)	1.19 (2.48)
Foreign Background	21.55 (15.20)	23.72 (16.34)	24.53 (16.16)	27.95 (18.47)
Active Certified Teachers	73.50 (13.45)	76.31 (12.31)	81.08 (14.31)	84.61 (12.60)
Postsecondary Parents	55.99 (16.16)	59.43 (16.05)	61.99 (12.04)	65.00 (12.85)
Female Students	48.48 (4.78)	48.74 (4.28)	52.19 (15.40)	52.66 (15.02)
Public Schools	3295	1318	1115	446
Private Schools	895	358	710	284
All Schools	4190	1676	1825	730

Notes: This table is inspired by Björkegren, Svaleryd, and Vlachos (2024). It presents averages and standard deviations (in parentheses) of school-level characteristics (expressed as percentages) before and after the Public Health Agency of Sweden's remote instruction recommendation. The pre-pandemic period covers academic years 2014/15 to 2018/19, and the post-pandemic period ranges from 2019/20 to 2020/21. Observation counts include repeated unique schools, i.e., five years in the pre-period and two years in the post-period. For a unique school count see "Balanced Panel" in Table 8. All variables are reported prior to propensity score matching.

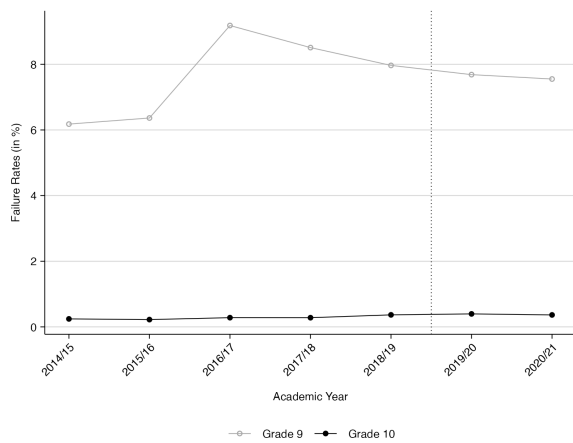


(a) Time Trends in Failure Rates

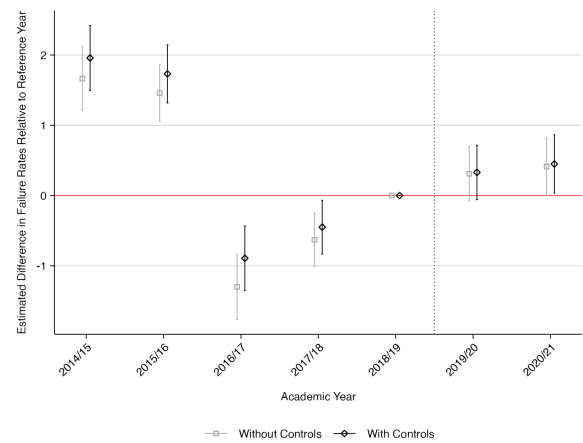
(b) Dynamic DiD Models

Figure 11 – Results for Subject Swedish Prior to PSM

Notes: (a) The figure shows the time trends in failure rates in subject Swedish by level of education for the academic years 2014/15 to 2020/21. The vertical reference line indicates the transition to remote instruction. (b) The figure shows the estimated difference in failure rates in subject Swedish between grade 9 and grade 10, relative to reference academic year 2019/20. Both models include school-level and time fixed effects. Controls are Foreign Background, Active Certified Teachers, Postsecondary Parents, and Female Students. 95 percent confidence interval are included. Standard errors are clustered at the school level.



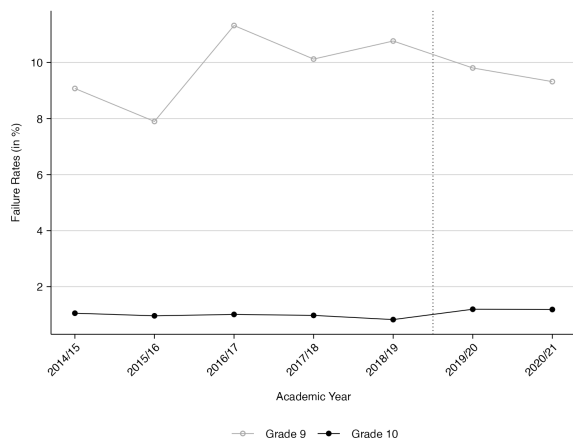
(a) Time Trends in Failure Rates



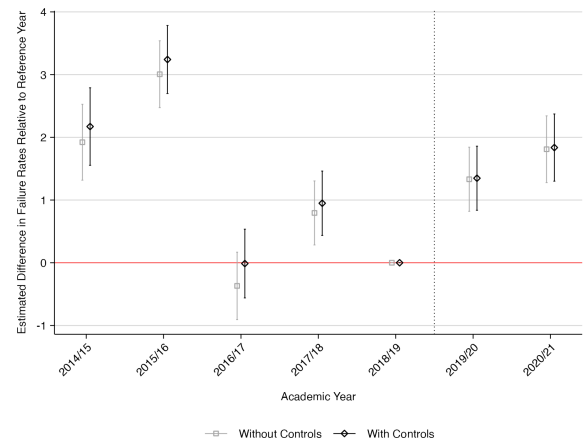
(b) Dynamic DiD Models

Figure 12 – Results for Subject English Prior to PSM

Notes: (a) The figure shows the time trends in failure rates in subject English by level of education for the academic years 2014/15 to 2020/21. The vertical reference line indicates the transition to remote instruction. (b) The figure shows the estimated difference in failure rates in subject English between grade 9 and grade 10, relative to reference academic year 2019/20. Both models include school-level and time fixed effects. Controls are Foreign Background, Active Certified Teachers, Postsecondary Parents, and Female Students. 95 percent confidence interval are included. Standard errors are clustered at the school level.



(a) Time Trends in Failure Rates



(b) Dynamic DiD Models

Figure 13 – Results for Subject Mathematics Prior to PSM

Notes: (a) The figure shows the time trends in failure rates in subject Mathematics by level of education for the academic years 2014/15 to 2020/21. The vertical reference line indicates the transition to remote instruction. (b) The figure shows the estimated difference in failure rates in subject Mathematics between grade 9 and grade 10, relative to reference academic year 2019/20. Both models include school-level and time fixed effects. Controls are Foreign Background, Active Certified Teachers, Postsecondary Parents, and Female Students. 95 percent confidence interval are included. Standard errors are clustered at the school level.

E Heterogeneity Analysis: Subject English

Table 14 – Heterogeneous Effects, Subject English

	Failure Rates in Subject English									
	Private (1)	Public (2)	Foreign Background		Postsecondary Parents		Female Students		Active Certified Teachers	
			High	Low	High	Low	High	Low	High	Low
2014/15 × Grade 10	−0.284 (0.480)	1.657** (0.656)	1.176 (3.234)	0.989*** (0.378)	1.003** (0.399)	0.048 (1.246)	0.541 (0.714)	1.230** (0.622)	1.141** (0.493)	1.691 (2.293)
2015/16 × Grade 10	0.312 (0.452)	1.949*** (0.518)	2.395 (2.246)	1.244*** (0.342)	1.233*** (0.385)	0.760 (0.978)	1.541** (0.631)	1.289** (0.535)	1.452*** (0.376)	2.285 (2.704)
2016/17 × Grade 10	−0.127 (0.436)	−1.397** (0.553)	0.704 (1.895)	−0.966** (0.392)	−0.262 (0.366)	−3.474*** (1.130)	−0.097 (0.646)	−1.363*** (0.524)	−0.944** (0.399)	0.577 (2.149)
2017/18 × Grade 10	−0.641 (0.406)	−0.740 (0.479)	0.346 (2.214)	−0.849*** (0.312)	−0.446 (0.301)	−1.824* (1.049)	−0.499 (0.577)	−0.857* (0.438)	−0.764** (0.338)	1.448 (2.131)
(Reference Year)	0	0	0	0	0	0	0	0	0	0
2019/20 × Grade 10	0.157 (0.396)	0.385 (0.464)	−0.397 (2.098)	0.296 (0.311)	0.423 (0.316)	−0.603 (1.018)	0.314 (0.599)	−0.014 (0.428)	0.223 (0.330)	1.238 (1.948)
2020/21 × Grade 10	−0.426 (0.493)	0.585 (0.511)	−1.264 (1.897)	0.406 (0.372)	0.349 (0.355)	−1.235 (1.211)	0.403 (0.523)	−0.280 (0.526)	−0.042 (0.369)	6.022** (2.921)
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,274	2,100	322	3,052	2,621	753	1,611	1,763	3,208	166

Notes: This table presents heterogeneous treatment effects by school ownership and composition characteristics for Subject English. For Foreign Background, Post-secondary Parents, Female Students and Active Certified Teachers, the sample was split as follows: "High" refers to schools with shares above 50%, and "Low" to those with shares at or below 50%. The reference academic year is 2018/19 and is excluded. Standard errors are clustered at the school level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Description of AI Usage

For the creation of this thesis, ChatGPT has aided the production of Python and R code for appending the data sets and for the formatting of tables, graphs and equations. It has also been used for formatting in Overleaf. The reason for using generative AI in these ways has been to improve efficiency. Each line of code has been thoroughly examined to understand what the consequences are of different choices in the code, and alterations have been made where appropriate. We have also used Scribbr for the creation of our reference list, although we are unsure if this is classified as generative AI. Incomplete and incorrect entries were altered manually.

We have avoided the use of generative AI in situations where we have considered there to be a risk that it may create errors. For instance, we faced a formatting problem for the reference list which we did not trust generative AI to handle, so we fixed the issue manually. We also turned off the generative AI included in Overleaf, as we wanted to produce the text ourselves. There may however have been additional uses of generative AI that we are unaware of due to the increasing prevalence of generative AI. Based on our experience using AI in the ways described above, we consider generative AI to be useful for repetitive tasks that do not require complex understanding, as time that otherwise would have been used on manual tasks can be freed up and used for other aspects of the thesis writing process.