Norwegian School of Economics
Bergen, Autumn 2017

The Effects of Technology on Students' Academic Performance

Rollout of Individual Laptops in Norwegian Upper Secondary Schools

Stine Victoria Stakkestad & Guro Fladvad Størdal

Supervisors: Patrick Dylan Bennett & Aline Bütikofer

Master Thesis, Economics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis is written at the Norwegian School of Economics as a part of our Master in Economics. It has been rewarding and motivating to study a topic of great importance in today's society. The close teamwork has been a joyful and highly valuable experience.

We would like to thank our supervisors Patrick Dylan Bennett and Aline Bütikofer for the useful feedback and engagement throughout this master thesis. We are grateful for the opportunity to work with such inspirational and dedicated researchers. Furthermore, we would like to thank contributors from the Norwegian Directorate for Education and Training, Statistics Norway, as well as County Authorities and Upper Secondary Schools throughout Norway.

Bergen, December 2017

Stine Victoria Stakkestad

Guro Fladvad Størdal

Abstract

Different technologies have been implemented in the educational system in Norway over the last decade and it has been a subject of debate whether the use of technology enhances students' educational outcomes. The aim of this master thesis is therefore to analyze the causal effect of the one-to-one laptop program in upper secondary education in Norway on the performance in three common core subjects: first-choice form of Norwegian, second-choice form of Norwegian, and English. The analysis is performed on a sample of 289 upper secondary schools in the time period from 2003 to 2016. We exploit data on average grades at school level and the rollout of the one-to-one laptop program across the country by using a generalized difference-in-difference approach and an event study specification. The results of this study indicate no clear benefits of technology use on academic performance in upper secondary education as no statistically significant effects are found. However, the true effect might be attenuated as the impact of laptops on students' academic performance is complex, i.e. there are both positive and negative effects, and performance is only reported as an average at school level.

The results presented in this thesis can be an important contribution to the literature in this field as little research has been conducted in Norway to interpret the causal relationship between technology and educational outcomes. The findings can hopefully inspire future research in the field to increase the knowledge on technology-led education. Moreover, it may also function as input for future decision making in the Norwegian educational system.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Purpose and Motivation

Norway is one of the countries with the most digitized educational systems in Europe (European Commission, 2013). Already in 2012, Norway had one-to-one coverage of internet-connected laptops in upper secondary school, while the EU-average was fourteen students per laptop. The one-to-one laptop coverage requires substantial investment. This could be one of the explanations why the Norwegian government spends 60 percent more money per student in upper secondary education than the average of countries in the Organisation for Economic Co-operation and Development (OECD) (Norwegian Directorate for Education and Training, 2017e). However, Norwegian students do not perform much better than the OECD-average in mathematics, science or reading, indicating that high spending per student cannot automatically be equated with high performance (OECD, 2015a). This motivated us to study the effect of technology use on academic performance in upper secondary school.

During the last ten years, there has been much focus on the place of technology in education (Hatlevik et al., 2013). Already in 2013, more than 80 percent of second-year students in upper secondary education reported that they used laptops always or regularly during class. In August 2017, the Norwegian Government presented the new digitization strategy for education, which aims to increase technology literacy among the students to make them more equipped to handle the future (Norwegian Ministry of Education and Research, 2017). The strategy states that educational institutions should be in a leading position in digitization as digital literacy will result in better labor market outcomes, and investments in technology at an early stage may increase returns to education drastically.

The new knowledge society is based on exploiting technological advancement (Erstad et al., 2005). Gaining access to digital tools in education including laptops will, therefore, be important. However, the Programme for International Student Assessment (PISA) conducted in 70 countries shows that students from countries which have heavily invested in educational information and communications technology (ICT) do not perform better than students in other countries (OECD, 2015b).

It is a paradox that extensive investments in digital infrastructure have been made while there is no clear guidance on how to take advantage of the digital devices in teaching (Hatlevik et al., 2009). In the report *ICT in Norwegian Education 2013* Hatlevik et al. (2013) show that almost half of the students in the second year of upper secondary school find the use of laptops in classroom disturbing for their educational outcomes. Further, about half of the students agree that they use too much time on non-educational activities and that the laptop (or tablet) displace time from activities that are conducive to learning. To fully exploit the benefits of laptop use in education it must complement existing teaching methods and improve the quality of education (OECD, 2015b).

The heavy investments in computers in school are made at the expense of traditional inputs such as educating better teachers and improved learning material (OECD, 2015b). For the program to be economically efficient it should yield better educational outcomes than alternative investments. The aim of this thesis is to estimate if the laptops is a good way of spending scarce resources.

1.2. Research Question

In this paper, we will exploit the national rollout of the one-to-one laptop program in upper secondary education to analyze the efficiency of technology use. Hence, our research question is as follows:

"What is the effect of use of individual laptops in upper secondary education on students' academic performance?"

Technology is a broad term and there are many technological developments initiated in the Norwegian school system, for example access to computer rooms, use of smart boards, or specific software used for computer-aided instruction (Bulman and Fairlie, 2016). In this paper, we have chosen to focus on the effect of the one-to-one laptop program in Norway, where each student at a school get access to an individual laptop. Hereinafter referred to as the laptop program.

Basic ICT literacy is one of the most important learning objectives in upper secondary education (Norwegian Ministry of Education and Research, 2006). Meaning that the laptops are not solely meant to support traditional subjects, but also make students familiar with technology in general. The scope of this paper is, however, limited to studying the impact on learning outcomes in traditional subjects. The benefit of students explicitly increasing their ICT literacy will therefore not be discussed in this paper, but should also be considered when evaluating the program. In addition, although it would be interesting to study the long-term effects of the laptop program, our school-level data only allows us to follow the students for a maximum of three years. This paper will therefore focus on the short-term effects of the laptop program.

The remainder of this thesis is organized as follows. In Chapter 2, we present the background including the main elements of the ICT policy in upper secondary education. Chapter 3 describes previous relevant literature in this field. Data used in the study are presented in Chapter 4, followed by an explanation of the empirical model chosen in Chapter 5. Furthermore, describing our results in Chapter 6, including a sensitivity analysis of the results. In Chapter 7, we discuss our findings and lastly, Chapter 8 concludes the thesis.

2. Background

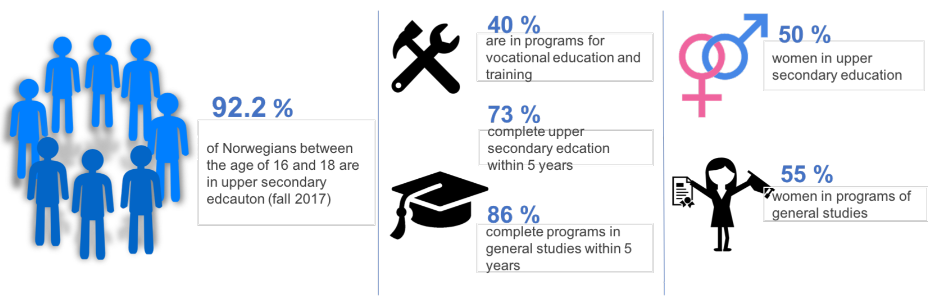
In this chapter, we will first describe the Norwegian upper secondary educational system in brief. Then we present the major lines of the Norwegian ICT-policy in education, including a description of the laptop program implementation.

2.1. Norwegian Upper Secondary Educational System

The county authorities are responsible for upper secondary education in Norway (Norwegian Ministry of Education and Research, 2007). The government has formulated a binding framework of legislation and regulations, including the National Curriculum. Within this framework, county authorities, as well as schools and teachers, can influence the education and training.

While grades 1–10 (primary and lower secondary school) are compulsory in Norway, upper secondary school is voluntary (Norwegian Ministry of Education and Research, 2012). However, it is an explicit political priority to increase attendance and completion in upper secondary education. In the autumn of 2017, 92.2 percent of all students aged 16 to 18 were enrolled in upper secondary education and 73 percent of the students completed their education within 5 years, see Figure 2.1 (Statistics Norway, 2017b). The educational pathway in upper secondary education comprises of three or four years and is organized into thirteen educational programs, five programs for general studies and eighth vocational programs (Norwegian Directorate for Education and Training, 2016b). About 60 percent of the students enter one of the programs for specialization in general studies, while the remaining 40 percent are entering one of the vocational education programs (Statistics Norway, 2017c).

Figure 2.1 – Statistics on Upper Secondary Education



Notes: Enrollment in vocational education programs, completion rate and female enrollment. Data from Statistics Norway (2017b).

2.2. Technology in the Educational System

Since 1990, technology-led educational development has been one of the main areas of focus for the Norwegian Government. *The Knowledge Promotion* which was put into force in 2006 states that digital literacy should be a core competency for all students (Norwegian Ministry of Education and Research, 2006).[1] This is formulated as one of five fundamental skills students should acquire during their basic education. The other four are oral skills, reading, writing and numeric skills. After the introduction of digitization in *the Knowledge Promotion*, the intention has been to gradually include technology use in all subjects (Hatlevik et al., 2009).

The digital infrastructure, including individual laptops, is the foundation for a good digital learning environment (Norwegian Ministry of Education and Research, 2017). Laptop use in education will increase opportunities for students on several dimensions and affect all subjects. First and foremost, it will give students increased access to information which will increase available literature, as well as challenge students' judgment on reliability. Furthermore, students will get access to programs that can enhance their learning outcome, for example by using educational software that corrects grammar or software helping students solve and understand mathematical problems, like the use of GeoGebra.[2] Overall, the use of laptop in education can make students able to focus on more complex problems, which in turn can increase students' performance.

However, access to laptops and students controlling their own usage require discipline. In addition to students being disturbed by the laptop, a major challenge is efficient use of technological devices in school (Hatlevik et al., 2013). Students confirm that the laptops are used more for entertainment, such as social media, music, and games, during class rather than for educational purposes. Moreover, the results from ITU Monitor 2013 showed that the benchmarks for digital literacy, in general, were not achieved. Thus, one of the main objectives of the digitization strategy of 2017 is to better utilize technology in order to actually improve the quality of upper secondary education (Norwegian Ministry of Education and Research, 2017).

Another main focus of the new strategy is proper training of teachers (Norwegian Ministry of Education and Research, 2017). As of today, the use of technical devices in class is not always an integrated part of the teaching and may pose a threat, rather than an improvement, to traditional learning. Hence, to improve educational outcome as a result of increased ICT

[1] *The Knowledge Promotion* is in Norwegian known as *Kunnskapsløftet*.

[2] GeoGebra is an interactive mathematical application for geometry, statistics, calculus and algebra.

usage, the new digitization strategy states that the teacher education program should ensure that all candidates obtain required professional digital literacy. In addition, all teachers should be offered further education to strengthen their digital literacy.

2.2.1 Implementation of Laptops

Upper secondary schools are operated by the county authorities (Knudsen, 2008). They therefore usually have to follow directions from the county authority regarding decisions made on provision and usage of laptops. However, implementation of the program is done individually at school level and some schools have thus implemented the laptop program prior to the decision of the county authority. The reason for this is that a few schools were chosen as a part of a pilot project at county or national level, and that a few upper secondary schools had available funds to initiate the project earlier than other schools. This means that the arrangement of the program differs from county to county, but there are also some variations within the counties.

Already in 1999, Nesodden upper secondary school in the county of Akershus introduced the laptop program as a part of a national pilot project (Frølich and Vestby, 2003). The project was evaluated as successful and as a result counties started providing students with laptops. The period of introduction was mainly from 2005 to 2010. As mentioned, some upper secondary schools chose to implement the laptop program prior to the county authority decision, but also a few schools chose to implement it later.

There are a variety of reasons for the differences between the counties in the timing of laptop program implementation. Some prioritized to develop the infrastructure, including power supply and wireless network, before the laptop program was implemented. Others prioritized to increase teachers' technological and digital competence. This was to increase teachers' abilities to utilize laptops to complement traditional classroom instruction, and hence improve the educational outcomes. Despite the differences, conversations with county authorities throughout Norway gave no indication of a systematic pattern in these priorities between those who were early and those who were late adopters of the laptop program.

In addition, counties opted for different schemes when implementing the laptop program. One option was for laptops to be purchased and distributed either by the school or the county authority and the students pay a certain amount each year to rent it. Another way was to provide students with subsidized deals and make them purchase laptops themselves. Both schemes allow students to bring the device home after school.

3. Literature Review

There is a comprehensive amount of studies on the theme of technology-led educational development. These studies focus on the effect on both social and educational aspects of use of technology in school and at home. In the following chapter, we will present relevant literature for our study. To our knowledge, there are no empirical studies interpreting the causal relationship of technology in education on academic performance in Norway.

Bulman and Fairlie (2016) address existing literature on technology and education in the paper *Technology and Education: Computers, Software, and the Internet*. In summary, investments in technology have ambiguous educational impact, and most often limited educational gains. Investments in technology in education are commonly divided into three categories; general investments in ICT in school, individual student laptops, and educational software, i.e. e-learning. We will mainly focus on the second form of investment, which is currently the biggest trend within educational policy. However, we will first present some studies on general ICT investments in education.

3.1. Effect of General ICT in Education

A randomized controlled trial was executed in California where over thousand computers were distributed randomly for free to children attending 6th-10th grade to use at home (Fairlie and Kalil, 2016). They find that the children who are given computers are more likely to have a social networking site, but also spend more time interacting with friends in person. There are no causal effects found concerning educational outcomes and only a small positive benefit to children's social development.

Faber et al. (2015) study the effects of government upgrades in ICT, by increasing the internet connection speeds, on children's school performance in England. They argue that the upgrades are randomly made across the country and that they can thus exploit exogenous variation to estimate the causal effect. They link the test scores of primary and secondary students to the availability of ICT at their home address and find that it has zero effect on students' educational attainment or learning productivity.

Malamud and Pop-Eleches (2011) examine a Romanian voucher program where 35,000 vouchers worth 200 euros were distributed to subsidize the purchase of home computers for low-income students enrolled in Romania's public schools. The study uses a regression discontinuity design to estimate the causal effect on academic achievement, cognitive skills, computer skills, and various non-cognitive outcomes. They find that the treated group achieve significantly lower grades in Math, English, and Romanian, but significantly higher score in a test of computer skills and in self-reported measures of computer fluency.

3.2. Effect of Laptop Programs in School

The first large-scale one-to-one laptop program initiated for educational purposes was implemented in the state of Maine in 2002 (Silvernail and Gritter, 2007). All students and teachers in 7th and 8th grade in the state were given laptops. Silvernail and Gritter (2007) perform a study that compares tests in writing performance written by hand in 2000 with the same form of test done after the introduction of laptops on the computer in 2005. The writing performance is increased by one-third of a standard deviation, but other types of tests do not seem to be affected. However, the simple comparisons made in the research might not be thorough enough to claim causality.

Suhr et al. (2010) study the effect of the introduction of a one-to-one laptop program for 4th and 5th graders in a Californian school district. They utilize a quasi-experimental research design. After two years the students who received a laptop perform better than non-laptop students in English language arts and tests measuring writing strategies and literary response and analysis.

The Texas laptop program was initiated in 21 schools in the state and the effects are studied by Shapley et al. (2009). The treatment group was matched with a suitable control group consisting of schools that did not receive laptops on several parameters such as school size, district, and proportion of minorities. The study uses difference-in-difference and finds some positive impact on the reading skills in some of the grades. There are no effects discovered in writing performance.

Furthermore, a study by Cristia et al. (2017) examines *the One Laptop per Child Program* conducted in Peru. This program aims to provide laptops to children to use at school and at home to improve learning in one of the poorest regions of the world. The paper focuses on the randomized controlled trial in Peru and exploits the randomized nature of the government policy program. Data were collected 15 months after the introduction and the study tests whether there is an increase in human capital accumulation. The program led to substantial increase in the usage of computers both at school and home. In conclusion, there is no significant effect on tests in neither math nor language courses. There is, however, a small effect on students' cognitive skills.

3.3. Summary of Existing Literature

Based on the literature presented in this chapter, there is little effect on students' performance of both general ICT investments in education and introduction of laptop programs. There is some evidence of an increase in computer and cognitive skills, but evidence of spill-over effects on other subjects is weak. However, the effects are ambiguous and so far it is hard to draw any conclusion in the field of study. In addition, the known literature mostly focus on the effect in lower educational levels and often on students from lower socioeconomic background than the average in Norway. Thus, the expectations for this study are ambiguous.

4. Data Description

This thesis aims to investigate the effect of the laptop program on students' academic performance in upper secondary. We have therefore constructed a panel data set where grades over time for upper secondary schools are linked with laptop program implementation year and a set of control variables. As the program is mainly rolled out from 2005 to 2010, the time period for this study is the school years 2003/2004 to 2015/2016. Data are mainly drawn from two sources: Norwegian Directorate for Education and Training and Statistics Norway.[1]

In this chapter we will first describe the data on upper secondary schools selected for this study, then we present the data on academic results and the selection of subjects. Thirdly, data on treatment year are presented, followed by data on control variables. Lastly, descriptive statistics will be outlined.

4.1. Data on Schools

As of autumn 2017, there are 419 upper secondary schools in Norway. 328 of these are public schools with an average of 536 students in each school (Norwegian Directorate for Education and Training, 2017e). The sample consists of 289 upper secondary schools for the school years 2003/2004 to 2015/2016 and is collected from the School Portal administered by the Norwegian Directorate for Education and Training.[2] In the remainder of the thesis we will refer to the school year of 2003/2004 as 2003 and so forth.

Data are collected on schools with students in General Study Programme only, while data from Vocational Education Programme (VET) and Supplementary Programme for General University and College Admissions Certification are left out. The rollout of the laptop program in VET and supplementary education often differ from the rollout in general studies. In addition, the timing of implementation and use of laptops in these study programs are less organized and more difficult to obtain precise information on. As a result data on implementation year are

[1] Norwegian Directorate for Education and Training (Utdanningsdirektoratet) is responsible for the development of kindergarten and primary and secondary education. Statistics Norway (Statistisk sentralbyrå) has overall responsibility for providing statistics on Norwegian society.

[2] The School Portal is a web-based portal for presenting relevant data on primary and secondary education (School Portal, 2017).

either missing or non-reliable. Hence, we will not be able to estimate the precise causal effect and the few schools that offer only VET and supplementary education are therefore not included. In addition, all private schools are excluded from the collected data. This is due to lack of data on grades and implementation year, as well as differences in regulations compared to public schools.

We have made several restrictions to the collected data. Firstly, only schools that have existed in the full time period, from 2003 to 2016, are included. The majority of schools opening during the chosen period are supplied with laptops from the opening year and, thus, there are no trends in grades before the adoption of the program. Likewise, most schools that closed during the period never received laptops and, thus, there are no trends in grades after the adoption. Only a few schools are lost due to this restriction. Secondly, the schools that started the laptop program before 2003 are excluded. The reasoning behind this choice is identical to the reason for the first restriction, i.e. no pre-trends. This applies to three upper secondary schools in the original data set. Thirdly, schools that have less than five observations on grades in total over the time period are removed from the data. The reason is that few observations will cause less precision when estimating the trends in grades before and after the laptop program implementation. Missing observations can indicate that exams were not carried out in the upper secondary school in the given year or not published because of privacy restrictions as there were too few students conducting the exam. This restriction is important in order to get valid trends, however, almost 30 schools are lost due to this.

4.2. Data on Academic Results

We have collected data on academic results from The Norwegian Directorate for Education and Training. Grades are collected as a total average score at school level and divided into averages for males and females. Hence, an observation is defined as the average grade in a given school obtained in a given subject in a given year. Referring to Section 4.1, we are operating with calender year and the observations for average grade in 2015 refers to the exam held at the end of school year 2015/2016.

The grading scale is ordinal and integer (Forskrift til Opplæringslova, 2009, §3-4). The scale goes from one to six, where one is failing, while two is the lowest grade students can obtain and still pass. Description of the grading scale used for upper secondary education is as follows:

- **Grade 6** – The student has an exceptionally high degree of competence in the subject.
- **Grade 5** – The student has a very high degree of competence in the subject.
- **Grade 4** – The student has a high degree of competence in the subject.
- **Grade 3** – The student has a fair degree of competence in the subject.
- **Grade 2** – The student has a low degree of competence in the subject.
- **Grade 1** – The student has a very low degree of competence in the subject.

In upper secondary education, the students receive two types of grades: final assessment grades and examination grades (Norwegian Directorate for Education and Training, 2016a). The final assessment grades are decided by the teachers in the respective subjects at the end of the school year based on a variety of assessments throughout the year. Hence, final achievement grades are based on the performance over time and teachers deciding the grades are only exposed to one school and might observe systematic improvements after the introduction of the laptop (Norwegian Directorate for Education and Training, 2017d). This can imply that final achievement grades are likely to be biased because evaluation will be less objective, and thus weakens the causal interpretation of the effect of laptops. The final achievement grades are therefore not considered to be a good measure on academic performance to detect the effect of implementation of laptops in Norwegian educational system.

The examination grade, which can be oral or written, are held at the end of the school year and the grade is decided by an external sensor. Written exams are given at a national level and created by 2-6 experts appointed by the Norwegian Ministry of Education and Research (Norwegian Directorate for Education and Training, 2017b). Hence, they are independent of school and county which ensures that all students are tested at the same difficulty level. Furthermore, the written exam is handed in anonymously and graded by two different sensors. The sensors are required to attend a sensor seminar before the assessment process so everyone has a common understanding of what is expected of the student. Thus, the grading is done with national guidance at county level making results comparable across schools and counties. In contrast, oral exams are locally given, thus the assessment is less objective (Norwegian Directorate for Education and Training, 2017a). In addition, it is likely that laptops will impact student performance in written exams to a greater extent compared to oral exams. Based on these arguments, grades on written exams are used as the measure of academic performance in this study.

Selection of Subjects

We aim to study the effect of technology on general academic performance in upper secondary education. Hence, we want to collect data on written exams in common core subjects for the program of general studies, which are Norwegian, English, first year mathematics and other foreign languages (Norwegian Directorate for Education and Training, 2016c). For both mathematics and other foreign languages the data are either missing for the period from 2003 to 2007 or suffer from poor quality. We have, therefore, collected data on grades in first-choice form of Norwegian, second-choice form of Norwegian and English.[3]

The number of exams the students need to take during their upper secondary education varies (Norwegian Directorate for Education and Training, 2016c). In the first year, about 20 percent of the students at each school are randomly selected to take one exam, either written or oral. Among these first-year courses are English. In the second year, all students conduct one exam, either oral or written, in a randomly drawn subject. In the final year, students conduct four exams, three written and one oral. Among the written exams, first-choice form of Norwegian is mandatory for all upper secondary students in general studies. The remaining exams are randomly assigned by the county authorities. Second-choice form of Norwegian can be one of these. It is important to note that the assignment of exams needs to be evenly divided into subjects and schools over time.

4.3. Data on Treatment Year

This paper analyzes the effect of technology use, represented by the laptop program, on academic results. Hence, a school is considered as treated from the year the students receive an individual laptop at school through such a program. It does not take into consideration whether the school had computer rooms or class set of laptops for loan in a random selection of classes prior to the program.

To collect information about the layout of the laptop program in each upper secondary school, dialogues with all county authorities were started. While some counties had concrete information about when the program was implemented in each school, others did not have any information. As previously stated, some schools chose to start the implementation prior to the

[3] First-choice form of Norwegian is either "Bokmål" or "Nynorsk" and second-choice form of Norwegian is the opposite of the first-choice form.

county standard. Therefore, in counties with a degree of uncertainty, all the upper secondary schools were contacted to collect more reliable information.

In order to check the validity of the information given by the county authorities, about 25 percent of the schools in each of the remaining counties were contacted as well. These schools were randomly selected from the sample. As these schools mainly confirmed the information given from the county authorities, the data are considered as valid. Nevertheless, there might be some errors in the observations, but as we have conducted validation checks, these errors are most likely randomly distributed throughout the data set and, hence, will not bias the results.

Implementation of Program

There are several different models for implementation of the laptop program. The most common model is that only first-year students receive laptops in the initial year of the program. These students keep the laptop throughout their time in upper secondary education. Hence, three years after the initiation there is complete one-to-one coverage of student laptops at the school. Another common model is that all students in all years at the given upper secondary school receive laptops immediately. Hence, there is complete one-to-one coverage already in the implementation year. Furthermore, there are several other less common models for implementing the laptop program starting with second or third-year students.

By exploiting data on treatment year and implementation model, we have defined dummy variables indicating whether first and third-year students have an individual laptop at the time of the exams. In addition, data on how long students have been exposed to the laptop program are collected through this information, which can be a maximum of three years.

4.4. Data on Control Variables

We have chosen to include six control variables in our analysis. Firstly, data on the number of students in each school from the period of 2003 to 2015 are collected from the School Portal. It is divided into the number of boys and girls in each school and based on this, we have constructed a variable on the share of boys at each school. Data on dropout rate and data on *the Pupil Survey* are collected from the School Portal. Both are for the time period from 2006 to 2015 at the school level. *The Pupil Survey* provides data on the student's well-being and learning environment at each school, and includes the following variables: well-being with teachers, social well-being, student's democracy, physical learning environment, bullying at school, motivation, professional guidance, participation in decisions regarding own education, career guidance, and academic challenge. The scale used in the survey is ordinal and integer from one to five. Lastly, data on expenditure per student per county in Norwegian Kroner (NOK) and student-to-teacher ratio per county are collected from Statistics Norway for the time period 2003 to 2015.[4]

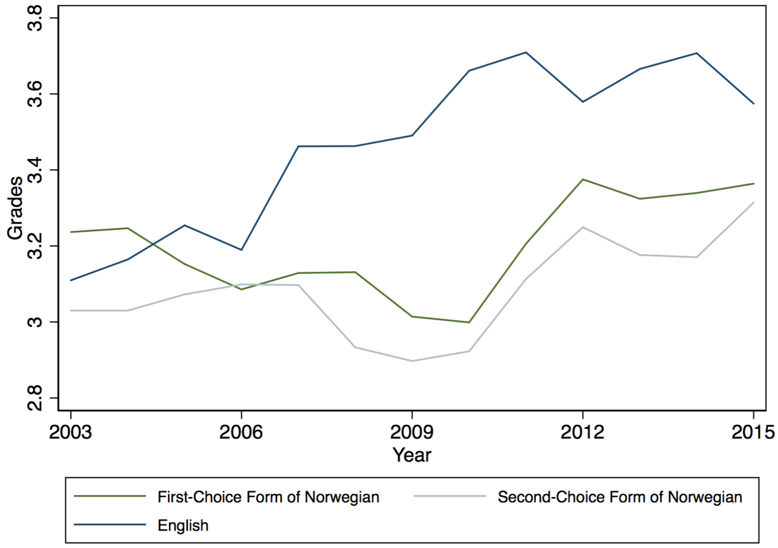
As the data on dropout rate and *the Pupil Survey* only run from 2006, there are systematic missing observations for the period from 2003 to 2005. Observations prior to 2006 are therefore replaced with the mean of the four following years. This is necessary to estimate trends in the period prior to the implementation for the early adopters. As these characteristics are relatively stable over time, the mean value for the four following years will be a suitable estimate. For the remaining missing observations on background characteristics throughout the data, we include a dummy variable indicating that the variable is missing to keep the sample constant across specifications with and without control variables.

4.5. Descriptive Statistics

Figure 4.1 shows the development of average grades in first-choice form of Norwegian, second-choice form of Norwegian, and English over time. The average grades tend to fluctuate around 3.2. Grades in first-choice and second-choice form of Norwegian have a lower and more stable trend throughout the time period, while English grades tend to have an increasing trend over the period of interest. Furthermore, Figure A1 in Appendix shows the development of average grades in the three subjects over time for each of the 19 counties. We observe that there are differences between the counties over time. For example, in the county of Finnmark there is a lot of variation in the performance between the subjects, while in the county of Oppland there is relatively little variance between the subjects. English stands out as the subject with the greatest variation from year to year at county level with several spikes and troughs, examples are the counties of Sør-Trøndelag and Telemark. The variation within each county confirms that grades are not normally distributed at county level, but at national level.

[4]Data are gathered from Statistics Norway's StatBank on Upper Secondary Education (Statistics Norway, 2017d).

Figure 4.1 – Average Grades



Notes: The figure plots the national average of grades in first-choice Form of Norwegian, second-choice Form of Norwegian, and English from 2003 to 2015. The grading scale is ordinal from one to six. Data are provided from the Norwegian Directorate for Education and Training's School Portal (2017).

See Table 4.1 for descriptive statistics on each subject for the full sample, as well as subsamples for male and female. For the full sample there are more than 3,000 observations on first-choice and second-choice form of Norwegian, while there are about 1,600 observations on English. Hence, the sample is quite small for the latter subject. This is, however, reasonable due to the fact that English is a subject in first grade where in total only 20 percent of the students at each school are selected to take one exam. As the observations are average grades at school level, there are no extreme observations of one or six, the variation in grades for the subjects is from 1.4 to 5.3. The standard deviation is about 0.4 grade points for the two Norwegian subjects, while it is 0.5 grade points for English. Furthermore, the grades are highest in English in all samples. This is true for the average grade, as well as the maximum and minimum observation. Lastly, on average female students seem to perform better in all subject compared to their male peers.

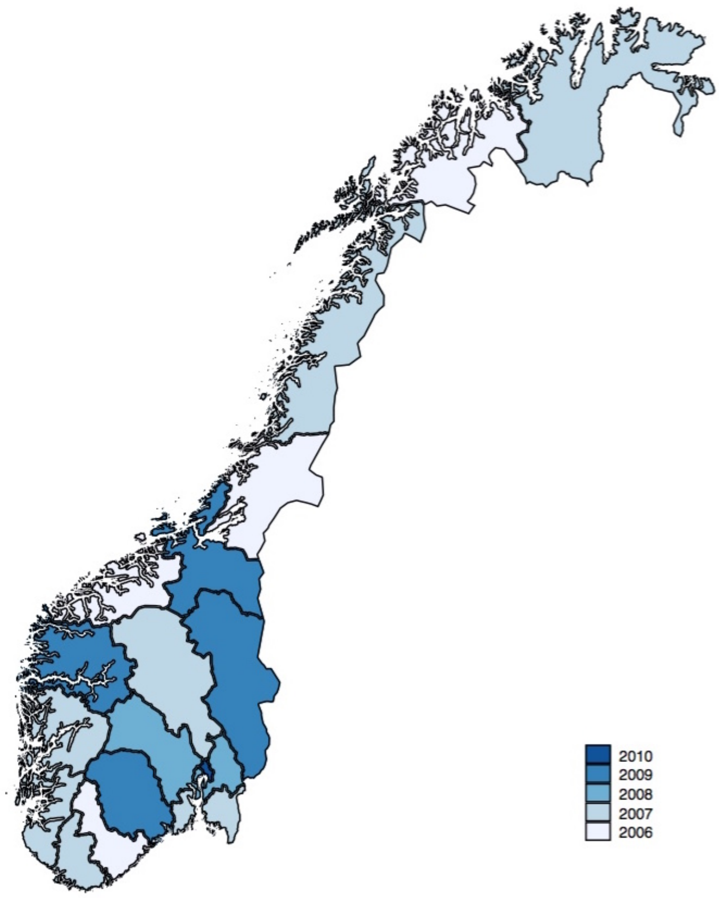
Table 4.1 – Descriptive Statistics

	Full Sample	Male Sample	Female Sample
First-choice Form of Norwegian			
Mean	3.1926	3.0765	3.3344
Std.Dev.	0.3921	0.3952	0.3951
Max	4.6	4.4	4.7
Min	1.5	1.6	1.6
Observations	3219	2779	2881
Second-choice Form of Norwegian			
Mean	3.0789	2.9489	3.2279
Std.Dev	0.4270	0.4305	0.4255
Max	4.8	4.5	4.9
Min	1.4	1.4	1.5
Observations	3052	2456	2594
English			
Mean	3.4371	3.4779	3.5060
Std.Dev	0.5118	0.5382	0.5153
Max	5.1	5.3	5.3
Min	1.7	1.8	2.1
Observations	1592	772	870

Notes: The table displays average grades from upper secondary education in Norway for the three common core subjects: first-choice form of Norwegian, second-choice form of Norwegian, and English. The time period of interest is 2003 to 2015. Descriptive statistics are divided into three samples: the full sample, male and female. The grading scale is ordinal from one to six. Data are provided from the Norwegian Directorate for Education and Training's School Portal (2017).

Figure 4.2 is a geographical representation of the average implementation year of the laptop program for first-year students of upper secondary education in each county in Norway. No clear patterns in the rollout of laptops across counties can be observed, which makes the basis for our identification strategy outlined in Chapter 5 (Empirical Approach). The average treatment year per county varies from 2006 to 2010, and the national average year of initiation is 2007. This is also shown in Figure A2 in Appendix, which plots the number of upper secondary schools implementing the laptop program each year, i.e. treatment year for each school. From 2004 there is an increase in the number of upper secondary schools implementing the laptop program, and from 2005 to 2007 the number almost doubles each year. Then the number of implementations decreases from 2007, and almost no new upper secondary schools have implemented the program after 2012.

Figure 4.2 – Rollout of Laptop Program by County



Notes: The map displays Norway's 19 counties. Different shades of blue indicate the average year of implementation of the laptop program to first grade students. Average implementation year in the counties vary from 2006 to 2010. Brighter colors indicates earlier implementation.

5. Empirical Approach

The aim of this thesis is to estimate the causal effect of the use of individual laptops in upper secondary school on students' academic performance. We utilize a rollout strategy and an event study specification to estimate average effects on academic results in all treated upper secondary schools. Our identification strategy exploits the variation in the timing of the laptop program implementation in upper secondary education. We assume that while the timing varies, it is independent of factors that affect our outcome variables and thus creates an exogenous shock to these.

5.1. Rollout

The rollout strategy is a generalized difference-in-difference approach, exploiting the rollout of individual laptops across upper secondary schools over time. Papers utilizing this approach are for example Akerman et al. (2015) and Bütikofer et al. (2016). We define a treatment and a control group based on the timing of the laptop implementation. All schools enter into the analysis at the same time, but the treatment period differ for each individual school. The control group consists of schools that at the given point in time have not implemented the laptop program, while the treated schools are those who implement the laptop program in the given year or have already implemented it. We estimate the following reduced form model:

$$y_{ict} = \alpha + \gamma D_{it} + \delta C_{it} + \beta X_{it} + \sigma X_{ct} + \phi_t + \tau_c + \epsilon_{ict}, \tag{5.1}$$

where y_{ict} is the outcome of interest for upper secondary school i in county c at time t. α is a constant term. D_{it} is a dummy variable equal to one if the laptop program is implemented in school i at time t, and zero otherwise. γ shows the effect of technology on academic results and is one of our key coefficient. Furthermore, C is an ordinal variable taking the value from zero to three depending on the number of years the student has been exposed to the laptop program. δ thus shows the effect of the length of treatment on academic results and is our second key coefficient. X_{it} is a set of characteristics at school level that is time variant. This includes the total number of students, share of boys, score on the different variables in *the Pupil*

Survey and dropout rate. X_{ct} is a set of time variant characteristics for the county including student-teacher ratio at the beginning of the school year and expenditure per student.

Furthermore, we have included unobserved time fixed-effects to control for common year specific events, indicated with ϕ_t. Such shocks could be that the Norwegian Directorate for Education and Training give an exam that is either particularly easy or difficult one year. τ_c is a set of indicator variables at county level allowing for time-invariant geographical factors to differ between the counties. An example could be that schools in counties where opposite political parties have power perform systematically different.

If unobserved year or county fixed effects that are correlated with the outcome variable are excluded from the model, it will lead to omitted variable bias (OVB) (Angrist and Pischke, 2015). Also including the control variables, X_{it} and X_{ct}, helps to avoid OVB by controlling for effects that could cause post-treatment differences between treated schools and the control group. Including these variables contributes to the conditional mean independence assumption that the error term should be independent of the variable of interest.

Test of Key Identifying Assumption

The empirical approach exploits the variation in the timing of the treatment, allowing for findings of average treatment effects on the outcome of interest. Hence, the rollout strategy relies on the assumption that timing of the laptop program implementation is uncorrelated with other determinants of academic performance. Ideally, we want to test these other determinants, for example mother and farther education. However, as we exploit school-level data this is not possible. Therefore, we will instead investigate whether there are any relevant policy changes in the period that could be correlated with the timing of the laptop implementation. In addition, we will graphically check whether implementation of laptop program seems to be determined by trends in observable characteristics of counties. Although the absence of such a determination is not proof of the identifying assumption, it does provide some suggestive evidence of whether it could be true.

As upper secondary schools are governed by the county authorities the laptop implementation is mainly determined by their decisions. Hence, the policy changes of interest are those creating differences across counties. To our knowledge, there are no major policy changes that might influence grades in the chosen period that will harm the identification strategy. *The Knowledge Promotion*, which was introduced in 2006, established new guidelines for education in upper secondary, including changes of curricula and school structure. However, this is a nationwide reform affecting all counties at the same time (Norwegian Ministry of Education and Research, 2006).

Furthermore, we graphically check if large differences in the expenditure per student coincide with the implementation of the laptop program. Studies such as *Pennies from heaven?* by Hægeland et al. (2007) show that school resources may affect student performance, and is thus a concern to the identification strategy. A sharp increase in expenditure per student prior to the implementation year might, thus, violate the key identifying assumption. Figure A3 in Appendix displays each county's expenditure per student for the period 2003 to 2015. The vertical line marks the average year of laptop program implementation in each county. The figure displays different patterns across counties, which will be discussed in the following paragraph.

Firstly, some experience a steady increase throughout the full period, such as the county of Oslo, or an increasing trend over a longer period before the implementation, as observed for the county of Telemark. Others experience a relatively constant spending until the initiation year of the laptop program, such as the counties of Troms and Østfold. Lastly, the most worrisome pattern is counties experiencing a change in trend just prior to the initial year. This is true for the counties of Akershus, Buskerud, Finnmark, Hedmark, and Sogn og Fjordane. The counties of Buskerud, Finnmark, and Hedmark experience a decreasing trend in expenditure per student before the trend turns just prior the initial year. The trend for the counties of Akershus and Sogn og Fjordane is relatively flat with a distinct increase just prior to the initial year. As expenditure per student could affect students' academic performance this might violate the key identifying assumption. As a further test, we will also run the main specification when excluding these five counties from the sample, see Subsection 6.1.1. A final remark is that all counties experience a steady increase in expenditure per student after implementing the laptop program. Since this is a common trend for all counties, it is not a concern that is likely to bias our estimates.

Other observable background characteristics do not cause great concern in relation to determining the implementation of laptops. The teacher-to-student ratio is relatively constant for all counties throughout the period. The remaining background characteristics are observed at school level. Because the timing of laptop program implementation is relatively stable for schools within the same county, these characteristics are less likely to impact the timing of implementation. Therefore, the key identifying assumption is therefore not tested on these. In addition it is worth noting that the results from testing the key identifying assumption on dropout rate and *the Pupil Survey* would provide highly biased results as the missing observations for the period from 2003 to 2005 are replaced by an average of the four following years.

5.2. Event Study

In order to test the assumption of independence in the timing of laptop implementation, the pre-trends are of interest (Angrist and Pischke, 2008). For this purpose, an event study specification is used.[1] Papers utilizing this approach are, for example, Bütikofer et al. (2012) and Bailey and Goodman-Bacon (2014). A control and a treatment group are defined for each year of our time period and the control group must in every sample be a credible candidate for the counterfactual outcome of the treatment group. The method visually presents the differences in estimates of the outcome variables in the two groups over time (Kose et al., 2015).

Evaluating whether there are parallel pre-trends is done by studying the anticipatory values of the variables of interest. We utilize the following event-study specification for this purpose:

$$y_{ict} = \alpha + \sum_{\iota=0}^{m} \omega_{-\iota} D_{i,t-\iota} + \sum_{\iota=1}^{q} \omega_{+\iota} D_{i,t+\iota} + \beta X_{it} + \sigma X_{ct} + \phi_t + \tau_c + \epsilon_{ict}, \tag{5.2}$$

where m $(\omega_{-1}, \omega_{-2}, ..., \omega_{-m})$ are the anticipatory effects and represents the academic results in the years prior to the laptop program. q $(\omega_{+1}, \omega_{+2}, ..., \omega_{+q})$ are the lagged effects, i.e. the results in the years after the treatment. X_{it} and X_{ct} are time variant control variables at school and county level respectively, similar to the rollout specification. ϕ_t is time fixed effects and τ_c is a set of county fixed effects. The year prior to the initial year of the laptop program will be omitted as a control year.

In cases where the anticipatory effects are significant, it indicates that there are differential pre-trends in the treatment and control group and that the independence assumption does not hold (Asteriou and Hall, 2011). In our case, this could be that laptops are given to upper secondary schools that have experienced an increase in academic results prior to implementation. It is important to note that the test is not evidence for parallel trends, but rather that we cannot rule out the existence of parallel trends. However, if $\omega_{-\iota}$ is insignificant for all $\iota < 0$ no evidence for differential pre-trends is found.

[1] This is also known as Granger causality testing.

If the anticipatory effects are insignificant, the lagged values in the model can have a causal interpretation. In cases where $\omega_{+\iota}$ is different from zero for all $\iota \geq 0$, $\omega_{+\iota}$ is the causal effect of the laptop program in the year ι after treatment. Thus, how the effect on academic grades evolves over time can be observed.

5.3. Specifications

Operating with panel data structure there are several potential issues that need to be addressed (Angrist and Pischke, 2008). Serial correlated error term is a common issue and indicates that the standard errors cannot be used for statistical inference. This occurs when the observations are not independent of each other and the values of the outcome variables in periods close by are correlated. In our model it is likely that average grades for a given school i in year t are relatively similar to the observations in year $t+1$ and $t+2$, and so on.

Another problem that can occur is clustering, meaning that the data have group structure (Angrist and Pischke, 2008). In this set-up it can for example be that observations from upper secondary schools within the same county are correlated because they are regulated by the same authorities and, hence, are exposed to similar learning environments.

Both serial and intra-group correlation can be adjusted for by clustering the standard errors (Angrist and Pischke, 2008). If we cluster the standard errors at school level, we can control for serial correlation over time. However, the variation in standard errors between schools needs to be random, i.e. no correlation between clusters of schools, such as within county. By clustering at a higher level we can control for this. The disadvantage will, however, be fewer clusters which means fewer random residuals to determine the standard errors. When the number of clusters is low, problems with inference arises as well. Angrist and Pischke (2008) determine 42 as a minimum of clusters for the inference to be reliable. Clustering on county results in only 19 clusters. In these specifications, standard errors are thus clustered at the school level and made robust to heteroscedasticity.[2]

[2] See Subsection 6.1.1 for clustering at a higher level.

6. Empirical Analysis

In the following chapter, we present our main findings, as well as a sensitivity analysis. We estimate the average effects of individual laptops on academic grades in three common core subjects of upper secondary education: first-choice form of Norwegian, second-choice form of Norwegian, and English. The main findings are divided into two parts, results from the rollout specification and results from the event study.

6.1. Main Findings

All results are reduced form estimates, measuring the intention to treat (ITT) effect, i.e. the effect of being assigned to treatment (Angrist and Pischke, 2015). However, since we use data on general studies only and the adoption of the laptop program was similar for most students within this study program at a given school, the ITT estimates and the treatment on the treated (TOT) estimates will be almost the same.

The ordinal and integer grading scale has few outcomes, running from one to six. Thus, it requires quite a large change in performance in order for a student's grade to increase from for example three to four. This indicates that if changes in performance are not large enough to make a shift in the examination grade, the change will not be uncovered. If the scale was from 1 to 100, it would be more sensitive to small changes and hence require less improvement to make a shift in grade.

6.1.1 Rollout

Table 6.1 presents the main estimated results of the laptop program on student academic performance in upper secondary education using the rollout specification from Equation 5.1. Each estimate is from a different regression. Column (1) shows the effect of a laptop program and the effect of length of treatment on academic results in first-choice form of Norwegian, while Column (2) presents the estimates on second-choice form of Norwegian. Column (3) present the effect of implementing a laptop program on English. Because the exam in English is held in the first year, the length of treatment is equal for all students and hence this variable is omitted from the specification for this subject.

Panel A presents the results from our preferred specification where data on number of students in each school, share of boys, dropout rate, score on the variables in *the Pupil Survey*, expenditure per student per county and student-to-teacher ratio are included as control variables. Panel B provides the results without individual or county specific control variables. Both specifications include fixed effects for time and county, and standard errors are clustered at the school level. Estimates with standard errors clustered at county level are presented in Table A3 in Appendix. The standard errors of the estimates for the three outcome variables in this specification are relatively similar in size to the standard errors in our baseline model.

Overall, we cannot find significant effects of the laptop program on students' academic performance in any of the three common core subjects in panel A. The signs of the estimates on the third-year subjects, first-choice and second-choice form of Norwegian, are positive, while the estimated effect on the first-year subject, English, is negative. This might be due to the fact that third-year students are using their laptops more efficiently for educational purposes, leading to more positive outcomes. However, none of the estimates are significant so we cannot draw any conclusions.

Panel B will challenge the validity of our baseline model and helps determine whether we estimate the effect of interest. The estimates on all subjects change, but the changes are small and within the confidence interval of the baseline specification. The most worrisome change is that the sign is altered for the estimated effect of the laptop program on first-choice form of Norwegian. Removing the control variables changes the estimated effect from 0.0080 in our baseline model to -0.0001. However, the estimate is not significantly different from the baseline specification.

In addition, we estimate the model where only one of the independent variables of interest are included at a time on the third-year subjects: first-choice form of Norwegian and second-choice form of Norwegian, see Table A4 in Appendix. Panel A presents the results when estimating the effect of the laptop program exclusively, i.e. only the effect of dummy variable D_{it}. All control variables and other specifications are identical to the baseline model. Similarly, Panel B presents the effect of length of treatment, the ordinal variable C_{it}. Estimates in Panel A and Panel B for both subjects are positive and still insignificant, and neither are significantly different from the estimates in the baseline specification.

Table 6.1 – Rollout Estimates

	(1) First-Choice form of Norwegian	(2) Second-Choice form of Norwegian	(3) English
Panel A: Baseline			
Laptop program	0.0080	0.0494	-0.0405
	(0.0894)	(0.1090)	(0.0620)
Length of treatment	0.0099	-0.0081	
	(0.0347)	(0.0416)	
Controls	Yes	Yes	Yes
No. of clusters	288	287	273
Observations	3219	3052	1592
Panel B: No Control variables			
	(1) First-Choice form of Norwegian	(2) Second-Choice form of Norwegian	(3) English
Laptop Program	-0.0001	0.0259	-0.0580
	(0.1070)	(0.1140)	(0.0679)
Length of treatment	0.0067	-0.0086	
	(0.0412)	(0.0452)	
Controls	No	No	No
No. of clusters	288	287	273
Observations	3219	3052	1592

Significance levels: *** 1% level, ** 5% level, * 10% level

Notes: Each parameter is from a separate regression of the outcome variables, which are average grades, on the laptop program and length of treatment, based on the model in Equation 5.1. The grading scale is ordinal from one to six. Robust standard errors adjusted for clustering at the level of the school are shown in parentheses. The sample includes schools in the time period 2003 to 2015. All specifications include a set of county and time fixed effects. Additional control variables in Panel A: number of students, share of boys, dropout rate, score on each variable in *the Pupil Survey*, expenditure per student per county and student-to-teacher ratio.

Referring to the discussion on key identifying assumption in Section 5.1, Table A5 in Appendix displays the results of the regression excluding the counties with increasing trend in expenditure per student prior to the implementation of the laptop program. The estimates are relatively similar and still insignificant. Even though this is not proof of the key identifying assumption, it does provide some evidence that it holds.

Hetereogenity

The descriptive statistics in Table 4.1 show that female students on average perform better in all subjects compared to their male peers. This makes it interesting to test whether the effects differ if we divide the sample into male and female students. The estimated effects of the laptop program and length of treatment when the sample is divided into subsamples are presented in Table A6 in Appendix. Panel A present the estimates for male, while estimates for female are presented in panel B. The results are insignificant for both subsamples. However, we observe that estimates are different for the two genders. For example, in first and second-choice form of Norwegian, the sign is negative for male students and positive for female students, but the estimates are not significantly different between the genders.

6.1.2 Event Study

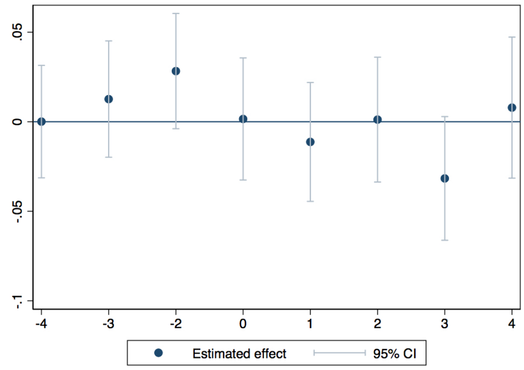
To test whether the implementation of the laptop program is independent of specific pre-trends, we utilize the event study specification from Equation 5.2. First, we check whether the assumption on pre-trends holds, before we study the lagged effects, which indicate whether there is an effect on the academic outcomes over time. The sample is restricted to the upper secondary schools where only first-year students received laptops in the initial year of the program.

Figure 6.1 visually plots the event study estimates of the effects of the laptop program on first-choice form of Norwegian and second-choice form of Norwegian, and English, including the 95 percent confidence interval for three anticipatory and four lagged effects. The set of control variables is similar to the rollout specification, and the standard errors are robust and clustered at school level.

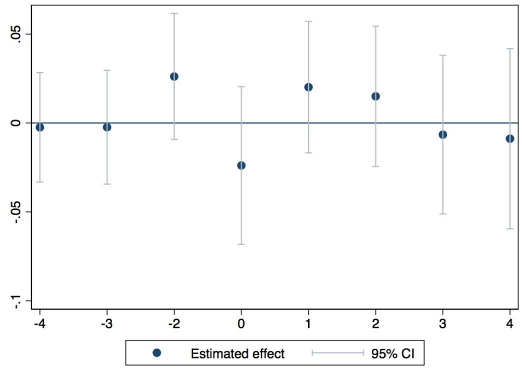
For both first-choice and second-choice form of Norwegian all the anticipatory effects are insignificant, indicating that the parallel trends assumption holds in the pre-period. However, the lagging effects are also insignificant, which means that no evidence of the laptop program affecting academic performance is found. In these two subjects the exam is undertaken in the third year, thus there should be no effects before year 2 in our analysis. Although the effects are insignificant, first-choice form of Norwegian tends to have an increasing trend in grades before year 0, while no such pattern is observed later and the variation from year to year seem more random after year 0. For example, there can be observed a larger, but insignificant, decrease in year 3, which is one year post treatment.

For English, there is a significant and negative effect three years after the implementation of laptops. However, the anticipatory effects in year two are also significant indicating different pre-trends, i.e. the treatment group have higher grades two years prior to the initiation compared to the control group. Hence, we cannot interpret the lagged effects as causal. In addition, the effect in the post-period does not persist over time and the drastic decrease appears to be somewhat random. English seems to be the subject where the results are most volatile throughout the period of interest. This could be because there are less students undertaking the exam each year and, hence, the sample is smaller and less stable compared to the two Norwegian courses.

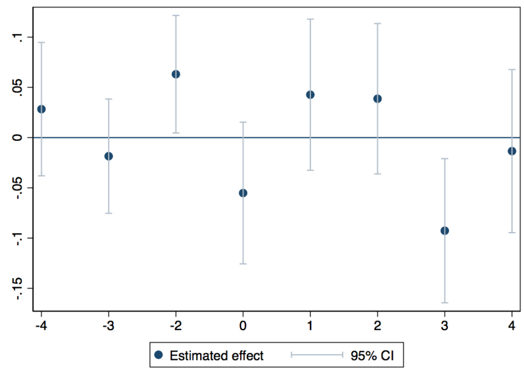
Figure 6.1 – Event Study Estimates



(a) First-Choice form of Norwegian



(b) Second-Choice form of Norwegian



(c) English

Notes: The figures plot the anticipatory $\omega_{-\iota}$ and post-treatment effects $\omega_{+\iota}$ from the event-study specification in Equation 5.2, as well as the 95 percent confidence interval. The set of control variables includes county and time fixed effects, as well as number of students, share of boys, dropout rate, score on each variable in *the Pupil Survey*, expenditure per student per county and student-to-teacher ratio. The standard errors are clustered at school level and robust. The period of interest is 2003 to 2015. Figure (a) shows the effects on first-choice form of Norwegian, figure (b) shows the effects on second-choice form of Norwegian, and figure (c) shows the effects on English.

6.2. Sensitivity Analysis

In this section, various sensitivity analyses are presented. This will test the robustness of our empirical results and might expose weaknesses in the model. Given that we find no significant effects and large standard errors, we expect to get relatively similar estimates in these specifications. First we test whether the results are robust to the exclusion of extreme observations. Then, we test the exclusion of uncertain observations. Further, we exclude counties one-by-one, and lastly, we check the robustness by estimating the specification with differential time trends.

Excluding Outliers

With few observations, Wooldridge (2012) emphasizes that the regression estimates can be sensitive to extreme observations. Hence, estimating the specification without extreme observations will be helpful to identify whether these drive the results. However, with a small sample, we must also be careful not to exclude too many observations as the precision of our estimates could be compromised. In Panel A of Table A7 in Appendix, we examine the robustness by excluding the upper and lower 5th percentile of observations of the output variables in the baseline specification. The estimates on the laptop program in all columns have changed signs compared to the results from the baseline model; first-choice and second-choice form of Norwegian have both negative coefficients, while English has a positive effect. The estimates do, however, remain insignificant and within the confidence interval of the baseline specification. This is as expected because we are operating with average grades at school level and, hence, there should be no extreme outliers.

Excluding Schools with Uncertain Treatment Year

Upper secondary schools that had no clear structure of laptop implementation or had changed the staff responsible for the laptop program in the later years are marked as uncertain in the data set. Panel B of Table A7 in Appendix present the estimates when the observations with a higher degree of uncertainty are excluded. Performing this analysis can potentially exclude observations with false information from the data and hence helps to check whether this uncertainty biases the results of our baseline model. However, none of the estimates on the three outcome variables are significantly different from the baseline estimates.

Excluding Counties One at a Time

As discussed in Section 4.5, Figure A1 in Appendix displays differences in grades between the counties over time and shows that there is great variation in some subjects. By excluding counties one at a time, these large fluctuations existing within specific counties are eliminated from the sample and, hence, we are able to observe whether this alters our results. The results are reported in Figure A4 in Appendix, the excluded county is indicated at the x-axis. Overall, there are no statistically significant effects in the estimates when removing one county at the time from our model, which indicates that the results are relatively consistent across the counties. However, there seems to be some variation, for example, we see that the estimated effects of the laptop program on average grades for second-choice form of Norwegian increase slightly, but not significantly when removing the county of Hedmark.

Differential Time Trends

To distinguish the effect of the laptop program from differential persistent county-specific time trends, we include a new parameter. Estimates including this variable are likely to be more robust (Angrist and Pischke, 2008). Differential time trends could occur if there are time driven underlying trends in academic performance that differ between the counties. An example is changes in educational policies at the county level in different years. Because we have a sufficient amount of groups, we can include this variable allowing for potential systematic differing trends between the counties. The following specification thus tests the robustness for county-specific time trends:

$$y_{ict} = \alpha + \gamma D_{it} + \delta C_{it} + \beta X_{it} + \sigma X_{ct} + \rho_c t + \phi_t + \tau_c + \epsilon_{ict}, \tag{6.1}$$

where ρ_c is the coefficient of a county-specific time trend multiplied by a linear time trend variable, t. The remainder of the variables are similar to the specification in Equation 5.1. If the estimate of the variable of interest, γ, is significantly different from the baseline specification there might be pre-existing linear time trends that should be controlled for. As shown in Table A8 in Appendix there are no significant effects in any of the subjects and none of the estimates are significantly different from the baseline model even if some signs are altered.

6.3. Summary of Results

We find no significant effects of the laptop program on students' academic performance. Removing control variables alters the estimates compared to our baseline model, but the effect remains insignificant and is within the confidence interval of the baseline specification. Hence, it has no further implication. The results are also emphasized by the sensitivity analysis. The estimates remain statistically insignificant and do not change much when performing different robustness checks. This implies that the baseline estimates are robust to exclusion of observations and to changes in the specification.

7. Discussion

In this chapter, we will first discuss our main findings. Then we will address some limitations to our analysis regarding data and empirical approach. Finally, implications of the study and proposals for further research are presented.

7.1. Discussion of the Results

The results show no significant effect for neither of the three chosen subjects: first-choice form of Norwegian, second-choice form of Norwegian and English. This indicates that students' academic performance neither increase nor decrease as a result of the laptop program. However, as will be discussed in the following paragraphs, the effect of technology use on students is complex. Detecting spill-over effects on traditional subjects using school average as outcome variables would require an unambiguous effect for a sufficient share of students at a given school. The limitations to the sampled data, which will be discussed in Section 7.2, might prevent us from detecting such changes.

Technology, represented by the laptop program, can affect students' academic performance through several channels. Firstly, it increases the access to information, for example, it becomes easier for students to do online research and teachers can provide course content to students digitally. Secondly, it is likely to provide students with digital tools that can improve learning and performance. Examples are dictionaries, programs correcting grammar, calculators, and graphical mathematics tools. Thirdly, it might increase opportunities to communicate and collaborate with other students and take advantage of each others knowledge. All of the above will strengthen educational outcomes. However, students might also get distracted by the use of a laptop in school. The increased access to entertainment such as social media, games, and the world wide web creates a tempting and easily accessible online environment for the students during class.

The impact of technology use on a student's academic performance depends on various social factors. First, the academic motivation will be crucial for how the student benefits from the access to a laptop. While motivated students may achieve on average better grades, non-motivated students may achieve on average lower grades due to the easy access to distractions on laptops during class. In addition, the socioeconomic background might cause students to be affected differently. For example, some students can take advantage of previous knowledge with use of home computers, as well as the fact that parental monitoring and supervision might differ between students from different socioeconomic backgrounds.

The channels and social factors described above are pulling in different directions, and one could argue that the expected effect of the laptop program is ambiguous. Hence, concluding that there are no significant effects of the laptop program on students' academic performance in general could seem reasonable. However, performing the analysis on individual-level data could make it possible to detect the effect of technology use on academic performance on each individual student rather than the aggregated effect at school level. In addition, as discussed in Section 6.1, our estimates might suffer from the fact that the grading scale is integer and only runs from one to six. Thus, it would require a large change in a student's performance to shift the grade.

7.2. Limitations to the Data Set

The first limitation we want to emphasize is that the collected data might suffer from omitted variable bias (OVB). This occurs when variables that have an effect on academic outcomes and are correlated with the timing of the implementation of laptops are omitted from the equation (Angrist and Pischke, 2015). Although adding control variables and fixed effects for county and time reduce the potential threat of OVB, the bias can still be viable and will be discussed in the next two paragraphs.

In 2009, all upper secondary schools had the opportunity to carry out digital exams, and by 2012 the norm was for all exams to be digitized (Norwegian Directorate for Education and Training, 2011). Our data set contains information on whether the students have received an individual laptop, but not which schools carried out exams digitally or written by hand. We assume that the timing of the laptop implementation at schools and digital exams are highly correlated. Due to the fact that laptops were gradually implemented, students across the country faced different ways of conducting the same exam. There is reason to believe that an exam written on a laptop might have a positive effect on performance. For instance, it is easier to structure the paper when the exam is digital and the sensor will not be affected by the quality of the handwriting. Due to this, our estimates could be biased. However, the Norwegian Directorate for Education and Training has strict guidelines for fair censorship (Norwegian Directorate for Education and

Training, 2017c). This should indicate that they have higher requirements when it comes to structure and grammar for digital exams. We, however, believe that this can be considered a weakness in this study.

In addition, educational performance is decided by a person's individual characteristics. We have not sampled data at individual level for neither academic performance nor the control variables. Individual characteristics that could have improved the analysis are parents income and parents education which are associated with a child's abilities and academic outcome. However, data on these variables are only available to us at the county and municipality level. Because the students can influence the decision of which upper secondary school to attend and there are large discrepancies within a county, adding these variables would not improve our analysis. The lack of these variables at individual level could, however, cause OVB in the specification.

A second limitation is that we only have access to data on average grades at school level. Access to individual data could have enabled us to uncover differences between students, as discussed in Section 7.1. Academic level, motivational level and socioeconomic background are examples of factors that will cause a student's academic performance to be affected differently by the use of technology. Hence, using school averages as a measure of academic performance might prevent us from detecting these individual effects. This is because the school-level average grade will remain the same despite the fact that some individuals improve their performance, while others experience a decrease in performance.

A third limitation is that there are missing observations in our data, both related to grades in subjects over time and control variables. First, missing observations in the outcome variables will cause less precision to our estimates, and might bias our results (Wooldridge, 2012). The extent of the problem depends on the reason why they are missing. If the missing observations are consistent throughout the data set or random, it is not likely to create a problem. As discussed in Section 4.2, exams are randomly assigned by the county authorities, and the assignment needs to be evenly divided into subjects and schools over time and thus randomness in missing observations is likely to be true. However, the model loses statistical power due to a reduced sample. Second, the systematic missing variables in some of the background characteristics from 2003 to 2005, discussed in Section 4.4, is a weakness in our model. We aim to reduce the harm of this by replacing these missing observations with the best possible estimate, which we consider to be the average of the four following years. In addition, there are some missing observations in the remaining years of the data on the background characteristics, but as these are considered to be consistent throughout the data set or random, it is not likely to bias the results.

Referring to the discussion in Section 6.2 about uncertain observations, there might be a form of misreporting in regards to treatment year causing a fourth limitation to our data. We have estimated that just above 10 percent of the observations on treatment year have a larger probability of being misreported. However, we tested the impact of this on our specification without observing any significant changes in the estimates. In addition, some unidentified schools gave individual classes access to laptops before officially implementing the laptop program. This implies that the treatment is not necessarily a binary relationship at school level. Hence, both the risk of misreporting regarding treatment year and unsystematic adoption of the laptop program in some schools can reduce the precision of our estimates. This threatens the internal validity of the model.

A last limitation to the data is that it only allows us to study short-term effects of the program. Students stay in upper secondary school for an average of three years. As we have accesses to school-level data, not individual-level data, this limits us from tracking the same students over longer periods of time. Students in our data set have either been exposed to the laptop for one year when undertaking exams in English, or for up to three years when undertaking the exams in first-choice and second-choice form of Norwegian. This enables us to estimate the short-term effects of the laptop program, but not the long-term effects. Interesting long-term outcomes could be selection of studies in higher education or labor market outcomes.

7.3. Limitations to the Empirical Approach

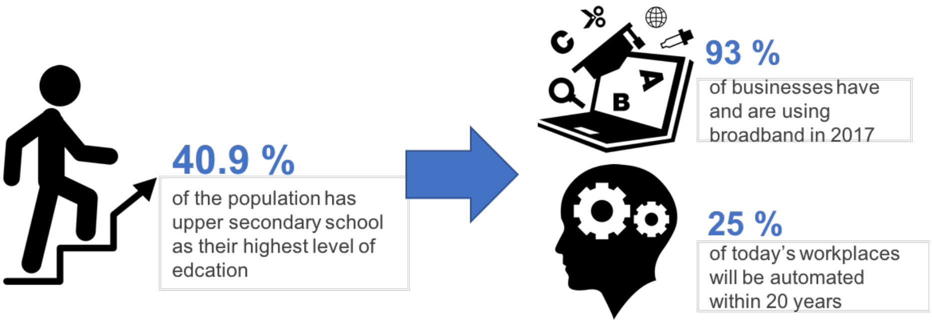
There might be limitations to our estimation strategy that undermines the validity of our results and violates the causal inference. We identify two main problems concerning our selected empirical approach. Firstly, there might be endogeneity between the laptop use in upper secondary schools and unobservable determinants of academic results. The rollout of the laptop program must be completely random, but this is hard to test and in most cases not completely true. An example of a violation can be that some counties have programs for teachers to make them better able to facilitate the students' laptop use as a part of the teaching.

Secondly, our thesis aims to find the effect of the use of technology in upper secondary school. We have assumed that the uptake rate in the upper secondary schools that initiate the laptop program is high, indicating that TOT effects and ITT effects are quite similar. However, as discussed in Section 7.2, the treatment is not necessarily a binary relationship. In addition, we do not have information on the degree of laptop use after the students are assigned a laptop. Thus, we have no information on the uptake rate, and cannot claim that we estimate more than the ITT effects. If the laptop program is not a strong predictor of actual laptop use in school, the effects from our analysis are severely underestimated.

7.4. Implications of the Study

Schools investing heavily in technology should not expect a drastic increase in students' academic outcomes (Bulman and Fairlie, 2016). Our main findings of no significant effects of technology use on educational outcomes are in line with previous studies discussed in the Literature Review (Chapter 3). Despite these findings, one cannot conclude that the laptop program has been an unsuccessful investment. There are other measures, which are not captured in academic performance, that are more likely to be affected, but more difficult to quantify and measure. As mentioned in the introduction, there is a rapid change in today's labor market and digital skills are essential, see Figure 7.1. Thus, making students more equipped to explicitly handle the future labor market through digital education is invaluable, but is not something we have been able to measure.

Figure 7.1 – Education and Labor Market Statistics



Notes: Digitization of upper secondary education is required to make students more equipped to handle the future labor market. Data from Ådne Cappelen et al. (2013), Statistics Norway (2017a) and Statistics Norway (2017b)

The decision makers have to consider whether the investments devoted to the laptop program are more efficient than expenditure on other educational inputs. The Government has to be thorough when deciding how to spend its scarce resources. Analyzing the effects of current technology initiatives in education can have implications for how educational policy is designed in the future. The Government can choose to invest less money in technological equipment if studies show that it has little or no effect on students' academic performance, i.e. is economically inefficient. However, it is difficult to judge whether the laptop program has been successful or not without being able to measure the students' digital skills and associated spill-over effects at individual level. The importance of digital skills is underlined by *the Knowledge Promotion*, and it is difficult to fulfill this objective without directly investing in technological equipment for students (Norwegian Ministry of Education and Research, 2006). Thus, a potential way of assessing the project's economic efficiency is to compare to other forms of technological investments such as computer rooms.

7.5. Further Research

Our study is limited to analyzing the short-term effects on the average grades in schools in three common core subjects: first-choice form of Norwegian, second-choice form of Norwegian and English. Hence, for further research it will be interesting to analyze other outcome variables and long-term effects. First and foremost, measures that capture digital literacy are of interest, but also variables on academic results in courses such as mathematics, other foreign languages, as well as subjects in social sciences including history and religion. In addition, it would be interesting to examine whether there are more students that choose technology oriented study fields as a result of being introduced to technology in upper secondary school. If the students obtain high digital competence before choosing higher education, the share of students applying for more technological study programs is likely to increase.

Furthermore, studying the effects of similar initiatives on other populations, such as students in lower levels of education and in other countries, could be of interest. Measuring the effect of laptops on younger students and their long-run outcomes would be highly relevant as younger students might react differently to the introduction of technology in education. Additionally, the effect of a laptop program might have a positive effect in other, less digitized, countries. However, the studies on less digitized countries presented in the Literature Review (Chapter 3) have similar findings to the ones presented in this paper.

Lastly, it would be interesting to study the effects of the laptop program at individual level. This would make it possible to control for more determinants of academic performance and hence obtain more precise estimates. This could make it possible to study the long-term effects of technology use in education, and hence whether the initiative makes individuals more capable of handling the future labor market, by measuring labor market outcomes.

8. Conclusion

This thesis aims to answer the following research question:

"What is the effect of use of individual laptops in upper secondary education on students' academic performance?"

The laptop program which was initiated in Norway during the 2000s contributed to making the upper secondary education the most digitized among the OECD-countries. In addition, Norway is the country with second highest expenditures per upper secondary student (OECD, 2017). To analyze whether the investments in technology in education are efficient use of public resources, we exploit the differences in the timing of the laptop program implementation across the country to estimate the causal effects on academic performance. For this purpose we use a generalized difference-in-difference approach and an event study specification. Performance is measured as an average grade at school level in three common core subjects of general studies in upper secondary education: first-choice form of Norwegian, second-choice form of Norwegian, and English. The sample consists of 289 public upper secondary schools for the school years 2003/2004 to 2015/2016.

The results presented in this thesis indicate that the use of individual laptops in the upper secondary education has no significant effect on academic performance. These findings are in line with previous literature in this field. It is therefore relevant to assume that the implementation of technology in the educational system in Norway has little effect on academic performance, at least when it is quantified as average grades at school level in the three common core subjects studied here. However, we suspect that a potential significant effect of the laptop program can be masked by contradictory effects of laptop use due to differences in for example motivation and socioeconomic background for students in the same school. As the data analyzed are average grades at school level, such individual differences in the response to the use of laptops will not be uncovered.

This paper is limited to analyzing the short-term effects on three traditional subjects. However, the findings can inspire future research in obtaining knowledge on long-term effects of technology use on other educational outcomes. In addition, the field of study is especially important when designing educational policy to prepare students for the future labour market.

9. Bibliography

Akerman, A., Gaarder, I., and Mogstad, M. (2015). The Skill Complementarity of Broadband Internet. *The Quarterly Journal of Economics*, 130(4):pp. 1781–824.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricists Companion* (1st edition). Princeton and Oxford: Princeton University Press.

Angrist, J. D. and Pischke, J.-S. (2015). *Mastering Metrics: The Path from Cause to Effect* (1st edition). Princeton and Oxford: Princeton University Press.

Asteriou, D. and Hall, S. G. (2011). *Applied Econometrics* (2nd edition). Palgrave Macmillan.

Bailey, M. J. and Goodman-Bacon, A. (2014). The War on Poverty's Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans. *National Bureau of Economic Research*, NBER Working Paper No. 20653.

Bulman, G. and Fairlie, R. W. (2016). Technology and Education: Computers, Software and the Internet. *National Bureau of Economic Research*, NBER Working Paper No. 22237.

Bütikofer, A., Løken, K. V., and Salvanes, K. G. (2012). Long-Term Consequences of Access to Well-child Visits. *Economics of Education Review*, 31(5):pp. 601–614.

Bütikofer, A., Mølland, E., and Salvanes, K. G. (2016). Childhood Nutrition and Labor Market Outcomes: Evidence from a School Breakfast Program. NHH Dept. of Economics Discussion Paper No. 15/2016.

Cristia, J., Ibarraran, P., Cueto, S., Santiago, A., and Severín, E. (2017). Technology and Child Development: Evidence from the One Laptop per Child Program. *American Economic Journal: Applied Economics*, 9:pp. 295–320.

Erstad, O., Kløvstad, V., Kristiansen, T., and Søby, M. (2005). *Digital Skole Hver Dag – om Helhetlig Utvikling av Digital Kompetanse i Grunnopplæringen* (Norwegian Centre for ICT in Education). Retrieved 20 September 2017 from `https://www.udir.no/Upload/Rapporter/5/ITU_rapport.pdf?epslanguage=no`.

European Commission (2013). Survey of Schools: ICT in Education Benchmarking Access, Use and Attitudes to Technology in Europe's Schools. Retrieved 28 September 2017 from https://ec.europa.eu/digital-single-market/sites/digital-agenda/files/KK-31-13-401-EN-N.pdf.

Faber, B., Sanchis-Guarner, R., and Weinhardt, F. (2015). ICT and Education: Evidence from Student Home Addresses. *National Bureau of Economic Research*, NBER Working Paper No. 21306.

Fairlie, R. W. and Kalil, A. (2016). The Effect of Computers on Children's Social Development and School Participation: Evidence from a Randomized Control Experiment. *National Bureau of Economic Research*, NBER Working Paper No. 22907.

Forskrift til Opplæringslova (2009). Forskrift 1 juli 2009 nr. 964 om Individuell Vurdering i Grunnskolen og i Vidaregåande Opplæring.

Frølich, T. H. and Vestby, G. M. (2003). *Ingen vei tilbake. Innovative læringsmiljøer med bruk av IKTerfaringsrapport fra Nesodden videregående skole* (Norwegian Centre for ICT in Education). Retrieved 8 December 2017 from http://www.ituarkiv.no/filearchive/fil_ITU_Rapport_14.pdf.

Hatlevik, O. E., Egeberg, G., Gumundsdóttir, G. B., Loftsgarden, M., and Loi, M. (2013). *Monitor Skole 2013 - Om Digital Kompetanse og Erfaringer med Bruk av IKT i Skolen* (Norwegian Centre for ICT in Education). Retrieved 10 September 2017 from https://iktsenteret.no/sites/iktsenteret.no/files/attachments/monitor_skole_2013_4des.pdf.

Hatlevik, O. E., Ottestad, G., Skaug, J. H., Kløvstad, V., and Berge, O. (2009). *Monitor 2009 - Skolens Digitale Tilstand* (Norwegian Centre for ICT in Education). Retrieved 10 Spetember 2017 from https://iktsenteret.no/sites/iktsenteret.no/files/attachments/itu_monitor_2009_fullstendig_rapport.pdf.

Hægeland, T., Raaum, O., and Salvanes, K. G. (2007). Pennies from heaven? - Using exogenous tax variation to identify effects of school resources on pupil achievement. NHH Dept. of Economics Discussion Paper No. 15/2016.

Knudsen, O. M. (2008). Her er PC-tilbudet i ditt fylke. *Tek.no.* Retrieved 9 October 2017 from https://www.tek.no/artikler/her_er_pc-tilbudet_i_ditt_fylke/83752.

Kose, E., Kuka, E., and Shenhav, N. (2015). Women's Enfranchisement and Children's Education: the Long-Run Impact og the U.S Suffrage Movement. IZA Discussion Paper No. 10148, Institute for the Study of Labor (IZA).

Malamud, O. and Pop-Eleches, C. (2011). Home Computer Use and the Development of Human Capital. *The Quarterly Journal of Economics*, 126(2):pp. 987–1027.

Norwegian Directorate for Education and Training (2011). *Erfaringer og Vurdering av Eksamen 2011 og 2012*. Retrieved 11 November 2017 from https://www.udir.no/globalassets/upload/eksamen/4/erfaringer_og_vurderinger_av_eksamen_2011_2012.pdf.

Norwegian Directorate for Education and Training (2016a). Karakterer i Videregående Opplæring. Published 16 November 2016. Retrieved 10 October 2017 from https://www.udir.no/tall-og-forskning/finn-forskning/tema/karakterer/karakterer-i-videregaende.

Norwegian Directorate for Education and Training (2016b). Rammeverk for Eksamen. Retrieved 10 October 2017 from https://www.udir.no/utdanningslopet/videregaende-opplaring/.

Norwegian Directorate for Education and Training (2016c). Trekkordning ved Eksamen for Grunnskole og Videregående Opplæring. Published 7 November 2016. Retrieved 5 October 2017 from https://www.udir.no/regelverk-og-tilsyn/finn-regelverk/etter-tema/eksamen/Udir-4-2016-trekkordning-ved-eksamen.

Norwegian Directorate for Education and Training (2017a). Eksamen. Retrieved 7 October 2017 from https://www.udir.no/eksamen-og-prover/eksamen/.

Norwegian Directorate for Education and Training (2017b). Rammeverk for Eksamen. Retrieved 5 October 2017 from https://www.udir.no/eksamen-og-prover/eksamen/rammeverk-eksamen/3.organiseringen-av-arbeidet-med-eksamen/.

Norwegian Directorate for Education and Training (2017c). Rammeverk for Eksamen - 4. Kvalitetssikring av Eksamen og Sensur. Retrieved 21 November 2017 from https://www.udir.no/eksamen-og-prover/eksamen/rammeverk-eksamen/4.kvalitetssikring-av-eksamen-og-sensur/.

Norwegian Directorate for Education and Training (2017d). Standpunktvurdering. Published 5 October 2017. Retrieved 10 October 2017 from https://www.udir.no/laring-og-trivsel/vurdering/sluttvurdering/standpunktvurdering/.

Norwegian Directorate for Education and Training (2017e). *Utdanningsspeilet - Tall og Analyse* *av Barnehager og Grunnopplæringen i Norge*. Retrieved 29 September 2017 from http://utdanningsspeilet.udir.no/2017/.

Norwegian Ministry of Education and Research (2006). *The Knowledge Promotion*. Retrieved 17 October 2017 from https://www.regjeringen.no/globalassets/upload/kilde/kd/bro/2006/0002/ddd/pdfv/292311-kunnskapsloftet2006_engelsk_ii.pdf.

Norwegian Ministry of Education and Research (2007). *Education – from Kindergarten to Adult Education*. Retrieved 7 October 2017 from https://www.udir.no/Upload/Brosjyrer/5/Education_in_Norway.pdf?epslanguage=no.

Norwegian Ministry of Education and Research (2012). *The Education Mirror 2012: Analysis of Primary and Lower and Upper Secondary Education in Norway*. Retrieved 10 October 2017 from https://www.udir.no/Upload/Rapporter/2012/US2012_ENG_nettversjon.pdf.

Norwegian Ministry of Education and Research (2017). *Framtid, Fornyelse og Digitalisering - Digitaliseringsstrategi for Grunnopplæringen 2017–2021*. Retrieved 15 October 2017 from https://www.regjeringen.no/contentassets/dc02a65c18a7464db394766247e5f5fc/kd_framtid_fornyelse_digitalisering_nett.pdf.

OECD (2015a). *PISA 2015 Key Findings for Norway*. Retrieved 6 December 2017 from http://www.oecd.org/norway/pisa-2015-norway.htm.

OECD (2015b). *Students, Computer and Learning: Making the Connection*. OECD Publishing, Paris. Retrieved 25 October 2017 from http://dx.doi.org/10.1787/9789264239555-en.

OECD (2017). Education spending. Retrieved 3 October 2017 from https://data.oecd.org/eduresource/education-spending.htm.

School Portal (2017). Rapportbygger - Videregående Skole. Retrieved 10 September 2017 from https://skoleporten.udir.no/rapportbygger.

Shapley, K., Sheehan, D., Sturges, K., Caranikas-Walker, F., Huntsberger, B., and Maloney, C. (2009). Evaluation of the Texas Technology Immersion Pilot: Final Outcomes for a Four-Year Study 2004-05 to 2007-08. (Prepared for Texas Education Agency). Retrieved 11 September 2017 from https://files.eric.ed.gov/fulltext/ED536296.pdf.

Silvernail, D. L. and Gritter, A. K. (2007). Maine's Middle School Laptop Program: Creating

Better Writers. Center for Educational Policy, Applied Research and Evaluation (CEPARE), University of Southern Maine, Gorham.

Statistics Norway (2017a). ICT Usage in Enterprises. Published 19 August 2017. Retrieved 13 September 2017 from `https://www.ssb.no/teknologi-og-innovasjon/statistikker/iktbrukn`.

Statistics Norway (2017b). Key Figures on Education. Published 26 September 2017. Retrieved 7 November 2017 from `https://www.ssb.no/en/utdanning/nokkeltall`.

Statistics Norway (2017c). Upper Secondary Education. Published 14 March 2017. Retrieved 1 September 2017 from `https://www.ssb.no/en/utdanning/statistikker/vgu`.

Statistics Norway (2017d). Videregående opplæring - nøkkeltall. Retrieved 10 September 2017 from `https://www.ssb.no/statistikkbanken/selectvarval/Define.asp?subjectcode=&ProductId=&MainTable=Kostra2F8299CVid&nvl=&PLanguage=0&nyTmpVar=true&CMSSubjectArea=utdanning&KortNavnWeb=vgu&StatVariant=&checked=true`.

Suhr, K. A., Hernandez, D. A., Grimes, D., and Warschauer, M. (2010). Laptops and Fourth-Grade Literacy: Assisting the Jump over the Fourth-Grade Slump. *Journal of Technology, Learning, and Assessment*, 9(5). Retrieved 8 September 2017 from http://www.jtla.org.

Wooldridge, J. M. (2012). *Introduction to Econometrics* (Europe, Middle East & Africa Edition). Hampshire: South-Western Cengage Learning.

Ådne Cappelen, Gjefsen, H., Gjelsvik, M., Holm, I., and Stølen, N. M. (2013). *Forecasting Demand and Supply of Labour by Education* (Statistics Norway). Retrieved 2 September 2017 from `https://www.ssb.no/en/arbeid-og-lonn/artikler-og-publikasjoner/_attachment/144096?_ts=141cb387d78`.

10. Appendix

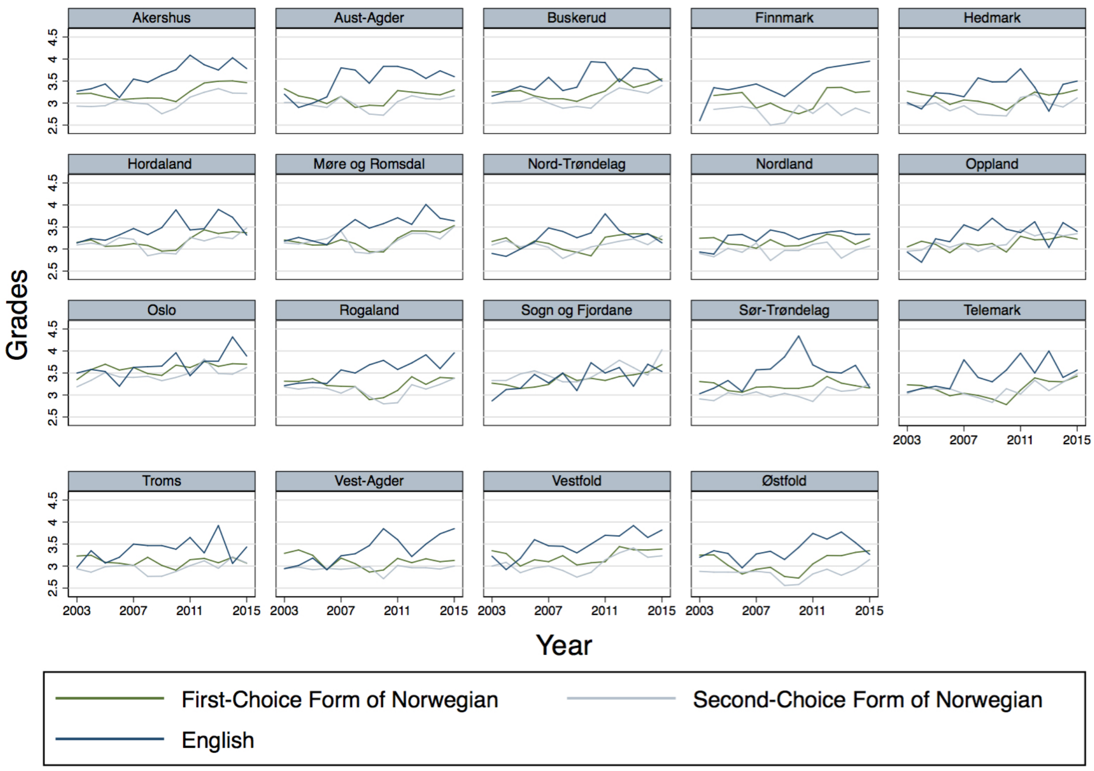
Table A1 – List of Acronyms

ICT	Information and Communication Technology
ITT	Intention-to-Treat
NOK	Norwegian Kroners
OECD	Organisation for Economic Co-operation and Development
OVB	Omitted Variable Bias
PISA	Programme for International Student Assessment
SSB	Statistics Norway
TOT	Treatment-on the-Treated
VET	Vocational Education Programme and Training

Table A2 – List of Control Variables

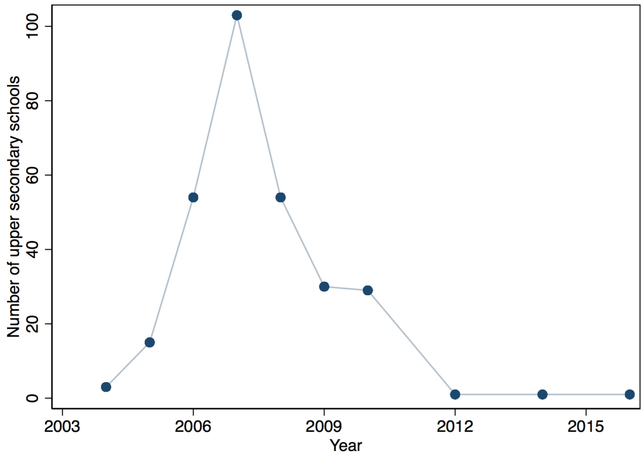
	Time Period	Description
School Characteristics		
Number of students	2003-2015	Displays the number of students in each school at the beginning of every school year and is divided into the number of boys and girls. Data are collected from the School Portal.
Boyshare	2003-2015	Based on the number of boys and girls at each school we have constructed a variable displaying the share of boys at each school.
Dropout rate	2006-2015	Displays the yearly dropout rate at each school. Data are collected from the School Portal.
The Pupil Survey	2006-2015	Displays data on the student's well-being and learning environment at each school. It measurs the following variables: well-being with teachers, social well-being, student's democracy, physical learning environment, bullying at school, motivation, professional guidance, participation in decisions regarding own education, career guidance, and academic challenge. The scale used in the survey is ordinal from 1 to 5. Data are collected as an average at school level and are collected from the School Portal.
County Characteristics		
Expenditure per student	2003-2015	Displays the expenditure per student per school year in upper secondary education in each county. The variable is given in NOK. Data are collected from Statistics Norway.
Student-to-teacher ratio	2003-2015	Displays how many students there are per teacher in a county. Data are collected from Statistics Norway.

Figure A1 – Average Grades by County



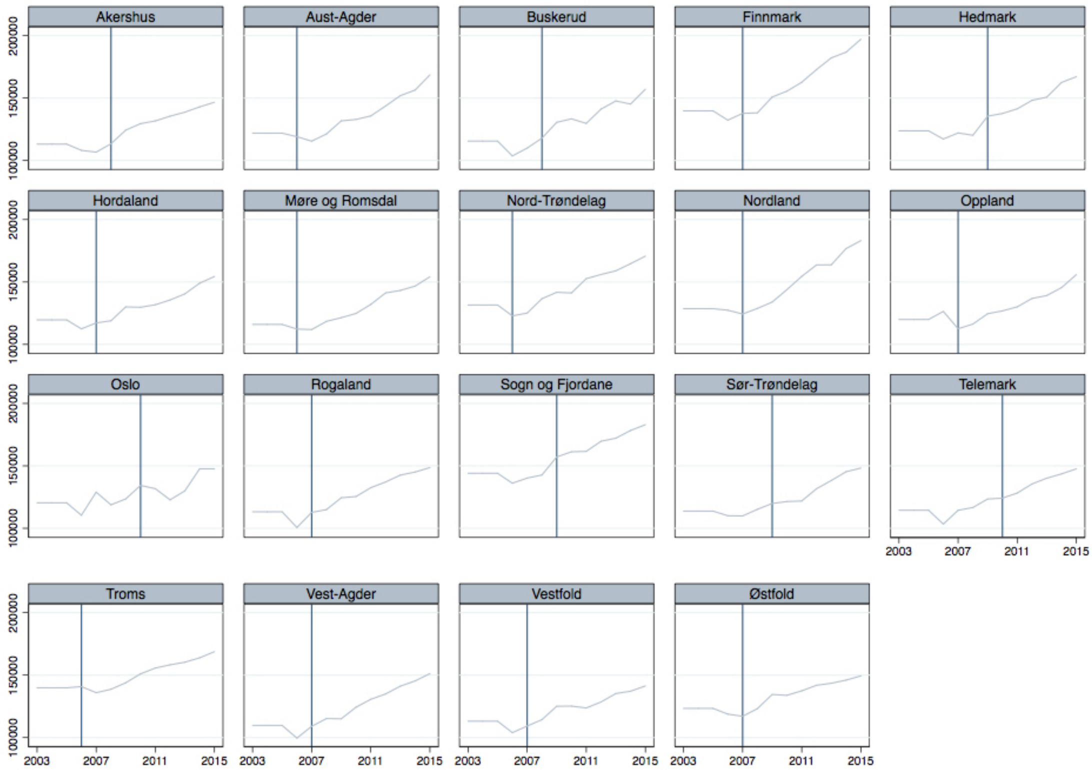
Notes: The figure displays average grades in first-choice form of Norwegian, second-choice form of Norwegian and English by county from 2003 to 2015. The grading scale is ordinal from one to six. Data are provided from the Norwegian Directorate for Education and Training's School Portal (2017).

Figure A2 – Number of Schools Implementing the Laptop Program



Notes: The figure shows the number of schools implementing the laptop program each year in the period 2003 to 2015.

Figure A3 – Expenditure per Student per County



Notes: The figure displays school expenditure in NOK per student by county for the time period 2003 to 2015. The vertical line marks the year the laptop program on average was introduced in each county. Data are collected from Statistics Norway (2017d).

Table A3 – Rollout Estimates: Standard Errors Clustered at County Level

	(1) First-Choice form of Norwegian	(2) Second-Choice form of Norwegian	(3) English
Laptop program	0.0080	0.0494	-0.0405
	(0.1010)	(0.1130)	(0.0668)
Length of treatment	0.0099	-0.0081	
	(0.0391)	(0.0476)	
Controls	Yes	Yes	Yes
No. of clusters	19	19	19
Observations	3219	3052	1592

Significance levels: *** 1% level, ** 5% level, * 10% level

Notes: Each parameter is from a separate regression of the outcome variables, which are average grades, on the laptop program and length of treatment, based on the model in Equation 5.1. The grading scale is ordinal from one to six. Robust standard errors adjusted for clustering at the level of the county are shown in parentheses. The sample includes schools in the time period 2003 to 2015. All specifications include a set of county and time fixed effects, and the following control variables: number of students, share of boys, dropout rate, score on each variable in *the Pupil Survey*, expenditure per student per county and student-to-teacher ratio.

Table A4 – Rollout Estimates: One Independent Variable of Interest.

	Panel A: Laptop Program	
	(1)	(2)
	First-Choice form of Norwegian	Second-Choice form of Norwegian
Laptop program	0.0354	0.0271
	(0.0308)	(0.0371)
Controls	Yes	Yes
No. of clusters	288	287
Observations	3219	3052
	Panel B: Length of treatment	
	(1)	(2)
	First-Choice form of Norwegian	Second-Choice form of Norwegian
Length of treatment	0.0126	0.0088
	(0.0113)	(0.0136)
Controls	Yes	Yes
No. of clusters	288	287
Observations	3219	3052

Significance levels: *** 1% level, ** 5% level, * 10% level

Notes: Each parameter is from a separate regression of the outcome variables, which are average grades, on the laptop program or the length of treatment, based on the model in Equation 5.1. The grading scale is ordinal from one to six. Robust standard errors adjusted for clustering at the level of the school are shown in parentheses. The sample includes schools in the time period 2003 to 2015. All specifications include a set of county and time fixed effects, and the following control variables: number of students, share of boys, dropout rate, score on each variable in *the Pupil Survey*, expenditure per student per county and student-to-teacher ratio. Panel A provides the estimated effects of laptop program only; Panel B provides the estimated effects of length of treatment only.

Table A5 – Test of Key Identifying Assumptions

	(1) First-Choice form of Norwegian	(2) Second-Choice form of Norwegian	(3) English
Laptop program	-0.1045	0.0050	-0.1095
	(0.1492)	(0.1789)	(0.1666)
Length of treatment	0.0525	-0.0009	
	(0.0542)	(0.0632)	
Controls	Yes	Yes	Yes
No. of clusters	204	193	150
Observations	1024	938	444

Significance levels: *** 1% level, ** 5% level, * 10% level

Notes: Each parameter is from a separate regression of the outcome variables, which are average grades, on the laptop program and length of treatment, based on the model in Equation 5.1. The counties of Akershus, Buskerud, Finnmark, Hedmark and Sogn og Fjordane are excluded. The grading scale is ordinal from one to six. Robust standard errors adjusted for clustering at the level of the school are shown in parentheses. The sample includes schools in the time period 2003 to 2015. All specifications include a set of county and time fixed effects, and the following control variables: number of students, share of boys, dropout rate, score on each variable in *the Pupil Survey*, expenditure per student per county and student-to-teacher ratio.

Table A6 – Rollout Estimates: Divided by Gender

		Panel A: Male	
	(1)	(2)	(3)
	First-Choice form of Norwegian	Second-Choice form of Norwegian	English
Laptop program	-0.0358	-0.0094	0.0056
	(0.0859)	(0.1170)	(0.0810)
Length of treatment	0.0319	0.0211	
	(0.0328)	(0.0429)	
Controls	Yes	Yes	Yes
No. of clusters	271	260	217
Observations	2779	2456	772
		Panel B: Female	
	(1)	(2)	(3)
	First-Choice form of Norwegian	Second-Choice form of Norwegian	English
Laptop Program	0.0066	0.0616	0.0847
	(0.0981)	(0.1420)	(0.0935)
Length of treatment	0.0020	-0.0289	
	(0.0380)	(0.0510)	
Controls	Yes	Yes	Yes
No. of clusters	283	275	223
Observations	2881	2594	870

Significance levels: *** 1% level, ** 5% level, * 10% level

Notes: Each parameter is from a separate regression of the outcome variables, which are average grades, on the laptop program and length of treatment, based on the model in Equation 5.1. The grading scale is ordinal from one to six. Robust standard errors adjusted for clustering at the level of the school are shown in parentheses. The sample includes schools in the time period 2003 to 2015. All specifications include a set of county and time fixed effects, and the following control variables: number of students, share of boys, dropout rate, score on each variable in *the Pupil Survey*, expenditure per student per county and student-to-teacher ratio. Panel A provides the estimated effects for male only; Panel B provides the estimated effects for female only.

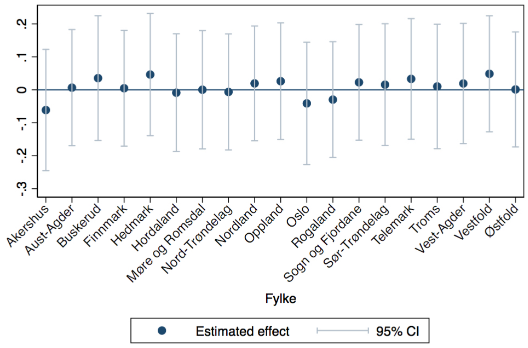
Table A7 – Regression Sensitivity Analysis

	(1) First-Choice form of Norwegian	(2) Second-Choice form of Norwegian	(3) English
Panel A: Excluding Outliers			
Laptop program	-0.0564	-0.1052	0.0122
	(0.0860)	(0.1138)	(0.0539)
Length of treatment	0.0306	0.0378	
	(0.0305)	(0.0410)	
Controls	Yes	Yes	Yes
No. of clusters	285	286	272
Observations	2814	2688	1448
Panel B: Excluding Uncertain Observations			
	(1) First-Choice form of Norwegian	(2) Second-Choice form of Norwegian	(3) English
Laptop Program	0.0292	0.0855	-0.0937
	(0.0909)	(0.1090)	(0.0623)
Length of treatment	0.0043	-0.0267	
	(0.0326)	(0.0396)	
Controls	Yes	Yes	Yes
No. of clusters	263	262	248
Observations	2922	2760	1413

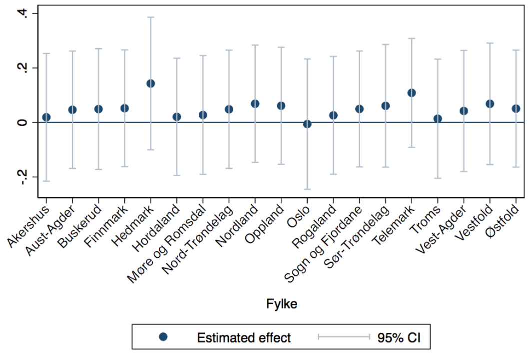
Significance levels: *** 1% level, ** 5% level, * 10% level

Notes: Each parameter is from a separate regression of the outcome variables, which are average grades, on the laptop program and length of treatment, based on the model in Equation 5.1. The grading scale is ordinal from one to six. Robust standard errors adjusted for clustering at the level of the school are shown in parentheses. The sample includes schools in the time period 2003 to 2015. All specifications include a set of county and time fixed effects, and the following control variables: number of students, share of boys, dropout rate, score on each variable in *the Pupil Survey*, expenditure per student per county and student-to-teacher ratio. Panel A provides the estimated effects when upper and lower 5% percentile are excluded; Panel B provides the estimated effects when uncertain observations are excluded.

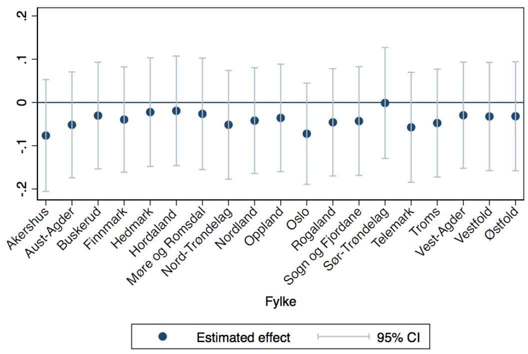
Figure A4 – Regression Sensitivity Analysis: Leaving Out Counties One at a Time



(a) First-Choice form of Norwegian



(b) Second-Choice form of Norwegian



(c) English

Notes: Each plot in the figure is from a separate regression where each county is excluded one at a time. The county excluded is indicated on the x-axis. The plots display estimated results from Equation 5.1, as well as the 95 percent confidence interval. The set of control variables includes county and time fixed effects, as well as number of students, share of boys, dropout rate, score on each variable in *the Pupil Survey*, expenditure per student per county and student-to-teacher ratio. The standard errors are cluster at school level and robust. The period of interest is 2003 to 2015. Figure (a) shows the effects on first-choice form of Norwegian, figure (b) shows the effects on second-choice form of Norwegian, and figure (c) shows the effects on English.

Table A8 – Regression Sensitivity Analysis: County-Specific Linear Time Trends

	(1) First-Choice form of Norwegian	(2) Second-Choice form of Norwegian	(3) English
Laptop program	-0.0093	0.0208	-0.0364
	(0.0911)	(0.1099)	(0.0614)
Length of treatment	0.0164	0.0060	
	(0.0355)	(0.0420)	
Controls	Yes	Yes	Yes
No. of clusters	288	287	273
Observations	3219	3052	1592

Significance levels: *** 1% level, ** 5% level, * 10% level

Notes: Each parameter is from a separate regression of the outcome variables, which are average grades, on the laptop program and length of treatment, based on the model in Equation 6.1. The grading scale is ordinal from one to six. Robust standard errors adjusted for clustering at the level of the school are shown in parentheses. The sample includes schools in the time period 2003 to 2015. All specifications include a set of county and time fixed effects, and the following control variables: number of students, share of boys, dropout rate, score on each variable in *the Pupil Survey*, expenditure per student per county and student-to-teacher ratio.