Linnæus University
Sweden

Master Thesis

Consequences of distance learning
- *Effects of the Corona pandemic on students' grades*

Author: Jonas Martinsson
Supervisor: Joakim Jansson
Examiner: Tobias König
Term: VT21
Subject: Economics
Level: Master Thesis
Course code: 4NA06E

Linnæus University
Sweden

Abstract

The Corona pandemic and Covid-19 has affected the entire globe where we had to adapt to a new way of living and a new standard of doing things. For the universities in Sweden and around the world this generally meant that the universities had to switch to online teaching and digital examinations to the largest extent possible. Even though we have lived in this new kind of every day we know barely anything about how this has affected the grades of the students, and what we can learn about it to do things more effectively in the future.

In this paper, I examine the possible effects on grades for students at the university level by using newly collected data from Linnaeus University's department of economics and statistics.

Results show that the grades after the decision to switch to online teaching and digital exams in March of 2020, overall became higher for both males and females but no significant difference between the genders were found.

Key words

Linnæus University
Sweden

Acknowledgments

Linnæus University
Sweden

Table of contents

1 Introduction

During the first quarter of 2020, a pandemic developed when the new virus SARS-CoV-2 hit the world, which meant that society as a whole had to adapt. This meant that, for instance, universities had to change teaching and examinations. On March 16th, Linnaeus University, along with many other universities around the globe during this month, made the decision to switch to digital teaching and exams as far as possible following recommendations from the Swedish Public Health Agency[1]. The decision also meant that all teaching and examinations would take place online to the highest extent possible, which led to students being even more responsible for their own studies. Since this global pandemic has affected us all in some way, it is an area that is important to investigate further, to help us if we face another situation like this but also if there are some part that could be more effective than it was before the pandemic, such as the teachers pedagogical approach mentioned by Habragoda (2020).

Even though this decision was made at numerous universities and affected numerous students we know barely anything about how these decisions affected the quality of the education and therefore also the grades of the students. Especially since early studies like Engzell, Frey & Verhagen (2021) and Todrican (2020) show that many institutions had a hard time to adapt to the online teaching.

Along with this, the digitalization made it impossible for exams in halls where anonymous exams were conducted to a large extent, so the exams now again became non-anonymous. In the autumn of 2016, to prevent discrimination and have a fairer grading of the students, the head of school at Linnaeus University decided on a grading reform with start on the spring semester of 2017[2]. The purpose with the reform was to remove the identity of the test takers on standard exams, so that the students only wrote a code on the taken exams with no other revealing information.

---

[1]Press release from the Head of School at Linnaeus University regarding the decision to switch to online teaching and examinations:
https://lnu.se/mot-linneuniversitetet/aktuellt/nyheter/2020/linneuniversitetet-gar-over-till-digital-undervisning-och-tentamen-med-anledning-av-corona/

[2] https://lnu.se/mot-linneuniversitetet/aktuellt/nyheter/2016/anonyma-salstentor/

Since both students and teachers now had to adapt to a new standard, where teachers had to develop a new skill in terms of tutoring from distance, and students who now have even more responsibility over their own studies, there are reasons to believe that this also affects the grades of the students. With this as a background, this study will examine and see if there are effects in grades as a result of this digitalization and distance education between males and females. Looking at some previous studies done regarding grading and girls and females' grades in elementary and secondary school in Sweden, they usually perform better than boys and males. Despite this, they generally have lower wages in the labour market which can be a pretty hard evidence against discrimination of females in the labour market. It might also induce discrimination against boys and males in grading, which makes the area of gender differences interesting for research.

To investigate this eventual effect, I will use new data from Linnaeus University's department of economics and statistics and compare exams graded both anonymously and non-anonymously between the fall term of 2015 up to the spring term of 2021 at the same university to see if there is any effect in grades due to the corona pandemic and also to see if there are any differences between males and females. I will use this shock, the pandemic, as a natural policy treatment in a difference-in-difference model.

Results shows a clear effect on the grades after the decision to switch to online teaching and digital exams in March of 2020, were it overall became higher for both males and females. Results did not, however, show any significant difference between the genders.

A regression on the examination activities that only included multiple-choice questions was also made in order to control for discrimination in grading. Here we could also see an increase in grades for both males and females, but additionally, these results also show that females benefited more from this kind of examinations. The changes in grades that were found might however depend on different reasons.

Regarding grades on activities that is graded non-anonymously in a normal state as well, results show that students actually received higher grades over time, although to a small extent. However, we cannot say that we have a difference between the genders and the results are more likely implying that the grades regarding these activities have been rather stable over time. These results could be an indicator that students actually been studying harder over time, which would have been affecting all kinds of examinations.

As stated above, the changes in grades that were found might depend on different reasons, and one eventual problem with the non-anonymous exams as a result of the online education is that of discrimination. So, this could be a possible explanation for some part of the changes in students' grades.

Discrimination and bias have also always existed in our society and for several years studies have been done within this area, such as Ahmed, Andersson & Hammarstedt (2010), Åberg & Boström (2020) and Bengtsson, Tyrefors & Iverman (2012) as well as Ahmed & Hammarstedt (2020) and Carlsson & Rooth (2007) where discrimination is demonstrated in various fields in our society. While the previous studies focus on ethnical discrimination, studies like Lumley (2015), Triana et al. (2019), Alteri (2016) and Bilan et al. (2020) focuses on gender discrimination.

Lumley (2015) shows how gender equality has evolved conceptually and provides an overview of the educational initiatives in place for raising gender awareness and Alteri (2016), the author looks into the Acts of Patient Protection and Affordable Care Act and the Health Care and Education Reconciliation Act that were signed into law while Triana et al (2019), authors examine how the context influences the relationship between employees' perceptions of gender discrimination and outcomes at work.

Investigating the presence of discrimination is important since discrimination or bias can be a reason for a future set back in the labour market where the individual is subsequently discriminated against in terms of salary and employment. According to the theories from Becker ([1957] 1971) and Phelps (1972), all types of discrimination is either taste-based or statistical. One could then argue that statistical discrimination would not exist at the university level, where this paper is conducted, as they are not the same type of profit-maximizing organization as traditional firms, but an examiner may unintentionally give a grade that is either higher or lower than the specific individual actually deserves depending on the type of group the individual belongs to. This gives us reason to believe that it might also be present in the school system since we know from previous studies, such as Tyrefors et al. (2019), Tyrefors, Höglin & Johannesson (2011) and Feld et al. (2015) that bias from the graders comes into play when the grader has information on the exam taker. Therefore, it was reasonable to believe that also the non-anonymous exams would affect the grades of the students.

One reason that results did not show any signs of discrimination in this case might be because of our relatively small sample but also that this paper is done at an institution that researches within these areas itself, indicating that the teachers at this institution has more awareness regarding discrimination and education.

Other explanations for the results found is that certain individuals may co-op with the distance tutoring in a better way than others and some teachers might tackle this situation better than other teaches which also might affect the students' taking the courses. There are also reasons to believe that cheating has increased since now, all the exams are taken at home with no supervision over the students.

Studies within the area of cheating, like Simha (2014), McKendall et al (2020), Lento et al (2018) and Case et al (2019) examines the extent of cheating and why students do it. The findings of Case et al (2019) shows that participation in e-cheating is on the rise and that two-thirds of students perceive that it is easy to cheat. Their findings also shows that males are more prone to cheat than females which is in line with the findings of Lento et al (2018) that examines differences between male and female accounting faculty members' perceptions of academic dishonesty and their uses of controls to prevent academically dishonest behaviour.

Since the main reason for this study is to examine differences in grading, as a result of the pandemics non-anonymous exams, between males and females, the research question for this essay will be: *"Do the at home teaching as a result of the Corona-pandemic have an effect on student grades at the university level? "*

As this is still an ongoing pandemic, there are still many unexplored areas regarding Corona's effect on student grades so far. Despite the ongoing pandemic, it is still measurable as this area does not suffer from the same information lag in the same way as other areas do, since we can measure and compare the grades right away. Early studies done in this area, such as Chaudhary (2020), Rigo (2020), Todrican (2020) and Root & Rehfeldt (2021) show that many teachers lack experience regarding online teaching and that schools have had a hard time to adapt to the online education. Jugănaru (2020) also reveals that many students, the longer the pandemic has lasted, actually miss the physical school experience. This shows that there may be reasons to believe that students' grades have been affected by the pandemic, which has also demonstrated in the Netherlands in a study of Engzell, Frey & Verhagen (2021).

As stated above, the decision to switch to online teaching was made at numerous universities and affected numerous students but still, we know nothing about how these decisions affected the quality of the education and therefore also the grades of the students. Therefore, the results found in this study will contribute to this area with the help from this newly collected data.

Since the brief theory behind this kind of study is mentioned in the introduction the planned structure of the essay would be as follows: In part two we will go through the literature used as a base for this study. In the third part I will describe the methodological framework used and the fourth part will explain and describe the data used. The fifth and sixth part will consist of the results found and the conclusions drawn from it.

2 Literature Review

During the previous year there have been some studies made regarding the effect of the pandemic in the area of education. Studies like Engzell, Frey & Verhagen (2021), Habragoda (2020) and Todrican (2020) show that many institutions had a hard time to adapt to the online teaching suggesting that this might affect the students' grades, since student and teachers had to adapt to a new kind of tutoring. As stated in the introduction, this global pandemic has affected us all in some way, it is an area that is important to investigate further, to help us if we face a new situation like this but also if there are some part that could be more effective than it was before the pandemic. While the early studies made contributes to this study, this study can contribute with new results and information regarding the Coronas effects on students' grades at the university level as well as other limitations with online home exams.

Studies on discrimination, both gender and ethnic discrimination, have been made over the years, both within the school system and also in other areas, as stated above. As these types of discrimination exist throughout our society, these studies can provide some information and hint at current explanations regarding the possible effects of the grading as a result of the Corona pandemic. At the same time, this study can contribute with new information regarding discrimination due to the corona pandemic.

So, in the following section, I will briefly present and describe some of the studies that have been done previously that I find interesting and important for this paper, divided first into literature regarding effects of Corona on Education and then regarding discrimination within the school system.

2.1 Literature regarding effects of Corona on Education

In a study, Engzell, Frey & Verhagen (2021) evaluate the effect of school closures due to the Corona Pandemic on primary school performance in the Netherlands. Results reveal a learning loss of about 3 percentile points or 0.08 standard deviations. Losses were up to 60% larger among students from less-educated homes, confirming worries about the uneven toll of the pandemic on children and families. The findings imply that students made little or no progress while learning from home and suggest losses even larger in countries with weaker infrastructure or longer school closures.

Vlachos, Hertegård & Svaleryd (2021) compared Swedish upper-secondary schools that moved to online instruction, to lower-secondary schools that remained open since parents and teachers were differently exposed to open and closed schools, but otherwise facing similar conditions. They found that, among parents, exposure to open rather than closed schools resulted in a small increase in PCR-confirmed. The results for parents indicate that keeping lower-secondary schools open had minor consequences for the overall transmission of SARS-CoV-2 in society.

In the paper Rigo (2020), the author deals with the use of Edmodo (an educational social network) as a distance learning substitution tool in teaching English within university English courses during the first wave of COVID-19 in Slovakia. The findings present students' perception of Edmondo pointing to its benefits, weaknesses, and overall effectiveness as a communication platform within distance learning.

In the study of Chaudhary (2020), the author examines the effect of information & communication technology (ICT) orientation on university's brand in times of COVID-19. Results show that the perceived service quality acts as a significant but negative moderator in the relationship between ICT orientation and student's satisfaction and further reveal that the trust acts as a positive moderator in the relationship between student's satisfaction and university's brand performance.

In the paper Todrican (2020), the author analyses the impact that the spread of Coronavirus has on students in Bucharest during the period of emergency. Results suggests that the reopening of school units and the resumption of the usual activities will be an unrealistic expectation and that the educational system is and will be affected by the evolution of the SARS-CoV-2 epidemic.

In Root & Rehfeldt (2021) the authors of the paper talks about the rapid enrolment in online courses and environmental and biological determinates, such as Covid-19, that challenge how universities respond to education. And that this way of learning has empirical support from the past, Skinner (1968), where the investigations evaluated the validity of programmed instruction in online courses, and student perceptions of the online modalities used.

In their paper, Bogoslov & Lungu (2020), summarize the main obstacles regarding the distance learning processes in Europe, due to the Corona pandemic. Conclusions states that geographical locations represent critical factors with less developed socio-economic environments hardly able to engage in remote learning. As well, the pedagogical side is not fully prepared for a change that involves the transition of traditional learning in the digital environment.

In the research study Boyer-Davis (2020) the author tried to understand technostress in higher education and the consequences to faculty. Results show a significant difference in overall faculty perceived technostress scores before the Covid-19 pandemic and during the health crisis.

In a paper, Habragoda (2020) the author attempts to explore key challenges confronted with the implementation of online pedagogical approach in Sri Lanka, due to the Corona pandemic. Results show that lack of training in pedagogy for online teaching is the severest challenge faced by many institutions since most instructors were new to online teaching. The authors states that instructors must receive proper professional training and development to have higher expectations and to adapt their teaching to appropriate online teaching strategies.

In the paper Jugănaru (2020), the author draws attention to an analysis of the online education trends at the international level following the Corona pandemic. Results shows that at first, students were delighted with the idea of staying at home and continuing their school activities. Gradually, most students have stated that they want to return to the form of face-to-face educational activity. Beliefs that the progress in information and communications technology will be increasingly assimilated into the classic system of higher education training creates a need to identify and develop the types of teaching activities which can be done better online than in the traditional way.

In the article McMullen et al. (2020), the authors summarize tips to teach online students effectively following the Corona pandemic. According to the authors, instructors will need to adopt strategies to orient new learners, humanize the online environment, and maintain engaging course material. If instructors do this with success, then they will be able to reach and educate students wherever they may be.

In another study, Sales et al (2020) the authors analyse the perceptions of Social Sciences faculty about their own information and digital competence and that of their students, as well as the changes that have occurred as a consequence the coronavirus pandemic. The results show that the faculty members participating in this study have a critical vision of the information and digital competence of the student body, indicating that they lack capacity for evaluation, critical use, and communication. The authors also state that the effort made towards the online development of teaching as a consequence of the pandemic has been a shock for faculty, who declare a positive attitude towards reviewing their own digital informative practices to adapt their teaching.

2.2 Literature regarding gender and ethnical differences in the schooling system

In their article, Buchmann, DiPrete & McDaniel (2008) the authors review the empirical research and theoretical perspectives on gender inequalities in educational performance and attainment from early childhood to young adulthood and concludes by recommending three directions for future research. First interdisciplinary efforts to understand gender differences in cognitive development and non-cognitive abilities in early childhood. Then research on the structure and practices of schooling, and finally analyses of how gender differences might amplify other kinds of inequalities, such as racial, ethnic, class, or nativity inequalities.

In Dunn-Jensen et.al (2016) the authors present an experiential activity that management instructors can use to help students understand and appreciate the reality and power of unconscious bias. The basic framework of this activity can then easily be adapted to focus on other types of unconscious bias and stereotyping to help students who will become leaders in organizations.

In this study, Jurges & Schneider (2011) the authors used data from the German PISA 2000 extension study and analysed the effect of relative age on track recommendations and actual choice. Younger pupils and boys were less likely to be recommended to and enrolled in the academic track and the authors found no convincing evidence that postponing the recommendation by lengthening primary school by 2 years reduces the age or gender bias.

Lavy (2008) tests to see if there were gender stereotyping and discrimination by Israeli high-school teachers. Lavy found that male students were discriminated against in each subject which was contrary to expectations based on previous studies. This widened the achievement difference between females and males, since females typically outperform boys in every subject.

In another study, Lavy & Angrist (2009), the authors tried to use cash incentives to increase certification rates among low achieving students. This was a certificate necessary for post-secondary schooling. Their results showed a substantial increase in certification among girl, which lead to an increased likelihood of college attendance, but had no effect on boys.

Tyrefors & Jansson (2020) presents evidence of female university students benefiting from being graded anonymously, this in conflict with previous results found at secondary school level. The authors found an effect of lower grades due to anonymization for both genders, but also that the anonymization raised female grades relative to male. They conclude with stating that they found evidence of overall bias against female students, also in contrary to most of the literature din before, at lower schooling level.

In their article, Tyrefors et.al. (2019) the authors compare students' final mathematics grade with new data from a high school introductory test score in mathematics. Both the grades and the test were based on the same syllabus and the found results show a substantial grading difference with grading bias against boys.

Tyrefors, Höglin & Johannesson (2011) studied discrimination against boys in Swedish high schools which is interesting since girls are the ones typically having higher grades and outperform boys in school but generally have lower wages later on the labour market. The authors found that female students had higher grades than male students in both the non-blind test score as well as the blinded test score. There was also a clear tendency of an overall downgrading for both genders in the blind grading, but they found no evidence of discrimination against neither boys nor girls in grading.

In the paper, Keng (2020), the author uses administrative data from a public university in Taiwan to examine gender bias in teaching evaluations. They test for statistical discrimination against female instructors and results show that statistical discrimination is a significant source of gender bias in teaching evaluations. The results also suggest that the gender gap in teaching evaluations changes over time as male and female students evaluate male and female instructors differentially.

In this paper, Huskanovic & Nur (2020) the authors investigate how anonymous exams impact the grades for foreign-named students and whether there are differences in this treatment effect against them when considering the gender of the graders. Results show that there was no negative treatment effect against foreign-named students, meaning that the foreign-named students did not gain from the reform.

In their paper, Tyrefors, Höglin & Johannesson (2015), analyse national tests in the Swedish language graded both non-blindly by the student's own teacher and then again blindly, without any identifying information. The found results showed that, on average, the non-blind test score were 8% higher for the students with Swedish backgrounds compared to the foreign background students and that the non-blind grading favoured the students with Swedish backgrounds where the test scores increased by 16% compared to the 4% increase in the non-blind test score for students with foreign background.

Feld et al. (2015) tried in their paper to investigate if differential arise because agents discriminate against others or because they favour their own kind, exophobia or endophilia. They found favouritism but no discrimination by nationality nor by gender.

In this paper, Van Ewijk (2011), the author tries to explain why previous research shows that ethnic minority students perform poorer in school when they are taught by teachers belonging to the ethnic majority. The found results show that such a direct grading bias does not occur but instead finds indirect evidence for alternative explanations, such as that teachers report lower expectations that likely affect their behaviour towards minority students.

3 Methodological framework

I will use the decision to conduct teaching and examinations digitally and at a distance taken at the department of Economics and Statistics at Linnaeus University in March 2020 and use data from the same university which consists of graded activities in the courses at the department of Economics and Statistics from the fall term of 2015 to the spring term of 2021.

To perform this test, I will use a difference-in-difference method (DID), which is a quasi-experimental technique and the technique that I find most usable to estimate the effect that the digital exam has on a student's grade.

The DID is a method used to study the changes that occur as a result of a specific regulation or reform, in this case, the decisions to have tutoring and exams from home. In order to use the DID-method, you have to use data for different groups (males and females), and then compare them to each other over time. Here, I also have the underlying assumption that the trends are parallel between the two groups. The effects can then be revealed by comparing how the grades between males and females changes after the introduction of anonymous exams and then at home teaching following the Corona pandemic.

The method is able to find causal effects; the approach is an uncomplicated way of measuring how a specific policy change can lead to an effect. In this case, this would mean that we can investigate causal effects between males and females following the policies to first introduce anonymous exams with the start in January 2017 and then to transit to online teaching and digital examinations in March of 2020. Since the method has these characteristics, I found it the most suitable method to use to help me find the effects of the corona pandemic.

The hypothesis is that females will benefit more from being graded anonymous than males. This hypothesis is based on results from previous studies, like Tyrefors & Jansson (2020) that presented evidence of female students, also at the university level, benefiting from being graded anonymously. This is also consistent with theories of stereotyping.

To investigate this, I will use an interaction model and my estimating equation is:

Testscore$_i$ = α + γFemale$_i$ + δ$_1$Anonymous + δ$_2$After + β$_1$Anonymous*Female$_i$ + β$_2$After*Female$_i$ + ε$_i$

…where α is the intercept, Anonymous is a dummy that takes the value of 1 if the exam is taken within the period from the spring term of 2017 to March 20 2020, and the value of 0 if the exam is taken during another time period. After is also a dummy that takes the value of 1 if the exam is taken after March of 2020, and the value of 0 if it is taken before that time period, and Female is a dummy that takes the value of 1 if the exam taker is a female and the value of 0 if it is a male and ε is the error term. The regressions are also clustered towards the number of individuals in the dataset instead of number of total observations. This reduced the standard errors implying there is negative correlation within cluster, so we instead see a cancelation of variation when the residuals are summed over clusters.

In this model, α is interpreted as how students is treated in grading before 2017. δ$_1$ belonging to the coefficient "Anonymous" tells us how the students is treated during the time period from 2017-01-01 up to 2020-03-15 and δ$_2$ belonging to the coefficient "After" tells us how these students were treated during the time period after the 16th of March 2020. Therefore, the other coefficients are interpreted as deviations from this group. Since we have excluded males from the variables as we have a dummy for females, we could also interpret the coefficients as that the male students constitute our baseline as well as the coefficients "Anonymous" and "After", while the interactions estimates the effect between males and females, if it exists.

So, the coefficients for the interactions, the β's, are the coefficients that are of most importance. This since it takes both the specific time period and if the exam taker is female into account, so these coefficients measure the total effect of anonymous exams and the transition to online teaching. With existence of effects in grading between the genders, we would receive negative/positive β's in the interaction, meaning that females receive lower/higher grades after the exams started to be held anonymous or at home compared to the male students.

If no change in grades is present, the β's will be 0, therefore, we interpret the β's as the effect for females during that time period. Since we interpret the coefficients as deviations from α, this means that if the β's is negative, the individual receives a lower grade than the score should have generated compared to our "male" group, and vice versa if the β's was to be positive.

With the use of this type of data I acknowledge, as stated above, that the eventual effects found might not only be due to discrimination, but also that some groups and individuals manage studies on their own better than others.

One way of trying to get around the problem of the discrimination effect and might also be a better objective way to measure knowledge, is with the help of data on multiple-choice questions. This because with multiple-choice questions there is only one correct answer, without any further explanatory text, so therefore there is no room for discrimination in the grading. Also, the increased risk of cheating might also affect the eventual changes in the results found as well as another part that is important to consider the teacher's ability and knowledge in hosting online teaching.

Even though students at the university level are adults and able to research and study without being reliable on the teachers in the same way as students at lower levels in school are at the normal state, a decrease in the quality of lectures might also lead to a decrease in the results of the students.

I do though acknowledge the limitations to estimate the effects on a specific individual since I don't have data on the same exams graded both blindly and non-blindly. There might also be examples of gender discrimination that differ within and between the different groups which might overestimate the possible effects in certain groups.

4 Data

As stated above, I will use newly collected data on graded activities at Linnaeus University from the fall term of 2015 to the spring term of 2021 from the department of Economics and Statistics and its courses. During this period, there were two main grading systems in the Swedish University System regarding this department: G (pass), VG (pass with distinction) and U (fail); and the Bologna grading scale, which had to be implemented from the fall of 2008 at the latest. I will standardize the scores from these grading systems separately by subtracting the mean and divide it by the standard deviation, to make the different grading systems comparable.

Regarding the G-VG system I have given the grade VG a numeric score of 5, which is the highest, and G a score of 2,5. The grade U was rewarded with the score of 0. From the Bologna grading scale, which uses the grades from A-F, the grade A was given a numeric score of 5, grade B the score of 4, grade C the score of 3. Grade D got a numeric score of 2 and E a score of 1. Both the grade F and Fx got the score of 0.

I will test anonymous and non-anonymous grading using a regression model including interactions and measure the effect via differences between males and females. The data is collected from administrative system Ladok and contains information on the date of the exam and the gender of the exam taker. There is also other basic information regarding the course and type of examination along with the grade of the exam. The personal identity number is then also changed to a fictional identity number in order to keep the exam takers anonymous.

Table 1. Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Female	42,176	.4730178	.4992774	0	1
After	42,176	.2463012	.4308611	0	1
Anonymous	42,176	.5507398	.4974247	0	1
newid	42,176	2031.426	981.1129	1	3691
score	42,176	-2.99e-07	.9999882	-2.54534	3.190076

Summary statistics is provided in Table 1 on the previous page. In this table we can see that we have a portion of female students of 47.3 % and that as much as 55.1 % of the exams is taken during the period where anonymous exams were conducted at the Linnaeus University. The summary statistics also provide us with the information that 24.6 % of the observed exams is done after the decision to switch to online teaching and does therefore leaves us with 20.2 % of the exams taken up to the end of 2016. The max value of 3691 of the variable "newid" tells us how many unique individuals that makes up the total number of 42 176 observations.

We also have statistics on the standardized score. As stated above, I have given each grade a numeric score and then taken each of these observed scores, subtracted it with the mean and then divided it with the standard deviation to obtain these standardized scores. Due to the standardization we see a negative minimum score since if the observed numeric score is below the average, the standardized score will be negative, and if it is higher than the average, the standardized score will be positive. The standardization also gives us a mean very close to 0 (the mean of a standardized variable will rarely be exactly 0 due to rounding errors), and a standard deviation of 1, as is expected.

Table 2. Frequency

	Frequency	Percent	Cum.
Before 2017	8 560	20.30	20.30
Anonymous	23 228	55.07	75.37
After	10 388	24.63	100.00
Total	42 176	100.00	100.00

The table 2 that is showed above gives us a summary of the frequency in each group belonging to the different time period. With a total of 42 176 observations we have relatively high frequencies in each of the groups. We can see that we have most individuals in the "Anonymous group", belonging to the time period of 2017 up to 15th of March 2020. We have a rather similar frequency of students in the other groups with 8 560 observations before the reform with anonymous exams and 10 388 observations from the period following the pandemic. A specific table for each group can be seen in table A1-3 in the appendix.

5 Results

Table 3 shows us the results from the OLS regression where the standardized score is the dependent variable. The regression is also clustered towards the number of individuals in the sample and in this first regression I have excluded those types of examinations such as thesis works and oral presentations which is non-anonymous in normal cases. As stated above, if no changes in grades were present, all the β's would be 0, so we interpreted the β's as the effect of change for each group. If the β's was to be negative, the female individual receives a lower grade than the score should have generated compared to our "male" group.

Table 3. Results from regression with anonymous exams

VARIABLES	Anonymous exams
Female	0.0765
	(0.0532)
Anonymous	0.0750*
	(0.0384)
After	0.206***
	(0.0446)
anonymousXfemale	-0.0301
	(0.0558)
afterXfemale	0.0793
	(0.0631)
Constant	-0.141***
	(0.0375)
Observations	31,397
R-squared	0.009

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

If we begin by looking at the significance of the coefficients, we can see that only the constant and the coefficient "After" are significant at the 1%-level, and that "Anonymous" is significant at the 5%-level implying that these are the only coefficients we can say have a statistically significant effect in our model. The other coefficients are therefore insignificant, saying that we cannot interpret them being as not being zero, but not saying they are zero as well.

So in this model, when looking at the female students, the coefficients are insignificant so we cannot interpret them as that the female students generated higher grades than males before the reform of anonymous exams in 2017 and later that the female students actually received lower grades after this reform as well as receiving higher grades again after the switch to online teaching.

If we go over to the coefficients "Anonymous" and "After", we can see that the grades for these students became slightly higher when exams became graded anonymously and to an even higher extent when the University switched to online teaching in March of 2020. So, we can say that students generally benefited from being graded anonymously as well as benefited from the switch to online teaching and digital exams, but we cannot say that we have a statistically significant difference between male and female students.

All of the coefficients are also relatively small which indicates that its impact on the dependent variable is rather small.

This is also confirmed by our low R^2-value of 0.009 which shows the variation in the dependent variable that is explained by the explanatory variables, in our case, how the score of an exam is explained by the students gender and within which time period the exam was taken and graded. With an R^2-value of 0.009 only 0,9% of the variation in the score is explained by these variables.

Although, we have an F-value, collected from the appendix where we can find the original output from all regressions in Table A5-7, yielding F=22.84. The p-value associated with this F-value is very small (0.0000) so we can say that the independent variables reliably predict the dependent variable.

Table 4. Results from regression with non-anonymous exams

VARIABLES	Non-anonymous exams
Female	-0.0372
	(0.0673)
Anonymous	0.122**
	(0.0508)
After	0.0163
	(0.0664)
anonymousXfemale	-0.000350
	(0.0746)
afterXfemale	0.0765
	(0.0916)
Constant	-0.0290
	(0.0470)
Observations	10,779
R-squared	0.004

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In the above regression we instead see the results regarding other kind of graded activities that is not graded anonymous in normal cases such as oral examinations and assignments. This regression is done as a comparison to the one with only anonymously graded activities.

Here, "Anonymous" is the only coefficient that could be considered statistically significant, even if it's only at the 10%-level. This implying that students actually received higher grades during this time period on activities that where not graded anonymously. However, we cannot say that we have a difference between the genders and the results are more likely implying that the grades regarding these activities have been rather stable over time.

We also have an even lower R^2-value of 0.0038 compared to that of 0.009 in the previous regression and a significantly lower F-value of 3.05 and an associate p-value of 0.0095.

Table 5. Results from regression with multiple-choice questions

VARIABLES	Multiple-choice questions
Female	-0.0275
	(0.0684)
Anonymous	0.227***
	(0.0512)
After	0.155**
	(0.0618)
anonymousXfemale	0.0363
	(0.0736)
afterXfemale	0.167*
	(0.0873)
Constant	0.0749
	(0.0485)
Observations	10,895
R-squared	0.010

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

In this regression we see the results regarding the graded activities called "Duggor" which consists of multiple-choice questions. This is done because with multiple-choice questions there is only one correct answer, without any further explanatory text, so therefore there is no room for discrimination in the grading. Here, the coefficient belonging to "Anonymous" are the only one significant at the 1%-level, and that of "After" is significant at the 5%-level while "afterXfemale" is significant at the 10%-level.

If we begin with our group of male students here as well, represented by the constant (α) and the coefficients "Anonymous" and "After", we can see that the grades for these students became higher when exams became graded anonymously. Male students also gained when the university switched to online teaching in March of 2020.

Regarding the female students we can say that we have an effect with the at home exams due to the Corona pandemic, to an even higher extent compared to males. We cannot, however, with statistical significance say that they received higher grades once the grading became anonymous nor that the female students generated lower grades than males before the reform that started in 2017. We have a low R^2-value here as well, of 0.0101, although higher than in the other regressions, and an F-value 10.35. In this regression as well, we have a low p-value associated to the F-value.

6 Conclusions

A difference-in-difference method has been used to try to find eventual effects on students' grade due to the online teaching as a result of the Corona pandemic at the university level in Sweden and to see if there were any differences between male and female students. The coefficients of interest in the regression models were on beforehand the β's (the interactions) that showed the differences between the genders a well as the coefficient "After" that showed how students generally performed since the decision to have digital, online exams. The results from the previous section showed us that we cannot say that there is a difference in grading between the genders. However, looking at the coefficients regarding the grades after the decision to switch to online teaching and digital exams in March of 2020, the grades went up in all the regression models, implying that students benefited from this switch.

Even though no findings of differences in grading between genders, one way to control for discrimination, for example, was to make a regression on the activities that only included multiple choice questions. This because since then, there is only one correct answer, without any further explanatory text, so therefore there is no room for discrimination in the grading. Looking at that regression we saw that the grades for both male and female students became higher when exams became graded anonymously and then even higher when the University switched to online teaching in March of 2020, with female students benefiting more from this compared to males. So, regarding the multiple-choice questions, both female and male gained from the decision to conduct online teaching and digital exams.

Looking at the results from table 3 where the regression includes activities that is supposed to be graded anonymously in a normal state, such as hall exams, we saw that both male and female received higher grades after the introduction of anonymous grading as well as the switch to digital examination, but we cannot say that there were an statistically significant difference between genders. With anonymously graded hall exams, handwriting is the only way to bias the grade, so this is interesting since there are studies, like Thomas & Vaught (2020) and Burr (2002) showing that a person with better handwriting often benefits from this and that females generally have better handwriting compared to males, so gender discrimination could have lived on even after the reform to introduce anonymous exams.

Since the decision to conduct the exams from distance, all standard examinations within this dataset instead became open book exams where the students were allowed to receive help from course literature during the exam. This, in addition to teachers possibly not adjusting the exams to be conducted at home to the correct extent, during this presumed temporarily period of time, allow for a student to perform a higher grade even though the actual knowledge of the student might not be at the same standard as shown in the exam. As stated, we saw that students benefited from the online examinations. In addition to the open book exams, the possibility to cheat has increased with these kind of examinations where students easily can help each other out and as stated above, studies have shown that many students take the chance when given and that males actually are more prone to cheat than females, so this might also be an explanation for the even higher grades after this switch.

If we compare these results to the ones found in the study in the Netherlands by Engzell, Frey & Verhagen (2021), they found a learning loss of about 3 percentile points or 0.08 standard deviations and that students made little or no progress while learning from home and suggest losses even larger in countries with weaker infrastructure or longer school closures. This study is done on students aged 8-11 and, in this age, students are more dependent on the teachers and may not have the same requirements to complete studies on their own but requires more support from others. So, this should affect the grades to a greater extent than students at the university level who are adults and generally should be able to take greater responsibility.

There might also be such of an easy explanation that during this time, due to the pandemic, students hasn't been able to live the typical student life with social meetings and with restrictions regarding bars and restaurants, maybe they just have studied more since there are not as much else to do?

As seen in the result section, we stated that the results overall became higher after the switch to online teaching and digital examinations. This outcome was also some sort of predicted when one considers all those different explanations for differences in grades that one could see, although in the presence of eventual discrimination we expected to see lower results with the re-introduction of non-anonymously graded exams. We couldn't however see any statistically significant difference between males and females' grades.

Even though there are many benefits from the technological evolution and the possibilities of effectiveness with online teaching, we see that the institutions have a lot to learn before they

can take advantage of these possibilities. This might be an area for possible future research, to see how the universities can embrace the positive parts from online teaching and how professors could improve their pedagogical skills.

With the regressions used and results found I acknowledge that there are some limitations in this paper, which might suggest further research within these areas. In this paper I have not been able to estimate the effects on a specific individual since I don't have data on the same exams graded both blindly and non-blindly. These results might however strengthen the reason to believe that there should be anonymous grading to the highest extent possible and that hall examinations for standard exams should be re-introduced when possible to prevent further cheating and get a more overall fairness of the students' grades.

One could also extend this study in the future by looking at differences between the multiple-choice examinations and other kind of examinations to see the grading differences in anonymous grading as well, with the previous studies regarding handwriting skills in the back of the mind.

The last possible areas for future research might also be, since we now soon can see the end of this pandemic, to look into differences in grading ones the school now goes back to its normal state. Then one could state if these increasing results found in this study was more likely due to cheating and the fact that it was open book exams or if the students actually did study more and increased their actual knowledge.

It might also be interesting to see how different levels in the schooling system was affected. With the knowledge from the study in the Netherlands it might be interesting to see how the elementary schools in Sweden was affected by the pandemic since they had to take on certain restrictions but was not closed down as in the Netherlands. Which students in the schooling system has been affected the most?

This leads up to the students in the Swedish High Schools who has been on and off regarding teaching online and at distance versus actually being at school. How has these students been affected in grades? These students are supposed to go on to the Universities or into the labour market after their graduation, has this affected their possibilities and opportunities?

So, as the pandemic moves towards its end and the world prepares to go back to some kind of normality, we still have many unsearched and unanswered questions of how this past year has affected us.

References

Ahmed, A., Andersson, L. and Hammarstedt, M. (2010) Can discrimination in the housing market be reduced by increasing the information about the applicants? *Land Economics, 86:1, 79-90*

Ahmed, A. and Hammarstedt, M. (2020) Ethnic discrimination in contacts with public authorities: a correspondence test among Swedish municipalities, *Applied Economics Letters, 27:17, 1391-1394*

Alteri, Ashley M. (2016) Gender-Based Discrimination in the Workplace. *Review of Public Personnel Administration. Sep2016, Vol. 36 Issue 3, 264-282.*

Bengtsson, R., Tyrefors, B., and Iverman, E. (2012) Gender and ethnic discrimination in the rental housing market. *Applied Economics Letters, 2012:19, 1–5*

Bilan, Y., et. al (2020) Gender discrimination and its links with compensations and benefits practices in enterprises. *Entrepreneurial Business and Economic Review 2020, 8:3, 190-203*

Bogoslov, I., and Lungu, A. (2020) Facing the new learning normality: Europe at a glance in the context of coronavirus pandemic. *Revista Economică 72:1 (2020). 25-36*

Boyer-Davis, S. (2020) Technostress in Higher Education: An examination of faculty perceptions before and during the Covid-19 pandemic. *Journal of Business and Accounting, Vol 13:1, 42-58*

Buchmann, C., DiPrete, T., and McDaniel, A. (2008) Gender Inequalities in Education. *Annual Review of Sociology 2008:34, 319-337.*

Burr, V. (2002) Judging Gender from Samples of Adult Handwriting: Accuracy and Use of Cues. *Journal of Social Psychology. Dec 2002, 142:6, p691-700.*

Carlsson, M., and Rooth, D. (2007) Evidence of Ethnic Discrimination in the Swedish Labor Market Using Experimental Data. *ZA Discussion Paper No. 2281*

Case, C., King, D., Case, J. (2019) E-Cheating and undergraduate Business students: Trends and the role of gender. *Journal of Business & Behavioral Sciences. Spring 2019, 31:1, p102-113.*

Chaudhary, M. (2020) Enhancing University's Brand Performance during the COVID-19 Outbreak: The Role of ICT Orientation, Perceived Service Quality, Trust, and Student's Satisfaction. *Pakistan Journal of Commerce & Social Sciences, 2020, 14:3, 629-651.*

Dunn-Jensen, L. et.al (2016) Revealing Gender Bias: An Experiential Exercise. *Organization Management Journal, 13:2, 101-114*

Engzell, P., Frey, A., and Verhagen, M. (2021) Learning loss due to school closures during the COVID-19 pandemic. *PNAS 2021, 118:17*

Feld, J., Salamanca, N., and Hamermesh, D. (2015) Endophilia or Exophobia: Beyond Discrimination. *The Economic Journal, 126:8, 1503–1527*

Habragoda, B. (2020) Challenges of Implementing Online Teaching in Universities during Covid-19 Global Pandemic: A Developing Country's Perspective. *International Journal of Management, Accounting and Economic, 7:11, 675-682*

Huskanovic, J., and Nur, B. (2020) Discrimination against foreign named students - a study based on a Swedish University

Jugănaru, M. (2020) From Traditional Education to the Online Training Process in Romanian Higher Education – an Experiment Generated by the New Coronavirus Pandemic. *"Ovidius" University Annals, Economic Sciences Series, Issue 1/2020, 676-681*

Jurges, H. and Schneider, K. (2011) Why Young Boys Stumble: Early Tracking, Age and Gender Bias in the German School System. *German Economic Review 12(4): 371–394*

Keng, S. (2020) Gender bias and statistical discrimination against female instructors in student evaluations of teaching. *Labour Economics. Oct2020, Vol. 66.*

Lavy, V. (2008) Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of Public Economics 92 (2008) 2083-2105*

Lavy, V. and Angrist, J. (2009) The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial. *American Economic Review 2009, 99:4, 1384–1414*

Lento, C., Sayed, N., Bujaki, M. (2018) Sex role socialization and perceptions of student academic dishonesty by male and female accounting faculty. *Accounting Education. Feb 2018, 27:1, p1-26.*

Lumley, K. (2015) Without fear or favour, affection or ill will: addressing gender bias in NSW judicial education. *The Judicial Review, 2014:12, 63-78*

McKendall, M., Klein, H., Levenburg, N., de la Rosa, D. (2020) College Student Cheating and Perceived Instructor Fairness. *Journal of the Academy of Business Education. Fall 2010, Vol. 11, p14-32.*

McMullen, A., Bursuc, V., Doval, C., Grant, S., Grossberg, J., Jones, I., Mae, V., Brooks, W., and Willey, S. (2020) Strategies for Effective Online Teaching in Higher Education. *Business Education Innovation Journal, vol. 12:2, 40-48*

Rigo, F. (2020) Distance Learning with Edmodo During the Coronavirus School Closures. *Marketing Identity 2020, 537-546.*

Root, W., and Rehfeldt, R. (2021) Towards a Modern-Day Teaching Machine: The Synthesis of Programmed Instruction and Online Education. *Psychological Record, Mar2021, 71:1, 85-94.*

Sales, D., Cuevas-Cerveró, A., and Gómez-Hernandéz, J. (2020) Perspectives on the information and digital competence of Social Sciences students and faculty before and during lockdown due to Covid-19. *Profesional de la información, v. 29, n. 4, e290423.*

Simha, A. (2014) Cheating in College - Why Students Do It and What Educators Can Do About It. *Academy of Management Learning & Education. Mar 2014, 13:1, p143-145.*

Thomas, S., Vaught, S. (2020) The Write Stuff: What the Evidence Says About Using Handwriting Analysis in Hiring. *SAM Advanced Management Journal (07497075). Fall 2002, 66:4, p31-35.*

Todrican, M. (2020) Sustainable Education System in the Face of the COVID-19 Pandemic. *Ovidius University Annals, Series Economic Sciences 2020, 20:2, 776-782*

Triana, M., et. al (2019) Perceived Workplace Gender Discrimination and Employee Consequences: A Meta-Analysis and Complementary Studies Considering Country Context. *Journal of Management, 45:6, July 2019, 2419–2447*

Tyrefors et.al. (2019) Gender grading bias in junior high school mathematics. *Applied Economics Letters, 27:11, 915-919*

Tyrefors, B., Höglin, E. and Johannesson, M. (2011) Are boys discriminated in Swedish high schools? *Economics of Education Review 30 (2011) 682-690*

Tyrefors, B., Höglin, E. and Johannesson, M. (2015) Discrimination against students with foreign backgrounds: Evidence from grading in Swedish public high schools. *Education Economics, 23:6, 660-676*

Tyrefors, B., and Jansson, J. (2019) Gender grading bias at the university level: quasi-experimental evidence from an anonymous grading reform

Tyrefors, B., and Jansson, J. (2020) The Genius is a Male. Stereotypes and Same-Sex Bias in Exam Grading in Economics at Stockholm University

Van Ewijk, R (2011) Same work, lower grade. *Economics of Education Review 30 (2011) 1045– 1058*

Vlachos, J., Hertegård, E., & Svaleryd, H. B. (2021) The effects of school closures on SARS-CoV-2 among parents and teachers. *Proceedings of the National Academy of Sciences, 118(9).*

Åberg & Boström (2020) Ethnic Discrimination, A study on Swedish Municipalities

Appendix

Table A1. Summary over variable "Anonymous"

```
Exams taken
from 170101
   to 200315        Freq.       Percent        Cum.
─────────────┼──────────────────────────────────────
          0       18,948         44.93        44.93
          1       23,228         55.07       100.00
─────────────┼──────────────────────────────────────
      Total       42,176        100.00
```

Table A1 shows us the frequency regarding the dummy variable
"Anonymous" where the observations valued 1 show us how many
observations the dataset contains within this time period. In this case, 23 228
exams where taken within the time period 170101-200315.

Table A2. Summary over variable "After"

```
Exams taken
      after
     200316        Freq.       Percent        Cum.
─────────────┼──────────────────────────────────────
          0       31,788         75.37        75.37
          1       10,388         24.63       100.00
─────────────┼──────────────────────────────────────
      Total       42,176        100.00
```

Table A2 shows us the frequency regarding the dummy variable "After"
where the observations valued 1 show us how many observations the dataset
contains within this time period. In this case, 10 388 exams where taken after
200316.

Table A3. Summary over variable "Female"

Dummy for female exam takers	Freq.	Percent	Cum.
0	22,226	52.70	52.70
1	19,950	47.30	100.00
Total	42,176	100.00	

Table A3 shows us the frequency regarding the dummy variable "Female" where the observations valued 1 show us how many observations the dataset contains where the student is of female gender. In this case, 19 950 exams takers where female.

Table A4. Results from regression with all exams

```
Linear regression                          Number of obs   =    42,176
                                           F(5, 3690)      =     15.31
                                           Prob > F        =    0.0000
                                           R-squared       =    0.0056
                                           Root MSE        =    .99726
```

(Std. Err. adjusted for 3,691 clusters in newid)

score	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
Female	.0510871	.0448127	1.14	0.254	-.0367729	.1389471
Anonymous	.0927551	.0320754	2.89	0.004	.0298679	.1556423
After	.1604166	.0389212	4.12	0.000	.0841074	.2367258
anonymousXfemale	-.0281942	.0470489	-0.60	0.549	-.1204387	.0640502
afterXfemale	.07637	.0549854	1.39	0.165	-.0314348	.1841749
_cons	-.1162672	.0316283	-3.68	0.000	-.178278	-.0542565

In the regression in Table 5 we have all observations included, activities supposed to be graded anonymous as well as activities not supposed. We have a low R^2-value of 0.0056 and a F-value yielding F=15.31. The p-value associated with this F-value is 0.0000 so according to this we can say that the independænt variables reliably predict the dependent variable.

Table A5. Original output from regression with anonymous exams

```
Linear regression                              Number of obs    =      31,397
                                               F(5, 3570)       =       22.84
                                               Prob > F         =      0.0000
                                               R-squared        =      0.0090
                                               Root MSE         =        1.04

                              (Std. Err. adjusted for 3,571 clusters in newid)
```

score	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
Female	.0765169	.0531793	1.44	0.150	-.0277479	.1807817
Anonymous	.0750427	.0383806	1.96	0.051	-.0002074	.1502928
After	.206279	.0445563	4.63	0.000	.1189206	.2936374
anonymousXfemale	-.0300651	.055799	-0.54	0.590	-.1394662	.0793359
afterXfemale	.0793489	.06314	1.26	0.209	-.0444452	.2031429
_cons	-.1413962	.0375492	-3.77	0.000	-.2150163	-.0677761

Table A6. Original output from regression with non-anonymous exams

```
Linear regression                              Number of obs    =      10,779
                                               F(5, 2889)       =        3.05
                                               Prob > F         =      0.0095
                                               R-squared        =      0.0038
                                               Root MSE         =       .8544

                              (Std. Err. adjusted for 2,890 clusters in newid)
```

score	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
Female	-.0371848	.0673382	-0.55	0.581	-.1692207	.0948511
Anonymous	.1217751	.0507615	2.40	0.017	.0222426	.2213076
After	.0162759	.066395	0.25	0.806	-.1139103	.1464622
anonymousXfemale	-.0003503	.0746171	-0.00	0.996	-.1466585	.1459578
afterXfemale	.0764988	.0915633	0.84	0.404	-.1030372	.2560347
_cons	-.0289868	.0470079	-0.62	0.538	-.1211592	.0631856

Table A7. Original output from regression with multiple choice questions

```
Linear regression                          Number of obs   =      10,895
                                           F(5, 1906)      =       10.35
                                           Prob > F        =      0.0000
                                           R-squared       =      0.0101
                                           Root MSE        =      1.0869

                                 (Std. Err. adjusted for 1,907 clusters in newid)
```

score	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
Female	-.0274656	.0683583	-0.40	0.688	-.1615305	.1065993
Anonymous	.2266781	.0512057	4.43	0.000	.126253	.3271032
After	.1549333	.0617827	2.51	0.012	.0337645	.2761021
anonymousXfemale	.0363138	.0736413	0.49	0.622	-.1081122	.1807399
afterXfemale	.1668256	.0873231	1.91	0.056	-.0044333	.3380845
_cons	.0748998	.0485273	1.54	0.123	-.0202725	.170072