



Universitetet
i Stavanger

UIS BUSINESS SCHOOL

MASTER'S THESIS

STUDY PROGRAM:

Business Administration

THESIS IS WRITTEN IN THE FOLLOWING
SPECIALIZATION/SUBJECT:

Economic Analysis

IS THE ASSIGNMENT CONFIDENTIAL?
(**NB!** Use the red form for confidential theses)

TITLE:

The effect from Reform 97 on the high school graduation gender gap

AUTHOR(S)

SUPERVISOR:

Ingeborg Foldøy Solli

Candidate number:

4016

.....

4047

.....

Name:

Hanne Ulset

.....

Karina Vigdel

.....



Universitetet
i Stavanger

Handelshøgskolen ved
Universitetet i Stavanger
Stavanger, spring 2019



The effect from Reform 97 on the high school graduation gender gap

Hanne Ulset and Karina Vigdel
Supervisor: Ingeborg Foldøy Solli

Master Thesis
Business Administration – Economic Analysis

Preface

The basis for this research stems from our own passion for higher education and illuminates the fact that some students drop out of school before graduating from high school.

Educational policies undoubtedly change the Norwegian school system, and this paper has been written in cooperation with the Business School at The University of Stavanger to investigate these effects. It has been an interesting, educational and challenging period. We would like to thank our supervisor Ingeborg Foldøy Solli for all the help and feedback during this period, and also send a thank you to Trond Pedersen for all help regarding Microdata.

Abstract

In 1997, Reform 97 was introduced in the Norwegian school system. The reform made it mandatory for children to start school at the age of six instead of seven, resulting in expanding the primary school with one year, now lasting ten years instead of nine. Using Norwegian register data collected from Microdata, we investigate if the reform has had any effect on the high school graduation gender gap, focusing on whether males have been negatively affected compared to females. Using a difference-in-difference and binary logit regression, we investigate the effect from the reform on the gender gap. We find an increase in the gender gap after the reform was implemented, but after including several control variables to the regression analysis, our results show no significant effect from Reform 97. We conclude that there must be other factors explaining the increase in the gender gap.

Table of Contents

Tables.....	1
Tables in the paper	1
Appendix tables.....	2
Figures.....	3
1. Introduction.....	4
2. Existing Literature.....	6
2.1 Gender differences and gender gap	6
2.2 Reform 97	8
3. Institutional Background.....	9
4. Labour theory and hypothesis	11
4.1 Gender gap-the difference between male and female	12
4.2 Gender gap and Reform 97.....	13
4.3 The rate of time preference.....	13
5. Empirical strategy.....	16
5.1 Difference-in-difference	16
5.2 Binary logistic regression	19
6. Data and Sample Description.....	21
6.1 Microdata	21
6.2 Variables.....	22
7. Empirical Results.....	23
7.1 Descriptive results.....	23
7.2 Main results.....	30
7.2.1 Short binary regression model	30
7.2.2 Full binary regression model.....	32
7.2.3 Subsamples.....	37
8. Conclusion.....	41
References	42

Tables

Tables in the paper

Table 1: Share of students graduating high school, shown as a fraction of birth cohort each year. Robust standard errors in parentheses.....	23
Table 2: Share of female and male students graduating high school, shown as a fraction of birth cohort each year. Robust standard errors in parentheses.....	25
Table 3: Mean values of the different variables in our data set, divided into pre- and post-segments. Variables explained in chapter 3. Robust standard errors in parentheses.	26
Table 4: Mean values of the different variables in our data set, divided into immigration status; Norwegian, Western and non-Western immigrants, as well as before and after (pre and post) the reform was implemented. Robust standard errors in parentheses.....	28
Table 5: Binary logit regression output. Model 1 regresses probability of high school graduation on post, male and male*post. Model 2 includes parents' education, Model 3 siblings and Model 4 immigration status. All variables from Model 1-4 is included in Model 5. *, ** and *** denote significance at 10 percent, 5 percent and 1 percent level respectively. Robust standard errors in parentheses.....	30
Table 6: Full binary logit regression model regressing graduation rate on year of birth and gender, extended into 5 models with variables on parents' education, number of siblings and immigration status. *, ** and *** denote significance at 10 percent, 5 percent and 1 percent level respectively. Robust standard errors in parentheses.	33
Table 7: Short binary logit regression model showing the probability of graduating high school dependent on gender and whether they were exposed to the treatment or not (post). Subsamples showing regression output for Norwegian, Western and non-Western students. *, ** and *** denote significance at 10 percent, 5 percent and 1 percent level respectively. Robust standard errors in parentheses.....	37
Table 8: Subsample analysis. Binary logit regression analysis divided into parents' education level, immigration status and the number of siblings. *, ** and *** denote significance at 10 percent, 5 percent and 1 percent level respectively. Robust standard errors in parentheses. ..	39

Appendix tables

Appendix 1: Microdata	i
Appendix 2: PISA.....	iii
Appendix 3: Descriptive information	iv
Appendix 4: Binary logit regression output.....	xi
Appendix 5: Short binary logit output, subsample on immigration status	xiv
Appendix 6: Full binary logit regression output.....	xv
Appendix 7: Full binary logit regression output, subsamples.....	xx
Appendix 8: Info on excluded variables	xxxiii

Figures

Figure 1: Men and Women 16 years and older with higher education, 2017 12

Figure 2: Difference-in-difference estimate shown by arrows between two graphs across time.
..... 17

Figure 3: Students' high school graduation rate in Norway, years 1987-1995. Numbers from
Table 1. 23

Figure 4: Graduation rate for students in Norway in the years 1987-1995, divided into gender.
Numbers from Table 2. 25

Figure 5: Difference in gender in graduation rate for students in Norway in the years 1987-
1995. Cacluated as difference between female and male in Table 2. 25

Figure 6: Mean values from the data sample, as well as the values before and after Reform 97.
Numbers from Table 3. 26

Figure 7: Showing the shared of different variables in our data set, divided into immigration
segments and before and after Reform 97 was implemented. 28

1. Introduction

A gender gap is the difference in how males and females perform and the choices they make that reflect on what they achieve. There has always been a gender gap and we see a gap in educational attainment¹. According to SSB², 66 166 students in the period from 2012-2017 graduated from high school. Out of these, there were 18 774 females graduating from general studies compared to 15 621 males. Looking at students graduating from vocational study program, the number of males exceeds number of females, 18 611 compared to 13 169. Looking at the same numbers at the earlier period, from 2010 to 2015, we find the same trend. More females than males are graduating from general studies, 17 742 compared to 14 734, and more males are graduating from vocational study program, 18 041 compared to 13 320³.

In 1997, a new reform was introduced in the Norwegian school system which led to some changes as of the school year 1997-98. Reform 97 expanded the primary school from being nine years of schooling to be ten years of schooling⁴. This extra year of schooling required children to start school at an age of six instead of seven. One of the main goals of the reform was to make equal learning possibilities for all students, independent on social and economic background. In the aftermath some critics have stated that the reform developed a theoretical-based school, ending up being a better fit for females than for males⁵. This criticism is interesting, as it suggests that Reform 97 may have resulted in leaving one of the genders worse off, in contrast to the goal. This master thesis will look closer into the effects of Reform 97. Are students better prepared for high school and higher education after an extra year of schooling, or does this theoretical school give the students distaste for further education, especially among males? We want to find out if Reform 97 contributed to expanding the high school graduation gender gap. The research problem for this master thesis is:

“Has Reform 97 resulted in expanding the gap between males’ and females’ high school graduation rate?”

¹ (World Economic Forum, 2018)

² (Statistics Norway, 2018)

³ (Statistics Norway, 2016)

⁴ (Kirke-, utdannings- og forskningsdepartementet, 1996)

⁵ (Haug, 2003)

Register data from the whole Norwegian population will try to answer this question. The data are collected from Microdata.no, a collaboration between NSD (Norwegian Centre for Research Data) and SSB (Statistics Norway). This paper use students born between 1987 and 1995.

This paper consists of 8 chapters. Chapter 2 presents existing literature on gender differences and Reform 97, while chapter 3 display the institutional background, looking more closely into the reform and the changes it contained. Further, chapter 4 will present the labour theory and our research question. Our empirical strategy is presented in chapter 5 and in chapter 6 we will present Microdata and our data set, as well as the variables that were used in our analysis. In chapter 7 the empirical results will be presented and discussed, before a conclusion is made in chapter 8.

2. Existing Literature

The fundamental differences between males and females have been studied for many years and have been analysed using different methods and perspectives. Most of the research are from abroad, and therefore does not consider Reform 97, but we assume that it follows the same principles and theories.

2.1 Gender differences and gender gap

Risk aversity is one of the features often separating males and females, and studies have found that women tend to be more risk averse than men, meaning that men are more willing to take risky decisions than women are⁶. When looking at risk taking, most researchers refer to goals, values, options, and outcomes. Goals and values, among other things, determine the kinds of outputs pursued by individuals. This could be a choice between good grades in school versus being popular with friends. These goals and values will also determine the kinds of options that individuals are considering, for example studying versus socializing⁷. Looking closer into educational environment, Pontiell reports the fear of failure as the most significant negative emotion as this leads to more cautious behaviour patterns. In addition to the risk with education and failure, an educated person is not guaranteed to get a relevant job after graduating. This might have an impact on whether they attend school or not. As females are found to be more risk averse than males, according to this research, this should indicate that male students are more likely to graduate from high school than females.

Alan Feingold examines the gender differences in personality traits through meta-analysis⁸. Combining several other studies in the literature of gender differences in personality (1958-1992), he finds attributes differ between males and females. For example, males were found to be more assertive and had slightly higher self-esteem than females, while females scored higher on extraversion, anxiety, trust and tender-mindedness. The differences were generally constant across ages, years of data collection, nations and educational levels. This is relevant for our research, as it shows a fundamental difference between males and females.

To test the difference in learning between male and female students, there has also been conducted experiments. At a British Columbia, private, all-female high school, Angela Josette

⁶ (Bymes, Miller, & William , 1999)

⁷ (Pontiell, 2003)

⁸ (Feingold, 1994)

Magon conducted an experiment on her students⁹. This experiment was executed by dividing a science class into two and given lessons that were designed to target either males or females. In the female-designed classes, the teacher used a soft and gentle voice and allowed females to voluntarily answer without pressure. The classroom had a slightly warm temperature, decorated with plants and warm colours. In the male-designed classes, the teacher had a louder voice with brief and concise instructions. The classroom had a lower temperature than the female-designed classroom and the room was without distractions. The teacher also focused on actively learning by games and used learning materials like maps and diagrams. The results indicated that engagement and enjoyment of lessons do not always correlate better learning. With the all-female setting, the literature strategy aimed at teaching females, produced higher achievement for the female students than the strategies targeted to teaching males. Having a one-size fits all type of strategy for teaching children at school may therefore not be a very successful approach. If these findings are true for both males and females, there is reason to believe that Reform 97 may have affected the two genders differently.

More researchers have discovered the differences between males and females. Michael Gurian and Kathy Stevens has written a book where they find that males and females are learning differently due to biological reasons; their brains are built differently¹⁰. They find that generally, a male's brain develops later than a female's brain, making females more receptive for early schooling than males. The research also find that males prefer mechanical and structural thinking. In addition, males tend to be less efficient multitaskers and better at learning through task and project focus, not only by reading and writing.

To summarize, existing literature shows that males appreciate an active learning method while females seem to manage a more theoretic school day. If this existing literature is right, it is interesting to investigate how the Norwegian populations has responded to Reform 97.

⁹ (Magon, 2009)

¹⁰ (Gurian & Stevens, 2010)

2.2 Reform 97

The evaluation of Reform 97 that was conducted by Peder Haug in cooperation with the Norwegian research Council, is one of the most comprehensive report that has been made on the effect of the reform¹¹. They find that the school, after the reform, has a high activity level, but that the learning is lower. The evaluation questions the change specially for certain groups of students where the effect has been negative. The goal of having equal learning possibilities for all students independent on gender, parental economy, residency, capabilities, preconditions and cultural and linguistic background does not seem to have been met, according to the evaluation. The most distinct effect is that females systematically did better than males on almost all subjects. They also find that many students with a different native language from Norwegian, do worse in school. The same effect is found for students from families with low education compared to those with parents with higher education.

Studies on intensive programs promoting early learning have shown positive effects on children from disadvantaged families' development. However, we know less about how larger universal learning programs can influence children's development. Drange, Havnes and Sandsør studies the effect of Reform 97 on children's grades when graduating from high school within standard time and if the reform affected the students' probability of graduating from high school within standard time, or if more students chose academical specialization because of the reform¹². Before the reform, many families decided to send their kids to preschool and therefore didn't experience the big difference before and after the reform. Because of this, Drange, Havnes and Sandsør have estimated the effect of the reform on the little group that did not have a preschool available before 1997. Even if many of these children came from families with low income and education, their results show that the reform barely had any effect on the students' development.

The evaluation of the reform claims that the school system is more theoretical and that the main goals of the reform has not been met. Based on the existing researches on Reform 97, we expect to find that males have a harder time trying to adapt to the Norwegian school system. Has Reform 97 resulted in a school system better fitting for females than males?

¹¹ (Haug, 2003)

¹² (Drange, Havnes, & Sandsør, 2012)

3. Institutional Background

The school level in Norway consists of three institutions, primary school, secondary school (together called elementary school) and high school. The first level of the Norwegian school system has a duration of 7 years. Children start in August the year they turn 6 years old and leave at an age of 12/13. The next level, the secondary school, lasts for three years. At this stage students are for the first time graded in all subjects. The students are now ready to attend high school. In this paper, high school are the same as “Videregående” in the Norwegian school system. High school is then the school students attend after graduating from 10 years of primary school and before they possibly attend higher education at universities.

Until 1997, children were attending primary school in August the year they turned 7 years. The year before attending school, some children got the opportunity to attend a kindergarten pre-school program. The pre-school program was supposed to prepare the children for enrolment in the primary school. The children were introduced to simple learning through playtime. The pre-school program was not obligatory so not every 6-years old in Norway had the opportunity to participate. In 1992-93, 80% of all 6-year olds had a voluntary pre-school offer¹³.

The fact that not all children had the opportunity to attend pre-school, started to worry the Norwegian Government. They were worry about children enter primary school on different footings. Children from families of low education and low income did not often attend pre-school programs and therefore had a different starting point when attending school, compared to those who attended. A new school reform was proposed in 1993 and passed the Norwegian Parliament in 1994. The reform was implemented in August 1997 and got the name Reform 97¹³. Reform 97 made it obligatory for all 6-year olds to attend school in a similar pre-school program. All 6-year olds where now getting equal educational offer, instead of an voluntarily pre-school program, regardless of their residence and their household economy.

The new mandatory kindergarten program was aimed to give the children the best combination of kindergarten traditions and school. The 6-years old were structured in groups of 20 with two kindergarten teachers and were exposed to learning through play four hours a day. The new school curriculum that followed Reform 97, strengthened the focus for

¹³ (Kirke-, utdannings- og forskningsdepartementet, 1993)

developing social, language and physical skills through free play and “learning-by-playing”¹⁴. The elementary school was, together with the children’s parents, supposed to take responsibility for the children’s teaching and childhood environment. One of the main goals is to prepare children to manage future challenges and to motivate children for further learning and understanding¹⁵.

In year 2000, Norwegian 15-year olds were tested, for the first time, in an international school test called PISA (Program for International Student Assessment). The test is completed through a cooperation between the 36 countries that make up the OECD¹⁶; an organisation working together for economic cooperation and development. The students are tested in mathematics, science and reading. In 2006, Norwegian students’ results were disconcerting, scoring below the mean in every tested subject¹⁷. Something had to be done with the Norwegian school-system and a reform called “Kunnskapsløftet” was introduced in 2006. This reform implied changes in the curriculum in all classes, from the first class in elementary school to the last year of high school. The main goal was improving learning outcomes for all students¹⁸.

Reform 97 is somehow replaced by Kunnskapsløftet 2006 and its new goals. The desire to improve the Norwegian school results may have made the school more theoretical, possibly a better fit for females than for males¹⁹. There may be reason to believe that the reform from 2006 could have an effect on our results.

¹⁴ (Drange, Havnes, & Sandsør, 2012)

¹⁵ (Kirke-, utdnings- og forskningsdepartementet, 1993)

¹⁶ <https://www.oecd.org/about/members-and-partners/>

¹⁷ (Utdanningsdirektoratet, 2007)

¹⁸ (Regjeringen.no, 2019)

¹⁹ (Haug, 2003)

4. Labour theory and hypothesis

A form of gender gap has always existed, whether in terms of pay, political representation or educational attainment. Until the middle of the 1960's, there was a broad consensus in Norway that society should be built on the nuclear family. People got married in an early age; the man had paid work while the woman worked at home. The so-called housewife era took a turn during the 1960's. The small feminist opposition grew bigger and more females attended higher education. A more positive view of gender equality emerged²⁰.

Until the end of the 1950's, males and females were divided into different classes. For the females, the school was a place to learn how to be a good housewife. Housekeeping was a compulsory subject for females which resulted in fewer hours in theory subjects compared to the males. People started to protest against this and in the end of the 1950's, a common school law was introduced in Norway²⁰. Nine-years of mandatory school for all children were introduced in 1969. Males and females were now attending the same classes with the same rights and duties. Slowly but surely it became more and more common for women to attend higher education. In recent times, education among males and females has taken a turnaround. From almost non educated females in the 1950's to more females than males with higher education in 2008-2017²⁰.

Figure 1 shows the share of males and females graduating from high school. The green bar shows the proportion of males with higher education (percent), while the purple bar shows the same for females, both with percentage rate on left side. The yellow line shows the relationship between men and women with higher education (scores). This score is calculated as a share of females divided by share of males, giving a percentage relationship between the two. The recent educational history in Norway shows that fewer males than females attend higher education. In the past 70 years the Norwegian school has changed from being a place more suited for the males to become a place where males seem to enjoy less than females. We will investigate to what extent Reform 97 has made school harder for males than for the females in the Norwegian school system.

²⁰ (Lønnå, Kvinners rettigheter i Norge fra 1945 til 1990-årene, 2017)

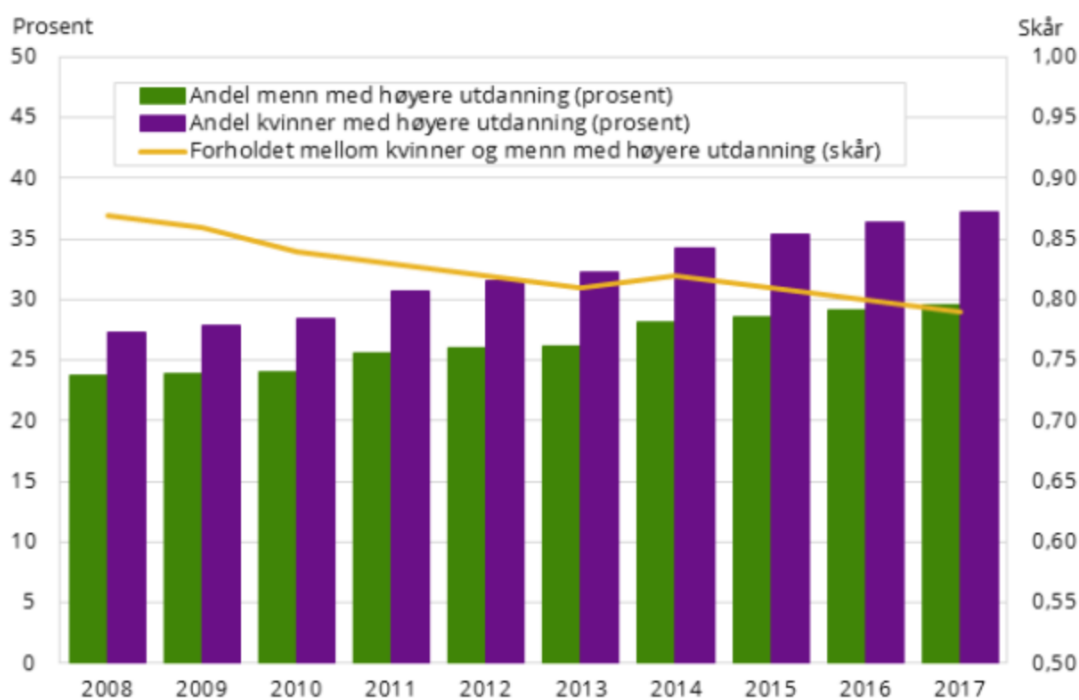


Figure 1: Men and Women 16 years and older with higher education, 2017

4.1 Gender gap-the difference between male and female

There are several expectations associated with being male or female. These expectations vary across countries and cultures and has changed the past decades in the females' favour, at least in some countries. Parts of the expectations of being a male or a female will be learned through socialization process. Females will often be seen as nurturing and emotional while male are socialized into more instrumental roles as providing economic support for the family. In addition to these social definitions and expectations, research show that it exists biological differences between males and females.

Scientists has found that males' and females' brains are simple built differently that makes them learn differently²¹. The research found that male brains develops later than female brains and that they prefer different types of learning methods. Males enjoy active learning by doing, while females learn better than males through reading and writing. This cognitive brain differences are backed up by a study done by Magon²². In her investigation, she compares females' success in female-designed classes versus male-designed classes. Using a soft and

²¹ (Gurian & Stevens, 2010)

²² (Magon, 2009)

gentle voice, a warm and cosy classroom putting no pressure on the females, resulted in better results. Males and females are different by nature, and it seems like the biological differences can result in learning differences between the gender. Summarized, we assume that males have worse prospects than females to manage a theoretical school with less focus on activity and learning by doing, compared to the female's prospects.

4.2 Gender gap and Reform 97

The evaluation of Reform 97 showed that the males were getting worse off at school compared to females²³. We believe that males are more impatient than females and that they enjoy more physical learning. The school system in Norway has developed to be a theoretical based school and the research on the males and female's brains support our theory that Reform 97 will affect males more negatively compared to females. Presented in the introduction chapter, numbers from SSB²⁴ show that males are choosing a more active educational program. The research on males and female's brain can help us understand why more females are graduating from high school and attends higher education.

4.3 The rate of time preference

An interesting topic under labour theory is human capital. Human capital is, according to Borjas, the unique set of abilities and acquired skills that individuals bring into the labour market²⁵. Some workers obtain a lot of schooling while others choose to start working in an early age. Workers who finish high school and attend to higher education are willing to give up some earnings today as they expect to be rewarded with higher earnings in the future. The present value of discounted utility can be calculated as follows:

$$PV = \frac{y}{(1+r)^t}$$

The model uses the rate of time preference, r , to discount the utility of the cash flow. As education is associated with higher earnings and more consumption later, it is interesting to investigate why some people are motivated to attend higher education while others drop out from school at an early age. When attending school, the students will face some years of low income but expect higher income in the future. When dropping out or not attend school, the workers will first face higher income than those attending school but may never experience

²³ (Haug, 2003)

²⁴ (Statistics Norway, 2018)

²⁵ (Borjas, 2016)

the same high wage increase in the future. When deciding whether to obtain higher education or not, the concept of present value is used. The present value of an individual's utility for attending high school can be presented like this:

$$PV_{HS} = w_{HS} + \frac{w_{HS}}{(1+r)} + \frac{w_{HS}}{(1+r)^2} + \dots + \frac{w_{HS}}{(1+r)^{46}}$$

where w_{HS} is the wage a person face after graduating from high school in its 46 terms of working life, from 18-64 years old. These 46 terms of a working life are discounted by the time preference rate, r . An important clarification is that it is not the lifetime wage stream that is discounted, but the benefit of having money today, compared to tomorrow.

The time preference rate, r , is a subjective discount rate and shows how costly it is for an individual to attend school. The perceived cost of investing in school is the main difference between males and females when they choose to attend school or not. The costliness will be based on the persons capabilities and patience. A person with a high time preference rate is "impatient" and will appreciate money today more than money tomorrow. A person of lower rate of time preference is more patient and provident. The higher the rate, the less likely a person is to attend higher education.

For students graduating from high school, they can either start working or attend higher education. For a person attending college, the present value could look like this:

$$PV_{COL} = -H - \frac{H}{(1+r)} - \frac{H}{(1+r)^2} - \frac{H}{(1+r)^3} + \frac{w_{COL}}{(1+r)^4} + \dots + \frac{w_{COL}}{(1+r)^{46}}$$

where H shows the direct costs of attending college, discounted three times by the time preference rate, r . When attending school, you also give up three years of paid work. w_{COL} , also discounted by the time preference rate in 46 terms, is the expected lifetime earnings after graduating from college. The post-college wage is expected to be higher than the post-high school wage. A person will attend school if it maximizes the present value of lifetime earnings²⁶. Higher education will be profitable, if the present value of college exceeds the present value of attending high school. If the assumption that males have become more impatient and education are more expensive for them, this will mean that it is less likely to be profitable for them to attend school.

²⁶ (Borjas, 2016)

Attending school will be perceived as costly for students that are not academically strong. The evaluation of Reform 97 showed that male students were worse off after the reform, compared to females. Is there reason to believe that the reform changed the male students' time preference rate when it comes to education? Has the institutional change negatively affected their ability to learn? Their cost of attending school will then be high; their time preference rate has increased. If this is the case, males have become even more impatient after the introduction of Reform 97, compared to females.

As mentioned earlier in this chapter, the biological differences between males and females make a theoretic school more facilitated for the females. The evaluation of Reform 97 reveals a school system that has become more theoretical over the last couple of years. Based on the brain research, we believe that a more theoretical school may have had a negative impact on male's time preference rate. The research shows that female's brains are more objectionable for theoretic learning, and this is what the numbers from Statistics Norway confirms²⁷. The difference between males and females graduating from general studies at high school, are more than 9% in the females' favour in the period 2012-2017. Looking at vocational study program in the same period, it is the other way around. Males exceed females with over 17%.

This paper investigates if Reform 97 has led to fewer males graduating from high school, compared to females. If this is to be the case, the rate of time preference for males has increased more than the females' rates. In other words, more males than females are dropping out from high school after the introduction of Reform 97. This theory underlies this paper's research question: Has Reform 97 resulted in expanding the gender gap between males' and females' high school graduation rate? We will either find a positive or negative effect on the gender gap or find no significant effect. These two possible outcomes can be presented as hypothesis:

H1: Reform 97 has had a positive or negative effect on the gender gap regarding graduation from high school.

H0: No effect.

²⁷ (Statistics Norway, 2018)

5. Empirical strategy

This paper is an attempt to find a possible increase or decrease in the gap between males and females graduating from high school due to Reform 97. To do so, this paper uses a difference-in-difference model (DID) and a regression analysis. To conduct these analyses, we use a data set of quantitative data with an extensive research design.

5.1 Difference-in-difference

The data that is used in this paper are cross-sectional data. This means that samples are drawn from the same population across time; before and after Reform 97, to identify the effect of the treatment. To explore and control for the systematic trend in the difference over the time dimension in the data, we introduce time explicitly in the model specification²⁸. Each individual is observed before the introduction of Reform 97, $t_0 < k$ and after the introduction, $t_1 > k$. Let d_{it} denote the treatment status of individual i at time t . The treatment status will look like this:

$$d_i = \begin{cases} 1 & \text{if } d_{it} = 1 \\ 0 & \text{otherwise} \end{cases}$$

Where d_i shows the treatment status; whether the individual is born to start school before or after the reform. $d_i = 1$ is the treatment group, born in 1991 or later, while $d_i = 0$ is the control group, born before year 1991.

To find the change in gender gap after the reform, we need to exclude the systematic gender gap that already exists in the data set. Therefore, in addition to time, we need to control for the general gender differences. Each individual is observed across time, registered either as male or female. Let m_i denote the gender status for individual i , where the $m_i = 1$ is a male individual and $m_i = 0$ is female.

$$m_i = \begin{cases} 1 & \text{if male} \\ 0 & \text{otherwise} \end{cases}$$

A method to measure the treatment on the treated is by using a difference-in-difference approach. Let $\eta_{M,1991}$ be the treatment effect for male students born to start school after the

²⁸ (Blundell & Costa , 2009)

reform, denoted M for male and 1991 for birthyear. $\eta_{M,1991}$ takes a value of 1 if the individual is male ($m_i = 1$) and in the treated group; born in 1991 or after ($d_i = 1$):

$$\eta_{M,1991} = \begin{cases} 1 & \text{if } m_i = d_i = 1 \\ 0 & \text{otherwise} \end{cases}$$

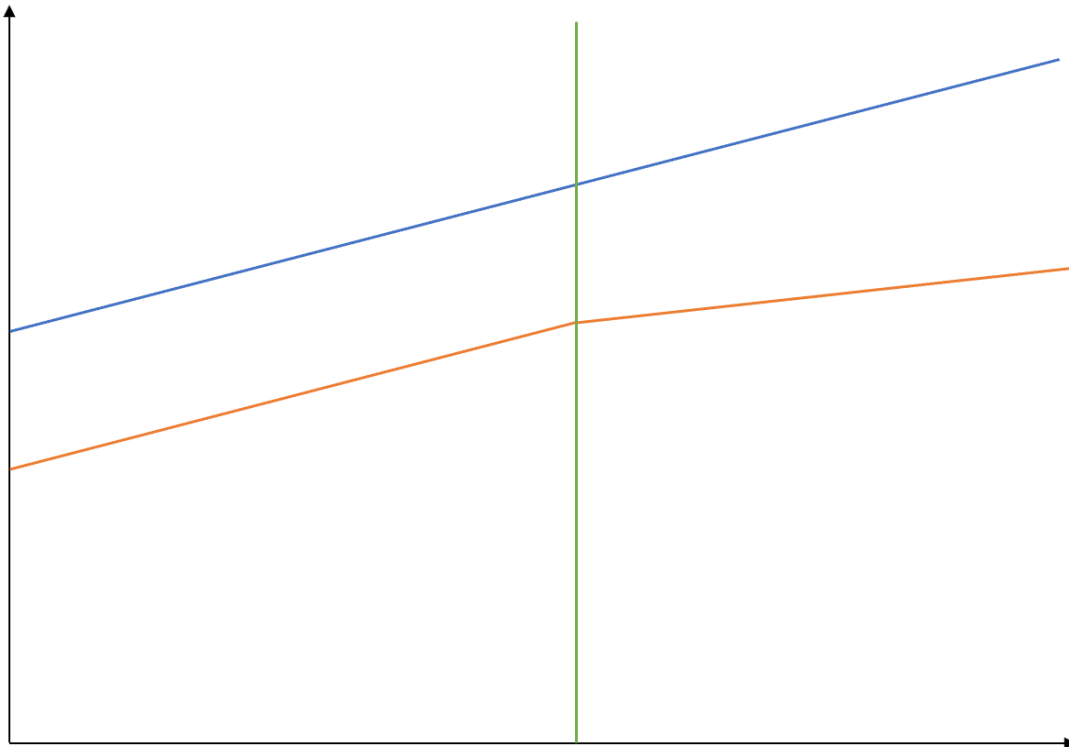


Figure 2: Difference-in-difference estimate shown by arrows between two graphs across time.

To investigate whether there has been a change in the gender gap regarding high school graduation rate, we use the following difference-in-difference calculation:

$$\eta_{M,1991} = (GR_{M,1991} - GR_{F,1991}) - (GR_{M,1990} - GR_{F,1990})$$

Comparing the two years closest to the treatment, we look at the direct affect for the 1990 cohort versus the first treated cohort born in 1991. $\eta_{M,1991}$ is the difference-in-difference coefficient, denoted with M for male and 1991 for birthyear. The formula calculates the gender difference in year 1991 compared to the year before. The first parenthesis: $(GR_{M,1991} - GR_{F,1991})$ represent the high school graduation gender gap in year 1991. The second parenthesis: $(GR_{M,1990} - GR_{F,1990})$ represent the same gender gap, but for those born in year 1990. Both the 1990 and the 1991 cohort started school in 1997, but the 1991 cohort was then 6 years old, while the 1990 cohort was 7; starting directly in second grade.

For Reform 97 to have had an effect on the high school graduation gender gap, we expect the DID coefficient to be statistically significant from birthyear 1991 and after, while showing no significant effect in the birthyears prior to this. Assuming male students have a lower probability of graduating from high school, an increase in the gender gap is shown by the coefficient being negative, the opposite for a decrease in the gender gap.

In this paper we wish to find any possible differences in the graduation rate between males and females that are non-treated: Born between 1987 and 1990, and those treated: Born after. We want to investigate the average effect of the treatment on the treated. This is done by removing unobservable individual effects and common macro effects by relying on two important assumptions²⁹:

1. common time effects across groups
2. no systematic composition changes within each group

For the treatment to be the only factor that is changing over time, these two assumptions have to be met. This is necessary for the coefficients to be unbiased.

This paper will compare the probability of students graduating from high school born before and in 1990 to those born in 1991 and after: The non-treated group consists of those born between year 1987 and 1990, starting school at age 7, while the treated group are those born between 1991 and 1995 who started school at the age of 6 and thus was affected by Reform 97. There may be heterogenous effects in the analysis, implying that the effects will differ across different parts of the population.

²⁹ (Blundell & Costa , 2009)

5.2 Binary logistic regression

The main purpose of a regression analysis is to look at relations between different variables and how they are correlated with each other³⁰. A binary logistic regression is used when you have a binary dependent variable having two possible outcomes, in our case 1 or 0; graduated or not. Since our dependent variable represents whether students have graduated from high school or not; taking value 1 if they have graduated and 0 otherwise, we will use binary logistic regression to find the DID coefficient. A logit regression is different from an OLS regression in the way its output is interpreted. When using logit, we look at the probability for the dependent variable, $Y = 1$, as a function of the independent variables. This means that an increase in one of the independent variables makes the outcome of the dependent variable, $y=1$, more or less likely. It is the sign of the coefficient that are interpreted, not the magnitude. This is because different models have different scales of coefficient.

This paper's binary logistic regression model will look like this:

$$P(Y_i = 1) = \alpha + \sum_{k=1987}^{1995} \beta_1 \text{birthyear} + \beta_2 \text{male} + \beta_3 \text{male} * \text{birthyear} + \sum \beta_4 X + \varepsilon_i$$

The regression formula predicts the probability P that an individual has graduated from high school ($Y_i = 1$) as a function of birth year (*birthyear*), gender (*male*), the DID coefficient (*male * birthyear*) as well as other control variables (X). β_1 is constructed as dummy variables for each year of birth the objects can be born; in our case between year 1987 and 1995, taking a value of 1 if object is born said year, and 0 otherwise. β_2 is a dummy variable taking a value of 1 if object is male, 0 otherwise. These two coefficients explain the general graduation increase and gender gap, but not the change in the gender gap over time; this effect is explained by our DID coefficient: β_3 . The β_3 coefficient is constructed as a dummy variable taking a value of 1 if object is male and born in 1991 or after. If Reform 97 has increased or decreased the high school graduation gender gap, the analysis will show a statistically significant effect in β_3 , the DID coefficient. If the gap is closed (decreased) we expect the variable to be negative, while if the gap has increased, we expect the variable to be positive. To achieve a good and robust analysis, we include more than one independent variable. This is to investigate if there exists other characteristics that may have an impact on the gender gap.

³⁰ (Thrane, 2003)

The betas in the equation show how the probability that $Y = 1$ changes if the variable that the beta represents occurs. In this paper, this means that the betas will say something about the probability for a student to graduate from high school dependent on the year of birth β_1 , gender β_2 , and other independent variables β_4 . This paper will especially investigate any difference between males and females dropping out of high school, and if the relationship between these has changed due to implementing the reform. If it is the case that fewer males graduate from high school compared to females, the regression line will be steeper for the males because beta represents the slope of the regression line.

The significance level for each of the independent variables will be read out from the p-value. The p-value says something about the probability that the independent variables will explain the variance in the dependent variable and how significant it is. The p-value always takes a number between 0-1 and the closer to 0, the better. Using models with significance level, it is normal to use a 1%, 5% or 10% significance level, which let us know how explanatory each independent variable is. If some of the independent variables have a p-value higher than 10%, they are not statistically significant which means that the variable is not explanatory. Then the null hypothesis is rejected; this independent variable does not explain the variation in the dependent variable.

It is not likely that the dependent variable is only affected by the model's included independent variables. The main goal with a socioeconomic regression analysis is to reveal not all the independent variables that affect the dependent variable, but to find the most important variables that affect the independent variable the most³¹. The variables that are not included in the regression analysis but still has an effect on the dependent variable, are caught up by the residual, ε_i . In addition to this, the residual also represents measurement errors and inaccuracy. The residual should be as small as possible, if not, the regression represents an inaccurate and less credible analyse.

³¹ (Midtbø, 2007)

6. Data and Sample Description

This chapter will present the data set that is used to prepare the analysis. This paper uses register data collected from microdata.no, a collaboration between The Norwegian Centre for Research Data and Statistics Norway. Only summary statistics will be presented, as the register data are confidential.

6.1 Microdata

Microdata.no has been developed in collaboration between the Norwegian Centre for Research Data (NSD) and Statistics Norway (SSB). The service provides researchers and students with approved research institutions access to use register data from SSB. Microdata has available data on population, education, labour market, and social security. The database consists of individual demographic variables like the year of birth, number of siblings, immigration status and socioeconomic variables like education. The individual data have information on the Norwegian population from 1900 to 2016. All data are anonymizing, and privacy policy is taken care of.

This data set was chosen for this paper as register data are raw and individual data that is a good basis for further research. It contains interesting and relevant variables collected from the whole population as year of birth, gender, immigration background, number of siblings and parents' educational level.

6.2 Variables

In this sub chapter, we will announce all variables that are included in the binary logit regression model. We will explain how the variables were conducted and what information the variable gives us. Some variables are divided into categories which will also be explained. All the information on the variables is collected from the main source of the data set, Microdata³². The sample selection consists of students in Norway born from year 1987 to 1995.

Graduation: Constructed as a dummy variable, taking a value of 1 if object graduated from high school within the age of 20, 0 otherwise.

Gender: Constructed as a dummy variable, taking a value of 1 if object if male, 0 otherwise.

Year of birth: Constructed as dummy variables for each year 1987-1995, taking a value of 1 if object is born in said year and 0 otherwise, where the year 1987 is the base year.

Parents' education level: Constructed as dummy variables for both mothers' and fathers' education level. Taking a value of 1 if the parent has university degree, 0 otherwise. Missing values have been gathered in a separate dummy variable for when parent's education level is unknown.

Number of siblings: Constructed as dummy variables for 0, 1, 2 and 3 and more siblings, taking a value of 1 if object has 0, 1, 2 or 3 or more siblings respectively, 0 otherwise. Missing values have been gathered in a separate dummy variable for when number of siblings is unknown.

Immigration status: Constructed as dummy variables for Norwegian, Western and non-Western³³ immigrants, taking a value of 1 if object is Norwegian, Western or non-Western respectively, 0 otherwise.

32 (Microdata, 2017)

33 Western: consisting of countries in the EU/EEA, USA, Canada, Australia and New Zealand.

Non-Western: consisting of countries in Asia, Africa, Latin-America, Oceania except Australia and New Zealand and Europe without EU/EEA. (Statistics Norway, 2011)

7. Empirical Results

7.1 Descriptive results

Before presenting the difference-in-difference coefficient, we will present the descriptive results. Table 1 below shows the number of students graduating high school within the year they turn 20, shown as a fraction of the birth cohort. For example, the number 0.51 in 1988 shows that 51% of the birth cohort born in 1988 graduated high school within the age of 20.

1987	1988	1989	1990	1991	1992	1993	1994	1995
0.495	0.512	0.536	0.547	0.562	0.571	0.578	0.612	0.627
(0.5)	(0.5)	(0.499)	(0.498)	(0.496)	(0.495)	(0.494)	(0.487)	(0.484)

Table 1: Share of students graduating high school, shown as a fraction of birth cohort each year³⁴. Robust standard errors in parentheses.

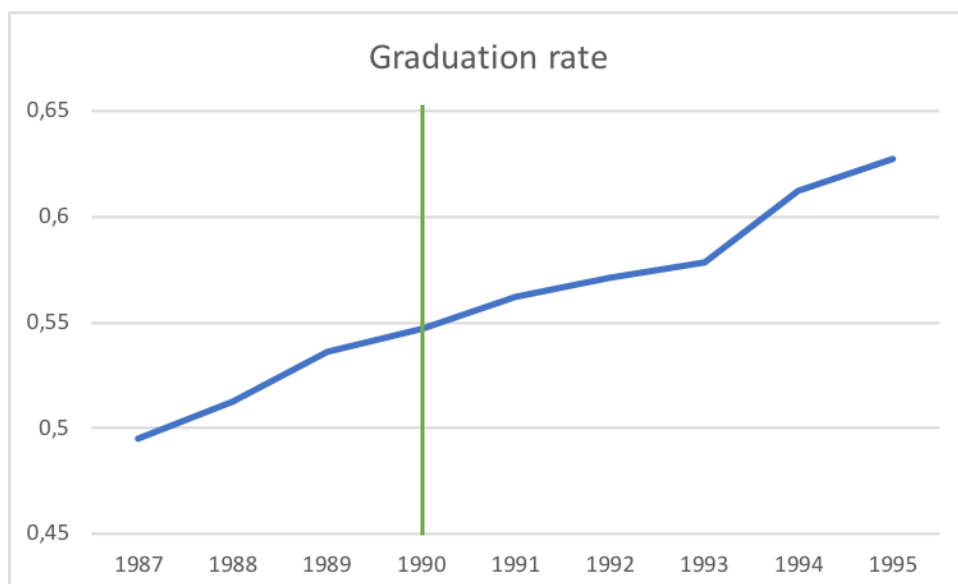


Figure 3: Students' high school graduation rate in Norway, years 1987-1995. Numbers from Table 1.

Figure 3 shows a graph of the output found in Table 1 above. The general trend found in Figure 3 shows a steady increase in students' graduation rate over time from birthyear 1987 to 1995. We see the trend flattening out around from year 1989 until 1993, but then there is a sudden larger increase from 1993 to 1994 from the value 58% to 61%. The share of the birth cohort graduating high school range from 49% in 1987 to 63% in 1995.

³⁴ Complete numbers can be found in Appendix 3.

When looking at Table 1 and Figure 3, it seems like Reform 97 did not have any effect on the overall graduation rate for students in Norway, as there seems to be an even increase in the graduation rate from 1987 to 1995. We do, however, find a slight jump from the 1993- to the 1994-cohort. Although it can be argued that this jump might be a delayed effect from Reform 97, this is an effect we should have observed with the 1991 cohort, as it was the first to be exposed to the change. The likeliest case is that this jump was not caused by Reform 97, but something else later in time.

In Table 2 below we present the information from Table 1 divided into gender. Table 2 shows the fraction of females and males born each year that graduated high school within the age of 20. For example, the number 0.63 in 1992 in Table 2 shows that 63% of females born in 1992 graduated high school within the age of 20, compared to 52% for the male students.

Figure 4 presents Table 2 as graphs. The graphs show that both female and male students' graduation rate had a positive trend, increasing every year from 1987 to 1995. We also find the same sudden increase from 1993 to 1994. The figure shows the effect from Reform 97 from birthyear 1990 to 1991 and it seems that the graph showing the males' graduation rate flattens out while the females' graduation rate keeps increasing. This effect on gender gap is shown in Figure 5, where we find a stable gender difference of around 10% in the years 1987-1990 and an increase in the gender gap in 1991 and after.

The green line shown in both Figure 4 and 5, represents the point of implementation of Reform 97. From this point we find the possible change in the high school graduation gender gap. In Figure 4, the change will be shown as an increase or decrease in the distance between the two graphs (for males and females). If the reform had an effect on the gender gap, we will find a new increase or decrease in the graph in Figure 5 from the green line (after year 1990). Seen together with Figure 4, we read that the increase in graduation gender gap is explained by females' graduation rate increasing more than males' graduation rate.

	1987	1988	1989	1990	1991	1992	1993	1994	1995
Female	0.548 (0.498)	0.562 (0.496)	0.588 (0.492)	0.598 (0.4903)	0.62 (0.485)	0.632 (0.4824)	0.643 (0.479)	0.683 (0.465)	0.698 (0.459)
Male	0.443 (0.497)	0.464 (0.499)	0.486 (0.5)	0.497 (0.5)	0.507 (0.5)	0.515 (0.5)	0.517 (0.5)	0.546 (0.498)	0.562 (0.496)

Table 2: Share of female and male students graduating high school, shown as a fraction of birth cohort each year³⁵. Robust standard errors in parentheses.

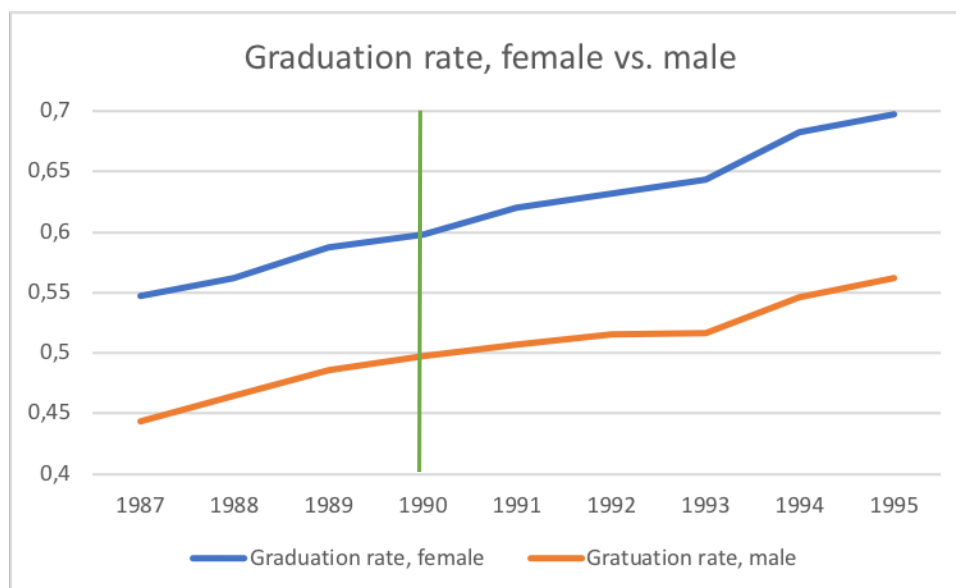


Figure 4: Graduation rate for students in Norway in the years 1987-1995, divided into gender. Numbers from Table 2.

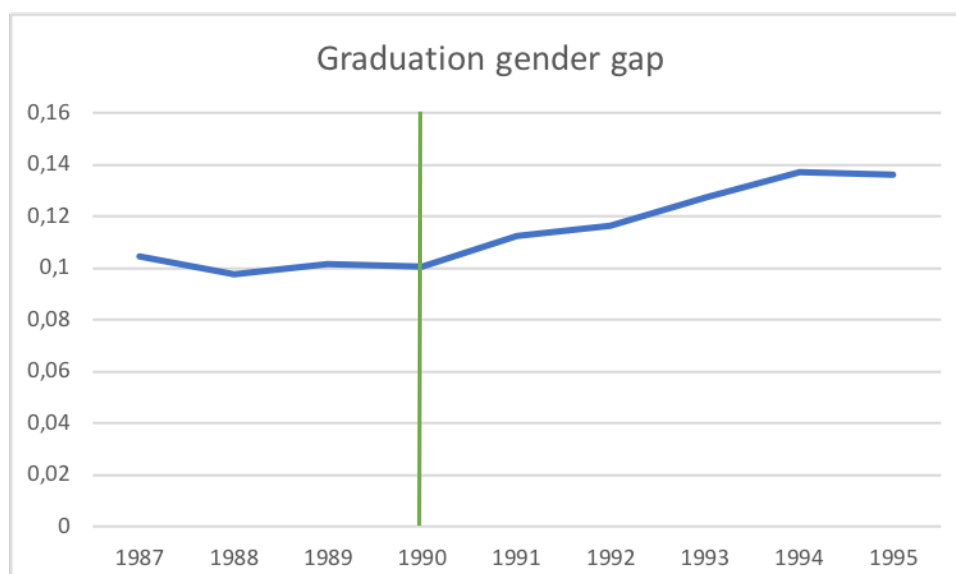


Figure 5: Difference in gender in graduation rate for students in Norway in the years 1987-1995. Calculated as difference between female and male in Table 2.

³⁵ Complete numbers can be found in Appendix 3.

	All	Pre	Post
Male	0.512 (0.5)	0.508 (0.5)	0.516 (0.50)
Mother education	0.287 (0.453)	0.250 (0.433)	0.319 (0.466)
Father education	0.240 (0.427)	0.219 (0.413)	0.259 (0.438)
Mother edu unknown	0.142 (0.349)	0.197 (0.397)	0.096 (0.295)
Father edu unknown	0.158 (0.365)	0.211 (0.408)	0.113 (0.317)
Only child	0.116 (0.320)	0.130 (0.336)	0.105 (0.306)
1 sibling	0.355 (0.479)	0.329 (0.470)	0.378 (0.485)
2 siblings	0.270 (0.444)	0.242 (0.428)	0.294 (0.456)
3 siblings or more	0.102 (0.303)	0.094 (0.292)	0.109 (0.312)
Siblings unknown	0.156 (0.363)	0.205 (0.403)	0.114 (0.318)
Norwegian	0.809 (0.393)	0.762 (0.428)	0.849 (0.358)
Western	0.089 (0.285)	0.118 (0.323)	0.064 (0.245)
Non-Western	0.102 (0.302)	0.120 (0.120)	0.087 (0.281)

Table 3: Mean values of the different variables in our data set, divided into pre- and post- segments³⁶. Variables explained in chapter 3. Robust standard errors in parentheses.

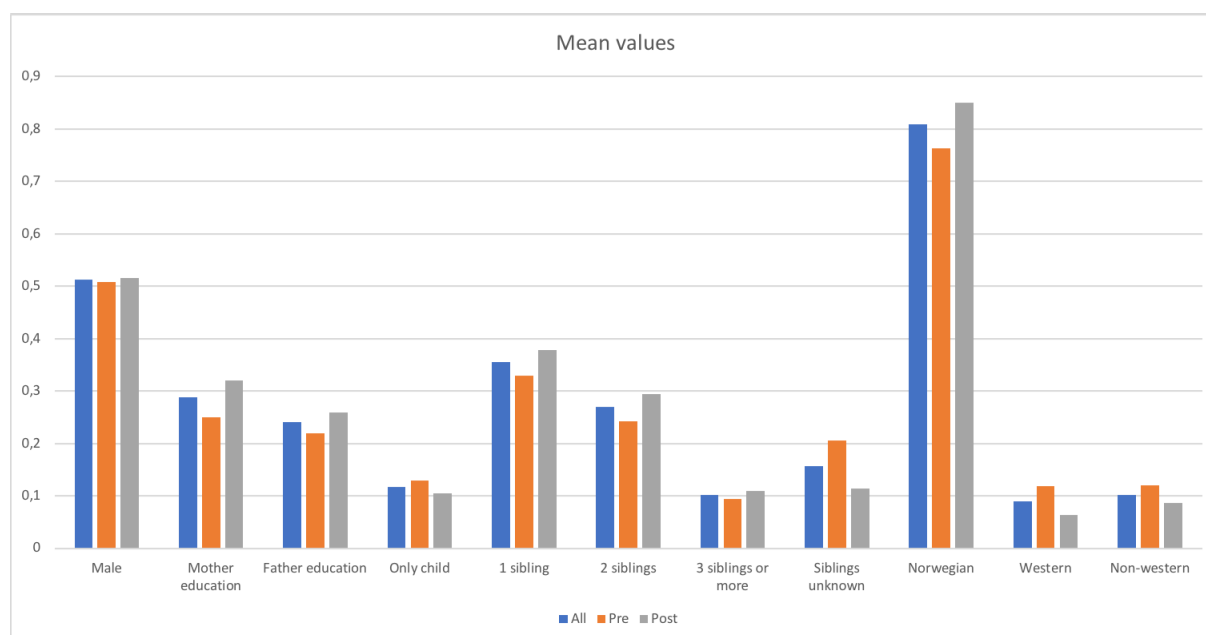


Figure 6: Mean values from the data sample, as well as the values before and after Reform 97. Numbers from Table 3.

³⁶ Complete numbers can be found in Appendix 3.

The descriptive statistics found in Table 3 above present mean values and standard deviation for our independent variables. The table presents the overall mean values as well as birth cohorts treated (post) and non-treated (pre) by the reform. The share of male students have not changed dramatically. There is a steady share of both genders of around 50% each; although we see that the share of male students is generally slightly higher.

The share of students whose mothers have obtained a higher education is higher than the share of students whose fathers have a higher education. The trend over time for both cases is positive; the share of parents having higher education is higher after the implementation of Reform 97 (post), than before (pre). We find a slight increase in parents' education level; from 25% to 32% for mothers, and 22% to 26% for fathers. Interesting observation is that the share of mothers obtaining higher education has increased more than the fathers' have. The objects whose parents have unknown education are not presented here, but will be included in the further regression analysis.

Regarding siblings, the table shows that most students have 1 sibling, both before and after the reform. There seems to be a general increase in the number of siblings over time, as the variables *1 sibling*, *2 siblings*, and *3 siblings or more* are all higher in the post segment. There is also a share of the objects who have not registered number of siblings of around 16%, but this has decreased over time. We assume that this will not affect our results dramatically.

We have divided our data set into three segments, according to where the objects are born. The largest part is the Norwegian segment, which has increased from 76% before to 85% after the reform was implemented. The immigration segments; type 1 for Western countries and type 2 for non-Western countries, have both decreased in the same period.

We have created three variables for missing information; mothers' and fathers' education unknown and number of siblings unknown. The decrease in the unknown parents' education level variable is a positive, because it means that more information has been collected. We see the same trend for the siblings' unknown variable, decreasing from 20% to 11%.

	Norwegian		Western		Non-Western	
	Pre	Post	Pre	Post	Pre	Post
Graduation high school	0.630 (0.483)	0.649 (0.477)	0.170 (0.376)	0.260 (0.438)	0.186 (0.389)	0.254 (0.435)
Male	0.513 (0.5)	0.516 (0.50)	0.523 (0.50)	0.50 (0.5)	0.46 (0.498)	0.529 (0.499)
Mother education	0.309 (0.462)	0.353 (0.478)	0.047 (0.212)	0.143 (0.35)	0.071 (0.257)	0.118 (0.322)
Father education	0.268 (0.443)	0.282 (0.450)	0.038 (0.190)	0.113 (0.317)	0.084 (0.278)	0.133 (0.34)
Mother edu unknown	N.A.	N.A.	0.891 (0.312)	0.672 (0.47)	0.727 (0.446)	0.559 (0.497)
Father edu unknown	0.018 (0.133)	0.016 (0.124)	0.92 (0.272)	0.731 (0.444)	0.745 (0.436)	0.615 (0.487)
Only child	0.160 (0.366)	0.116 (0.320)	0.021 (0.143)	0.035 (0.184)	0.045 (0.207)	0.047 (0.211)
1 sibling	0.410 (0.492)	0.424 (0.494)	0.043 (0.203)	0.103 (0.304)	0.096 (0.295)	0.126 (0.332)
2 siblings	0.302 (0.459)	0.333 (0.471)	0.028 (0.166)	0.070 (0.255)	0.070 (0.256)	0.082 (0.275)
3 siblings or more	0.107 (0.309)	0.114 (0.318)	0.013 (0.114)	0.032 (0.177)	0.095 (0.294)	0.12 (0.32)
Siblings unknown	0.021 (0.145)	0.014 (0.116)	0.895 (0.307)	0.76 (0.427)	0.693 (0.461)	0.625 (0.484)

Table 4: Mean values of the different variables in our data set, divided into immigration status; Norwegian, Western and non-Western³⁷ immigrants, as well as before and after (pre and post) the reform was implemented³⁸. Robust standard errors in parentheses.

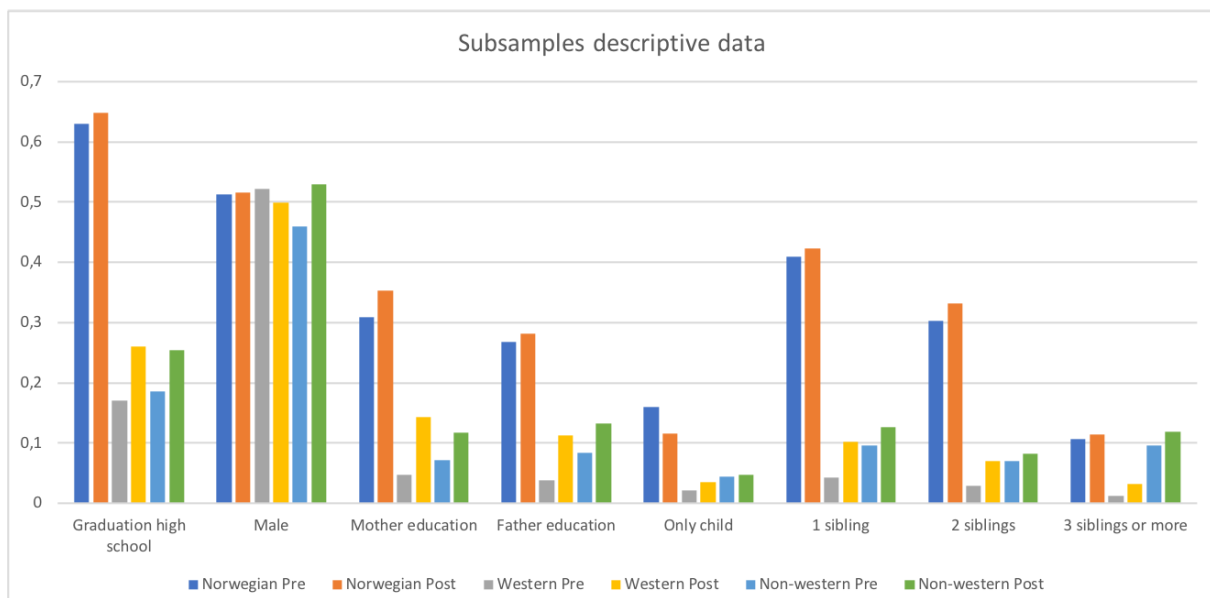


Figure 7: Showing the shared of different variables in our data set, divided into immigration segments and before and after Reform 97 was implemented.

³⁷ (Statistics Norway, 2011)

³⁸ Complete numbers can be found in Appendix 3.

The graduation rate presented in Table 4 and Figure 7 above shows a large difference between Norwegian and immigrated students. While Norwegian students on average graduate 63-65% of the time, students from Western countries only graduate around 17-26% and students from non-Western countries have almost the same graduation rate; around 19-25%. This is a considerable difference, larger than we expected, but there might be disturbing factors in the data set explaining this variation.

As expected, the gender rate is around 50% for all immigration segments. The only noticeable diversity from this is in the non-Western segment, where the share of male students was lower than 50%, around 46% before the reform was implemented, and increased to 53% after the reform. We believe that this should, however, not have any noteworthy impact on our analysis.

Looking at mothers' education level, Table 4 shows an increase for all immigration segments over time. The change is larger for immigrants, especially for immigrants from Western countries with an increase from 5% pre to 14% post reform. We see the same growth for fathers' education level over time, increasing for all segments, but also here largest for Western countries. We are aware that both of these variables may have been affected by disturbing factors in the data set, as well as missing information for some objects. Some of the increase in parents' education may be explained by the decrease in the parents' education unknown variables.

Table 4 indicates that Norwegian students have both more siblings as well as they are more often only children, which of course is not possible. This is explained by the high share of unknown number of siblings' variable for immigrants; 90% for Western and 69% for non-Western immigrants. This makes it difficult to compare the actual number of siblings across immigration segments. The size of the unknown variables decreases over time, to 76% for Western and 63% for non-Western, but is still a very large share compared to the Norwegian siblings' unknown share. However, we still observe an increase in number of siblings over time across all immigration segments. Because the Norwegian segment is such a big segment, and because the number of siblings is not the main focus of our problem, we have decided to include the number of siblings' variable in our analysis even after facing these issues. We will also conduct a subsample regression analysis to check for the difference in immigration

segments. This way we will be able to look at the Norwegian segment only and avoid the number of siblings' issue.

7.2 Main results

If Reform 97 had an effect on the gender gap in the graduation rate from high school, we expect to see significant results, either positive; the gender gap has decreased and males' compared to females' graduation rate has increased, or negative; the gender gap has increased and males' compared to females' graduation rate has decreased. This gender gap increase will be shown by the interaction variable male*post or male*birthyear, our β_3 from the binary regression model presented in chapter 5.2.

7.2.1 Short binary regression model

	Model 1	Model 2	Model 3	Model 4	Model 5
Post	0.341*** (0.007)	0.064*** (0.008)	0.114*** (0.008)	0.186*** (0.008)	0.057*** (0.008)
Male	-0.406*** (0.007)	-0.569*** (0.008)	-0.542*** (0.008)	-0.499*** (0.008)	-0.583*** (0.009)
Male*post	-0.117*** (0.010)	-0.049*** (0.011)	-0.047*** (0.011)	-0.082*** (0.011)	-0.042*** (0.011)
N	633081	633081	633081	633081	633081
Pseudo R ²	0.013	0.158	0.123	0.099	0.168

Table 5: Binary logit regression output. Model 1 regresses probability of high school graduation on post, male and male*post. Model 2 includes parents' education, Model 3 siblings and Model 4 immigration status. All variables from Model 1-4 is included in Model 5. *, ** and *** denote significance at 10 percent, 5 percent and 1 percent level respectively. Robust standard errors in parentheses.

We have conducted a short regression model presented in Table 5 above. Model 1 is a binary logit regression showing graduation as a function of treated objects (shown as post, taking a value of 1 if object is born between 1991 and 1995, 0 otherwise), gender (shown as male, taking a value of 1 if male, 0 otherwise) and the DID coefficient (shown as male*post, taking a value of 1 if object is male and born between 1991-1995, 0 otherwise).

In Model 2 we have included both parent's education level as well as missing variables in the regression, not presented in Table 5³⁹. We can see from Model 2 that the output changes substantially when we include these control variables. The same problem occurs in Model 3,

³⁹ All output from regression in Table 5 can be found in the Appendix 4.

where we have included dummy variables for number of siblings as well as a missing variable, not presented in Table 5³⁴. Model 4 includes dummy variables for immigration status (Norwegian, Western, non-Western). We see that this model, too, changes the output from the regression. This model produces an output more similar to Model 1, than Model 2 and 3 does. In Model 5 all variables have been included; parents' education, number of siblings and immigration status. From the output from Model 2-4 we expect Model 5 to have different output compared to Model 1. Model 5 confirms that assumption and we see changes across all models in the table.

Common to all models is a positive value in β_1 (birthyear) and negative values for both β_2 (male) and β_3 (male*post). β_1 shows the general trend over time, meaning that students born in 1991 or after (post) have a higher probability of graduating from high school than those born before 1991 (pre). β_2 shows that on average male students have a lower probability of graduating from high school than female students. These two coefficients explain the general graduation increase and gender gap, but not the change in the gender gap over time; this effect is explained by our DID – the β_3 coefficient (male*post). This coefficient is negative, showing an increase in the gender gap after Reform 97 was implemented. In addition, all output is significant at 1% level. However, because the output changes substantially between the models, we conclude that our results are not robust.

7.2.2 Full binary regression model

If Reform 97 had a negative effect on males' graduation rate, we expect a negative and significant β_3 from year 1991 and after; the years before should show no effect as they are not treated.

Table 6, presented below, shows the complete data set binary logit regression models explaining the graduation rate based on birthyear; β_1 , gender; β_2 , the DID coefficient; β_3 and several control variables; β_4 . The effect on the gender gap from the reform is shown by the β_3 from year 1991. Model 6 shows a more general model, looking at the effect of gender and birthyear on graduation rate. Model 7 includes parents' education to see how this affects the graduation rate. Model 8 looks at number of siblings, which are divided into three segments; 1 sibling, 2 siblings and 3 or more siblings, with only children being the control group. Model 9 look at immigration status, also divided into three segments; Norwegian, Western and non-Western. The last model, Model 10, includes all independent variables mentioned above. When dividing the data sample into five models, we do a robustness test; adding more and more control variables to check how it impacts the output in our model. We investigate the effects from adding more control variables on the DID coefficient.

7.2.2.1 Model 6: General model

In Model 6 found in Table 6 below we look at the effect from birthyear (not presented in the table, gender and the DID coefficient on the high school graduation rate. The model is from a logit regression model. New binary variables for birthyears, ranging from 1987 to 1995, have been created, each taking a value of 1 if object is born said year and 0 otherwise. The male coefficient takes a value of 1 if object is male, 0 otherwise. The birthyear*male coefficient takes the value of 1 if object is born in said year *and* is male, 0 otherwise. Female students born in 1987 is our base group for the analysis.

	Model 6	Model 7	Model 8	Model 9	Model 10
1987*male					
1988*male	0.027 (0.021)	0.014 (0.024)	0.037 (0.023)	0.023 (0.023)	0.022 (0.024)
1989*male	0.008 (0.02)	0.008 (0.024)	0.038 (0.023)	0.007 (0.023)	0.023 (0.024)
1990*male	0.011 (0.021)	0.023 (0.024)	0.049** (0.023)	0.006 (0.023)	0.039 (0.02)
1991*male	-0.041* (0.021)	-0.013 (0.024)	0.019 (0.023)	-0.027 (0.023)	0.006 (0.024)
1992*male	-0.059** (0.022)	-0.003 (0.024)	0.017 (0.024)	-0.039* (0.023)	0.013 (0.024)
1993*male	-0.106*** (0.022)	-0.017 (0.024)	0.007 (0.024)	-0.056** (0.023)	0.004 (0.024)
1994*male	-0.165*** (0.022)	-0.088*** (0.024)	-0.071*** (0.024)	-0.131*** (0.023)	-0.073*** (0.03)
1995*male	-0.171*** (0.022)	-0.076*** (0.024)	-0.062*** (0.024)	-0.121*** (0.024)	-0.061** (0.03)
Male	-0.419*** (0.022)	-0.580*** (0.017)	-0.574*** (0.017)	-0.509*** (0.016)	-0.604*** (0.017)
Mother education		0.694*** (0.007)			0.670*** (0.007)
Father education		0.699*** (0.008)			0.710*** (0.008)
Mother edu unknown		-1.512*** (0.016)			-0.702*** (0.020)
Father edu unknown		-0.885*** (0.014)			-0.382*** (0.016)
1 sibling			0.355*** (0.009)		0.293*** (0.009)
2 siblings			0.409*** (0.01)		0.295*** (0.009)
3+ siblings			0.090*** (0.011)		0.093*** (0.012)
Siblings unknown			-2.214*** (0.012)		-0.962*** (0.018)
Western				-1.935*** (0.011)	-0.05*** (0.018)
Non-Western				-1.882*** (0.010)	-0.509*** (0.014)
N	633081	633081	633081	633081	633081
Pseudo R ²	0.015	0.158	0.124	0.099	0.168

Table 6: Full binary logit regression model regressing graduation rate on year of birth and gender, extended into 5 models with variables on parents' education, number of siblings and immigration status⁴⁰. *, ** and *** denote significance at 10 percent, 5 percent and 1 percent level respectively. Robust standard errors in parentheses.

⁴⁰ Output for all regression in Table 6 can be found in the appendix 6.

From Model 6 we find a general negative trend on the gender gap, shown in the interaction variables, from the year 1991. In agreement with our expectations, we see no change in the gender gap before year 1991. Before the treatment in 1991, the interaction variables show positive, but not statistically significant, values, and turn to negative values in 1991, increasing over time to 1995. This means that the high school graduation gender gap was decreasing, or closing, for students born in 1987-1990, but then started to increase.

Model 6 implies that Reform 97 had a significant negative effect on the high school graduation gender gap; the gender gap increased and male students were worse off after the reform. However, this model only studies birthyear and gender and may therefore not explain much of the variation in the probability of graduating high school. What happens when we include other factors, control variables, in the model?

7.2.2.2 Model 7: Parents' education level

In Model 7 presented in Table 6 we investigate the effect of including parents' education level on the high school graduation gender gap. This is our first robustness test, checking if our results in Model 6 changes when we introduce parents' education level as a control variable. We find that including the parents' education level have some effect on the results. The general gender difference is still negative; the share of male students graduating from high school is lower than for female students, and we find a negative trend over time on the male*birthyear coefficients. Here the β_3 coefficients are not statistically significant until year 1994, which makes us unable to reject our null hypothesis; there is no effect from the reform on the high school graduation gender gap. The effect from the parents' education itself is positive and significant at 1% level. The parents' education level unknown shows a large and highly significant, negative effect on object's high school graduation rate. This can be explained by the amount of missing information about immigrants' parents educational background.

7.2.2.3 Model 8: Number of siblings

The control variables for number of siblings is included in Model 8 found in Table 6. The number of siblings variables are created as five dummy variables; no siblings, one sibling, two siblings, three or more siblings and siblings unknown, each taking a value of 1 if the object has no siblings, one sibling, two siblings, three or more siblings or siblings is unknown,

and 0 otherwise. Including these control variables in our model changes the outcome on our DID coefficients slightly. The output of the regression has changed from including the number of siblings' control variables, both in magnitude and significance.

The number of siblings variables alone show some interesting effects on high school graduation rate. All variables showing 1 siblings or more have positive effects on high school graduation rate, while siblings unknown shows a strong negative effect. Comparing the three siblings coefficients, we see that having one sibling and having two siblings highly increases the probability of the object graduating from high school, while having three siblings or more has a weaker positive effect. Regarding the unknown siblings variable, it is interesting to note that most of these objects are from one of the two immigration segments, and it is likely that much of this variance can be explained using the immigration segments, rather than number of siblings.

7.2.2.4 Model 9: Immigration status

The immigration status is divided into three segments: Norwegian; taking a value of 1 if object is born in Norway and 0 otherwise, Western; taking a value of 1 if object is born in a Western country and 0 otherwise, and non-Western; taking a value of 1 if object is born in a non-Western country and 0 otherwise.

Compared to Model 6 we don't see any major differences on our DID estimates when adding the immigration status to the model. The 1991*male coefficient is no longer statistically significant at 10% level, but the magnitude and general signs of the coefficients are close to equal in Model 9 as Model 6. We find a negative turn in the gender gap trend from year 1991, which agrees with our theory that male students are worse off after Reform 97. This effect increases over time and we see that the later the birthyear, the higher the statistical significance of the coefficient.

The Western and non-Western immigration variables alone affect the high school graduation rate negatively, and the values are statistically significant at 1% level. This means that immigrated students have a lower probability of graduating from high school than Norwegians.

7.2.2.5 Model 10: All control variables

In Model 10 all the control variables from the Models 7-9 are included in the binary logit regression model. The model is not robust across models, as the result changes quite a lot when we add more and more control variables to the model. The model captures a strong negative effect from the male coefficient, β_2 , as well as similar effects from the other control variables in our model. The male*birthyear coefficients, β_3 , show a negative trend from year 1994, and the results are not statistically significant until that year. If the reform affected the gender gap, we would see this negative and significant effect from birthyear 1991. We can therefore not reject our null hypothesis; Reform 97 had no effect on the high school graduation gender gap.

Studying the control variables found in Model 10, we see that the parents' education level variables have almost the same effect as they showed in Model 7, both in sign and size. We see, however, a slight decrease in the magnitude of the parents' education unknown variable. This variation has now been caught up in other independent variables in our model. We see the same trend on the number of siblings' variables. The effect has been lowered in size for almost all the number of siblings' variables (except for three siblings or more), while the largest and most noticeable change is the one for unknown number of siblings. This variation has now been caught up in other independent variables, such as the immigration status as previously discussed.

Other factors than those included in our models may affect the change in the high school graduation gender gap, i.e. the new reform implemented in 2006, Kunnskapsløftet, may have had disturbing effects on our results. Some of the objects in our data set were still in school when this reform was implemented, and the effects captured by our regression may be a result from the Kunnskapsløftet reform, rather than Reform 97 that we want to investigate.

Only Model 6 shows the expected result that the gender gap has increased as a result of Reform 97. Including control variables resulted in dramatical changes in our estimates, indicating that our results are not robust. There might be heterogeneous effects; different effects on different parts of the sample. Do male students with higher educated parents have a higher probability of graduating high school than those without? Or do non-Western immigrants struggle more compared to other students when it comes to graduating high school? To address these questions, we conduct a subsample analysis.

7.2.3 Subsamples

In a subsample analysis, we look at different parts of the whole sample individually. The sample is divided into groups, subsamples, based on similar characteristics; for example, immigration status.

7.2.3.1 Short model subsample

	Model 11		
	Norwegian	Western	Non-Western
Post	0.091*** (0.009)	0.563*** (0.029)	0.600*** (0.026)
Male	-0.580*** (0.009)	-0.146*** (0.029)	-0.198*** (0.028)
Male*post	-0.012 (0.012)	-0.061 (0.042)	-0.397*** (0.039)
N	512415	56219	64447
Pseudo R ²	0.015	0.013	0.014

Table 7: Short binary logit regression model showing the probability of graduating high school dependent on gender and whether they were exposed to the treatment or not (post). Subsamples showing regression output for Norwegian, Western and non-Western⁴¹ students⁴². *, ** and *** denote significance at 10 percent, 5 percent and 1 percent level respectively. Robust standard errors in parentheses.

Table 7 above shows output from three different binary logit regressions for the different subsamples: Norwegian, Western and non-Western students. Based on the non-robust output found in Table 5 and 6, we found reason to believe that immigration status highly affects the independent variables, β 's, in our model. We want to test if different immigration statuses impact the β 's differently; does the probability of graduating from high school change depending on where a student is born?

Model 11 is a binary logit regression model, explaining the probability of graduating from high school dependent on birthyear, β_1 , gender, β_2 , and the interaction variable of the two, β_3 . It is divided into subsamples for Norwegian, Western and non-Western students. The change over time, shown as post, indicate that the graduation rate for immigrants have a higher growth than for Norwegian students. Comparing genders, we see that males generally have a lower probability of graduating from high school in all segments, but the effect is larger for the non-immigrants. The effect from the variables post, β_1 , and male, β_2 , are significant at 1% level.

⁴¹ (Statistics Norway, 2011)

⁴² Output for all regression in Table 7 can be found in the appendix 5.

The gender gap for non-Western students are most negatively affected by Reform 97. Only for non-Western students is this β_3 statistically significant (at 1% level). The model indicates that the reform had no significant effect on the gender gap in high school graduation for Norwegian and Western students. The evaluation of Reform 97, conducted by Haug⁴³, found that immigrated students were worse off after the reform. Our results indicate that this effect was stronger for non-Western immigrants. From this we can argue that male non-Western immigrants are more vulnerable for the change from Reform 97.

7.2.3.2 Full model subsample

Table 8 presents three different subsamples from our sample: parents education level, immigration status and number of siblings⁴⁴. Though number of siblings and parents' education level has shown to not give any usable information due to missing data, we have decided to present the results of these regressions in the table and briefly comment the results.

Model 12.1 and 12.2 for mother's and father's education level is divided into high and low; high if parent has university degree and low if not⁴⁵. The β_3 's (birthyear*male) show only positive or no effect from the reform on the gender gap. This contradicts the actual change in the gender gap, as the change in gender gap for the whole sample is negative (gap increasing). This can be explained by the missing data for parents' education level. From Appendix 7 we see that the negative trend in the gender gap is explained by the parents' education level being unknown. As presented in Table 4, immigrants have the highest share of missing information on parents' education level, ranging from 56% to 92%. In comparison, parents' education level for Norwegian students has less than 2% unknown data. Thus, we can argue that the negative trend in the gender gap from parents' education level is indirectly explained by immigration status.

⁴³ (Haug, 2003)

⁴⁴ Output from all regression in Table 8 can be found in Appendix 7.

⁴⁵ Regression output for subsample parent's education level unknown is found in Appendix 7.

	Model 12.1: Mother education		Model 12.2: Father education		Model 13: Immigration			Model 14: Number of siblings		
	Low	High	Low	High	Non-Western	Western	Norwegian	0	1	2 or more
1987										
1988*male	0.051* (0,030)	0.041 (0,054)	0.045 (0,029)	-0.008 (0,058)	-0.150* (0,087)	0.026 (0,086)	0.048* (0,027)	0.003 (0,059)	0.034 (0,041)	0.089** (0,042)
1989*male	0.065** (0,030)	0.088 (0,053)	0.075** (0,029)	0.011 (0,058)	-0.284*** (0,088)	-0.131 (0,087)	0.079*** (0,027)	0.115* (0,060)	0.043* (0,040)	0.087** (0,041)
1990*male	0.084*** (0,029)	0.118** (0,053)	0.093*** (0,029)	0.042 (0,057)	-0.265*** (0,089)	-0.150* (0,090)	0.102*** (0,026)	0.07 (0,061)	0.077 (0,040)	0.114*** (0,041)
1991*male	0.053* (0,029)	0.061 (0,052)	0.056* (0,029)	0.008 (0,057)	-0.281*** (0,090)	-0.075 (0,093)	0.066** (0,026)	0.047 (0,062)	0.017 (0,040)	0.103** (0,040)
1992*male	0.055* (0,030)	0.121** (0,052)	0.083*** (0,029)	-0.021 (0,056)	-0.336*** (0,091)	-0.250** (0,098)	0.084*** (0,026)	0.187*** (0,063)	0.010 (0,040)	0.122*** (0,040)
1993*male	0.085*** (0,03)	0.042 (0,051)	0.055* (0,029)	0.062 (0,056)	-0.30*** (0,093)	-0.326*** (0,104)	0.077*** (0,026)	0.144** (0,065)	0.029 (0,040)	0.105*** (0,040)
1994*male	-0.021 (0,030)	0.046 (0,051)	-0.002 (0,029)	-0.056 (0,056)	-0.464*** (0,093)	-0.408*** (0,105)	0.013 (0,027)	-0.002 (0,066)	-0.037 (0,040)	0.051 (0,041)
1995*male	0.018 (0,030)	-0.015 (0,051)	0.015 (0,030)	-0.052 (0,056)	-0.364*** (0,091)	-0.259** (0,105)	0.016 (0,027)	-0.03 (0,066)	-0.025 (0,040)	0.069* (0,041)
Male	-0.665*** (0,021)	-0.710*** (0,040)	-0.679*** (0,021)	-0.622*** (0,041)	-0.349*** (0,061)	-0.105* (0,060)	-0.686*** (0,019)	-0.668*** (0,041)	-0.669*** (0,029)	-0.712*** (0,030)
Mother education			0.675*** (0,009)	0.635*** (0,013)	0.456*** (0,034)	0.584*** (0,043)	0.684*** (0,007)	0.577*** (0,019)	0.626*** (0,011)	0.745*** (0,011)
Father education	0.726*** (0,010)	0.697*** (0,012)			0.329*** (0,034)	0.575*** (0,050)	0.738*** (0,008)	0.702*** (0,021)	0.702*** (0,012)	0.734*** (0,012)
Mother edu unknown			-0.736*** (0,033)	-0.487*** (0,060)	-0.756*** (0,032)	-0.927*** (0,048)		-0.638*** (0,076)		-0.573*** (0,037)
Father edu unknown	-0.469*** (0,021)	-0.125*** (0,036)			-0.462*** (0,031)	-0.326*** (0,044)	-0.552*** (0,023)	-0.117*** (0,037)	-0.539*** (0,031)	-0.516*** (0,030)
1 sibling	0.285*** (0,010)	0.334*** (0,018)	0.312*** (0,010)	0.328*** (0,021)	-0.037*** (0,046)	0.198*** (0,065)	0.308*** (0,009)			
2 siblings	0.261*** (0,011)	0.406*** (0,019)	0.303*** (0,011)	0.381*** (0,021)	-0.220*** (0,049)	0.162** (0,070)	0.314*** (0,010)			
3+ siblings	0.015 (0,014)	0.303*** (0,023)	0.086*** (0,013)	0.208*** (0,026)	-0.471*** (0,047)	0.008 (0,083)	0.117*** (0,012)			
Siblings unknown	-0.968** (0,024)	-0.744*** (0,042)	-0.953*** (0,025)	-0.948*** (0,045)	-1.377*** (0,047)	-0.788*** (0,065)	-1.111*** (0,026)			
Western	-0.022 (0,029)	-0.273*** (0,035)	-0.077** (0,032)	-0.283*** (0,039)				-0.15*** (0,057)	-0.167*** (0,037)	-0.225*** (0,037)
Non-Western	-0.393*** (0,018)	-0.793*** (0,030)	-0.375*** (0,020)	-0.870*** (0,028)				-0.211*** (0,043)	-0.502*** (0,026)	-0.703*** (0,022)
N	361083	181976	380701	152136	64447	56219	512415	73547	224995	235832
Pseudo R ²	0.043	0.051	0.045	0.056	0.213	0.180	0.071	0.050	0.059	0.079

Table 8: Subsample analysis. Binary logit regression analysis divided into parents' education level, immigration status and the number of siblings. *, ** and *** denote significance at 10 percent, 5 percent and 1 percent level respectively. Robust standard errors in parentheses.

We find the same issue regarding missing data for the number of siblings' variables. It is difficult to discuss the output from the subsamples in Model 14, as there is a substantial amount of missing data related to number of siblings for immigrated students. The β_3 's show only positive or statistical non-significant effects and does not cover the actual negative change in the gender gap. This negative effect is captured in the subsample for number of siblings unknown⁴⁶. For number of siblings' unknown, as parents' education level, the share of missing data is highest for immigrated students; 63%-90% compared to 2% for Norwegian students.

The most interesting finding so far is found in Model 13, where the effect from Reform 97 on the high school graduation gender gap captured by the model changes substantially across immigration segments. Starting by studying the immediate effect from the reform, β_3 1991*male indicates that the effect from the reform differed across the segments. The reform seems to have had a positive effect; decreasing the gender gap, for the Norwegian students, while there is no immediate effect on Western immigrants. The really interesting finding is that non-Western immigrants have a highly significant negative effect; indicating an increase, on the gender gap. Looking at all β_3 for the different subsamples, we find that Norwegian students experienced a general decrease in the high school graduation gender gap for all periods (all the β coefficients are positive), while non-Western students experienced the opposite: The high school graduation gender gap increased in all periods from year 1987 to 1995 (all β coefficients are negative). The effect is more difficult to conclude for Western immigrants, as the effect differs over time. We find a negative effect over time, though not all coefficients are statistically significant, compared to Norwegian students, but not as strong the effect we find for non-Western immigrants.

For Reform 97 to have caused the increase in high school graduation gender gap, we need to find no significant effect between year 1987 and 1990, as well as negative and significant effects in years 1991-1995. These effects were not captured in any of the subsamples in our model. This means that the negative change in the gender gap does not seem to have been caused by the reform.

⁴⁶ Subsample regression output for number of siblings' unknown is found in Appendix 7.

8. Conclusion

In 1997 the Norwegian government implemented a reform, Reform 97, that reduced the school starting age for children from seven to six years old. We use the implementation of this reform to investigate if the age when starting school affects the high school graduation gender gap. Children born in 1991 or later were affected by the reform, as they were six years old when the reform was implemented, forcing them to attend elementary school for ten years compared to nine before. Using a set of comprehensive, national register data from Microdata, we estimate a difference-in-difference estimate, using a binary logit regression model, which explores the differences in high school graduation gender gap due to the introduction of this reform. If the reform affected the gender gap, we expect this to show in the difference-in-difference estimates for all students born after 1991.

In our main analysis we find no clear patterns indicating that the reform had any impact on the gender gap in the high school graduation rate. We find, however, a clear increase in the gender gap in year 1991 and beyond. Our regression analysis show that male students generally have a lower probability of graduating from high school than female students and that there is a negative time trend in the graduation rate, but that this trend seems to be explained by other factors. In the regression there were positive effects from parents' higher education level as well as number of siblings, while having Western or non-Western immigration status gave a negative effect.

When dividing the sample into subsamples, we find a pattern indicating that immigration status has a higher explanatory power on the increase in the gender gap, than the reform. We see that the gender gap for Norwegian students are closing, while immigrated students' gender gap is increasing, especially for non-Western immigrants. This negative effect for non-Western students is significant throughout our data set and implies that this trend was happening independent on Reform 97.

This paper attempts to find an effect from Reform 97 on the high school graduation gender gap but fails to find any significant effect. The increase in the gender gap has to be explained by other factors. Our results can contribute to illuminating immigrated, especially male, students' accomplishments in Norwegian schooling. A natural extension of our work is to investigate the explaining factor on the gender gap increase for immigrants, non-Western immigrants in particular.

References

- Blundell, R., & Costa, M. D. (2009). *Alternative Approaches to Evaluation in Empirical Microeconomics*. Retrieved from University College London: <https://www.ucl.ac.uk/~uctp39a/Blundell-CostaDias-27-09-2008.pdf>
- Borjas, G. J. (2016). *Labor Economics*. New York: McGraw-Hill Education.
- Byrnes, J. P., Miller, D. C., & William, S. D. (1999). *Gender Differences in Risk Taking: A Meta Analysis*. Retrieved from American Psychological Association: <https://psycnet.apa.org/fulltext/1999-13573-004.html>
- Drange, N., Havnes, T., & Sandsør, A. M. (2012, June). *Kindergarten for all: Long run effects of a universal intervention*. Retrieved from Statistic Norway: <https://www.ssb.no/a/publikasjoner/pdf/DP/dp695.pdf>
- Feinglod, A. (1994). *Gender Differences in Personality: A Meta Analysis*. Retrieved from American Psychological Association: <https://psycnet.apa.org/fulltext/1999-13573-004.html>
- Ghauri, P., & Grønhaug, K. (2010). *Research Methods in Business Studies (4. utg.)*. New York: Financial Times Prentice Hall.
- Gurian, M., & Stevens, K. (2010). *Boys and girls learn differently! A guide for teachers and parents*. John Wiley & Sons.
- Haug, P. (2003, september). *Sluttrapport frå styret for Program for evaluering av Reform 97*. Retrieved from Norges forskningsråd.no: <https://www.forskningsradet.no/om-forskningsradet/publikasjoner/2003/evaluering-av-reform-972/>
- Johannessen, A., Tufte, P. A., & Christoffersen, L. (2010). *Introduksjon til samfunnsvitenskapelig metode (4. utg.)*. Oslo: Abstrakt forlag.
- Kirke-, utdannings- og forskningsdepartementet. (1996, November 14). *Reform 97- Dette er grunnskolereformen*. Retrieved from Regjeringen.no: <https://www.regjeringen.no/no/dokumentarkiv/regjeringen-brundtland-iii/kuf/veiledninger/1996/reform-97-dette-er-grunnskolereformen/id87403/>
- Kirke-, utdnings- og forskningsdepartementet. (1993). *vi smaa, en Alen lange*;
- Kjærnsli, M., & Jensen, F. (2016, December 6). *PISA 2015: Stø kurs*. Retrieved from Utdanningsdirektoratet: <https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/pisa-2015/>
- Kjærnsli, M., & Roe, A. (2010). *På rett spor. Norske elevers kompetanse i lesing, matematikk og naturfag i PISA 2009*. Oslo: Universitetet i Oslo ILS.
- Kjærnsli, M., Lie, S., Olsen, R. V., Roe, A., & Turmo, A. (2004). *Resultater fra PISA 2003*. Oslo: Universitetet i Oslo ILS, Utdanningsdirektoratet.
- Lønnå, E. (2017, November 28). *Kvinnens rettigheter i Norge fra 1945 til 1990-årene*. Retrieved from Store norske leksikon: https://snl.no/Kvinnens_rettigheter_i_Norge_fra_1945_til_1990-%C3%A5rene?fbclid=IwAR1V86NGXbr3UzN7JIEOGthfbuI7PMc3zcRAwxk7MYoY4dqGZOpO-NDhMQ
- Lønnå, E. (2018, May 25). *Kvinnens rettigheter i Norge fra 1945 til 1990-årene*. Retrieved from Store Norske Leksikon: https://snl.no/Kvinnens_rettigheter_i_Norge_fra_1945_til_1990-%C3%A5rene
- Magon, A. J. (2009, April). *Gender, the Brain and Education: Do Boys and Girls Learn Differently?* Retrieved from https://dspace.library.uvic.ca/bitstream/handle/1828/1411/magon_project_signatures_removed.pdf?sequence=1&isAllowed=y
- Microdata. (2017). *Variabeloversikt*. Retrieved from Microdata.no: <https://microdata.no/discovery>

- Midtbø, T. (2007). *Regresjonsanalyse for samfunnsvitere, med eksempler i SPSS*. Oslo: Universitetsforlaget.
- Mitra, D., & Zheng, A. (2017). *Pennsylvania's Best Investment: The Social and Economic Benefits of Public Education*. Pennsylvania: Education Law Center.
- Pontiell, J. A. (2003). Enhancers and inhibitors of teacher risk taking: A case study. *Peabody Journal of Education*, 5-24.
- Regjeringen.no. (2019, April 23). *Kunnskapsløftet*. Retrieved from Regjeringen.no: https://www.regjeringen.no/no/tema/utdanning/grunnopplaring/kunnskapsloftet/id534689/?fbclid=IwAR3Gtq9c6FkrSIRrekJgh_7ZgHGkEVM9a9SKp2G47tyoi1krhNOSsyYVjaI
- Roe, A., Solheim, R., & Kjærnsli, M. (2007, November 29). *PISA 2006: Svakere resultater i alle fag*. Retrieved from Utdanningsdirektoratet: <https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/PISA-2006-Svakere-resultater-i-alle-fag/>
- Statistics Norway. (2011, December 1). *Standard for land og statsborgerskap i personstatistikk*. Retrieved from Statistics Norway: <https://www.ssb.no/klass/klassifikasjoner/91/varianter/847>
- Statistics Norway. (2016, June 2). *Gjennomføring i videregående opplæring, 2010-2015*. Retrieved from Statistics Norway: <https://www.ssb.no/utdanning/statistikker/vgogjen/aar/2016-06-02>
- Statistics Norway. (2018, May 29). *Gjennomføring i videregående opplæring*. Retrieved from Statistics Norway: <https://www.ssb.no/utdanning/statistikker/vgogjen/aar>
- Statistics Norway. (2019, January 8). *Utdanningsgapet bare øker*. Retrieved from Statistics Norway: <https://www.ssb.no/befolkning/artikler-og-publikasjoner/utdanningsgapet-bare-oket>
- Thrane, C. (2003). *Regresjonsanalyse i praksis*. Kristiansand : Høyskoleforlaget AS .
- Universitetet i Oslo ILS, Utdanningsdirektoratet. (2004, December 6). *PISA 2003 - Rett spor eller ville veier?* Retrieved from Utdanningsdirektoratet: <https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/PISA-2003---resultatene-2004/>
- Utdannings- og forskningsdepartementet. (2005). *Kunnskapsløftet – reformen i grunnskole og videregående opplæring*. Utdannings- og forskningsdepartementet.
- Utdanningsdirektoratet. (2007, November 29). *PISA 2006: Svakere resultater i alle fag*. Retrieved from Udir: <https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/PISA-2006-Svakere-resultater-i-alle-fag/>
- Utdanningsdirektoratet. (2016, September 16). *PISA*. Retrieved from Utdanningsdirektoratet: <https://www.udir.no/tall-og-forskning/internasjonale-studier/pisa/>
- World Economic Forum. (2018). *The Global Gender Gap Report*. Geneva: World Economic Forum. Retrieved from http://www3.weforum.org/docs/WEF_GGGR_2018.pdf

Appendix

Appendix 1: Microdata

Microdata.no has been developed in collaboration between the Norwegian Centre for Research Data (NSD) and Statistics Norway (SSB). The service provides researchers and students with approved research institutions access to use register data from SSB. Microdata has available data on population, education, labor market, and social security. The database consists of individual demographic variables like the year of birth, number of siblings, immigration status and socioeconomic variables like education. The individual data have information on the Norwegian population from 1900 to 2016. All data are anonymizing, and privacy policy is taken care of.

This dataset was chosen for this paper as register data are raw and individual data that is a good basis for further research. It contains interesting and relevant variables collected from the whole population as year of birth, gender, immigration background, number of siblings and parents' educational level. In addition, the data set is extensive and quantitative, including a causal research design. These criteria for a dataset are crucial for answering this paper's research question.

We found Microdata as a well-functioning dataset, consisting of many different variables that we found important for this papers' research question. The dataset was very much user friendly as it was quite similar to other more known programs as STATA and SPSS. The people working at Microdata were always very helpful if we sometimes ran into some challenges.

However, the data set had some drawbacks as well. When exporting our worksheet to excel, negative numbers did not show. We had to make a detour, copying the worksheet to google sheets (googles answer to excel), and from there export it to excel. Another thing about microdata was that due to the privacy policy it was not possible to investigate any individual object in the sample. Doing the different commands, it was not possible to be sure that we got the result we wanted, you just had to trust that the command did the intended job. Another problem we faced with Microdata was for regression with subsamples. By including the whole model, some variables should not vary in a subsample regression (as they are either 1; the subsample, or 0; outside the subsample. i.e. the number of siblings' variables do not vary

when regressing only on only children). We also encountered a problem where there was not enough variation in the variable to give an output on a regression: mother's education level unknown did not vary for Norwegian students or for students with 1 sibling, this gave a technical error instead of giving output with N.A.

Appendix 2: PISA

In year 2000, Norwegian 15-year olds were tested, for the first time, in an international school test called PISA (Program for International Student Assessment). The test is completed through a cooperation between the 36 countries that make up the OECD; an organisation working together for economic cooperation and development. The students are tested in mathematics, science and reading.

The results from year 2003, showed that the Norwegian students scored below the results of other Nordic countries, especially in mathematics and science⁴⁷. In 2006, Norwegian students' results were disconcerting, scoring below the mean in every tested subject. Something had to be done with the Norwegian school-system and a reform called "Kunnskapsløftet" was introduced in 2006. This reform implied changes in the curriculum in all classes, from the first class in elementary school to the last year of high school. The main goal was improving learning outcomes for all students⁴⁸.

The new curriculum had a stronger focus on learning outcomes and had clear goals for the student's competences. Reading and writing as well as other basic skills, in all subjects, was important from the first year in school⁴⁹. The new curriculum resulted in brighter results on the PISA tests, where Norwegian students scored over the OECD-mean in all subjects in 2009⁵⁰.

Reform 97 is somehow replaced by Kunnskapsløftet 2006 and its new curriculum. The desire to improve the Norwegian school results may have made the school more theoretical, possibly a better fit for females than for males⁵¹. There may be reason to believe that the reform from 2006 could influence our results regarding Reform 97.

⁴⁷ (Kjærnsli, Lie, Olsen, Roe, & Turmo, 2004)

⁴⁸ (Regjeringen.no, 2019)

⁴⁹ (Utdannings- og forskningsdepartementet, 2005)

⁵⁰ (Kjærnsli & Roe, På rett spor. Norske elevers kompetanse i lesing, matematikk og naturfag i PISA 2009, 2010)

⁵¹ (Haug, 2003)

Appendix 3: Descriptive information

endeligssummarize male mother_education father_education søsken_0 søsken_1 søsken_2 søsken_3_5 if No rwegian ==1 & etter_1991 ==0

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
male	0.5129	0.4998	221503	0	1	0	1	1
mother_education	0.3093	0.4622	221503	0	1	0	0	1
father_education	0.268	0.4429	221503	0	1	0	0	1
søsken_0	0.1596	0.3662	221503	0	1	0	0	0
søsken_1	0.4101	0.4919	221503	0	1	0	0	1
søsken_2	0.3023	0.4593	221503	0	1	0	0	1
søsken_3_5	0.1066	0.3086	221503	0	1	0	0	0

endeligssummarize male mother_education father_education søsken_0 søsken_1 søsken_2 søsken_3_5 if No rwegian ==1 & etter_1991 ==1

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
male	0.5159	0.4997	290910	0	1	0	1	1
mother_education	0.3532	0.478	290910	0	1	0	0	1
father_education	0.2823	0.4501	290910	0	1	0	0	1
søsken_0	0.1161	0.3203	290910	0	1	0	0	0
søsken_1	0.4239	0.4942	290910	0	1	0	0	1
søsken_2	0.3325	0.4711	290910	0	1	0	0	1
søsken_3_5	0.1139	0.3176	290910	0	1	0	0	0

endeligssummarize male mother_education father_education søsken_0 søsken_1 søsken_2 søsken_3_5 if Immigrant_1==1 & etter_1991 ==0

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
male	0.5227	0.4995	34354	0	1	0	1	1
mother_education	0.0472	0.2122	34354	0	1	0	0	0
father_education	0.0377	0.1904	34354	0	1	0	0	0
søsken_0	0.021	0.1432	34354	0	1	0	0	0
søsken_1	0.0429	0.2026	34354	0	1	0	0	0
søsken_2	0.0284	0.1661	34354	0	1	0	0	0
søsken_3_5	0.0131	0.1138	34354	0	1	0	0	0

endeligssummarize male mother_education father_education søsken_0 søsken_1 søsken_2 søsken_3_5 if Immigrant_1==1 & etter_1991 ==1

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
male	0.4999	0.5	21862	0	1	0	0	1
mother_education	0.1428	0.3499	21862	0	1	0	0	0
father_education	0.1134	0.3171	21862	0	1	0	0	0
søsken_0	0.0351	0.184	21862	0	1	0	0	0
søsken_1	0.1028	0.3037	21862	0	1	0	0	0
søsken_2	0.0701	0.2553	21862	0	1	0	0	0
søsken_3_5	0.0322	0.1767	21862	0	1	0	0	0

endeligssummarize male mother_education father_education søsken_0 søsken_1 søsken_2 søsken_3_5 if Immigrant_2==1 & etter_1991 ==0

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
male	0.4596	0.4984	34751	0	1	0	0	1
mother_education	0.0713	0.2574	34751	0	1	0	0	0
father_education	0.0841	0.2775	34751	0	1	0	0	0
søsken_0	0.045	0.2072	34751	0	1	0	0	0
søsken_1	0.0964	0.2952	34751	0	1	0	0	0
søsken_2	0.0703	0.2557	34751	0	1	0	0	0
søsken_3_5	0.0954	0.2937	34751	0	1	0	0	0

endeligssummarize male mother_education father_education søsken_0 søsken_1 søsken_2 søsken_3_5 if Immigrant_2==1 & etter_1991 ==1

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
male	0.5289	0.4992	29697	0	1	0	1	1
mother_education	0.1175	0.3221	29697	0	1	0	0	0
father_education	0.1332	0.3398	29697	0	1	0	0	0
søsken_0	0.0466	0.2109	29697	0	1	0	0	0
søsken_1	0.1264	0.3323	29697	0	1	0	0	0
søsken_2	0.0824	0.275	29697	0	1	0	0	0
søsken_3_5	0.1196	0.3245	29697	0	1	0	0	0

- **endeligssummarize fullført_innen_20**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
fullført innen 20	0.559	0.4965	633084	0	1	0	1	1

- **endeligssummarize fullført_innen_20 if etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
fullført_innen_20	0.5227	0.4995	290604	0	1	0	1	1

- **endeligssummarize fullført_innen_20 if etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
fullført_innen_20	0.5897	0.4919	342482	0	1	0	1	1

- **endeligssummarize søsken_ukjent**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
søsken_ukjent	0.1559	0.3628	633084	0	1	0	0	0

- **endeligssummarize søsken_ukjent if etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
søsken_ukjent	0.2049	0.4036	290604	0	1	0	0	0

- **endeligssummarize søsken_ukjent if etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
søsken_ukjent	0.1143	0.3182	342482	0	1	0	0	0

- **endeligssummarize utdanning_mor_ukjent**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_mor_ukjent	0.1422	0.3493	633084	0	1	0	0	0

- **endeligssummarize utdanning_mor_ukjent if etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_mor_ukjent	0.1965	0.3973	290604	0	1	0	0	0

- **endeligssummarize utdanning_mor_ukjent if etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_mor_ukjent	0.0962	0.2948	342482	0	1	0	0	0

- **endeligssummarize utdanning_far_ukjent**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_far_ukjent	0.1583	0.3651	633084	0	1	0	0	0

- **endeligssummarize utdanning_far_ukjent if etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_far_ukjent	0.2114	0.4083	290604	0	1	0	0	0

- **endeligssummarize utdanning_far_ukjent if etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_far_ukjent	0.1133	0.3169	342482	0	1	0	0	0

- endeligssummarize fullført_innen_20 if Norwegian ==1 & etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
fullført innen 20	0.6302	0.4827	221503	0	1	0	1	1
- endeligssummarize fullført_innen_20 if Norwegian ==1 & etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
fullført_innen_20	0.6488	0.4773	290910	0	1	0	1	1
- endeligssummarize fullført_innen_20 if Immigrant_1==1 & etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
fullført_innen_20	0.1701	0.3757	34354	0	1	0	0	0
- endeligssummarize fullført_innen_20 if Immigrant_1==1 & etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
fullført_innen_20	0.2596	0.4384	21862	0	1	0	0	1
- endeligssummarize fullført_innen_20 if Immigrant_2==1 & etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
fullført innen 20	0.1863	0.3894	34751	0	1	0	0	0
- endeligssummarize fullført_innen_20 if Immigrant_2==1 & etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
fullført_innen_20	0.2538	0.4352	29697	0	1	0	0	1

- **endeligssummarize søsken_ukjent if Norwegian ==1 & etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
søsken_ukjent	0.0214	0.1447	221503	0	1	0	0	0
- **endeligssummarize søsken_ukjent if Norwegian ==1 & etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
søsken_ukjent	0.0137	0.1162	290910	0	1	0	0	0
- **endeligssummarize søsken_ukjent if Immigrant_1 ==1 & etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
søsken_ukjent	0.8946	0.307	34354	0	1	1	1	1
- **endeligssummarize søsken_ukjent if Immigrant_1 ==1 & etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
søsken_ukjent	0.7597	0.4272	21862	0	1	1	1	1
- **endeligssummarize søsken_ukjent if Immigrant_2 ==1 & etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
søsken_ukjent	0.6929	0.4613	34751	0	1	0	1	1
- **endeligssummarize søsken_ukjent if Immigrant_2 ==1 & etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
søsken_ukjent	0.6249	0.4842	29697	0	1	0	1	1

- **endeligssummarize utdanning_mor_ukjent if Norwegian ==1 & etter_1991 ==0**

Teknisk feil

- **endeligssummarize utdanning_mor_ukjent if Norwegian ==1 & etter_1991 ==1**

Teknisk feil

- **endeligssummarize utdanning_mor_ukjent if Immigrant_1 ==1 & etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_mor_ukjent	0.891	0.3116	34354	0	1	1	1	1

- **endeligssummarize utdanning_mor_ukjent if Immigrant_1 ==1 & etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_mor_ukjent	0.672	0.4695	21862	0	1	0	1	1

- **endeligssummarize utdanning_mor_ukjent if Immigrant_2 ==1 & etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_mor_ukjent	0.7271	0.4455	34751	0	1	0	1	1

- **endeligssummarize utdanning_mor_ukjent if Immigrant_2 ==1 & etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_mor_ukjent	0.5586	0.4966	29697	0	1	0	1	1

- **endeligssummarize utdanning_far_ukjent if Norwegian ==1 & etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_far_ukjent	0.0179	0.1325	221503	0	1	0	0	0

- **endeligssummarize utdanning_far_ukjent if Norwegian ==1 & etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_far_ukjent	0.0157	0.1243	290910	0	1	0	0	0

- **endeligssummarize utdanning_far_ukjent if Immigrant_1 ==1 & etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_far_ukjent	0.9197	0.2718	34354	0	1	1	1	1

- **endeligssummarize utdanning_far_ukjent if Immigrant_1 ==1 & etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_far_ukjent	0.7306	0.4437	21862	0	1	0	1	1

- **endeligssummarize utdanning_far_ukjent if Immigrant_2 ==1 & etter_1991 ==0**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_far_ukjent	0.7451	0.4358	34751	0	1	0	1	1

- **endeligssummarize utdanning_far_ukjent if Immigrant_2 ==1 & etter_1991 ==1**

Variabel	Gj.snitt	Std.avvik	Antall	Min	Maks	25%	50%	75%
utdanning_far_ukjent	0.6147	0.4867	29697	0	1	0	1	1

Appendix 4: Binary logit regression output

Model 1

- endeliglogit fullført_innen_20 etter_1991 male male_etter**

Antall iter4
 Log sans-4.2869182e+5
 Antall obs633081
 LR chi2(4)11425.1
 Prob > chi20
 Pseudo R20.01315

fullført_innen_20	Coef.	Std.feil	z	P> z 	[95% Konf. intervall]
etter_1991	0.34068	0.00743	45.8293	0	0.32611 0.35525
male	-0.40605	0.00746	-54.3705	0	-0.42069 -0.39141
male_etter	-0.11675	0.01025	-11.3851	4.96032e-30	-0.13685 -0.09665
Konst	0.29812	0.00534	55.7617	0	0.28764 0.3086

Model 2

- endeliglogit fullført_innen_20 etter_1991 male male_etter mother_education father_education utdanning_far_ukjent utdanning_mor_ukjent**

Antall iter6
 Log sans-3.6574881e+5
 Antall obs633081
 LR chi2(8)1.3731112e+5
 Prob > chi20
 Pseudo R20.15804

fullført_innen_20	Coef.	Std.feil	z	P> z 	[95% Konf. intervall]
etter_1991	0.06432	0.00837	7.678	1.61582e-14	0.0479 0.08075
male	-0.5688	0.00848	-66.9987	0	-0.58544 -0.55216
male_etter	-0.04916	0.0114	-4.3108	0.00001	-0.07151 -0.02681
mother_education	0.69605	0.00702	99.0425	0	0.68228 0.70983
father_education	0.69853	0.0076	91.894	0	0.68363 0.71343
utdanning_far_ukjent	-0.88535	0.01425	-62.1028	0	-0.91329 -0.85741
utdanning_mor_ukjent	-1.51701	0.01607	-94.3534	0	-1.54853 -1.4855
Konst	0.47372	0.00659	71.8835	0	0.46081 0.48664

Model 3

- **endeliglogit fullført_innen_20 etter_1991 male male_etter søsken_1 søsken_2 søsken_3_5 søsken_ukjent**

Antall iter6
 Log sans-3.8086992e+5
 Antall obs633081
 LR chi2(8)1.070689e+5
 Prob > chi20
 Pseudo R20.12323

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
etter_1991	0.1144	0.00819	13.9537	2.98466e-44	0.09833 0.13047
male	-0.54222	0.00824	-65.769	0	-0.55838 -0.52606
male_etter	-0.0471	0.01112	-4.23458	0.00002	-0.06891 -0.0253
søsken_1	0.35849	0.00876	40.9131	0	0.34131 0.37566
søsken_2	0.41217	0.00914	45.0851	0	0.39425 0.43008
søsken_3_5	0.09235	0.01104	8.36192	6.17016e-17	0.0707 0.11399
søsken_ukjent	-2.21722	0.01216	-182.243	0	-2.24106 -2.19337
Konst	0.51657	0.00911	56.6443	0	0.4987 0.53445

Model 4

- **endeliglogit fullført_innen_20 etter_1991 male male_etter Immigrant_1 Immigrant_2**

Antall iter5
 Log sans-3.9149321e+5
 Antall obs633081
 LR chi2(6)85822.3
 Prob > chi20
 Pseudo R20.09878

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
etter_1991	0.18551	0.00799	23.2177	3.01299e-119	0.16985 0.20117
male	-0.49913	0.00804	-62.063	0	-0.51489 -0.48336
male_etter	-0.08174	0.01092	-7.48338	7.24301e-14	-0.10315 -0.06033
Immigrant_1	-1.94407	0.01096	-177.281	0	-1.96557 -1.92258
Immigrant_2	-1.88771	0.01008	-187.117	0	-1.90748 -1.86794
Konst	0.76565	0.00602	127.025	0	0.75384 0.77747

Model 5

- **endeliglogit fullført_innen_20 etter_1991 male male_etter mother_education father_education utdanning_far_ukjent utdanning_mor_ukjent søsken_1 søsken_2 søsken_3_5søsken_ukjent Immigrant_1 Immigrant_2**

Antall iter6
 Log sans-3.6137711e+5
 Antall obs633081
 LR chi2(14)1.4605453e+5
 Prob > chi20
 Pseudo R20.1681

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
etter_1991	0.05678	0.00846	6.71138	1.9279e-11	0.04019 0.07336

fullført_innen_20	Coef.	Std.feil	z	P> z 	[95% Konf.	intervall]
male	-0.58316	0.00854	-68.2209	0	-0.59992	-0.56641
male_etter	-0.04208	0.01149	-3.66277	0.00024	-0.0646	-0.01956
mother_education	0.6723	0.00707	94.9743	0	0.65843	0.68618
father_education	0.71072	0.00767	92.6305	0	0.69568	0.72576
utdanning_far_ukjent	-0.38393	0.01618	-23.7202	2.22693e-124	-0.41566	-0.35221
utdanning_mor_ukjent	-0.70883	0.01965	-36.0561	1.10617e-284	-0.74736	-0.6703
søsken_1	0.29509	0.00903	32.6658	4.76692e-234	0.27738	0.3128
søsken_2	0.29691	0.00944	31.4305	7.74699e-217	0.2784	0.31543
søsken_3_5	0.09468	0.0115	8.22796	1.90421e-16	0.07213	0.11723
søsken_ukjent	-0.95663	0.01848	-51.7402	0	-0.99287	-0.9204
Immigrant_1	-0.04958	0.01776	-2.79082	0.00525	-0.0844	-0.01476
Immigrant_2	-0.50726	0.01388	-36.5208	5.17833e-292	-0.53448	-0.48004
Konst	0.29817	0.00959	31.0848	3.85878e-212	0.27937	0.31697

Appendix 5: Short binary logit output, subsample on immigration status

Model 11

endeliglogit fullført_innen_20 etter_1991 male male_etter if Norwegian ==1

Antall iter5
 Log sans-3.2951576e+5
 Antall obs512415
 LR chi2(4)10150
 Prob > chi20
 Pseudo R20.01516

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
etter_1991	0.09062	0.00889	10.1918	2.15665e-24	0.07319 0.10804
male	-0.57984	0.00893	-64.8895	0	-0.59735 -0.56232
male_etter	-0.01164	0.01192	-0.97598	0.32907	-0.03502 0.01173
Konst	0.84156	0.00663	126.825	0	0.82856 0.85457

endeliglogit fullført_innen_20 etter_1991 male male_etter if Immigrant_1 ==1

Antall iter6
 Log sans-28150.1
 Antall obs56219
 LR chi2(4)715.946
 Prob > chi27.31265e-155
 Pseudo R20.01255

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
etter_1991	0.56326	0.02942	19.1408	1.15321e-81	0.50559 0.62094
male	-0.14565	0.02873	-5.06941	3.99032e-7	-0.20196 -0.08933
male_etter	-0.06121	0.04221	-1.45015	0.14701	-0.14394 0.02151
Konst	-1.51068	0.02029	-74.4513	0	-1.55045 -1.47091

endeliglogit fullført_innen_20 etter_1991 male male_etter if Immigrant_2 ==1

Antall iter5
 Log sans-33262.5
 Antall obs64447
 LR chi2(4)969.152
 Prob > chi28.84598e-210
 Pseudo R20.01435

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
etter_1991	0.60029	0.02578	23.2763	7.69445e-120	0.54974 0.65083
male	-0.19785	0.02785	-7.10234	1.22657e-12	-0.25245 -0.14325
male_etter	-0.3965	0.03879	-10.2214	1.58954e-24	-0.47253 -0.32047
Konst	-1.38602	0.01824	-75.9745	0	-1.42178 -1.35027

Appendix 6: Full binary logit regression output

Model 6

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 male

Antall iter: 5

Log sans: -4.2796492e+5

Antall obs: 633081

LR chi2(18): 12878.9

Prob > chi20

Pseudo R2: 0.01482

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf.	intervall]
født_1988	0.05797	0.01508	3.8432	0.00012	0.02841	0.08754
født_1989	0.16392	0.01516	10.8117	3.02819e-27	0.1342	0.19363
født_1990	0.20657	0.01518	13.6072	3.627e-42	0.17682	0.23633
født_1991	0.29847	0.01539	19.3817	1.10134e-83	0.26828	0.32865
født_1992	0.34847	0.01559	22.3424	1.43129e-110	0.3179	0.37903
født_1993	0.40058	0.01574	25.4473	7.55733e-143	0.36972	0.43143
født_1994	0.57568	0.01603	35.9107	2.07856e-282	0.54426	0.6071
født_1995	0.64583	0.01618	39.9077	0	0.61411	0.67755
male_1988	0.02662	0.02118	1.25702	0.20874	-0.01488	0.06814
male_1989	0.00844	0.02121	0.39807	0.69057	-0.03312	0.05001
male_1990	0.01067	0.02122	0.50271	0.61516	-0.03093	0.05228
male_1991	-0.04105	0.02143	-1.91542	0.05543	-0.08306	0.00095
male_1992	-0.05918	0.02161	-2.7379	0.00618	-0.10154	-0.01681
male_1993	-0.10591	0.02179	-4.85949	0	-0.14863	-0.06319
male_1994	-0.16517	0.02205	-7.49063	6.85424e-14	-0.20839	-0.12195
male_1995	-0.17063	0.02217	-7.69447	1.42072e-14	-0.21409	-0.12716
male	-0.41864	0.01508	-27.75	1.74066e-169	-0.44821	-0.38907
Konst	0.1909	0.01071	17.8091	5.99701e-71	0.16989	0.21191

Model 7

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 mother_education father_education utdanning_mor_ukjent utdanning_far_ukjent

Antall iter: 6

Log sans: -3.6563518e+5

Antall obs: 633081

LR chi2(22): 1.3753838e+5

Prob > chi20

Pseudo R2: 0.1583

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf.	intervall]
født_1988	0.01256	0.01748	0.71853	0.47243	-0.02171	0.04684
født_1989	0.06073	0.0175	3.47017	0.00052	0.02643	0.09503
født_1990	0.0328	0.01742	1.88218	0.05981	-0.00135	0.06696
født_1991	0.04931	0.01755	2.80995	0.00495	0.01491	0.08372
født_1992	0.04117	0.01768	2.32768	0.01992	0.0065	0.07583
født_1993	0.02392	0.01772	1.34942	0.1772	-0.01082	0.05866
født_1994	0.16442	0.01796	9.15107	5.63698e-20	0.12921	0.19964
født_1995	0.18715	0.01803	10.3762	3.17942e-25	0.1518	0.2225
male_1988	0.0138	0.02417	0.57093	0.56804	-0.03357	0.06118
male_1989	0.00766	0.02411	0.31791	0.75054	-0.0396	0.05493
male_1990	0.02272	0.024	0.94667	0.3438	-0.02432	0.06977
male_1991	-0.01298	0.0241	-0.53877	0.59004	-0.06022	0.03425
male_1992	-0.00249	0.02421	-0.10289	0.91804	-0.04995	0.04496
male_1993	-0.0172	0.02429	-0.70815	0.47884	-0.06481	0.0304
male_1994	-0.08788	0.02446	-3.59168	0.00032	-0.13584	-0.03992
male_1995	-0.07575	0.0245	-3.09075	0.00199	-0.12378	-0.02771
male	-0.58	0.01729	-33.5339	1.54454e-246	-0.6139	-0.5461
mother_education	0.6942	0.00703	98.7376	0	0.68042	0.70798
father_education	0.69871	0.0076	91.9027	0	0.68381	0.71361
utdanning_mor_ukjent	-1.51218	0.01607	-94.049	0	-1.54369	-1.48067
utdanning_far_ukjent	-0.88467	0.01425	-62.0684	0	-0.91261	-0.85674
Konst	0.44652	0.01273	35.0729	1.74082e-269	0.42157	0.47147

Model 8

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 søsken_1 søsken_2 søsken_3_5 søsken_ukjent

Antall iter: 6

Log sans: -3.8064282e+5

Antall obs: 633081

LR chi2(22): 1.075231e+5

Prob > chi20

Pseudo R2: 0.12375

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	-0.01449	0.01703	-0.85079	0.39488	-0.04789 0.01889
født_1989	0.02544	0.01704	1.49269	0.13551	-0.00796 0.05886
født_1990	0.00772	0.01699	0.45482	0.64923	-0.02557 0.04103
født_1991	0.03522	0.01713	2.05535	0.03984	0.00163 0.0688
født_1992	0.05051	0.01729	2.92034	0.00349	0.01661 0.08441
født_1993	0.05018	0.01736	2.88945	0.00385	0.01614 0.08422
født_1994	0.21391	0.01765	12.1133	8.97141e-34	0.1793 0.24853
født_1995	0.26397	0.01777	14.8478	7.18567e-50	0.22912 0.29882
male_1988	0.03736	0.02348	1.59133	0.11153	-0.00865 0.08339
male_1989	0.03762	0.02342	1.6059	0.10829	-0.00829 0.08354
male_1990	0.04907	0.02334	2.10219	0.03553	0.00332 0.09482
male_1991	0.01933	0.02345	0.82449	0.40965	-0.02663 0.0653
male_1992	0.01699	0.02359	0.72026	0.47136	-0.02924 0.06323
male_1993	0.00692	0.0237	0.29226	0.77008	-0.03953 0.05338
male_1994	-0.07102	0.02393	-2.96706	0.003	-0.11794 -0.0241
male_1995	-0.06241	0.02403	-2.59631	0.00942	-0.10952 -0.01529
male	-0.57359	0.01681	-34.1124	4.82921e-255	-0.60654 -0.54063
søsken_1	0.35498	0.00877	40.4765	0	0.3378 0.37217
søsken_2	0.40901	0.00915	44.6912	0	0.39107 0.42695
søsken_3_5	0.0903	0.01105	8.17066	3.06704e-16	0.06864 0.11197
søsken_ukjent	-2.21418	0.01217	-181.908	0	-2.23803 -2.19032
Konst	0.51361	0.01392	36.8881	7.15238e-298	0.48632 0.5409

Model 9

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 Immigrant_1 Immigrant_2

Antall iter: 5

Log sans: -3.9119867e+5

Antall obs: 633081

LR chi2(20): 86411.4

Prob > chi20

Pseudo R2: 0.09945

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	0.02001	0.01639	1.22092	0.22211	-0.01211 0.05215
født_1989	0.09203	0.01645	5.59472	2.20966e-8	0.05979 0.12428
født_1990	0.09577	0.01643	5.82786	5.61417e-9	0.06356 0.12798
født_1991	0.14115	0.01661	8.4937	2.00155e-17	0.10858 0.17373
født_1992	0.17052	0.0168	10.1476	3.39267e-24	0.13759 0.20346
født_1993	0.17242	0.01689	10.2044	1.8943e-24	0.1393 0.20554
født_1994	0.33607	0.01717	19.5626	3.22078e-85	0.3024 0.36975
født_1995	0.38973	0.01731	22.5113	3.21663e-112	0.3558 0.42367
male_1988	0.02342	0.02282	1.02644	0.30468	-0.0213 0.06815
male_1989	0.00741	0.02281	0.3249	0.74525	-0.0373 0.05213
male_1990	0.0059	0.02277	0.25926	0.79543	-0.03873 0.05055
male_1991	-0.02657	0.02293	-1.15848	0.24666	-0.07153 0.01838
male_1992	-0.03925	0.0231	-1.69909	0.0893	-0.08453 0.00602
male_1993	-0.05634	0.02322	-2.42586	0.01527	-0.10186 -0.01082
male_1994	-0.13135	0.02345	-5.59932	2.15186e-8	-0.17733 -0.08537
male_1995	-0.12073	0.02357	-5.12162	3.02919e-7	-0.16693 -0.07452
male	-0.50863	0.01629	-31.2097	7.85343e-214	-0.54058 -0.47669
Immigrant_1	-1.93473	0.01097	-176.258	0	-1.95624 -1.91321
Immigrant_2	-1.88218	0.01009	-186.439	0	-1.90197 -1.86239
Konst	0.71183	0.01182	60.2212	0	0.68867 0.735

Model 10

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 mother_education father_education utdanning_mor_ukjent utdanning_far_ukjent søsken_1 søsken_2 søsken_3_5 søsken_ukjent Immigrant_1 Immigrant_2

Antall iter: 6

Log sans: -3.6125187e+5

Antall obs: 633081

LR chi2(28): 1.46305e+5

Prob > chi20

Pseudo R2: 0.16839

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	-0.00908	0.01765	-0.51467	0.60677	-0.04368 0.02551
født_1989	0.02546	0.01766	1.44188	0.14933	-0.00914 0.06008
født_1990	-0.00587	0.01759	-0.33397	0.73839	-0.04035 0.0286
født_1991	0.00927	0.01771	0.52369	0.60048	-0.02543 0.04398
født_1992	0.00613	0.01786	0.34361	0.73113	-0.02886 0.04114
født_1993	-0.01227	0.0179	-0.68532	0.49313	-0.04736 0.02282
født_1994	0.13699	0.01816	7.54121	4.65605e-14	0.10138 0.17259
født_1995	0.16914	0.01824	9.27057	1.85146e-20	0.13338 0.20491
male_1988	0.02196	0.02434	0.90222	0.36693	-0.02575 0.06967
male_1989	0.02249	0.02428	0.92626	0.35431	-0.0251 0.07008
male_1990	0.03912	0.02417	1.61833	0.10558	-0.00825 0.0865
male_1991	0.00553	0.02426	0.22821	0.81948	-0.04202 0.0531
male_1992	0.01273	0.02439	0.52204	0.60163	-0.03507 0.06054
male_1993	0.0035	0.02448	0.14297	0.8863	-0.04448 0.05148
male_1994	-0.07314	0.02467	-2.96383	0.00303	-0.12151 -0.02477
male_1995	-0.06118	0.02473	-2.47313	0.01339	-0.10967 -0.01269
male	-0.60429	0.01742	-34.6737	1.9592e-263	-0.63845 -0.57014
mother_education	0.6704	0.00708	94.667	0	0.65652 0.68428
father_education	0.71092	0.00767	92.6405	0	0.69588 0.72596
utdanning_mor_ukjent	-0.70202	0.01965	-35.7168	2.16306e-279	-0.74054 -0.66349
utdanning_far_ukjent	-0.38164	0.01618	-23.5835	5.68002e-123	-0.41335 -0.34992
søsken_1	0.293	0.00904	32.4099	1.98565e-230	0.27528 0.31072
søsken_2	0.29521	0.00945	31.2222	5.32364e-214	0.27668 0.31375
søsken_3_5	0.0934	0.01151	8.11148	5.00068e-16	0.07083 0.11596
søsken_ukjent	-0.96222	0.01849	-52.0195	0	-0.99847 -0.92596
Immigrant_1	-0.04971	0.01776	-2.79861	0.00513	-0.08453 -0.01489
Immigrant_2	-0.50871	0.01388	-36.6268	1.06874e-293	-0.53593 -0.48148
Konst	0.2969	0.01451	20.4598	4.90468e-93	0.26846 0.32534

Appendix 7: Full binary logit regression output, subsamples

Model 12.1

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 male_father_education utdanning_far_ukjent søsken_1 søsken_2 søsken_3_5 søsken_ukjent Immigrant_1 Immigrant_2 if mother_education_low ==1

Antall iter: 5
 Log sans: -2.365323e+5
 Antall obs: 361083
 LR chi2(26): 21311.2
 Prob > chi20
 Pseudo R2: 0.0431

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	-0.01223	0.02159	-0.56677	0.57086	-0.05456 0.03008
født_1989	0.00124	0.02155	0.05786	0.95385	-0.041 0.0435
født_1990	-0.02369	0.02141	-1.1061	0.26867	-0.06566 0.01828
født_1991	-0.01359	0.02147	-0.63291	0.52678	-0.05568 0.0285
født_1992	0.00071	0.02168	0.03319	0.97352	-0.04178 0.04322
født_1993	-0.02408	0.02169	-1.1099	0.26704	-0.06661 0.01844
født_1994	0.14615	0.02205	6.62798	3.40294e-11	0.10293 0.18937
født_1995	0.155	0.0221	7.01135	2.36025e-12	0.11167 0.19833
male_1988	0.05102	0.02968	1.71901	0.08561	-0.00715 0.1092
male_1989	0.0646	0.02955	2.18571	0.02883	0.00667 0.12253
male_1990	0.084	0.02935	2.8614	0.00421	0.02646 0.14154
male_1991	0.05302	0.02946	1.79957	0.07192	-0.00472 0.11077
male_1992	0.05492	0.02965	1.85239	0.06396	-0.00318 0.11304
male_1993	0.08494	0.02977	2.85348	0.00432	0.0266 0.14329
male_1994	-0.02062	0.03003	-0.68672	0.49225	-0.0795 0.03824
male_1995	0.018	0.03012	0.59758	0.55011	-0.04103 0.07704
male	-0.66497	0.02125	-31.2827	7.99996e-215	-0.70664 -0.62331
father_education	0.72617	0.01037	69.9683	0	0.70583 0.74652
utdanning_far_ukjent	-0.46905	0.02089	-22.4509	1.25392e-111	-0.51 -0.4281
søsken_1	0.28451	0.01048	27.1433	3.03228e-162	0.26396 0.30505
søsken_2	0.2614	0.01104	23.6589	9.54189e-124	0.23974 0.28305
søsken_3_5	0.01533	0.0135	1.13598	0.25596	-0.01112 0.0418
søsken_ukjent	-0.96832	0.02374	-40.783	0	-1.01486 -0.92178
Immigrant_1	-0.02227	0.0285	-0.78134	0.43459	-0.07814 0.03359
Immigrant_2	-0.39291	0.01809	-21.7172	1.41057e-104	-0.42837 -0.35745
Konst	0.33307	0.01733	19.2123	2.91662e-82	0.29909 0.36705

Model 12.1 cont.

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 father_education utdanning_far_ukjent søsken_1 søsken_2 søsken_3_5 søsken_ukjent Immigrant_1 Immigrant_2 if mother_education ==1

Antall iter: 6

Log sans: -93993.7

Antall obs: 181976

LR chi2(26): 10050.9

Prob > chi20

Pseudo R2: 0.05075

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	-0.046	0.0417	-1.10295	0.27004	-0.12775 0.03574
født_1989	-0.0185	0.0412	-0.44905	0.65339	-0.09925 0.06225
født_1990	-0.07451	0.04039	-1.84478	0.06506	-0.15368 0.00465
født_1991	-0.03548	0.04037	-0.87894	0.37942	-0.1146 0.04364
født_1992	-0.07681	0.03995	-1.92275	0.05451	-0.15511 0.00148
født_1993	-0.08887	0.03961	-2.2432	0.02488	-0.16652 -0.01122
født_1994	-0.02381	0.03951	-0.60261	0.54676	-0.10126 0.05363
født_1995	0.05306	0.03954	1.34193	0.17961	-0.02443 0.13057
male_1988	0.0409	0.05389	0.75906	0.44781	-0.06472 0.14653
male_1989	0.0877	0.05336	1.64344	0.10029	-0.01689 0.1923
male_1990	0.11793	0.05251	2.24555	0.02473	0.01499 0.22086
male_1991	0.0612	0.05217	1.17306	0.24076	-0.04105 0.16346
male_1992	0.12079	0.05181	2.33099	0.01975	0.01922 0.22235
male_1993	0.04215	0.05134	0.82113	0.41157	-0.05846 0.14278
male_1994	0.04594	0.05119	0.89739	0.3695	-0.05439 0.14627
male_1995	-0.01483	0.05101	-0.2909	0.77112	-0.11482 0.08514
male	-0.70998	0.03909	-18.1608	1.05337e-73	-0.78661 -0.63336
father_education	0.6971	0.01171	59.5043	0	0.67413 0.72006
utdanning_far_ukjent	-0.12482	0.03576	-3.48997	0.00048	-0.19492 -0.05472
søsken_1	0.33361	0.01812	18.4028	1.24616e-75	0.29808 0.36914
søsken_2	0.40549	0.01868	21.6974	2.1706e-104	0.36886 0.44212
søsken_3_5	0.30314	0.02342	12.9433	2.56093e-38	0.25724 0.34905
søsken_ukjent	-0.74375	0.04171	-17.8285	4.24048e-71	-0.82551 -0.66198
Immigrant_1	-0.27317	0.03543	-7.70878	1.27022e-14	-0.34263 -0.20372
Immigrant_2	-0.7925	0.02963	-26.7445	1.4303e-157	-0.85058 -0.73442
Konst	1.00186	0.03364	29.7797	7.14175e-195	0.93592 1.06779

Model 12.1 excluded

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 father_education utdanning_far_ukjent søsken_1 søsken_2 søsken_3_5 søsken_ukjent Immigrant_1 Immigrant_2 if utdanning_mor_ukjent ==1

Antall iter7
 Log sans-30348.1
 Antall obs90022
 LR chi2(26)3516.27
 Prob > chi20
 Pseudo R20.05475

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	0.04958	0.05072	0.97762	0.32826	-0.04982 0.149
født_1989	0.19864	0.05091	3.90153	0.00009	0.09885 0.29843
født_1990	0.16108	0.05309	3.03381	0.00241	0.05701 0.26515
født_1991	0.1432	0.05708	2.50851	0.01212	0.03131 0.25509
født_1992	0.06775	0.06121	1.10682	0.26836	-0.05222 0.18773
født_1993	0.04708	0.06646	0.70844	0.47867	-0.08317 0.17734
født_1994	0.31214	0.0675	4.62427	0	0.17984 0.44444
født_1995	0.43947	0.07118	6.17347	6.68069e-10	0.29995 0.579
male_1988	-0.13264	0.07361	-1.80187	0.07156	-0.27692 0.01163
male_1989	-0.27292	0.07498	-3.63959	0.00027	-0.41989 -0.12595
male_1990	-0.2556	0.0788	-3.24331	0.00118	-0.41006 -0.10114
male_1991	-0.2167	0.08356	-2.59333	0.0095	-0.38048 -0.05292
male_1992	-0.38548	0.091	-4.23602	0.00002	-0.56384 -0.20712
male_1993	-0.5094	0.09982	-5.10307	3.34171e-7	-0.70505 -0.31375
male_1994	-0.4897	0.10036	-4.87901	0	-0.68642 -0.29298
male_1995	-0.33322	0.10237	-3.25483	0.00113	-0.53387 -0.13256
male	-0.07629	0.05089	-1.49885	0.13391	-0.17605 0.02347
father_education	0.65028	0.06379	10.194	2.10743e-24	0.52526 0.77531
utdanning_far_ukjent	-0.3204	0.04544	-7.05009	1.78799e-12	-0.40948 -0.23133
søsken_1	0.18316	0.09083	2.01658	0.04373	0.00514 0.36119
søsken_2	-0.03082	0.09141	-0.33719	0.73596	-0.21 0.14834
søsken_3_5	-0.06902	0.08501	-0.81194	0.41682	-0.23565 0.09759
søsken_ukjent	-1.38134	0.07952	-17.3699	1.39312e-67	-1.53721 -1.22547
Immigrant_1	0.21873	0.05617	3.89362	0.00009	0.10863 0.32884
Immigrant_2	-0.29097	0.05296	-5.49368	3.93637e-8	-0.39478 -0.18716
Konst	-0.53489	0.08849	-6.04437	1.49988e-9	-0.70834 -0.36144

Model 12.2

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 male_mother_education utdanning_mor_ukjent søsken_1 søsken_2 søsken_3_5 søsken_ukjent Immigrant_1 Immigrant_2 if_father_education_low ==1

Antall iter: 5

Log sans: -2.4765414e+5

Antall obs: 380701

LR chi2(26): 23196.7

Prob > chi20

Pseudo R2: 0.04473

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	-0.01988	0.02155	-0.92232	0.35635	-0.06212 0.02236
født_1989	-0.00431	0.02146	-0.2009	0.84077	-0.04638 0.03775
født_1990	-0.03873	0.0213	-1.81869	0.06895	-0.08048 0.003
født_1991	-0.02955	0.02133	-1.38547	0.1659	-0.07136 0.01225
født_1992	-0.01797	0.02148	-0.83636	0.40294	-0.06009 0.02414
født_1993	-0.0221	0.02147	-1.02925	0.30335	-0.06419 0.01998
født_1994	0.11417	0.02172	5.25543	1.47674e-7	0.07159 0.15676
født_1995	0.1447	0.02177	6.64646	3.00221e-11	0.10202 0.18737
male_1988	0.0445	0.02949	1.50878	0.13135	-0.0133 0.10231
male_1989	0.07491	0.02931	2.55531	0.0106	0.01745 0.13237
male_1990	0.09281	0.02909	3.19065	0.00141	0.0358 0.14983
male_1991	0.05553	0.02909	1.90885	0.05628	-0.00148 0.11254
male_1992	0.08289	0.02924	2.83479	0.00458	0.02558 0.1402
male_1993	0.055	0.02931	1.87647	0.06058	-0.00244 0.11245
male_1994	-0.00226	0.02948	-0.07691	0.93868	-0.06004 0.05551
male_1995	0.0154	0.02953	0.52159	0.60194	-0.04248 0.07329
male	-0.67888	0.02118	-32.0444	2.62335e-225	-0.72041 -0.63736
mother_education	0.67498	0.00871	77.4169	0	0.6579 0.69207
utdanning_mor_ukjent	-0.73562	0.03349	-21.9651	6.21012e-107	-0.80126 -0.66998
søsken_1	0.31244	0.01032	30.2686	2.96286e-201	0.29221 0.33267
søsken_2	0.30312	0.01084	27.9599	4.98516e-172	0.28187 0.32437
søsken_3_5	0.08586	0.01331	6.44653	1.14432e-10	0.05975 0.11196
søsken_ukjent	-0.95312	0.02528	-37.6995	0	-1.00267 -0.90357
Immigrant_1	-0.07737	0.03188	-2.42676	0.01523	-0.13985 -0.01488
Immigrant_2	-0.37477	0.01998	-18.7513	1.88805e-78	-0.41395 -0.3356
Konst	0.31754	0.0173	18.348	3.42246e-75	0.28362 0.35146

Model 12.2 cont.

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 male_mother_education utdanning_mor_ukjent søsken_1 søsken_2 søsken_3_5 søsken_ukjent Immigrant_1 Immigrant_2 if father_education ==1

Antall iter: 6

Log sans: -75741.6

Antall obs: 152136

LR chi2(26): 8999.54

Prob > chi20

Pseudo R2: 0.05607

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	0.01629	0.04444	0.36666	0.71386	-0.0708 0.10339
født_1989	0.03887	0.04425	0.87846	0.37968	-0.04786 0.12561
født_1990	-0.00225	0.04348	-0.05196	0.95855	-0.08748 0.08296
født_1991	0.02935	0.04363	0.67266	0.50115	-0.05617 0.11487
født_1992	-0.00798	0.04332	-0.18436	0.85372	-0.0929 0.07693
født_1993	-0.07324	0.04288	-1.70784	0.08766	-0.15729 0.01081
født_1994	0.06817	0.0436	1.5637	0.11788	-0.01727 0.15363
født_1995	0.05305	0.04322	1.22736	0.21968	-0.03166 0.13776
male_1988	-0.00767	0.05778	-0.1329	0.89427	-0.12093 0.10557
male_1989	0.0107	0.05752	0.18612	0.85234	-0.10203 0.12344
male_1990	0.0416	0.05676	0.73302	0.46354	-0.06964 0.15285
male_1991	0.00808	0.05683	0.14227	0.88686	-0.1033 0.11947
male_1992	-0.02052	0.05631	-0.36441	0.71554	-0.1309 0.08985
male_1993	0.0621	0.05602	1.10847	0.26765	-0.0477 0.17191
male_1994	-0.05624	0.05645	-0.99625	0.31912	-0.16688 0.0544
male_1995	-0.05196	0.05598	-0.92826	0.35326	-0.1617 0.05776
male	-0.62172	0.04125	-15.0713	2.49943e-51	-0.70257 -0.54087
mother_education	0.63499	0.01301	48.7834	0	0.60948 0.6605
utdanning_mor_ukjent	-0.4874	0.05908	-8.24888	1.59881e-16	-0.60321 -0.37159
søsken_1	0.32831	0.02087	15.7299	9.43525e-56	0.2874 0.36921
søsken_2	0.38101	0.02142	17.7867	8.9574e-71	0.33903 0.423
søsken_3_5	0.20745	0.02567	8.08044	6.45307e-16	0.15713 0.25777
søsken_ukjent	-0.94824	0.04493	-21.1044	7.24386e-99	-1.0363 -0.86018
Immigrant_1	-0.28264	0.03913	-7.22145	5.14355e-13	-0.35935 -0.20593
Immigrant_2	-0.86999	0.02798	-31.091	3.18372e-212	-0.92483 -0.81514
Konst	1.02788	0.03592	28.6128	4.64335e-180	0.95747 1.09829

Model 12.2 excluded

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 male_mother_education utdanning_mor_ukjent søsken_1 søsken_2 søsken_3_5 søsken_ukjent Immigrant_1 Immigrant_2 if utdanning_far_ukjent ==1

Antall iter6
 Log sans-37444.3
 Antall obs100244
 LR chi2(26)10760.9
 Prob > chi20
 Pseudo R20.12563

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	-0.00683	0.04631	-0.14764	0.88262	-0.09762 0.08394
født_1989	0.12222	0.04659	2.62299	0.00871	0.03089 0.21354
født_1990	0.10633	0.04797	2.21659	0.02665	0.01231 0.20035
født_1991	0.1335	0.05024	2.65724	0.00787	0.03503 0.23198
født_1992	0.06661	0.05312	1.25405	0.20982	-0.03749 0.17073
født_1993	0.00429	0.05681	0.07556	0.93976	-0.10705 0.11564
født_1994	0.29371	0.05677	5.17305	2.30303e-7	0.18243 0.405
født_1995	0.47191	0.05867	8.04298	8.76785e-16	0.35691 0.58691
male_1988	-0.02117	0.06818	-0.31049	0.75618	-0.15481 0.11247
male_1989	-0.19423	0.06947	-2.79579	0.00517	-0.3304 -0.05806
male_1990	-0.16853	0.07174	-2.34922	0.01881	-0.30914 -0.02792
male_1991	-0.1448	0.07481	-1.93547	0.05293	-0.29143 0.00183
male_1992	-0.21116	0.07874	-2.6817	0.00732	-0.36549 -0.05683
male_1993	-0.30918	0.08397	-3.68177	0.00023	-0.47377 -0.14459
male_1994	-0.39089	0.08489	-4.6044	0	-0.55729 -0.2245
male_1995	-0.31663	0.08425	-3.75813	0.00017	-0.48177 -0.1515
male	-0.21443	0.0475	-4.51378	0	-0.30754 -0.12132
mother_education	0.78267	0.03673	21.3076	9.63779e-101	0.71068 0.85466
utdanning_mor_ukjent	-0.75371	0.03502	-21.5177	1.06169e-102	-0.82236 -0.68506
søsken_1	-0.1173	0.04338	-2.70383	0.00685	-0.20232 -0.03227
søsken_2	-0.26169	0.04908	-5.3313	9.75079e-8	-0.3579 -0.16548
søsken_3_5	-0.2745	0.05413	-5.07045	3.96873e-7	-0.38061 -0.16839
søsken_ukjent	-1.308	0.04491	-29.1231	1.83009e-186	-1.39603 -1.21997
Immigrant_1	0.22667	0.03535	6.41138	1.44204e-10	0.15738 0.29597
Immigrant_2	-0.24418	0.03336	-7.31923	2.49381e-13	-0.30956 -0.17879
Konst	-0.12437	0.04897	-2.53927	0.0111	-0.22036 -0.02837

Model 13

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 male_mother_education father_education utdanning_mor_ukjent utdanning_far_ukjent søsken_1 søsken_2 søsken_3_5 søsken_ukjent if Immigrant_2==1

Antall iter: 6

Log sans: -26574.1

Antall obs: 64447

LR chi2(26): 14345.9

Prob > chi20

Pseudo R2: 0.21255

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	0.05366	0.05802	0.92478	0.35507	-0.06006 0.16738
født_1989	0.2382	0.05827	4.08726	0.00004	0.12397 0.35243
født_1990	0.20708	0.05908	3.505	0.00045	0.09128 0.32288
født_1991	0.31804	0.06038	5.26666	1.38925e-7	0.19968 0.4364
født_1992	0.27747	0.06124	4.53061	0	0.15743 0.3975
født_1993	0.21765	0.06403	3.39899	0.00067	0.09214 0.34316
født_1994	0.52895	0.06324	8.36321	6.10323e-17	0.40499 0.65291
født_1995	0.64164	0.0633	10.1363	3.81036e-24	0.51757 0.76571
male_1988	-0.15026	0.08721	-1.72294	0.08489	-0.3212 0.02067
male_1989	-0.28427	0.08789	-3.23423	0.00121	-0.45654 -0.112
male_1990	-0.26544	0.0894	-2.96887	0.00298	-0.44068 -0.0902
male_1991	-0.28138	0.08987	-3.13078	0.00174	-0.45754 -0.10522
male_1992	-0.33559	0.09091	-3.69137	0.00022	-0.51378 -0.15741
male_1993	-0.29993	0.09347	-3.20865	0.00133	-0.48314 -0.11672
male_1994	-0.4635	0.09318	-4.97425	6.54991e-7	-0.64613 -0.28087
male_1995	-0.36404	0.09109	-3.99618	0.00006	-0.54259 -0.18549
male	-0.34888	0.06093	-5.72532	1.03235e-8	-0.46832 -0.22945
mother_education	0.45623	0.03424	13.3224	1.71276e-40	0.38911 0.52335
father_education	0.3288	0.03364	9.77154	1.49159e-22	0.26285 0.39475
utdanning_mor_ukjent	-0.75605	0.03208	-23.5647	8.87284e-123	-0.81894 -0.69317
utdanning_far_ukjent	-0.46173	0.03063	-15.074	2.40014e-51	-0.52176 -0.40169
søsken_1	-0.03685	0.04619	-0.79775	0.42501	-0.12739 0.05369
søsken_2	-0.22016	0.0493	-4.46578	0	-0.31679 -0.12353
søsken_3_5	-0.47058	0.04744	-9.91789	3.4803e-23	-0.56358 -0.37758
søsken_ukjent	-1.37681	0.04742	-29.034	2.44493e-185	-1.46975 -1.28386
Konst	0.13079	0.05821	2.24691	0.02464	0.0167 0.24489

Model 13 cont.

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 mother_education father_education utdanning_mor_ukjent utdanning_far_ukjent søsken_1 søsken_2 søsken_3_5 søsken_ukjent if Immigrant_1==1

Antall iter: 6

Log sans: -23375.2

Antall obs: 56219

LR chi2(26): 10265.6

Prob > chi20

Pseudo R2: 0.18004

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	0.00897	0.06205	0.14458	0.88504	-0.11264 0.13059
født_1989	0.09367	0.06199	1.511	0.13078	-0.02783 0.21518
født_1990	0.04055	0.06381	0.63556	0.52506	-0.08451 0.16563
født_1991	0.05489	0.06647	0.82587	0.40887	-0.07538 0.18517
født_1992	-0.00438	0.06884	-0.06369	0.9492	-0.13932 0.13054
født_1993	-0.02341	0.07241	-0.32337	0.74641	-0.16534 0.11851
født_1994	0.23957	0.07384	3.24424	0.00117	0.09483 0.3843
født_1995	0.32432	0.07665	4.2308	0.00002	0.17408 0.47457
male_1988	0.02568	0.08627	0.29767	0.76595	-0.1434 0.19476
male_1989	-0.13135	0.08737	-1.50334	0.13274	-0.30261 0.03989
male_1990	-0.15007	0.09049	-1.65834	0.09724	-0.32744 0.02729
male_1991	-0.07498	0.09324	-0.80418	0.42129	-0.25774 0.10777
male_1992	-0.25023	0.09785	-2.55714	0.01055	-0.44202 -0.05843
male_1993	-0.32584	0.10354	-3.14684	0.00165	-0.52879 -0.12289
male_1994	-0.40843	0.10488	-3.89414	0.00009	-0.614 -0.20286
male_1995	-0.25917	0.10544	-2.45784	0.01397	-0.46585 -0.0525
male	-0.10449	0.06024	-1.73443	0.08284	-0.22256 0.01358
mother_education	0.58421	0.0432	13.5225	1.15063e-41	0.49953 0.66889
father_education	0.57451	0.04984	11.5258	9.7768e-31	0.47682 0.67221
utdanning_mor_ukjent	-0.9268	0.04828	-19.1947	4.09262e-82	-1.02143 -0.83216
utdanning_far_ukjent	-0.32617	0.04368	-7.46631	8.24741e-14	-0.41179 -0.24055
søsken_1	0.19815	0.06497	3.04968	0.00229	0.0708 0.32551
søsken_2	0.16202	0.06991	2.3174	0.02048	0.02499 0.29906
søsken_3_5	0.00753	0.08319	0.09057	0.92782	-0.15552 0.1706
søsken_ukjent	-0.7877	0.06539	-12.0448	2.06406e-33	-0.91587 -0.65952
Konst	0.07042	0.07526	0.93563	0.34946	-0.07709 0.21794

Model 13 cont.

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 male_mother_education father_education utdanning_far_ukjent søsken_1 søsken_2 søsken_3_5 søsken_ukjent if Norwegian==1

Antall iter: 5
 Log sans: -3.1098043e+5
 Antall obs: 512415
 LR chi2(25): 47220.7
 Prob > chi20
 Pseudo R2: 0.07056

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	-0.02501	0.01977	-1.26502	0.20586	-0.06376 0.01374
født_1989	-0.01288	0.01969	-0.65421	0.51297	-0.05147 0.02571
født_1990	-0.04487	0.01952	-2.29825	0.02154	-0.08313 -0.0066
født_1991	-0.03949	0.01954	-2.02043	0.04333	-0.0778 -0.00118
født_1992	-0.03516	0.01968	-1.78629	0.07405	-0.07375 0.00341
født_1993	-0.05406	0.01963	-2.75407	0.00588	-0.09254 -0.01559
født_1994	0.07391	0.01988	3.71798	0.0002	0.03494 0.11287
født_1995	0.0956	0.01991	4.79959	0	0.05656 0.13464
male_1988	0.0476	0.02672	1.78104	0.0749	-0.00478 0.09998
male_1989	0.07869	0.02657	2.96158	0.00306	0.02661 0.13077
male_1990	0.10201	0.02635	3.8705	0.0001	0.05035 0.15367
male_1991	0.06573	0.02636	2.4936	0.01264	0.01406 0.1174
male_1992	0.08385	0.02648	3.16655	0.00154	0.03195 0.13575
male_1993	0.07686	0.02647	2.90287	0.00369	0.02496 0.12876
male_1994	0.01272	0.02665	0.47726	0.63317	-0.03951 0.06495
male_1995	0.01591	0.02669	0.59609	0.5511	-0.0364 0.06823
male	-0.68593	0.01921	-35.69	5.64091e-279	-0.7236 -0.64826
mother_education	0.68402	0.00736	92.9308	0	0.66959 0.69844
father_education	0.73811	0.00802	91.9562	0	0.72237 0.75384
utdanning_far_ukjent	-0.55238	0.02315	-23.8607	7.83689e-126	-0.59776 -0.50701
søsken_1	0.30799	0.00931	33.0481	1.65495e-239	0.28972 0.32626
søsken_2	0.3141	0.00974	32.2459	4.00335e-228	0.29501 0.3332
søsken_3_5	0.11701	0.01207	9.68779	3.39797e-22	0.09334 0.14068
søsken_ukjent	-1.11048	0.02575	-43.1151	0	-1.16096 -1.05999
Konst	0.32797	0.01589	20.6386	1.23474e-94	0.29682 0.35911

Note: Utdanning_mor_ukjent (mother education unknown) is excluded from this regression, as there are too few observations when Norwegian==1

Model 14

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 male_mother_education father_education utdanning_mor_ukjent utdanning_far_ukjent Immigrant_1 Immigrant_2 if søsken_0 ==1

Antall iter: 5

Log sans: -47788

Antall obs: 73547

LR chi2(24): 5002.17

Prob > chi20

Pseudo R2: 0.04973

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	-0.00445	0.04309	-0.10348	0.91757	-0.08892 0.08
født_1989	-0.01778	0.04401	-0.40394	0.68625	-0.10405 0.06849
født_1990	-0.05639	0.04439	-1.2704	0.20393	-0.14339 0.0306
født_1991	0.02105	0.04549	0.46287	0.64345	-0.0681 0.11022
født_1992	-0.10181	0.04629	-2.19942	0.02784	-0.19254 -0.01108
født_1993	-0.04829	0.04735	-1.01977	0.30783	-0.14111 0.04452
født_1994	0.0834	0.04867	1.7136	0.0866	-0.01199 0.17881
født_1995	0.13604	0.0487	2.7931	0.00522	0.04057 0.2315
male_1988	0.00275	0.05899	0.04672	0.96272	-0.11287 0.11838
male_1989	0.11492	0.06025	1.9073	0.05648	-0.00317 0.23302
male_1990	0.06984	0.06115	1.14206	0.25342	-0.05001 0.1897
male_1991	0.04652	0.0623	0.74668	0.45525	-0.07559 0.16865
male_1992	0.18685	0.06339	2.94729	0.0032	0.06259 0.31111
male_1993	0.14363	0.06508	2.20691	0.02731	0.01607 0.2712
male_1994	-0.00189	0.06622	-0.02867	0.97712	-0.1317 0.1279
male_1995	-0.02985	0.066	-0.45232	0.65103	-0.15922 0.09951
male	-0.66765	0.04128	-16.1724	7.89346e-59	-0.74856 -0.58673
mother_education	0.57646	0.01882	30.6201	6.60632e-206	0.53956 0.61336
father_education	0.70205	0.02107	33.3096	2.80172e-243	0.66074 0.74336
utdanning_mor_ukjent	-0.6382	0.07646	-8.34632	7.04175e-17	-0.78807 -0.48833
utdanning_far_ukjent	-0.11698	0.03667	-3.18969	0.00142	-0.18886 -0.04509
Immigrant_1	-0.14951	0.05729	-2.60968	0.00906	-0.2618 -0.03722
Immigrant_2	-0.21114	0.04282	-4.93051	8.20128e-7	-0.29507 -0.1272
Konst	0.33082	0.03047	10.8539	1.90887e-27	0.27108 0.39055

Model 14 cont.

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 mother_education father_education utdanning_far_ukjent Immigrant_1 Immigrant_2 if søsken_1 ==1

Antall iter: 5

Log sans: -1.3658414e+5

Antall obs: 224995

LR chi2(23): 16973.1

Prob > chi20

Pseudo R2: 0.05849

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf.	intervall]
født_1988	-0.01435	0.03025	-0.47468	0.63501	-0.07364	0.04493
født_1989	0.02166	0.03003	0.72149	0.4706	-0.03719	0.08052
født_1990	-0.02256	0.02963	-0.76141	0.44641	-0.08064	0.03551
født_1991	-0.00205	0.02976	-0.06892	0.94504	-0.06039	0.05628
født_1992	0.02815	0.02999	0.93862	0.34792	-0.03063	0.08694
født_1993	-0.00564	0.02977	-0.18968	0.84955	-0.06399	0.0527
født_1994	0.12435	0.03	4.14413	0.00003	0.06554	0.18317
født_1995	0.1569	0.0301	5.21108	1.87744e-7	0.09788	0.21591
male_1988	0.0341	0.0408	0.83592	0.40319	-0.04586	0.11407
male_1989	0.04344	0.04049	1.07289	0.28331	-0.03592	0.12282
male_1990	0.07681	0.04005	1.9179	0.05512	-0.00168	0.15531
male_1991	0.01682	0.04008	0.41978	0.67463	-0.06173	0.09539
male_1992	0.0102	0.04024	0.25348	0.79989	-0.06868	0.08908
male_1993	0.02887	0.04019	0.71836	0.47253	-0.0499	0.10764
male_1994	-0.03709	0.04025	-0.92143	0.35682	-0.11599	0.0418
male_1995	-0.02505	0.04033	-0.62129	0.5344	-0.10411	0.05399
male	-0.66878	0.02937	-22.7693	9.22255e-115	-0.72634	-0.61121
mother_education	0.62588	0.01105	56.6002	0	0.60421	0.64756
father_education	0.7019	0.01211	57.9592	0	0.67816	0.72563
utdanning_far_ukjent	-0.53881	0.03088	-17.4487	3.51563e-68	-0.59934	-0.47829
Immigrant_1	-0.16688	0.03698	-4.51258	0	-0.23937	-0.0944
Immigrant_2	-0.50221	0.02643	-18.9966	1.81872e-80	-0.55403	-0.4504
Konst	0.63229	0.02195	28.7932	2.60593e-182	0.58925	0.67533

Model 14 cont.

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 male_mother_education father_education utdanning_mor_ukjent utdanning_far_ukjent Immigrant_1 Immigrant_2 if søsken_2_5 ==1

Antall iter: 5

Log sans: -1.4137335e+5

Antall obs: 235832

LR chi2(24): 24103.5

Prob > chi20

Pseudo R2: 0.07855

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	-0.03933	0.03121	-1.25995	0.20768	-0.10052 0.02185
født_1989	-0.03675	0.03073	-1.19601	0.23169	-0.09699 0.02347
født_1990	-0.05917	0.03039	-1.94694	0.05154	-0.11873 0.00039
født_1991	-0.07449	0.03016	-2.46908	0.01354	-0.13362 -0.01535
født_1992	-0.07747	0.03022	-2.563	0.01037	-0.13671 -0.01822
født_1993	-0.10889	0.03016	-3.60929	0.0003	-0.16802 -0.04976
født_1994	0.03203	0.03062	1.04605	0.29553	-0.02798 0.09205
født_1995	0.02939	0.03065	0.95894	0.33758	-0.03068 0.08948
male_1988	0.08897	0.04177	2.12966	0.03319	0.00708 0.17085
male_1989	0.08726	0.04106	2.12489	0.03359	0.00677 0.16775
male_1990	0.11432	0.04056	2.81792	0.00483	0.0348 0.19383
male_1991	0.1032	0.04035	2.55757	0.01054	0.02411 0.18229
male_1992	0.1218	0.04039	3.01577	0.00256	0.04264 0.20097
male_1993	0.10506	0.04032	2.60558	0.00917	0.02603 0.18409
male_1994	0.0511	0.04074	1.25441	0.20969	-0.02874 0.13095
male_1995	0.06936	0.0408	1.69981	0.08916	-0.01061 0.14935
male	-0.71232	0.03046	-23.384	6.21549e-121	-0.77202 -0.65261
mother_education	0.74461	0.0109	68.2837	0	0.72324 0.76599
father_education	0.73422	0.01162	63.1459	0	0.71143 0.75701
utdanning_mor_ukjent	-0.57324	0.03657	-15.673	2.31086e-55	-0.64492 -0.50155
utdanning_far_ukjent	-0.51594	0.03003	-17.1786	3.83789e-66	-0.57481 -0.45708
Immigrant_1	-0.22528	0.03715	-6.06349	1.33198e-9	-0.2981 -0.15246
Immigrant_2	-0.70275	0.02196	-31.9932	1.35559e-224	-0.7458 -0.65969
Konst	0.60927	0.02309	26.384	2.09101e-153	0.56401 0.65453

Model 14 excluded

endeliglogit fullført_innen_20 født_1988 født_1989 født_1990 født_1991 født_1992 født_1993 født_1994 født_1995 male_1988 male_1989 male_1990 male_1991 male_1992 male_1993 male_1994 male_1995 mother_education father_education utdanning_mor_ukjent utdanning_far_ukjent Immigrant_1 Immigrant_2 if søsken_ukjent ==1

Antall iter: 6

Log sans: -35216.5

Antall obs: 98707

LR chi2(24): 5689.44

Prob > chi20

Pseudo R2: 0.07474

fullført_innen_20	Coef.	Std.feil	z	P> z	[95% Konf. intervall]
født_1988	0.01123	0.04837	0.23217	0.81639	-0.08358 0.10605
født_1989	0.15569	0.04911	3.1703	0.00152	0.05944 0.25195
født_1990	0.11781	0.05106	2.30728	0.02103	0.01773 0.21789
født_1991	0.11806	0.05399	2.18642	0.02878	0.01222 0.22389
født_1992	0.12438	0.05572	2.23195	0.02561	0.01515 0.2336
født_1993	0.09627	0.05863	1.64179	0.10063	-0.01865 0.2112
født_1994	0.39935	0.05716	6.98655	2.81726e-12	0.28732 0.51138
født_1995	0.55569	0.0574	9.68087	3.63574e-22	0.44318 0.66819
male_1988	-0.10412	0.07138	-1.45871	0.14464	-0.24403 0.03578
male_1989	-0.21427	0.07288	-2.94001	0.00328	-0.35712 -0.07142
male_1990	-0.1469	0.07541	-1.94802	0.05141	-0.2947 0.0009
male_1991	-0.10619	0.07875	-1.34838	0.17753	-0.26056 0.04816
male_1992	-0.3814	0.08347	-4.56925	0	-0.545 -0.2178
male_1993	-0.3783	0.08571	-4.41347	0.00001	-0.5463 -0.2103
male_1994	-0.49877	0.08409	-5.93123	3.00673e-9	-0.6636 -0.33395
male_1995	-0.3764	0.0812	-4.63511	0	-0.53556 -0.21723
male	-0.0879	0.0485	-1.81235	0.06993	-0.18296 0.00715
mother_education	0.71067	0.04379	16.2264	3.28148e-59	0.62483 0.79651
father_education	0.4595	0.04733	9.70825	2.7805e-22	0.36673 0.55227
utdanning_mor_ukjent	-0.94589	0.03721	-25.4169	1.6391e-142	-1.01883 -0.87295
utdanning_far_ukjent	-0.40147	0.03865	-10.3851	2.89593e-25	-0.47724 -0.3257
Immigrant_1	0.33148	0.04208	7.87606	3.37863e-15	0.24899 0.41397
Immigrant_2	-0.17659	0.04112	-4.29371	0.00001	-0.2572 -0.09598
Konst	-1.01415	0.04126	-24.5745	2.36288e-133	-1.09503 -0.93327

Appendix 8: Info on excluded variables

Endelig tabulate Norwegian utdanning_mor_ukjent

		<i>utdanning_mor_ukjent</i>		
		0	1	Total
<i>Norwegian</i>	0	33511	87152	120666
	1	509544	2864	512417
<i>Total</i>		543061	90019	633084

Endelig tabulate Norwegian utdanning_far_ukjent

		<i>utdanning_far_ukjent</i>		
		0	1	Total
<i>Norwegian</i>	0	28954	91721	120666
	1	503891	8528	512417
<i>Total</i>		532837	100240	633084

Endelig tabulate søsken_0 utdanning_far_ukjent

		<i>utdanning_far_ukjent</i>		
		0	1	Total
<i>søsken_0</i>	0	463401	96133	559533
	1	69442	4111	73548
<i>Total</i>		532837	100240	633084

Endelig tabulate søsken_0 utdanning_mor_ukjent

		<i>utdanning_mor_ukjent</i>		
		0	1	Total
<i>søsken_0</i>	0	470386	89144	559533
	1	72667	881	73548
<i>Total</i>		543061	90019	633084

Explains why mother education unknown is excluded in some regression analyses; as there are too few observations. Excluded in subsamples Norwegian and for only children.