The Compositional Effect

The Decline in Labor Market Participation of Low-Skilled Males

Tore Adam Reiremo

Master of Philosophy in Economics

University of Oslo

November 2016
The Compositional Effect

*The Decline in Labor Market Participation of Low-Skilled Males*
© Tore Adam Reiremo

2016

The Compositional Effect - The Decline in Labor Market Participation of Low-Skilled Males

Tore Adam Reiremo

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo
Abstract

In this Master’s Thesis I investigate the decline in participation rate of Norwegian males without upper secondary education born in the period 1963 and 1983. The main question is to what extent the decline of above 17 percent can be explained by changes in the distribution of individual characteristics for males without upper secondary education, or if there are factors that makes participating in the labor force less prosperous for these individuals. Using a rich data set consisting of every male born between 1963 and 1983, data on parent’s income and IQ-test scores, I find that observed characteristics fail to explain much of the large decline. Using a more general approach I show that one must accept quite strong assumptions for the compositional effect to account for more than 60 percent of the decline in participation rate, leaving the other 40 percent to be explained by other factors. This raises a series of concerns regarding the wellbeing and future for low-skilled males in Norway, and should be subject to future research.
Preface

This master’s thesis has been written at the Ragnar Frisch Centre for Economic Research, as a part of the project “Egalitarianism under pressure? New perspectives on inequality and social cohesion”. The thesis has only been possible because of the rich data set that has been made available through the Frisch Centre, acquired from Statistics Norway and Norwegian Armed Forces. The views and conclusions expressed in this thesis are of the author, and cannot be attributed to anyone else.

I would like to give my gratitude to all the people working at the Frisch Centre for the opportunity to write this thesis. An especial thank goes to my supervisor Simen Markussen for his good ideas and important guidance throughout the process. I am also grateful to Elise Luhr Diethrichson for essential support, proofreading and feedback, and lastly to Trygve Larsen Morset for important feedback and ideas. Any remaining mistakes are my own.
Table of Contents

1 Introduction ........................................................................................................................ 1

2 Utilized data ....................................................................................................................... 4
  2.1 The Data Set ................................................................................................................ 4
  2.2 The Population of Interest ........................................................................................... 5
  2.3 Variables ...................................................................................................................... 6
    2.3.1 Labor Market Participation ................................................................................ 6
    2.3.2 Parental Income ................................................................................................. 7
    2.3.3 Education ............................................................................................................ 8
    2.3.4 IQ-test scores ..................................................................................................... 8
    2.3.5 Other variables .................................................................................................. 10

3 Descriptive Statistics ........................................................................................................ 11
  3.1 Graduation and Participation Rates ............................................................................. 11
  3.1 The compositional effect ........................................................................................... 14
    3.1.1 Father’s income ............................................................................................... 14
    3.1.2 Trend in IQ-test score ...................................................................................... 16
    3.1.3 Sectors of employment .................................................................................... 18
  3.2 Missing Observations ................................................................................................ 19
  3.3 Summary Statistics .................................................................................................... 22

4 Theoretical and Empirical Background ............................................................................ 24
  4.1 Skill Biased Technological Change ............................................................................ 24
  4.2 Immigration ............................................................................................................... 26
  4.3 Trade .......................................................................................................................... 27

5 Analysis ............................................................................................................................ 29
  5.1 Matching: Controlling for Observed Characteristics................................................. 29
    5.1.1 The Estimand ...................................................................................................... 30
    5.1.2 Limitations ........................................................................................................ 31
    5.1.3 Defining the Matching Estimator ....................................................................... 33
    5.1.4 Choosing Matching Algorithm ......................................................................... 35
    5.1.5 Inclusion of Individual characteristics ............................................................ 36
    5.1.6 Assessing Balance ............................................................................................. 37
    5.1.7 Inference ............................................................................................................ 38
    5.1.8 Results ................................................................................................................ 40
    5.1.9 Balance statistics ............................................................................................... 44
    5.1.10 Robustness tests .............................................................................................. 47

VIII
5.1.11 Conclusion of the Matching Estimators ............................................................. 48

5.2 Non-parametric Bounds ............................................................................................. 49

5.2.1 Methodology ...................................................................................................... 49

5.2.2 The Upper Bound ............................................................................................... 52

5.2.3 Bounds of Joint Participation rate. ..................................................................... 53

5.2.4 Conclusion Based on the Non-parametric Bounds ............................................ 55

6 Conclusion ........................................................................................................................ 56

References ................................................................................................................................ 58

List of tables and figures

Table 1 – The various data sets ................................................................................................. 4

Figure 1 - Norwegian graduation rates ..................................................................................... 11

Figure 3 – Participation by males with different educational attainment. ............................ 13

Figure 4 – Participation by females with different educational attainment. ....................... 13

Figure 5 – Graduation rates and father’s decile .................................................................... 14

Figure 6 – Distribution of father’s income, low-skilled males ........................................... 15

Figure 7 – Distribution of father’s income – medium-skilled males .................................... 16

Figure 8 – IQ-test scores – low-skilled males ........................................................................ 17

Figure 9 – IQ-test scores. High-skilled males ........................................................................ 17

Table 2 – Sectors of employment ............................................................................................ 18

Figure 10 – Missing parental income .................................................................................... 20

Figure 11 – Missing IQ-test scores (assumed ability level omitted) ..................................... 21

Table 3 – Summary statistics for individuals with non-missing observation ...................... 23

Table 4 – Difference in participation. Before and after matching, relative to 1980 (fraction) 41

Figure 12 – Weighting by the odds ....................................................................................... 42

Figure 13 – Mahalanobis matching ........................................................................................ 42

Figure 14 - Nearest neighbor matching .................................................................................. 43

Figure 15 – Kernel matching ................................................................................................ 43

Table 5 – Common support .................................................................................................. 44

Figure 16 – Standardized difference in means ...................................................................... 45

Figure 17 – Variance ratio ...................................................................................................... 46

Figure 18 – Robustness tests ................................................................................................. 48

Table 6 – Never-takers, always-takers, defiers and compliers .............................................. 50

Figure 19 – Shares of Never-takers, compliers and Always-takers ..................................... 51

Figure 20 – Joint participation, assuming full participation of compliers ............................. 53

Figure 21 – Joint participation rate under different assumptions ........................................ 54

Figure 22 – Share of the decline in participation due to the compositional effect ............ 55

Table 7 – Robustness tests ................................................................................................. 61

Table 8 – Robustness tests ................................................................................................. 62
1 Introduction

The overall level of employment and welfare benefits have attracted a lot of attention in the Norwegian political debate over the past recent decades (Terum and Hatland 2014). The trend showing that an increased number of people have shifted from paid work to welfare benefits has raised concern on both sides of the political spectrum. The debate reflects the more general notion that is rooted in the Norwegian welfare system, namely the principle that those able to work should account for their own income, while the role of the welfare state is the security of those that are unable.

The Norwegian labor market has also attracted a lot of attention for the high participation rates of women, as well as its policy that encourages women with children to work. As I will show later the participation rates increased from 75 percent of women born in 1963 to 82 percent of women born in 1983, measured at the age of 30. Over the same period participation rates of males experienced a slight decline from around 91 to 87 percent. Further examination reveal that this decline in labor market participation seems to be mostly caused by males without upper secondary education: The participation rate of this group declined by almost 17 percentage points during the twenty-year period between 1963-1983, also measured at age 30.

Participation is usually defined as someone who is either working or looking for a job (Statistics Norway 2015). In that sense, whether a person participates or not is a pure choice for anyone who is healthy enough to do so. So why has the participation rate of thirty-year old males without upper secondary education declined by such a large amount? The theories of qualitative choice (Train 1986) highlights that such decisions are affected by the characteristics of the decision maker and the characteristics of the available alternatives. That is, we might expect the decline in participation to be either due to changes in the characteristics of low-skilled males or their prospects of being inside or outside of the labor market.

There are several economic theories that suggests that the labor market conditions for low-skilled individuals worsen over time, i.e. there is a reduction in the demand for low-skilled individuals or the quality of jobs available to such individuals. One such theory is the theory of skill-biased technological change that predicts that innovation and capital accumulation over time leads to a demand shift towards workers with a higher skill set. Some other theories suggest that low-skilled individuals are more exposed to competition from immigration and
trade, such that the jobs that used to be available to these individuals might be transferred to immigrants or moved abroad. A withdrawal from the labor market may also be due to improvements in the conditions of withdrawing from the labor market, for example due to changes in the welfare system. If some of these effects have been strong during the period between 1994 and 2013, one might expect that more low-skilled males withdrew from the labor market due to a reduction of wages, more time spent in unemployment, reduced job-quality or improved life quality outside of the labor market.

If the above-mentioned theories are responsible for the observed shift in participation rates, it may give rise to a series of concerns. One such concern is that, if a decrease in the demand for low-skilled workers are not met with higher graduation rates, this will lead to higher income inequality (Goldin and Katz). Further, if there is a gradual reduction of the number of jobs that are available to low-skilled males, the effect of labor market programs that are being implemented by the government, might wear off because there simply are not enough jobs. Besides, if low-skilled individuals are becoming more and more dependent on the welfare system to earn a living, this might lead to higher government expenditure in the future.

During the twenty-year period between 1963 and 1983, however, there has also been a significant expansion of graduation rates from upper secondary education. This suggests that the group of low-skilled individuals in the beginning of the period might be quite a different group than low-skilled males today. To be more precise, one might suspect that dropouts in the 1963-cohorts consists of more individuals with a high ability level, motivation and strong social background than the cohorts in 1983, a change that is usually labeled compositional change (Juhn and Potter 2006).

If this compositional effect is accountable for the reduction in participation, there is not the same reasons for concern about the labor market conditions of low-skilled males. In fact, if the compositional change is responsible for a large share of the decline, the changes on the individual level may be minuscule.

The scope of this Master’s thesis is therefore to investigate to what extent the compositional effect is accountable for the observed reduction in participation in the labor market. The thesis can be summarized in the following question: Can the decline in labor market participation of Norwegian males without upper secondary education be explained by a compositional effect?
To assess this question, I will use a register data set consisting of various data on the individual level to try and decompose the decline in the participation into a compositional effect and other effects. The decomposition will be based on a matching estimator on observed characteristics, used as a screening device to construct comparable cohorts of low-skilled males. The change in participation of these constructed samples may no-longer be attributed to changes in observed characteristics, since these no longer vary between cohorts.

The choice of the matching estimator is based both on its intuitive appeal, and the fact that matching does not rely on assumptions regarding a functional form. In addition, the matching estimator can be used to evaluate the participation rate over the whole distribution of covariates. The limitations of this approach are that there is little academic literature and theory on matching in settings where one are not trying to estimate a causal effect (Stuart 2010). In addition, there is some ambiguity regarding the formal inference and choice between a wide variety of estimators (Ibid.). I will get back to these limitations in the section describing the methodological choices made in this thesis.

Two important aspects that limits the scope of the analysis is the number of missing observations and potentially important unobserved characteristics. The presence of these problems will severely limit the interpretation of the results. I will therefore develop an approach that offer the opportunity to say something about the feasible size of the compositional effect under less strict assumptions.

I will start by introducing the data set and present the descriptive statistics in chapter two and tree respectively. This is because the main motivation behind the research question lies in these data. I will then discuss some of the theories that may explain the decline in the participation rate in some detail in chapter four, as well as to discuss to what extent they may explain the observed decline in participation. I will then go on to implement the decomposition of the participation rate in chapter five, before I conclude in chapter six.
2 Utilized data

2.1 The Data Set

The analysis in this thesis is based on register data of family background, educational attainment, labor and capital income, official unemployment registries, county of residence, immigration background and IQ-test scores. The data is available for every individual that was born between 1963 and 1983, and has lived in Norway at some point in time before 2013. All data is provided by Statistics Norway, except for the IQ-test scores that are collected and provided by the Norwegian Armed Forces. Every data set contains information on individual level and is summarized in Table 1.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Summary</th>
<th>Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics and family background</td>
<td>Immigration status, birth year, and identification, education level and demographics of parents.</td>
<td>Statistics Norway</td>
</tr>
<tr>
<td>Educational attainment</td>
<td>Levels of education: Lower and upper secondary education, higher education.</td>
<td>Statistics Norway</td>
</tr>
<tr>
<td>Labor and capital income</td>
<td>Gross wage income, sickness leave transfers and net capital income of taxes.</td>
<td>Statistics Norway</td>
</tr>
<tr>
<td>Unemployment</td>
<td>Official register of unemployment and vocationally disability.</td>
<td>Statistics Norway</td>
</tr>
<tr>
<td>IQ-test scores</td>
<td>Test scores ranking from one to nine, based on three different tests of ability.</td>
<td>Norwegian Armed Forces</td>
</tr>
</tbody>
</table>
2.2 The Population of Interest

The primary focus of this thesis will be restricted to Norwegian males at the age of thirty. I will motivate this restriction in this subchapter. Subsequently, I will show how many individuals that are left in the population of interest.

To isolate the effects on participation from gender and immigration specific effects, I will restrict the analysis to Norwegian born men with at least one Norwegian born parent. In addition, I will only be focusing on participation at the year of thirty both to simplify the analysis, and to highlight the differences between the cohorts.

The primary reason for why I do not include immigrants in the analysis, is because these individuals might be subject to some immigration specific effects, such as the effects of discrimination against immigrants by employers. Another reason is that the reasons for immigration may not be the same over the whole period. This would imply that unobserved characteristics of immigrants in the beginning of the sample period are not comparable to those at the end of the period. Lastly, the observations on immigrants seems to be subject to some registration problems, i.e. the number of immigrants with duplicate and contradictory observations are quite large.

While participation rates of women may be interesting as well, I will not include women in the main analysis of this thesis. This is because there is a large increase in the overall participation rate of women over the period, and I am not able to decompose this effect from the compositional effect and other effects that this thesis will try to highlight. Another reason is that screening of teenagers by the Norwegian Armed Forces were not mandatory for women born before 1997, and there are therefore very sparse data on women’s IQ-test scores.

Even though it would be possible to construct a panel data set that would follow individuals over a period of time, I have decided to only use participation at the age of thirty. In this way, I will highlight the changes in the cohort composition, and considerably simplify both the construction of the data set and the analysis.

The total population born in the period between 1963 and 1983 who were living in Norway at the age of thirty consists of 681.770 males and 661.165 females. After subtracting the immigrant population, these numbers are reduced to 599.141 males and 571.428 females, of
which 163,492 males and 149,190 females have not graduated from upper secondary education.

### 2.3 Variables

A few variables need some more elaboration than the brief introduction above. This is because some are constructed for the usage in this Master’s Thesis, and some have some aspects that need a more thorough introduction.

#### 2.3.1 Labor Market Participation

I will be using a simple definition of labor market participation constructed on observed income and unemployment indicators. I will define an individual as part of the labor force if he or she receives an annual income above two times the official public pension base rate (grunnbeløpet i folketrygden), as defined by the government in Norway. The basic income is used by the government for means-testing and calculations of welfare benefits, and was equal to 85,245 NOK in 2013. This measure usually grows more than the price level, yet slower than the median income in Norway. An individual is also going to be defined as being a part of the labor force if he or she is officially registered as an applicant for three or more months in a given year.

This definition of the labor force is a rough definition. It excludes those who registered as unemployed for a short period of time and individuals with low income. The motivation is to exclude levels of income that are unlikely to be the main source of income, and include individuals that are most likely to be serious about their job applications. One possible limitation is that some individuals are looking for a job without officially register as an applicant. That makes the level of unemployment difficult to compare to official numbers from Statistics Norway, and other EU and EEC countries, whom base their statistics on surveys (Statistics Norway 2015). These two methods for constructing unemployment statistics shows similar trends over time, and mostly differ for individuals under the age of 25.
2.3.2 Parental Income

Observed income of parents is going to be the most important indicator of social background that will be implemented in this thesis. I have made some strategic choices on how to implement these observations to utilize the data as efficient as possible.

The data that will be used for parental income is the same data used in the definition of labor market participation: The sum of wage and capital income. The most significant advantage of this data set is that it is observable for the whole period between 1967 and 2014.

The approach I am going to use is the same as Markussen and Røed (2016). These authors use the average income over a period of three years as a proxy for lifetime income of parents, and use this as a measure of social background. Studying the 1945-cohort, they observe that the correlation coefficient with lifetime income is greatest for income observed between the age of 45 and 55, with a coefficient of around 0.8.

To make sure that observed income is comparable, I will compare fathers to fathers, mothers to mothers, and compare income registered at the same age of the parents. To make sure that income received in two different years are comparable I will compare income in multiples of the Public pension base rate, in addition to normalizing the income such that the average will be equal zero for each year. This way, the level of registered income is less likely to be affected by growth in income.

One potential problem is that the number of missing variables might be high in the first or last cohorts. This is because some parents might be born too late or too early to be observed with the right age between 1967 and 2014. The way I have approached this problem is to use observed income over years that minimizes the share of missing observation for any cohort. That is, I minimize the maximum share of missing observations in any cohort. The result of this approach is to use the average over the years that mothers are between the age of 46 to 48 and fathers are between the age of 51 to 53.

Lastly, these measures are used to construct father’s and mother’s income percentile over every cohort. That is, I will construct percentiles for every cohort, independently for mothers and fathers, using the measure constructed above.
2.3.3 Education
Norwegian children start their compulsory education at the age of six, and finish with graduation from lower secondary school at the age of 16. After which attendance in upper secondary education (Norwegian “videregående skole”) is voluntary. Upper secondary education usually lasts for three years for students who prepare for university, or three to four years for vocational schooling (Haaland).

The educational reform of 1994 (“reform 94”) granted every individual that graduated from lower secondary education after 1993 the statutory right to attend upper secondary education. In addition, the definition for having graduated from upper secondary education was slightly changed (Statistics Norway 2006). Prior to the reform, everyone that finished at least two years of upper secondary education was defined to be on the upper secondary level, while the requirements were increased to three or four years after the reform, depending on the length of the course. Individuals born after 1978 are therefore slightly more seldom registered as on the upper secondary level than individuals born in 1978 of earlier. Consequently, a change in the definition of graduation results in a small shift in the observed trend in the graduation trend.

If individuals with two years of upper secondary education had an ability-level between those with more and less education, this change in definition would have led to a greater compositional change and decline in participation. This is because the earliest cohorts would then include some individuals with higher ability than if the post reform definition were used over the whole period, increasing the initial gap in composition and consequently the participation rate. Although this may be the cause of some of the observed gap, the change in participation does not have a very large impact on graduation rate, and the consequences of this change in definition are likely to be small.

2.3.4 IQ-test scores
The IQ-test scores are collected by the Norwegian Armed Forces as part of the conscription to the Norwegian military service. The test scores rank from one to nine and are collected during military conscription. Participation in these sessions were usually carried out when individuals were between 17 or 20 years old, and were compulsory to males and voluntary for females the period when data I use was collected. Not showing up for conscription were
associated with sanctions from the Norwegian Armed Forces or the Norwegian government, and a large share of Norwegian males have therefore participated.

The test statistics that will be implemented in the thesis are based on the sum of the results of three independent tests: arithmetic, figures and word comparison. The results are calculated from the completed tests whenever one or two of these tests are missing, without this being indicated in the data set. If none of the tests were carried out, the ability levels are stored as “A”, “B” or “C” depending on the assumed ability level being below, above or around average.

To what extent are these test scores a good measure of ability that would be relevant for participation in the labor market? First, IQ is usually not regarded as a perfect measure of ability. It does capture some aspects of intelligence, but leave out skills like emotional intelligence and perseverance that are relevant for labor market outcomes (Haaland 2013). In addition, some nuances disappear when the overall test results are computed: we are not able to separate individuals that performs well on different parts of the tests, and some information is lost due to truncation.

Another aspect that might influence the validity of the IQ data is that these tests are carried out during or after attendance to upper secondary education, and is likely to be affected by attendance to such education. This is most likely to be a problem if one tries to compare graduates to non-graduates, since the test scores would then be a consequence of the level of education. In my analysis, I will compare non-graduates from different cohorts to each other. If some of these individuals have partial attendance to upper secondary education, this may be a consequence of ability, and may render them more able to participate in the labor force later. If this is the case, then such partial attendance to upper secondary education represents some unobserved variation that is not some sort of bias, but rather variation of interest.

Lastly, the number of missing IQ observations are quite large and vary between cohorts. This will also affect the validity of the data, it is therefore an aspect that will be scrutinized further in chapter four.
2.3.5 Other variables

In addition to the variables mentioned above, I will also use county of residence, the age of parents at birth and parental level of education in some parts of the analysis.

First, I will use the age of mothers and fathers at the age of birth as some indicator for social background. If for example individuals that has spent a lot of time acquiring a career have children later in life, then this variable will pick up some of this variation. Secondly, I will use parent’s level of education as another measure of social background. The measure of parental education used in this data set are not separate for mothers and fathers, but equal to the level of educational attained by the parent with highest level of education.

The main problem with these two variables are that they systematically change over time: While the average age of mothers and fathers declines by around one year between the cohorts born in 1963 and 1983, the share of individuals that have at least one parent with upper secondary education or higher education increases over the whole period. These trends give rise to the question of whether or not individuals with the same observed background are comparable across cohorts. If one is interested in the causal effect of parent’s education on children’s education one may want to control for this variation. If, on the other hand, the interest is in parental education as some sort of proxy for parental ability, then one may expect the level of education to signal different ability level as parents become more and more educated. This concern is discussed further in chapter 5.1.5, when inclusion of individual characteristics in the estimation procedure is discussed.

In this chapter the data that is being used have been introduced, and some constructed variables have been defined. In the following chapter I will have a closer look at some observed trends in this data set, as well as to investigate the impact of missing variables further.
3 Descriptive Statistics

In this chapter I will introduce the observed graduation- and participation rates: the two statistics that give rise to the hypothesis of a compositional effect that this thesis will aim to investigate. I will also have a closer look at the distribution of father’s income and IQ-test scores in the population, as a first assessment of plausibility of the compositional effect. Lastly, this chapter will discuss missing observations, and further investigate their impact on the analysis carried out later in this thesis.

3.1 Graduation and Participation Rates

The Norwegian education system experienced a rapid increase in graduation rates that begun before the birth of the 1963-cohort, and slowed down with the cohorts born in the mid and late seventies (Figure 1). While around 50-60 percent of the 1963-cohort graduated from upper secondary education, graduation rates increased to around 80 percent for the cohorts born in the beginning of the eighties. Furthermore, while graduation rates were approximately six percentage points lower for women for the cohorts born in the beginning of the sixties, this gap was more than reversed for the cohorts born towards the end of this period.

![Graduation Rates](image)

Figure 1 - Norwegian graduation rates
During the same period, we observe a rapid increase in labor market participation by women, and a gradual decline in participation rate by males (Figure 2). If we break down these number by education, we observe that males without upper secondary education drive most of the change in participation: medium- and high-skilled men have a fairly stable participation rate over the cohorts born between 1963 and 1983 (Figure 3). Notice that men with only upper secondary education participate slightly more than those with higher education. This is probably because I have not included education in the definition of participation: some with higher education might be pursuing an even higher degree. Although this gives a somewhat biased view on the relative participation of medium- and high-skilled males, it is likely to be a small problem for low-skilled male individuals.

When we break down participation by education for women, most of the increase in participation seems to have disappeared: there is a slight increase in participation for low- and high-skilled women, and a decrease for low-skilled women by about 10 percentage points (Figure 4). The reason why the increase in participation disappears when we break down the numbers is because of the rapid increase in graduation during the period: There are simply more women with upper secondary education towards the end of the period, and this group weighs more in the average towards the end of the period.

![Participation Rates](image)

Figure 2 – Participation rates by males and females.
Figure 3 – Participation by males with different educational attainment.

Figure 4 – Participation by females with different educational attainment.
Notice also that there seems to be two periods of falling participation rates for low-skilled individuals of both genders: one gradual reduction that parallels the trend in graduation, and a sharp decline right after the 1978-cohort. The gradual decline might be explained by the gradual outflow of individuals, i.e. the compositional effect, while the rapid decrease occurs right after the Great Recession of 2008. It seems like the great recession affected the participation rates of low-skilled individuals particularly hard, and that these individuals have not been able to recover afterwards.

### 3.1 The compositional effect

#### 3.1.1 Father’s income

The question that arises from investigating these figures is why there is a decline in the participation for low-skilled individuals, males in particular. One clue that could answer this question is acquired by looking at the average father’s decile in conjunction with the graduation rate:

![Figure 5 – Graduation rates and father’s decile](image)

Recall that percentiles are constructed separately for each cohort, i.e. the average percentile is equal to 50 for any cohort. There is, however, an obvious decline in the average percentile for
both graduates and non-graduates. The reason for this has to be a flow of individuals between these two groups, i.e. a compositional effect. It is also clear that the change in average father’s income is largest for graduates. This suggests that the group that has flowed into graduation of upper secondary education are relatively average compared to non-graduates, but with a social background poorer than the average graduates.

We can break down the distribution of father’s income even further, as in Figure 6 and Figure 7. The most noticeable observation is that there is a clear and persistent selection into graduation related to social background. Secondly, the share of low-skilled individuals that have fathers around the lower spectrum in terms of income seems to increase over time. That is, there is clear evidence of a change in composition.

Figure 6 – Distribution of father’s income, low-skilled males
The distributions of mother’s income is similar to that of father’s income, i.e. upper secondary graduates are generally from families where mothers have higher income, and the concentration of individuals with poor mothers increases for nongraduates between the 1963 and 1983-cohorts. The graphs of mother’s income are therefore omitted.

### 3.1.2 Trend in IQ-test score

Since the number of missing IQ-test scores are large and vary over time, I have included them in the following graphs. Three facts can be observed from graphs of IQ-test scores over time. First, the individuals that graduate from upper secondary education clearly performs better on the IQ test. Secondly, the large and varying number of missing observations makes it difficult to say anything about the development of IQ-test score over time. The peak in the number of missing observations around the 1973-cohort is particularly interesting, although I am not aware of the reason for why there are so many missing observations for these cohorts. Lastly, the share of missing observations are clearly larger in the group of dropouts than graduates, a point I will get back to in the section on missing observations below.
Figure 8 – IQ-test scores – low-skilled males

Figure 9 – IQ-test scores. High-skilled males
### 3.1.3 Sectors of employment

If there has been a deterioration of the labor market conditions for low-skilled individuals over the period between 1993 and 2013, the years of observed participation, this should be connected to the sectors it is most likely that they will find work. The table below has been constructed by using numbers from the annual national accounts for 2013 (Statistics Norway 2016 (b)). The time series for employment broken down on level of education, gender and sector is unfortunately not available before 2008. I have therefore not been able to find comparable numbers from the first cohorts in the period.

#### Table 2 – Sectors of employment

<table>
<thead>
<tr>
<th>Sector of employment (percent)</th>
<th>Group</th>
<th>Low-skilled males</th>
<th>Medium-skilled males</th>
<th>Low-skilled females</th>
<th>Medium-skilled women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture, forestry and fishery</td>
<td>5,8</td>
<td>4,3</td>
<td>2,0</td>
<td>1,7</td>
<td></td>
</tr>
<tr>
<td>Secondary sector</td>
<td>33,0</td>
<td>39,8</td>
<td>8,5</td>
<td>8,6</td>
<td></td>
</tr>
<tr>
<td>Private services¹</td>
<td>47,6</td>
<td>41,2</td>
<td>46,7</td>
<td>37,7</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>1,3</td>
<td>1,7</td>
<td>3,8</td>
<td>4,9</td>
<td></td>
</tr>
<tr>
<td>Health and social services</td>
<td>6,6</td>
<td>4,9</td>
<td>29,4</td>
<td>35,5</td>
<td></td>
</tr>
<tr>
<td>Personal services</td>
<td>2,5</td>
<td>2,3</td>
<td>4,5</td>
<td>4,9</td>
<td></td>
</tr>
<tr>
<td>Public administration, defense and social insurance</td>
<td>3,2</td>
<td>5,8</td>
<td>5,0</td>
<td>6,5</td>
<td></td>
</tr>
</tbody>
</table>

¹Retail, hotels, restaurants, transport, financial services, business finance and real estate.

The sectors are grouped similar to the grouping of sectors in (Statistics Norway 2016 (c)),

The two main sectors that employed males in 2013 were the secondary sector and private services. Males without upper secondary education were more seldom working in the secondary sector, and conversely more often working in private services. The greatest difference is however between the employment patterns of males and females. Women are much less often employed in the secondary sector, and more often employed in health and social services.

The gender gap in employment patterns may be the reason why the changes in participation rates may have affected males and females differently. Since sector of employment does not include individuals outside the labor market, these data will not be utilized in the main analysis in this thesis. I will however discuss these numbers in combination with other theories that may explain the decline in participation in the next chapter.
### 3.2 Missing Observations

Although the identified population of 163,492 low-skilled Norwegians males constitutes the population of interest, missing observations will narrow the population further. These missing observations will pose three potential problems. First, I will not be able to perfectly observe or compare the composition of individuals over time. Secondly, the decline in participation may be different after removing individuals with missing observations, and hence the participation rate investigated in the analysis may not be representative for the decline in the whole population of low-skilled males. Lastly, and more problematic: observability might be correlated with some unobserved variables. If this is true, it may render two groups that are identical in terms of observed variables incomparable in terms of unobserved variables. I will try to address these potential problems after presenting the number of missing observations.

First, there are approximately 2,449 males with missing level of education. That constitutes around 0.4 percent of the total population of males born in the period between 1963 and 1983. These numbers decrease from around 1 percent for the cohorts born in the sixties to around 0.2 for the cohorts born in the beginning of the eighties. Although there are some downward trends in these data, I do not believe they possess much of a threat, simply because almost every individual has observed educational attainment.

Parental income is missing for 4.5 percent of low-skilled males, i.e. 4.5 percent of low-skilled males have either missing father’s income or missing mother’s income. There is a slight upward trend of missing parental income, with around 3 percent missing in the first cohorts and around 7 percent towards the end. Parental income is only missing if parents are born too early or too late, or if the id-number for mothers or fathers are missing. That is, these observations are missing for individuals with unknown fathers, individuals that are born in the last part of the sample period and that have particularly young parents, and individuals born in the sixties with old parents. These variables might themselves be correlated with social background and other relevant characteristics, and thus the observability of these variables pose a potential problem.

Having a look at the trend in missing parental income, as well as participation rates by low-skilled males with missing and non-missing parental income reveal two things. First, low-skilled individuals with non-missing parental income have a higher participation rate in every year except for one year, so there is a clear correlation between observable parental income
and participation. Secondly, the correlation between observability and participation seems to be persistent over the period, so the presence of missing parental income may not be such a big problem when we compare different cohorts, although this result is far from conclusive.

![Participation Rate](image)

Figure 10 – Missing parental income

IQ-test scores on the other hand, are missing for as many as 13.5 percent of low-skilled individuals. Figure 11 reveals that there are two mayor spikes of missing observations, one in 1973-1974 and one during the last two years in the sample period. Low-skilled individuals with unobserved IQ-test scores have a participation rate that is around 20 percentage points lower than low-skilled males with observed IQ-test scores, i.e. the correlation between observed IQ-test scores and participation is high. Lastly, the correlation seems to increase over the sample period, suggesting that the process that gives missing IQ-test scores at the beginning of the period may differ from the process that gives missing observations at the end of the period. Although I do not know the reason for why the number of missing observations changes with cohorts, one may think that for example the consequences to non-participation have changed. This is one of the reason why I have not compared the distribution of IQ-test scores in different cohorts: One cohort of low-skilled males may have a higher IQ-test score
than another cohort, simply because we observe IQ-test scores from different parts of the cohorts.

![Participation Rate](image)

Figure 11 – Missing IQ-test scores (assumed ability level omitted).

Lastly, comparing the participation rate of low-skilled males with observed IQ-test scores to those with assumed ability levels some systematic differences: While around 84 percent of those with IQ-test score equal to two participates in the labor force, around 80 percent of those with assumed ability above average participates. Those with assumed ability level systematically participates less often than those with an IQ-test score above the lowest level. I therefore conclude that those with assumed ability level at some level is not necessarily comparable to individuals with observed IQ-test score at the same level, and omit the former group in the analysis.

It is difficult to say whether missing observations makes comparison challenging, and I will not investigate this matter further. I will however note that the correlation between observability and participation as well as the quite large share of missing observations is a potential threat to the comparability of cohorts. After removing those individuals with missing observations, we are left with a non-random sample that will be analyzed in later chapters.
Some of the descriptive statistics that have been introduced so far for the whole population, will be replicated for the reduced sample in the next chapter.

### 3.3 Summary Statistics

Summary statistics for the remaining population of low-skilled males with non-missing observations are reported in Table 3 below. The total decline in participation for this population is still around 10 percentage points, so there is still a lot of variation that I will try to explain in the main analysis. Further, the number of individuals are reduced from around 10,000 individuals to almost 3500 for the last cohorts. In addition, there are two spikes of the number of missing observations: One in the cohorts 1972 and 1973, and a second spike for the last two-three cohorts. The last spike parallels a small increase in parent’s income percentile. This latter increase disappears if one does not include individuals with missing IQ-test score (not included), and is therefore a product of the increase in the number of missing observations towards the end of the sample-period.
Table 3 – Summary statistics for individuals with non-missing observation

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Participation rate</th>
<th>IQ-test score</th>
<th>Income percentile</th>
<th>Age</th>
<th>Income percentile</th>
<th>Age</th>
<th>Parental education&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Remaining share&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Number of individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1963</td>
<td>90.02</td>
<td>3.85</td>
<td>41.20</td>
<td>30.83</td>
<td>45.83</td>
<td>27.10</td>
<td>57.68</td>
<td>84.57</td>
<td>10.237</td>
</tr>
<tr>
<td>1964</td>
<td>90.19</td>
<td>3.83</td>
<td>41.36</td>
<td>30.66</td>
<td>45.23</td>
<td>26.97</td>
<td>58.21</td>
<td>85.85</td>
<td>10.525</td>
</tr>
<tr>
<td>1965</td>
<td>90.16</td>
<td>3.83</td>
<td>41.18</td>
<td>30.46</td>
<td>45.68</td>
<td>26.79</td>
<td>60.28</td>
<td>86.88</td>
<td>10.098</td>
</tr>
<tr>
<td>1966</td>
<td>89.68</td>
<td>3.87</td>
<td>40.16</td>
<td>29.97</td>
<td>44.92</td>
<td>26.41</td>
<td>60.54</td>
<td>87.52</td>
<td>9.833</td>
</tr>
<tr>
<td>1967</td>
<td>90.03</td>
<td>3.90</td>
<td>40.85</td>
<td>29.72</td>
<td>44.93</td>
<td>26.23</td>
<td>61.99</td>
<td>86.78</td>
<td>9.047</td>
</tr>
<tr>
<td>1968</td>
<td>88.62</td>
<td>3.88</td>
<td>40.34</td>
<td>29.40</td>
<td>44.47</td>
<td>25.92</td>
<td>63.45</td>
<td>85.98</td>
<td>8.531</td>
</tr>
<tr>
<td>1969</td>
<td>87.74</td>
<td>3.88</td>
<td>40.68</td>
<td>29.28</td>
<td>44.03</td>
<td>25.96</td>
<td>65.41</td>
<td>85.98</td>
<td>8.129</td>
</tr>
<tr>
<td>1970</td>
<td>87.34</td>
<td>3.86</td>
<td>40.42</td>
<td>28.93</td>
<td>43.87</td>
<td>25.73</td>
<td>66.86</td>
<td>84.07</td>
<td>7.425</td>
</tr>
<tr>
<td>1971</td>
<td>86.23</td>
<td>3.88</td>
<td>40.52</td>
<td>29.04</td>
<td>43.58</td>
<td>25.76</td>
<td>68.19</td>
<td>82.49</td>
<td>6.988</td>
</tr>
<tr>
<td>1972</td>
<td>85.07</td>
<td>3.88</td>
<td>40.31</td>
<td>28.73</td>
<td>43.32</td>
<td>25.50</td>
<td>69.66</td>
<td>82.36</td>
<td>6.451</td>
</tr>
<tr>
<td>1973</td>
<td>84.02</td>
<td>3.93</td>
<td>39.77</td>
<td>28.58</td>
<td>42.79</td>
<td>25.46</td>
<td>71.64</td>
<td>73.97</td>
<td>5.155</td>
</tr>
<tr>
<td>1974</td>
<td>81.16</td>
<td>3.98</td>
<td>39.54</td>
<td>28.38</td>
<td>43.18</td>
<td>25.25</td>
<td>73.46</td>
<td>77.04</td>
<td>5.011</td>
</tr>
<tr>
<td>1975</td>
<td>79.93</td>
<td>4.06</td>
<td>39.87</td>
<td>28.37</td>
<td>42.43</td>
<td>25.28</td>
<td>74.85</td>
<td>80.86</td>
<td>4.743</td>
</tr>
<tr>
<td>1976</td>
<td>80.83</td>
<td>4.02</td>
<td>39.46</td>
<td>28.44</td>
<td>41.73</td>
<td>25.36</td>
<td>76.72</td>
<td>81.16</td>
<td>4.433</td>
</tr>
<tr>
<td>1977</td>
<td>79.76</td>
<td>4.00</td>
<td>39.48</td>
<td>28.52</td>
<td>42.03</td>
<td>25.38</td>
<td>75.65</td>
<td>82.37</td>
<td>4.205</td>
</tr>
<tr>
<td>1978</td>
<td>81.18</td>
<td>4.08</td>
<td>40.10</td>
<td>28.76</td>
<td>42.52</td>
<td>25.62</td>
<td>78.07</td>
<td>82.16</td>
<td>4.282</td>
</tr>
<tr>
<td>1979</td>
<td>79.80</td>
<td>4.03</td>
<td>39.90</td>
<td>28.75</td>
<td>42.39</td>
<td>25.60</td>
<td>77.37</td>
<td>81.58</td>
<td>4.485</td>
</tr>
<tr>
<td>1980</td>
<td>75.95</td>
<td>4.04</td>
<td>39.58</td>
<td>28.69</td>
<td>42.53</td>
<td>25.46</td>
<td>75.90</td>
<td>81.26</td>
<td>4.145</td>
</tr>
<tr>
<td>1981</td>
<td>75.97</td>
<td>3.98</td>
<td>39.46</td>
<td>28.94</td>
<td>42.61</td>
<td>25.68</td>
<td>76.40</td>
<td>78.22</td>
<td>4.029</td>
</tr>
<tr>
<td>1982</td>
<td>77.52</td>
<td>3.98</td>
<td>39.98</td>
<td>29.02</td>
<td>41.83</td>
<td>25.79</td>
<td>77.65</td>
<td>74.02</td>
<td>3.910</td>
</tr>
<tr>
<td>1983</td>
<td>79.28</td>
<td>3.94</td>
<td>40.25</td>
<td>29.29</td>
<td>43.11</td>
<td>26.02</td>
<td>77.13</td>
<td>68.00</td>
<td>3.524</td>
</tr>
<tr>
<td>Total</td>
<td>85.57</td>
<td>3.91</td>
<td>40.39</td>
<td>29.41</td>
<td>43.92</td>
<td>26.03</td>
<td>67.26</td>
<td>82.69</td>
<td>135.186</td>
</tr>
</tbody>
</table>

All reported statistics are means except for the number of individuals.

<sup>1</sup>The share of individuals with at least one parent with graduation from upper secondary education

<sup>2</sup>Remaining share is the share of low-skilled males that have no missing observations and are included in this table.
4 Theoretical and Empirical Background

Before I move on to the analysis of the compositional effect, I will give a brief survey of some existing theories and observations that may be relevant in explaining the reduction in participation for low-skilled individuals. This is not supposed to be a comprehensive survey of every possible reason for the observed decline, but rather a brief review of some theories that have been important in economics or the public debate.

4.1 Skill Biased Technological Change

The theory of skill biased technological change is one of the most significant theories of relative demand for low- and high-skilled labor. Albeit technology might be complimentary to both high- and low-skilled workers, whenever technology is more complimentary to high-skilled human capital, technological innovation and capital accumulation will shift the demand towards workers with a higher skilled set over time. The book The Race Between Education and Technology (Goldin and Katz 2009) examines the case of the United States (US). The authors show how the relationship between technological development and shifts in relative labor supply can explain much of the evolution of relative wages throughout the twentieth century. Moreover, they find that while the skill biased technological change has been more or less constant for much of the twentieth century, the primary source of changing wage structure is that the supply of relative skill has not been able to keep up with the pace of technological growth. That is, the stagnation of the level of high school graduation goes a long way in explaining the increased wage disparities beginning in the eighties.

The stagnation of graduation levels in the US during the eighties is comparable to the stagnation in the Norwegian graduation trends during the nineties, and has potential to explain a possible change in the labor market participation of low-skilled individuals. Notice, however, that the implication of such mechanisms should be that the greatest change in the labor market should be located after the stagnation of graduation rates, i.e. after the supply of relative skill were no longer able to keep up with the technological evolution. However, Figure 3 (p. 13) reveals a different story: Most of the decline in participation where
happening at the same time as graduation rates continued to increase, with an exception for the sharp decline after the Great Recession.

In recent decades, a lot of attention have been focused on the impact of the digitalization of the modern-day economy. Author and Dorn (2013) finds that technological growth in the nineties differs from previous development in the US. They show that there has been a rapid growth in the demand for labor in the top and bottom of the skill-distribution, and a decline in the demand for medium-skill labor. This decline is explained as a consequence of digital technology being good at performing routine tasks. Non-routine tasks, on the other hand, still require a significant contribution from human labor. The result of these mechanism is a decline in manufacturing jobs relative to service sector jobs. Service sector jobs being for example janitors, health service, child care, and hairdressers.

Juhn and Potter (2006) examines the participation rates of males and females in the context of the US Labor market. They find that the skill-polarization that is also described by Author and Dorn (2013) seems to have favored the comparative advantage of female labor, relative to males. Further, Weinberg (2000) have found empirical evidence that investment in information technology at the industry level is positively correlated with demand for female labor.

The changes in demand for labor across sectors that is unraveled by Juhn and Potter (2006) and Author and Dorn (2013) may relate to the Norwegian labor market as well. Even though we do not observe what sectors individuals worked in before 2008, the share of males working in manufacturing far exceeded the share of females, and the contrary was true for the health sector and personal services (Table 2, p. 18). Hence, it is possible that there has been a shift in employment from sectors dominated by males to sectors dominated by women. On the other hand, there has not been an increase in the number of low-skilled women in the labor force (Figure 1 and Figure 4, pp. 11 - 13). It has, however been an increase in the number of women with graduation from upper secondary education in the labor force, so a combination of a gender-bias and skill-bias of technological change. Table 2 (p. 18) is unfortunately not finely tuned enough to say something about the employment patterns of low- and medium skilled males in relation to the change in the labor market described by Author and Dorn (2013).
4.2 Immigration

The Norwegian labor market experienced a rapid increase of labor immigration after the expansion of the European Union in 2004 (Ødegård 2014): In 2004, about 4,100 non-Nordic individuals immigrated to Norway looking for work, and this number increased to around 21,400 in 2014 (Statistics Norway 2016 (a)). The total number of labor immigrants in the period between 2004 and 2014 was about 208,200, in addition to 290,800 individuals that immigrated due to other reasons than looking for work. Non-Nordic citizens are not required to register to stay in Norway, and are therefore omitted in these statistics.

How would we expect the rapid increase in immigration to affect the Norwegian labor market? This depends on to what extent immigrants are substitutes for different groups of native workers. In terms of individual characteristics, one might assume that Eastern European immigrants competed with young males higher up the skill spectrum: three quarters where males, and about 80 percent where below the age of 40. In addition, as few as 12 percent had education below the upper secondary level, and 33 percent had some form of higher education. When we take into consideration that 17 percent of these immigrants’ education were unobserved, these numbers becomes less reliable.

That a lot of Eastern European immigrants have graduated from high school does not necessarily imply that they do not compete with low-skilled Norwegian males: Labor immigrants may not compete with high- and medium skilled Norwegians due to language barriers or other skill requirements. To see who these immigrants might compete with we may investigate what type of jobs they usually have, because Eastern European immigrants seem to have been highly concentrated around a few sectors in the economy. Around 50 percent worked in the secondary sector of the economy (e.g. construction and industry), while about 40 percent worked in service sector jobs like hotels, restaurants and retail (Statistics Norway 2016 (d)). These sectors employed a large share of both low-skilled and high-skilled Norwegian males around 2013 (Table 2 p. 18), so there is no a priori reason why they should compete more with either. Since women works in the secondary sector much more seldom than males, it may contribute to why participation rates of low-skilled women have declined less than their male counterparts.

Ødegård (2014) Investigates the impact of Eastern European immigration on the labor market associated with the shipbuilding industry, hotels, fishery and meat production: Four sectors
where a lot of Eastern Europeans have found work. She finds that the impact of labor immigration has led to more usage of temporary labor contracts, as well as less local recruitment and on-the-job training. These factors may have created more difficulties for native youth to find jobs, as well as discourage youth from looking for jobs in these sectors. On the contrary, she notes that these effects may not have a large impact on the short run, due to the low unemployment rates in Norway during the mid-nineties.

As in the last section I conclude that labor immigration may have caused the relative changes in participation between males and females, but that the impact on the relative labor market condition for males with and without upper secondary education is inconclusive.

### 4.3 Trade

Until recently, most of the empirical investigations of labor market effects of trade shocks concluded that such effect did not have a large impact on wages (Balsvik, Jensen et al. 2015). A recent paper by David, Dorn et al. (2013), however, estimates the effects of a significant rice of import from China on the US labor market. They find that this increase of trade between 1992 and 2007 had a significant effect on workers that worked in sectors that were exposed to trade from China. More specifically, they find that workers that worked in these sectors gained lower wages, and that these workers were more likely to drop out of the labor market.

Performing a similar analysis on the Norwegian labor market between 1996 and 2007, Balsvik, Jensen et al. (2015) finds that these effects are far smaller in the Norwegian labor market than the results in David, Dorn et al. (2013). This is mainly because the introduction of China to the world markets have increase the demand for Norwegian export goods. In addition, imports from China are more often used as intermediate inputs in manufacturing rather than as final goods. This results in reduced costs for importing firms, as well as the increased competition to Norwegian firms that produce these same intermediate inputs. Although they do not find an effect on wages, they do conclude that increased imports from China has pushed low-skilled workers into unemployment or out of the labor force.

One important aspect of the data reported by Balsvik, Jensen et al. (2015), is that most of the increase in imports from China occurred after 2001. That is, this trade shock is unlikely to have had a severe impact on the measured labor market participation for cohorts born before
1971, reported in chapter 3. On the other hand, if exports from China competed with Norwegian export goods before this, then it can also have had an effect of cohorts born before this.

In this chapter I have reviewed some of the literature on skill biased technological change, labor immigration and trade. I have also tried to relate these theories and associated empirical observations to the trends observed in chapter 3, and evaluated their potential relevance in the setting of declining participation rate of low-skilled males. Even though this review is not supposed to be a comprehensive review or rigorous analysis of these theories, I have showed that some aspects of these theories may explain at least part of the observed decline. I will exclusively focus on the compositional effect in the next chapter, and try to decompose the change in participation into the effect of compositional change and other effects. The change in participation that remains after controlling for the compositional effect is due to something else, of which the theories reviewed in this chapters are candidates.
5 Analysis

In this chapter I will use two different methods to decompose the decline in participation of low-skilled males into a compositional effect and other effects. The first approach will utilize a series of matching estimators to generate cohorts that have approximately identical distribution of observed characteristics. Comparing the change in participation of these cohorts will give a trend in participation where the compositional effect is mitigated. The latter approach abstracts from observed characteristics to avoid problems with missing observations and unobserved characteristics. This approach will let us say something about the feasible size of the compositional effect under quite weak assumptions.

5.1 Matching: Controlling for Observed Characteristics

The goal of the following sections is to introduce the matching estimator, assess its performance, and report the results of the matching approach. I will begin by specifying the underlying function that I am trying to approximate, before I elaborate on the limitations of this approach. Afterwards I will formally introduce the matching estimator, and explain how matching in this setting relates to the usual setting of estimating causal effects. I will then go on to discuss how to assess the performance of such estimators, as well as to discuss what characteristics to include in the matching approach. After that, I will discuss how to perform inference, before I report the results and robustness tests. Lastly, I will repeat the estimation with different sets of covariates, to see if the results are consistent across different inclusion of individual characteristics. I will base most of the analysis on a survey paper by Stuart (2010), which covers a wide variety of matching estimators, model selection, and diagnostics.
5.1.1 The Estimand

I am going to start off with an analogy to simplify the forthcoming discussion. Say that individual characteristics are only described by social background, and that social background is either poor or good. The share of each of these two groups may be exactly 50% in each cohort, but may vary among non-graduates due to changes in the selection into higher levels of education. Social background will be represented by the variable $X_i$ that takes the value one if individual $i$ has a good social background, and zero if not. Let labor market participation of any upper secondary dropout (non-graduate) be represented by the dummy variable $LM_{do}$ that is equal to one if that individual participates in the labor force. If we draw a non-graduate at random, the probability that this individual participates in the labor force will be equal to the participation rate, and affected by the likelihood that this individual has a good social background:

$$Pr(LMP_{do} = 1) = Pr(LMP_{do} = 1|X = 1) Pr(X_{do} = 1)$$
$$+ Pr(LMP_{do} = 1|X = 0) Pr(X_{do} = 0)$$

If we draw two individuals from two different cohorts at random, the probability that each of them participate might differ because there might be a higher share of individuals with good social background in one cohort than the other: A compositional effect. On the other hand, there might be some other factors that are different in the two cohorts, and this might affect the probability of participation given social background. The role of the matching estimator in this setting would be to construct cohorts where the share of individuals with poor and good social background are identical, and hence mitigate the compositional effect when comparing the participation rate of these two cohorts.

To generalize this to a setting where $X$ might contain more than one variable and possibly continuous variables, let $f_{X|do,y}(x)$ be the joint density function of the variables in $X$ over non-graduates in cohort $Y = y$. It is straightforward to show that the expected participation of any dropout is a function of the joint density function and the conditional expectation:

$$E(LM_{do,y}) = E[E(LMP_{do,y}|X)] = \int f_{X|do,y}(x) E[LM_{do}|X = x, Y = y] dx$$

Where the integral is over the support of $X$. The goal of the matching strategy is to fix the distribution of $X$ for several cohorts, such that the only source of changing participation is due
to any variables that are free to vary within the conditional expectation. That is, I will choose a “reference cohort” \( y' \) and a “source cohort” \( y \), and then construct a sample of the source cohort that has similar characteristics as the reference cohort. To accomplish this I will estimate what Frölich (2007) denotes the \textit{covariate-distribution adjusted mean}:

\[
\int f_{X|do,y'}(x)E[LMP_{do}|X = x, Y = y]dx \ , \ \forall y \neq y'
\]  

(5.1)

This expression can be interpreted as what the expected participation rate of dropouts in the source cohort would have been if their covariates where distributed as in the reference cohort. The interest in this thesis is not the covariate-distribution adjusted mean on its own, but on how much of the decline that is left after controlling for the change in composition. That is, the interest is in the difference between equation (5.1) and the participation rate in the reference cohort:

\[
E[LMP_{y'}] - \int f_{X|do,y}(x)E[LMP_{do}|X = x, Y = y]dx \ , \ \forall y \neq y'
\]  

(5.2)

This is the change in participation after controlling for the compositional effect. The residual change in participation between these two cohorts will be due to other effects than changed composition: If we observe a large shift in participation after fixing the distribution of covariates, the conclusion will be that there are other factors that explain the decline in participation over time. I will later refer to (5.2) as the \textit{composition fixed cohort effects} for intuition, and refer to equal distribution of covariates as balance.

The first term in (5.2) can be estimated using the observed participation rate of that cohort without much problem. Frölich (2007) shows that the last term (the covariate-distribution adjusted mean) can be estimated using matching in general, and propensity score matching more specifically. Before I move to the practical implementation of such estimators, however, I will have a look at some possible limitations of estimating the last term.

\subsection*{5.1.2 Limitations}

Due to the limited data set at hand, there are two main limitations that need to be addressed when we estimate (5.1): First, we face a setting with quite large number of missing observations in the dataset, and secondly, we only have data on a subset of the relevant individual characteristics.
All individuals with missing characteristics are dropped from the sample when we implement the matching estimator. Thus, we are only able to observe the compositional effect for a subsample of the whole population. If the observed sample is representative for the whole population, as if observability where random, this would not be a problem. This is because we could interpret the results as being representative for the whole population. As we have seen, however, this is not the case: observability of IQ-test scores is highly correlated with both graduation from upper secondary education and participation in the labor force (see chapter 3.2), and may also be correlated with unobserved characteristics.

To what extent is this a problem? As we have seen, there is a large decline in the participation rate in the subsample with non-missing variables, and these individuals make up around 80 percent of the population of low-skilled males (Table 3, p. 23). The results from the analysis of this subset is just as valid as far as we are able to balance all observed and unobserved characteristics. There are however no reasons for why we should conclude that these findings generalize to the resulting 20 percent of low-skilled males that are dropped due to missing observations.

The second problem is that there are probably some individual characteristics that are relevant for participation that are not included in the data set. When we balance the cohorts with respect to the observed characteristics, these characteristics are also free to vary within the sample. Some candidates for such variables are motivation, health and preferences. In general, I will only be able to observe the compositional effect with respect to observed characteristics. It is, however, possible to show that some variables can be balanced over time, even if they are unobserved. To see the condition that must hold for an unobserved variable to be balanced after matching on observables, let $X$ be the vector of observed characteristics, and $Z$ any single unobserved variable (proof in appendix A):

$$f_{Z|X,do,y'}(z|x) = f_{Z|X,do,y}(z|x)$$

That is, if the distribution of $Z$ given $X$ is the same in the two groups, the matching algorithm will balance $Z$ even if it is unobserved. We might expect this to be true if there is a strong and persistent correlation between the unobserved variable and the observed covariates, or if the distribution of $Z$ is independent of $X$ but identical in the two cohorts. If, on the other hand, an unobserved variable is uncorrelated with $X$, and the distribution of that variable is different in the groups we are comparing, that variable will be unbalanced in the two samples.
One aspect that decreases the likelihood of (5.3) holding is that we remove individuals with missing observations. It is therefore not sufficient for the condition to hold in the whole population, but rather in the reduced sample of individuals with non-missing observations. If $Z$ is correlated with observability, then the condition might fail to hold even if it does hold in the whole population.

### 5.1.3 Defining the Matching Estimator

The typical application of matching is to estimate causal effects (Stuart 2010), and the language associated with the method is also related to causal effects. In this section I will therefore briefly introduce the usual implementation of matching, and relate that to the implementation that I will use.

The theoretical background of matching usually starts with some treatment variable $T_i$ that usually takes on the value one or zero for treated and untreated individuals. This treatment variable is suspected to have a causally effect on an outcome variable $Y_i$, and $Y_i$ may be correlated with some other variables. The set up usually includes the concept of potential outcomes, i.e. for every individual $i$ there are two distinct outcomes, and the one that is realized is determined by the treatment:

$$Y_i = Y_i(0) + T_i(Y_i(1) - Y_i(0))$$

The causal effect of $T_i$ on $Y_i$ is defined as the difference between $Y_i(1)$ and $Y_i(0)$, and may be different for different individuals. The “fundamental problem of causal inference” (Stuart 2010) arises from the fact that any individual either receives treatment or no treatment, and therefore only one of these potential outcomes are realized. The unobserved potential outcome must therefore be estimated, and the matching estimator is one such estimator.

Matching estimators approaches this problem by matching any treated individual to one or more untreated individuals with similar characteristics, and vice versa for untreated individuals. If treatment is uncorrelated with potential outcomes conditional on characteristics, the difference in expected outcome for any matched group will only differ because of treatment or some remaining imbalance of characteristics. The estimated causal effect on individual $i$ is therefore the difference between its outcome and the weighted average outcome of $i$'s matched counterparts. The average of these estimates can be used to
construct the average causal effect on the treated, the average causal effect on the non-treated, or the average causal effect on the whole population by taking averages over different treatment groups. The practical implementation of such estimators typically involves assigning weights via the matching algorithm, such that the weighted distribution of covariates in the treated and untreated sample are approximately balanced.

In the setting of this thesis we are trying to balance covariates across cohorts. The “treatment” variable is analogous to birth year, and we are trying to balance the distribution of covariates over several cohorts. That is, we need to choose one “reference cohort” and one “source cohort”, and then find individuals in the source cohorts that are similar to the reference cohort. Since the data set will consist of 20 cohorts, we have one reference cohort and 19 source cohorts, so this two-cohort comparison must be repeated for every source cohort. After the balanced cohorts are generated, the composition fixed cohort effects in (5.2) are estimated by running a weighted regression of participation on cohort-dummies. The dummy for being in the reference cohort is going to be omitted, such that the remaining cohort dummies will represent the difference in participation to the reference cohort. I will also compare these results to the unweighted regression of participation to cohort dummies to say something about the magnitude of the compositional effect: The difference between these two measures will be equal to the compositional effect.

Matching usually relies on the assumption that potential outcomes are independent of treatment status conditional on the observed covariates, the conditional independence assumption (Angrist and Pischke 2008, chapter 3.2). Notice, however, that there is no reason to believe that birth year should have any direct causal effect on participation. The change in participation that is left after balancing individual characteristics will therefore be due to those variables that that creates the correlation between cohort affiliation and participation rate—which are the factors we are interested in. The conditional independence assumption is therefore not necessary in this application. This intuition is also confirmed by Frölich (2007).

Before we go on to implement these estimators there are several issues that needs clarification. First, there are several matching algorithm available, and little guidance on how to choose an optimal algorithm. Secondly, we need to assess to what extent out matching implementation are able to balance the covariates. Lastly, we need to clarify how we can do inference based on the matching results.
5.1.4 Choosing Matching Algorithm

One key aspect of the matching approach is that it may not be possible to generate perfectly balanced samples, as there may not be perfect matches for every treated individual. In such settings, matching involves a tradeoff between balancing different covariates, and different matching algorithms that are available approaches this problem differently. One problem is that there is little guidance on what matching algorithm that should be considered optimal (Stuart 2010). The way I have chosen to approach this problem is to report the results from four different estimators. This way I will ensure robustness with respect to what algorithm is used. That is, similar results will strengthen the conclusions drawn, while diverging results will make it more difficult to say something based on the approach.

The four algorithms I have chosen to use are nearest neighbor propensity score matching (with replacement), epanechnikov kernel matching on the propensity score, nearest neighbor mahalanobis metric matching and weighting by the odds. The last estimator is not a matching estimator as described above, but weights every source cohort to represent the reference cohort, and hence works similarly to matching (Stuart 2010).

Both nearest neighbor propensity score matching and kernel matching defines the distance between two individual in terms of their propensity score. The propensity score is, for every two-cohort comparison, the probability of being part of the reference cohort conditional on observed covariates. I will estimate the propensity score by using a logistic regression of birth year (indicator for being part of the reference cohort) on parental income, parental income squared, parents age, IQ-test score, and indicator variables that indicates that mother’s or father’s income is zero. Distance between any to individuals \(i \) and \(j \) are then defined to be \(|\hat{e}_i - \hat{e}_j|\), where \(\hat{e} \) is the estimated propensity score. Nearest neighbor matching matches every individual in the reference cohort to the closest individual in the source cohort, while kernel matching uses a weighted average of individuals in the neighborhood of the individual in the reference cohort.

Mahalanobis metric matching works the same way as nearest neighbor propensity score matching, but defines distance as \((X_i - X_j)\Sigma^{-1}(X_i - X_j)^\top\), where \(\Sigma\) is the full variance-covariance matrix of \(X\) in the source cohort. Here IQ-test scores are converted to a series of indicator variables, as recommended by Stuart (2010), while the square of parental income are omitted to avoid overfitting to parental income relative to the other covariates.
Lastly, weighting by the odds are implemented by weighting every individual in the source cohort by $\hat{e}_i/(1 - \hat{e}_i)$, where $\hat{e}$ are the same estimated propensity score as above. As with the other estimators the reference cohort are unweighted.

Nearest neighbor matching will be implemented by using the Stata-command `teffects`, while the user written command `psmatch2` (Leuven and Sianesi 2015) will be used to implement the mahalobis metric- and kernel matching.

One important aspect of matching is that we need more control variables than treated variables to be able to acquire sufficient balance (Stuart 2010). Hence, we need to choose a reference cohort towards the end of the sample. I will also try to choose a cohort that is as representative as possible, and hence avoid the last cohorts with especially many missing observations. The last cohort that does not have many missing variables is the cohort of 1980, and I will therefore use this as the reference cohort. and balance the covariates of earlier cohorts. The cohorts after 1980 are omitted since the change in participation and graduation are minuscule compared to the 1980-cohort. Notice that the estimators will be repeated for every cohort before the 1980-cohort, and are therefore not affected by omitting certain cohorts.

### 5.1.5 Inclusion of Individual characteristics

One important decision that needs to be addressed is what variables that should be included in the matching procedure. There are three reasons for why this is important. As described above, matching involves a trade off of between balancing different covariates. Hence, including additional variables might make it more difficult to balance other covariates. Secondly, I will address to what extent an observed variable reflects the compositional effect that we are interested in. Lastly, the population used in the matching approach will vary with variable inclusion, since those with any missing observations are dropped.

Clearly, talking about a compositional effect suggests that the set of individuals in one group is somehow different to the set of individuals in another group. In our setting, this can be due to one of two effects, or both. First, there might be intrinsic differences between any two cohorts. For example, the level of education of parents have increased over the time period we are looking at (see Table 3 p. 23). This, in turn, might make their children more or less likely
to finish high school or more or less likely to be a part of the labor force when they grow up. Secondly, a source of changing composition might be due to self-selection into higher levels of education.

The main motivation behind investigating the compositional effect was the effect of changed graduation, as described in chapter 3. That is, the main concern is about variables that reflects the relative distribution of individuals between the groups of graduates and non-graduates, rather than variables that change over time. This is reflected in the fact that parental income are relative measures within each cohort. In addition, the average IQ-level of cohorts are more or less constant over the period (see Sundet, Barlaug et al. (2004)), and are therefore comparable between cohorts. In contrast, the age of parents and their level of education changes over time. This raises the question whether two individuals with identical observations of parental education and age are comparable, or if these variables signal different types of parents across cohorts.

The result of these concerns is that I will not match on parental education, since this variable changes drastically over time (see Table 3, p.23). I have also decided to include the age of parents, since this variable may change since the income of young parents are missing in the first cohorts and the income of old parents are missing in the last cohorts (see section 2.3.2). Since these decisions may influence the results, I will also include the results from matching on alternative sets of covariates in chapter 5.1.11.

5.1.6 Assessing Balance

I will use two commonly used numerical diagnostics tools to assess how well the covariates are balanced before and after matching: the absolute standardized difference in mean (SDM) (also called the absolute standardized bias) and the variance ratio (VR) (Stuart 2010, McCaffrey, Griffin et al. 2013):

$$SDM: \left| \frac{\bar{X}_{y'} - \bar{X}_y}{\hat{\sigma}_{y'}} \right|$$

$$VR: \frac{\hat{\sigma}_{y}^2}{\hat{\sigma}_{y'}^2}$$
These will be estimated for every covariate, source cohort and matching algorithm. The samples are considered well balanced whenever the SDMs are close to zero and the VRs are close to one. A proposed rule of thumb is to consider two samples as balanced if the SDMs are smaller than 0.25 and the VRs are between 0.5 and 2 (Stuart 2010, McCaffrey, Griffin et al. 2013).

I will be using a variant of the SDM from McCaffrey, Griffin et al. (2013) to make sure that the SDMs are comparable for different cohort comparisons and matching algorithms. Their standardized difference uses the standard deviation from the reference cohort in the denominator. This measure is identical across different cohort comparisons and matching algorithms, and can therefore be used for comparison.

Another concern when implementing propensity score matching is whether there is sufficient overlap of the propensity score, a condition labeled common support (Stuart 2010). The condition is related to the more general concern regarding sufficient overlap of the covariates (Stuart 2010). When we perform the matching, there should be individuals in the source cohort with similar propensity score as every individual in the reference cohort to be able to construct comparable samples. Caliendo and Kopeinig (2008) shows that the condition can be summarized with the following condition:

$$\Pr(Y = y' | X) < 1$$

A usual solution to this problem is to drop individuals in the reference cohort with no source cohort individuals within some range of the propensity score, or to drop individuals with propensity score close to one. The drawback of these approaches are that the matched samples will be constructed against different reference cohorts, and different two-cohort comparisons may therefore not be comparable. I will instead keep every individual in the reference cohort, and report the maximum estimated propensity score for both cohorts, and assess whether overlap is plausible. I will consider any individual with propensity score above 0.8 as a violation of the common support, a boundary used in Caliendo and Kopeinig (2008).

### 5.1.7 Inference

To investigate whether the composition fixed cohort effects are significantly different from zero, we will need some estimation of the standard deviation of these estimates. These estimates will need to take account of several factors of the data and estimation procedure. In
this section I will account for how I will approach this problem and the choice of such estimators. I will also introduce a test of any compositional effect, that in essence tests the impact of the weights introduced by the matching estimators.

There are three concerns that needs to be addressed when standard errors are introduced. First, we cannot make the assumption that the residuals are uncorrelated to the regressors (homoscedasticity) when the outcome variable is a dummy variable (Angrist and Pischke 2008, Chapter 3). Secondly, there may be correlation between participation within areas of residence, if for example local labor market conditions affects participation, so clustered standard errors are likely to be appropriate (Angrist and Pischke 2008, Chapter 8.2). Lastly, the estimation of standard errors will need to take the into account that the matching algorithm and estimation of the propensity score when implementing matching (Stuart 2010). As far as I know, none of the analytic standard errors that have been developed take all of these concerns into account. One proposed solution is to bootstrap the standard error (Ibid.), but the size of the data set makes this infeasible: it would take between 10 and 50 days to iterate one matching estimator 50 times.

There are therefore no standard errors that are appropriate, as far as I know. I will instead compare different standard errors, and use the largest observed standard errors, such that the probability of a type 1 errors is as small as possible. That is, I will compare the analytical standard errors that are implemented in the mahalanobis metric matching (Abadie and Imbens 2006) and nearest neighbor matching (Ibid.) to the standard errors that are clustered on the county level, and report the standard errors that are largest. Kernel matching and weighting by the odds will be implemented with clustered standard errors only. It turns out that clustered standard errors are largest in every case. These standard errors take both clustering and the lack of heteroscedasticity into account, but treats the weights as exogenous.

I will also report a test of the opposite hypothesis. That is, I am going to test whether there is any compositional effect for any of the cohorts. This is equivalent to testing if the weights have any impact on the estimated participation rates. Such a test is developed in the literature on weighted analysis of survey data (Bollen, Biemer et al. 2016). To implement the test I will need to regress participation on dummy variables that indicate cohort affiliation:

\[ LMP_i = \alpha + \sum_j \beta_j \cdot \text{cohort}_i \cdot t_{i,j} \]
This will yield the participation rate if no weights are used, and the composition adjusted participation when weights are implemented. Let $\hat{\beta}$ be the vector of unweighted estimators, and $\hat{\beta}^{WLS}$ be the vector of composition adjusted participation. Fuller (2011) shows that the test of whether these to estimators have the same expectation can be performed by regression the outcome variable on the cohort-dummies and the interaction between these variables and the weights. The test will be a simple F-test of whether the coefficients on the weighted dummies are all zero.

### 5.1.8 Results

The results from the matching estimators are available in table 4. They are the remaining change in participation after removing the compositional effect from observed covariates. The results are also reported in graph 12-15, in which case the confidence intervals are equal to the composition fixed cohort effects plus or minus 1.96 standard errors. Hence, the change in participation is significant on the five percent level whenever the confidence interval does not cover the x-axis, which represents the (unweighted) participation rate in 1980.

There are several things to comment on these results. First, the change in participation are significantly different from zero for every cohort and with every matching algorithm. In fact, most of the composition adjusted participations are significantly different from zero on the one percent level. If we compare the unweighted change in participation with the matched change in participation, the matching algorithm seems to have made little difference to the participation rate, a result that is also consistent across all the different matching algorithms. Lastly, the test of any compositional effect gives varying results based on the matching estimator: The test is insignificantly different from zero for the mahalanobis estimator (on the five percent level), and significant on the one percent level for the weighting by the odds and kernel estimators.
<table>
<thead>
<tr>
<th>Cohort</th>
<th>Unweighted sample</th>
<th>Odds</th>
<th>Nearest neighbor</th>
<th>Mahalanobis</th>
<th>Epanechnikov</th>
</tr>
</thead>
<tbody>
<tr>
<td>1963</td>
<td>0.141</td>
<td>0.139***</td>
<td>0.135***</td>
<td>0.135***</td>
<td>0.141***</td>
</tr>
<tr>
<td></td>
<td>(0.0136)</td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.0139)</td>
<td></td>
</tr>
<tr>
<td>1964</td>
<td>0.142</td>
<td>0.143***</td>
<td>0.140***</td>
<td>0.152***</td>
<td>0.143***</td>
</tr>
<tr>
<td></td>
<td>(0.0134)</td>
<td>(0.0151)</td>
<td>(0.0151)</td>
<td>(0.0124)</td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>0.142</td>
<td>0.143***</td>
<td>0.142***</td>
<td>0.129***</td>
<td>0.143***</td>
</tr>
<tr>
<td></td>
<td>(0.0125)</td>
<td>(0.0138)</td>
<td>(0.0138)</td>
<td>(0.0126)</td>
<td></td>
</tr>
<tr>
<td>1966</td>
<td>0.137</td>
<td>0.141***</td>
<td>0.153***</td>
<td>0.147***</td>
<td>0.141***</td>
</tr>
<tr>
<td></td>
<td>(0.0138)</td>
<td>(0.0156)</td>
<td>(0.0156)</td>
<td>(0.0136)</td>
<td></td>
</tr>
<tr>
<td>1967</td>
<td>0.141</td>
<td>0.140***</td>
<td>0.147***</td>
<td>0.146***</td>
<td>0.142***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.0125)</td>
<td>(0.0125)</td>
<td>(0.0129)</td>
<td></td>
</tr>
<tr>
<td>1968</td>
<td>0.127</td>
<td>0.128***</td>
<td>0.140***</td>
<td>0.126***</td>
<td>0.130***</td>
</tr>
<tr>
<td></td>
<td>(0.0116)</td>
<td>(0.0119)</td>
<td>(0.0119)</td>
<td>(0.0119)</td>
<td></td>
</tr>
<tr>
<td>1969</td>
<td>0.118</td>
<td>0.120***</td>
<td>0.128***</td>
<td>0.122***</td>
<td>0.123***</td>
</tr>
<tr>
<td></td>
<td>(0.0144)</td>
<td>(0.0147)</td>
<td>(0.0147)</td>
<td>(0.0144)</td>
<td></td>
</tr>
<tr>
<td>1970</td>
<td>0.114</td>
<td>0.115***</td>
<td>0.120***</td>
<td>0.124***</td>
<td>0.119***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.0136)</td>
<td>(0.0136)</td>
<td>(0.0128)</td>
<td></td>
</tr>
<tr>
<td>1971</td>
<td>0.103</td>
<td>0.103***</td>
<td>0.107***</td>
<td>0.100***</td>
<td>0.105***</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.0111)</td>
<td>(0.0111)</td>
<td>(0.0118)</td>
<td></td>
</tr>
<tr>
<td>1972</td>
<td>0.091</td>
<td>0.093***</td>
<td>0.088***</td>
<td>0.102***</td>
<td>0.094***</td>
</tr>
<tr>
<td></td>
<td>(0.0117)</td>
<td>(0.0128)</td>
<td>(0.0128)</td>
<td>(0.0113)</td>
<td></td>
</tr>
<tr>
<td>1973</td>
<td>0.081</td>
<td>0.083***</td>
<td>0.088***</td>
<td>0.077***</td>
<td>0.084***</td>
</tr>
<tr>
<td></td>
<td>(0.0112)</td>
<td>(0.0138)</td>
<td>(0.0138)</td>
<td>(0.0107)</td>
<td></td>
</tr>
<tr>
<td>1974</td>
<td>0.052</td>
<td>0.054***</td>
<td>0.052***</td>
<td>0.063***</td>
<td>0.057***</td>
</tr>
<tr>
<td></td>
<td>(0.0122)</td>
<td>(0.0124)</td>
<td>(0.0124)</td>
<td>(0.0123)</td>
<td></td>
</tr>
<tr>
<td>1975</td>
<td>0.040</td>
<td>0.042***</td>
<td>0.046***</td>
<td>0.049***</td>
<td>0.046***</td>
</tr>
<tr>
<td></td>
<td>(0.0127)</td>
<td>(0.0174)</td>
<td>(0.0174)</td>
<td>(0.0127)</td>
<td></td>
</tr>
<tr>
<td>1976</td>
<td>0.049</td>
<td>0.048***</td>
<td>0.044***</td>
<td>0.047***</td>
<td>0.050***</td>
</tr>
<tr>
<td></td>
<td>(0.0135)</td>
<td>(0.0151)</td>
<td>(0.0151)</td>
<td>(0.0136)</td>
<td></td>
</tr>
<tr>
<td>1977</td>
<td>0.038</td>
<td>0.036***</td>
<td>0.049***</td>
<td>0.038***</td>
<td>0.035***</td>
</tr>
<tr>
<td></td>
<td>(0.0143)</td>
<td>(0.0144)</td>
<td>(0.0144)</td>
<td>(0.0136)</td>
<td></td>
</tr>
<tr>
<td>1978</td>
<td>0.052</td>
<td>0.053***</td>
<td>0.071***</td>
<td>0.060***</td>
<td>0.059***</td>
</tr>
<tr>
<td></td>
<td>(0.0105)</td>
<td>(0.0121)</td>
<td>(0.0121)</td>
<td>(0.0106)</td>
<td></td>
</tr>
<tr>
<td>1979</td>
<td>0.039</td>
<td>0.038***</td>
<td>0.033***</td>
<td>0.042***</td>
<td>0.038***</td>
</tr>
<tr>
<td></td>
<td>(0.0103)</td>
<td>(0.0146)</td>
<td>(0.0146)</td>
<td>(0.0104)</td>
<td></td>
</tr>
</tbody>
</table>

Test\(^1\)  
0.000  0.011  0.069  0.000

\(^1\)This is the p-value from the F-test of any compositional effect described in chapter 5.1.6. Test statistics are compared with the F distribution with 17 and 123.688 degrees of freedom.

\(^*\)p < 0.10, \(^**\)p < 0.05, \(^***\)p < 0.01

Standard errors in parenthesis. All standard errors are clustered on the county level. None of the standard errors take into account the estimation of the propensity score or other aspects of matching.
Figure 12 – Weighting by the odds

Figure 13 – Mahalanobis matching.
Figure 14 - Nearest neighbor matching

Figure 15 – Kernel matching.
5.1.9 Balance statistics

Before looking at the balance metrics, recall that balance is trustworthy whenever the standardized differences of means have an absolute value of less than 0.25 and the variance ratios are between 0.5 and 2. This is the case for every matching algorithm implemented. The balance statistics are also more balanced after matching than before, so the matching algorithms seem to be able to balance the characteristics between the cohorts. Notice also that parental wage and IQ-test scores are almost perfectly balanced when using weighting by the odds and mahalanobis matching. However, the required standardized difference in mean and variance ratio are mostly fulfilled for the unweighted cohorts as well. The exception is the age of parents, that trend downwards during the period (see Table 3, p. 23). That is, there does not seem to be much variation of observed characteristics in the sample, and this may be the reason why the matching estimators are unable to explain the decline in the participation rate.

The maximum estimated propensity scores that are reported in Table 5 are far below 0.8 that is required for common support. If we compare source and reference cohorts, it is also clear that the maximum propensity scores are higher for several of the source cohorts. Although it is not possible to draw firm conclusions based on these numbers, it seems plausible that there is simply not enough systematic difference in observed characteristics to estimate widely diverging propensity scores.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Source cohort</th>
<th>Reference cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>1963</td>
<td>0.65</td>
<td>0.56</td>
</tr>
<tr>
<td>1964</td>
<td>0.60</td>
<td>0.55</td>
</tr>
<tr>
<td>1965</td>
<td>0.63</td>
<td>0.60</td>
</tr>
<tr>
<td>1966</td>
<td>0.59</td>
<td>0.52</td>
</tr>
<tr>
<td>1967</td>
<td>0.58</td>
<td>0.53</td>
</tr>
<tr>
<td>1968</td>
<td>0.56</td>
<td>0.55</td>
</tr>
<tr>
<td>1969</td>
<td>0.55</td>
<td>0.52</td>
</tr>
<tr>
<td>1970</td>
<td>0.57</td>
<td>0.55</td>
</tr>
<tr>
<td>1971</td>
<td>0.57</td>
<td>0.56</td>
</tr>
<tr>
<td>1972</td>
<td>0.55</td>
<td>0.54</td>
</tr>
<tr>
<td>1973</td>
<td>0.55</td>
<td>0.55</td>
</tr>
<tr>
<td>1974</td>
<td>0.56</td>
<td>0.55</td>
</tr>
<tr>
<td>1975</td>
<td>0.57</td>
<td>0.59</td>
</tr>
<tr>
<td>1976</td>
<td>0.59</td>
<td>0.63</td>
</tr>
<tr>
<td>1977</td>
<td>0.58</td>
<td>0.62</td>
</tr>
<tr>
<td>1978</td>
<td>0.56</td>
<td>0.57</td>
</tr>
<tr>
<td>1979</td>
<td>0.53</td>
<td>0.53</td>
</tr>
</tbody>
</table>
Figure 16 – Standardized difference in means
Figure 17 – Variance ratio
5.1.10 Robustness tests

One of the weaknesses of the analysis so far is that the same covariates have been used in all matching algorithms, as well as the same estimations of the propensity scores. I will therefore perform a series of matching estimators that include alternative sets of covariates to investigate whether or not the results are robust to such decisions. I will use weighting by the odds for every such alternative specification, since this estimator seemed to perform well in terms of balancing the covariates. Another feature of using different specifications of control variables is that the number of individuals with missing observations will vary. Hence, the results will also reflect the effect of dropping different parts of the population, as described in section 5.1.2.

Thus far I have tried to balance parental wage, parental age, as well as IQ-test scores across cohorts. The alternative specifications reported in this section will include parental wage only, parental wage and IQ-test scores, and full set of covariates (parental wage, parental age, IQ-test scores and parental level of education). Since one concern is that parental age and education changes over time, I will also include the results of matching on the full set of covariates where these two variables are normalized. That is, the full set of relative covariates will include parental age and education after subtracting the cohort average of these variables.

The results from the robustness tests are included in figure 18, as well as Table 7 and Table 8 in appendix B. The results are almost identical as the former results, and are significantly different from zero with every implementation. Notice also that the composition fixed cohort effects are almost identical as the unweighted participation rates. This confirms our result that there is not much sign of a compositional effect in terms of observed variables.
Robustness Tests

Figure 18 – Robustness tests

5.1.11 Conclusion of the Matching Estimators

The matching estimators seems to be able to balance the covariates considerably compared to the distribution of covariates in the unmatched sample. Balancing the observed covariates does, however, not explain the decline in participation rate. There are several possible reasons why this is the case. First, one might conclude that there is not much sign of a compositional effect. But then the inability to explain the decline in participation may be because we only observe a subset of relevant individual characteristics. To conclude that there is no compositional effect would therefore to read too much out of the analysis. It is, however, possible to say that there is no compositional effect in terms of observed covariates. This is a rather surprising results, considering the evidence for a change in composition and that ability and social background are thought to be important predictors of labor market outcomes (Marks 2013, pp. 92-93)
5.2 Non-parametric Bounds

Although the results of the matching estimator are interesting, it only investigates the compositional effect in terms of observable characteristics for a subset of the population. To overcome these restrictiveness, I will utilize a more general analysis of the compositional effect in this subchapter.

The following analysis is similar to the non-parametric bounds analysis used to disentangle causal effects in for example De Haan (2008). The analysis will mostly be based on the theory of instrument variable analysis with heterogenous potential outcomes (Angrist and Pischke 2008, chapter 4.4). This is yet again an application of methodology that is associated with the estimation of causal effects, so I am going to spend some time to explain how I use this method. Fundamentally, the theory of instrument variables will be used for identification rather than estimation per se. The first section will be dedicated to introducing this methodology. I will subsequently use this method to estimate the compositional effect.

5.2.1 Methodology

The literature that tries to account for the negative correlation between employment levels and sickness absenteeism looks for a potential compositional effect in the labor force. In good times, a number of individuals flows into employment, and if these individuals have a particular high level of absenteeism, they might be accountable for the observed correlation. Using a similar line of reasoning lets us formulate an alternative to the approach in chapter 5.

First, note that the approach I am going to introduce will redirect our attention from the direct change in composition to the effects of increased graduation. More specifically, the motivation for the compositional effect was that some types of individuals graduated more often towards the end of the period we are investigating. This outflow of individuals might be accountable for both a change in the composition and the reduction of participation. What I will investigate in this chapter is what the participation rate of non-graduates might have been if there had not been an increase in graduation, effectively removing the compositional effect.

This line of reasoning relies on some correspondence between individuals in different cohorts. If we treat cohort affiliation as a set of instrument variables that pushes people into higher levels of education, we could say that the union of those who never graduates and those that
react to this instrument corresponds. If we could somehow remove this instrument, we would be left with corresponding individuals in every cohort.

To be more specific, let the cohort born in 1963 be the reference cohort that receives no instrument. Every subsequent cohort will be subject to some effects that encourages increased graduation, relative to the reference cohort. Some individuals would have graduated or dropped out regardless of birth year, while some individuals react to this “instrument”, and are pushed into graduation. In addition, one might think that some individuals drop out of upper secondary education because they are born later than 1963. I will henceforth denote the individuals based on the language in the literature on instrument variables (Angrist and Pischke 2008, chapter 4.4):

<table>
<thead>
<tr>
<th>Born in 1963:</th>
<th>Born in their respective year:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dropouts</td>
</tr>
<tr>
<td>Dropouts</td>
<td>Never-takers (nt)</td>
</tr>
<tr>
<td>Graduates</td>
<td>Defiers (d)</td>
</tr>
</tbody>
</table>

That is, instead of our former groups of dropouts and graduates, I will now operate with four groups. Let $\theta_i$ denote this group affiliation for individual $i$, such that $\theta_i \in \Theta = \{nt, at, c, d\}$.

Since we do not observe anyone’s graduation in the case where they are born in some other year, none if these groups are identifiable. If we assume that there are no never defiers, however, we know that every drop-out is also a never-taker. We can also deduce that every drop-out in the reference cohort is either a complier or a never-taker. In addition, the analysis is going to be based on the assumption that if we removed these cohort instruments, then we will be left with a similar set of cohorts, that are the same shares of the whole populations.

Is the assumption of no defiers a plausible assumption? First of all, if there is a process that jointly encourages graduation for some individuals and makes graduation more difficult for others, then the latter group will be defiers. This may be the case if the education system has become more selective while increasing the overall graduation rates. One can also imagine that an increasing number of individuals experience health issues that both decreases the
likelihood of graduation and participation in the labor force. This is the same kind of compositional change that affects whole cohorts, as described in the last chapter. On the other hand, the share of dropouts declines significantly over the time-period, so the share of defiers are probably much lower than any of the other groups. In essence, we need the dropouts in the source cohorts to be representative for a subset of the dropouts in the reference cohort, an assumption that is more likely.

It is straightforward to illustrate the three remaining groups with a graph of graduation rates if we assume that there are no defiers:

![Graph of Share of the Three Groups](image)

Figure 19 – Shares of Never-takers, compliers and Always-takers

Notice that the shares of compliers are equal to the change in graduation rates in relation to the reference cohort. What the instrument variable methodology has given us the opportunity to do is to identify two groups that corresponds across cohorts. That is, since the “cohort instrument” is what increases graduation rates, the joint group of compliers and never-takers corresponds when we compare different cohorts.

It is easy to show that the joint participation rate of never-takers and compliers in any cohort is equal to the weighted average of the participation of these two groups:
\[ E(LMP_{nt\&c,y}) = \Pr(\theta = nt)E(LMP_{nt,y}) + \Pr(\theta = c)E(LMP_{out,y}|e = do) \quad (5.4) \]

Where we implicitly condition on that the individual is either a never-taker or compliers. The condition \( e = do \) reflects that the interest is in the participation rate of compliers if they had no upper secondary education. The sample analogy of this equation is.

\[ \bar{LMP}_{nt\&c,y} = g_{nt,y}LMP_{nt,y} + g_cLMP_{c,y}^{do} \quad (5.5) \]

Where

\[ g_{i,y} = \frac{n_{i,y}}{n_{nt,y} + n_{c,y}}, \quad i \in \{nt, c\} \]

is the relative share of never-takers and compliers, and the superscript in (5.5) is the same as the condition in (5.4). The only unobserved variable to the right of the equality sign is the participation rate of compliers.

### 5.2.2 The Upper Bound

If the outflow of individuals were the only source of the change in participation, the joint participation rate of these two groups would be constant. Albeit we do not observe \( \bar{LMP}_{c}^{do} \), we can make different assumption regarding this counterfactual participation rate. To start off, observe that the highest possible joint participation rate would be if the participation of compliers were equal to one: \( \bar{LMP}_{c}^{do} = 1 \). Any observed participation above this level would be impossible under the no defiers assumption, while any joint participation rate under this level is feasible. To calculate this joint level, we simply execute (5.5) with \( \bar{LMP}_{c}^{do} = 1 \):

\[ \bar{LMP}_{nt\&c,y} = g_{nt,y}LMP_{nt,y} + g_cLMP_{c,y}^{do} \]
As we can see, a constant joint participation rate is possible in the beginning of the sample period, but not possible towards the end of the period. That is, the effect of an outflow of individuals can only account for the total decline in participation if there is a significant inflow of defiers.

### 5.2.3 Bounds of Joint Participation rate.

Albeit we have mapped out feasible joint participation rate, a full participation rate of compliers is fairly unlikely. To keep things simple, we could say that the participation rate is likely to be somewhere between the participation rate of actual dropouts and graduates. Likewise, we could assume that the participation rate of compliers where somewhere between the participation rate of graduates and full participation. Using this logic we could create three bands of assumptions regarding the participation rate:

Assumption 1: $\overline{LMP}_{do} \leq \overline{LMP}_c^d \leq \overline{LMP}_{gr}$

Assumption 2: $\overline{LMP}_{gr} < \overline{LMP}_c^d \leq 1$

Assumption 3: $1 < \overline{LMP}_c^d$
Albeit assumption 3 is impossible, the joint participation rate we acquire through this assumption is feasible if there is a significant impact of defiers. Putting these assumptions into the equation (5.5) gives us some interesting bands of joint participation under these assumptions:

![Participation Rate](image)

Figure 21 – Joint participation rate under different assumptions

Where the joint participation rate is bound from above to focus on the decline in participation. We can then calculate the share of the of the decline that is due to the increase in graduation:
Noting once more that the compositional effect cannot account for the whole decline in participation unless there is a violation of the no defiers-assumption. In addition, we see that the participation of compliers must be above the participation rate of graduates to be able to explain more than 60 percent of the decline in participation (relative to 1963).

### 5.2.4 Conclusion Based on the Non-parametric Bounds

In this section, I have investigated what the participation rate would have been if the graduation rates would have been constant. Under fairly weak assumptions I have shown that the change in graduation cannot account for the full decline in participation rate. Moreover, I have shown that one must assume that the individuals that have increasingly graduate must have a participation rate that is greater than the observed participation rates of graduates if they should account for more than 60 percent of the decline in participation of low-skilled males. Since this outflow of individuals are the main reason to investigate the compositional effect, the findings also support the results in the last section. Namely that the compositional effect is unlikely to have caused the full decline in participation for low-skilled males.
6 Conclusion

In this thesis I have investigated to what extent the decline in labor market participation rates of low-skilled males can be explained by a compositional effect. This theory competes with other theories that predicts a reduction in the labor market condition of these individuals.

Using a register data set covering every cohort born between 1963 and 1983, I have decomposed the decline of participation into a compositional effect and other effects. Using matching to control for the composition of different cohort, I have found that observed characteristics are unable to explain the decline in labor market participation. This does not necessarily imply that the impact of a compositional effect is absent, but may be a result of the large amount of missing observations and non-observed individual characteristics. This is however a surprising result, given that social background and observed ability are often thought to be important predictors of labor market outcomes.

Using the approach of non-parametric bounds, I have shown that one must accept quite strong assumptions for the compositional effect to account for more than 60 percent of the observed decline in labor market participation. For the compositional effect to explain the full decline, one must accept assumptions that seems unrealistic given the trend in graduation during the period.

Although both methods are subject to some constraints, they both point in the same direction. Namely that there is something going on that pushes low-skilled males out of the labor market. This may be some change that makes two cohorts different in some way, for example if there is an increase in health issues over time. It may also be because low-skilled males don’t find that they have the same opportunities in the labor market as before. Some theories suggest that technology or immigration affects the possibilities that these individuals face in the labor market. If this is the case, this raises concerns regarding the wellbeing of males without upper secondary education, their dependence on the welfare state, and the overall distribution of income.

To further investigating the compositional effect, I would recommend acquiring more data on relevant individual characteristics. One such potential variable that may contain some interesting variation is health records. In general, however, I believe the labor market condition of low-skilled individuals in Norway should acquire even more attention. Although
some research have been devoted to the labor market condition and participation rate of these individuals, more research into the role of technology, immigration and trade could shed more light on the observed trends in participation rates by Norwegian males without upper secondary education.
References


Statistics Norway (2016 (d)). "En demografisk beskrivelse av arbeidsinnvandrere fra EU/EØS og deres familier (A Demographic Description of Migrant Workers From EU / EEA and Their Families)."


Appendix A

Let $\overline{X}$ be the vector that consists of the observed characteristics $X$ and unobserved characteristics $Z$. The composition adjusted participation in (5.1) can then be rewritten as a product of the distribution of observed and unobserved characteristics:

$$
\int f_{\overline{X}|do,y}(x)E[LMP_{do}|X = x, Z = z, Y = y]dx
= \int f_{\overline{X}|do,y}f_{Z|X,do,y}(x)E[LMP_{do}|X = x, Z = z, Y = y]dx, \ \forall y \neq y'
$$

The function that will be approximated when $Z$ is unobserved will be the function where $Z$ is free to vary:

$$
\int f_{X|do,y}(x)E[LMP_{do}|X = x, Y = y]dx
= \int f_{X|do,y}(x)E[E[LMP_{do}|X = x, Z = z, Y = y]|Z = z]dx
= \int f_{X|do,y}(x)f_{Z|X,do,y}(x)E[LMP_{do}|X = x, Z = z, Y = y]dx
$$

Where the first equality follows from the law of iterated expectation. Notice that these two functions only differ in terms of the conditional distribution of $Z$, and that the equations are identical whenever $f_{Z|X,do,y}(x) = f_{Z|X,do,y}(x)$.
Appendix B

Table 7 – Robustness tests

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Wage only</th>
<th></th>
<th></th>
<th>Wage and IQ-test score</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unweighted sample</td>
<td>Matched sample</td>
<td>Unweighted sample</td>
<td>Matched sample</td>
<td></td>
</tr>
<tr>
<td>1963</td>
<td>0.166</td>
<td>0.161***</td>
<td>0.141</td>
<td>0.139***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0144)</td>
<td></td>
<td>(0.0138)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1964</td>
<td>0.166</td>
<td>0.164***</td>
<td>0.142</td>
<td>0.143***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td></td>
<td>(0.0126)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>0.165</td>
<td>0.163***</td>
<td>0.142</td>
<td>0.143***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td></td>
<td>(0.0126)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1966</td>
<td>0.161</td>
<td>0.162***</td>
<td>0.137</td>
<td>0.14***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td></td>
<td>(0.0139)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1967</td>
<td>0.16</td>
<td>0.156***</td>
<td>0.141</td>
<td>0.14***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0129)</td>
<td></td>
<td>(0.013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1968</td>
<td>0.141</td>
<td>0.141***</td>
<td>0.127</td>
<td>0.128***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0124)</td>
<td></td>
<td>(0.0117)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1969</td>
<td>0.132</td>
<td>0.131***</td>
<td>0.118</td>
<td>0.12***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0139)</td>
<td></td>
<td>(0.0144)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1970</td>
<td>0.127</td>
<td>0.124***</td>
<td>0.114</td>
<td>0.115***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0145)</td>
<td></td>
<td>(0.013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1971</td>
<td>0.109</td>
<td>0.105***</td>
<td>0.103</td>
<td>0.102***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0133)</td>
<td></td>
<td>(0.0119)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1972</td>
<td>0.098</td>
<td>0.096***</td>
<td>0.091</td>
<td>0.093***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0124)</td>
<td></td>
<td>(0.0117)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1973</td>
<td>0.082</td>
<td>0.081***</td>
<td>0.081</td>
<td>0.083***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0122)</td>
<td></td>
<td>(0.0111)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1974</td>
<td>0.057</td>
<td>0.056***</td>
<td>0.052</td>
<td>0.054***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0137)</td>
<td></td>
<td>(0.0121)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1975</td>
<td>0.042</td>
<td>0.041***</td>
<td>0.04</td>
<td>0.043***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0135)</td>
<td></td>
<td>(0.0127)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1976</td>
<td>0.049</td>
<td>0.048***</td>
<td>0.049</td>
<td>0.049***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0138)</td>
<td></td>
<td>(0.0135)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1977</td>
<td>0.04</td>
<td>0.038**</td>
<td>0.038</td>
<td>0.037**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0141)</td>
<td></td>
<td>(0.0143)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1978</td>
<td>0.054</td>
<td>0.053***</td>
<td>0.052</td>
<td>0.053**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0109)</td>
<td></td>
<td>(0.0105)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1979</td>
<td>0.036</td>
<td>0.035***</td>
<td>0.039</td>
<td>0.038**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td></td>
<td>(0.0102)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N 141.585  141.585  123.723  123.723

* p < 0.10, ** p < 0.05, *** p < 0.01
Standard errors in parenthesis. All standard errors are clustered on the county level. None of the standard errors take into account the estimation of the propensity score or other aspects of matching.
### Table 8 – Robustness tests

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Full set of covariates</th>
<th>Relative covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unweighted sample</td>
<td>Matched sample</td>
</tr>
<tr>
<td>1963</td>
<td>0,147 (0,0131)</td>
<td>0,139*** (0,0134)</td>
</tr>
<tr>
<td>1964</td>
<td>0,15 (0,0135)</td>
<td>0,147*** (0,0139)</td>
</tr>
<tr>
<td>1965</td>
<td>0,148 (0,0128)</td>
<td>0,144*** (0,0122)</td>
</tr>
<tr>
<td>1966</td>
<td>0,144 (0,0131)</td>
<td>0,14*** (0,0135)</td>
</tr>
<tr>
<td>1967</td>
<td>0,146 (0,0128)</td>
<td>0,139*** (0,0128)</td>
</tr>
<tr>
<td>1968</td>
<td>0,132 (0,0111)</td>
<td>0,13*** (0,0111)</td>
</tr>
<tr>
<td>1969</td>
<td>0,123 (0,0149)</td>
<td>0,121*** (0,0149)</td>
</tr>
<tr>
<td>1970</td>
<td>0,117 (0,0131)</td>
<td>0,112*** (0,0131)</td>
</tr>
<tr>
<td>1971</td>
<td>0,107 (0,0122)</td>
<td>0,104*** (0,0122)</td>
</tr>
<tr>
<td>1972</td>
<td>0,095 (0,0115)</td>
<td>0,093*** (0,0115)</td>
</tr>
<tr>
<td>1973</td>
<td>0,082 (0,0114)</td>
<td>0,083*** (0,0114)</td>
</tr>
<tr>
<td>1974</td>
<td>0,055 (0,0125)</td>
<td>0,053*** (0,0125)</td>
</tr>
<tr>
<td>1975</td>
<td>0,041 (0,0129)</td>
<td>0,041*** (0,0129)</td>
</tr>
<tr>
<td>1976</td>
<td>0,05 (0,0136)</td>
<td>0,049*** (0,0136)</td>
</tr>
<tr>
<td>1977</td>
<td>0,038 (0,0144)</td>
<td>0,036** (0,0145)</td>
</tr>
<tr>
<td>1978</td>
<td>0,053 (0,0105)</td>
<td>0,053*** (0,0104)</td>
</tr>
<tr>
<td>1979</td>
<td>0,038 (0,0103)</td>
<td>0,039*** (0,0102)</td>
</tr>
</tbody>
</table>

N = 121,539

*p < 0.10,  **p < 0.05,  ***p < 0.01

Standard errors in parenthesis. All standard errors are clustered on the county level. None of the standard errors take into account the estimation of the propensity score or other aspects of matching.