NHH

Norwegian School of Economics
Bergen, Spring 2020

Gender Differences in the Labour Market: Explaining the Gender Wage Gap

Empirical Evidence from Norway

Marte Mellesdal Dehli

Supervisor: Astrid Kunze

Master thesis, Economics and Business Administration

Major: Economics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis constitutes 30 ECTS points of my major in Economics and is the final project concluding my master´s degree at the Norwegian School of Economics. The process of writing this thesis has been challenging, but the knowledge gained is invaluable. I feel fortunate to write about such an important topic as gender equality and to have the opportunity to use register data through micordata.no.

I wish to express my sincere appreciation to my supervisor, Astrid Kunze, for inspirational discussions and excellent support throughout this process. I am grateful for your time and guidance. Lastly, I want to thank friends and family for their great support along the way of writing this thesis.

Norwegian School of Economics

Bergen, June 2020

Marte Mellesdal Dehli

Abstract

This master´s thesis aims to investigate potential sources of the gender wage gap using Norwegian register data on the full population. First, I seek to understand to what extent traditional human capital factors and other work-related characteristics contribute to the gender wage gap. Using a traditional Oaxaca-Blinder decomposition method, I decompose the male-female differentials from a cross-sectional perspective based on the study by Blau and Kahn (2017). Second, I attempt to understand whether gender inequality is due to children and if there is a motherhood penalty in earnings by adopting the event study approach suggested by Kleven, Landais, and Søgaard (2018). By controlling for maternal age and calendar year, the event study allows for capturing the effect of children on female and male wages over time.

The O-B decomposition reveals that conventional human capital factors in aggregate decrease the gender wage gap, while gender segregation in industries increases the gender wage gap by a small share. As a result, most of the gender wage gap is due to unexplained factors, which calls for a discussion for other potential explanations of the gender wage gap. The event study reveals a significant child penalty in earnings for mothers, implying that children have significant impacts on wages. This motherhood penalty suggests negative selection into work and labour market adjustments around the childbirth. Whether the drop in female wages is due to unprofitable choices or discrimination is hard to establish, but it might be that both of them play a role to some degree.

Keywords – Gender equality, gender wage gap, event study, motherhood penalty, microdata.no

Contents

List of Figures

List of Tables

1 Introduction

In recent decades, the roles of women and men have converged, which has caused gender differences to decrease substantially in many developed countries. More women have entered the labour market, and the educational level of women has increased substantially (Becker, 2009). The converging roles between women and men have led to a narrowing of gender differences in, among others, labour force participation, life-time labour force experience, occupation and education (Goldin, 2014; Goldin, Katz, & Kuziemko, 2006). There has also been a convergence in earnings, which will be the main focus of this thesis.

Equal pay for equal work is not only a legal requirement but also a measure for a fairer society. Since women represent half of the world's population, empowering women is crucial to increase productivity and economic growth (UN, 2020). In a recent report, the OECD (2018) states that the Nordic countries have been leaders in the development of gender equality, and the high proportion of women in working life has benefited these countries both socially and economically. The report also emphasises that the Nordic countries have come further in achieving gender equality compared to other OECD countries (OECD, 2018).

Nevertheless, mandatory wage transparency and gender pay reporting requirements have given new insight into the degree of equal pay between men and women, and there is still a long way to go to attain full gender equality, even for the Nordic countries. Although Norway is one of the most gender-equal countries, and the workforce contains almost as many women as men, full equality in wages is yet to be achieved. By 2019, women's average monthly wages accounted for only 87.6 per cent of men's wages (Askvik, 2020). The differences were most significant for full-time workers and people with higher education (Kristoffersen, 2017).

So why do we observe these persistent wage disparities between women and men? Even though the female participation rate has increased, part-time work is still far more common among women. In addition, men and women tend to have different occupations and work in different industries. Statistics Norway shows that the majority of women work in health, social work and education, while most men are in manufacturing and construction (SSB, 2020). Women also tend to make different human capital investments, causing disparities in skill-sets and work experience (Blau, Ferber, & Winkler, 2014). Policies and measures such as parental leave, child care subsidies and gender-specific anti-discrimination laws have been implemented to ensure women's rights when they have children (Jaumotte, 2003). As we will see later in the thesis, these policies might work contrary to their purpose by increasing the differences between women and men.

In 1993, men were entitled to paid parental leave of four weeks. The paternity quota remained constant for more than a decade. Finally, in the period 2005-2013, there was an increase in the paternity quota from five to fourteen weeks (Hamre, 2017). Since the paternity quota intended to ensure a more equal distribution of care taking between mothers and fathers, it is most appropriate to study wage differences in the period after 2005. Using register data collected from microdata.no (henceforth Microdata) on the full population, the thesis aims to provide new insight into gender differences in wages. The data includes information about the population, education, earnings and work characteristics, in addition to data on fertility and parental leave. Since the availability of these variables differs, the focus of this thesis will be particularly on 2006 and 2007.

Using two complementary statistical methods, the thesis aims to investigate gender differences from different points of view. The first part of the empirical analysis decomposes the gender wage gap from a cross-sectional perspective on the full population for 2006 and 2007. The second part turn to investigate children´s impacts on wages using an event study approach over the 2003-2015 period for those individuals who have their first child during 2006-2007. Based on register data for the full population in Norway and two complementary statistical methods, the thesis attempts to answer the following research question:

What are the sources of the persistent gender wage gap,
and is gender inequality due to children?

The thesis contributes to the growing literature on the extent of and potential explanations behind the gender wage gap. Using Norwegian register data on the full population allows for assessing high-quality data from a large sample (full population) with a minimal probability for measurement errors. This is a substantial advantage compared to studies using surveys. Moreover, the data enables to study non-mother to mother transitions for a specific year they give birth, which will be exploited further in the study. Previous research has studied children's impact on wages[1], but the number of Norwegian event studies on the topic is limited. Thus, the thesis also makes a methodological contribution by using an event study approach[2] to analyse children's impacts on wages for the individuals who work full-time in either public or private sector.

The empirical analysis is based on two particular studies which I replicate to test existing economic explanations. First, I decompose gender differences in wages according to the study of Blau and Kahn (2017) by using a traditional Oaxaca-Blinder method[3]. Second, I conduct the event study based on the paper of Kleven et al. (2018) to investigate children's impacts the wage trajectories. The replication studies allow for comparing the results of the empirical analysis to previous evidence, and accordingly, for new interpretations of the empirical evidence from Norway.

Norway has a well-developed and well-documented system for register data and having access to such data is valuable for the research purpose. Microdata allows for access to unique register data instantaneously and to explore and organise the data accordingly. Since Microdata was launched in 2018, few studies have used data from Microdata in empirical research. Thus, the thesis is an innovative contribution to the existing literature. Microdata allows for both cross-sectional and panel data sets, which makes it possible to analyse gender differences in wages from different points of view. A detailed note on the user-friendliness as well as personal experiences with Microdata is available in the Appendix.

The remainder of the thesis is structured as follows: Chapter 2 provides a summary of the economics behind potential explanations of the gender wage gap. Chapter 3 presents an overview of the background on the gender wage gap and the Oaxaca-Blinder methodology. Chapter 4 provides an overview of the background of the event study approach and the empirical strategy of the event study. Chapter 5 describes the data and sample selection. Chapter 6 presents the results of the empirical analysis, while Chapter 7 discusses the findings followed by limitations and suggestions for future research. Chapter 8 concludes the findings of the study.

[1] See e.g., Kunze (2015), Kunze (2018), Hardoy and Schøne (2008).
[2] See Section 4.2 for details on the event study approach.
[3] See Section 3.2 for details on the Oaxaca-Blinder methodology.

2 Economic explanations and hypotheses

Before going into the study in detail, some of the main economic explanations behind gender differences in wages will be outlined. A clear understanding of the economic concepts helps to understand why gender wage gaps occur in the economy. Women are generally considered to be paid less than men. A gender wage gap or a gender pay gap arises when women and men are paid differently, and can be determined as the difference between the median earnings of women relative to the median earnings of men (OECD, 2020a). Gender differences in wages can also be measured as female earnings divided by male earnings, known as the gender pay ratio, showing how much a woman makes compared to a man (Blau et al., 2014).

2.1 Summary of economic explanations

The human capital model provides the primary supply-side explanation of gender differences in wages and helps explain why some choose to invest in human capital while others do not (Blau et al., 2014). The two primary human capital factors include education and labour market experience, and human capital theory suggests that earnings rise with additional education or on-job-training because of the productivity-enhancing effects of education and work experience (Becker, 2009; Mincer, 1962). Gender differences in these areas can produce substantial differences in earnings between women and men (Mincer & Polachek, 1974). Further, the human capital model provides insight into the expected working life of women and men. Given traditional gender roles, many women anticipate shorter and more disrupted work lives than men. For this reason, women choose to invest in fewer years of education and work experience. Based on the human capital model, we expect the gender wage gap to decline if women increase their investments in human capital relative to men.

Occupational segregation also contributes to the gender wage gap. The divergence between women and men in types of jobs is referred to as occupational segregation in the labour market. Women are on average more likely to work in low-paid occupations and industries and men on average in higher-paid occupations and industries, which may reflect substantial job barriers or different job preferences between women and men

(Kunze, 2018). Gender segregation in a human capital perspective is quite straightforward. Given the traditional division of labour, women are expected to select occupations and industries that require less investment in education and on-job-training, as they anticipate a shorter and less continuous work career than men (Blau et al., 2014).

Women's greater responsibility for children is an essential factor in explaining the gender wage gap. The effect of family on women's wages is quite different than in the case of men. This finding is known as the motherhood penalty (Blau et al., 2014; Budig & England, 2001). Waldfogel (1998) introduced the term "family-gap" by showing that women with children earn less than women without children. Likewise, Kunze (2015) showed that women with children are less likely to be promoted, indicating a family gap in career promotion. There are several explanations of the family gap or the motherhood penalty, and it may be that all of them play a role to some degree.

Women with children are less likely to participate in the labour market than men and women without children, and when they do participate, they tend to work fewer hours and earn lower hourly wages (Sigle-Rushton & Waldfogel, 2007). Combining family life with working life can be challenging for many women. In order to take care of the children, women might not return to their full-time positions after giving birth, but instead change to part-time positions or even withdraw from the labour market. This finding is known as negative selection into work (Kunze, 2008). As a consequence, mothers can miss out on promotions, bonuses and other career opportunities.

Likewise, women might not return to their previous employer after childbirth if their employer does not provide adequate maternity leave. Instead, women may switch to more child-friendly firms (Hotz, Johansson, & Karimi, 2017). The desire for time flexibility due to the arrival of children can cause women to change work to the public sector or firms that require less working hours and overtime. Flexibility at work comes typically at a high price, and women might have to give up the ability to climb career ladders over the long-run in order to take care of the children (Goldin, 2014; Hotz et al., 2017).

The extent of family-friendly policies, including the access to parental leave, is another explanation for the existence of the gender wage gap. The purpose of work-facilitating policies is to ensure a better family-work balance, to encourage job continuity after birth and to increase women's earnings, by providing the right to return to a previous position of employment and by offering financial support (Ejrnæs & Kunze, 2013). However, studies have found that extended parental leave increases the proportion of women who never return to work and decreases employment and earnings in the short-run (Lalive & Zweimüller, 2009). Parental leave can, therefore, be an obstacle for mothers as work interruptions related to childbirth might lead to loss of human capital and weaker labour market prospects, which affect mothers' wages directly (Lalive, Schlosser, Steinhauer, & Zweimüller, 2014).

Labour market discrimination against mothers is another explanation that contributes to the gender wage gap. Such discrimination suggests that women face differential treatment based on parental status (Correll, Benard, & Paik, 2007). For instance, employers might perceive mothers as less productive compared to non-mothers due to their greater responsibilities at home and constraints on work schedules (Blau et al., 2014). Consequently, employers would place mothers in less rewarding jobs, promoting them less and paying them less within jobs (Budig & England, 2001).

Overall, there are several potential explanations behind the gender wage gap. The main economic explanations include human capital, occupational segregation and the effects of women's responsibility for childbearing such as nonrandom selection, demand for family-friendly jobs, extended parental leave and labour market discrimination by employers. Since the roles of women have changed dramatically in recent decades, it is appropriate to start by investigating to what extent traditional human capital factors can explain the gender wage gap. Thus, the first part of the empirical analysis decomposes the male-female differentials in wages adopting the Oaxaca-Blinder method suggested by Blau and Kahn (2017).

The second part of the analysis attempts to understand whether gender inequality is due to children and if there is a motherhood penalty in earnings. Since the event study approach does a better job at capturing the impacts of children than the traditional Oaxaca-Blinder decomposition, I investigate children's impacts on female and male wages using the event study approach suggested by Kleven et al. (2018). By controlling for maternal age and calendar year, the event study allows for capturing the effects of children on wages. Likewise, the time dimension allows for observing how mothers' wages evolve according to the varying needs of the child.

2.2 Hypothesis development

Summarising the main economic explanations has yielded insight into potential drivers of the gender wage gap. Different studies have looked at the contribution of traditional human capital factors to the gender wage gap. Particularly, Blau and Kahn (2017) have investigated the change in the importance of conventional human capital factors for the US over the 1980-2010 period. Table A3.1 in the Appendix displays the decomposition by Blau and Kahn. The tables shows that traditional human capital variables taken together contribute less to the gender wage gap in 2010 than in 1980 because of women's increased investments in human capital, particularly in education. However, gender differences in occupation and industry continued to be important in 2010. Based on these findings, we make the following predictions:

Hypothesis 1a: There is a persistent gender wage gap in Norway partly due to gender segregation in industries.

Hypothesis 1b: The decomposition proposed by Blau and Kahn is relevant to understand the gender wage gap in the Norwegian labour market.

Further evidence presents explanations outside traditional human capital factors, including motherhood penalties and family-gaps, suggesting that women with children are paid less than women without children (Blau et al., 2014; Budig & England, 2001; Kunze, 2015; Waldfogel, 1998). Previous research has revealed a drop in women's wages around the birth of the first child. Kunze and Ejrnæs (2004) revealed a dip in women's real wages shortly before giving birth and a drop of 10 to 20 % after finishing maternity leave and returning to the labour market, while Kleven et al. (2018) showed that women experience an immediate drop in gross earnings of almost 30%. Since Kleven et al. (2018) provides the basis for the event study of the thesis, we make the following predictions:

Hypothesis 2a: The event study provides new insight into children's contribution to the gender wage gap.

Hypothesis 2b: The arrival of the first child leads to a child penalty in earnings for mothers similar to the Danish finding of almost 30%.

3 Decomposing the gender wage gap

The preceding chapter provided a summary of the main economic explanations behind the gender wage gap. This chapter, however, presents empirical evidence on the sources of gender differences in earnings and discusses reasons for why the gender wage gap has declined over the past decades. Lastly, the Oaxaca-Blinder method and the corresponding empirical strategy are presented.

3.1 Background on the gender wage gap

3.1.1 *Evidence on the sources of gender differences in wages*

Economists and social scientists have long attempted to find new empirical evidence on the sources of the gender wage gap. Traditionally, when studying the gender wage gap, researchers have focused on human capital (schooling and work experience), the family division of labour, compensating wage differentials, discrimination, and issues relating to selection into the labour force (Blau & Kahn, 2017). Despite various findings among the studies due to different method of analysis and data, several studies discovered that differences in qualifications or (potential) productivity could not solely explain the gender wage gap (Blau et al., 2014; Blau & Kahn, 2008).

A widely used method to study the gender wage gap is to decompose gender differences into two parts: the part of the gender differences that is due to differences in human capital or other qualifications and the part that cannot be explained by such factors. In a recent study, Blau and Kahn (2017) used Panel Study of Income Dynamics (PSID) over the 1980-2010 period to provide new empirical evidence on the extent of and trends in the gender wage gap. Notably, they showed that the gender wage gap in the United States improved substantially over the 1980-2010 period and that the gender wage gap was more persistent at the top of the wage distribution than elsewhere.

By providing evidence on the importance of gender differences in productivity of traditional human capital factors and possible labour-market discrimination, Blau and Kahn (2017) showed that women's improvement in qualifications relative to men contributed substantially to the narrowing of the gender wage gap. In particular, women's improvements in education, experience, professional representation and shortfall in union coverage played an essential role in decreasing the gender wage gap, but also the decline in the unexplained portion contributed to the narrowing of the gender wage gap. By 2010, however, they found that conventional human capital factors explained little of the gender wage gap in the aggregate, but that other factors such as occupation and industry continued to be significant. These findings suggest that gender differences in work-related characteristics have fallen in importance and that human capital factors are only part of the story. What then causes the remaining gender wage gap?

3.1.2 *Possible sources of the unexplained gender wage gap*

According to traditional analysis, such as the Blau-Kahn analysis, the existence of an unexplained gap is consistent with discrimination against women in the labour market (Blau et al., 2014). However, this does not mean that the entire unexplained gap is due to discrimination. One possible source of the unexplained gender wage gap is soft skills such as attitudes towards negotiation, competition and risk, but also the fact that women potentially place a lesser value on money and work than men (Fortin, 2008). Another highly possible source of the unexplained gap is children. A large and growing literature indicates that women's labour-market outcomes might be negatively affected by motherhood (Blau & Kahn, 2017; Sigle-Rushton & Waldfogel, 2007; Waldfogel, 1998). Evidence implies that the gender wage gap increases after the birth of the first child because of women's labour market adjustments around childbirth (Kunze, 2018). Kleven et al. (2018) found in their study that almost all of the remaining gender inequality can be attributed to children.

3.1.3 *The declining gender wage gap*

As outlined, the gender wage gap has declined significantly over the past decades. Substantial evidence shows that women have steadily increased their levels of education, now surpassing men in several countries (Becker, 2009). More women have also entered the labour force over the past decades, causing the traditional gender roles to converge (Goldin, 2014). Since women have become more similar to men, in terms of qualifications and productivity, conventional human capital factors might have problems explaining the persistent gender wage gap that we observe in many countries today.

To investigate what role traditional human capital factors play in explaining the gender wage gap in Norway, we replicate the decomposition conducted by Blau and Kahn (2017) by using Norwegian register data. The decomposition analysis is built upon the Oaxaca-Blinder method (Blinder, 1973; Oaxaca, 1973), which have been the main empirical workhorse for analysing the gender wage gap over the past few decades. In Chapter 4, we turn to investigate children's impacts on wages as an alternative approach for studying the gender wage gap.

3.2 Oaxaca-Blinder decomposition method

The Oaxaca-Blinder decomposition method is a starting point for investigating the gender wage gap. The method provides further detail on the contribution of particular labour-market characteristics to the gender wage gap (Gardeazabal & Ugidos, 2004). The Oaxaca-Blinder decomposition aims to explain gender wage differentials at the mean by decomposing the gap into differences in labour-market characteristics and the effects on these characteristics (Blinder, 1973; Oaxaca, 1973). The latter is known as the unexplained gender wage gap and has been earlier interpreted as a measure of discrimination, which suggests unequal pay for equally qualified workers. Based on Blau and Kahn (2017), the following equations describe the Oaxaca-Blinder decomposition. First, I estimate separate male (m) and female (f) ordinary least squares (OLS) wage regressions for individual i, in year t:

$$Y_m = X_m B_m + u_m \tag{3.1}$$

$$Y_f = X_f B_f + u_f \tag{3.2}$$

Where Y denotes the log of wages, X denotes a vector of explanatory variables including individual human capital characteristics (education and experience), regional and industrial dummy variables, B denotes a vector of coefficients and u denotes an error term. Given Equation 3.1 and 3.2, the wage differentials or the total gender wage gap can be computed. I assume b_m and b_f to be the OLS estimates of B_m and B_f. The bars denote the mean of the variables. Since OLS with a constant term yields residuals with a zero mean, the difference in the mean can be written as:

$$\bar{Y}_m - \bar{Y}_f = \; b_m \bar{X}_m \; + \; b_f \bar{X}_f = \; b_m(\bar{X}_m - \bar{X}_f) + \bar{X}_f(b_m - b_f) \qquad (3.3)$$

Where the first term of the right hand side of Equation 3.3 displays the explained gender wage gap, which is the impact of gender differences in the explanatory variables X, evaluated using the current year male OLS coefficients for the corresponding variables, b_m. The second term displays the unexplained gender wage gap. Unexplained gender differences are differences in returns evaluated using the current year female residual for the corresponding variables, X_f.

The empirical analysis performs a stepwise Oaxaca-Blinder decomposition in order to decompose wage differentials at the mean. The first step includes estimating separate female and male OLS wage regressions in line with Equation 3.1 and 3.2 using the human capital and the full specification. The second step employs the estimated OLS coefficients obtained from the regression specifications, together with descriptive statistics retrieved from Microdata (see Table 5.5 and Table 5.6), to calculate the gender wage gap according to Equation 3.3.

Regression 1: Human-capital specification

The first regression specification applied in the empirical analysis estimates the relationship between log wages and human capital variables. For year t, I estimate separate OLS regressions for male and female:

$$ln\, W_i^g = \; \beta_0 + \beta_1 Educ_i^g + \beta_2 Exper_i^g + \beta_3 (Exper^2)_i^g + \beta_4 Metro_i^g + \gamma_r Region_{r,i}^g + \epsilon_i^g \quad (3.4)$$

Where subscript i denotes individuals and subscript g denotes gender (male and female).

The specification includes traditional human capital characteristics including education, experience and experience squared in addition to dummy variables for large cities (metro) and regions. β_1 and β_2 display the estimated percentage point change in wages by one additional year of education and potential experience respectively, while β_3 display the estimated coefficient of the non-linear effect of potential experience on wages. β_4 displays the estimated percentage point change in wages if individual i is working in a large city with at least 100,000 inhabitants (otherwise zero). The dummy variable region consists of five regional groups related to residency where Eastern Norway is omitted from the regression to serve as the reference category. Thus, γ_r measures the proportionate difference in wages for the regional dummies relative to Eastern Norway, holding all other factors constant.

Regression 2: Full specification

The second regression specification is an extension of the first specification. Industry dummies are added to the human capital variables in order to control for industry effects. For year t, I estimate the full specification as follows:

$$\begin{aligned} ln\ W_i^g = \ & \beta_0 \ + \ \beta_1 Educ_i^g \ + \ \beta_2 Exper_i^g \ + \ \beta_3 (Exper^2)_i^g \\ & + \ \beta_4 Metro_i^g + \ \gamma_r Region_{r,i}^g \ + \ \delta_s Industry_{s,i}^g \ + \ \epsilon_i^g \end{aligned} \tag{3.5}$$

Where the second last term denotes industry dummies. There are thirteen distinct industry groups to which each individual can be assigned. Each individual is only assigned to one group at a time, and the corresponding group takes the value of 1 (otherwise 0). δ_s displays the estimated expected change in wages for a specific industry group relative to the reference group. Industry group D, manufacturing, serves as the industry reference group.

Explaining the gender wage gap at the mean

The final step of the Oaxaca-Blinder decomposition is to determine and measure the explained and the unexplained portions of the gender wage gap, equivalent to Equation 3.3. Due to limitations of Microdata, I use Excel to manually calculate the final step of the decomposition. Accordingly, the estimated OLS coefficients from both regression specifications and the means of the explanatory variables retrieved from the descriptive statistics in Microdata are used to decompose the gender wage gap in detail. The entries are the male-female differential in the indicated variables multiplied by the current year male OLS coefficients for the corresponding variable. The coefficients from the male wage regression serve as the reference group. Since the expectation is that males do not to get discriminated, the coefficients from the wage regression represent the actual return and thus constitutes the weights (Kunze, 2008). The total unexplained gap is the mean female residual from the male OLS wage equation.

4 Event study: impacts of children

This chapter presents the event study literature and methodology. The first part presents the theoretical and empirical background, whereas the second part presents the empirical strategy and identification. The purpose of the event study is to identify the impacts of children on the gender wage gap.

4.1 Background on the event study approach

An event study is a statistical method which is frequently used in finance and economics to measure effects of particular economic events on the value of firms or particular labour market outcomes. The event study approach allows for identifying changes in economic outcomes around the birth of the first child. Accordingly, the event study approach has grown in popularity, expanding the literature on the effects of parenthood on labour market trajectories. Event studies enable researchers to study the impacts of children on gender inequality between women with and without children, between men with and without children, and finally between women and men with children. The latter constitutes the main focus of this thesis.

Among the event studies that have contributed the most to the literature is the study of children and gender inequality conducted by Kleven et al. (2018). This study provides the basis for the second part of the empirical analysis later in the thesis. Kleven et al. (2018) estimate the impact of children on the labour market trajectories of women relative to men by using a quasi-experimental event study approach and full-population administrative data from Denmark from 1980 to 2013. To examine the full implications of children for gender inequality, they adopt an event study based on sharp changes around the birth of the first child. The empirical identification includes full sets of age dummies and year dummies to control non-parametrically for underlying life-cycle trends and time trends such as wage inflation and business cycles. A balanced panel of parents is observed each year from five years before the childbirth until ten years after the childbirth. For a broad set of labour market outcomes, the results report large and sharp effects of children in a negative direction. First, the results show that the impact of children on women is large and persistent and that the birth of the first child creates a gender gap in earnings of around 20% in the long run. Underlying this earnings penalty, Kleven et al. (2018) suggest sharp impacts of children on labour force participation, hours worked, wages rate, occupation, sector and company choices. Second, the results show that the fraction of child-related gender inequality has increased dramatically over time and that almost all of the remaining gender inequality is due to children.

To highlight other important event studies on children and gender inequality a natural starting point is a study performed by Angelov, Johansson, and Lindahl (2016), which estimates the short and long-term effects of entering parenthood on the gender gaps in income and wages. The study aims to compare income and wage trajectories of women relative to men before and after parenthood using an event study approach and administrative data from Sweden during the period from 1986-2008. Focusing on the within-couple gap, Angelov et al. (2016) find that the male-female gender gap in income and wages have increased in the long run. The results also reveal a large drop in working hours after the birth of the first child even though the Swedish market is a highly flexible labour market.

Furthermore, Lucifora, Meurs, and Villar (2017) analyse the impact of first childbirth on earnings and careers in an internal labour market, namely a family-friendly French company using a panel of personnel records from 2005 to 2016. This study focuses on comparing parents with and without children. Despite the family-friendly institutional context, the results show that women's labour market outcomes are affected mainly by the birth of the first child, while fatherhood does not significantly impact men's wages or careers. One year after birth, women's total pay and individual bonuses diverge substantially, and the drop is persistent showing no evidence of a catching-up trend. Lucifora et al. (2017) suggest that mothers' reduction in working hours, increase in hours of absence and decrease in extra-time are possible explanations of these results.

In summary, the event studies suggest a child earnings penalty for mothers, experiencing a drop in wages around the birth of the first child. In contrast, the event studies suggest that men are unaffected of the arrival of the first child.

4.1.1 *O-B decomposition vs. event study using childbirths*

Using event studies to analyse the gender wage gap represents a departure from standard gender gap decompositions both in the variation used for identification and in terms of the question asked (Kleven et al., 2018). While standard gender gap decompositions usually use cross-sectional variation in labour market characteristics, excluding children because of the choice of having children is endogenous, event studies exploit within-person variation in the timing of childbirth. The focus of standard decomposition approaches includes statistically measuring whether men and women receive unequal pay for equal work while controlling for human capital and labour market variables. In contrast, the focus of event study approaches includes estimating the impact of children on gender inequality not controlling for labour market characteristics that are transmission mechanisms for children. Kleven et al. (2018) suggest that even if standard decompositions show perfectly equal pay for equal work (a zero gender wage gap), event study approaches may still detect large child-related gender inequality.

4.1.2 *Comparisons of methods*

Kleven et al. (2018) contribute to the literature by comparing standard event study estimates to more sophisticated event study approaches that use control groups or instruments for childbirth. Their findings reveal some significant benefits associated with the use of the event study approach. By controlling non-parametrically for age and time trends, the event study approach does an excellent job of identifying child penalties also in the long-run. The event study approach provides the opportunity to follow the full dynamic trajectory of the effects, and it is very accurate as it uses individual-level variation in the timing of the birth of the first child. According to Angelov et al. (2016), the event study approach has the additional advantage of drawing direct inference on the average gender gap rather than on female earnings or wages, which makes it possible to control for observed and unobserved features of the spouse. A drawback of the event study approach is that the ordinary least squares (OLS) method produces biased estimates because of either omitted variables or reverse causality. Parenthood decision may be endogenous if it depends on current and expected earnings trajectories Angelov et al. (2016).

4.1.3 *Norwegian event studies*

Although event studies have become more popular over the past decades, there is still a limited number of event studies in the literature using Norwegian data to analyse the impact of children on wages. Bütikofer, Jensen, and Salvanes (2018) is an exception using Norwegian registry data from 1989 to 2000, to study the role of parenthood on gender gap among top earners. In particular, they study the effect of parenthood on the careers of women relative to men with graduate MBA, law, STEM or medical degrees. The results suggest that women in professions with more nonlinear wage structures, such as those requiring MBA and law degrees, suffer from a larger and more persistent child earnings penalty, in contrast to women in professions with a more linear wage structure, such as STEM and medicine. Other Norwegian studies examining children's impact on wages include Kunze (2015), Kunze (2018), Hardoy and Schøne (2008). Although these are not event studies, they contribute to understanding the gender wage gap in the Norwegian labour market.

Overall, few event studies use Norwegian data to analyse the impact of children on the gender gap. Bütikofer et al. (2018) only study the wage impact of having a child for top earners, making it difficult to define the child earnings penalty for the average worker. Thus, by using Norwegian register data for the full population and an event study approach, the empirical analysis of the thesis aims to expand on the research contributing to new knowledge on the topic.

4.2 Empirical strategy and identification

The event study approach allows to statistically measure changes in wages around the event, which is the birth of the first child. Although fertility choices are not exogenous, limiting the sample to men and women with children, the estimated gender wage gap after giving birth is independent of expected wage trajectories for men and women, and only depends on a potential omitted variable bias (Angelov et al., 2016; Bütikofer et al., 2018). Thus, the focus of the empirical analysis is to determine the effects of children women and men with children.

The empirical analysis includes a balanced panel of parents every year, starting three years before the birth of the first child until eight years after the childbirth. The time window, therefore, spans from -3 to +8. For each parent, the birth year of the first child is denoted as $t = 0$ and index all years relative to this year. I estimate the following linear OLS wage regression for individual i of gender g in calendar year s and at event time t, separately for men and women:

$$Y^g_{ist} = \alpha^g + \sum_{j \neq -2} \delta^g_t \cdot D[j = t] + \sum_k \beta^g_k \cdot D[k = Age_{is}] + \lambda^g_s + \epsilon^g_{ist} \qquad (4.1)$$

Where Y is the outcome of interest and denotes log wages. The second term on the right-hand side of the equation denotes a set of event time dummies, where the time dummy in period t is equal to 1 if the wage of the individual is observed at year t relative to the event of the birth of the first child. The event times $t = -3...0, ...8$ are included in the model, while event time $t = -2$ is omitted to serve as the reference category. Thus, the interpretation of all other dummies becomes relative to the omitted variable. δ displays the estimated percentage point change in wages at event time t relative to $t = -2$, for each gender. For $t < 0$, δ captures the pre-child effect, whereas for $t > 0$, δ captures the post-child effect. The latter effect is the effect of interest and measures how female and male wages evolve overtime after the birth of the first child. The third term denotes a set of age dummies which equal 1 when the individual is within a specific group of age, in a specific year. The parents aged 30-39 are omitted as the reference category to avoid the dummy variable trap. Λ denotes calendar year fixed effects, while ϵ denotes an error term. The model includes age dummies and year fixed effects to control for the fact that men often are a few years older than women when they have their first child and to control for underlying macro trends in the economy.

5 Data

This chapter presents the data used in the empirical analysis. The first sections give a description of the data employed and how registers and variables have been merged and created using Microdata. Further, the last sections include the sample selection, variable description and descriptive statistics. Since the empirical analysis is twofold, the details of both models are presented separately.

5.1 Data description

The analysis is based on register data for the full population in Norway collected using Microdata. The data combines several registers, linked at the individual level via personal identification numbers, and includes data from the education register, the population register, the tax and earnings register and labour market data. Thus, Microdata allows for access to rich and accurate data for the entire population, which is a significant advantage of the empirical analysis. In order to investigate the sources of the gender wage gap using two complementary methods, different data sets are created for each specific method.

First, I employ cross-sectional data from 2006 and 2007 to estimate the separate female and male OLS wage regressions, according to the Oaxaca-Blinder method. The data combines variables from the education register and the tax and earnings register and contains information on both the population and labour market characteristics, such as earnings, education, experience, region and industry. As shown in the next section, these variables have been generated merging various register variables that are available in Microdata. New variables have also been generated, including age and full-time. Full-time is here defined as the agreed/expected working hours of at least 30 hours per week (Microdata, 2020). The final data sets for the O-B decomposition include a total of 1,250,867 and 1,281,208 observations in 2006 and 2007, respectively.

Second, I link tax register data with population data to conduct an event study of children's impact on female and male wages. The data contains information about the population, labour market status, taxable income, agreed/expected working hours, industry, family member and fertility. We construct a balanced panel that includes all mothers and fathers who had their first child between 2006 and 2007, who are observed annually from three years before birth until eight years after birth. Taxable income and working hours are imported as panel data over 2003-2015 to generate event time dummies, age dummies and years. Additionally, new dummy variables for full-time work, private and public sector are generated. The sectors are classified into public and private sectors based on the industry of the main employer. An overview of the industrial classification is provided in Table A1.3 in the Appendix. The original data set contains 143,149 mothers and fathers and a total of 1,775,212 observations. From this data set, smaller data sets are created to investigate the impacts of children in detail for specific groups.

5.2 Microdata

As indicated, all the data employed in the thesis are collected from Microdata. Since the register data cannot be viewed, downloaded or extracted due to privacy protection requirements, I process and analyse all available register data using the analytical platform provided by Microdata. In order to prevent the most extreme values from being visible or influencing the analyses, the 1% highest and 1% lowest values are replaced by the limit value for the last and first percentile, respectively.

Since a considerable part of the work has consisted of merging registers and creating new variables, I describe how this has been accomplished. Note that all regression results, tables and figures are created manually using Excel, as Microdata has a limited set of functions for the output. For all data sets, the first step is to import the population by gender into an empty data set. Each individual has a unique personal identification number which is used to link various register variables and years. The next step depends on the method of analysis.

The O-B decomposition is a cross-sectional analysis, and hence the variables for 2006 and 2007 have to be imported in separate data sets. In these data sets resident status, labour market status, working hours, industry and birth year are imported to define the samples. Birth year is used to determine the individual's age. After defining the sample, the next step is to import and create all the variables included in the wage regressions, that is, log wages, education, experience, region and industry variables. Wages are generated by importing details on taxable income from the tax and earnings register. Education, however, is collected from the education register and is classified according to the Norwegian classification standard (NUS2000). In order to generate years of education, dummies representing the Norwegian educational system are created and converted into years. Experience is created as a function of age and education and represents potential labour market experience. Regions and metropolitan area, as well as industry, are generated as dummy variables. Since the wage regressions are regular regression analysis using cross-sectional data, Microdata allows for both regression specifications to be run at once.

The event study, on the other hand, is a panel data study which requires the data to be organised differently than by regular regression analysis. Panel data is created using a single import panel command. Variables cannot be imported more than once into the same panel data set, nor can ordinary cross-sectional data be mixed with panel data. Consequently, the samples need to be defined before merging them into a new data set with panel data for taxable income and working hours over the 2003-2015 period. In this way, individuals and years are merged using the personal identification number. Since the dating format indicates the number of days measured from 01.01.1970, a new variable has to be generated in order to retrieve the calendar years. Finally, all the remaining variables are merged into the panel data set to create the regression specification. Since regressions only can be run once per data set, the data is duplicated into new data sets for each sample before running the regressions separately for women and men.

5.3 Sample selection

This section describes the sample selection of the two models. Even though the preceding section introduced the samples, this section gives a more detailed description of the sample selection process. The selection process is presented separately for each model.

Model 1: O-B decomposition

The sample selection is based on the sample used by Blau and Kahn (2017). In order to compare the results later, it is useful to have as similar samples as possible. Thus, the final samples include full-time, non-farm wage and salary workers ages 25-64 with at least eight weeks of employment (excluding self-employed). Table 5.1 provides an overview of the selection process, showing how the final samples are defined for each year.

The samples are derived from the full population of Norway, containing approximately five million people. From the population, the samples include only those who are residents, who earn wages and salary, who work more than 30 hours per week, who are aged 25-64 and who are non-farm workers. All missing values of the imported variables are removed from the data set. Also, self-employed, unemployed and individuals outside the workforce are excluded from the samples.

Since the decomposition is a cross-sectional analysis, the samples differ to a small extent in the respective years. However, the sample selection process is the same for both years, resulting in quite similar numbers of observations. The final samples, which provide the basis for the decomposition of the gender wage gap, contains 1,250,870 and 1,281,207 observations in 2006 and 2007, respectively. The share of women is around 40% in both samples.

Table 5.1: Sample selection (O-B decomposition)

	2006		**2007**	
	Number of observations	Removed observations	Number of observations	Removed observations
(1) Keep only residents	4,578,009		4,578,009	
(2) Keep only wage and salary workers	1,968,489	2,609,520	2,022,428	2,555,581
(3) Keep only full-time workers	1,342,565	625,924	1,383,363	639,065
(4) Remove industry missing values	1,342,523	42	1,383,318	45
(5) Keep only non-farm workers	1,342,372	151	1,383,287	31
(6) Keep only ages 25-64	1,260,876	81,496	1,291,706	91,581
(7) Remove wage missing values	1,260,108	768	1,290,872	834
(8) Remove education missing values	1,250,869	9,239	1,281,219	9,653
(9) Remove municipality missing values	1,250,867	2	1,281,208	11
Final sample	**1,250,867**		**1,281,208**	

Model 2: Event study

Table 5.2 displays the selection process for each sample in the event study. The left side of the table shows the sample selection of the original sample and the corresponding subsamples. In contrast, the right side of the table shows the selection of the sample, including individuals with only one child during the time window, referred to as the restricted sample. The study focuses on full-time workers instead of part-time workers because of the number of observations. Since the sample includes mothers and fathers of two birth cohorts, the number of individuals working part-time is limited, which could bias the results.

The original sample is created by importing the full population and by keeping only mothers and fathers who had their first child in 2006 or 2007. The sample includes mothers and fathers aged over 20 years. The sample does not restrict these individuals to have only one child throughout the time window. From the original sample, subsamples are created to analyse how wages evolve in different groups. Mainly, there are three subsamples, including full-time workers, full-time workers in the public sector and full-time workers in the private sector. Finally, a more restricted sample is examined, consisting of those individuals who had their first and only child during 2006-2007. This sample is not created based on the original sample but instead on the full population.

Table 5.2 shows that the original sample contains 143,149 individuals, while the restricted sample contains only 20,760 individuals. After cleaning the samples and merging personal identification numbers ("p-id") of all individual with years, the original sample consists of 1,775,212 observations. By keeping only those individuals working full-time, the sample decreases to 1,069,625 observations. Full-time workers are then divided into different subsamples based on whether they work in the public or private sector. These samples include a total number of 342,895 and 725,977 observations, respectively. Since not every industry is assigned to one of the two sectors, the total number of observations is less than the sum of all mothers and fathers who work full-time. Finally, a sample of individuals who have their first and only child between 2006-2007 is created. The final restricted sample includes 156,211 observations and is constructed, keeping only residents over 20 years. All missing values have been removed. The restricted sample includes fewer observations than the original sample because of a more extensive cleaning process.

Table 5.2: Sample selection (Event study)

	All individuals			Only 1 child	
	Number of observations	Removed observations		Number of observations	Removed observations
(1) Keep only individuals who have their first child between 2006-2007	143,149		(1) Keep only individuals who have their first and only child between 2006-2007	20,760	
(2) Merge p-id and year	1,860,991		(2) Remove industry missing values	14,575	6,185
(3) Keep only ages 20-69	1,775,212	85,779	(3) Merge p-id and year	189,411	
Final sample (All individuals)	**1,775,212**		(4) Remove wage missing values	179,063	10,348
(4) Keep only full-time workers	1,069,625	705,587	(5) Remove work hours missing values	157,391	21,672
Final sample (Full-time)	**1,069,625**		(6) Keep only residents	157,311	80
(5a) Keep only public sector	342,895	726,730	(7) Keep only ages > 20	156,211	1,100
Final sample (Public sector)	**342,895**		Final sample	**156,211**	
(5b) Keep only private sector	725,977	343,648			
Final sample (Private sector)	**725,977**				

Table 5.3 displays all the samples by gender. In the original sample and the corresponding subsamples, the share of women is around 44%. In the more restricted sample, the share of women is around 53%. When merging the personal identification numbers with years, the share of women in the total number of observations for the original sample and the restricted sample turn out to be around 44% for both samples. However, the shares of each gender in the remaining subsamples change when merging the individuals into the panel data set. For full-time workers, the share of women is around 35%. The share of women in the public sector is 58% and 24% in the private sector.

Table 5.3: Samples by gender (Event study)

	All individuals		Full-time		Public sector		Private sector		Only 1 child	
Individuals										
Mothers	63,401	44%	63,401	44%	63,401	44%	63,401	44%	10,919	53%
Fathers	79,752	56%	79,752	56%	79,752	56%	79,752	56%	9,840	47%
Sum	143,149	100%	143,149	100%	143,149	100%	143,149	100%	20,760	100%
Observations										
Female	776,446	44%	375,583	35%	199,793	58%	17,2885	24%	68,986	44%
Male	998,772	56%	694,036	65%	143,105	42%	553,095	76%	87,213	56%
Sum	1,775,212	100%	1,069,625	100%	342,895	100%	725,977	100%	156,211	100%

5.4 Variable description

This section provides a detailed description of the variables included in the analysis. Furthermore, there is a discussion on the relevance of these variables. In the Appendix, Table A1.1 displays all variables included in the O-B decomposition, whereas Table A1.2 provides an overview of the variables included in the event study.

5.4.1 *Dependent variable*

Log wages are utilised as the dependent variable in both analyses. Wages are constructed by importing details from the tax and earnings register on taxable income at the individual level. Wages include salaries, taxable benefits in kind and sickness and childbirth benefits during the calendar year. Moreover, the variable contains only residents per 31.12 each year and exclude people with missing values or values of zero, which means that everyone registered with annual earnings is included, also those who are under the minimum wage. The advantage of using logarithmic scales is the semi-elastic interpretation of the regression coefficients. By using log wages, we can identify the percentage point change in wages caused by specific variables. Since we aim to study gender differences in wages, the only labour-market outcome of interest is log wages.

5.4.2 *Explanatory variables*

Unlike the dependent variable, the explanatory variables differ between the two models. This section presents the independent variables and the control variables utilised in the empirical analysis for each method, separately.

Model 1: O-B decomposition

The O-B decomposition utilises human capital variables and other work characteristics in line with the explanatory variables that Blau and Kahn (2017) included in their study. Education and labour market experience serve as the independent variables in both the human capital and the full specification. In the human capital specification, experience squared, metropolitan area and region dummies are included as control variables. The full specification is an extension of the human capital specification, which means industry variables are added as a control variable. In order to replicate the US decomposition of the gender wage gap, these variables are merged into broader terms, including education, experience, region and industry variables.

Education denotes years of education and shows how much an additional year of education increases wages. The variable is created using data on the highest educational level obtained by each individual in the data set according to the Norwegian classification standard (NUS2000), assigned by values in terms of years[4]. *Experience* refers to potential labour market experience, measured as age minus education years minus seven, and shows how much an additional year of work experience increases wages. *Experience*^{2} is a control variable which is added to control for the non-linear relationship between experience and wages. *Metro* and *region* are dummy variables which determine whether the person works in a large city or not, and in which of the five regions of Norway, the person is a resident. In the decomposition, these variables make up the region variable. Metropolitan area is defined as a large city with at least 100,000 inhabitants[5]. Further, Norway is divided into Northern, Trøndelag, Western and Southern, based on the old municipalities of Norway. Eastern Norway serves as the reference category since most of the population live in this region.

Industry includes thirteen industry dummies and is classified according to the standard industrial classification from 2002 (SN2002). Table A1.3 in the Appendix provides an overview of all represented industries. Industry D (manufacturing) is excluded from the regression and serves as the reference category. In contrast to Blau and Kahn (2017), this decomposition only adds industry variables to the full specification, not occupation variables. Since occupations are not available in Microdata before 2009, it is to possible to control for this factor for now. This issue is a substantial drawback of our empirical analysis. The exclusion of occupations could potentially lead to omitted variable bias.

Model 2: Event study

The event study includes a set of event time dummies as the only independent variable. In particular, it includes twelve event time dummies which span from three years before childbirth until eight years after childbirth. The event time -2 serves as the reference category and is omitted from the regression. In this way, the event time dummies show

[4]Explained further in Section 5.2.

[5]Oslo, Bergen, Stavanger/Sandnes, Trondheim, Fredrikstad/Sarpsborg and Drammen (SSB, 2019).

how female and male wages evolve before and after the birth of the first child relative to wages two years before childbirth.

Furthermore, *age dummies* and *year* are included to control for underlying trends which could affect the outcome of the event time coefficients. The inclusion of the control variables is motivated based on the event study of Kleven et al. (2018) and improves the estimates of female and male wages. In particular, age dummies and year are added to control non-parametrically for underlying life-cycle trends. On average, men tend to be a couple of years older than women when having their first child. The variable year is included to control for underlying time trends such as wage inflation and business cycles. Unlike Kleven et al. (2018), the regression specification includes years as a linear trend instead of years as a dummy variable. When including year dummies, the regression coefficients become insignificant, possibly due to the number of birth cohorts (2006 and 2007).

5.5 Descriptive statistics

This part displays the summary statistics of samples and variables used in the empirical analysis. I start by outlining the summary statistics of the O-B decomposition before I present the summary statistics of the event study. Since Microdata replaces the 1% highest and 1% lowest values, the mean values and the standard deviations are affected correspondingly. Besides, statistical noise is added to variables such as earnings when using the commands tabulate or summarise to obtain frequency tables and summary statistics for a sample. However, the variables will be adjusted proportionally so that the average numbers are unaffected.

Model 1: O-B decomposition

Table 5.4 displays the summary statistics of the samples examined in 2006 and 2007. The statistics show that the average age across the samples is almost 44 for each gender, which implies that the samples include a good adult selection. However, the average wages vary substantially between women and men by approximately NOK 100,000 each year, which is in line with the expectations and the reason for why it is important to study the gender wage gap. In detail, women earn on average NOK 349,561 and 370,908, while men earn on average NOK 456,657 and NOK 488,090 in 2006 and 2007, respectively.

Table 5.4: Summary statistics of samples (O-B decomposition)

			2006					
Variable	**Mean**	**Std.dev**	**Obs.**	**Min**	**Max**	**25%**	**50%**	**75%**
Female								
Age	43.7327	10.2803	489098	25	63	35	43	52
Birth year	1962.2673	10.2803	489098	1943	1981	1954	1963	1971
Wages	349561.0757	115104.6949	489098	121599	829746	278695.25	329889.5	394259
Log wages	12.7157	0.3107	489098	11.7085	13.6289	12.5379	12.7065	12.8848
Male								
Age	43.7712	10.34633928	761772	25	63	35	43	52
Birth year	1962.2288	10.34633928	761772	1943	1981	1954	1963	1971
Wages	456656.6715	189285.615	761772	128146	1178073	336902	407316.5	519874.25
Log wages	12.9588	0.3750	761772	11.7609	13.9794	12.7275	12.9173	13.1613

			2007					
Variable	**Mean**	**Std.dev**	**Obs.**	**Min**	**Max**	**25%**	**50%**	**75%**
Female								
Age	43.7691	10.2970	506559	25	63	35	43	52
Birth year	1963.2309	10.2970	506559	1944	1982	1955	1964	1972
Wages	370907.5483	124258.9881	506559	130815	896078	294538.5	348844	416138.5
Log wages	12.7741	0.3121	506559	11.7815	13.7058	12.5932	12.7624	12.9388
Male								
Age	43.8571	10.42328757	774640	25	63	35	43	52
Birth year	1963.1429	10.42328757	774640	1944	1982	1955	1964	1972
Wages	488090.3674	205566.8361	774640	137753	1277359	357899.75	433999.5	556954.25
Log wages	13.0236	0.37902853	774640	11.8332	14.0603	12.7880	12.9808	13.2302

Table 5.5 provides summary statistics of the variables included in the empirical analysis for 2006, while Table 5.6 provides summary statistics of the variables included in the empirical analysis for 2007. These tables display the mean values of the variables used to measure differences in labour-market characteristics in the Oaxaca-Blinder decomposition. The results of the summary statistics are quite similar for 2006 and 2007. In particular, the tables show that women exceed men in years of education. Women have, on average more than 14 years of education, while men have, on average less than 14 years of education. Despite the female advantage in education, men exceed women in labour market experience. Men have, on average, almost 23 years of experience, while women have, on average, around 22.5 years of experience.

Furthermore, the statistics show that women are more likely to work in a large city than men. Around 41% of all women and around 38% of all men work in a metropolitan area. Men and women seem to be equally distributed across the regions of Norway, but appear to work in very different industries. Women are, on average, more likely to work within education and health and social work. At the same time, men are, on average, more likely to work within wholesale and retail trade and real estate, renting and business activities. This finding is in line with what Statistics Norway has found, and suggest a gender-segregated labour market.

Table 5.5: Summary statistics of variables for 2006

			Male									Female					
Variable	**Mean**	**Std.dev**	**Obs.**	**Min**	**Max**	**25%**	**50%**	**75%**	**Variable**	**Mean**	**Std.dev**	**Obs.**	**Min**	**Max**	**25%**	**50%**	**75%**
Education	13.8447	2.3822	761772	10	18	13	13	16	Education	14.2357	2.3382	489098	10	18	13	13	16
Experience	22.9006	10.5770	761772	4	43	14	23	31	Experience	22.4847	10.8520	489098	3	43	13	22	31
Experience^2	636.3105	504.5806	761772	16	1849	196	529	961	Experience^2	623.3278	505.3035	489098	9	1849	169	484	961
Metro area	0.3745	0.4840	761772	0	1	0	0	1	Metro area	0.4086	0.4916	489098	0	1	0	0	1
Northern	0.0905	0.2869	761772	0	1	0	0	0	Northern	0.0993	0.2990	489098	0	1	0	0	0
Trøndelag	0.0870	0.2818	761772	0	1	0	0	0	Trøndelag	0.0822	0.2747	489098	0	1	0	0	0
Western	0.2716	0.4448	761772	0	1	0	0	1	Western	0.2409	0.4276	489098	0	1	0	0	0
Southern	0.0550	0.2280	761772	0	1	0	0	0	Southern	0.0405	0.1970	489098	0	1	0	0	0
Eastern	0.3790	0.4851	761772	0	1	0	0	1	Eastern	0.3804	0.4855	489098	0	1	0	0	1
Industry C	0.0440	0.2052	761772	0	1	0	0	0	Industry C	0.0149	0.1211	489098	0	1	0	0	0
Industry D	0.2014	0.4010	761772	0	1	0	0	0	Industry D	0.0821	0.2745	489098	0	1	0	0	0
Industry E	0.0141	0.1179	761772	0	1	0	0	0	Industry E	0.0000	0.0000	489098	0	0	0	0	0
Industry F	0.1177	0.3223	761772	0	1	0	0	0	Industry F	0.0115	0.1068	489098	0	1	0	0	0
Industry G	0.1425	0.3495	761772	0	1	0	0	0	Industry G	0.1144	0.3183	489098	0	1	0	0	0
Industry H	0.0118	0.1081	761772	0	1	0	0	0	Industry H	0.0224	0.1479	489098	0	1	0	0	0
Industry I	0.0967	0.2956	761772	0	1	0	0	0	Industry I	0.0462	0.2100	489098	0	1	0	0	0
Industry J	0.0262	0.1598	761772	0	1	0	0	0	Industry J	0.0352	0.1843	489098	0	1	0	0	0
Industry K	0.1219	0.3272	761772	0	1	0	0	0	Industry K	0.0981	0.2975	489098	0	1	0	0	0
Industry L	0.0782	0.2685	761772	0	1	0	0	0	Industry L	0.0998	0.2997	489098	0	1	0	0	0
Industry M	0.0598	0.2371	761772	0	1	0	0	0	Industry M	0.1353	0.3421	489098	0	1	0	0	0
Industry N	0.0545	0.2270	761772	0	1	0	0	0	Industry N	0.2952	0.4561	489098	0	1	0	0	1
Industry O	0.0311	0.1736	761772	0	1	0	0	0	Industry O	0.0399	0.1958	489098	0	1	0	0	0

Table 5.6: Summary statistics of variables for 2007

			Male									Female					
Variable	**Mean**	**Std.dev**	**Obs.**	**Min**	**Max**	**25%**	**50%**	**75%**	**Variable**	**Mean**	**Std.dev**	**Obs.**	**Min**	**Max**	**25%**	**50%**	**75%**
Education	13.8572	2.3861	774640	10	18	13	13	16	Education	14.2737	2.3444	506559	10	18	13	15	16
Experience	22.9653	10.6652	774640	3	43	14	23	32	Experience	22.4837	10.8692	506559	3	43	13	22	31
Experience^2	641.1503	508.9282	774640	9	1849	196	529	1024	Experience^2	623.6572	506.4069	506559	9	1849	169	484	961
Metro area	0.3750	0.4841	774640	0	1	0	0	1	Metro area	0.4046	0.4908	506559	0	1	0	0	1
Northern	0.0899	0.2861	774640	0	1	0	0	0	Northern	0.0980	0.2973	506559	0	1	0	0	0
Trøndelag	0.0870	0.2819	774640	0	1	0	0	0	Trøndelag	0.0817	0.2738	506559	0	1	0	0	0
Western	0.2719	0.4449	774640	0	1	0	0	1	Western	0.2431	0.4290	506559	0	1	0	0	0
Southern	0.0553	0.2286	774640	0	1	0	0	0	Southern	0.0426	0.2020	506559	0	1	0	0	0
Eastern	0.3780	0.4849	774640	0	1	0	0	1	Eastern	0.3795	0.4853	506559	0	1	0	0	1
Industry C	0.0446	0.2065	774640	0	1	0	0	0	Industry C	0.0157	0.1242	506559	0	1	0	0	0
Industry D	0.1980	0.3985	774640	0	1	0	0	0	Industry D	0.0810	0.2728	506559	0	1	0	0	0
Industry E	0.0131	0.1137	774640	0	1	0	0	0	Industry E	0.0000	0.0000	506559	0	0	0	0	0
Industry F	0.1209	0.3261	774640	0	1	0	0	0	Industry F	0.0121	0.1095	506559	0	1	0	0	0
Industry G	0.1414	0.3484	774640	0	1	0	0	0	Industry G	0.1129	0.3164	506559	0	1	0	0	0
Industry H	0.0116	0.1072	774640	0	1	0	0	0	Industry H	0.0219	0.1463	506559	0	1	0	0	0
Industry I	0.0970	0.2960	774640	0	1	0	0	0	Industry I	0.0458	0.2090	506559	0	1	0	0	0
Industry J	0.0268	0.1614	774640	0	1	0	0	0	Industry J	0.0350	0.1837	506559	0	1	0	0	0
Industry K	0.1264	0.3323	774640	0	1	0	0	0	Industry K	0.1013	0.3017	506559	0	1	0	0	0
Industry L	0.0769	0.2664	774640	0	1	0	0	0	Industry L	0.0997	0.2996	506559	0	1	0	0	0
Industry M	0.0589	0.2354	774640	0	1	0	0	0	Industry M	0.1349	0.3416	506559	0	1	0	0	0
Industry N	0.0537	0.2254	774640	0	1	0	0	0	Industry N	0.2956	0.4563	506559	0	1	0	0	1
Industry O	0.0307	0.1724	774640	0	1	0	0	0	Industry O	0.0394	0.1945	506559	0	1	0	0	0

Model 2: Event study

Table 5.7 shows the summary statistics of the two birth cohorts for both samples by gender. The table shows that around 49% of women and men in the original sample have their first child in the first birth cohort (2006) and that the remaining 51% of the sample have their first child in the second birth cohort (2007). The shares are quite similar for the individuals who have their one and only child during 2006-2007. Around 49% of women in the restricted sample have their first and only child in the first birth cohort, while the remaining 51% have their first and only child within the second birth cohort. Similarly, around 48% of men in the restricted sample have their first and only child in the first birth cohort, while the remaining 52% have their first and only child within the second birth cohort.

Table 5.7: Summary statistics of birth cohorts

		All individuals (Original sample)						
Variable	**Mean**	**Std.dev**	**Obs.**	**Min**	**Max**	**25%**	**50%**	**75%**
Female								
Child06	0.4934	0.5	776439	0	1	0	0	1
Child07	0.5066	0.5	776439	0	1	0	1	1
Male								
Child06	0.4938	0.5	998777	0	1	0	0	1
Child07	0.5062	0.5	998777	0	1	0	1	1

		Individuals with only one child (Restricted sample)						
Variable	**Mean**	**Std.dev**	**Obs.**	**Min**	**Mix**	**25%**	**50%**	**75%**
Female								
Child06	0.4853	0.4998	68987	0	1	0	0	1
Child07	0.5147	0.4998	68987	0	1	0	1	1
Male								
Child06	0.4815	0.4997	87221	0	1	0	0	1
Child07	0.5185	0.4997	87221	0	1	0	1	1

Table 5.8 shows the summary statistics of the original sample, including all individuals, and the more restricted sample including individuals who had their first and only child during 2006-2007. For each sample, the table provides summary statistics for mothers and fathers separately. The table shows that women on average are a few years younger than men, which is why it is necessary to control for age in the regression specification. More specifically, women are, on average, almost 35 years old, while men, on average, are almost 38 years old.

Further, the table shows that women have lower average wages than men, which is in line with the observations of the O-B sample. Female wages are on average NOK 298,351 and NOK 338,273 in the respective samples, while male wages are on average NOK 435,534 and NOK 461,372 for the original and the restricted samples, respectively. The statistics show that women are more likely to work part-time than men, but most women still work full-time. Additionally, women tend to work in the public sector rather than the private sector as opposed to men who have a significant share in the private sector. These findings are quite interesting as they are consistent with numbers from Statistics Norway.

Table 5.8: Summary statistics of samples (Event study)

		All individuals (Original sample)						
Variable	**Mean**	**Std.dev**	**Obs.**	**Min**	**Max**	**25%**	**50%**	**75%**
Female								
Age	34.9079	9.8288	776439	21	66	28	33	39
Birthyear	1974.1844	9.3016	776439	1944	1987	1971	1976	1980
Wages	298351.0971	175513.7419	667891	3802	891804	179795.25	293539	394781.5
Log wages	12.3007	1.0175	667891	8.2433	13.7010	12.0996	12.5898	12.8861
Part-time	0.2276	0.4193	776439	0	1	0	0	0
Full-time	0.4837	0.4997	776439	0	1	0	0	1
Public sector	0.4730	0.4993	776439	0	1	0	0	1
Private sector	0.3959	0.4890	776439	0	1	0	0	1
Male								
Age	37.9930	9.7119	998777	21	66	31	37	43
Birthyear	1971.0333	9.0837	998777	1944	1986	1967	1973	1977
Wages	435533.6313	249251.4053	874616	4000	1181361	290760	411942	557741
Log wages	12.6917	1.0180	874616	8.2940	13.9822	12.5803	12.9286	13.2316
Part-time	0.0843	0.2779	998777	0	1	0	0	0
Full-time	0.6949	0.4605	998777	0	1	0	1	1
Public sector	0.2007	0.4005	998777	0	1	0	0	0
Private sector	0.7409	0.4381	998777	0	1	0	1	1

		Individuals with only one child (Restricted sample)						
Variable	**Mean**	**Std.dev**	**Obs.**	**Min**	**Max**	**25%**	**50%**	**75%**
Female								
Age	34.8285	7.4373	68987	21	57	29	35	40
Birthyear	1973.7044	6.8016	68987	1951	1986	1969	1974	1979
Wages	338273.4402	166342.2051	68987	24063	932478	235224.25	321774	420540
Log wages	12.5827	0.6153	68987	10.0884	13.7456	12.3683	12.6816	12.9493
Male								
Age	38.0064	8.2446	87221	22	59	32	38	43
Birthyear	1970.5140	7.6863	87221	1949	1985	1966	1971	1976
Wages	461371.9726	222748.1391	87221	19908	1164444	328498.25	426306	562728.25
Log wages	12.8889	0.6510	87221	9.8989	13.9678	12.7023	12.9629	13.2406

Table 5.9 displays the summary statistics of the variables included in the event study for the original and restricted samples, separately. The means of the event time dummies suggest a balanced panel of individuals as the shares do not differ significantly over time. In the original sample, the shares are quite stable at almost 7.8 % over the twelve years, which implies that we do not have a problem with uneven numbers of observations that an unbalanced panel would have had. Even though the shares are varying from 4% to 9%

in the restricted sample, the panel appears to be quite balanced.

Table 5.9: Summary statistics of variables (Event study)

		All individuals (Original sample)						
Variable	**Mean**	**Std.dev**	**Obs.**	**Min**	**Max**	**25%**	**50%**	**75%**
Log wages	12.5228	1.0379	1542503	8.2433	13.9822	12.3442	12.7804	13.0979
t = -3	0.0735	0.2610	1775212	0	1	0	0	0
t = -2	0.0753	0.2639	1775212	0	1	0	0	0
t = -1	0.0768	0.2662	1775212	0	1	0	0	0
t = 0	0.0778	0.2679	1775212	0	1	0	0	0
t = 1	0.0779	0.2680	1775212	0	1	0	0	0
t = 2	0.0779	0.2680	1775212	0	1	0	0	0
t = 3	0.0778	0.2679	1775212	0	1	0	0	0
t = 4	0.0778	0.2678	1775212	0	1	0	0	0
t = 5	0.0777	0.2677	1775212	0	1	0	0	0
t = 6	0.0776	0.2675	1775212	0	1	0	0	0
t = 7	0.0779	0.2680	1775212	0	1	0	0	0
t = 8	0.0778	0.2679	1775212	0	1	0	0	0
Age20	0.2452	0.4302	1775212	0	1	0	0	0
Age30	0.4416	0.4966	1775212	0	1	0	0	1
Age40	0.2005	0.4004	1775212	0	1	0	0	0
Age50	0.0749	0.2633	1775212	0	1	0	0	0
Age60	0.0378	0.1907	1775212	0	1	0	0	0
Year	2009.0552	3.7200	1775212	2003	2015	2006	2009	2012

	Individuals with only one child (Restricted sample)							
Variable	**Mean**	**Std.dev**	**Obs.**	**Min**	**Max**	**25%**	**50%**	**75%**
Log wages	12.7535	0.6575	156211	9.8989	13.9678	12.5396	12.8432	13.1231
t = -3	0.0774	0.2672	156211	0	1	0	0	0
t = -2	0.0814	0.2735	156211	0	1	0	0	0
t = -1	0.0861	0.2805	156211	0	1	0	0	0
t = 0	0.0900	0.2862	156211	0	1	0	0	0
t = 1	0.0879	0.2832	156211	0	1	0	0	0
t = 2	0.0846	0.2782	156211	0	1	0	0	0
t = 3	0.0832	0.2762	156211	0	1	0	0	0
t = 4	0.0828	0.2756	156211	0	1	0	0	0
t = 5	0.0828	0.2756	156211	0	1	0	0	0
t = 6	0.0830	0.2758	156211	0	1	0	0	0
t = 7	0.0827	0.2755	156211	0	1	0	0	0
t = 8	0.0401	0.1961	156211	0	1	0	0	0
Age20	0.2000	0.4000	156211	0	1	0	0	0
Age30	0.4576	0.4982	156211	0	1	0	0	1
Age40	0.2764	0.4472	156211	0	1	0	0	1
Age50	0.0661	0.2484	156211	0	1	0	0	0
Year	2008.5330	3.4038	156211	2003	2014	2006	2008	2011

6 Results

This chapter presents the main findings from the empirical analysis. The first part reports the results of the Oaxaca-Blinder decomposition, whereas the second part reports the event study results. The last and final part summarises the main findings of the two separate models. A discussion of the main findings is presented in Chapter 7.

6.1 O-B decomposition results

This section presents the estimates of the contribution of particular labour market variables and other work-related characteristics to the gender wage gap. Table 6.1 displays the portion of the total gender wage gap in 2006 and 2007 accounted for by gender differences in each group of variables for both the human capital and the full specification, based on the Oaxaca-Blinder decomposition method. In particular, the table shows the male-female differences in the means of each variable multiplied by the corresponding male coefficients from the year wage regression. Table A2.1 in the Appendix displays the corresponding regression results.

Using Norwegian register data, I have replicated the US decomposition of the gender wage gap conducted by Blau and Kahn (2017) for 1980 and 2010. In the following, I present the decomposition results for Norway as well as providing a comparison of the results. Recall that the Blau-Kahn analysis is presented in Table A3.1 in the Appendix.

6.1.1 *O-B decomposition results for Norway*

In Table 6.1, Panel A displays the contribution of traditional human-capital variables to the gender wage gap (education, experience and region). In particular, the female advantage in education reverses the gender wage gap as women exceed men in educational attainment, causing women's relative wages to rise and the gender wage gap to decrease. Due to women's higher level of education, the corresponding variables account for a negative 10% and a negative 9% in 2006 and 2007, respectively. Also, the small female advantage in the region variables reduce the gap and account for a negative 1% in 2006. The region variables show the distribution of workers across regions. If women and men are equally distributed, the gender wage gap does not increase. This small female advantage in the region variables suggests that the average wage for women is higher in regions employing a larger share of women. In particular, women are more likely to work in high-paid jobs in cities, which increases their wages indirectly. By 2007, the region variables do not favour any gender, thus accounting for 0% of the gender wage gap. This finding suggests that in 2007, women and men were equally distributed across regions, not causing the gender wage gap to increase.

In contrast, the male advantage in the labour market experience increases the gender wage gap and accounts for 2% of the gap in both 2006 and 2007. Even though women exceed men in educational attainment, women lag in labour market experience. This finding implies that men have more years of work experience than women. Overall, the human-capital variables point in different directions as education narrows the gender wage gap, while experience increases the gap. When including the region variables, the total explained gap suggests a narrowing of the total gender wage gap. The explained portion of the gap accounts for a negative 8% and a negative 7% in 2006 and 2007, respectively. Due to these narrowing effects, the unexplained gender wage gap accounts for 108% (0.2631 log points) and 107% (0.3122 log points) of the total gender wage gap.

Panel B shows the decomposition using the full specification. Mainly, it displays the contribution of conventional human capital characteristics, region and industry variables to the gender wage gap. The effects of the human capital and region variables are quite similar to the results in panel A, implying that the impact of these measures of human capital operates primarily within industries. Consequently, the contribution of these characteristics does not substantially change when we add industries to the regression specification. In detail, Panel B shows a substantial male advantage in the industry variables, accounting for 26% and 27% of the gender wage gap in 2006 and 2007, respectively. This male advantage in the industry variables implies that many women work in low-paid industries as opposed to men, which eventually increases the gender wage gap.

In summary, education and region variables have a positive effect on the gender wage gap, which causes the gap to decrease. On the other hand, labour market experience and industry variables have adverse effects on the gender wage gap, causing differentials between women and men to increase. Due to the negative effects of the industry variables, the explained gender wage gap accounts for 17% and 18% in 2006 and 2007, respectively. Thus, when adding industry variables, the explained portion of the gender wage gap increases substantially. Compared to the human capital specification, the unexplained portion of the gender wage gap has decreased and account for 83% (0.2025 log points) and 82% (0.2051 log points) in 2006 and 2007, respectively. The reduction in the unexplained gender wage gap indicates that gender segregation in industries is an essential factor in explaining the Norwegian gender wage gap as women and men tend to work in different industries. Nevertheless, there is a significant portion that cannot be explained by either human capital, regions or industries variables.

Table 6.1: Decomposition of the gender wage gap in 2006 and 2007

	2006		2007	
	Log points	Per cent of gender gap	Log points	Per cent of gender gap
Variables	(1)	(2)	(3)	(4)
Panel A. Human-capital specification				
Education variables	-0.0235	-10%	-0.0248	-9%
Experience variables	0.0051	2%	0.0045	2%
Region variables	-0.0015	-1%	-0.0008	0%
Total explained	-0.0199	-8%	-0.0211	-7%
Total unexplained	0.2631	108%	0.3122	107%
Total pay gap	0.2432	100%	0.2911	100%
Panel B. Full specification				
Education variables	-0.0256	-11%	-0.0273	-11%
Experience variables	0.0052	2%	0.0048	2%
Region variables	-0.0025	-1%	-0.0017	-1%
Industry variables	0.0631	26%	0.0686	27%
Total explained	0.0403	17%	0.0444	18%
Total unexplained	0.2025	83%	0.2051	82%
Total pay gap	0.2428	100%	0.2495	100%

6.1.2 *Comparison of results*

As detailed in Panel A, the female advantage in traditional human capital variables slightly increases women's wages, causing the gender wage gap to decrease. Notably, women's higher levels of education contribute significantly to the narrowing of the gender wage gap. The female advantage is in line with what Blau and Kahn (2017) found for the US, where women increased their educational attainment over the 1980-2010 period, causing the contribution of education to change from 2.7% of the gender wage gap to a negative 7.9%. Despite the female improvement of educational attainment in the US, women seem to be slightly better educated relative to men in Norway. Moreover, the small female advantage we observe in the region variables in 2006, cannot be found for the US as region variables favoured men over women in both 1980 and 2010.

Furthermore, there is a small male advantage in labour market experience in both 2006 and 2007 by 2%. Compared to Blau and Kahn (2017), the contribution of labour market experience to the gender wage gap differs substantially between the US and Norway. In the US, the labour market experience accounted for 23.9% and 15.9% in 1980 and 2010, respectively. This comparison shows that gender differences in experience are substantially higher in the US than in Norway. Specifically, this suggests that Norwegian women do not lack work experience relative to men as much as American women, indicating that women participate more in the Norwegian labour market. This finding is consistent with the expectations to female employment rates which differ between the US and Norway (Kunze, 2018; OECD, 2020b). Overall, the explained part of the gender wage gap looks quite different for Norway than for the United States. While obtaining a reversed explained gender wage gap, suggesting a female advantage in human capital factors, Blau and Kahn (2017) showed a substantial male advantage in the explained gap for the US, accounting for 28.6% and 14.8%of the total gap in 1980 and 2010, respectively. As a consequent of the difference in the explained portions, a more significant unexplained gap in Norway is observed.

As observed in Panel B, the human capital variables do not change substantially in the full specification, suggesting that the impact of human capital variables generally operate within industries. Likewise, Blau and Kahn (2017) showed that human capital variables accounted for almost the same amount in the full specification. In contrast, the full specification of the US decomposition included occupation and industry variables as well as unionisation in addition to the human capital variables. Thus, the comparison of the full specification between the decompositions becomes challenging. However, Blau and Kahn (2017) showed that the explained portion of the gender wage gap increased significantly and accounted for 51.5% (0.2459 log points) and 62.0% (0.1434 logs points) of the total gender wage gap in 1980 and 2010. These results are quite different from the results in Panel B. As most people in Norway belong to a trade union, the most significant difference in the full specification, therefore, becomes the occupation variables. Despite the differences in the specifications, there is a more significant unexplained gap in the decomposition of the Norwegian gender wage gap. The big question, therefore, becomes whether the unexplained gap is due to lacking variables, labour market discrimination or other unexplained factors.

Overall, women's investments in human capital have increased, making women surpass men in educational attainment in both the US and Norway. However, women still lag men in labour market experience, but to a more considerable extent in the US. While region variables indicate a small female advantage in the decomposition, the US decomposition showed a male advantage in region variables. The contribution of human capital variables does not significantly change in the full specification. However, the explained portion of the gender gap increases in the full specification for both countries. Unlike the US decomposition, the full specification is only extended with industry variables. Although the industry variables account for a significant share of the gap, the lack of occupation variables is a drawback of the model. The lack of variables may lead to a potential problem of omitted variable bias and makes the comparison of the results challenging.

6.2 Event study results

The event study presents the impacts of children on the trajectory of wages for men and women in Norway. Figure 6.1-6.5 display the event time coefficients before and after the birth of the first child, estimated from Equation 4.1, for both genders. Table A2.2 in the Appendix displays the separate female and male regression results, including the event time coefficients. The figures provide details on how female and male wages evolve separately relative to two years before childbirth. Due to normalisation, the graphs start at zero in event time -2 and show the relative changes in wages. Thus, the event study allows for determining the extent to which children impact the gender wage gap.

Figure 6.1 displays the gender-specific impacts of children on wages for the full population, including all mothers and fathers having their first child during 2006-2007. The vertical axis shows wages relative to two years before childbirth in terms of log points, while the horizontal axis shows event time in terms of calendar years. The horizontal axis covers the time window, which starts three years before childbirth and ends eight years after the childbirth. Event time -3 is not included due to insignificant regression results. The vertical orange line represents the timing of the childbirth, denoted as $t = 0$, the year in which the individual has his/her first child.

In particular, Figure 6.1 suggests that female wages begin to diverge one year before the birth of the first child and that a significant drop occurs at the timing of the childbirth. Afterwards, female wages seem to converge relatively quickly in the first years after birth, causing female wages to almost return to the same level it was two years before birth. Moreover, Figure 6.1 suggests that men's wages increase in the year before and after the childbirth but eventually start to decrease three years after the childbirth, continuing until the end of the time window. Thus, men appear to be unaffected around the birth of the first child. Overall, the converging curves suggest only a short-term child penalty for mothers in which women's wages recover within eight years after the childbirth.

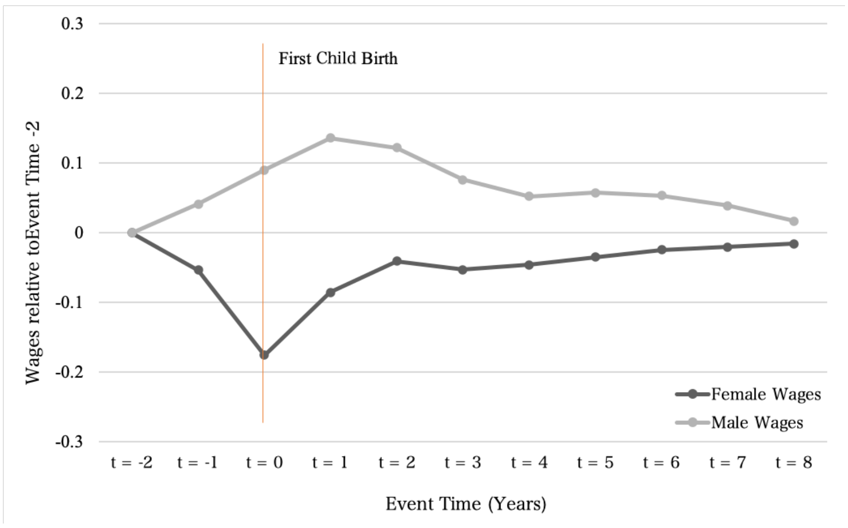

Figure 6.1: All individuals (mothers and fathers)

Figure 6.2 shows the impacts of children on wages for full-time working individuals from the full population. For female wages, the curve starts to diverge one year before the childbirth and continue to drop until the year the first child is born. By the childbirth, the marginal change in women's wages is almost 20 percentage points less on average than two years before childbirth. Thus, the curve suggests a severe child penalty in wages for mothers. In contrast to 6.1, the child penalty in earnings persists throughout the time window. Although female wages start to increase the year after birth, growth already declines three years after birth, causing a negative trend in female wages until the end of the time window. Thus, female wages do not return to pre-child levels, suggesting that the child penalty is permanent in the time window we observe. This finding is in contrast to what the female curve shows in Figure 6.1, where the female child penalty was only temporary, and female wages came almost back to pre-child levels. The finding is striking and suggests that women who work full-time suffer from a more substantial child penalty.

Moreover, Figure 6.2 shows steady growth in male wages, where the wages increased the most around the birth of the first child. Compared to the wage level two years before birth, the male curve suggests that men´s wages marginally increased by 10-15 percentage points in the years after the first childbirth. Due to the negative trend in the female curve and the positive trend in the male curve, wage differences between women and men increase from two years after childbirth continuing until the end of the time window. These trends are different from what we find in Figure 6.1, where the curves almost return to original pre-child wage levels. In sum, Figure 6.2 shows a severe and persistent child-penalty for mothers who work full-time. As the male curve shows an upward sloping trend, the figure suggests an increasing gender wage gap. Finally, Figure 6.2 implies that the impacts of children on female and male wages are more considerable for individuals who work full-time, compared to the previous figure which also includes individuals who work part-time.

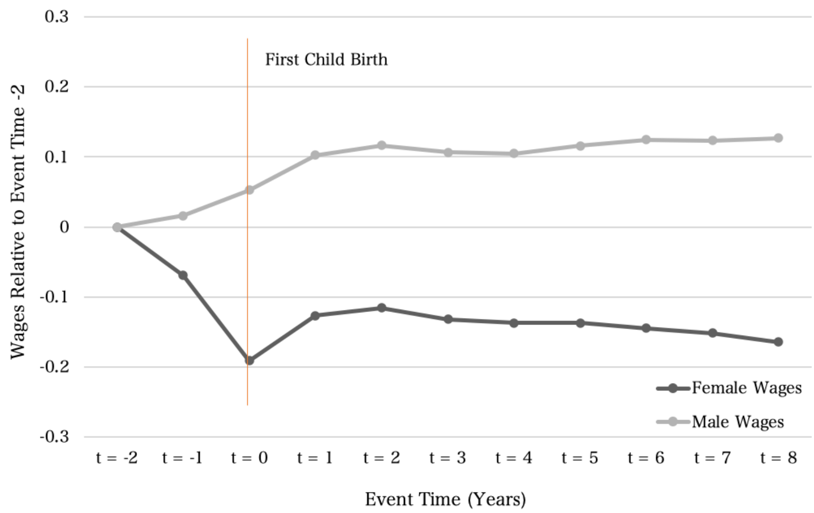

Figure 6.2: Full-time working individuals

Figure 6.3 displays the impacts of children on wages for full-time working individuals in the public sector. The female curve suggests that female wages begin to diverge one year before the childbirth and continue to decrease until the birth of the first child. By the childbirth, the relative change in women's wages is almost 20 percentage points less than two years before the birth of the first child. The female curve implies a short-term increase in female wages the two first years after birth. However, from two years after birth and onwards, the curve indicates a significant negative trend in female wages. Eight years after birth, female wages appear to be lower than in the year the child was born. This substantial decline in female wages over time is a striking finding, suggesting that women's wages are worse off when the child is of school-age then when the child was newborn.

Furthermore, Figure 6.3 shows that male wages are almost unaffected before childbirth but start to increase the year the child is born and continue to grow steadily throughout the time window. Thus, the diverging curves suggest that gender differences in wages increase accordingly. Compared to Figure 6.2, these curves diverge to a more considerable extent, indicating a severe child penalty for mothers in the public sector. However, as in the previous figure, the picture looks quite different compared to Figure 6.1. While female and male wages almost return to the original levels in 6.1, these curves diverge substantially causing the gender wage gap to increase between mothers and fathers working full-time in the public sector. The conflicting trends suggest that individuals working full-time in the public sector might be more affected by children. Overall, the female curve suggests a severe and persistent child penalty for mothers who work full-time in the public sector, which never recovers during the time window as the female and male curves diverge over time.

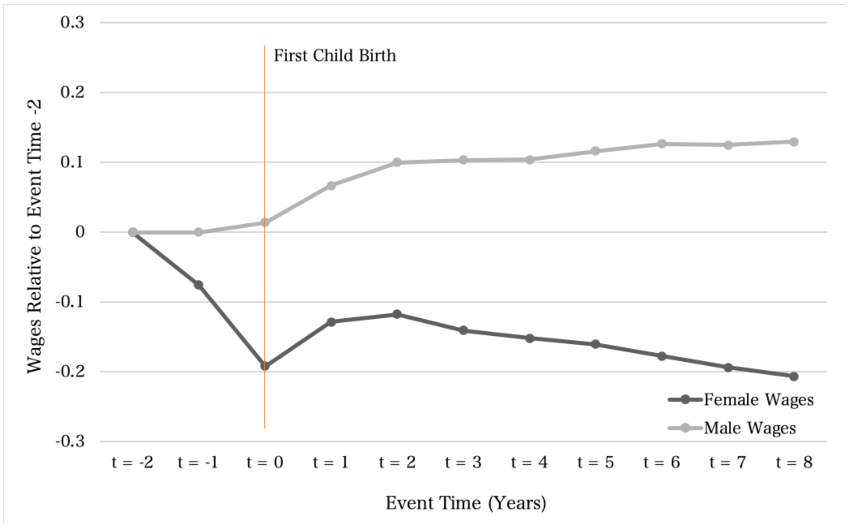

Figure 6.3: Full-time working individuals in public sector

Figure 6.4 shows the impacts of children on wages for full-time working individuals in the private sector. The female curve suggests a significant child penalty for mothers who work full-time in the private sector, starting one year before the childbirth. By the birth year, women´s wages marginally increased by 20 percentage points relative to two years before birth. The drop in female wages is the largest observed, which implies that mothers working full-time in the private sector suffer the most from having their first child. Despite a short-term increase in female wages in the first years after the childbirth, female wages start to decline already two years after birth and continues throughout the time window. Nevertheless, female wages diverge at a lower rate in the private sector compared to the public sector in the second half of the total time window.

In contrast, Figure 6.4 shows that male wages begin to increase the year before birth and continue to grow steadily in the first years after the birth before stabilising around 10 percentage points above pe-child wages. Overall, compared to Figure 6.2 and Figure 6.3, the curves appear to be more stable in the second half of the time window. Compared to Figure 6.1, the figure suggests the same as the three previous graphs; women working full-time seem to suffer a more severe and persistent child penalty.

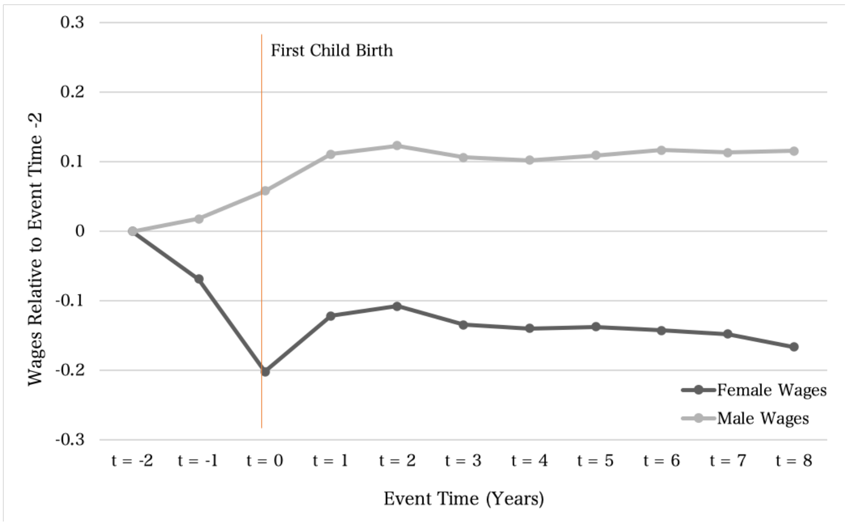

Figure 6.4: Full-time working individuals in private sector

Finally, Figure 6.5 shows the impacts of children on wages for individuals who have their first and only child during 2006-2007. Figure 6.5 shows a significant drop in female wages starting one year before childbirth continuing until the year the child is born. This figure suggests a child penalty for mothers regardless of whether they work part-time or full-time, which is different from what be observed in Figure 6.1. Besides, the figure shows that male wages start to increase the year before childbirth with the highest marginal growth rate in the years around the birth, which confirms that fathers are not negatively affected of the arrival of a child.

Overall, Figure 6.5 shows that the evolution of female and male wages are quite similar to the three previous figures, even though female wages converge to a slightly high level in the second part of the time window. This finding suggests that in Figure 6.2-6.4 some of the decline post-birth is driven by a second or further childbirths. While the three preceding figures show only full-time workers, this figure shows both full-time and part-time workers. Compared to Figure 6.1, the patterns of the wage curves differ substantially, yielding conflicting results about the impact of children on wages. This finding is striking and is a drawback of the empirical analysis as it seems implausible that part-time work explains this.

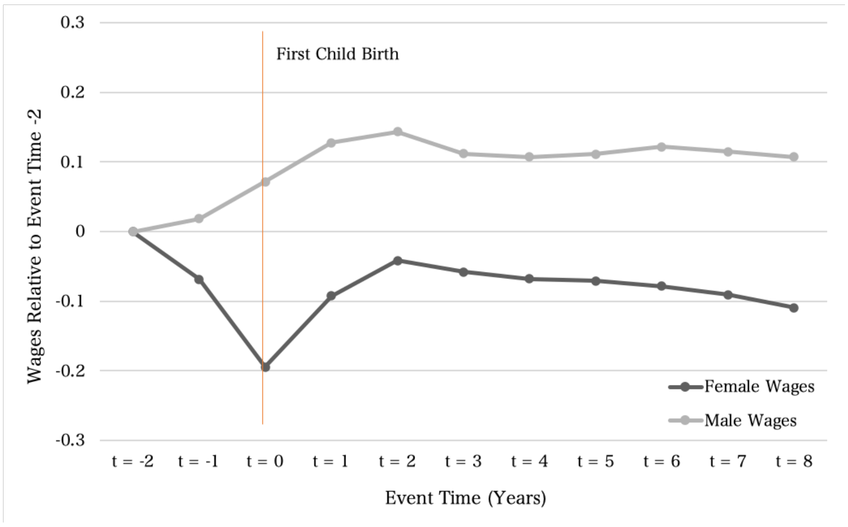

Figure 6.5: Wage and salary working individuals with only one child

Overall, Figures 6.1-6.5 show a substantial child penalty for mothers around the birth of the first child. In the aftermath of the childbirth, female wages evolve differently across the samples. Figure 6.1 shows only a short-term child penalty for mothers, while Figures 6.2-6.5 show persistent child penalties for mothers. The latter is line with what Kleven et al. (2018) found in their study for Denmark. In contrast to women, men seem to be unaffected from having their first child. The event study also shows that mothers working full-time suffer from an extensive child penalty in both the public and the private sector. However, when studying both full-time and part-time mothers, the results are conflicting. Finally, children have substantial impacts on female and male wages and serve as a potential explanation of the gender wage gap.

6.3 Summary of results

The empirical analysis includes two replication studies. As a starting point for analysing the gender wage gap, I decompose the gender wage gap based on the US study conducted by Blau and Kahn (2017). Then, I investigate whether gender inequality is due to children based on the Danish study by Kleven et al. (2018).

Using a traditional Oaxaca-Blinder decomposition method from a cross-sectional perspective, I show that traditional human capital factors in the aggregate decrease the gender wage gap in both 2006 and 2007, which is not in line with what Blau and Kahn (2017) found in their study. When adding industry variables in the full specification, the explained portion of the gender gap increases substantially in both years, suggesting that the impacts of the measures of human capital operate primarily within industries. The first hypothesis suggests that the gender wage gap is partly due to gender segregation. The decomposition confirms that a small portion of the gender wage gap is due to women working in low-paid industries. The second hypothesis suggests that the Blau-Kahn analysis is relevant for understanding the gender wage gap, also for the Norwegian labour market. As shown, conventional human capital factors decreased the gender wage gap, resulting in a more substantial unexplained component compared to the US. Whether the model suffers from a potential omitted variable bias or not, the decomposition provides insight into the contribution of human capital factors and industry factors.

To investigate the gender wage gap further, I conduct an event study analysing the impacts of children on the trajectory of wages for men and women. In particular, I study how the wages evolve for the individuals who had their first child between 2006-2007. The third hypothesis suggests that the event study gives new insight into children's impacts on the gender wage gap. As shown, there are substantial changes in female and male wages around the birth of the first child. Mainly, women's wages appear to be affected significantly by the arrival of a child, suggesting a motherhood penalty in earnings. These findings are essential when human capital has become decreasingly important and yields insight into alternative explanations. The fourth hypothesis suggests that the arrival of the first child leads to a child penalty in earnings of almost 30%. The event study results reveal a significant child penalty in wages for mothers in the years around the birth of the first child of around 20 percentage points. However, female wages evolve differently in the aftermath of the childbirth, causing the child penalty to vary across samples. Overall, the motherhood penalty observed in all graphs suggests that children have negative impacts on the gender wage gap, increasing gender differences in earnings over time.

7 Discussion

The main goal of the study has been to answer the research question: *"What are the sources of the persistent gender wage gap, and is gender inequality due to children?"*. In order to answer this research question, I have conducted a traditional Oaxaca-Blinder decomposition based on Blau and Kahn (2017) and an event study of children's impacts on women and men's wages based on Kleven et al. (2018). This chapter discusses the main findings in light of previous literature and the main economic explanations, and lastly discusses the study's limitations and suggestions for future research. The results of the empirical analysis are detailed in Chapter 6.

7.1 Discussion of empirical strategy and findings

This thesis contributes to the existing literature by providing insight into the extent of the gender wage gap from different points of view. The empirical analysis is based on Norwegian register data for the full population, which allows for analysing the gender wage gap in depth. Mainly, full-population register data allows for studying non-mother to mother transitions for the specific year they gave birth. When pooling birth cohorts, I might compensate for the small sample size, but also introduce differences and potential problems measuring changes around the event of having a child. These could be among others, changing macroeconomic factors and work experience.

The Oaxaca-Blinder decomposition shows that human capital factors in the aggregate decrease the gender wage gap in 2006 and 2007. Despite the small male advantage in labour market experience, women's higher levels in education cause the gender wage gap to narrow. When controlling for region variables in the human capital specification, the explained portion of the gender wage gap accounts for -8% and -7% in 2006 and 2007, while when controlling for region variables and industry variables in the full specification, the explained portion accounts for 17% and 18% in 2006 and 2007, respectively. As the contribution of human capital variables does not change across the specifications, the increase in the explained portion suggests that the impacts of the measures of human capital operate primarily within industries. Even though the decomposition reveals essential gender differences in industry, the unexplained portion of the gender wage gap persists and accounts for 83% and 82% in 2006 and 2007, respectively. Compared to the US decomposition by Blau and Kahn (2017), human capital factors contribute less to the gender wage gap in Norway, causing the unexplained gap to be substantially higher.

The O-B decomposition analysis yields interesting insight into the gender wage gap. In particular, human capital factors taken together decrease the gender wage gap in both specifications. Compared to Blau and Kahn (2017), the contribution of human capital factors to the gender wage gap is significantly lower in Norway than in the US. These findings are consistent with the expectations when looking at female employment rates in the US and Norway detailed in Kunze (2018). Another interesting finding is the contribution of industry variables to the gender wage gap. While the human capital specification suggests a female advantage in the explained portion of the gap, the full explanation implies a male advantage when controlling for industry variables. This finding indicates that parts of the gender wage gap are driven by gender differences within industries, which is consistent with the explanation of pronounced gender segregation in the Norwegian labour market, as well as what Blau and Kahn (2017) found in their study. Thus, gender differences in wages are partly driven by the fact that women, on average, tend to work in low-paid industries, while men work in high-paid industries.

Although the proportion of the explained gap increased in the full specification when controlling for industry, most of the gender wage gap is due to unexplained factors in both 2006 and 2007. Compared to Blau and Kahn (2017), a more significant unexplained gap is observed in this decomposition than for the US. The big question is, therefore, whether the unexplained gap is due to the different wages regressions, labour market discrimination or other factors that the O-B decomposition does not capture. The decomposition lacks variables compared to the US decomposition, and the fact the specification do not control for gender differences in occupations is a considerable drawback of the model. Since industries most likely are correlated with occupations, there might be a problem of omitted variables bias, which could lead to biased and inconsistent coefficient estimates and make the interpretations and comparisons challenging. Despite the potential problem of omitted variable bias, the O-B decomposition method suggest that the unexplained gap is due to labour market discrimination. In that case, the decomposition suggests that labour market discrimination accounts for more than 100% of the total gender wage gap in the human capital specification and more than 80% of the total gender wage gap the full specification. However, it seems odd that discrimination represents such a large proportion of the total gender wage gap. This suggests that there might be other explanations beyond traditional supply-side factors, gender segregation and labour market discrimination that contribute to the gender wage gap.

Women and men can be different in ways that are important for their wage levels and which are legitimate explanations for the remaining differences we observe in the O-B decomposition. For instance, gender disparities can be due to differences in motivation, competitive instinct and self-confidence, which are challenging to test statistically. However, the fact that the unexplained wage gap remains unchanged over time in many developed countries can indicate that some of the gender differences contributing to the gender wage gap cannot be easily changed. This calls for a discussion of whether children can explain gender differences in earnings. Using an event study approach, I explore whether children can serve as a potential source of the gender wage gap by investigating children's impacts on female and male wages for individuals who had their first child during 2006-2007. The event study approach allows for taking advantage of the time dimension, following mothers and fathers for eight years after the childbirth until the child becomes of school-age.

The event study results provide interesting insights into children's impacts on female and male wages over time. In particular, there is a substantial child penalty in earnings for mothers around the birth of the first child for all samples examined, while at the same time men are unaffected. By the birth year, female wages appear to have marginally decreased by around 20 percentage points relative to two years before childbirth, for all samples, which is nearly in line with what Kleven et al. (2018) found in the Danish study. Nevertheless, this number seems quite large and calls for a discussion of whether the empirical analysis captures other effects than just children's impact. After the childbirth, women's wage evolves differently across the samples. There is a short-term child penalty in earnings for mothers in the sample, including all individuals, whereas there is a more severe child penalty for mothers in the remaining samples. This finding yields conflicting results which imply that children have both temporary and persistent impacts on female and male wages. Thus, the question becomes which of the graphs display the right picture of children´s impacts on wages.

So what are the reasons for the drop in female wages around the birth of the first child? The event study focuses primarily on full-time working mothers, both in the public and private sector. The results suggest that mothers who work full-time suffer from an extensive child penalty in earnings that persists during the time window. Previous research has shown that there is a strong relationship between the selection process of returning to work and the wage process around childbirth. Ejrnæs and Kunze (2013) found evidence for negative selection into work, suggesting that mothers who suffer from relatively significant wage losses around childbirth are more likely to return to full-time employment after the arrival of the child. These findings suggest that the observed drop in female wages can be due to negative selection as women might have switched from full-time to part-time positions when having their first child, or reduced their working hours substantially. Angelov et al. (2016) showed that there was a substantial drop in working hours related to parenthood, which supports the argument that women adjusts their working hours when having their first child.

Nevertheless, non-random selection into work can potentially bias the samples used to compute the gender wage gap because of women's noncontinuous working life (Kunze, 2018). This introduces the potential problem that the motherhood penalty can be overestimated since the specification does not control for the fact that women might not work full-time over the entire time window. Since the motherhood penalty in wages appears to be around 20 %, it might be that the estimates are overestimated. Even though the estimates might be overestimated, there is strong evidence for a motherhood penalty in wages. A motherhood penalty might cause women lose a substantial amount of earnings in the long-run.

Moreover, the drop in female wages can also be due to women changing jobs to more child-friendly firms or industries around the birth of the first child. As shown, women's wages start to drop the year before the childbirth for all samples, indicating that women make labour market adjustments during pregnancy. Goldin (2014) argues that many women tend to change to more flexible firms and industries when having their first child, but that flexible jobs come at a high price. This means that women might give up on high-paid jobs to achieve flexibility at work. Substantial evidence also shows that extended parental leave might affect women's wages negatively increasing the likelihood of women never returning to work. Unfortunately, the event study cannot capture whether the parental leave is the reason for the wage drop, but becomes an interesting question for further research.

Finally, the question becomes whether or not women make unprofitable choices. As shown in the O-B decomposition, women exceed men in educational attainment, which means that gender differences in earnings are not due to women lacking education. However, the decomposition shows that women lag in labour market experience and that there is gender segregation in the Norwegian labour market. This suggests that women might make unprofitable choices, but does not say anything about why they choose to work in such industries. The substantial motherhood penalty observed in the event study might imply that women choose to work in low-paid industries and occupations or choose to reduce working hours as a result of having their first child. Nevertheless, it is quite challenging to distinguish whether these are voluntary choices or whether women are discriminated by employers.

Overall, the O-B decomposition and the event study yield interesting insights into the gender wage gap. Using Norwegian register data and two different complementary methods, potential explanations behind the gender wage gap have been discussed. While the first method provides insight into the decreasing importance of traditional supply-side factors, the second method provides insight into children's contribution to gender inequality. Even though Kleven et al. (2018) showed that almost all remaining gender inequality is due to children, the empirical strategy cannot determine whether children can explain the unexplained gap obtained in the OB decomposition. Nevertheless, we see that children have significant impacts on female and male wages and that mothers particularly suffer from a severe child penalty. Thus indicates that children is a crucial source of the gender wage gap. As shown, the significant motherhood penalty can be a result of labour market adjustments related to childbirth, such as negative selection into work and changes in occupations and industries. Parental leave and labour market discrimination are other potential explanations.

7.2 Limitations and suggestions for future research

Using Microdata has provided many possibilities for analysing highly detailed register data. However, there are also certain limitations associated with processing and analysing the register variables in Microdata. A more detailed note on the user experience is available in the Appendix. Moreover, there are other challenges associated with empirical strategy as well. In the following, I highlight the limitations of the thesis in the form of suggestions for further research.

The main drawback of the O-B decomposition is the lack of variables. Since the analysis employs data from 2006-2007, and the variables for occupations are not available until 2009, occupations are left out of the decomposition. The absence of this variable makes it difficult to compare the results to Blau and Kahn (2017) who do control for occupations. Additionally, the analysis utilises potential experience instead of actual experience, which means that interruptions in work-life such as parental leave are not taken into account. Thus, the O-B decomposition would be more precise if the regression specification controlled for occupations and used actual experience instead of potential experience.

The event study is only a partial analysis of the impacts of children on female and male wages. The next step would be a more comprehensive analysis of the effects of children, for instance, by analysing the impacts by the number of children, or by linking the event study results to the O-B decomposition, showing the fraction of child-related gender inequality in a dynamic O-B decomposition. It would also have been appropriate to study full-time workers with one child relative to full-time-workers with more than one child.

Another limitation is the conflicting results found in the event study. Figure 6.1 suggest only a short-term motherhood penalty, while the other figures imply substantial motherhood penalties over time. Even though the subsamples are based on the original sample, the results differ substantially. These findings make it more challenging to conclude to what extent children impact female and male wages, and it seems implausible that part-time work explains this. Plotting additional samples would help to understand why these results differ. For instance, it could be interesting to study the differences between women working full-time relative to women in part-time employment.

8 Conclusion

In this thesis, I have studied the gender wage gap using Norwegian register data collected from Microdata. As a starting point, I have replicated the US decomposition by Blau and Kahn (2017) from a cross-sectional perspective, decomposing gender differences in wages for 2006 and 2007, into a component accounted for by differences in traditional labour market characteristics and an unexplained component. Second, I have adopted the event study methodology suggested by Kleven et al. (2018) in order to investigate children's impact on female and male wages. Using an event study approach allows for observing how female and male wages have evolved over the 2003-2015 period for those who had their first child during 2006-2007.

The Oaxaca-Blinder decomposition has shown that traditional human capital factors have become less important and that human capital variables in the aggregate decrease the gender wage gap. The industry variables, on the other hand, suggest a gender-segregated labour market, which increases the gender wage gap. However, the decomposition showed that most of the gender wage gap is due to unexplained factors, indicating labour market discrimination against women. Since conventional human capital factors can no longer explain the gender wage gap, I turn to investigate alternative explanations beyond traditional supply-side factors using an event study approach to study children´s impact on female and male wages.

The event study has provided insight into the impacts of children on female and male wages. Even though the event study is only a first-step analysis of children's impacts, the event time coefficients indicate that mothers who work full-time suffer from a significant and persistent motherhood penalty in earnings, while fathers appear to be unaffected. The motherhood penalty suggest a drop in female wages of around 20% in the birth year. When studying both full-time and part-time workers, the results are conflicting, suggesting a short-term as well as a persistent child penalty in the wages of mothers.

As discussed, there are several potential explanations to the drop in female wages around the birth of the first child. Women might reduce working hours, change to part-time positions or even withdraw for the labour force as a result of having the first child. As women's wages start to drop one year before childbirth, women seem to make labour market adjustments in preparation for childbirth. Additionally, women might change to more family-friendly jobs to achieve flexibility which can cause their wages to decrease substantially. Also, extended parental leave seems to affect women's wages negatively. Thus, the arrival of the first child has some significant impacts on women´s wages in the short-term. Whether mothers make unprofitable choices or are victims of labour market discrimination is difficult to determine in this study. Nevertheless, it seems that children can explain parts of the gender wage gap, together with gender segregation in the labour market.

References

Angelov, N., Johansson, P., & Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics*, *34*(3), 545–579. doi: https://dx.doi.org/10.1086/684851

Askvik, T. (2020). *Saktere nedgang i lønnsforskjellene mellom kvinner og menn.* Statistics Norway. Retrieved from `https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/saktere-nedgang-i-lonnsforskjellene-mellom-kvinner-og-menn`

Becker, G. S. (2009). *Human capital: A theoretical and empirical analysis, with special reference to education.* University of Chicago press.

Blau, F. D., Ferber, M. A., & Winkler, A. E. (2014). *The economics of women, men and work.* Pearson Higher Ed.

Blau, F. D., & Kahn, L. M. (2008). Women's work and wages. *The new Palgrave dictionary of economics*, *8*, 762–72.

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, *55*(3), 789-865.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, 436-455.

Budig, M. J., & England, P. (2001). The wage penalty for motherhood. *American sociological review*, 204–225.

Bütikofer, A., Jensen, S., & Salvanes, K. G. (2018). The role of parenthood on the gender gap among top earners. *European Economic Review*, *109*, 103–123. doi: https://doi.org/10.1016/j.euroecorev.2018.05.008

Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American journal of sociology*, *112*(5), 1297–1338.

Ejrnæs, M., & Kunze, A. (2013). Work and wage dynamics around childbirth. *The Scandinavian Journal of Economics*, *115*(3), 856–877.

Fortin, N. M. (2008). The gender wage gap among young adults in the united states the importance of money versus people. *Journal of Human Resources*, *43*(4), 884–918.

Gardeazabal, J., & Ugidos, A. (2004). More on identification in detailed wage decompositions. *Review of Economics and Statistics*, *86*(4), 1034–1036.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, *104*(4), 1091-1119. doi: https://doi.org/110.1257/aer.104.4.1091

Goldin, C., Katz, L. F., & Kuziemko, I. (2006). *The homecoming of american college women: The reversal of the college gender gap* (Working Paper No. 12139). National Bureau of Economic Research. doi: https://doi.org/110.3386/w12139

Hamre, K. (2017). *Fedrekvoten – mer populær enn noen gang.* Statistics Norway. Retrieved from https://www.ssb.no/befolkning/artikler-og-publikasjoner/fedrekvoten-mer-populaer-enn-noen-gang--298200

Hardoy, I., & Schøne, P. (2008). The family gap and family friendly policies: the case of norway. *Applied Economics*, *40*(22), 2857–2871.

Hotz, V. J., Johansson, P., & Karimi, A. (2017). *Parenthood, family friendly workplaces, and the gender gaps in early work careers* (NBER Working Papers No. 24173). National Bureau of Economic Research, Inc.

Jaumotte, F. (2003). Female labour force participation: past trends and main determinants in oecd countries. *OECD Economics Department Working Papers*, *376*.

Kleven, H., Landais, C., & Søgaard, J. E. (2018). *Children and gender inequality: Evidence from denmark* (Working Paper No. 24219). National Bureau of Economic Research. doi: https://doi.org/10.3386/w24219

Kristoffersen, S. (2017). *Lønnsforskjellene mellom kvinner og menn vedvarer.* Statistics Norway. Retrieved from https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/lonnsforskjellene-mellom-kvinner-og-menn-vedvarer--331542

Kunze, A. (2008). Gender wage gap studies: consistency and decomposition. *Empirical Economics*, *35*(1), 63-76. doi: https://doi.org/10.1007/s00181-007-0143-4

Kunze, A. (2015). The family gap in career progression. *Research in Labor Economics*, *41*, 115–142.

Kunze, A. (2018). The gender wage gap in developed countries. *The Oxford Handbook of Women and the Economy*, 369.

Kunze, A., & Ejrnæs, M. (2004). Wage dips and drops around first birth [IZA Discussion Paper]. (1011).

Lalive, R., Schlosser, A., Steinhauer, A., & Zweimüller, J. (2014). Parental leave and mothers' careers: The relative importance of job protection and cash benefits. *Review of Economic Studies*, *81*(1), 219–265.

Lalive, R., & Zweimüller, J. (2009). How does parental leave affect fertility and return to work? evidence from two natural experiments. *The Quarterly Journal of Economics*, *124*(3), 1363–1402.

Lucifora, C., Meurs, D., & Villar, E. (2017). *Children, earnings and careers in an internal labor market.* Retrieved from http://www.aiel.it/cms/cms-files/submission/all20170831111248.pdf

Microdata. (2020). *Forventet/avtalt arbeidstid, intervaller.* Author. Retrieved 2020-06-15, from https://microdata.no/discovery/variable/query=arbeidsti&name=REGSYS_ARBTID&date=1972-01-01&revision=1

Mincer, J. (1962). On-the-job training: Costs, returns, and some implications. *Journal of* *political Economy*, *70* (5, Part 2), 50–79.

Mincer, J., & Polachek, S. (1974). Family investments in human capital: Earnings of women. *Journal of political Economy*, *82* (2, Part 2), 76–108.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, 693–709.

OECD. (2018). *Is the last mile the longest? economic gains from gender equality in nordic countries*. OECD Publishing. doi: https://doi.org/10.1787/9789264300040-en

OECD. (2020a). *Earnings and wages - gender wage gap - oecd data.* Retrieved from `https://data.oecd.org/earnwage/gender-wage-gap.htm`

OECD. (2020b). *Employment rate (indicator).* doi: https://doi.org/10.1787/1de68a9b-en

Sigle-Rushton, W., & Waldfogel, J. (2007). Motherhood and women's earnings in anglo-american, continental european, and nordic countries. *Feminist economics*, *13* (2), 55–91.

SSB. (2019). *Tettsteders befolkning og areal.* Author. Retrieved from `https://www.ssb.no/befolkning/statistikker/beftett/aar`

SSB. (2020). *Likestilling.* Statistics Norway. Retrieved 2020-06-03, from `https://www.ssb.no/befolkning/faktaside/likestilling`

UN. (2020). *Gender equality.* Retrieved 2020-06-03, from `https://www.un.org/en/sections/issues-depth/gender-equality/`

Waldfogel, J. (1998). Understanding the" family gap" in pay for women with children. *Journal of economic Perspectives*, *12* (1), 137–156.

Appendix

A1 Variables

Table A1.1: Variables (O-B decomposition)

Variable name	Microdata code	Description
A. Sample selection		
gender	BEFOLKNING_KJOENN	Imports all individuals into dataset by gender
male		1: Male, 0: Female
female		1: Female, 0: Male
year_of_birth	BEFOLKNING_FOEDSELS_AAR_MND	Birth year and month. Used to generate birth year.
birth_year		Birth year. Used to generate age.
age		Age of individuals
regstat	BEFOLKNING_REGSTAT	Whether the individual is resident
labstat	REGSYS_YRKSTAT	Whether the individual earns wages
fulltime	REGSYS_ARBTID	Whether the individual work full-time
industry	REGSYS_NARING_SN2002	Industry. Used to generate industry dummies.
B. Regression		
wages	INNTEKT_WLONN	Income salary
lnw		Logarithm of wages
educ_level	NUDB_BU	Level of education. Used to generate education.
educ		Years of education (dummy variables)
exper		Labour-market experience
exper2		Experience squared
region	BOSATTEFDT_BOSTED	Residence municipality (dummy variables)
metro	REGSYS_ARBKOMM	Metropolitan area (dummy variable)
industry dummies		Industry main employer

Table A1.2: Variables (Event study)

Variable name	Microdata code	Description
A. Sample selection		
gender	BEFOLKNING_KJOENN	Imports all individuals into dataset by gender
male		1: Male, 0: Female
female		1: Female, 0: Male
year_of_birth	BEFOLKNING_FOEDSELS_AAR_MND	Birth year and month. Used to generate birth year.
birthyear		Birth year. Used to generate age.
children	BEFOLKNING_BARN_I_HUSH	Number of children aged 0-17
child06		All individuals having their first child in 2006
child07		All individuals having their first child in 2007
industry	REGSYS_NARING_SN2002	Industry. Used to generate industry dummies.
public		Industries within public sector
private		Industries within private sector
workhours	REGSYS_ARBTID	Expected/agreed working hours, intervals.
full_time		Working at least 30 hours a week
part_time		Working less than 30 hours a week
regstat	BEFOLKNING_REGSTAT	Whether the individual is resident
year		Calendar years
age		Age of individuals
B. Regression		
wages	INNTEKT_WLONN	Income salary
lnw		Logarithm of wages
event time dummies		Time relative to the birth of the first child
age dummies		Groups of ages

Table A1.3: Industrial classification

SSB Class	Code	Sector
A - Agriculture, hunting and forestry	0-4	Private
B - Fishing	5	Private
C - Mining and quarrying	10-14	Private
D - Manufacturing	15-37	Private
E - Electricity, gas and water supply	40-41	None
F - Construction	45	Private
G - Wholesale and retail trade, repair of motor vehicles, motorcycles and personal and household goods	50-52	Private
H - Hotels and restaurants	55	Private
I - Transport, storage and communication	60-64	Private
J - Financial intermediation	65-67	Private
K - Real estate, renting and business activities	70-74	Private
L - Public administration and defence, compulsory and social security	75	Public
M - Eductaion	80	Public
N - Health and social work	85	Public
O - Other community, social and personal service activities	90-93	Public
P - Private households with employed persons	95	None
Q - Extra-territorial organisations and bodies	99	None
Z - Unkown	0	None

A2 Regression tables

Table A2.1: Regression results (O-B decomposition)

	2006				**2007**			
	Human capital specification		Full specification		Human capital specification		Full specification	
	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Education	0.0602***	0.0608***	0.0654***	0.0652***	0.0596***	0.0599***	0.0655***	0.0651***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Experience	0.0289***	0.0236***	0.0289***	0.0235***	0.0286***	0.0235***	0.0289***	0.0237***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002	(0.0002)	(0.0002)
Experience^2	-0.0005***	-0.0004***	-0.0005***	-0.0004***	-0.0005***	-0.0004***	-0.0005***	-0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Metropolitan area	0.0694***	0.0826***	0.0690***	0.0704***	0.0686***	0.0837***	0.0667***	0.0697***
	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0009)	(0.0008)	(0.0008)	(0.0008)
Region dummies	YES	YES	YES	YES	YES	YES	YES	YES
Industry dummies	NO	NO	YES	YES	NO	NO	YES	YES
Constant	11.7807***	11.5655***	11.7289***	11.5770***	11.8556***	11.6334***	11.7885***	11.6371***
	(0.0031)	(0.0033)	(0.0032)	(0.0035)	(0.0031)	(0.0033)	(0.0032)	(0.0035)
Observations	761776	489094	761776	489094	774644	506563	774644	506563
R^2	0.18511	0.23069	0.25723	0.28809	0.17909	0.22481	0.25600	0.29154
Adjusted R^2	0.18510	0.23068	0.25721	0.28806	0.17909	0.22480	0.25598	0.29151

Standard errors in parentheses

***p<0.01 **p<0.05 *p<0.1

Table A2.2: Regression results (Event study)

	Dependent variable: Log wages									
	Model 1 (All individuals)		Model 2 (Full-time)		Model 3 (Public sector)		Model 4 (Private sector)		Model 5 (Only 1 child)	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
t = -3	0.0034	0.0038	0.001	0.0038	0.0005	0.0023	0.0018	0.0052	0.0022	-0.0018
	(0.0034)	(0.0042)	(0.0021)	(0.0023)	(0.0040)	(0.0028)	(0.0024)	(0.0037)	(0.0061)	(0.0071)
t = -1	0.0412***	-0.0538***	0.0163***	-0.0690***	0.0009	-0.0756***	0.0179***	-0.0690***	0.0186**	-0.0685***
	(0.0032)	(0.0039)	(0.0030)	(0.0033)	(0.0057)	(0.0039)	(0.0034)	(0.0051)	(0.0087)	(0.0100)
t = 0	0.0899***	-0.1756***	0.0524***	-0.1912***	0.0137*	-0.1917***	0.0580***	-0.2019***	0.0714***	-0.1947***
	(0.0032)	(0.0039)	(0.0041)	(0.0046)	(0.0079)	(0.0054)	(0.0047)	(0.0072)	(0.0120)	(0.0139)
t = 1	0.1360***	-0.0856***	0.1022***	-0.1264***	0.0664***	-0.1287***	0.1107***	-0.1219***	0.1276***	-0.0927***
	(0.0032)	(0.0039)	(0.0054)	(0.0060)	(0.0104)	(0.0071)	(0.0061)	(0.0094)	(0.0157)	(0.0182)
t = 2	0.1218***	-0.0409***	0.1163***	-0.1157***	0.0995***	-0.1175***	0.1228***	-0.1078***	0.1432***	-0.0414*
	(0.0033)	(0.0040)	(0.0067)	(0.0074)	(0.0129)	(0.0088)	(0.0076)	(0.0117)	(0.0195)	(0.0226)
t = 3	0.0763***	-0.0527***	0.1066***	-0.1321***	0.1030***	-0.1406***	0.1059***	-0.1342***	0.1121***	-0.0582**
	(0.0034)	(0.0041)	(0.0081)	(0.0089)	(0.0155)	(0.0106)	(0.0092)	(0.0141)	(0.0234)	(0.0272)
t = 4	0.0524***	-0.0458***	0.1051***	-0.1369***	0.1040***	-0.1519***	0.1020***	-0.1397***	0.1070***	-0.0681**
	(0.0035)	(0.0042)	(0.0095)	(0.0105)	(0.0182)	(0.0124)	(0.0107)	(0.0165)	(0.0273)	(0.0317)
t = 5	0.0574***	-0.0348***	0.1160***	-0.1373***	0.1163***	-0.1608***	0.1090***	-0.1376***	0.1115***	-0.0710*
	(0.0037)	(0.0045)	(0.0109)	(0.0120)	(0.0208)	(0.0143)	(0.0123)	(0.0189)	(0.0313)	(0.0364)
t = 6	0.0537***	-0.0243***	0.1245***	-0.1447***	0.1262***	-0.1776***	0.1166***	-0.1426***	0.12165***	-0.0784*
	(0.0039)	(0.0047)	(0.0122)	0.0135***	(0.0235)	(0.0161)	(0.0139)	(0.0213)	(0.0353)	(0.0410)
t = 7	0.0390***	-0.0205***	0.1235***	-0.1515***	0.1246***	-0.1935***	0.1132***	-0.1476***	0.1149***	-0.0907**
	(0.0041)	(0.0050)	(0.0136)	(0.0151)	(0.0261)	(0.0179)	(0.0154)	(0.0237)	(0.0392)	(0.0456)
t = 8	0.0167***	-0.0158***	0.1270***	-0.1641***	0.1294***	-0.2065***	0.1153***	-0.1662***	0.1073**	-0.1093**
	(0.0043)	(0.0053)	(0.0150)	(0.0167)	(0.0288)	(0.0198)	(0.0170)	(0.0262)	(0.0433)	(0.0504)
Age dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year	0.0627***	0.0598***	0.0488***	0.0690***	0.0472***	0.0726***	0.0503***	0.0695***	0.0489***	0.0664***
	(0.0005)	(0.0005)	(0.0014)	(0.0016)	(0.0027)	(0.0019)	(0.0016)	(0.0025)	(0.0041)	(0.0047)
Constant	-113.3129***	-107.7871***	-85.0714***	-125.8415***	-81.8489***	-132.9649***	-88.0713***	-126.6616***	-85.3404***	-120.5856***
	(0.9164)	(1.1101)	(2.8315)	(3.1478)	(5.4371)	(3.7491)	(3.2083)	(4.9498)	(8.1240)	(9.4703)
Observations	874613	667892	665127	369204	139010	196549	529397	170403	87218	68990
R^2 i	0.1208	0.1016	0.2698	0.3038	0.3362	0.3633	0.2663	0.2798	0.2460	0.2366
R^2 between	0.0451	0.0114	-0.0021	-0.0564	-0.0074	-0.0101	0.0079	-0.0431	0.0135	0.0150
R^2 total	0.0777	0.0582	0.1080	0.1195	0.1333	0.1684	0.1063	0.1087	0.1043	0.1146

Standard errors in parentheses

***p<0.01 **p<0.05 *p<0.1

A3 Results of the Blau-Kahn study

Table A3.1: Decomposition of the gender wage gap by Blau and Kahn

	1980		2010	
	Log points	Per cent of gender gap	Log points	Per cent of gender gap
Variables	(1)	(2)	(3)	(4)
Panel A. Human-capital specification				
Education variables	0.0129	2.7%	-0.0185	-8.0%
Experience variables	0.1141	23.9%	0.0370	16.0%
Region variables	0.0019	0.4%	0.0003	0.1%
Race variables	0.0076	1.6%	0.0153	6.6%
Total explained	0.1365	28.6%	0.0342	14.8%
Total unexplained	0.3405	71.4%	0.1972	85.2%
Total pay gap	0.4770	100.0%	0.2314	100.0%
Panel B. Full specification				
Education variables	0.0123	2.6%	-0.0137	-5.9%
Experience variables	0.1005	21.1%	0.0325	14.0%
Region variables	0.0001	0.0%	0.0008	0.3%
Race variables	0.0067	1.4%	0.0099	4.3%
Unionization	0.0298	6.2%	-0.0030	-1.3%
Industry variables	0.0457	9.6%	0.0407	17.6%
Occupation variables	0.0509	10.7%	0.0762	32.9%
Total explained	0.2459	51.6%	0.1434	62.0%
Total unexplained	0.2312	48.5%	0.0880	38.0%
Total pay gap	0.4770	100.0%	0.2314	100.0%

Source: Blau and Kahn (2017)

Note: Sample includes full time non-farm wage and salary workers ages 25–64 with at least twenty-six weeks of employment. Entries are the male–female differential in the indicated variables multiplied by the current year male log wage coefficients for the corresponding variables. The total unexplained gap is the mean female residual from the male log wage equation.

A4 Microdata user experience

This section provides a summary of my experiences using Microdata for my thesis. The purpose is to share knowledge and experience with others who intend to use Microdata for future research. In addition to providing insight into the user-friendliness, I also highlight the opportunities and limitations of Microdata.

Microdata is an online service developed by the Norwegian Centre for Research Data (NSD) and Statistics Norway, which provide instant access to unique register data on the full population in Norway. These registers include the Norwegian National Registry, the National Education Database (NUDB), the Register for Personal Tax Payers, Labour market data and FD-Trygd (event history database). Privacy is protected, allowing researchers to process and analyse the data without being able to view or gain knowledge of personal data. All output is subject to confidentiality measures, including restrictions for the sample size (more than 1000 individuals per sample), winsorisation of outliers and statistical noise added into descriptive statistics.

Since the register data cannot be viewed, downloaded or extracted due to privacy protection requirements, I have processed and analysed available register variables using the online analytical platform. This platform offers a data processing function, with support for population delimitations, linking and the development of new variables. In this way, register variables on the full population can be merged to create high-quality data sets linking individuals and years via the personal identification number. For statistical purposes, the platform offers a wide range of functions including among others, descriptive statistics, linear regression analysis and panel data analysis.

In order to perform panel data analysis, the data must be organised differently than in ordinary regression analysis. Panel data sets become very large if the entire population is included in the data set. Thus, the panel data function requires a small sample from the population to be merged with years; otherwise, the online server has a problem running the analysis. As shown in preceding sections, Microdata cannot import multiple times into the same panel data set, nor can ordinary cross-sectional data be mixed with panel data. This is a challenge associated with using Microdata as it becomes time demanding to organise the data sets. However, the panel data function in Microdata provides the opportunity to include the time component in the analysis making the data sets larger and with higher quality.

Since all the output is subject to confidentiality, Microdata does not provide functions for making plots or regression tables. The regression function prints the regression results which can be copied into Excel. As a result, I have created all the regression tables, summary statistics and figures manually using Excel. This has been a time demanding process, and there is a small risk that things may have gone wrong. However, I have run all the regression several times to check that the results are correctly constructed.

Overall, it has been a great experience using Microdata to collect, generate and analyse register data. The analytical platform is easy to use and has similar commands to other statistical software. It takes time to get to know the commands and the analytical tools, but ones you know how Microdata works the process is quite joyful. Be aware that Microdata takes some time to run the scripts, and when errors occur, nothing is telling you what the errors mean. However, Microdata provides a user manual which is very useful describing how the platform works and all the commands included. Microdata also provides user support which has been very handy and helpful for questions regarding technicalities and opportunities Microdata offers.

Although it has been time demanding to organise the data, especially the panel data sets, working in Microdata has increased my programming and econometric skills. It has been challenging, but very exciting, to create my own data sets based on available register data from Statistics Norway. It was very satisfying to see that the data sets could be used to analyse the gender wage gap and that the thesis contributes by providing new insight into the Norwegian labour market. I strongly recommend students and other researchers to use Microdata as many interesting register variables can be used for research. Also, new variables are continually being added, and the platform is continuously being improved.