Universitetet i Stavanger

DET SAMFUNNSVITENSKAPELIGE FAKULTET,
HANDELSHØGSKOLEN VED UIS

MASTEROPPGAVE

STUDIEPROGRAM: Økonomi og administrasjon	OPPGAVEN ER SKREVET INNEN FØLGENDE SPESIALISERINGSRETNING: Økonomisk analyse ER OPPGAVEN KONFIDENSIELL? (**NB**! Bruk rødt skjema ved konfidensiell oppgave)

TITTEL:
The change in Norway's Cash-for-Care program and its effect on maternal labor supply

ENGELSK TITTEL:
The change in Norway's Cash-for-Care program and its effect on maternal labor supply

FORFATTER(E)		VEILEDER:
Studentnummer: 216139 ……	Navn: Diana Hoang Nguyen ……	Venke Furre Haaland

OPPGAVEN ER MOTTATT I TO – 2 – INNBUNDNE EKSEMPLARER

Stavanger, ……/…… 2016 Underskrift administrasjon:……

Preface

This paper marks the end of my master's degree program at the Business School at the University of Stavanger. The writing process has been challenging at times, but it has mostly been very interesting and educational.

First, I would like to thank my supervisor, Venke Furre Haaland, for all the great support, feedback, and discussions this spring. I would also like to thank Professor Mari Rege for her helpful comments and advice, which helped me decide on this subject.

The data applied in the analysis in this publication are based on Labor Force Survey 2008, 1st quarter - 2014, 4th quarter. The data are provided by Statistics Norway, and prepared and made available by the NSD - Norwegian Centre for Research Data. Neither Statistics Norway, nor NSD are responsible for the analysis/interpretation of the data presented here.

Abstract

This paper investigates how a withdrawal of the Cash-for-Care program in Norway affected the mothers' employment decision. In 2012, the Norwegian government removed 2-year-old children's eligibility for the Cash-for-Care benefit. This reform was expected to incentivize the mothers of 2-year-olds to enter the labor market or encourage them to work more. I employ a difference-in-differences method, which exploits the variation in mothers' exposure to the reform. This is to see how the reform affected the mothers of 2-year-olds compared to mothers of older children, who were not affected by the reform at all.

The main results show a positive effect in the short run, when the children are 1- or 2-year-olds, but this is not statistically significant. On the other hand, my findings suggest a stronger effect in the longer run. I found a significant increase in the probability of being a full-time worker for mothers of children aged 3 years old at the end of the year.

However, the results from the placebo analyses are statistically significant, which threatens the identifying assumption. This indicates that the full-time employment trends between mothers of 2-year-olds and mothers of older children are not parallel in the pre-reform era. Therefore, the results from this analysis might not be valid and it makes them difficult to interpret.

Table of contents

List of tables

List of figures

Appendix list

1 Introduction

Since the 1970s, the female labor force participation rate has been increasing, it grew from around 50 % to 80 % for women between the ages of 25 and 54 (Statistics Norway, 2016b). There are similar strong increases in other OECD countries, such as the United States and other Nordic countries. By 2012, the female labor participation rate for the same age group had increased to about 75 % in the United States (OECD Statistics, 2016b).

Such a development in Norway may have been thanks to the many family policy reforms, where one of the goals was to help families combine work and family life. Several studies have analyzed these reforms and its potential effects on maternal supply. Finseraas, Hardoy, and Schøne (2015) found a significant increase in mothers' employment after the school reform in 1997, when the school starting age was lowered from 7 to 6. According to Havnes and Mogstad (2011), the expansion of public subsidized daycare facilities, which started in 1975, had little effect on mothers' labor supply. The expansion seems to have mainly led to a decrease in the use of informal child care instead (p. 1461). A paper analyzing the introduction of a 4-week paternal quota in 1993 for the paid parental leave presented similar results. The paternal quota did not seem to have changed the paternal nor the maternal labor supply (Cools, Fiva, & Kirkebøen, 2015).

The Norwegian government[1] introduced the Cash-for-Care (CFC) program in August 1998. In contrast to the other family policy reforms, this program might have incentivized mothers to work less. Empirical results from previous studies suggest that the CFC program reduced the maternal labor supply (Naz, 2004; Rønsen, 2009; Schøne, 2004b). Parents included in the program[2] could receive a tax-free payment from the government if they did not utilize publicly subsidized day care (St.prp. nr. 53 (1997-98), 1998). Thereby parents could have a choice between working or staying at home with their children, while still receiving a certain income.

In August 2012, the government[3] changed the program so that the benefit would no longer be eligible for 2-year-olds, but only for 1-year-olds. One political argument for the change was to increase female labor supply and promote gender equality. Another argument was to increase attendance in daycare centers to promote integration (Prop. 8 L (2011–2012), 2011, pp. 1-2).

[1] Bondevik's first cabinet.
[2] Parents of 1-year-olds were included from the start, parents of 2-year-olds were added to the program in 1999.
[3] Stoltenberg's second cabinet.

In this paper, I investigate how removing 2-year-olds from the CFC program influenced maternal labor supply. Did the removal of eligibility give the intended effect of increasing the mothers' full-time employment rate, working hours or both? I want to find out how the affected mothers adapted to the change in the short run and the longer run, whether they made the decision at the extensive[4] margin or at the intensive[5] margin.

This change in the CFC program could affect the maternal labor supply through at least two different channels. First, the removal of eligibility decreased the relative price of public subsidized daycare centers. This is because they no longer had to forego the cash benefit if they used subsidized care. According to previous studies, higher childcare costs affect married women's labor participation rate negatively (Blau & Robins, 1988; Powell, 2002; Ribar, 1992). This would indicate that the 2012 reform affected the mothers' labor supply positively. Second, mothers who did not use publicly subsidized day care on a full-time basis suffered a decrease in their non-labor income. A mother allocates her time between three different activities: work in the market, household production, and leisure activity. If her non-labor income decreases, there will be a negative income effect, which can cause a reduction in her demand for leisure, if this is a normal good. This means that she will have more time left to work in the market, which would increase her labor supply.

Several studies have analyzed the introduction of the CFC program; both Naz (2004) and Drange (2012) found a reduction in maternal supply in the short run, through a difference-in-differences estimation. Naz reported a decrease of working hours by 2.85 hours, while Drange reported a 4-percentage point decrease in the probability of being a full-time worker. On the other hand, very few studies have investigated the effect of the change in the CFC program in 2012, and this motivated me to do an analysis on this. Dahl (2014) used a logistic regression model and found a 3-percentage point increase in the probability to work for mothers of 2-year-olds in the short run after the 2012 reform. This paper differs from Dahl's study by using a difference-in-differences method, additionally I analyze the effect both in the short run and in the longer run. Furthermore, I investigate whether a change in labor supply was mainly thanks to a change at the extensive or intensive margin.

Exploring the causal effect of the removal of eligibility for a government transfer on female employment is challenging because of omitted variable bias. For instance, recipients and

[4] Being full-time employed or not.
[5] Changing the number of working hours.

eligible non-recipients of such a transfer might differ in ways that we cannot observe. In this paper, I utilize a natural experiment, the CFC reform in 2012, where the government removed a universal cash benefit to parents of 2-year-olds. This benefit incentivized mothers to leave the labor force or reduce their working hours. As the reform was universal to all women with a 2-year-old child, selection into treatment will not bias my empirical results.

Some parents had their period of eligibility for the CFC benefit shortened from 23[6] to 11[7] months. This CFC reform did not affect women whose children reached the age of 3 years before August 2012, but it did affect women with younger children. This allows me to control for other factors than the CFC reform that may have affected maternal labor supply. Thereby I employ a difference-in-differences method that exploits the variation in mothers' exposure to the reform. The identifying assumption is that without the existence of this CFC reform, the time trends in the full-time employment rates would be similar for both groups of mothers.

For this study, I used data from the Norwegian Labor Force Survey, provided by the Norwegian Centre for Research Data (NSD). The dataset includes outcome variables such as full-time employment and the number of weekly working hours. Background characteristics such as age, education, and marital status are also available. I focus on data collected from 2010 to 2014, which are about two years before and after the change in the CFC program occurred.

In my main analysis, I did find evidence that supports the argument that the 2012 reform affected maternal labor supply positively, especially at the extensive margin. The results suggest a stronger effect for children aged 3 years old, compared to children aged 2 years old at the end of the year. The probability for a mother whose child was 2 or 3 years old appears to increase by 9.91 percentage points, which is a significant increase. However, results from the placebo analyses threaten the identifying assumption; this indicates that the main results may not be valid. As for my subsample analyses, I did not find any evidence to support the argument that the reform affected mothers differently depending on education or marital status.

The rest of my paper is organized as follows: Section 2 will present the institutional settings, with special focus on the CFC program. While section 3 presents the labor supply theory and the hypotheses, which I will test in this paper. In section 4, I will discuss the previous

[6] The parents of children aged 13 to 35 months old used to be eligible.
[7] The parents of children aged 13 to 23 months old are currently eligible.

literature concerning this program and the mechanisms related to it. Section 5 describes the empirical method I have chosen, followed by a presentation of the data and sample I have used in section 6. The main, subsample and other results are presented and discussed in section 7. Lastly, section 8 will summarize and conclude the paper.

2 Institutional settings

2.1 The Cash-for-Care program

In 1998, the Norwegian government introduced the CFC program, which first included 1-year-olds and expanded to include 2-year-olds shortly after. The CFC benefit was a universal benefit any family could claim as long as they fulfilled certain requirements. It was a tax-free payment from the government to parents whose child did not attend a publicly subsidized daycare facility on a full-time basis. Parents who did not use publicly subsidized day care could receive the full[8] benefit, while parents who used it on a part-time basis could receive part of the full benefit (St.prp. nr. 53 (1997-98), 1998).

One of the main goals of this program was to give the parents more freedom in choosing what childcare arrangement to give their child (St.prp. nr. 53 (1997-98), 1998). This CFC benefit can financially help those who wish to stay at home to take care of their child themselves. State subsidies were given to daycare facilities, and this worked as an indirect public transfer to families who used publicly subsidized day care. Families who did not choose this type of child care could not receive this benefit. Therefore, introducing the CFC benefit[9] will help distribute public transfers equally between families, regardless of their choice of childcare arrangement. Lastly, this program can contribute towards the goal of having available places at daycare centers for every family who wants to use public day care.

The CFC program[10] was one of the central issues during Norway's parliamentary election in 1997. The debates regarding this and the EU issue were the most comprehensive political debates in the 1990s (Vollset, 2011, p. 243). What type of child care is the best for the children was the main topic of the CFC program debate. Should all parents freely choose the childcare arrangement they wanted and would they choose the one that was the best for their child? People were worried not every family could afford to have their child attend a public subsidized daycare center. In August 1997, the parental payments for a place in a daycare center could be as high as 36 000 NOK[11] per year (St.prp. nr. 53 (1997-98), 1998).

Opponents for this program argued that it is not a positive program for gender equality, because the most likely candidates to stay at home were the mothers rather than the fathers. There might not only be a short-run effect on the mothers' labor supply, but it could also

[8] 3000 NOK per month (Vollset, 2011, p. 245)
[9] The size is almost equivalent to the state subsidy for a place at a daycare center.
[10] Proposed by the center coalition.
[11] For a public daycare center, the payments could be up to 29 000 NOK per year.

affect their career in the long run. The fathers may have to work even more to make up for the income loss from the mother, and the result would be even less time for these fathers to take care of their child. Another concern was the decline in demand for daycare centers that would happen, which could lead to some centers closing down. Simultaneously, there would be a rise in demand for private childminders, a type of child care that had no supervision from the state. Therefore, they cannot be assured that someone with the needed skills and qualifications were taking care of the children (Hellevik & Koren, 2000, pp. 10-11). They were also worried it could lead to no attendance from specific children[12], which the child welfare (Bufetat) recommended should be attending a daycare center (Vollset, 2011, p. 248).

The center coalition[13] implemented the CFC program in August 1998, less than a year after they had come into power. There was a sliding scale of the CFC benefit consisting of four different levels, where parents could receive part of the benefit if they used subsidized care on a part-time basis. Table 1 describes how the scale and the size of the benefit based on daycare center attendance have had small changes over the years.

Table 1: CFC benefits per month prior to August 2012

Daycare center attendance per week agreement	**No attendance**	**15 hours or less**	**16-20 hours**	**21-30 hours**		**31 hours or more**
Percentage of the benefit	**100 %**	**80 %**	**60 %**	**45 %**		**0 %**
August 1st 1998	3000	2400	1800	1350		0
January 1st 1999	2263	1810	1357	1018		0
Daycare center attendance per week agreement	**No attendance**	**8 hours or less**	**9-16 hours**	**17-24 hours**	**25-32 hours**	**33 hours or more**
Percentage of the benefit	**100 %**	**80 %**	**60 %**	**40 %**	**20 %**	**0 %**
August 1st 1999	2263	1810	1357	905	453	0
January 1st 2000	1810	2400	1800	1200	600	0
August 1st 2003	3657	2926	2194	1463	732	0
January 1st 2006	3303	2642	1982	1321	661	0

Note: The table shows the change in the scale and size of the cash benefit from August 1, 1998 to January 1, 2006. The benefit size is measured in NOK. Source: (NAV, 2003, 2005)

The number of CFC recipients has decreased drastically over the years. The share of recipients among eligible parents declined from 73 % in 1999 to 27 % in 2009 (Hirsch, 2010). The percentage of recipients among parents of 1-year-olds have always been higher than among parents of 2-year-olds. Additionally, mothers were more likely to be CFC recipient than what the fathers were; 95.9 % of the recipients were women in 2004 (Daugstad, 2006). Parents with a non-western background seems to make use of the CFC program more often

[12] Such as children with a different mother tongue than Norwegian.
[13] Now called Bondevik's first cabinet.

than parents with a western background do. In 2007, the share of recipients among parents from Asia, Africa, South America, or European countries such as Turkey and Poland was the highest (Bakken & Myklebø, 2010, pp. 47-48).

The debates regarding the CFC program is still ongoing even after the implementation. Seeing how the share of recipients among eligible families with a migrant background has been growing through the years, while the share among all eligible families has been on the decline, the critics argue that the program prevents integration. They believe that the program leads to bad integration for both the children and mothers with a non-western background. Supporters of this program on the other hand, have faith that the families will make the decisions that are the best for them. This program gives the parents the chance to spend more time with their children and the supporters want to prioritize the families' needs.

As previously mentioned, there has been evidence of the CFC program's negative effect on maternal labor supply from empirical research (Håkonsen, 2001; Rønsen, 2000; Schøne, 2004b). Because of this, many government committees recommended a change, where 2-year-olds would be excluded from the program (Prop. 8 L (2011–2012), 2011, p. 2). Additionally, they were worried the children would less likely attend any public day care, the longer time the parents could receive the allowance. This would be bad for integration, especially for children whose mother tongue is not Norwegian.

Therefore, in September 2011, the Norwegian government declared their decision to remove 2-year olds from the CFC program starting from August 2012. They also changed the sliding scale of the program from five to two levels. From now on, you could receive either the full CFC benefit or half of the benefit. Additionally, the government increased the size of the allowance for the youngest 1-year-olds, which means that the size of the benefit from then on also depended on the child's age. See Table 2.

Table 2: CFC benefits per month from August 1, 2012 to August 1, 2014

Daycare center attendance per week agreement	**No attendance**	**19 hours or less**	**20 hours or more**
Percentage of the benefit	**100 %**	**50 %**	**0 %**
Child aged 13-18 months	5000	2500	0
Child aged 19-23 months	3303	1652	0

Note: The table shows the size of the cash benefit, which is measured in NOK, after the change happened in August 1, 2012.
Source: (NAV, 2003)

Table 3 describes which cohorts were either partly or fully affected by the removal of the CFC eligibility. Each cell represents the child's age in a given year and we can follow each cohort from year to year, by moving diagonally downwards to the right. The numbers in the cells inform us how many months of CFC eligibility the child from a specific cohort could possibly have in that given year. The dark gray cells represent the cohorts, which were fully treated; they did not have any eligibility while they were 2 years old. The light gray cells represent the cohorts that were partly treated. Some of the children in these cohorts did have eligibility as 2-year-olds, while some of them did not. Lastly, the white cells represent the cohorts with no treatment at all, which means that they were all eligible until they turned 3 years old.

Table 3: Months of eligibility

Age of the child	2009	2010	2011	2012	2013	2014
Age 1	0-11	0-11	0-11	0-11	0-11	0-11
Age 2	12	12	12	7-11	0-11	0-11
Age 3	0-11	0-11	0-11	0-7	0	0
Age 4	0	0	0	0	0	0
Age 5	0	0	0	0	0	0

Notes: The table[14] describes the nature of the treatment. Each cell shows how many months of eligibility, a child turning a certain age, could have in a specific year. The dark gray cells represent those who were fully treated and had no eligibility while they were two years old, whereas the light gray cells represent those who were partly treated. The last cohort with no treatment was born in 2008, while the first cohort, which was fully treated, was born in 2011. The children who were partly treated were born in 2009 or 2010.

2.2 Female labor supply in Norway

As mentioned in the introduction, the share of women working in Norway kept rising from the 1970s. The share of cohabiting or married women with children between 0 and 15 years old who were working rose from 74 % in 1991 to 81 % in 2004 (Tronstad, 2007, p. 10). Norway had one of the highest labor participation rates of women aged between 25 and 54 among the OECD countries in 1998, the year the CFC program was implemented (OECD Statistics, 2016b). In 2012, the year the change in the program occurred, female labor participation in Norway remained relatively high. There was a small increase in its participation rate, while the part-time employment rate for women aged 24-54 had declined. It used to be 31 % and above the OECD average in 1998, however it decreased to 20 % in 2012 (OECD Statistics, 2016a). There are gender differences in occupation choice and women are generally overrepresented in the health and social care sectors. Women also tend to work in the public sector, 58 % of public sector workers were women in 2005 (Tronstad, 2007, p. 39).

[14] Inspired by Drange and Rege (2013).

2.3 Other family policy changes

Parents of young children can receive three possible public transfers from the government; parental benefit, child benefit, and the CFC benefit. Single parents are entitled to additional public transfers, such as transitional benefit and childcare benefit. Changes in such transfers occurring around the same time as the change in the CFC program could potentially challenge my identification.

The length of the paternal quota of the parental benefit has changed several times since its introduction. The fathers whose children are born on the same day the change occurred or later, will be eligible for that specific length. The quota increased from 6 to 10 weeks in 2009 and there was yet another increase in 2011, from 10 to 12 weeks. These fathers can choose to use this quota anytime from the child is 6 weeks old up until he or she has turned 3 years old. Consequently, the quota increase affected part of the treatment group with younger children, but it did not affect anyone in the control group, which consists of older children. However, previous studies regarding this subject have not found any significant effect on maternal labor supply (Cools et al., 2015; Rege & Solli, 2013).

The transitional benefit is for single parents who have insufficient income, and the maximum amount was equal to 2G[15] in 2012. The recipients are required to be active through work or education once their youngest child has reached a certain age. The minimum activity requirement is for the single parents to work or study part-time, and this activity has been required once the child becomes 36 months old since 1998. This changed in 2012, when the recipients had to follow this requirement once the child was 12 months old. The government's reason for this reform was to help the single parents become self-supporting faster, and for them to have a stronger attachment to the labor market (Prop. 7 L (2011–2012), 2011, p. 8). This change only applied for the recipients who applied for this transitional benefit after December 31, 2011. According to Bjørnstad (2015), this change in the activity acquirement led to a 7-percentage point increase in the employment rate for the single parents of 1- or 2-year-olds. However, it is not certain how much this change affected the treatment group because their children were all born before 2012. For my analysis, I have included a control variable, which will indicate whether the mother is single or not. Furthermore, I will do a subsample analysis to compare the effect of the reform on the single mothers and married or cohabitant mothers, respectively.

[15] National insurance scheme basic amount, 1G = 82,122 NOK in 2012 (Norwegian Tax Administration, 2015).

3 Theory and hypotheses

The removal of eligibility could affect female labor supply through at least two different mechanisms, through a decrease in the non-labor income or a decrease in the relative price of public day care. In sections 3.1 through 3.3, I will describe a theoretical model, which illustrates how a mother makes her labor supply decisions. While section 3.4 explains why I expect an increase in maternal supply after the 2012 reform.

3.1 The basic model of labor-leisure choice

One of the models typically used to analyze labor supply behavior is the neoclassical model of labor-leisure choice (Borjas, 2013, p. 27). We assume that people wish to maximize their utility and satisfaction and every individual receives utility from consumption of goods and leisure. An additional assumption is that they will allocate their available time between work in the market and leisure (non-market) activity.

However, according to Mincer (1962, p. 65), we should be distinguishing between household production and leisure activity, especially for married women. This is because household work is an activity many women spend a lot of time on after they are married. Labor supply models sometimes omit household production because of the difficulty with categorizing certain activities at home. Taking care of your children can be said to be household work instead of leisure activity, but is playing with them also household work (Gronau, 1977, p. 2)? Another reason for the omission is that leisure activity and household production have the same opportunity cost[16]. However, Gronau continued saying how there are findings, showing that these two activities react differently to changes in the same variables, such as having a child.

3.2 Specialization

In a multiperson household[17], gains can be made through specialization. This can occur if a member has the greater comparative advantage in the market over their partner, while their partner is relatively more productive at home (Becker, 1980, p. 33). Someone who has a relatively higher marginal product in the market compared to at home, will more likely be specializing at working outside of home. Likewise, someone who has a relatively higher marginal product at home is more likely to do household work.

[16] Their offered or market wage rate.
[17] Formed by marriage or cohabitation.

Wives are more likely to have the relatively higher marginal productivity at home. This is because of the biological differences between men and women, the gender pay gap, and how they have traditionally been raised with different expectations (Becker, 1980, p. 37). Therefore, when a child is born into a family, it will usually affect the mother's labor supply negatively. Angrist and Evans (1998) found results supporting this. They used an instrumental variable method to analyze the relationship between having children and the labor supply of the parents. Their results indicate that children led to a decline in female labor supply, men on the other hand, had little labor supply response to the increase in family size. Another paper reported that an increase in the number of children affected a mother's reservation wage[18] positively, but this effect diminished with the child's age (Gronau, 1973). Naz (2004) investigated specialization in relation to the CFC program and found an increase in specialization as an effect. Women had less working hours after the reform, while men had an insignificant increase in theirs.

3.3 A Gronau-like model

In this section, I will describe a model, which is based on the Gronau model of time allocation and home production (Gronau, 1977, p. 6), where household production is included. We will assume that every individual will maximize his or her utility, which is summarized by a strictly concave utility function. In one of the models of family labor supply mentioned by Killingsworth (1986, pp. 131-132), every member's utility function will depend on the whole family's consumption of goods C and his or her own leisure time L.

(1) Utility function: $U(C, L, \tau)$

The preference parameter τ shows the trade-off between consumption of goods and leisure. The individual is constrained by the available time and the income they have, which is represented by a budget line. In this model, the purchased goods and the home produced goods are perfect substitutes, which means that the individual is indifferent to the combination of goods he or she consumes. For simplicity's sake, we will assume that the price for each good at the market is equal to 1.

(2) $C = C_N + C_H$

[18] Also called the asking wage, the minimum wage at which an individual is willing to accept entering the labor market.

C_N = goods purchased[19] at the market at a price P per unit of good

C_H = goods produced at home, such as clothes made in a home workshop, home-cooked food or child care, which is expressed by a household production function:

(3) $C_H = f(H)$

This function is represented by a production curve which is concave to the origin; this means that it has a decreasing marginal productivity, where $f' > 0$ and $f'' < 0$ (Gronau, 1977, p. 7). The production curve differs from person to person because their marginal productivity at home is different.

The money budget consists of the mother's earned income, which depends on her hourly offered wage rate (w) and working hours (N), and an unearned and non-labor income V. For the mother's budget constraint, we will assume that the earnings of her husband or partner is exogenous. Thus, the man's labor supply is given and his earned income is included in her non-labor income V. This is a normal assumption found in empirical literature[20] about the labor supply of young children's parents. The available hours (T) will be allocated between leisure (L), work in the market (N) and household production (H).

The money budget:

(4) $P \times C_N = w \times N + V$

The time budget:

(5) $T = H + N + L$

By putting N alone on the LHS in (5) and substituting it into (4), we can get the full constraint:

(6) $P \times C_N + w(H + L) = w \times T + V$

Rewriting the utility function by substituting (3) into (2) and then (2) into (1), we get:

(7) $U = U(C_N + f(H), L, \tau)$

In the end, this will be the utility maximization problem:

[19] With the available money one has.
[20] See (Averett, Peters, & Waldman, 1997; Ribar, 1992)

(8) $Max\ U = U(C_N + f(H), L, \tau)\ s.t.\ P \times C_N + w(H + L) = w \times T + V$

The optimal point for the individual will be where his or her highest possible indifference curve is tangent to either the household production curve or the budget line. This is illustrated in Figure 1, where the worker adapts at a point where the slope of the budget line is equal to the slope of the indifference curve that provides the highest possible utility. The marginal rate of substitution between goods and leisure will be equal either to the marginal product of work at home and their real reservation wage (9), or to their real wage rate (10).

(9) $$\frac{\frac{\delta U}{\delta L}}{\frac{\delta U}{\delta C}} = f'(H) = \frac{w*}{P}$$

(10) $$\frac{\frac{\delta U}{\delta L}}{\frac{\delta U}{\delta C}} = \frac{w}{P}$$

Figure 1: A worker will adapt where their utility is maximized.

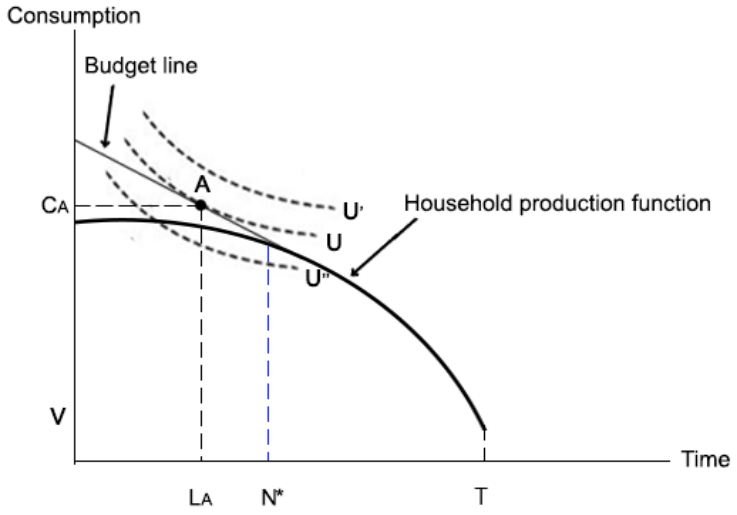


Figure 1. This utility maximizing individual, with a non-labor income V and a real wage rate w/P, will maximize her utility by adapting at point A. At this point, her indifference curve, with the highest possible utility, is tangent to her budget line. She chooses to have L_A hours of leisure, N - L_A working hours, T – N* hours in household production and a consumption of C_A.

Each utility function is expressed by indifference curves that are strictly convex to the origin, the shape of the curves are different for everyone. The reason for this is that there are differences in how people trade-off between consumption of goods C and leisure L. This will not only be dependent on their own unique preferences, but also on other factors such as how many young children they have or their preferences for childcare arrangements. A rather steep curve[21] indicates that the mother prefers consumption of leisure activity relatively more than

[21] See indifference curve U_B in Figure 2.

consumption of goods. They will therefore require a high wage to give up an hour of leisure. A rather flat curve[22] indicates that the mother prefers consumption of goods relatively more than leisure activity. They have a lower reservation wage compared to someone with steeper indifference curves.

3.4 Effects of the CFC reform in 2012

The removal of eligibility for 2-year olds caused a change in the mothers' budget line and reduced the relative price of publicly subsidized day care.

3.4.1 Effects on the budget line

There are three different types of childcare arrangements: (1) publicly subsidized day care, (2) informal caretakers, and (3) parental care. Every mother has a preference and we will initially assume that the CFC benefit did not affect her preference for either (1) or (2). This is something I will discuss later in the next section.

First, we will separately look at how the change in the CFC program affected the mothers who preferred (2) and (3), respectively. The cash benefit was a non-labor income for working mothers who prefer having their children taken care of by informal caretakers such as childminders or relatives. It is also a non-labor income for mothers who would choose to stay at home to take care of their children no matter if there was a cash benefit or not. Their reservation wage is higher than their offered or market wage, which is why they will not join the labor market (Heckman, 1974, p. 679). This is because they require a relatively high wage to give up an hour of their leisure time. However, their offered wage rate is so low that they will adapt at a point where their working hours will be zero[23]. Therefore, they ended up suffering an income loss after the withdrawal of the CFC benefit in 2012.

A negative non-labor income change will move the non-labor income V in Figure 1 downwards. This causes a vertical shift downwards of the budget line and there will be a negative income effect. Because of this shift, the worker will adapt at a different optimal point. According to the theory of consumer choice, a negative income effect decreases demand for normal goods and increases demand for inferior goods (Snyder & Nicholson, 2012, p. 136). Therefore, if leisure is a normal good, the mothers will want to consume less leisure after the reform.

[22] See indifference curve U_A in Figure 2.
[23] See scenario b) in Figure 2.

The effect of a non-labor income on working hours seems to vary based on previous studies; I will go into more details regarding this later on in section 4.2. However, from earlier studies about the introduction of the CFC program, we can see how an increase in the mothers' non-labor income led to a reduction in their labor supply (Drange & Rege, 2013; Naz, 2004; Schøne, 2004b). This indicates that leisure is a normal good for the mothers of 1- and 2-year-olds.

Figure 2: For mothers who use informal caretakers and regular stay-at-home mothers

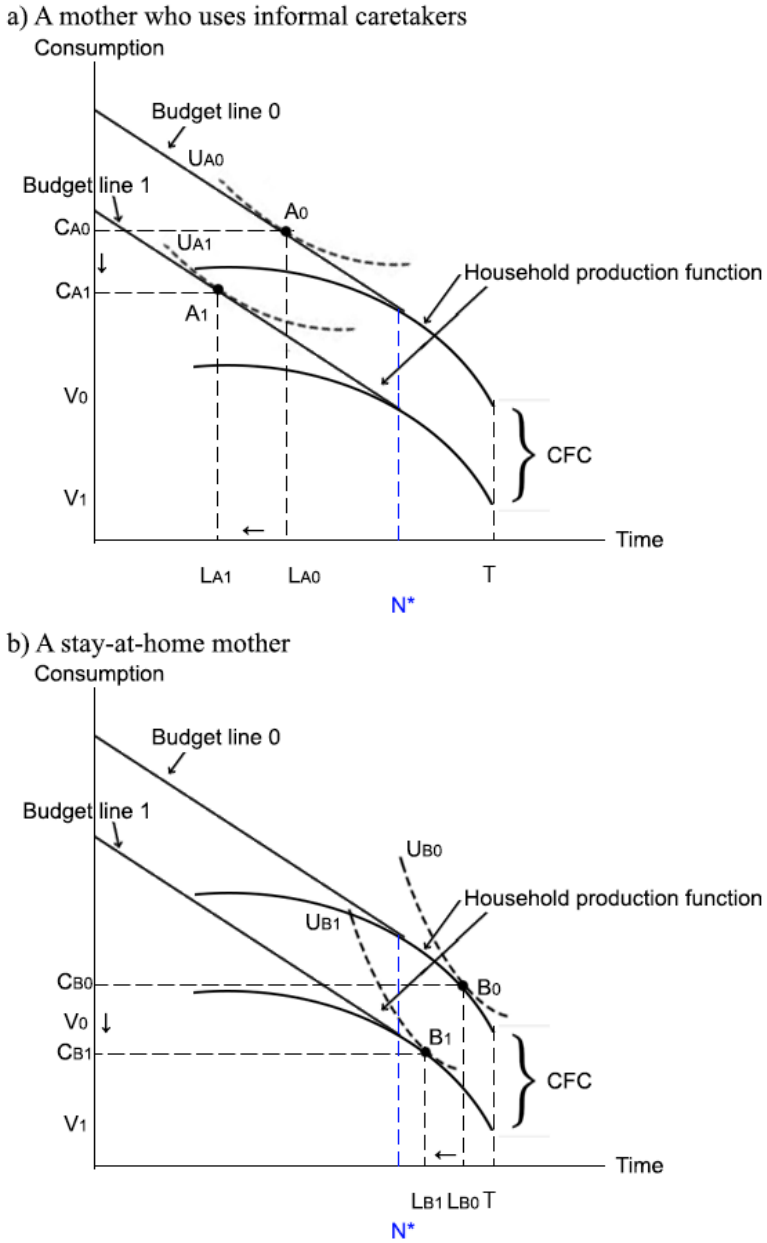


Figure 2. Scenario a) illustrates the optimization of a mother who uses informal child care, while scenario b) illustrates the optimization of a stay-at-home mother. A utility maximizing mother, who was working, will move from point A_0 to point A_1 after the reduction in non-labor income, from V_0 to V_1. Both her leisure hours and consumption of goods decrease, while her working hours increase and hours in household production remain the same. A utility maximizing mother, who was not working, will move from point B_0 to point B_1 after the reduction in non-labor income. Her consumption of goods and leisure time decrease and her hours in household production increase, while her working hours remain the same.

Figure 2 shows the removal of the cash benefit as a parallel shift of the budget line, from Budget line 0 to Budget line 1 for these types of mothers. A working mother who uses informal child care (A) will adapt at A_1, which leads to an increase in her working hours. Conversely, a stay-at-home mother (B), who in this case has a reservation wage higher than her offered wage[24], will still stay out of the labor force and her working hours remain the same. Both mothers reduce their consumption of goods and leisure hours. This is because of the negative income effect, which decreases their demand for normal goods.

As for the parents who use public day care, the mothers must choose between work and staying at home to take care of their child. Let us assume that an hour of labor equals an hour of attendance at the public daycare center. Budget line 0 in Figure 3 represents their budget line if they were still eligible for the CFC benefit. They will not be able to receive the cash benefit if they work full-time[25]. Those who work part-time will receive part of the benefit; the size will gradually decrease as their child's degree of public daycare attendance increases. Because of this, the "wage rate" for a part-time worker ended up being lower than for a full-time worker, although they were offered the same wage rate. Budget line 1 in Figure 3 represents the mothers' budget after the government removed the eligibility for the CFC benefit. The "wage rate" of the part-time working mothers will become equal to their offered wage rate; they will experience an increase in their "wage rate". Simultaneously, they will suffer a decline in their non-labor income. There will be a substitution effect, because an increase in the wage rate increases the opportunity cost of leisure, which will decrease the demand for leisure. On the other hand, the income effect will work in both directions. The "wage rate" increase leads to a positive effect, while the decline in non-labor income leads to a negative effect if leisure is a normal good.

Figure 3 illustrates examples where the substitution effect, the negative income effect, or both dominate the positive income effect. A part-time working mother (A) will increase her working hours and become a full-time worker after the policy change. Similarly, a mother (B), who used to stay at home, will enter the labor force after the removal of her eligibility. This is because her reservation wage has been met after the increase of her "wage rate". Both mothers will reduce their hours of leisure because the substitution effect and the negative income effect lead to a decline in their demand for leisure.

[24] The slope of the budget line is the negative value of the real offered wage rate.
[25] These mothers would adapt somewhere at the top left in figure 3, which is a part of both budget lines.

Figure 3: For mothers who use publicly subsidized day care

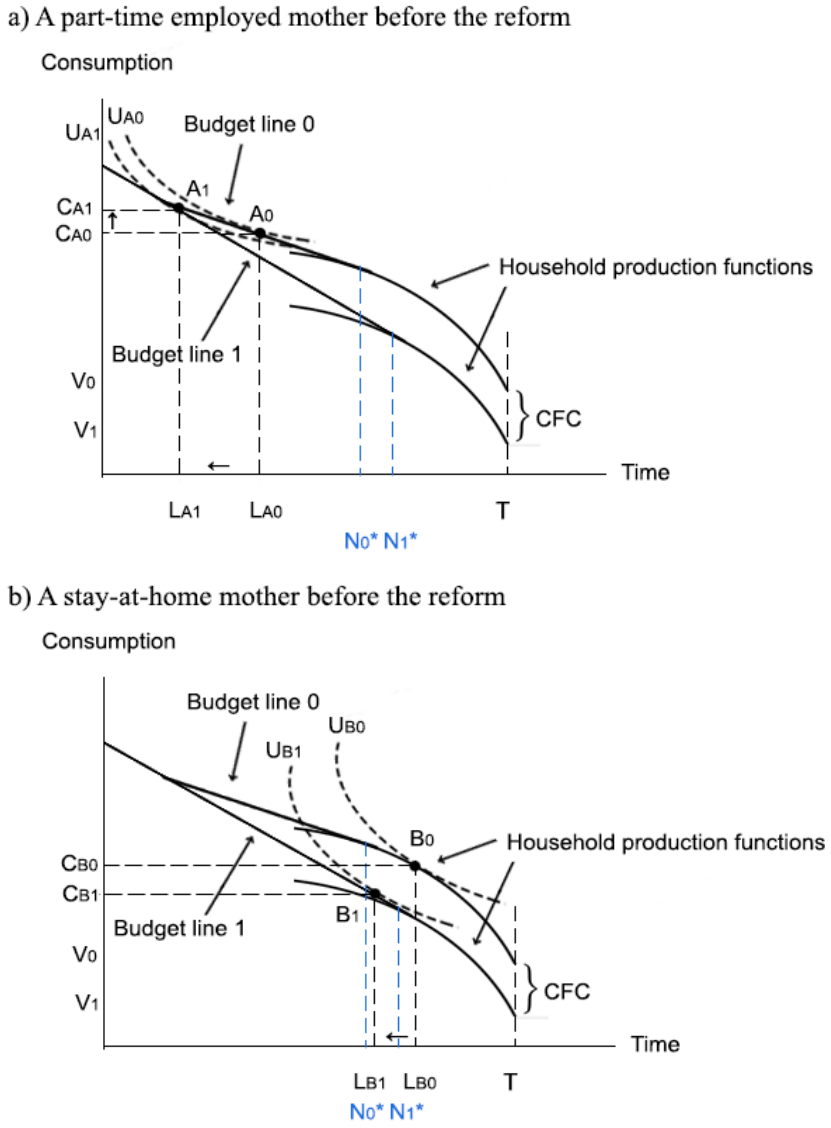


Figure 3[26]. Scenario a) illustrates the optimization of a mother who was a part-time worker before the reform, while scenario b) illustrates the optimization of a mother who was not working before the reform. A utility maximizing mother A, will move from point A_0 to point A_1 after the removal of eligibility. Her consumption of goods and working hours increase and she will become a full-time worker, while her hours in household production and leisure time decrease. A utility maximizing mother B, will move from point B_0 to B_1 and her consumption of goods and leisure hours decrease. Both her hours in household production and working hours increase, which means that she will enter the labor force.

3.4.2 Relative price of childcare arrangements

We have assumed so far in the previous section that the removal of the CFC benefit did not affect the parents' preference for childcare arrangements. However, the relative price of public day care did decrease after the parents no longer had the option of receiving the cash benefit. This is because those who were eligible had to forego the CFC benefit if they wanted to use publicly subsidized day care. Therefore, this benefit can be said to be an opportunity cost, which made the childcare costs in a publicly subsidized daycare center relatively more

[26] Figure inspired by Schøne (2004b).

expensive[27] than childcare services from other caretakers. This means that if these two childcare arrangements are seen as perfect substitutes, then the removal of eligibility will make more mothers go from informal child care to public child care.

Most of the previous studies have found results showing that childcare costs do affect maternal labor supply (Blau & Robins, 1988; Kimmel, 1998; Ribar, 1992). Mothers, who prefer parental care and think this is the best for their children, might increase their incentive to join the labor market due to the decrease in relative price of public day care.

Both the labor supply theory and the earlier studies about childcare costs predict that the CFC reform in 2012 will have positive effect on maternal labor supply. Therefore, we can derive the first hypothesis in this paper:

H1: ***The change in the CFC program affects the full-time employment of the mothers of 2-year-olds.***

3.5 Different types of mothers

The labor supply will vary between mothers depending on their demographic characteristics. This section will discuss two of the characteristics that may be relevant.

3.5.1 Educational level

There have been previous investigations about the relationship between a married woman's education and her labor supply. One of the studies implied that women that are more educated worked more frequently and had more working hours than less educated women (Heckman, 1974). His results also suggested that increased years of schooling led to an increase of both the offered wage and reservation wage. Likewise, Gronau (1973) found out that increased education had an effect on the mother's value of household productivity. His findings indicate that a mother with a college degree has a relatively higher shadow price of time[28] than someone less educated.

Therefore, a more educated mother compared to a less educated mother will have a steeper budget line[29] because of the higher offered wage. Furthermore, a more educated mother will have steeper indifference curves[30] because of the higher reservation wage. The reason for this

[27] I assume that the cost of publicly subsidized day care was higher than the cost for other caretakers minus the CFC benefit before the 2012 reform.
[28] Her reservation wage.
[29] Budget line B in figure 4.
[30] The indifference curves U_B in figure 4.

is that the mother's offered wage rate determines the slope of the budget line, whereas a higher reservation wage indicates that the mother demands a higher wage for giving up an hour of leisure. This means that the mother prefers leisure activity relatively more than consumption of goods, as previously discussed in section 3.3.

Figure 4: Different budget lines and indifference curves depending on educational level

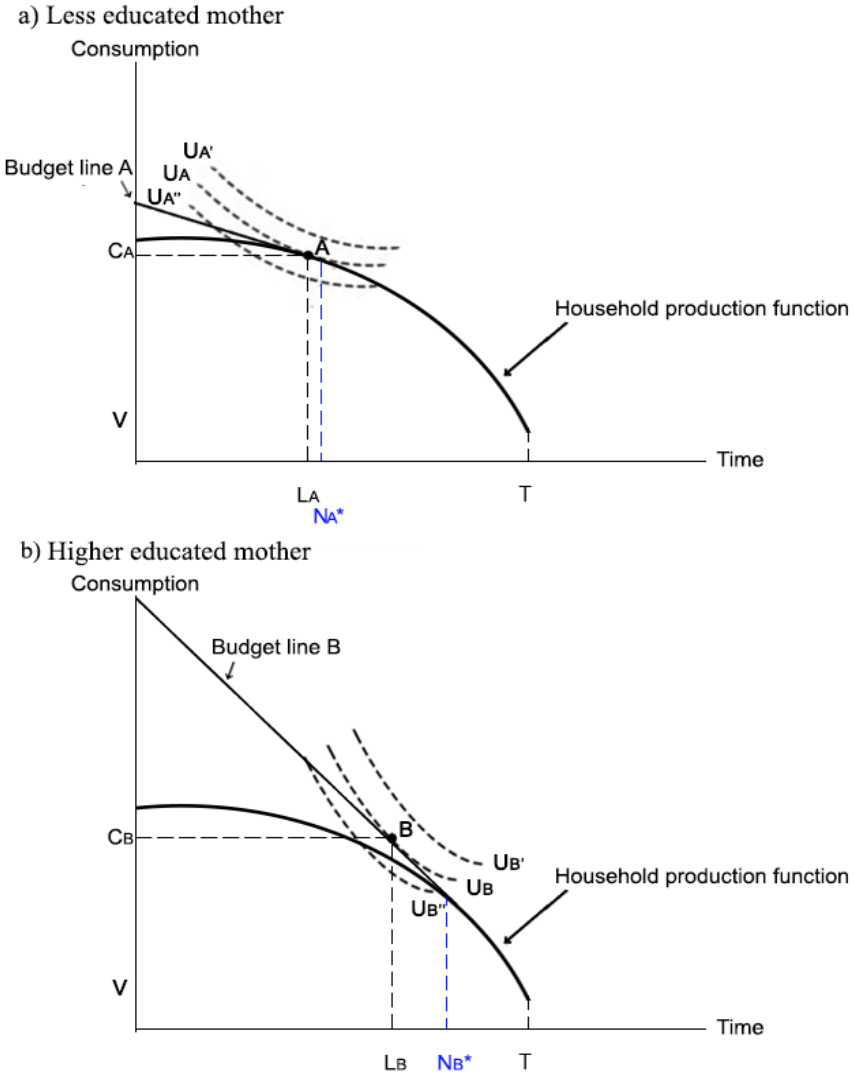


Figure 4. Scenario a) illustrates the optimization of the less educated mother (A), while scenario b) illustrates the optimization of the higher educated mother (B). Mother A will adapt at point A, where she chooses to have L_A hours of leisure, $T - N_A^*$ hours in household production, $N_A^* - L_A$ working hours and a consumption of C_A. Mother B adapts at point B, where she chooses to have L_B hours of leisure, $T - N_B^*$ hours in household production, $N_B - L_B$ working hours and a consumption of C_B. In this particular case, the higher educated mother has more hours of leisure, works more in the market and less at home, than what the less educated mother does.

Figure 4 illustrates an example of two mothers with the same non-labor income, but with different educational levels. They will adapt at different points because of the different slopes[31] of their budget lines and the difference in the shape of their indifference curves. Consequently, the effect of the change in the CFC program can be stronger for mothers with

[31] Their wage rate.

more schooling or it can have a stronger effect on mothers with less schooling. Figure 4 presents an example where the higher educated mother works at the market more than what the less educated mother does.

Based on the relationship between education and the labor supply of mothers, I have my second hypothesis:

H2: ***The change in the CFC program affected mothers of 2-year-olds differently depending on their educational level.***

3.5.2 Marital status

Although single mothers do not have their husband's or partner's income to rely on, they are still eligible for some public transfers, which are specific to single parents. As mentioned in section 2.3, a single mother of a 2-year-old can receive a transitional benefit, but this will start to decrease as they earn more than one fourth of their yearly benefit (NAV, 2016). This is similar to how the CFC benefit decreases with increased hours of public daycare attendance. Some municipalities encourage single parents to work by giving them priority in access to publicly subsidized daycare centers. Additionally, they are eligible for a childcare benefit, which reduces the cost of public day care. This makes it more attractive for single parents to use this type of child care. However, the childcare costs will be a bigger fraction of the income for a single mother (Anderson & Levine, 1999, p. 11). The reason for this is that the income of a family headed by a single parent will be on average lower than for a family with two parents.

From a previous study based on Statistics Norway's survey of living conditions from 2000 and 2001, the results showed that single mothers earned on average less than other mothers (Ugreninov, 2003). While the employment rate of single mothers is lower for all women between 25 and 44, their fraction of full-time workers is higher (Kjeldstad & Rønsen, 2002, p. 31). There also seems to be differences in labor supply between married and cohabitant mothers.

Figure 5 shows an example of a single mother and a married or cohabitant mother with the same offered wage rate. If the single mother receives a transitional benefit, as depicted by scenario a), both mothers will in this case have the same initial non-labor income. However, since the transitional benefit decreases with the single mother's income after she has reached a certain threshold income, their budget lines will not be similar. The single mother's budget

line A will bend and have a lower slope compared to the other mother. We can see that the single mother ends up working N* - L_A hours, which is less than how much the married or cohabitant mother works. Conversely, if the single mother does not receive any transitional benefit, as represented by scenario b), she ends up working more than what the other mother does instead. This time, both budget lines will be similar, but the non-labor income will be lower for the single mother than for the married or cohabitant mother.

Figure 5: A single mother and a married or cohabitant mother

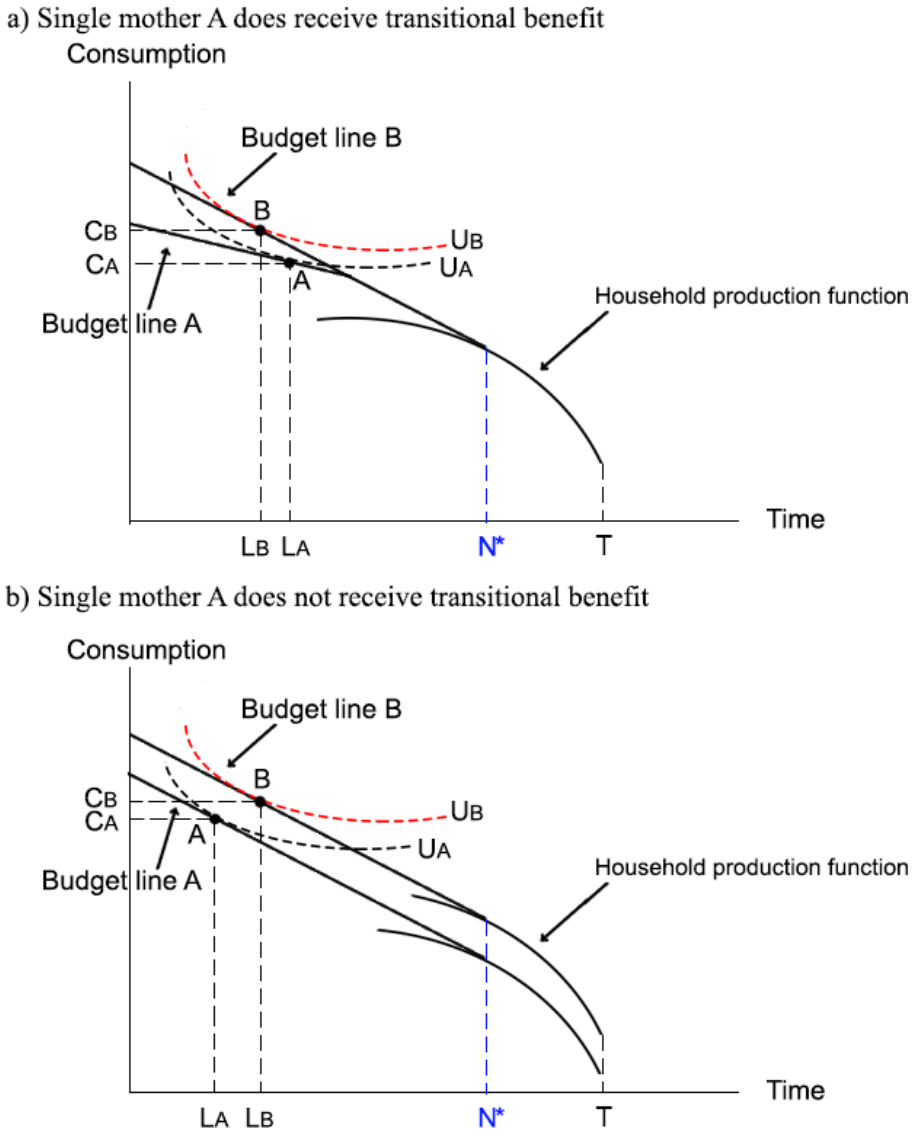


Figure 5. Scenario a) presents a single mother (A), who receives transitional benefit and a married or cohabitant mother (B) with the same offered wage rate and non-labor income. They will have different budget lines because the transitional benefit decreases with increased earnings after the single mother's earnings have surpassed 0.5 G[32]. Their optimal points will not be the same although they have the same preferences between leisure and consumption of goods because of the different budget lines. Mother A chooses to have N* - L_A working hours, which is less than that of mother B, who chooses to have N* - L_B working hours. Scenario b) presents the same mothers, but this time, the single mother (A) does not receive transitional benefit, therefore she will have a lower non-labor income than what the married or cohabitant mother has. They will adapt at different points because of this and mother A ends up working more than what mother B does.

[32] National insurance scheme basic amount.

Based on the discussion in this section, I have my third and last hypothesis:

H3: ***The change in the CFC program affected mothers of 2-year-olds differently depending on their marital status.***

4 Previous literature

In this section, I will discuss previous empirical literature regarding the CFC program and the related mechanisms from the labor supply theory mentioned in the previous section.

4.1 The effects of the CFC program

Ever since the CFC reform in 1998, many have analyzed its effect on the maternal labor supply. There have also been studies on other subjects possibly affected by the reform such as children's educational outcomes (Bettinger, Hægeland, & Rege, 2014) and choice of child care (Hellevik, 2000).

4.1.1 The short-run effect of the CFC program on labor supply

Håkonson et al. (2001), who used data from the Norwegian Labor Force Survey, estimated the short-run effect on the eligible mothers' labor supply to be a reduction of about 3,700 man-years, which was equal to a 7.5 % decrease (p. 35). Additionally, they did not see any effect on part-time work. This could be because there were mothers who decided to work part-time, but also part-time working mothers who decided to leave the labor force.

Naz (2004) used a difference-in-differences method, to analyze the effects of the CFC program on specialization and the parents' labor supply. She compared the change in labor supply from a pre-reform period to a post-reform period for two different groups. One of the groups were parents of children in CFC eligible age, while the other group consisted of parents with older children, who were too old to be eligible for the cash benefit. The children in the treatment group were aged 1-3 years, while the control group consisted of children aged 4-6 years. The results showed a 2.85 decrease in the wife's working hours, while the husband's working hours barely changed (Naz, 2004, p. 377). Since Naz (2004) used data from living standard surveys from spring 1998 and 1999, her sample consisted of partly treated children[33]. This was not the case for Drange (2012), who was able to estimate the effect on parents of fully treated children. This is because she had a bigger dataset with many more observations and observation years. She utilized a similar difference-in-differences method and found similar results. Drange measured the difference in labor supply from a pre-reform period to a post-reform period for parents of 2-year-olds, and compared it with the difference for parents with 5-year-olds. The results did not show any effect on the fathers' labor participation, but did suggest a reduction in the mothers' labor force participation. She

[33] The CFC program did not include 2-year-olds until January 1999, which means that they were partly treated. See section 2.1.

found a 4-percentage point decrease in full-time employment, which implied that 13 % of the full-time employed mothers had reduced their labor supply because the mean full-time employment rate was 31.5 %. Additionally, she found a 2.4 percentage point decrease in employment for mothers. These findings would indicate that many mothers did go from working full-time to working part-time. (Drange, 2012, p. 22).

Schøne (2004b, p. 719) did a triple difference estimation, where he followed a group of treated mothers from before childbirth until 2 years after birth. He did the same for the mothers in the control group, who gave birth to their child 3 years earlier than the treated mothers did. By observing these two groups in different years, he managed to compare mothers with children of the same age; this was the first difference-in-differences. He measured an additional difference-in-differences between mothers with older children in the same observational years. None of these older children had the opportunity to be eligible for the CFC benefit, and by comparing these two difference-in-differences, he estimated a triple difference estimator. The results showed a 4 % decline in labor participation, while the annual hours were reduced by 3 %. This could imply that the mothers responded more at the extensive margin than at the intensive margin. Through a multivariate analysis based on two surveys conducted before and after the reform, Rønsen (2000) found less use of subsidized care and an increase in the mothers' probability to not work. The probability reduction was especially big for teachers, while most nurses seemed to have gone from working full-time to part-time.

In earlier research about the CFC reform in relation to education, Naz (2004) reported that the reform led to a reduction in working hours for both highly educated mothers and mothers without a university degree. The difference was not statistically significant although the mothers with more schooling had a bigger reduction than the less educated. Rønsen (2000) on the other hand, found a small decrease in work probability for all mothers expect for those with the highest university degree.

4.1.2 The long-run effect of the CFC program on labor supply

The studies that analyzed the long-run effect of the CFC program had mixed results. Drange and Rege (2013) wanted to investigate if mothers leaving the labor force temporarily would affect their future careers. They still found a significant effect when the child had turned 4 years old; a year after they were no longer able to receive the cash benefit. The full-time employment rate of the mothers of 4-year-olds increased with 2 percentage points. However,

there was no longer any significant effect for mothers of 5-year-olds and up. In addition, this long-run effect seems to have affected mostly mothers without a university degree or mothers who had earnings below median before the program. In another study, Schøne (2004a) also found the negative effect to be dissipating over the years. However, unlike Drange and Rege, he did not find any significant long-run effect; there was only a 1-percentage point decrease in the labor participation of mothers of 3-year-olds.

Rønsen (2009) wanted to investigate the CFC reform's effect on maternal supply in a longer time perspective. She wanted to see whether the effect became stronger or weaker several years after the introduction of the CFC program compared to when the government first introduced it. She found the decline in weekly labor supply from spring 1998 to 2002 to be 3.75 hours, compared to 2.4 hours from spring 1998 to 1999. This seemed to imply that the long-term effects are stronger than the short-term. One of the weaknesses with this study is that there might have been other reforms in the period from 1999 to 2002, which affected the female labor supply. This weakness could have biased the estimates and made it seem like the long-term effects were stronger than what they actually were.

4.1.3 The effect of the CFC reform in 2012

There are not many previous studies about the removal of eligibility for the CFC benefit, but Dahl (2014) did find a short-run positive effect on maternal labor supply. He analyzed the effect the 2012 reform had on mothers of 2-year-olds by using a logistic regression model. His findings suggested that the mothers' probability to work increased with 3 percentage points after the reform. However, he ran separate regressions for the treatment group and the control group; the estimated 3-percentage point increase is therefore not only caused by the policy change. This is because he also needs to compare it to the change in probability to work for the control group, which was estimated to be a 1-percentage point increase. Dahl only analyzed parents who were working before the birth of their child, which means that he did not take into account the effect the 2012 reform had on parents who were not working before the childbirth. Furthermore, the control group consisted of people who did not have children in CFC eligible age, but who were the same age as the parents in his treatment group. There is a possibility that someone in this control group did not have any children, and people with children might not be comparable with people who have no children.

4.2 Non-labor income

As mentioned in section 3.4.1, the CFC benefit was an increase in non-labor income for certain mothers, and with the removal of eligibility, these mothers suffered a non-labor income loss. In previous research regarding the effect of an increase of a non-labor income on labor supply, Holtz-Eakin, Joulfaian and Rosen (1993) investigated the relationship between inheritance and labor supply. The labor force participation of both single individuals and married people fells as the size of the inheritance increased. Imbens, Rubin and Sacerdote (2001) analyzed the impact of lottery prizes on labor supply using data on lottery winners in Massachusetts from the 1980s. They found a significant reduction in labor supply from those who won a large amount of money, but they had no such result for winners of small prizes. Another paper looked at the exogenous increase in salaries of the members of the European Parliament, which happened in 2009 (Mocan & Altindag, 2013). Their salaries were not dependent on attendance to the Parliament, so the increase could be seen as an increase in non-labor income. The results showed a statistically significant negative relationship between an increase in salaries and labor supply.

4.3 Childcare costs

As mentioned previously, the removal of the CFC benefit led to a decrease in the relative price of public day care, which made public subsidized daycare centers more attractive to the parents. Empirical evidence has shown that there is a negative correlation between childcare costs and maternal labor supply. Blau and Robins (1988) found out that higher childcare costs raised the probability of the mother not working. This was the case even if an informal caretaker was taking care of her child. They estimated the childcare price elasticity of labor supply to be equal to -0.38 (p. 379). An even stronger effect was found by Ribar (1992, p. 156), where the hourly childcare price elasticity with respect to employment was estimated to be -0.74 for married women.

Another study also found that the mothers' decision to work was sensitive to childcare costs (Kimmel, 1998). This applied to both married and single mothers, but the latter was less sensitive compared to the former. Powell (2002, p. 123) discovered that decreasing the costs of one type of child care had a significant positive effect on the probability of the mother working and using that specific type of child care. Having a 10 % price subsidy for formal child care led to a 12 % increase in employment rate.

5 Empirical method

5.1 A causal relationship

In this paper, I want to find the causal effects of the policy change of the CFC program on maternal supply. Although two variables are correlated with each other, it does not necessarily mean that they have a causal relationship, where A causes changes in B. An increase in the mothers' labor supply, after a reform, does not necessarily mean that the reform was the reason for the increase. Other factors may have played a role in the resulting outcome. Thus, we need to separate the effects that would have happened anyway without the reform, from the effect caused by the CFC reform in 2012.

In the potential outcomes framework concept, developed by Rubin (1974), an individual has two possible outcomes, which will depend on whether he or she has been treated or not. This can be described[34] more precisely as:

$$Y_{1i}\ if\ D_i = 1$$
$$Y_{0i}\ if\ D_i = 0$$

Where Y_{1i} and Y_{0i} are the two potential outcomes, while the binary variable D_i is the treatment variable. The counterfactual case is a "what if" situation, where we could actually see what the outcome would be if a treated individual was actually not treated. This is not possible because an individual cannot have both been exposed and not been exposed to the very same treatment. Therefore, we can observe only one of the potential outcomes on a specific mother and this is called the fundamental problem of causal inference (Holland, 1986, p. 947). It is impossible to observe the causal effects of a treatment directly. A randomized experiment is the most reliable type to estimate the treatment effects. The reason is that individuals will be randomly selected for treatment, which means that there will be no selection bias. The mothers who no longer are eligible for the CFC benefit can be said to be a random selection. Mothers who work were just as likely to be treated, as mothers who stay at home. There will be an unbiased distribution of variables, which makes the groups comparable.

5.2 Quasi experiment

Since the change in the CFC program has already happened, it will not be possible for us to assign the treatment to people randomly. Every mother of a 2-year-old got the treatment

[34] Angrist and Pischke (2008, p. 13)

because it would have been unethical and controversial if only some of these mothers became non-eligible. What we have is a quasi-natural experiment with an eligibility cutoff, where the treated mothers are those who no longer were eligible for the cash benefit after their child turned 2 years old.

5.3 Difference-in-differences

One of the quasi experiment designs is the difference-in-differences method, where you can compare changes in outcomes between groups before and after a policy change. In this case, it would be mothers of 2-year-olds before and after the removal of eligibility. It is a normal approach used to measure the causal effects of a policy change. In a natural experiment, the control and treatment group have to be comparable and have similar properties[35] pre-treatment. Since the policy change was nationwide, it is not possible to have a control group consisting of mothers of 2-year olds who were eligible in the post-period. Instead, I will have mothers of older children, who were not affected by the CFC reform in 2012. These children will be 5 years old or older because they are old enough to have been entitled to the cash benefit for 2 years. Consequently, the 2012 reform did not have any effect on these older children.

Table 3 from section 2.1 describes the nature of the treatment, the 2009 and 2010 cohorts were partly treated, which means that some had the full or part of the treatment, while some did not get any treatment. Since the dataset provides only the birth year of the child, it is not possible to know who in these cohorts were fully treated. This is also why I cannot have 2012 as the post-reform year of observation. The 2011 cohort was the first cohort to be fully treated, so they will be a part of the post-reform treatment group. I will need observations from when they are currently receiving the treatment and this would be when the child is 2 years old. Therefore, I will use the observations from 2013 and 2014, the years the 2011 cohort turns 2 and 3 years old, respectively. The cohort will be either 1 or 2 years old in 2013 and 2 or 3 years old in 2014. The pre-reform treatment group will consist of children who did not receive any treatment at all; this would be the 2008 cohort, which was the last cohort with no treatment. They will be observed in 2010 and 2011, the years they reach the ages of 2 and 3 years, respectively.

[35] Blundell and Dias (2009)

I will compare labor supply variables of the post-treatment[36] mothers in 2013 or 2014 to the pre-treatment[37] mothers in 2010 or 2011. To make up for the time difference, I will compare the difference between the mothers of 2-year olds with the difference between the mothers of older children, who had no treatment at all. As mentioned earlier, the children in the control group must have been old enough to have not been affected by the 2012 reform, which means that they have to be born in 2008 or earlier. To increase sample size and thereby statistical power, I will include several cohorts in the pre-and post-control group. This is different from the pre- and post-treatment groups, which have one cohort in each. The pre-reform control group will be composed of the 2001-2004 cohorts, while the post-reform control group will be composed of the 2004-2007 cohorts. Thus, the children in the control group will turn 6-9 years old in 2010 or 2013.

The outcome variables I will use to analyze the mothers' labor decisions are weekly working hours and full-time employment.

5.3.1 Difference-in-differences estimator

I will estimate the following difference-in-differences estimator:

$$\delta_a = (Y_{a,y+3} - Y_{a,y}) - (Y_{a+4|a+5|a+6|a+7,y+3} - Y_{a+4|a+5|a+6|a+7,y})$$

Where Y is the outcome variable, the subscript a is the age of the treated child at the end of the year, while y denotes the first year of observation. A specific example is the difference-in-differences estimator for mothers of the children who turned age 2 in the year they were observed:

$$\delta_2 = (Y_{2,2013} - Y_{2,2010}) - (Y_{6|7|8|9,2013} - Y_{6|7|8|9,2010})$$

The first difference is the difference in the outcome variables between mothers of 2-year olds, before and after the reform. The mothers in 2010 were still entitled to the CFC benefit, while the mothers in 2013 were no longer eligible. The second difference is the difference in the outcome variables between mothers of 6- to 9-year-olds; this is the control group, which was not treated. So by subtracting this difference, we can control for other factors that have affected the labor market during these years.

[36] Those who were entitled to the cash benefit for only 11 months, up until their child reached age 2.
[37] Those who were entitled to the cash benefit for 23 months, up until their child reached age 3.

5.3.2 Identifying assumption

By using the difference-in-differences method, we can measure the average effect of the treatment on the treatment group. This relies on an important identifying assumption; the assumption of common trend, which posits that both groups would follow parallel time trends without the presence of the treatment. That no other exogenous variables affecting the outcome than the policy change should differ in the two groups[38]. Additionally, we have to take into account the pre-existing differences between the control and treatment group.

If the assumption of common trend holds, then the change in the average full-time employment rate among mothers of 2-year-olds and mothers of 6- to 9-year-olds from 2010 to 2013 would be the same if none of them got any treatment. Figure 6 illustrates this assumption; the trend of the control group is parallel with the striped line, which represents the assumed trend of the treatment group in the counterfactual case[39]. It is assumed that it is the treatment that causes the trends to become non-parallel. However, there might be observed characteristics, which can influence the mother's outcome variables differently. Therefore, any estimated differences in the outcome variables may have been caused by the compositional differences between the groups of mothers, instead of by the reform. This is why these characteristics will be included as variables in my regression analyses.

Figure 6: Difference-in-differences estimation

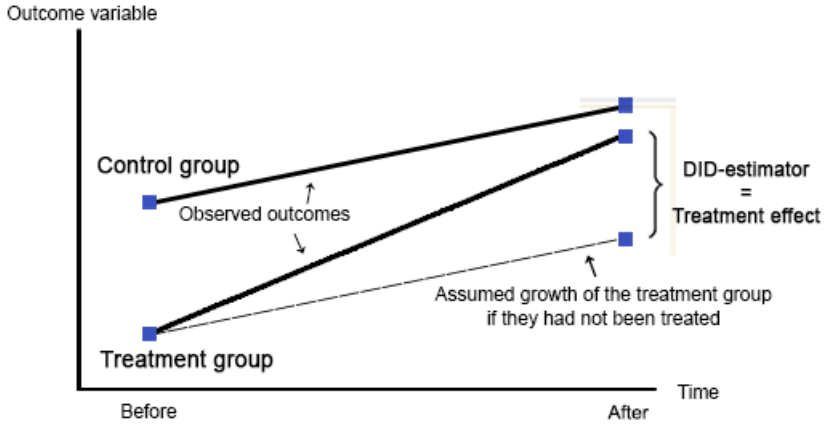


Figure 6. The DD estimator is the difference between the changes in the outcomes of the control and treatment group. The second black bold line is the actual trend of the treatment group, while the striped line is the assumed trend in the counterfactual case.

5.3.3 Possible threats to the common trend assumption

There is a chance that the growth in the outcome variables differs between the treatment and control groups. This is why we will have to compare the past trends prior to the policy

[38] Rubin (1974, p. 698).
[39] A hypothetical case where the policy change in the CFC program had not happened.

change, to see if they were parallel or not. One possibility is that the business cycles affect mothers of younger children differently than how they affect the mothers of older children. After the financial crisis in 2008, the unemployment rate rose in Norway and the growth lasted until the beginning of 2010[40], where it started decreasing. The financial crisis may have had a long lasting effect on certain mothers. This would be a problem if it affected one of the groups of mothers, for example the mothers of 2-year-olds, more than the other group.

Another possible concern is that the control group consists of school-age children, while the treatment group consists of younger children. There may be changes in school that will affect only the mothers of the older children and not the mothers of younger children. While there is a maximum limit for the cost of having children in daycare centers, there is no such thing for children who attend the after school program, Skolefritidsordning (SFO).

These are possible threats, which could lead to trends that would not be parallel between 2010 and 2013, even if the CFC reform in 2012 had not happened.

5.3.4 The linear regression model

I will estimate the δ_a estimators with the following linear regression model:

$$Y_{a,i} = \beta_0 + \beta_1 Age_{a,i} + \beta_2 Post_{2011+a,i} + \beta_3 X_i + \delta_a (Age_{a,i} \times Post_{2011+a,i}) + \varepsilon_i$$

$Y_{a,i}$ is the outcome variable of individual *i* with a child aged *a* in time *t*. This is a multiple linear regression model if the outcome variable is weekly working hours and a linear probability model if the outcome variable is full-time employment[41]. $Age_{a,i}$ is a dummy variable indicating whether the mother's child is one of the younger children[42] or not. $Post_i$ is a dummy variable indicating whether the mother is observed in the year 2011 + *a* or not, which is a period after the policy change. It takes the value of 1 if the mother was observed in 2013 or 2014. The interaction term between these two variables informs us that the mother is treated by having a younger child in a post-reform period if its value is 1. The accompanying coefficient δ_a, is the estimator we are interested in, it measures the effect of the treatment on the outcome variable, holding other factors constant. X_i is a vector with control variables that may affect the mother's labor supply and are more closely explained in section 6.3. The error term ε_i represents all other unobserved explanatory variables that affect the outcome variable.

[40] See Figure A1 of the appendix.
[41] A binary dependent variable.
[42] The children who reached age 2 in the first observation year and age 3 in the next observation year.

5.3.5 Alternative nonlinear regression model

A linear probability model is a linear regression model with a binary variable, such as full-time employment, as the dependent variable. One of the weaknesses with a model like this is that the estimated probability can be higher than one or lower than zero (Aldrich & Nelson, 1984). Therefore, I will run additional logistic regressions to account for this weakness, and see if the mean marginal effects based on these estimates are similar to the estimates from the linear regression model.

I will run the following standard logistic regression:

$$p_i = P(\gamma_i = 1 \mid Age_{a,i}, Post_{2011+a,i}, X_i)$$

$$p_i = \frac{1}{1 + e^{-z}}$$

Where p_i is the probability of being full-time employed and

$$z = log\ \frac{p_i}{1 - p_i} = \beta_0 + \beta_1 Age_{a,i} + \beta_2 Post_{2011+a,i} + \beta_3 X_i + \delta_a (Age_{a,i} \times Post_{2011+a,i})$$

6 Data

I will use the data from the Norwegian Labor Force Survey (AKU), provided by NSD. These surveys, conducted by Statistics Norway (SSB), produce quarterly anonymous data. There are about 20,000 respondents, between the ages of 16 and 74, for each quarterly survey. A number of households are randomly chosen among all the households in every municipality in Norway. The household members in the relevant age group will take part in the surveys. Each respondent is obligated to take this quarterly survey 8 times in a row, which means that the respondents will be observed for 2 years in a row. The data provides mostly labor and employment related information such as employment status, whether they are a full-time or part-time worker and the number of working hours in a specific reference week. It also includes individual demographic information gathered from Statistics Norway and the personal register. The demographic variables consist of age, gender, marital status, educational level, birth year of their youngest child and the number of children under 16.

6.1 The sample

As outlined in section 5.3, the main analytic sample will consist of children turning 2 years old or 6-9 years old in either 2010 or 2013 and children turning 3 years old or 7-10 years old in either 2011 or 2014.

The sample will be restricted to children with no younger siblings to avoid the indirect treatment someone in the control group may have gotten through a younger sibling. As we increase the age of the youngest child, the number of observations will decrease. A group of mothers with a 1- or 2-year-old as their youngest child will be bigger in size than a group of mothers with a group of mothers with a 5- or 6-year-old as their youngest child. The reason for this is that this child will be the last child for many mothers in both groups, but the former group is more likely to consist of mothers who just had their first child and who will have more[43] later on.

To sum up, I first sort out women whose youngest child reaches the ages of 2 or 6-9 years in either 2010 or 2013. Secondly, I sort out women with a youngest child reaching the ages of 3 or 7-10 years in either 2011 or 2014. Table 4 shows how the sample is divided into eight groups of mothers spread in four observation years. The sample has four treatment subgroups

[43] The fertility rate in Norway is above one, which means that women in Norway have more than one child on average (Statistics Norway, 2016a).

and four control subgroups, which consist of mothers of nine cohorts of children born in 2001-2008 or 2011.

Table 4: The treatment and control groups

	2010	**2013**	**2011**	**2014**
Treatment	Pre	Post	Pre	Post
Control	Pre	Post	Pre	Post

Notes: The sample consists of eight subgroups, which are observed in four different years. There are four subgroups in each analysis (2010 vs 2013 and 2011 vs 2014).

I restrict my sample to include observations from 2010-2011 and 2013-2014. Thus, I have data from about 2 years prior to the policy change and about 2 years after the policy change. I will be able to see if the removal of eligibility had an effect in the short run and in the longer run. Additionally, having observations prior to the reform allows me to explore the validity of the common trend assumption.

Furthermore, I have excluded mothers who had an unknown educational level. Likewise, mothers with missing values in control variables such as marital status and municipality type have been excluded. In the end, I excluded 101 observations due to missing values, which reduced the sample size by about 2 %. There is a possibility that the respondents with missing values reacted differently to the change in the CFC program than how the rest of the sample did. However, since they constitute such a low percentage of the whole sample, there might not be such a sample selection bias.

Since this is a pseudo panel data, the observations from quarter to quarter will not necessarily be from the same individuals. Some individuals will have one observation per year, while others may have up to four observations per year. The estimated robust standard errors may be biased and misleading if we include several observations per individual per year. This is because the variability between observations of the same individual is usually smaller than the variability between observations from different individuals. The background and independent variables of an individual will not likely change much from quarter to quarter in a specific year. Their education level will most likely stay the same, and so will the variable that indicates whether she has been treated or not. Since this is the case, the observed outcome variables in each quarter can be correlated with each other. This is called the serial correlation problem, where the variables of an individual are correlated over time (Angrist & Pischke, 2008, p. 294). Because of this, I need to either use the mean of the observations per individual or have one observation per individual randomly picked for the analysis.

Since the dataset does not include an ID variable to identify each individual, it is not possible to calculate the means of the observations per individual. Instead, I have decided to pick the observations based on the variable, which informs us the number of times the respondent has answered one of these quarterly surveys. Table 5 shows my method for selecting the observations, which will ensure that there will be only one observation per individual, per year.

Table 5: Number of times an individual has answered one of the quarterly surveys

Quarter	**1**	**2**	**3**	**4**
Number of times the respondent has done the survey	1, 2, 8	1, 4, 8	1, 6, 8	1, 8

Notes: The table shows which observations I have included in my sample from each specific quarter. This is to make sure that there is only one observation per individual per year. The number of times the respondent has done the survey includes the current survey that they are answering.

In the end, the sample consists of 4842 observations, where the age range of the mothers is between 19 and 56 years old.

6.2 The outcome variables

The main outcome variable is the full-time employment variable, which is a dummy variable. The mother is given the value of 0 if she is a part-time worker or not working, and a value of 1 if she is a full-time worker. The survey has classified those who answered 37 working hours or more per week as a full-time worker. An additional question, about whether they work full-time or part-time, were asked to respondents who answered between 32 and 36 working hours.

The variable measuring the weekly working hours is a discrete numerical variable. The number of working hours includes both the main work and the secondary work, if the respondent has one. The measure of weekly working hours is the number of hours that the mother and her employer have agreed on in the work agreement. This is more stable than another measure[44] provided in the dataset, since labor demand can vary from season to season. The number of hours in the work agreement will not vary depending on which quarter the mother was surveyed in.

6.3 The control variables

As discussed in section 5.3.2, I assume the control and treatment groups have a common trend in the counterfactual case. There may be compositional differences between the groups of

[44] The respondent answer how many hours she has actually worked in a specific reference week.

mothers, which can lead to biased estimates. To account for this, I will include the following variables in the regressions.

The educational variables consist of four educational levels[45], which is represented by dummy variables: mandatory school, high school, short[46] higher education, and long[47] higher education. These variables state the maximum educational level a mother has. For example, if a mother has some high school education or is a high school graduate, she will be placed in the second educational level. Only the last three dummy variables are included in the regressions due to collinearity. The marital status variable is a dummy variable, where the value of 1 indicates that the mother is single; she could be a divorcée, a widower, or someone who has never been married. The mother is given the value of 0 if she is married or cohabitating with someone. Ideally, these control variables would all have been measured before the 2012 reform, to ensure that they are not endogenous to the change in the CFC program[48]. However, this is not possible with the dataset I have, hence I will assume that the reform did not affect the mothers' educational level or marital status.

The mother's age variable describes the mother's age at the end of the year. Older mothers will more likely have more work experience and a higher attachment to the labor market. The number of children under 16 is a discrete variable, and has values 1-8. Lastly, the municipality variable is a dummy variable, which tells us what type of municipality the mother lives in. Value of 1 indicates that it is central, while value of 0 indicates that it is less central. A central municipality has a higher population than a less central municipality.

6.4 Limitations

The sample size is very small and therefore, it might not ensure a representative distribution of the population, which consists of mothers of 2-year-olds. This could be a threat to external validity[49] if the sample is not representative of the population you intend to draw conclusions on. A small sample size may also lead to high standard errors, which again leads to non-statistically significant estimates.

[45] The classification of the different educational levels is similar to Statistic Norway's classification (Statistics Norway, 2015)
[46] Up to four years.
[47] More than four years.
[48] There is a possibility that the mother's educational level or marital status was affected by the change in the CFC program.
[49] The degree a study's results can be generalized to the population (Steckler & McLeroy, 2008, p. 9).

I cannot be certain of a child's specific age at the time of observation. The reason for this is that the mothers are observed in all quarters of the year, but the dataset provides only the birth year of the child. Hence, the observations of the treatment group in the first post-reform period did not only strictly consist of 2-year-olds, the currently treated children. Because of this, it is not possible to estimate the isolated effect on the mothers of 2-year-olds. Instead, I will estimate the separate effect on the mothers of children who reach the ages of 2 or 3 years in the year of observation, respectively.

Additionally, the estimates may have been more precise if more control variables were included in the regressions, but the dataset only provides a limited number of background variables.

6.5 Summary statistics

Table 6 provides the summary statistics of the outcome and control variables by group and period. Panel A reports the means of the outcome variables, while Panel B reports the means of the control variables.

The means in Panel A are calculated using data from each pair of observation years separately, either from 2010 and 2013 or from 2011 and 2014. This is to see how these variables changes from 2010 to 2013 or from 2011 to 2014, respectively. First pair of observation years shows the difference in labor supply for mothers of children who are either 2 or 6-9 years old at the end of the year. Likewise, the second pair of observations shows the difference for mothers of children who are either 3 or 7-10 years old at the end of the year. For example, the first row only uses data from 2010 and 2013 and shows the difference between the pre- and post-period for the treatment group and control group. The treatment group in this case consists of the cohorts that reached age 2 in these two specific years, while the control group consists of the cohorts who reached the ages of 6-9 years.

Labor supply theory predicts that the outcome variables increase after the reform for the treatment group. Looking at Panel A, you do see an increase for the treatment group from the pre-reform period to the post-reform period. There was also mostly an increase for the control group, except for the working hours of the mothers of children aged 7-10 years old at the end of the year. The changes seems to have been higher within the treatment group compared to within the control group. Furthermore, the difference between the two differences in each row

is the unadjusted[50] difference-in-differences estimate, which is positive in all four rows. The estimates were especially high for the mothers of children who reached the ages of 3 or 7-10 years during the year they were observed.

For the control variables, it is necessary for them to be similar within each group for them to be comparable. Therefore, the differences between mothers of younger children, the 2008 and 2011 cohorts, need to be small. Likewise, the differences between mothers of older children, the 2001-2004 and 2004-2007 cohorts cannot be too big. I have only included the observations from the first pair of observation years: 2010 and 2013. Looking at the differences within a group[51] between the pre- and post-reform periods, the change in the control variables are mostly small, but they seem to have been bigger for the control group. Whereas the mothers of the later cohorts seem to live in less central municipalities after the reform, the mothers of the earlier cohorts seem to live in municipalities that are more central in the post-periods. We can see that the age of the mothers was very stable in the treatment group unlike in the control group, which experienced a higher increase.

When it comes to education, the mothers in the control group have increased the percentage of mothers with a short higher education. This has been in the expense of the mothers with a lower education, mostly at the high school level, where we can see a big decrease. Likewise, we can see that the percentage of mothers with high school education in the treatment group has decreased, but the percentage of those with mandatory education has barely changed. On the other hand, there is an increase in the percentage of mothers with a higher education, both short and long.

To account for the observed differences and to check how sensitive the estimates are to the inclusion of the control variables, I will be adding them stepwise in my main analysis. First, I will add the child characteristic, then the mother characteristics, lastly I will add the municipality type the mother lives in.

[50] Estimated without including any of the control variables.
[51] Either the treatment group, which consists of the 2008 and 2011 cohorts or the control group, which consists of the 2001, 2002, 2003, 2004, 2005, 2006 and 2007 cohorts.

Table 6: Summary statistics with the differences within each group

	Treatment group			Control group		
	Pre-reform Born in 2008	Post-reform Born in 2011	Diff	Pre-reform Born in 2001/2/3/4	Post-reform Born in 2004/5/6/7	Diff
Panel A						
Outcome variables:						
Full-time employment ages 2 & 6-9	0.453 (0.498)	0.506 (0.501)	0.053	0.521 (0.500)	0.545 (0.498)	0.024
Full time employment ages 3 & 7-10	0.417 (0.494)	0.545 (0.499)	0.128	0.548 (0.498)	0.563 (0.496)	0.015
Agreed working hours ages 2 & 6-9	32.34 (7.929)	32.79 (8.788)	0.45	33.27 (8.444)	33.49 (8.971)	0.22
Agreed working hours ages 3 & 7-10	31.87 (9.217)	33.03 (9.044)	1.16	33.83 (9.263)	33.76 (10.04)	-0.07
N	718	657		1792	1675	
	Observed in 2010	Observed in 2013		Observed in 2010	Observed in 2013	
Panel B						
Control variables:						
Mother's age	32.87 (5.201)	32.86 (5.084)	-0.01	39.42 (5.059)	39.73 (5.154)	0.31
Number of children under 16	1.859 (0.882)	1.898 (0.853)	0.039	1.977 (0.756)	2.032 (0.755)	0.055
Mandatory school	0.152 (0.359)	0.153 (0.360)	0.001	0.160 (0.367)	0.145 (0.352)	-0.015
High School	0.314 (0.465)	0.293 (0.456)	-0.021	0.390 (0.488)	0.346 (0.476)	-0.044
Short higher education	0.401 (0.491)	0.413 (0.493)	0.012	(0.353) (0.478)	0.411 (0.492)	0.058
Long higher education	0.134 (0.341)	0.141 (0.348)	0.007	0.096 (0.295)	0.097 (0.297)	0.001
Single mother	0.113 (0.316)	0.078 (0.268)	-0.035	0.204 (0.403)	0.143 (0.350)	-0.061
Central municipality	0.793 (0.406)	0.757 (0.429)	-0.036	0.730 (0.444)	0.778 (0.416)	0.048
N	382	334		945	863	

Notes: Mean coefficients; Sd in parentheses. The shaded cells show the difference within each group, from a pre-reform period to a post-reform period. The ages in Panel A refer to the ages of the mothers' children at the end of the year they were observed.

7 Empirical results

7.1 Past trends

As previously mentioned in section 5.3.3, I will do a graphical test to check and see if the assumption of common trend holds. The graphs show the full-time employment or the working hours of mothers with younger children and of mothers with older children between 2008 and 2012. In this test, I will compare the past trends of the outcome variables of the treatment and control groups up until the CFC reform in 2012. If the identifying assumption holds, the trends should be parallel from year to year except for the year that the reform takes place. The trend for the younger children is expected to have a higher positive slope than the trend for the older children from 2011 to 2012.

Figure 7: Full-time employment trends

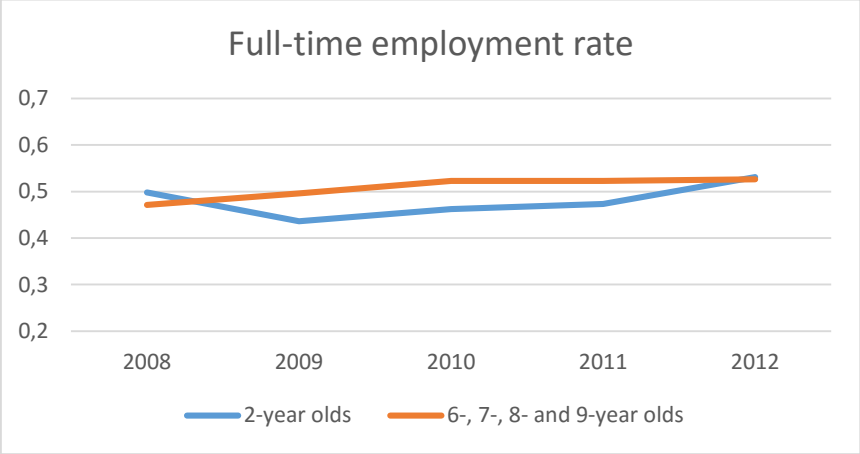


Figure 7. The graph shows the full-time employment trend of the mothers of children aged 2 years old at the end of the year, compared to the trend of the mothers of children aged 6-9 years old at the end of the year.

Figure 8: Weekly working hours trends

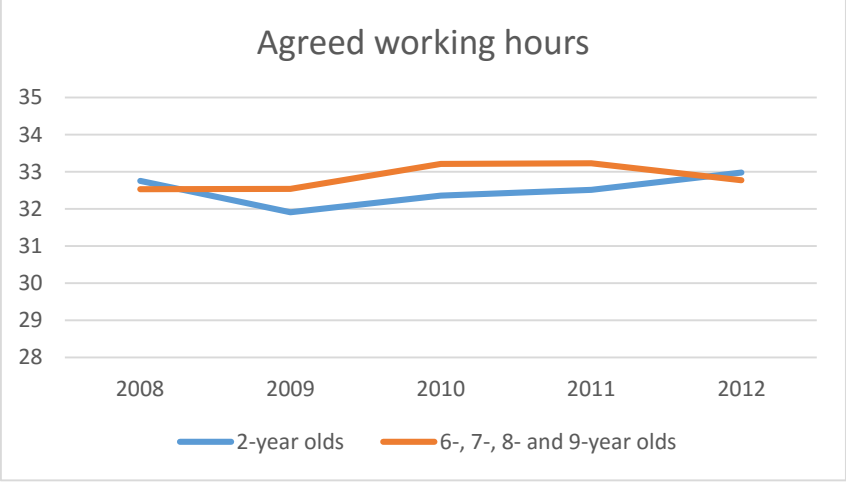


Figure 8. The graph shows the trend of the average number of weekly working hours of the mothers of children 2 years old at the end of the year, compared to the trend of the mothers of children aged 6-9 years old at the end of the year.

As we can see from the graphs in Figures 7 and 8, the change in the outcome variables from 2008 to 2009 had different directions for the two groups of mothers. The full-time employment went down for the mothers of younger children, while it increased for the mothers of older children. There is a similar development in the working hours for the mothers of younger children, while this variable remained stable for the mothers of older children. This may threaten our assumption of common trend, but an explanation for the difference could be the financial crisis, which happened in 2008. The mother in a household is typically the secondary earner of the household, while her husband or partner is the primary earner. During a recession, if the primary earner loses their job or suffers an income loss, the secondary earner may have to join the labor force or increase their working hours to make up for that loss. Previous studies found evidence supporting this phenomenon, which is called the added worker effect. The wives' response to their husband's unemployment was to enter the labor force if they were not already a part of it (Moehling, 2001), or to increase their working hours (Heckman & Macurdy, 1980).

This effect may have been stronger for the mothers with the older children, which is why their outcome variables did not decrease as they did for the mothers with younger children. Maybe the reason for this is that the mothers of 1- and 2-year-olds could still receive income through the CFC program, while staying at home. Additionally, the older children can go to school, which is free, while the younger children have to attend daycare centers, or be taken care of by childminders, which the parents have to pay for. Alternatively, they may let their relatives or someone else take care of their child for free, but only some of them have this option.

It seems like the trends are otherwise parallel from 2009 onwards until the reform happened in 2012. The outcome variables for the mothers of the younger children increased from 2011 to 2012, but we do not see similar changes for the mothers of the older children. Their full-time employment rate remained stable, while their average number of weekly working hours declined. This could indicate that the increase in the full-time employment rate and the working hours for the mothers of 2-year-olds was partly due to the treatment.

7.2 Main results

7.2.1 Full-time employment

Table 7 provides the first part of the main results; the first four models use the linear regression model from section 5.3.4, with full-time employment as the outcome variable. Panel A reports the short-run effect of the 2012 reform on full-time employment, which in

this paper refers to when the children in the treatment group were either 1- or 2-year-olds. This means that the treated children either were currently undergoing treatment or were waiting to receive the treatment. Panel B reports the effect in the longer run, which in this paper refers to when the children in the treatment group were either 2- or 3-year olds. These children were either currently undergoing treatment or had already received the treatment.

Table 7: The effect of the CFC reform in 2012 on the mothers' full-time employment

	(1) Full-time	(2) Full-time	(3) Full-time	(4) Full-time	(5) Full-time
Panel A:					
Age 2	0.0291	0.0281	0.0352	0.0307	0.137
	(0.0442)	(0.0439)	(0.0422)	(0.0422)	(0.186)
Mean marginal effect					0.0313
N	2524	2524	2524	2524	2524
Panel B:					
Age 3	0.113*	0.111*	0.100*	0.0991*	0.433*
	(0.0457)	(0.0457)	(0.0438)	(0.0439)	(0.194)
Mean marginal effect					0.0988*
N	2318	2318	2318	2318	2318
Included control variables:					
Child characteristic		X	X	X	X
Mother characteristics			X	X	X
Municipality type				X	X

Notes: Robust standard errors in parentheses. Significance level: + p<0.10, * p<0.05, ** p<0.01
The outcome variable is whether the mother is full-time employed or not. Models 1 to 4 are linear regression models; Model 1 is run without covariates, while the following models have covariates added stepwise. Model 2 includes the child characteristic (number of children under 16), while Model 3 adds the mother's characteristics (age, education, marital status). Lastly, Model 4 adds the municipality type the mother lives in (central or not). Model 5 is a logistic regression model, which includes the same control variables as Model 4. Each row shows the difference-in-differences estimators estimated with the mothers of a child in a specific age group as the treatment group. The ages refer to the younger children's age at the end of the year they were observed.

The table suggests a positive effect from the CFC reform in 2012 on maternal labor supply as expected from labor supply theory, seeing how all the estimates are positive. Removing the eligibility of the CFC benefit seems to have induced the mothers of younger children to become a full-time worker. This aligns well with the theory outlined in section 3.4.1, where we saw that the removal of eligibility led to a decrease in non-labor income for the mothers of 2-year-olds, who were not working. Simultaneously, mothers of 2-year-olds, who were working part-time, experienced both an increase in the slope of the budget line and a decline in their non-labor income. The increase in opportunity cost for leisure and the negative

income effect will both lead to a decrease in demand for leisure if it is a normal good. As a result, there will be more hours available to allocate to the labor market.

Model 1 presents the unadjusted difference in difference estimates, where none of the control variables is included. Control variables are added stepwise in the three following models to check how sensitive the estimates are to these variables. Adding control variables do not change the estimates too much; this indicates that the observable compositional changes among the different groups are not very big[52]. Model 4 may be the most reliable model since adding control variables seems to improve the precision of the difference-in-differences estimators.

The positive estimates indicate that the increase in labor supply for mothers of younger children were higher than the increase in labor supply for mothers of older children. The table suggests that the reform had a positive effect on the mothers' full-time employment. However, these estimates are not statistically significant, and an explanation for this could be that only part of the post-reform treatment group were currently receiving the treatment. The short-run estimates are around 3 percentage points, which is about 6 % of the mean full-time employment rate. This effect is less than the short-run effect found in previous studies about the introduction of the CFC program (Drange, 2012; Drange & Rege, 2013). However, the estimated percentage point increase is higher than the short-run effect Dahl (2014) had found, when he had considered the control group.

The effect seems stronger a year later, when every child was either 2 years old or 3 years old. We can see that the estimates are statistically significant at the 5 % level, when the treatment group consists of children turning 3 years old in the year of observation. Model 4 estimates that the treatment increased the probability of a mother with a 2- or 3-year-old child being full-time employed with 9.91 percentage points. This is substantial since it is almost one fifth of the mean full-time employment rate. There are no previous studies about the long-run effect of the 2012 reform, but the estimated effect is still very big if we compare them to results from studies about the introduction of the CFC program (Drange & Rege, 2013; Schøne, 2004a).

There is a possibility for a linear regression model to predict probability values higher than one or lower than zero. Therefore, a nonlinear regression model has also been estimated as a

[52] This also suggests that the unobserved compositional changes are not very big.

robustness test, to see if its mean marginal effects are similar to the estimates from Model 4. Model 5 reports the mean marginal effects based on the estimates from the logistic regression model mentioned in section 5.3.5, which includes the same control variables as Model 4. The estimated mean marginal effect in Panel B is a 9.88-percentage point increase in the probability of being full-time employed. This is statistically significant at the 5 % level and very close to the estimate from the linear regression model. Similar results are reported in Panel A, where the estimates from the linear and logistic regression models are 3.07 and 3.13 percentage points, respectively.

We can see that all the estimates in Panel B are all statistically significant at the 5 % level. These results are consistent with hypothesis H1 from section 3.4.2, which says that the full-time employment of mothers of 2-year-olds was affected by the CFC reform in 2012. Therefore, I can the reject the null hypothesis of there being no treatment effect on the treated mothers' full-time employment. The change in the CFC program seems to have induced a change in the maternal labor supply at the extensive margin, where the probability of being full-time employed increased.

7.2.2 Weekly working hours

Second part of the main results is reported in Table 8, where the outcome variable is the number of working hours per week. While the previous part analyzed the change in maternal labor supply at the extensive margin[53], this part analyzes the change at the intensive margin[54]. Using working hours as an outcome variable in addition to full-time employment is also a form of robustness testing. The estimates are expected to be positive and follow the same pattern as Table 7, where the effect becomes stronger in the longer run.

Table 8 has the same setup as Table 7, but this table excludes Model 5 because the variable measuring weekly working hours is not a binary variable. The results do support the first part of the main results as it shows a similar pattern to Table 7. All the estimates are positive, which indicates that the treatment had a positive effect on the mothers' weekly working hours. Model 4 estimates the 2012 reform to increase the weekly working hours by 0.248 hours in Panel A and 0.947 hours in Panel B, which is only a 0.8 % and 2.9 % increase. However, the standard errors are high, which leads to imprecise results, where none of the estimates is statistically significant. The estimates are higher in Panel B compared to Panel A, which

[53] Deciding on whether to be a full-time worker or not.
[54] Deciding on how many hours to work when they are employed.

indicates that the longer run effect is stronger than the short-run effect for mothers of 2-year-olds. Additionally, we can see that adding control variables will lead to small changes in the estimates, which again indicates that the compositional changes among the different groups are not very big.

Table 8: The effect of the CFC reform in 2012 on the mothers' weekly working hours

	(1) Working hours	(2) Working hours	(3) Working hours	(4) Working hours
Panel A:				
Age 2	0.240	0.272	0.332	0.248
	(0.820)	(0.812)	(0.793)	(0.791)
N	2155	2155	2155	2155
Panel B:				
Age 3	1.253	1.197	0.952	0.947
	(0.928)	(0.925)	(0.898)	(0.898)
N	2011	2011	2011	2011
Included control variables:				
Child characteristic		X	X	X
Mother characteristics			X	X
Municipality type				X

Notes: Robust standard errors in parentheses. Significance level: + p<0.10, * p<0.05, ** p<0.01
The outcome variable is the mother's agreed weekly working hours. All the models are linear regression models; Model 1 is run without covariates. The following models have covariates added stepwise. Model 2 includes the child characteristic (number of children under 16), while Model 3 adds the mother's characteristics (age, education, marital status). Lastly, Model 4 adds the municipality type the mother lives in (central or not). Each row shows the difference-in-differences estimators estimated with the mothers of a child in a specific age as the treatment group. The ages refer to the younger children's age at the end of the year they were observed.

Tables 7 and 8 suggest that the effect on maternal labor supply is stronger at the extensive margin than at the intensive margin. These findings indicate that a large part of the increase in labor supply was thanks to more mothers deciding to be full-time employed, whether by joining the labor market or by going from part-time to full-time employment.

7.3 Placebo test

An additional test I have done to check the assumption of common trend is to estimate placebo difference-in-differences. I will do a similar analysis to my main analysis, using the same regression as Model 4 from section 7.2, where the only difference is to move the whole analysis back or forward in time. In this analysis I ensure that everyone in both my pre-treatment group and post-treatment group will either all be treated or not treated at all. The purpose of this analysis is to see if there are non-parallel trends in full-time employment between mothers of 2-year-olds and mothers of older children.

Yearly data from the Norwegian Labor Force Survey, which I have available, is from 2008 to 2014. This allows me to estimate placebo difference-in-differences estimators using the observations from 2008 and 2011. Notably, the 2009 and 2010 cohorts are partly treated[55] and thereby cannot be part of these analyses. For the placebo test, the "treatment" group will be composed of mothers whose youngest child turns 3 years old in either 2008 or 2011. This would be the 2005 and 2008 cohorts, which received no treatment from the change in the CFC program. The control group will be composed of mothers whose youngest child turns 7-10 years old in either 2008 or 2011. None of these cohorts were treated, just like the "treatment" group. If the assumption of common trend holds, I expect the difference-in-differences estimators to be small and close to zero or statistically insignificant, because no one had any actual treatment in this placebo test.

Table 9: Placebo difference-in-differences estimators

	(1) FT	(2) FT	(3) Working hours	(4) Working hours
Age 3	-0.0873* (0.0436)	-0.0845+ (0.0437	-1.447 (0.894)	-1.418 (0.896)
N	2504	2504	2161	2161

Notes: Robust standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01
Models 1 and 3 include the same control variables as Model 4 in Table 7: mother's total number of children under 16, her age, education, marital status and the municipality type she lives in. Models 2 and 4 have an additional control variable; the male unemployment rate in each municipality type. The first two models have full-time employment as the outcome variable, while the last two models have agreed weekly working hours as the outcome variable. The age refers to the younger children's age at the end of the year they were observed.

From Table 9, Models 1 and 3 show the difference-in-differences estimators using the same regression as the one in the main analysis. We can see that the estimates are not close to zero, in Model 1, we can even see a statistically significant and negative treatment effect at the 5 % level. These estimates suggest that the change in full-time employment for the mothers with younger children is different from the change for the mothers with older children, that the latter had either a higher increase or a lower decrease than the former. The results from this placebo test threaten my assumption of common trend and could indicate that the empirical results in Tables 7 and 8 are not valid.

The explanation for the negative estimates could be the same explanation from section 7. The treatment group consists of mothers whose youngest child was either 2 or 3 years old, which means that part of the treatment group were entitled to the CFC benefit. Therefore, these mothers may have had a weaker added worker effect than what the mothers with older

[55] See Table 3 in section 2.1.

children have had.

From Figure A1 of the appendix, you can see the development of the unemployment rate in Norway for men and women, respectively. The increase in the unemployment rate was clearly much higher for men than for women from 2008 to 2011, which could partly be explained by the added worker effect. Mattingly and Smith (2010) analyzed the labor supply of women whose husband stopped working because of the financial crisis in 2008. According to them, mothers with children aged 5 or younger were less likely to enter the labor force to find a job than what mothers with older children were. Therefore, a recession may affect the labor supply of the treatment group with younger children differently compared to the labor supply of the control group with older children. The labor supply from mothers of older children would increase more or decrease less compared to that of mothers of younger children. It seems like the estimates will be biased if one of the observation years occurred right after the start of a recession. The estimated treatment effect will possibly be either more negative or less positive than the true treatment effect. In the worst-case scenario, the estimates may even be negative when the true effects are actually positive.

It is likely that the difference in added worker effect between the two groups of mothers is only a partial explanation for the findings in these placebo analyses, because the magnitude of the estimates from Table 9 is very high. To control for the added worker effect, an ideal control variable would be the male unemployment rate in the municipalities the mothers live in. The dataset does not tell us which municipality each mother lives in, however, we do know which type of municipality she lives in. Therefore, an additional control variable is added, to see if the male unemployment rates in central and less central municipalities will affect the estimates or not. We can see from Models 2 and 4 from Table 9 that they do make the estimates less negative, although it does not change much. This again threatens my identifying assumption of a common trend and suggests that the main results from this paper may not be valid and they will therefore be hard to interpret.

7.4 Subsample analyses

To test my hypotheses H2 and H3 outlined in sections 3.5.1 and 3.5.2, I will divide my sample into subsamples by different categorical variables.

In my first subsample analysis, I divide my sample into two subsamples by the mothers' educational level. This is to test hypothesis H2, which says that the mothers are affected by the CFC reform in 2012 differently depending on their educational level. The higher educated

mothers constitute 50.2 % of the whole sample and they consist of mothers who have had at least some education at a college or a university. The other subsample consists of the remaining less educated mothers. They either are high school graduates or have had less education than that; some of them have only finished mandatory school.

In my second subsample analysis, I divide my sample into two subsamples by their marital status. This is to test hypothesis H3, which says that the mothers are affected by the CFC reform in 2012 differently depending on their marital status. One subsample consists of the single mothers and constitutes 15.5 % of the sample. The rest of the mothers have a partner, through either marriage or cohabitation, and make up the other subsample.

7.4.1 Educational level

Tables 10 and 11 report the results from the first subsample analysis, where I have run pooled regressions, using a fully interacted[56] model. These two tables have the same setup, where Model 1 estimates the short-run effect, while Model 2 estimates effect in the longer run. I ran an additional regression for each model to estimate the difference in effects on the subsamples, and this is reported in the last row of each table. As mentioned in section 3.5.1, women with different educational levels are likely to respond differently to policy reforms related to the labor market. According to labor supply theory, the higher educated mothers are expected to have steeper budget lines and indifference curves than less educated mothers. Therefore, their optimal labor decision might be different from the decision of a less educated mother. It is uncertain which group of mothers will be affected more strongly by the change in the CFC program, as the results from previous studies regarding this program and educational level have been conflicting[57].

Table 10 provides the treatment effect on the probability of being full-time employed and it seems like the treatment affected the mothers in different directions depending on their educational level. The treatment seems to have had a negative effect for the higher educated and a positive effect for the less educated when the children are 1 or 2 years old, see model 1. However, this changes when the children are one year older, Model 2 estimates the treatment effect to be positive for higher educated mothers. It is estimated to increase the probability of the higher educated mother being full-time employed with 16.3 percentage points, and this is

[56] A model where the estimates would be the same if I ran regressions separately for each subsample.
[57] See previous literature in section 4.1.1.

statistically significant at the 1 % level. On the other hand, the treatment appears to have no significant effect on the less educated mothers when the children are 2 or 3 years old.

Table 10: Subsample analysis – education: Full-time employment

	(1) Full-time employment Age 2	(2) Full-time employment Age 3
Higher education	-0.00694	0.163**
	(0.0585)	(0.0602)
Lower education	0.0682	0.00194
	(0.0610)	(0.0643)
Difference	-0.0751	0.144
	(0.0845)	(0.0881)
N	2524	2318

Notes: Robust standard errors in parentheses. Significance level: + p<0.10, * p<0.05, ** p<0.01
Model 1 uses observations from 2010 and 2013, while Model 2 uses observations from 2011 and 2014. The outcome variable is whether the mother is full-time employed or not and both models include the control variables from Model 4 in Table 7. Last row informs us whether the effect of the treatment on one subsample is statistically significantly different from the effect on the other subsample. The ages refer to the younger children's age at the end of the year they were observed.

There is a similar pattern for the higher educated mothers in Table 11, which estimates the treatment effect on the mothers' weekly working hours. Model 2 estimates the treatment to increase the higher educated mother's working hours with 1.86 hours per week, this is statistically significant at the 10 % level. For the less educated mothers, we can see a similar drastic change in the effect of the treatment, but in the opposite direction. None of the estimates is statistically significant, but the treatment appears to go from being positive to negative, when the children are one year older.

We can see that the treatment effect on the higher educated mothers is not statistically different from the effect on the less educated mothers. These results do not support my second hypothesis and I fail to reject the null hypothesis of there being no difference in effects on the respective subsamples. Similar results were found in earlier studies concerning the introduction of the CFC program. Drange and Rege (2013) estimated similar treatment effects across educational levels in the short run. Likewise, Naz (2004) found no difference in the reduction of working hours between mothers with and without a university degree.

Table 11: Subsample analysis – education: Weekly working hours

	(1) Working hours Age 2	(2) Working hours Age 3
Higher education	-0.433	1.855+
	(0.891)	(1.046)
Lower education	1.128	-0.515
	(1.445)	(1.597)
Difference	-1.562	2.369
	(1.698)	(1.909)
N	2155	2011

Notes: Robust standard errors in parentheses. Significance level: + p<0.10, * p<0.05, ** p<0.01
Model 1 uses observations from 2010 and 2013, while Model 2 uses observations from 2011 and 2014. Both models have the mother's agreed weekly working hours as the outcome variable and both include the control variables from Model 4 in Table 7. Last row informs us whether the effect of the treatment on one subsample is statistically significantly different from the effect on the other subsample. The ages refer to the younger children's age at the end of the year they were observed.

7.4.2 Marital status

Similar to the first subsample analysis, I have run pooled regressions using a fully interacted model. Tables 12 and 13 have the same setup as Tables 10 and 11, and report the effects of the CFC reform in 2012 in the short run and in the longer run on each subsample. As previously mentioned in section 3.5.2, single mothers tend to have different budget lines compared to mothers with a partner. This can lead to different reactions to a reform related to the labor market, depending on their marital status.

Table 12: Subsample analysis – marital status: Full-time employment

	(1) Full-time employment Age 2	(2) Full-time employment Age 3
Single	0.172	0.148
	(0.120)	(0.113)
Married or cohabitant	0.00220	0.0870+
	(0.0453)	(0.0472)
Difference	0.170	0.0607
	(0.128)	(0.123)
N	2524	2318

Notes: Robust standard errors in parentheses. Significance level: + p<0.10, * p<0.05, ** p<0.01
Model 1 uses observations from 2010 and 2013, while model 2 use observations from 2011 and 2014. The outcome variable is whether the mother is full-time employed or not and both models include the control variables from Model 4 in Table 7. Last row informs us whether the effect of the treatment on one subsample is statistically significantly different from the effect on the other subsample. The ages refer to the younger children's age at the end of the year they were observed.

The treatment seems to have had a positive effect on the probability of full-time employment for both subsamples of mothers, see Table 12. The 2012 reform appears to have affected single mothers more, seeing how the estimates are larger for them than for the married or cohabitant mothers. However, only one of the estimates is statistically significant and it is the estimated treatment effect on married or cohabitant mothers when their child turned age 3 in the year they were observed. It is estimated to increase the probability of a mother of a 2- or 3-year-old being full-time employed by 8.7 percentage points, and this is statistically significant at the 10 % level.

Table 13 suggests that the treatment has increased the weekly working hours of all mothers, although none of the estimates is statistically significant. Corresponding to the results from Table 12, the treatment seems to have had a bigger effect on the single mothers compared to on the married or cohabitant mothers.

Table 13: Subsample analysis – marital status: Weekly working hours

	(1) Working hours Age 2	(2) Working hours Age 3
Single	1.731	0.817
	(3.446)	(3.540)
Married or cohabitant	0.111	0.761
	(0.821)	(0.929)
Difference	1.620	0.0557
	(3.543)	(3.660)
N	2155	2011

Notes: Robust standard errors in parentheses. Significance level: + p<0.10, * p<0.05, ** p<0.01
Model 1 uses observations from 2010 and 2013, while model 2 use observations from 2011 and 2014. Both models have the mother's agreed weekly working hours as the outcome variable and both include the control variables from Model 4 in Table 7. Last row informs us whether the effect of the treatment on one subsample is statistically significantly different from the effect on the other subsample. The ages refer to the younger children's age at the end of the year they were observed.

The estimated differences in the treatment effect on married or cohabitant mothers and on single mothers are not statistically significant for both outcome variables. I do not have sufficient evidence to reject the null hypothesis of there being no difference in effects on the respective subsamples because the results are not consistent with hypothesis H3. An explanation for why the estimated differences are not significant could be the high standard errors and the small sample. One reason the treatment effect seems to be bigger on the single mothers, may be because they do not have a partner's income to rely on, so they have a lower

non-labor income. Another reason could be the change in 2012 regarding the transitional benefit and its activity requirement. As mentioned in section 2.3, the single mothers were required to be active once the child was 12 months old since 2012, compared to the earlier threshold age, which was 36 months old. If this was the reason, it may have biased the main results and made the estimated effect more positive than the actual effect of the treatment. Nevertheless, the estimated effects on the married or cohabitant mothers are similar in size and direction as the estimates from the main analysis.

8 Summary and conclusion

The purpose of this thesis was to investigate the causal effect of the CFC reform in 2012 on the labor supply of mothers whose children were 2 years old. This reform removed the 2-year-olds' eligibility for a cash benefit, which parents could receive if they did not use publicly subsidized day care. I use a difference-in-differences method to analyze this by comparing the effect of the reform on two groups of mothers who had different exposures to the reform. The main results show a significant increase in the maternal labor supply at the extensive margin. The effect seems to have been stronger in the longer run, when the children were 2- or 3-year-olds, compared to the effect in the short run, when they were 1- or 2-year-olds. The probability of being full-time employed for mothers whose children were 2 or 3 years old was estimated to increase by 9.91 percentage points, which is a substantial change.

The placebo test results seem to threaten the identifying assumption, which is crucial for difference-in-differences estimations: the assumption of common trend. The estimated negative effect may have been a consequence of the financial crisis in 2007, because the added worker effect seems to be stronger for mothers of older children than for mothers of younger children. Hence, the estimates cannot be trusted if the mothers were observed right after the start of a recession, which is not the case in my analysis. However, it is likely that the difference in added worker effect only partially explains the findings in the placebo analyses, because the magnitude of these estimates is very high. This threatens the assumption of common trend and indicates that my main results may not be valid.

The subsample results suggest that higher educated and less educated mothers reacted to the reform differently, although the difference is not statistically significant. Single mothers also seem to increase their labor supply more than married or cohabitant mothers do after the reform. However, there is a possibility that the results had an upward bias from the change in the activity requirement for the transitional benefit, which single parents are eligible to receive.

Of my three hypotheses outlined in section 3, only the first hypothesis, H1 (*The change in the CFC program affects the full-time employment of the mothers of 2-year-olds*), is supported by the results from the regression analyses. The CFC reform in 2012 appears to have induced mothers of 2-year-olds, who were part-time workers or outside the labor force, to become full-time workers. I found no sufficient evidence to support hypotheses H2 (*The change in the CFC program affected mothers of 2-year-olds differently depending on their educational*

level) and H3 (*The change in the CFC program affected mothers of 2-year-olds differently depending on their marital status*). The sample may have been too small, which led to high standard errors and imprecise estimates.

One suggestion for a further study is to use a similar approach to this paper, but have a sample from a richer data set, such as registry data provided by Statistics Norway. In this way, the sample size will be larger and there will be a longer time series to test the identifying assumption. Another suggestion is to analyze the effect of the change in the CFC program in the long run, several years after the mother has been treated. It might also be interesting to investigate if the change affected mothers of western and non-western background differently, in both the short run and the long run.

References

Aldrich, J. H., & Nelson, F. D. (1984). *Linear probability, logit, and probit models* (Vol. 45). Beverly Hills: Sage.

Anderson, P. M., & Levine, P. B. (1999). *Child Care and Mothers' Employment Decisions.* (NBER Working Paper No. 7058). doi:10.3386/w7058

Angrist, J. D., & Evans, W. N. (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *The American Economic Review, 88*(3), 450-477. doi:10.3386/w5778

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics : an empiricist's companion*. Princeton: Princeton University Press.

Averett, S. L., Peters, H. E., & Waldman, D. M. (1997). Tax Credits, Labor Supply, and Child Care. *Review of Economics and Statistics, 79*(1), 125-135. doi:10.1162/003465397556467

Bakken, F., & Myklebø, S. (2010). Nedgang i bruk av kontantstøtte. [Decline in the use of the Cash-for-Care program]. *Arbeid og velferd, 2010*(1), 45-50. Retrieved from https://www.nav.no/

Becker, G. S. (1980). *Treatise on the Family*. Cambridge: Cambridge, MA, USA: Harvard University Press.

Bettinger, E., Hægeland, T., & Rege, M. (2014). Home with mom: the effects of stay-at-home parents on children's long-run educational outcomes. *Journal of Labor Economics, 32*(3), 443-467. doi:10.1086/675070

Bjørnstad, A. F. (2015). Økt sysselsetting blant mottakere av overgangsstønad. [Increased employment among recipients of the transitional benefit]. *Arbeid og velferd, 2015*(2), 95-104. Retrieved from https://www.nav.no/

Blau, D. M., & Robins, P. K. (1988). Child-care costs and family labor supply. *The Review of Economics and Statistics*, 374-381. doi:10.2307/1926774

Blundell, R., & Dias, M. (2009). Alternative Approaches to Evaluation in Empirical Microeconomics. *The Journal of Human Resources, 44*(3), 565-640. doi:10.1353/jhr.2009.0009

Borjas, G. J. (2013). *Labor economics* (6th ed. ed.). New York: McGraw-Hill.

Cools, S., Fiva, J. H., & Kirkebøen, L. J. (2015). Causal Effects of Paternity Leave on Children and Parents. *Scandinavian Journal of Economics, 117*(3), 801-828. doi:10.1111/sjoe.12113

Dahl, E. S. (2014). Fleire foreldre er i arbeid etter endringane i kontantsøtteordninga. [More parents work after the changes in the CFC program]. *Arbeid og velferd, 2014*(1), 25-36. Retrieved from https://www.nav.no/

Daugstad, G. (2006). *Omfang av bruk av kontantstøtte blant barn med ikke-vestlig innvandrerbakgrunn* [The extent of Cash-for-Care program usage among children with a non-western immigrant background] (Report No. 2006/26). Retrieved from Statistics Norway website: https://www.ssb.no/a/publikasjoner/pdf/rapp_200626/rapp_200626.pdf

Drange, N. (2012). *Crowding out Dad? The Effect of a Cash-for-Care Subsidy on Family time Allocation* (UiS Working Paper in Economics and Finance No. 2012/3). University of Stavanger.

Drange, N., & Rege, M. (2013). Trapped at home: The effect of mothers' temporary labor market exits on their subsequent work career. *Labour Economics, 24*, 125-136. doi:10.1016/j.labeco.2013.08.003

Finseraas, H., Hardoy, I., & Schøne, P. (2015, August). *Free Childcare and Mothers' Labour Supply: Evidence Using a School Starting Age Reform*. Paper presented at the European Economic Association Congress, Mannheim.

Gronau, R. (1973). The Effect of Children on the Housewife's Value of Time. *Journal of Political Economy, 81*(2), S168-199. doi:10.1086/260160

Gronau, R. (1977). Leisure, Home Production, and Work--the Theory of the Allocation of Time Revisited. *The Journal of Political Economy, 85*(6), 1099-1123. doi:10.1086/260629

Havnes, T., & Mogstad, M. (2011). Money for nothing? Universal child care and maternal employment. *Journal of Public Economics, 95*(11), 1455-1465. doi:10.1016/j.jpubeco.2011.05.016

Heckman, J. (1974). Shadow Prices, Market Wages, and Labor Supply. *Econometrica, 42*(4), 679-694. doi:10.2307/1913937

Heckman, J., & Macurdy, T. E. (1980). A Life Cycle Model of Female Labour Supply. *The Review of Economic Studies, 47*(1), 47-74. doi:10.2307/2297103

Hellevik, T. (2000). *Småbarnsforeldres yrkesdeltakelse og valg av barnetilsyn før og etter kontantstøttens innføring* [The employment of young children's parents and their choice of child care before and after the introduction of the Cash-for-Care program] (NOVA Report No. 2/00). Oslo: Norwegian Social Research.

Hellevik, T., & Koren, C. (2000). *Parents between work and care : effects of the new cash benefit scheme* (Vol. 7/00). Oslo: Norwegian Social Research.

Hirsch, A. A. (2010). Færre barn med kontantstøtte : brukere av kontantstøtte 1999-2009. [Less children with Cash-for-Care benefit : Users of the Cash-for-Care program 1999-2009]. *Samfunnsspeillet, 24*(4), 2-6. Retrieved from http://www.ssb.no/a/samfunnsspeilet/index.html

Holland, P. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association, 81*(396), 945-960. doi:10.1080/01621459.1986.10478354

Holtz-Eakin, D., Joulfaian, D., & Rosen, H. S. (1993). The Carnegie Conjecture: Some Empirical Evidence. *The Quarterly Journal of Economics, 108*(2), 413-435. doi:10.2307/2118337

Håkonsen, L. (2001). *Kontantstøtten : effekter på arbeidstilbud og inntektsfordeling* [The Cash-for-Care benefit : effects on labor supply and income distribution] (Report No. 2001/5). Retrieved from Statistics Norway website: https://www.ssb.no/a/publikasjoner/pdf/rapp_200105/rapp_200105.pdf

Imbens, G. W., Rubin, D. B., & Sacerdote, B. I. (2001). Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players. *American Economic Review, 91*(4), 778-794. doi:10.1257/aer.91.4.778

Killingsworth, M. R., & Heckman, J. (1986). Female Labor supply: A survey. In O. Ashenfelter & R. Layard (Eds.), *Handbook of Labor Economics* (Vol. 1, pp. 103-204). doi: 10.1016/S1573-4463(86)01005-2

Kimmel, J. (1998). Child Care Costs as a Barrier to Employment for Single and Married Mothers. *Review of Economics and Statistics, 80*(2), 287-299. doi:10.1162/003465398557384

Kjeldstad, R., & Rønsen, M. (2002). *Enslige foreldre på arbeidsmarkedet 1980-1999 : en sammenligning med gifte mødre og fedre*. [Single parents at the labor market 1980-1999 : a comparison with married mothers and fathers]. Statistical Analyses (1993 : online), Vol. 49.

Mattingly, M. J., & Smith, K. E. (2010). Changes in Wives' Employment When Husbands Stop Working: A Recession-Prosperity Comparison. *Family Relations, 59*(4), 343-357. doi:10.1111/j.1741-3729.2010.00607.x

Mincer, J. (1962). Labor force participation of married women: A study of labor supply. In Universities-National Bureau Committee for Economic Research (Ed.), *Aspects of labor economics* (pp. 63-105). Retrieved from http://www.nber.org/chapters/c0603.pdf

Mocan, N., & Altindag, D. T. (2013). Salaries and Work Effort: An Analysis of the European Union Parliamentarians. *Economic Journal, 123*(573), 1130-1167. doi:10.1111/ecoj.12056

Moehling, C. M. (2001). Women's work and men's unemployment. *The Journal of Economic History, 61*(4), 926-949. doi:10.1017/s0022050701042036

NAV. (2003, 2016.01.16). Vedlegg 1 til Hovednr. 55 Nr. 2 - Beregning av underholdsbidrag til barn: Sjabloner og andre satser [Appendix 1 to main no. 55 No. 2 - Calculation of child support: Fitting child-care expenses and other rates]. Retrieved from https://www.nav.no/rettskildene/Forside/Hovednummer/hovednummer-55-barnelova-gjelder-fra-01.10.2003

NAV. (2005). *Trygdestatistisk årbok 2005*. [Insurance Statistics Yearbook 2005]. Retrieved from https://www.nav.no/no/NAV+og+samfunn/Statistikk/Flere+statistikkomrader/Relatert+informasjon/trygdestatistisk-%C3%A5rbok-2002-2005

NAV. (2016, 2016.05.27). Transitional benefit. Retrieved from https://www.nav.no/en/Home/Benefits+and+services/Relatert+informasjon/transitional-benefit#chapter-1

Naz, G. (2004). The impact of cash-benefit reform on parents' labour force participation. *Journal of Population Economics, 17*(2), 369-383. doi:10.1007/s00148-003-0157-y

Norwegian Tax Administration. (2015, 2015.05.01). National Insurance scheme basic amount. Retrieved from http://www.skatteetaten.no/en/Rates/National-Insurance-scheme-basic-amount/

OECD Statistics. (2016a). *Incidence of FTPT employment - common definition* [Database]. doi:10.1787/data-00299-en

OECD Statistics. (2016b). *Labour force statistics by sex and age: indicators* [Database]. doi:10.1787/data-00310-en

Powell, L. M. (2002). Joint Labor Supply and Childcare Choice Decisions of Married Mothers. *The Journal of Human Resources, 37*(1), 106-128. doi:10.2307/3069605

Prop. 7 L (2011–2012). (2011). *Endringar i folketrygdlova og enkelte andre lover (samleproposisjon hausten 2011)*. [Changes in the National Insurance Act and other acts (proposition fall 2011)]. Oslo: Ministry of Labour and Social Affairs.

Prop. 8 L (2011–2012). (2011). *Endringer i kontantstøtteloven: Tilråding fra Barne-, likestillings- og inkluderingsdepartementet 4. november 2011, godkjent i statsråd samme dag: (Regjeringen Stoltenberg II).* [Changes in the Cash Benefit Act: Recommendation from the Ministry of Children, Equality and Social Inclusion November 4, 2011, approved by the Council of State the same day, (Stoltenberg II Cabinet)]. Oslo: Ministry of Children, Equality and Social Inclusion.

Rege, M., & Solli, I. F. (2013). The impact of paternity leave on fathers' future earnings. *Demography, 50*(6), 2255-2277. doi:10.1007/s13524-013-0233-1

Ribar, D. C. (1992). Child Care and the Labor Supply of Married Women: Reduced Form Evidence. *The Journal of Human Resources, 27*(1), 134-165. doi:10.2307/145915

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology, 66*(5), 688-701. doi:10.1037/h0037350

Rønsen, M. (2000). *Impacts on women's work and child care choices of cash-for-care programs*. Documents (Statistics Norway : online), Vol. 2000/13. Retrieved from http://www.ssb.no/a/histstat/doc/doc_200013.pdf

Rønsen, M. (2009). Long-term Effects of Cash for Childcare on Mothers' Labour Supply. *Labour: Review of Labour Economics and Industrial Relations, 23*(3), 507-533. doi:10.1111/j.1467-9914.2009.00456.x

Schøne, P. (2004a). Kontantstøtten og mødres arbeidstilbud: Varig effekt eller retur til arbeid? [The Cash-for-Care benefit and maternal labor supply : long-lasting effect or return to work?]. *Norsk økonomisk tidsskrift, 118*, 1-21. Retrieved from http://samfunnsokonomene.no/tidsskrifter/norsk-oko-tidsskrift/

Schøne, P. (2004b). Labour supply effects of a cash-for-care subsidy. *Journal of Population Economics, 17*(4), 703-727. doi:10.1007/s00148-003-0176-8

Snyder, C., & Nicholson, W. (2012). *Microeconomic theory : basic principles and extensions* (11th ed. ed.). Australia: South-Western Cengage Learning.

St.prp. nr. 53 (1997-98). (1998). *Innføring av kontantstøtte til småbarnsforeldre: Tilråding fra Barne- og familiedepartementet av 17. april 1998, godkjent i statsråd samme dag: (Regjeringen Bondevik I).* [Introduction of the Cash-for-Care benefit to parents of young children: Recommendation from the Minisitry of Children and Family Affairs April 17, 1998, approved by the Council of State the same day: (Bondevik I Cabinet)]. Oslo: Ministry of Children and Family Affairs.

Statistics Norway. (2015, 2015.06.18). Population´s level of education, 1 October 2014. Retrieved from https://www.ssb.no/en/utdanning/statistikker/utniv

Statistics Norway. (2016a). Table: 04232: Total fertility rate, women (C). Retrieved from http://www.ssb.no/en/statistikkbanken

Statistics Norway. (2016b). Table: 05111: Population, by labour force status, age and sex. Retrieved from http://www.ssb.no/en/statistikkbanken

Steckler, A., & McLeroy, K. R. (2008). The importance of external validity. *American Journal of Public Health, 98*(1), 9-10. doi:10.2105/ajph.2007.126847

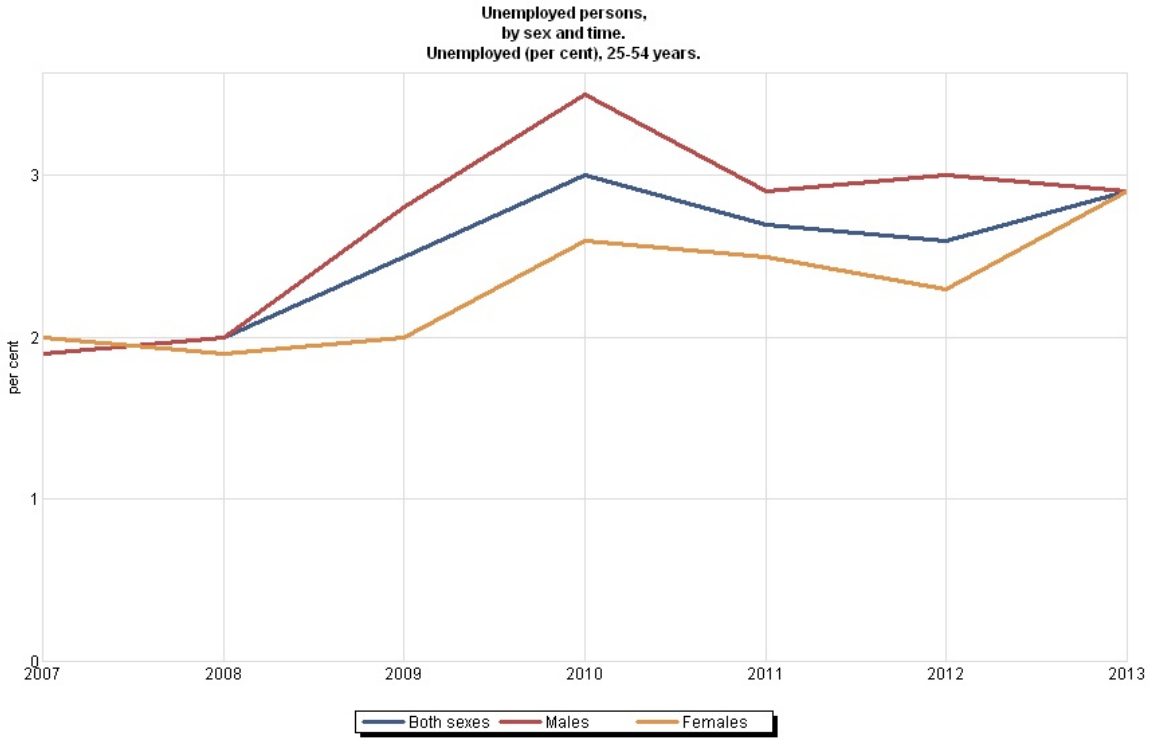
Tronstad, K. R. (2007). *Fordelingen av økonomiske ressurser mellom kvinner og menn : Inntekt, sysselsetting og tidsbruk* [The distribution of economic resources between women and men : income, employment and time allocation] (Report No. 2007/1). Retrieved from Statistics Norway website: http://www.ssb.no/a/publikasjoner/pdf/rapp_200535/rapp_200535.pdf

Ugreninov, E. (2003). Hvorfor jobber så få alenemødre? [Why do so few single mothers work?]. *Samfunnsspeilet, 17*(6), 33-37. Retrieved from http://www.ssb.no/a/samfunnsspeilet/index.html

Vollset, G. (2011). *Familiepolitikkens historie - 1970 til 2000* [The history of the family policy – 1970 to 2000] (NOVA Report No. 1/11). Retrieved from Norwegian Social Research website: http://www.nova.no/asset/4479/1/4479_1.pdf

Appendix

Figure A1: The unemployment rate from 2007 to 2013



Source: Statistics Norway. (2016). Table: 08517: Unemployed persons, by age and sex. Retrieved from https://www.ssb.no/statistikkbanken/selecttable/hovedtabellHjem.asp?KortNavnWeb=aku&CMSSubjectArea=arbeid-og-lonn&PLanguage=1&checked=true