Three essays on the evolution and on policy implications of working hours constraints

Dissertation zur Erlangung des Grades eines
Doktors der Wirtschaftswissenschaft

eingereicht an der

Fakultät für Wirtschaftswissenschaften
der Universität Regensburg

vorgelegt von:
Franziska Charlotte ZIMMERT

Berichterstatter:
Prof. Dr. Enzo WEBER
Universität Regensburg

Prof. Dr. Gesine STEPHAN
Friedrich-Alexander-Universität
Erlangen-Nürnberg

Tag der Disputation: 02. Juni 2020

Contents

List of Tables

List of Figures

List of Abbreviations

AIPW	**A**ugmented **I**nverse **P**robability **W**eighting
ATET	**A**verage **T**reatment **E**ffect on the **T**reated
CATET	**C**onditional **A**verage **T**reatment **E**ffect on the **T**reated
DiD	**D**ifference-**i**n-**D**ifferences
GSOEP	**G**erman **So**cio-**E**conomic **P**anel
HCA	**H**ome **C**are **A**llowances
IPW	**I**nverse **P**robability **W**eighting
KiföG	**Ki**nder**fö**rderungs**g**esetz
nd	**n**o **d**iscrepancy
oe	**o**ver**e**mployed
RD	**R**egression **D**iscontinuity
sd	**s**tandard **d**eviation
se	**s**tandard **e**rror
SSC	**S**ocial **S**ecurity **C**ontributions
SUTVA	**S**table **U**nit **T**reatment **V**alue **A**ssumption
TzBfG	**T**eilzeit- und **B**e**f**ristungs**g**esetz
ue	**u**nder**e**mployed

Chapter 1

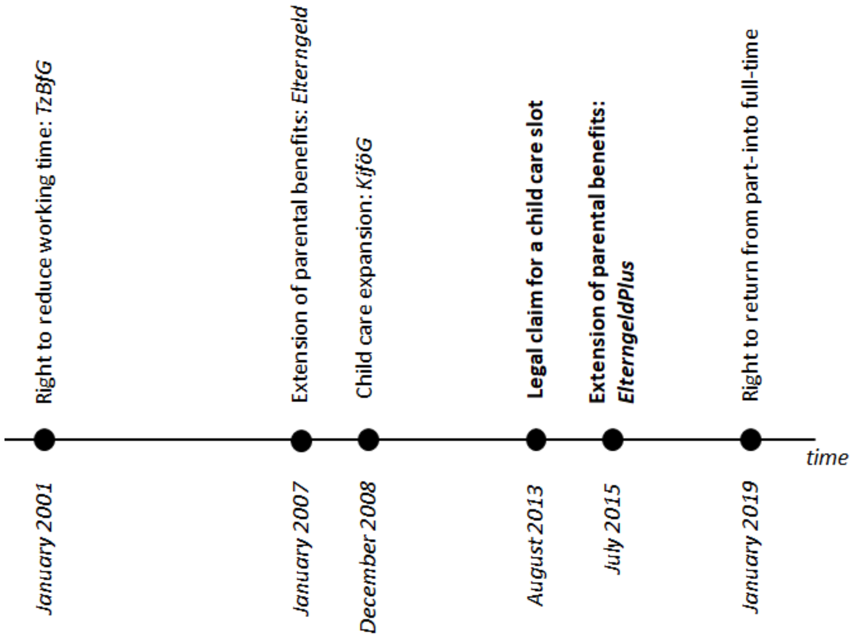
Introduction

Labor supply has gained importance in connection with the prediction of a decrease in Germany's labor force potential (Fuchs et al., 2016). In this regard, the current debate especially focuses on mothers for whom the employment potential is considered to be high. Issues like the reconciliation of family and work life and flexible working time arrangements rank high on the political agenda. Examples for policies having become effective over the last years vary from the introduction of the right to reduce working time in 2001 to the extension of parental leave eligibility in 2007 or the expansion of subsidized early child care culminating in the legal claim for a child care slot for children younger than three years old in 2013 (compare Figure 1.1). Apart from the political framework, employers are also increasingly making use of flexible working time arrangements like working hours accounts (Ellguth et al., 2018). While flexible working time measures are mainly employer-oriented and allow for sudden labor demand adjustments, employee-friendly arrangements are expected to become more important in the future (Zapf and Weber, 2017).
This thesis examines individual labor supply at different life stages with a special focus on working hour preferences and maternal employment. It consists of three articles each represented by a chapter. As a starting point, the first article analyzes the factors contributing to the evolution of working hour discrepancies. Since the empirical results show that especially mothers are concerned by those discrepancies, Chapters 3 and 4 evaluate different family policies with the potential to avoid and solve working hour discrepancies of young mothers. These policies relate to the availability of public child care and paid maternity leave.

1.1 Agreed, actual and preferred working hours

Before going into detail further, the concepts of agreed, actual and preferred working hours have to be explained. Agreed working hours typically refer to contracted or usually performed working hours while actual working hours also depict temporal fluctuations such as overtime hours. Working hour preferences represent the individual notion of what is desirable and are generally questioned in surveys. In Germany, the Socio-Economic Panel (GSOEP) and the Microcencus are the main data sources. Both surveys condition working hour preferences on income such that

FIGURE 1.1: Selected German reforms between 2001 and 2019



Abbreviations: Teilzeit- und Befristungsgesetz (TzBfG), Kinderförderungsgesetz (KiföG).
Source: Own representation.

FIGURE 1.2: Working hour discrepancies over time in %

(a) Women (b) Men

Notes: The sample includes between 4,778 and 14,877 observations per year. Self-employed, apprentices, interns and individuals completing their civilian or military service are not considered. Actual working hours (per week) include over time hours. Preferred working hours (per week) are not observed in 1996. Underemployment: preferred-actual hours> 2.5, overemployment: preferred-actual hours < -2.5, no/small discrepancy: preferred-actual hours ≥ -2.5 & preferred-actual hours ≤ 2.5.
Source: Own calculations based on SOEPlong v34, 1991-2017. Weighted analysis.

respondents are free to indicate their preferences without internalizing any other restrictions. Other differences in the questionnaire design of the GSOEP and Microcensus can lead to statistical variation. E.g., the Microcensus, in contrast to the GSOEP, filters the survey question on working hour preferences. Before indicating the amount of desirable hours, the respondent has to answer if she or he wants to change agreed working hours. Hence, respondents of the GSOEP may feel free to also indicate small changes such that the GSOEP statistics are expected to be an upper bound for working hour discrepancies (Holst and Bringmann, 2016). Moreover, in contrast to underemployment (wish for an increase of hours), the indication of overemployment (wish for a decline of hours) is voluntary in the Microcensus. Thus, the share of overemployed may be underestimated. This thesis uses both data from the GSOEP and Microcensus: Chapter 2 exploits the panel structure of the GSOEP whereas Chapter 3 makes use of the larger sample size of the Microcensus. As the research question of the latter article focuses on young mothers, the underestimation of overemployment is expected to be less severe.

The unifying result of the two data sources is that actual or agreed and preferred working hours do not necessarily coincide leading to overemployment or underemployment. As highlighted by Figure 1.2, using GSOEP data, more than one half of German employees express preference for working less or more while the majority belongs to the first group. Furthermore, male hour discrepancies, although on a high level, turn out to be mostly stable since German reunification. The share of underemployed women, however, has risen over the last years motivating the emphasis of this thesis. The concept of working hour preferences has the potential to broaden the perspective and to offer complete insight in individual labor supply. Adjusting actual working hours for those currently underemployed increases the aggregate work volume (Ehing, 2014). Furthermore, employees affected by working hour discrepancies show lower levels of life, health and work satisfaction (Grözinger et al., 2008). Hence, realizing or avoiding working hour discrepancies can have strong welfare effects (Bryan, 2007).

If the majority of German employees experience working hour discrepancies, which factors can explain their occurrence? Actual and preferred working hours are subject to individual, family and employer interests and thus, change accordingly but not necessarily in line with each other. They depend on the individual life or career stage that are shaped by occurring life events. Especially the birth of a child shapes the decision on the division of work within a household (e.g., Schulz and Blossfeld, 2006: for Germany). Although gender roles have converged for the last decades, differences between genders in terms of paid working hours and unpaid housework hours remain (Wanger, 2015). Maternal preference for an hour reduction might be only temporary, however it can mark future employment including career opportunities, earnings and social prospects after retiring. Goldin (2014) finds for the United States that maternal labor supply is even further decreasing some years after the birth of a child. This kind of path dependence is also present in Germany where traditional employment patterns reinforce the longer a couple is married (Schulz and Blossfeld, 2006). While women usually reduce employment after the birth of a child, empirical studies result in slight paternal compensations by offering more hours to offset the decline in household income (Drago et al., 2009; Pollmann-Schult and Reynolds, 2017; Reynolds and Johnson, 2012). Similarly, young fathers express only small preference for working hour reductions. However, cohort comparisons provide evidence that attitudes have changed as younger cohorts of fathers reduce actual working hours by one to two weekly hours in case the partner holds a full-time job (Pollmann-Schult and Reynolds, 2017). Nevertheless, these results highlight the strong family dependence of female employment careers.

1.2 Structure of this thesis

This thesis builds on these findings and is structured in three essays as follows. The first article analyzes the development of hour discrepancies by focusing on the most relevant household and job characteristics related to the creation and resolution of hours constraints. As creating or solving an hour discrepancy can also depend on the time already spent in the state of not having or having a discrepancy respectively, the empirical analysis is based on a discrete duration model controlling for individual fixed effects and using annual panel data from the German Socio-Economic Panel (GSOEP). The findings show that the occupational context, i.e., the individual job autonomy, is related to the evolution of working hour discrepancies for both women and men. In higher job positions individuals are more likely to become and remain overemployed. In contrast, the importance of household factors demonstrates gender differences and reveals that motherhood is linked to a lower probability for becoming underemployed, but the probability to leave this state is also smaller.

Hence, the following two chapters examine the subgroup of young mothers who have been in the focus of two recent social policies. Using the exogenous nature of these reforms, both studies pursue the identification of causal effects. The second article evaluates the effectiveness of the German child care expansion for under three-year-olds culminating in a legal claim for a child care slot (compare Figure 1.1, written in bold) in the context of female labor supply. Complementary to the first article, this chapter concentrates on the adjustment of agreed versus preferred working hours as the availability of low-cost external child care might have affected them differently. E.g., underemployed mothers might have responded to the reform by an increase of agreed working hours. Going back to Figure 1.2a, one can indeed detect a slight decrease of female underemployment after the legal claim for subsidized child care became effective in 2013. To rule out any spurious correlation, the article uses the exogenous rise of child care provision for difference-in-differences estimation and compares districts with a large increase of the child care coverage rate to those with a smaller child care expansion. The findings suggest that on average agreed and preferred working hours increase in response to the reform and that

the effect size is quite similar amounting to about five hours per week. Interestingly, only cohabiting mothers are characterized by a larger increase of agreed working hours in contrast to preferred working hours. This finding emphasizes the potential of child care provision in societies with a traditional division of household labor. In Germany, most mothers have a part-time working contract while the father works in full-time (Wanger, 2015).
Promoting an early return to work in part-time after childbirth is the main objective of the parental leave reform in 2015 (compare Figure 1.1, written in bold). The last chapter evaluates its effect on maternal employment with a special focus on the working time pattern. As prior studies mainly focus on the timing of the return to work and thus, the extensive employment margin (e.g., Baker and Milligan, 2008; Dahl et al., 2016), this article contributes to the literature by analyzing the intensive margin. In 2015, the German government decided to double the maximum receipt duration of a part-time subsidy right after the birth of a child. The dynamic optimization problem developed in the article proposes that the policy has an ambiguous effect on the decision when to return to work, but it makes part-time work more attractive relative to full-time work. This incentive may imply worse medium to long run employment prospects if mothers remain (involuntarily) part-time employed. Long working hours are related to better career opportunities and even considered to be one of the "last chapter" (Goldin, 2014) for reducing the gender wage gap. However, the empirical findings of the article cannot confirm such a part-time trap caused by the policy up to the child's second birthday. The reform rather yields additional employment of about two percentage points up to the first birthday. These effects are mainly driven by part-time employment of those mothers who would have also returned in part-time in absence of the reform, as the results do not show a decrease of full-time employment. The machine learning augmented estimation strategy also allows to estimate heterogenous effects. The heterogeneity analysis demonstrates that medium-earners and prior part-time working mothers have the strongest response to the new policy. Unfortunately, the administrative character of the data does not allow to examine individual working hour preferences which could further inform on those mothers unwilling to take up the part-time subsidy. Hence, the article cannot definitely answer if social norms, the lack of child care facilities or too low financial incentives are the driving forces.

The following three chapters include the articles as intended for publication in scientific journals. A conclusion ends this thesis.

Chapter 2

The creation and resolution of discrepancies between preferred and actual working hours over the life course

Joint with Prof. Dr. Enzo Weber[ab]

Abstract: This article contributes to the analysis of working hour discrepancies, i.e., under- and overemployment, by exploring how they emerge and resolve with special consideration of the household context. It uses a rich longitudinal data set, the German Socio-economic Panel, for a discrete duration analysis controlling for unobserved heterogeneity. We focus on the most relevant household and job characteristics. Findings suggest that job autonomy plays a crucial role for the creation and resolution of discrepancies. We especially contribute to previous studies by also examining path dependence and find that both the creation and resolution of discrepancies are characterized by positive duration dependence, but by negative occurrence dependence.

Keywords: working hour preferences, working hour discrepancies, household context, life course, working-time arrangements

Acknowledgements: The authors are grateful to Susanne Wanger, Ines Zapf and Johann Ludsteck for helpful suggestions and valuable input. The authors benefited also from comments from participants of the Doctoral Workshop on Applied Econometrics at the University of Strasbourg and the Jour fixe of the Graduate Programme of the Institute for Employment Research (IAB) and the School of Business and Economics of the University of Erlangen-Nuremberg (FAU).

[a]Institute for Employment Research (IAB) Nuremberg; Regensburger Strasse 104, 90478 Nuremberg, Germany

[b]University of Regensburg; Universitätsstrasse 31, 93053 Regensburg, Germany

2.1 Introduction

Discrepancies between preferred and actual working hours are a common phenomenon in industrialized countries (Reynolds, 2003, 2004; Stier and Lewin-Epstein, 2003). Empirical studies show that a discrepancy between working hour wishes and actual hours does not only deteriorate life, health or work satisfaction, but realization of working hour preferences can also strengthen the employment potential which is especially important in aging societies (Ehing, 2014). Hence, impeding the creation or supporting the resolution of working hour discrepancies can have positive welfare effects (Bryan, 2007). The underlying study seeks to further inform these debates by providing evidence on the dynamics of hour discrepancies, i.e., creation and resolution, in a household context. Several studies for different countries agree that the family context is one key determinant in addition to job and firm characteristics (Drago et al., 2005; Ehing, 2014; Fagan, 2001; Merz, 2002; Pollmann-Schult, 2009; Reynolds, 2003). Especially children are a determinant for under- and overemployment, i.e., the wish for an hour increase or decrease respectively. Empirical findings suggest that mothers are less likely to be underemployed while fathers do not prefer an hour reduction (Ehing, 2014; Pollmann-Schult, 2009). Gender disparities also show up concerning the presence of a partner. Single women tend to be under- rather than overemployed, but men without a partner have a lower probability for wanting an increase in labor supply. These findings emphasize that men and women are differently affected by time and monetary constraints imposed on the household. Apart from the family background higher levels of education and income determine overemployment (Pollmann-Schult, 2009; Reynolds, 2003) whereas underemployment is characterized by medium levels of education and low incomes (Ehing, 2014). While explaining the presence of working hour discrepancies, these studies take a cross-sectional point of view, i.e., they neglect the development of working hour discrepancies over time. Reynolds and Aletraris (2006, 2010) analyze the creation and resolution mechanism of an hour discrepancy using Australian and US data respectively. Reynolds and Aletraris (2006) emphasize that both a change in preferred and/or actual hours contribute to creating and solving over- or underemployment, but preferred hours are of higher importance. Furthermore, both studies find that a discrepancy of preferred and actual hours persists over time, especially the desire for fewer hours is hard to implement. This article contributes to the existing literature on working hour discrepancies in two ways. Firstly, by exploiting rich panel data, the German Socio-economic Panel (GSOEP, 1985-2016), for a longitudinal life course approach, which also allows a detailed view on the household and its employment situation. Beyond that, as a methodological advancement, by strengthening causal interpretations as the GSOEP not only enables to examine the individual development of working hour discrepancies over a long time period, but also allows to control for unobserved individual characteristics and cohort effects. In a discrete duration analysis (Allison, 1982), taking unobserved heterogeneity into account, this article examines how different individual and household characteristics contribute to the creation and resolution of working hour discrepancies over time. Hence, not the presence of working hour discrepancies, but their development is analyzed. In this context, the panel structure enables to consider path dependence. The German labor market serves as an interesting example as it is a country where the traditional employment pattern is still wide spread providing potential for working hour discrepancies (Wanger, 2015). The findings suggest that the individual job autonomy is one of the main driving forces for the creation and resolution of working hour discrepancies. Further interesting results concern the path dependence of working hour discrepancies. Both the creation and resolution of under- and overemployment become more likely the longer the current spell continues, but less likely the more spells already occurred in the past. The paper proceeds as follows: Section 2.2 deals with theoretical considerations. Section 2.3 includes a description of the data and of the estimation strategy. The regression results can be found in Section 2.4. The last section concludes with a discussion.

2.2 Theoretical considerations and hypotheses

Standard labor supply theory suggests that individuals are free to choose their working hours according to their preferences. Deviating from neoclassical considerations, economists highlight the existence of market imperfections (e.g., Bryan, 2007). Sociologists emphasize the role of changing preferences for justifying working hour discrepancies (Clarkberg and Moen, 2001; Reynolds and Aletraris, 2006). A common feature is that both approaches suggest that individuals are differently affected by a discrepancy of preferred and actual working hours dependent on their life stage including, e.g., the formation of the household, marriage, and the education of children. Thus, working hour discrepancies should not only be examined from the individual's perspective, but enclose the broader household context. The analysis provides a broad perspective and focuses on five main factors: the family composition, institutional constraints, the individual occupational position, the individual career stage and duration-related characteristics of the working hour discrepancy. For better readability, Table 2.1 summarizes the proposed hypotheses.

2.2.1 Family composition

Longitudinal research on the development of working hour discrepancies over time is scarce. However, a change in the life situation affects preferred hours (Campbell and van Wanrooy, 2013). Events in an individual's life like the arrival and departure of children are examples for altering working hour preferences. Discrepancies are likely to be created if an adjustment of the actual number of hours is hard to implement.

Social role models are an important factor for explaining traditional employment patterns that imply a full-time working man whose partner supplies a reduced amount of hours and has the main responsibility for the household. The majority of German women states the reason for their part-time employment to be family duties while the most important factor for men is that a full-time job cannot be found (Wanger, 2015). Although the employment rate of women has risen over the last decades, a major female conflict stems from reconciling housework and job (Hochschild and Machung, 1989) providing potential for working hour discrepancies. Men also face expectations in terms of male breadwinning which is considered to be crucial for compensating potential female income losses (Kaufman and Uhlenberg, 2000) or for the masculine identity (Potuchek, 1997). Hence, normative and time or monetary interdependencies within the household can cause both women and men not to supply the amount of hours they actually want to provide.

As women are more likely to suffer from the conflict of being simultaneously the ideal homemaker and worker, mothers should be even more affected by working time discrepancies than childless women (Reynolds, 2004). Suppose a full-time employed mother carries out the bulk of the domestic work including the care for children. *If her children are younger, she is more willing to reduce her working hours, and thus, an hours constraint should evolve with a lower probability. However, when children grow older, working preferences rise again resulting in a higher (lower) probability for getting under- (over-) employed in comparison with childless women.*

Apart from varying preferences due to changed life situations, resignation or settling can also be of importance in consideration of the resolution of hour discrepancies. It describes the circumstance individuals develop a preference for the working hours they can get (Reynolds and Aletraris, 2006). *Underemployed mothers might be more willing to adapt to their lower actual hours which helps solving the discrepancy while an adjustment of preferences is harder to achieve for overemployed mothers.*

Fathers face different social expectations. Although gender roles have been changing, traditional employment patterns persist (Wanger, 2015). Men are supposed to financially support their families (Potuchek, 1997) whereas a preference for an hour reduction might be interpreted as a lack of job commitment (Fagan, 2001). *Hence, fathers should be more (less) likely to end up in overemployment (underemployment) compared to childless men. On the other hand, solving a preference for less hours should be harder and thus, overemployment of fathers be characterized by a higher persistence. In contrast, underemployment is expected to be solved easier as actual working hours can adjust to higher preferences.*

2.2.2 Institutional constraints and interventions

We focus on reform effects in two important policy fields. Firstly, Germany underwent a large expansion of especially early child care facilities over the last 20 years (legal claim to kindergarten in 1996, *Tagesbetreuungsausbaugesetz* in 2005, *Tagesförderungsgesetz* in 2008). As increased availability and lower prices of child care incentivizes employment, but also reduces inter-role conflicts (Greenhaus and Beutell,

TABLE 2.1: Summary of hypotheses

	Creation of under-employed	Creation of over-employed	Resolution of under-employed	Resolution of over-employed
Children (*Reference* No children)				
Mothers of young children	−	−	+	−
Mothers of older children	+	−	+	−
Fathers	−	+	+	−
Child care expansion	−	−	+	+
Legal claim for a part-time job	+	−	−	+
Job autonomy	−	+	+	−
Career stages				
Earlier stages	+	+	−	−
Middle stages (*Reference*)				
Later stages	+	+	−	−
Path dependence				
Duration dependence	−	−	+	+
Occurrence dependence	+	+	−	−

Notes: + suggests a higher probability for creating/solving a discrepancy. − suggests a lower probability for creating/solving a discrepancy.
Source: Own representation.

1985) and creates cultural acceptance (Zoch and Hondralis, 2017) *we expect that working hour discrepancies are less likely to emerge and easier to become solved*.
Secondly, since 2001, German employees in firms with more than 15 employees have a legal claim for a part-time job independently from their family background. *We expect that this institutional change reduces the risk of becoming and staying overemployed. However, it has also the potential for making the creation and persistence of underemployment more likely by specifying a fixed amount of hours one cannot easily increase at a later point in time.*

2.2.3 Job characteristics

The divergence of working hour wishes and actual hours can also be expected to differ with respect to the individual occupational position. The normative signalling power of long working hours is especially high in professional and managerial positions characterized by non-standard tasks the results of which are hard to assess (Landers et al., 1996). *Thus, overemployment should more likely emerge and also persist in such positions compared to lower-rank jobs. For the same reasoning, underemployment is supposed to occur less likely in high-rank occupations. Furthermore, it is expected that the resolution of underemployment is more difficult for lower job positions and worse chances to change the employer as these characteristics deterioate one's bargaining position* (Reynolds and Aletraris, 2010).

2.2.4 Career stages

Furthermore, besides family duties and role models, market imperfections like asymmetric information can explain why actual working hours diverge from the preferred amount of hours. As long working hours serve as a signal of productivity to the employer, employees offer working hours that exceed their preferences (Sousa-Poza and Ziegler, 2003). Long working hours are especially important when individuals suffer from financial insecurity or the lack of job alternatives which forces them to accept job conditions they would otherwise reject (Stewart and Swaffield, 1997). However, job insecurity also matters for accepting and remaining in jobs if preferences exceed actual hours. This argumentation particularly holds during early life stages when employees still have to prove themselves or pursue a promotion and have less financial resources or shortly before retiring with worse reemployment chances and an increasing risk of health restrictions (Gielen, 2009). *Therefore, employees should be more likely to create and less likely to solve a working hour discrepancy in earlier and late phases compared to middle stages*.

2.2.5 Path dependence

The data and methodological approach allow for analyzing issues of path dependence. Two different forms of path dependence are distinguished. The first one relates to the duration of the current spell (duration dependence), the second one to the number of spells occurred in the past (occurrence dependence) (compare Heckman and Borjas, 1980). Regarding duration dependence, one may expect that individuals sort themselves into the state of (not) having a discrepancy. Hence, transitions into under- or overemployment (i.e., creation) are assumed to be negatively related to the duration spent in a non-discrepancy state. *I.e., the longer preferences match to actual working hours, the less likely would under- or overemployment occur*. For the resolution of discrepancies, discouragement or resignation may matter leading to an adjustment of preferences the longer a discrepancy lasts (Reynolds and Aletraris, 2006). However, from the perspective of utility theory, marginal costs of a discrepancy would increase the longer the spell already lasts, so that efforts to adjust actual working hours would increase, too. *Thus, the more time spent in under- or overemployment, the more likely employees are to leave this state either by an adjustment of preferred or actual working hours (positive duration dependence)*. While both channels cannot work at the same time for the same person, we will investigate the relevance of both adjustment mechanisms. As for the transition into unemployment, *past experience of working hour discrepancies is expected to increase the probability for having another spell of under- or overemployment, since individuals are more willing to accept bad job offers or working time arrangements* (Gibbons and Katz, 1991). For the same reasoning, *persistence of discrepancies is more likely for those with previous spells*.

2.3 Data, variables and estimation strategy

2.3.1 Data

To evaluate working hour discrepancies and their dynamics over time, panel data giving information on preferred and actual working hours over a long time span is needed. The GSOEP as an annual repeated household survey fulfills both criteria (see Wagner et al., 2007: for more details). Conducted since 1984, the GSOEP firstly only covered West-German households. After the German reunification also East-German households were interviewed and included in the analysis. The survey is designed to cover both economic and sociological questions such as the current life situation, employment, income and health, but also attitudes and different concepts of satisfaction. It has the great advantage that not only individual data is a hand, but also information on other household members which allows to approach the topic from a comprehensive household context. All individuals older than 16 years in the period from 1985 until 2016 are included. The waves of 1984 and 1996 have to be omitted as they do not contain information on working hour wishes. Extreme values of more than 80 hours per week (actual or preferred), as well as discrepancies exceeding a difference of 70 hours are dropped.

2.3.2 Outcome variables, data preparation and estimation strategy

Currently employed respondents are asked the following questions about their preferred and actual working hours: "If you could choose your own work hours, taking into account that your income would change according to the number of hours, how many hours would you want to work per week?" and "How many hours do you generally work per week, including any overtime?". Hence, actual hours diverge from agreed hours by including overtime.[1] The wording of these questions turns out to be meaningful, e.g., filtering the question on working hour preferences influences the amount of hours the respondent indicates (Holst and Bringmann, 2016). Stating a preference on working hours might furthermore be complex as individuals evaluate different background circumstances like the household income and household duties simultaneously. Hence, Campbell and van Wanrooy (2013) emphasize to consider preferences not as pre-determined and stable values. A working hour discrepancy $discr_{i,t}$ for individual i at time t is defined as the difference of desired and actual hours exceeding a threshold x such that underemployed respondents have a positive and overemployed employees a negative discrepancy. The threshold of 2.5 weekly hours in the baseline estimations is in line with previous studies (Knaus and Otterbach, 2019) and will turn out to be robust. The binary outcome variables

[1] Marginal and self-employed do not indicate agreed working hours. As these groups are included in the analysis, it is relied on the measure actual working hours.

indicate the creation ($discr\_cre_{i,t}$) and resolution ($discr\_res_{i,t}$) of a working hour discrepancy conditioned on the previous survey year:

$$discr\_cre_{i,t} = \begin{cases} 1 \text{ if } |discr_{i,t}| \geq x \text{ and } |discr_{i,t-1}| < x \\ 0 \text{ else} \end{cases}$$

and

$$discr\_res_{i,t} = \begin{cases} 1 \text{ if } |discr_{i,t}| < x \text{ and } |discr_{i,t-1}| \geq x \\ 0 \text{ else .} \end{cases}$$

The last lines of Panel A in Table 2.2 show that women have in equal shares no discrepancy or are overemployed. Most men are overemployed while the numbers also demonstrate that women are more often underemployed compared to men. These findings are very similar to other European surveys like the British Household Panel Survey (Bryan, 2007). Besides, women have similar working hour wishes independent from having no discrepancy or being under- or overemployed. For those experiencing a discrepancy, the absolute difference between preferred and actual hours amounts to about 10 weekly hours which is comparable to the male hour discrepancy. However, underemployed men have a weekly working hour wish of about 43 hours while for the overemployed it amounts to only 37 hours.

As the focus of the analysis lies on the emergence and resolution of a working hour discrepancy over time, the original panel data set has to be transformed into spell data. That means for those individuals for whom a discrepancy evolves, preferred and actual hours have to coincide at the first period of the spell. Table 2.3 represents possible preparation examples for two individuals like in Willett and Singer (1995). Individual 1 experiences two spells of an discrepancy creation, whereas the first spell of individual 2 is right-censored and not characterized by a discrepancy creation. For the resolution of a discrepancy, preferred and actual hours diverge at the beginning of the spell and data is prepared analogously. Getting non-employed is not considered as a resolution mechanism.

Panel B of Table 2.2 shows how many individuals create or solve a discrepancy in each period. In the first period both *discr_cre* and *discr_res* equal zero as a starting point. One period later, e.g., 1,547 women (about 61 percent of those women ever becoming underemployed and currently being in the second period) have become underemployed. This share is quite similarly decreasing by gender and discrepancy type such that those with long duration without discrepancy are on average less likely to create one. However, those with long discrepancy duration are also less likely to leave this state. A striking finding is that among underemployed men the percentage for leaving this state is relatively high in the second period (about 72 percent), but lower for leaving overemployment (about 49 percent). To examine to what extent these unconditional correlations are related to other variables of interest we use duration analysis.

TABLE 2.2: Descriptive statistics of outcome variables

	Women				Men			
Panel A	Hour distribution over discrepancy types							
	nd	ue	oe		nd	ue	oe	
Preferred working hours *Mean*	29.91	30.20	30.35		39.52	43.15	37.26	
Actual working hours *Mean*	30.17	19.83	40.55		39.77	32.17	47.89	
Difference between preferred and actual hours *Mean*	-0.26	10.37	-10.20		-0.25	10.98	-10.63	
N	39,920	16,046	39,839		46,598	9,801	59,228	
%	41.67	16.75	41.58		40.30	8.48	51.22	
Panel B	Number of individuals creating/solving a discrepancy over time							
	Creation		Resolution		Creation		Resolution	
	$discr\_cre = 1$		$discr\_res = 1$		$discr\_cre = 1$		$discr\_res = 1$	
	ue	oe	ue	oe	ue	oe	ue	oe
1st period N	0	0	0	0	0	0	0	0
%	0	0	0	0	0	0	0	0
2nd period N	1,547	3,375	1,772	2,775	1,008	4,859	1,511	3,763
%	60.91	59.95	63.81	53.77	57.11	59.25	72.44	48.79
3rd period N	418	1,021	497	892	324	1,457	304	1,433
%	47.61	49.76	54.32	41.18	47.72	48.23	56.19	39.79
4th period N	163	416	171	433	129	576	109	664
%	39.37	44.59	45.36	37.04	41.21	40.39	50.00	33.62
$\geq$ 5th period N	169	392	158	483	117	612	91	874
%	28.94	36.10	36.66	27.62	33.82	34.63	41.55	26.61
N	2,297	5,204	2,598	4,583	1,578	7,504	2,015	6,734

Notes: nd=no discrepancy, ue=underemployed, oe=overemployed.
Source: Own calculations based on GSOEP v33.1, 1985-2016. Pooled analysis in Panel A.

TABLE 2.3: Preparation as person-spell-period data set

individual	*spell*	*period*	*discr* (hours)	*discr_cre* (binary)
1	1	1	1.5	0
1	1	2	3.5	1
1	2	1	2.0	0
1	2	2	2.0	0
1	2	3	4.0	1
2	1	1	-0.5	0
2	1	2	1.5	0
2	1	3	0.5	0

Notes: *discr* measures the difference of preferred and actual weekly hours. *discr_cre* depicts whether a discrepancy has become created (1) or not (0). Analog data preparation for the resolution of hour discrepancies.
Source: Representation as in Willett and Singer (1995).

It allows to analyze the dynamics of working hour discrepancies in dependence from various factors. As annual panel data is at hand, a discrete duration analysis (Allison, 1982) is conducted where the dependent variables $discr\_cre_{it}$ and $discr\_res_{it}$ are binary indicators for creating and solving the discrepancy. In order to take unobserved heterogeneity into account, the fixed-effects or conditional logit estimator (Chamberlain, 1984) is used. Like the fixed effects estimator, the conditional logit estimator differences time-constant variables including unobserved characteristics out. Considering the role of social norms for the division of labor within the household, this property of the estimator is valuable as it is difficult to find a suitable proxy for the normative aspect. In this context social values are as well important. As attitudes, especially concerning the working time arrangement of men and women, have likely changed between the different cohorts included in the data set, the elimination of such factors is important for getting unbiased estimates. The probability for a positive outcome of $discr\_cre_{i,t}$ or $discr\_res_{i,t}$ is

$$
\begin{aligned}
P(discr\_cre_{i,t} = 1|X_1, ...X_T, c_i) = \Lambda(X_t\beta + c_i) \qquad &= \frac{e^{X_t\beta + c_i}}{1 + e^{X_t\beta + c_i}};\\
P(discr\_res_{i,t} = 1|X_1, ...X_T, c_i) = \Lambda(X_t\gamma + c_i) \qquad &= \frac{e^{X_t\gamma + c_i}}{1 + e^{X_t\gamma + c_i}}
\end{aligned}
$$

where $\Lambda(\cdot)$ denotes the logistic distribution, X a matrix of regressors and c_i individual-specific, time-constant factors. The contribution of an observation to the likelihood function depends on whether the outcome variable changes at least once, e.g., with only two observational periods $T = 2$, the probability for $discr\_cre_{i,2} = 1$ [$discr\_res_{i,2} =$

1] conditional on $discr\_cre_{i,1} + discr\_cre_{i,2} = 1$ [$discr\_res_{i,1} + discr\_res_{i,2} = 1$] becomes

$$\begin{aligned}
&P(discr_cre_{i,2} = 1 | X_1, X_2, c_i, discr_cre_{i,1} + discr_cre_{i,2} = 1) \\
&= \Lambda((X_2 - X_1)\beta) \\
&[P(discr_res_{i,2} = 1 | X_1, X_2, c_i, discr_res_{i,1} + discr_res_{i,2} = 1) \\
&= \Lambda((X_2 - X_1)\gamma)]
\end{aligned}$$

which is independent from c_i. Alternatively, one may estimate a competing risk model with multinomial logistic regression (compare Reynolds and Aletraris, 2010) that can also differentiate between transitions from under- to overemployment and vice versa. However, this kind of transition was found to be rare (about four percent of all changes for both kinds).

2.3.3 Explanatory variables

The explanatory variables of interest include different characteristics considering the individual him-/herself and the household he/she lives in. While the discussion of results will concentrate on the proposed hypotheses, we will provide a complete description of other included covariates in this section.

Firstly, a variable for the life course dimension is defined depicting important transitions in an individual's working life (Settersten Jr and Mayer, 1997) as they are the learning phase, the beginning of the working career, the establishment in the job, a middle phase and the years before and after retirement. The learning phase is created upon the question if the respondent is currently receiving education or training (vocational and further training or university) up to an age of 36 years. Once the learning phase has passed, individuals change to the three years-lasting stage of the career start which always refers to the highest level of education achieved. Hence, for persons with a vocational degree who decide to go to college, the career start will be postponed to the period after university. By the same token, breaks of unemployment after the learning phase are not taken into account. The phase of establishment in the working life lasts for five years after the stage of the career start. It is followed by the middle stage that is divided into two parts at the age of 45. The phase before retirement is defined upon the age and it includes individuals of 56 years and older. Workers older than 65 years are captured in the retirement phase and considered separately, as working beyond the statutory retirement age is supposed to be characterized by special conditions such as financial needs or high motivation. As respondents can enter the survey at each life stage, there are cases where phases cannot be determined successively starting from the learning phase. For those the weighted median age for each survey year of the persons from the already successively determined career start stage is used. According to the achieved educational level, the median age assigns life stage membership.

Furthermore, not only a categorical variable representing the children's age is included, but also the daily hours of child care provided by the parents themselves (coded 0 for childless individuals) and a measure of institutional child care, depicting whether the youngest child is in part- or full-time care or not in institutional care, are controlled for. In addition, the daily hours of housekeeping describe the hours spent for unpaid work. The daily hours for child care and housekeeping might be subject to an endogeneity problem as these variables can be determined simultaneously with the dependent variable. While this problem cannot be definitely solved, potential biases are mitigated by instrumenting those variables with their first lag, i.e., linear predictions of the first stage regression are inserted in the second stage[2].
To depict the individual and the partner's occupational position, the autonomy within the job (a generated variable strongly correlated with the job classification ISCO or the Prestige Scale of Treiman, 1976) is used. It describes the complexity or differentiation of tasks and responsibilities connected with them. The duration spent (un-)constrained is depicted by two variables. One measures the length of the spell until the discrepancy has occurred or been solved while the first and second period as well as the periods exceeding the fifth are grouped due to the small number of observations with long duration. The second depicts the number of spells that have occurred before the current spell. Again, more than two or three spells are grouped in categories.
Furthermore, we consider important institutional changes over the last years, i.e., the legal claim to work part-time and the expansion of child care facilities, which also allows to account for systematic differences across East and West Germany. For the latter aspect the binary indicator for young children is interacted with dummy variables standing for important periods defined by child care reforms (legal claim to kindergarten in 1996, *Tagesbetreuungsausbaugesetz* in 2005, *Tagesförderungsgesetz* in 2008). To take the legal claim to work part-time in firms with more than 15 employees into account, an interaction of the post-reform years with firm size that is greater than 20 employees, the next available threshold in the GSOEP, is considered. Hence, the resulting estimate gives the coefficient for treated employees.
Apart from these characteristics, other aspects of the individual, her/his partner, the firm side and the labor market are included as control variables. Education (no degree, vocational or university degree) and the gross wage are considered. The latter is based on the gross monthly individual income divided by the agreed working hours per month. Overtime allowances are considered in this calculation with a factor of 1.25.[3] Tenure and experience in full- or part-time work and in unemployment are included to depict the employment history. Besides the partner's occupational autonomy, her/his characteristics are represented by the employment status and the daily hours spent on child care and housekeeping. These variables are interacted

[2]We run linear regressions of each potential endogenous variable on all other included covariates. The resulting predictions are used in the second stage.

[3]The agreed monthly working hours are generated by multiplying the weekly hours with the factor 4.348.

with a binary indicator for the presence/activity of the partner and the partner's hours spent on child care and housekeeping are instrumented by their first lags.[4] A dummy for changing the job within or to another firm and the number of employees depict the firm side. Occupational labor market tightness, i.e., the number of vacancies divided by the number of unemployed control for the occupation-specific labor market situation. These numbers are taken from official statistics of the Federal Employment Agency according to the job classification system KldB88 and KldB10. Besides, a regional dummy for East Germany as well as a quadratic time trend are included. Summary statistics of the most relevant explanatory variables can be found in Table A.1 in the appendix.

2.4 Estimation results

Tables 2.4 and 2.5 depict the estimation results for becoming and leaving over- and underemployed conditioned on gender. About 7,000 observations of about 2,000 women are included in the sample of underemployment. In the male sample one can observe about 1,000 men and 5,000 data points. The overemployed samples are larger with about 15,000 observations of almost 4,000 women and more than 23,000 observations of 5,000 men. Each individual included in the sample experiences the creation or resolution of a discrepancy at least once and between 26 (resolution of male underemployment) and 52 percent (resolution of male overemployment) of them at least twice. The estimation is based on individual within-variation over time so that standard errors might be large for coefficients of variables with less changes. This applies to the children's age (about 5 percent of switches), the educational degree (1 to 2 percent of switches) and the presence of a partner (4 to 7 percent of switches). The analysis also contains a single fully interacted model that includes both under- and overemployed to identify statistically significant differences between both groups. These are indicated by an italic odds ratio in Tables 2.4 and 2.5.

2.4.1 Creation of a working hour discrepancy

Familiy characteristics

The first rows of Table 2.4 show that the odds ratios of having children reveals disparities between genders. In general, children are linked to a lower probability for women to become under- and overemployed. The odds ratios are strongly pronounced for mothers of younger children. This finding contradicts the results of Reynolds and Johnson (2012) for the US who find that the transition from no to one child increases the size of a discrepancy. Other transitions within the family of their

[4] Wunder and Heineck (2013) additionally look at the partner's hour discrepancy and Reynolds (2014) at hour preferences couples have for each other. Partner spillovers, however, are not in the focus of this paper.

analysis have less explanatory power. Fathers have a higher probability for getting underemployed when their children become older, but the odds ratio is not statistically significant. Thus, the expectation how discrepancies evolve for parents can only be supported for mothers of young children and female overemployment in general. Expectations do not hold for fathers.

Institutional constraints and interventions

Although the expansion of subsidized child care shows an increase of children institutionally cared for, the interaction of a dummy depicting stages for the expansion of child care with a dummy for children younger than six years old does not hint at changes regarding the creation of under- and overemployment which does not support expectations.[5] However, the pure, not interacted coefficient of institutional child care seems to matter. When children enter part-time care, their parents are more likely to become under- or overemployed, e.g., the odds for the creation of female underemployment are 4.7 times higher compared to a full-time slot. The coefficient is less strongly measured for the lack of care facilities. A possible reason is selectivity leading parents with a lower work commitment or for whom child care costs are too high to care for their children on their own.

Secondly, since 2001 there is a legal claim to work part-time in firms with more than 15 employees. We consider an interaction of the post-reform years with firm size that is greater than 20 employees, the next available threshold in the GSOEP and find that women affected by the legal claim to work part-time have on average a lower odds to get underemployed (0.628). Other interaction terms show no statistical relevance.

Job characteristics

Considering the occupational characteristics, it becomes obvious that reaching a higher level of occupational autonomy leads to a higher probability for becoming overemployed. The strong odds ratio of the latter is likely to be connected to peer pressure and weakly delimited workload, but can also be seen in the context of certain individuals having preferences for a steeper career path, which involves both higher autonomy and long working hours. The emergence of female underemployment is quite independent from the occupational autonomy. In contrast, the odds of getting underemployed are significantly lower for men switching into higher positions than into jobs of lower autonomy. To sum up, the results indicate a time conflict for jobs of higher autonomy and responsibility which supports the proposed hypotheses.

[5] For better readability, these results are not shown in Table 2.4. See Appendix A, Table A.2 for additional estimation results.

Career stages

The creation of a discrepancy follows a hump-shaped pattern with regard to the career stages. It is less likely to become under- or overemployed when switching to later career stages compared to the middle stage. Women are less likely to become underemployed and men less likely to become overemployed when starting the career. However, for women the odds of an overemployment creation are higher than in the middle stage, e.g., 1.6 times higher during establishing. For men the odds for becoming underemployed are also higher when they start their careers, but the coefficient is less precisely measured. Hence, the expectation that the occurrence of working hour discrepancies is more likely during earlier life stages can only be supported for the creation of female overemployment. Against expectations, both women and men show lower creation probabilities at later stages in life.

Path dependence

Further interesting results concern the life course dimension. The results hint at positive duration dependence, i.e., the more time spent in a non-discrepancy state, the more likely the occurrence of a discrepancy which does not support expectations. Besides, for those with multiple spells, i.e., those who have already solved a discrepancy, the emergence of another discrepancy is considerably less likely. This suggests that they do not show a higher willingness to accept bad working time arrangements.

Adjustment margin

In a second specification it is analyzed which of the adjustment mechanisms, preferred or actual working hours, prevails by introducing dummies for an increase/ decrease (decrease/increase) of preferred/actual hours in case of underemployment (overemployment). While these dummies are obviously endogenous with regard to the left hand side, the regression exercise is taken as descriptive evidence as in Reynolds and Aletraris (2006). Similar to these authors, the odds ratios show that a change of preferred hours is more important than a change of actual hours (about 1.4 to 2.2 times), but that both margins matter for the creation of hour discrepancies. As more than one quarter of observations within each sample (women/men, over-/underemployment) are characterized by an adjustment of preferred and/or actual hours, both turn out to be empirically relevant. The number of changes of preferred and actual hours is similar apart from the creation of female underemployment where the adjustment of preferences is stronger pronounced.

2.4.2 Resolution of a working hour discrepancy

Besides the creation of hour discrepancies, the life-course-oriented setting further allows analyzing the factors influencing how an existing discrepancy can be solved.

TABLE 2.4: Estimation results for the creation of a discrepancy

	Women		Men		Women		Men	
	Underemployment				Overemployment			
1 Family characteristics: Children (*Reference* No children)								
Children le6	0.0769**	(-2.11)	0.762	(-0.40)	0.309***	(-2.62)	0.710	(-1.04)
Children le10	0.257	(-1.02)	1.139	(-0.18)	0.365**	(-2.21)	0.773	(-0.72)
Children le15	0.275	(-0.98)	1.457	(-0.50)	0.383**	(-2.21)	0.874	(-0.40)
2 Job characteristic: Occupational autonomy (*Reference* Middle=3)								
Apprenticeship	0.533	(-1.56)	0.883	(-0.34)	0.489***	(-2.60)	0.611**	(-2.09)
Low=1	1.311	(-1.41)	1.037	(-0.16)	0.614***	(-3.11)	0.914	(-0.79)
2	1.236*	(-1.65)	1.107	(-0.53)	0.854*	(-1.80)	0.890	(-1.38)
4	1.342	(-1.25)	0.687*	(-1.70)	1.289**	(-2.35)	1.168*	(-1.87)
High=5	1.232	(-0.26)	0.565	(-1.12)	1.188	(-0.69)	1.529***	(-2.61)
3 Career stages (*Reference* Middle stage up to 45 years)								
Learning stage	0.135***	(-2.99)	0.871	(-0.27)	0.618	(-1.50)	0.507***	(-3.07)
Career start	0.329***	(-2.74)	1.229	(-0.50)	1.168	(-0.68)	0.525***	(-3.72)
Establishing	0.892	(-0.39)	1.182	(-0.54)	1.557**	(-2.48)	0.881	(-1.07)
Middle stage up to 55 years	1.098	(-0.48)	0.889	(-0.36)	0.778*	(-1.86)	0.744***	(-2.75)
Pre-retirement	1.233	(-0.62)	0.724	(-0.58)	0.741	(-1.41)	0.707**	(-2.08)
Retirement	0.704	(-0.52)	0.116	(-0.82)	0.329**	(-2.13)	0.235***	(-2.98)
4 Path dependence								
Period (*Reference* 1st and 2nd period)								
3rd period	6.139***	(-17.32)	5.680***	(-12.73)	4.917***	(-23.77)	4.931***	(-29.74)
4th period	10.870***	(-13.70)	9.156***	(-10.46)	7.448***	(-17.67)	6.263***	(-21.91)
5th period	14.040***	(-9.99)	10.80***	(-7.90)	9.933***	(-13.29)	7.402***	(-17.98)
Spell (*Reference* 1st spell)								
2nd spell	0.041***	(-14.17)	0.028***	(-11.23)	0.0819***	(-20.75)	0.102***	(-26.79)
3rd spell					0.0105***	(-18.35)	0.0144***	(-25.11)
N	7,303		5,121		15,884		23,343	
n	1,895		1,335		3,950		5,454	

Notes: Exponentiated coefficients (odds ratios) of fixed effects-logit estimation. Instead of providing marginal effects, odds ratios are indicated as they do not require plugging in a value for the unobserved component. The odds ratio gives the multiplicative value for the odds if the explanatory variable increases by one unit. t-values in parentheses. Standard errors are bootstrapped with 1,000 replications. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.010$. Other than listed explanatory variables are previously mentioned.
Abbreviations: Children le6 (le10, le15) means younger than 7 (11, 16) years old.
Source: Own calculations based on GSOEP v33.1, 1985-2016.

TABLE 2.5: Estimation results for the resolution of a discrepancy

	Women		Men		Women		Men	
	Underemployment				Overemployment			
1 Family characteristics: Children (*Reference* No children)								
Children le6	0.0665**	(-2.56)	2.698	(-1.25)	0.574	(-1.09)	1.192	(-0.53)
Children le10	0.312	(-1.10)	1.547	(-0.56)	0.736	(-0.60)	1.310	(-0.75)
Children le15	0.284	(-1.23)	1.888	(-0.84)	0.719	(-0.67)	1.199	(-0.53)
2 Job characteristic: Occupational autonomy (*Reference* Middle=3)								
Apprenticeship	1.569	-1.46	1.548	(-1.21)	0.523**	(-2.12)	1.054	(-0.19)
Low=1	0.586***	(-2.92)	0.812	(-0.91)	1.164	(-0.95)	1.277**	(-1.99)
2	0.896	(-0.90)	0.912	(-0.47)	1.084	(-0.90)	1.137	(-1.46)
4	1.125	(-0.60)	1.226	(-1.05)	0.729***	(-2.66)	0.825**	(-2.05)
High=5	0.749	(-0.26)	2.211	(-1.64)	1.281	(-0.82)	0.792	(-1.46)
3 Career stages (*Reference* Middle stage up to 45 years)								
Learning stage	0.377*	(-1.66)	0.255***	(-2.69)	0.989	(-0.03)	1.130	(-0.49)
Career start	0.894	(-0.26)	0.377**	(-2.27)	1.160	(-0.58)	1.200	(-1.05)
Establishing	1.206	-0.630	0.697	(-1.22)	1.759***	(-3.02)	1.054	(-0.43)
Middle stage up to 55 years	0.926	(-0.37)	0.765	(-0.96)	0.933	(-0.50)	0.818*	(-1.82)
Pre-retirement	1.098	(-0.27)	0.581	(-1.06)	0.916	(-0.42)	0.741*	(-1.70)
Retirement	3.040	(-1.05)	0.169	(-1.38)	1.861	(-1.21)	1.003	(-0.01)
4 Path dependence								
Period (*Reference* 1st and 2nd period)								
3rd period	7.359***	(-18.10)	8.228***	(-13.47)	3.939***	(-21.23)	4.061***	(-28.17)
4th period	12.50***	(-14.39)	21.98***	(-10.76	5.488***	(-17.7)	5.162***	(-21.43))
5th period	16.67***	(-10.59)	62.60***	(-8.54)	6.024***	(-12.52)	5.641***	(-17.26)
Spell (*Reference* 1st spell)								
2nd spell	0.042***	(-13.28)	0.054***	(-8.83)	0.071***	(-18.29)	0.073***	(-25.54)
3rd spell					0.009***	(-16.69)	0.008***	(-24.27)
N	7,545		5,333		15,964		25,041	
n	2,165		1,702		3,493		4,894	

Notes: Exponentiated coefficients (odds ratios) of fixed effects-logit estimation. Instead of providing marginal effects, odds ratios are indicated as they do not require plugging in a value for the unobserved component. The odds ratio gives the multiplicative value for the odds if the explanatory variable increases by one unit. t-values in parentheses. Standard errors are bootstrapped with 1,000 replications. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.010$. Other than listed explanatory variables are previously mentioned.
Abbreviations: Children le6 (le10, le15) means younger than 7 (11, 16) years old.
Source: Own calculations based on GSOEP v33.1, 1985-2016.

Thus, the following analysis will turn to the impact of the variables introduced above on the transition from a discrepancy to a non-discrepancy state (Table 2.5).

Family characteristics

The impact of children on the resolution of working hour discrepancies provides some interesting insights. In general, children impede the resolution of female working hour discrepancies, but especially underemployed mothers are concerned. Hence, the findings do not support the hypothesis that an adjustment of preferences, i.e., settling, plays an important role for underemployed mothers. Children make the resolution of hours constraints more likely for men, but the odds ratios are not statistically different from one. Thus, the coefficients act in the expected way for underemployed fathers while the hypothesis cannot be supported for fathers with a preference for an hour reduction.

Institutional constraints and interventions

Like for the creation of hour discrepancies we find that the expansion of child care does not change the odds ratios of the child-related variables and hence, there is on average no evidence for the effectiveness of the child care expansion in terms of solving hour discrepancies. In contrast to the creation of hour discrepancies, also the not interacted measure for child care lacks significance. However, we find that housekeeping and for women also time spent on child care keep from solving a discrepancy of both underemployment (0.924 and 0.963) and overemployment (0.882** and 0.934*).[6]
Moreover, we do neither find evidence for an increase of the probability to resolve overemployment nor for the persistence of underemployment after the legal claim to work part-time came into force in firms with more than 15 employees.

Job characteristics

Concerning the job characteristics, the resolution is less likely for underemployed in low prestige jobs and overemployed in high prestige jobs. These odds ratios also turn out to statistically significantly differ between under- and overemployed. In the former case a lack of flexibility and low negotiation power are plausible reasons. Like for the creation of overemployment, peer pressure or delimited workload might prove relevant for solving overemployment. Hence, these findings support the proposed hypotheses.

[6]For better readability, these results are not shown in Table 2.5. See Appendix A, Table A.3 for additional estimation results.

Career stages

Continuing with the impact of the career stages, women have a higher probability for quitting overemployment in earlier life stages than during the middle phase. Recapitulating that also the creation of overemployment was more likely during these stages, women turn out to be more prone to switches into and out of overemployment in the early phases. This finding does not correspond to expectations that here, the resolution of discrepancies is harder to be achieved. However, it demonstrates that persistence of female overemployment is most problematic in the middle of the working life. The pattern of the female underemployed is similar from the establishing stage onwards, but the odds ratios are smaller and not significant. Male underemployed have a lower probability for solving their discrepancy during earlier stages. Hence, underemployment represents a substantial problem for men in their early career, with those affected facing severe constraints which supports expectations for this group. We furthermore find that overemployed men have difficulties solving a discrepancy before retiring. This finding suggests the importance of flexible retiring schemes encouraging to decrease actual working hours.

Path dependence and adjustment margin

Like expected, multiple spells prevent constrained individuals from solving a discrepancy. Interestingly, while the reemergence of a discrepancy is less likely, once an additional spell occurs, it is hard to end. Furthermore, the longer the discrepancy has already lasted, the more likely is its resolution supporting positive duration dependence. The odds ratio is considerably stronger for under- than for overemployment. Regarding the explanation, the further analysis seeks to discriminate between resignation and increasing efforts to adjust actual working time.

In general, as for the creation of hour discrepancies, the adjustment of preferred hours contributes more to the resolution of discrepancies (1.4 to 2.1 times as much). About 28 up to 34 percent of the observations contains an adjustment of preferred and/or actual hours, where for the resolution of discrepancies the number of changes of actual hours is slightly higher compared to changes of preferred hours (except for underemployed men). Also along the duration of the spell, descriptive findings show that changes of preferences and actual working hours are similarly distributed. By the same token, the odds ratios of interactions of the duration variable with the binary indicators for a change of preferred and actual working hours do not statistically significantly differ from each other in almost all cases.[7] In conclusion, both adjustment margins - resignation and increasing effort to solve discrepancies - matter for resolution.

[7] Only for the resolution of underemployment, the adjustment of actual hours tends to gain importance the longer the discrepancy has already lasted.

2.4.3 Robustness

As a robustness check logistic regressions that do not control for unobserved heterogeneity are estimated. Especially concerning the career stages and the children's age there are some differences to the fixed effects-logit estimates which hints at an endogeneity problem for these variables. The differences of the variables considering the duration, period and spell, are the most striking. The logit estimates of period (spell) turn out to be smaller (greater) than in the fixed effects-logit estimation. Unlike to ordinary least squares estimation, the bias of logit estimates caused by unobserved heterogeneity cannot only be explained by (i) the correlation of the endogenous regressor with the residual and the correlation of the dependent variable with the residual, but also by (ii) the residual's variance (Mood, 2010). Even if (i) can be neglected, the coefficient will be downward biased due to (ii). The residual may contain individual characteristics like having a strong tendency for being constrained caused by health and satisfaction issues or workplace conditions (sorting effect). Then this "discrepancy type" will have a positive correlation with the creation of a discrepancy. Besides, it will be negatively correlated with period, i.e., the discrepancy occurs to an earlier point in time, and positively with spell, i.e., the number of spells increases for "discrepancy types". Hence, the first part of the bias (i) can be explained by an overall negative bias for the coefficient of period and an upward bias for the coefficient of spell in case of logit estimation that does not account for unobserved heterogeneity. An analogous argumentation holds for the resolution of a discrepancy. Thus, applying a logit estimation is not sufficient in the given analysis. Furthermore, a Hausman test suggests preference for the fixed effects-logit estimator as opposed to logit estimation for both creating and solving over- or underemployment.

A second specification varies in the bandwidth of the discrepancy interval. Both a smaller and larger discrepancy interval of 1.5 and 3.5 hours are tested and we find no strong differences to the baseline estimation. The odds ratios of the underemployed male subsamples turn out to differ the most which can result from limited within variation and small sample size. Furthermore, a specification that drops left-censored spells is tested which slightly changes the size of the odds ratios. The same holds for dropping observations before the German reunification in 1990 and observations with very small or long durations between subsequent interviews.

Additionally, the daily hours spent on care for relatives (instrumented by its first lag) is added as further regressor which is only available since 2001. Care duties are both associated with the creation and resolution of female underemployment, i.e., the odds for a creation (resolution) are 1.2 (0.7) times higher (lower) for an increase in care duties by one hour per day.

2.5 Discussion and conclusion

This article contributes to the existing literature on working hour discrepancies in examining the discrepancy of preferred and actual hours from a longitudinal approach, i.e., stressing how discrepancies emerge and resolve over time. Particularly, the definition of career stages gives the analysis a life-course orientation that has been predominantly neglected in the research of working hour discrepancies. Furthermore, the data structure allows to observe individuals over a long time horizon of 30 years, representing an advantage over existing studies that only conduct a two-wave comparison.

Our main findings concern four different aspects. Firstly, we find that mothers are differently affected by hour discrepancies in comparison with childless women. Although children are linked to a lower probability for the creation of hours constraints, those mothers experiencing constraints are more likely to get stuck. This finding especially holds for underemployed mothers of young children and suggests that traditional role models still prevail. Apart from social norms, regulations on the German tax and health insurance foster traditional employment patterns for married couples. Policies that encourage an equal employment pattern as well as increasing the supply of institutional child care can help avoiding and solving hour discrepancies for currently employed mothers. Although the results cast doubt on the effectiveness of the expansion of mainly part-time child care slots over the last twenty years, Chapter 3 will show that the introduction of a legal claim for a child care slot has a similar effect on preferred and agreed hours such that their discrepancy is not affected. Moreover, our findings show that actual take up, i.e., the transition from full-time to part-time care is related to a higher probability for getting underemployed.

Secondly, the job autonomy is one of the main determinants for becoming constrained and leaving this state. The creation and persistence of overemployment are related to a higher job autonomy while the opposite pattern holds for underemployment. Higher job autonomy is often linked with a steeper career path that expects long working hours as signal of motivation and performance. On the contrary, low negotiation power and less flexibility in the context of long-term working contracts is relevant for the development of underemployment. As, thirdly, the career stages also contribute to the creation and resolution of hour discrepancies, the findings suggest the importance of flexible working time arrangements not only for certain job positions, but also in dependence from different life stages (compare Gielen, 2009). E.g., we find that overemployed men face severe constraints in reducing actual working hours before retiring. Hence, this applies both to the amount of hours, which is still often subject to a strict full-time part-time divide, and the timing that sets conditions for working hour preferences.

Fourthly, this article examines path dependence of working hour discrepancies and finds that the creation of hours constraints exhibit positive duration dependence, but negative occurrence dependence. Thus, the results do not support the hypothesis of individuals sorting into the state of having discrepancies. Moreover, we find that the longer the constrained spell already lasts, the more likely individuals are to leave this state. However, with an increasing number of previous discrepancy spells, individuals are more likely to remain in under- or overemployment. These findings highlight that individuals are in a constant flux of creating and solving working hour discrepancies.

Reynolds and Aletraris (2006, 2010) additionally find out more about the adaption of preferred and/or actual working hours. However, these studies do not take unobserved heterogeneity into account that might especially occur if social norms are important. Hence, the presented study benefits also from a full panel structure in controlling for attitudes, norms and cohort effects that influence the working behavior. To understand the creation and resolution of working hours constraints in detail, future research may focus on how hour wishes and actual hours adjust over the life course. The life course perspective is especially fruitful in the context of different assumptions made by economists and sociologists on the adaption of preferences and actual hours. Economic theory highlights the role of varying actual hours adapting according to individual preferences while sociologists also emphasize the possibility of changing preferences. Additionally, knowledge on which of the adjustment mechanisms prevails can give further advice for strengthening the employment potential.

Chapter 3

Early child care and the employment potential of mothers: Evidence from semi-parametric difference-in-differences estimation

Abstract: This paper examines the effect of an expansion of subsidized early child care on maternal labor market outcomes. It contributes to the literature by analyzing preferred working hours. Semi-parametric difference-in-differences estimation based on survey data from the German Microcensus gives positive effects on the employment rate, as well as on agreed and preferred working hours. As agreed and preferred working hours adjust in line with each other, expansion of early child care can tap labor market potentials beyond those of currently underemployed mothers. Moreover, conditional effects show that especially better educated and cohabiting mothers respond to the reform.

Keywords: early child care, maternal labor supply, semi-parametric difference-in-differences, subsidized child care, working hour preferences

Acknowledgements: I gratefully acknowledge the valuable advice provided by Gesine Stephan, Enzo Weber, Susanne Wanger and Michael Zimmert as well as by participants of the German Economic Association's Annual Conference 2019 and of the Workshop on Labour Economics organized by the IAAEU in Trier. My special thanks go to the Research Data Centre of the German Federal Statistical Office in Düsseldorf and Fürth for the remote data access, in particular to Arne Schömann, Andreas Nickl, Alina Krauss and Lisa-Marie Jannaschk.

3.1 Introduction

Employment rates and working hours in industrialized countries vary strongly across gender for which the family background is often considered to be a main driving force (OECD, 2017). While male careers are less life-course dependent, women more often withdraw from the labor market or reduce their working hours after giving birth to a child. Hence, policymakers advocate an expansion of publicly subsidized child care in order to strengthen the employment potential in aging societies. Indeed, the female employment rate turns out to be higher in countries such as the Scandinavian states where child care is sufficiently provided. However, empirical studies cannot unanimously support a positive causal relationship between subsidized child care and female employment outcomes. I address this issue by evaluating not only the effect of low-cost subsidized child care on the employment share and agreed weekly working hours, but I further inform these debates by also examining underlying working hour preferences.

The article contributes to the existing literature in three different ways. Firstly, it extends the analysis to working hour preferences and the mismatch between agreed and preferred working hours. Working hour discrepancies are quite common in industrialized countries as the previous chapter and other studies suggest (Drago et al., 2005; Ehing, 2014; Fagan, 2001; Merz, 2002; Pollmann-Schult, 2009; Reynolds, 2003, 2004). Hence, evaluating if the availability of subsidized child care can affect working hour discrepancies is important in ageing societies as fulfilling a preference for more or less hours has positive effects on the employment potential and on individual life, health or work measures (Ehing, 2014; Matiaske et al., 2017). I use a rich data set from the German Microcensus which is a one percent representative sample of German households. The repeated cross-sections contain information on the household composition and its economic and social background and the data allows to examine over- and underemployment as well as individual working hour preferences. Furthermore, the focus is on early child care (children less than three years old) on which there is less empirical evidence on compared to preschool institutions. Thirdly, instead of applying a linear OLS estimator, a two-stage semi-parametric difference-in-differences (DiD) estimation procedure proposed by Abadie (2005) is used such that the linear form assumption in the outcome equation does not need to hold and common support between treated and control group can be enforced. Moreover, the approach allows to infer heterogenous treatment effects. There is a growing literature on evaluating the effectiveness of subsidized child care not only on parental, mainly maternal outcomes (e.g., Andresen and Havnes, 2019; Bauernschuster and Schlotter, 2015; Cascio, 2009; Gelbach, 2002; Havnes and Mogstad, 2011; Yamaguchi et al., 2018), but also on the child's development (e.g., Duflo, 2001; Felfe et al., 2015; Felfe and Lalive, 2018) and fertility (Bauernschuster et al., 2016). Many of these empirical studies rely on identification strategies that exploit exogenous variation resulting from quasi-experiments. In line with these studies I use the expansion of subsidized child care in Germany, induced by the introduction of a legal claim for a child care slot, to examine maternal employment.

I analyze the German labor market as an interesting example for the persistence of traditional employment patterns. Although the female employment rate converges to the male employment rate, there is strong variation when further conditioning on motherhood for both the extensive and the intensive margin. In 2011, almost one half of the childless couples both worked full-time while this was only the case for 22 percent of the couples with children (Wanger, 2015). In almost 20 percent of the families, the mother is not employed and the father works full-time (14 percent for childless couples). The majority of parents is characterized by a full-time working father while the mother holds a part-time position. About one quarter of part-time working women states the care for children or for people in need of care to be the reason for the employment status. Hence, the reform implemented in 2013 had a high potential to increase female employment both in terms of the extensive and intensive margin. Especially involuntarily underemployed mothers might have raised agreed hours.

In 2008, the German government formulated a law for the expansion of subsidized child care for children aged one to three (*Kinderförderungsgesetz KiföG*) culminating in a legal claim for a child care slot from August 2013 onwards. I use the exogenous variation of the expansion of subsidized child care induced by the reform to compare districts in which the coverage rate increased significantly (the treated or high-intensity group) with those for which the coverage rate changed only by a small amount (the control or low-intensity group). To be more concrete, I follow the approach of Bauernschuster et al. (2016), Felfe et al. (2015) and Havnes and Mogstad (2011) who exploit spatial variation of German districts, Spanish states and Norwegian districts respectively for which the child care coverage expanded differently after the legal framework had changed. The authors define control and treatment group by dividing the observational units at the median of the percentage point change in the coverage rate. Thus, the DiD strategy compares labor market outcomes of mothers with children aged up to three years in treated districts with those where child care increases to a lesser extent before and after the legal claim came into force.

The resulting intention-to-treat estimates give a positive impact both on the extensive and intensive margin. Mothers of up to three-year-olds in districts with a large increase of the child care coverage rate have a 5.7 percentage points higher employment rate after the reform than their counterparts in districts with a lower expansion of subsidized child care. Agreed and preferred working hours are on average about five hours per week higher and change similarly such that their mismatch is not affected. The results are robust to several sensitivity checks. Especially the common trend for treated and control group in the absence of the reform seems to hold. I furthermore show that the estimates are higher for better educated mothers and that the adjustment mechanism of agreed and preferred working hours differs for

cohabiting mothers.
The paper proceeds as follows: The next section gives an overview on previous empirical studies and describes theoretical considerations. Section 3.3 explains the institutional background of the German child care system including its reform and how it is exploited for the estimation strategy. Furthermore, the data is presented. The estimation results can be found in Section 3.4. The last section concludes.

3.2 Child care availability and maternal employment

3.2.1 Related empirical findings

Estimating the causal effect of publicly financed child care on employment outcomes suffers from several difficulties. One is that its price and the availability of informal child care provided by the family are often insufficiently observed (Havnes and Mogstad, 2011). Another problem is the endogeneity of child care availability and costs to employment measures. Hence, most studies apply quasi-experimental designs that benefit from exogenous variation induced by a policy reform or an instrumental variable (for a review see Morrissey, 2017). However, the empirical results strongly differ between countries depending on the economic conditions before the reform was implemented, the population under consideration and the organization of child care including private, public and informal arrangements. The bandwidth of the effect of more generous child care varies from positive (Andresen and Havnes, 2019; Baker et al., 2008; Bauernschuster and Schlotter, 2015; Berlinski and Galiani, 2007; Berlinski et al., 2011; Fendel and Jochimsen, 2017; Fitzpatrick, 2012; Gelbach, 2002; Geyer et al., 2015; Lefebvre and Merrigan, 2008; Nollenberger and Rodríguez-Planas, 2011; Schlosser, 2005; Yamaguchi et al., 2018) to negligibly small or insignificant coefficients (Bettendorf et al., 2015; Cascio, 2009; Givord and Marbot, 2015; Goux and Maurin, 2010; Havnes and Mogstad, 2011; Lundin et al., 2008).
Gelbach (2002) uses an instrumental variable approach to estimate the effect of public school enrollment by exploiting quarter of birth regulations for the US. He estimates a positive effect on the employment rate and on weekly hours for single mothers while the coefficient is slightly smaller for married women. Fitzpatrick (2012) finds only a positive effect for single mothers in the US with a regression discontinuity (RD) design that is as well characterized by a child's eligibility to kindergarten. Berlinski et al. (2011) apply a RD design for Argentina where kindergarten enrollment is defined by a cut-off date. Women whose youngest child attends kindergarten have a higher employment probability, also in full-time, and weekly hours rise on average by 7.8.
The majority of empirical studies uses quasi-experiments for a DiD design. Schlosser (2005) evaluates a reform that affected Arab mothers of children aged three to four in Israel. She finds that free public preschool increased maternal employment by 8.1

percentage points and average weekly hours by 2.8. Berlinski and Galiani (2007) estimate positive employment effects for Argentinean mothers of children aged three to five. The authors exploit a preschool construction program taking place in the mid 1990s. The staggered introduction of subsidized child care in the Canadian province Quebec was found to increase female employment by 7.7 percentage points (Baker et al., 2008) which is in line with Lefebvre and Merrigan (2008) who evaluate the same reform and also find a positive effect on working hours. Positive effects can also be found for Spain (Nollenberger and Rodríguez-Planas, 2011) and Germany (Bauernschuster and Schlotter, 2015; Fendel and Jochimsen, 2017; Geyer et al., 2015). Bauernschuster and Schlotter (2015) show that the transition to kindergarten defined by cut-off rules is related to an increase in labor force participation by 36.6 percentage points and in average weekly hours by 14.3, i.e., by 23.2 percent. Fendel and Jochimsen (2017) find positive short-term effects on the maternal labor force participation for the child care reform of August 2013 including the legal claim for child care and the introduction of home care allowances. With a microsimulation study Geyer et al. (2015) demonstrate that universal child care has large, positive effects for children older than one year. Other studies evaluating labor market responses of mothers with children younger than three years old result in positive effects for cohabiting mothers characterized by a shift to full-time employment (Andresen and Havnes, 2019: for Norway).

In contrast, Lundin et al. (2008), Givord and Marbot (2015) and Havnes and Mogstad (2011) find estimates for maternal (full-time) employment in Sweden, France and Norway that are close to zero. The latter article evaluates the expansion of child care availability for three to six year old children and suggests that public child care mainly crowded out informal arrangements.

Referring to this, ambiguous findings from preliminary empirical work might also stem from the ignorance of underlying preferences. Lundin et al. (2008) and Givord and Marbot (2015) might have find no effects in the context of an already high share of working mothers whose preferred and agreed working hours potentially match. Countries with lower maternal employment which show positive responses to the availability of subsidized child care could be those with a higher share of underemployed women adjusting agreed to preferred working hours. In line with these considerations several authors emphasize the role of adjusting preferences in case of occuring life events like the birth of a child (Campbell and van Wanrooy, 2013; Drago et al., 2005; Reynolds and Johnson, 2012). Reynolds and Johnson (2012) evaluate how the number of children living in the household affects preferred and actual working hours for the US and find that the birth of the first child is related to a larger drop of female working hour preferences compared to actual working hours. The impact on male working hours does not statistically significantly differ from zero. This finding is in line with Drago et al. (2005) who evaluate working hour preferences for Australian employees and conclude that women are more sensitive to changing life conditions than men. Chapter 2 examines the mismatch dynamics considering

household and job characteristics and finds suggestive evidence that the lack of institutional care arrangements may foster the creation of working hour discrepancies. However, the mentioned studies do not examine the direct effect of subsidizing child care on maternal working hours or neglect the adjustment mechanism (agreed versus preferred working hours).
Before turning to the empirical analysis, the next section is dedicated to a theoretical discourse how preferred working hours may be affected by the expansion of institutional child care.

3.2.2 Theoretical considerations

The examined reform mainly focused on the availability instead of the affordability of child care (Kreyenfeld and Hank, 2000) which in turn can be considered as an implicit subsidy (Berlinski and Galiani, 2007). The legal claim introduced in August 2013 guaranteed parents at least part-time care (four hours per day). Neoclassical economic and sociological theories predict an increase for female labor force participation whenever child care costs decrease. However, the effect on the intensive margin remains ambiguous and represents a weighted average of the substitution and income effect (Gelbach, 2002). Moreover, the overall effect is determined by the degree to which public care crowds out other care arrangements. Previous studies show that the impact on the extensive margin is negligible if women substitute informal or private with institutional, subsidized arrangements (Havnes and Mogstad, 2011; Yamaguchi et al., 2018) or the female labor market participation is already high (Givord and Marbot, 2015; Lundin et al., 2008).
Furthermore, sociological theory predicts that family policies encouraging female employment shape social norms (Gangl and Ziefle, 2015; Zoch and Hondralis, 2017). Hence, working mothers feel more accepted if they use institutional care resulting in an increase of female employment. The availability of subsidized child care can have different effects on preferred and actual working hours. Neoclassical theory assumes perfect labor markets on which the absence of frictions equalizes working hour preferences and actual hours. However, a mismatch can occur whenever social or occupational constraints prevent employees from supplying the preferred hours (Drago et al., 2005; Ehing, 2014; Fagan, 2001; Merz, 2002; Pollmann-Schult, 2009; Reynolds, 2003, 2004). The availability of child care has the potential to decrease this discrepancy while the adjustment of preferred and/or agreed hours depends on the state of being under-, overemployed or unconstrained before the reform came into effect.
Before further going into detail, the concept of working hour preferences has to be explained. In general, the formulation of the survey question on preferred working hours differentiates between two concepts of hours constraints. Although most surveys on working hour preferences consider earnings adjustments, one has to distinguish if respondents are free to indicate their preferences or if they take other constraints like the care for children into account. Campbell and van Wanrooy (2013)

suggest for further clarification that closed-ended questions on working hour preferences can be followed up by questions on the feasibility of preferences or on the constraints preventing from adjusting to the respondent's preferences. These are exactly the kind of questions the German Microcensus used in this article adds to the indication on working hour preferences. Thus, respondents indicating the wish for a change of working hour preferences are likely to freely choose the amount of preferred working hours and give information on the feasibility and potential external constraints in other related survey questions. In reference to the research question of this article, mothers with preference for supplying more working hours, but who are restricted by the lack of external child care offers, are not supposed to internalize the child care constraint when stating their desired working hours.

In this regard, underemployed women are expected to adjust their agreed hours to their preferred amount as the availability of subsidized child care lowers time and monetary constraints. Furthermore, institutional child care can attenuate interrole conflicts between family and occupational requirements (Greenhaus and Beutell, 1985). Hence, overemployed mothers are supposed to adjust a working hour mismatch by an increase in preferred hours. Finally, if unconstrained women adjust their agreed hours due to the availability of subsidized child care, the change should go in line with an adjustment of their preferences. Due to the repeated cross-sections, the analysis does not allow to examine the individual adjustment of working hours and cannot directly test the proposed mechanisms. However, mean changes of preferred and agreed working hours hint at distributional shifts that can reveal more on adjusting underlying preferences when combined with findings on changes of the share of unconstrained or under- and overemployed mothers. Given that mothers are constrained due to an insufficient supply of external child care, the reform can theoretically lead to the following four scenarios. Firstly, only average agreed working hours rise which should show up in a lower share of underemployed mothers. Secondly, only the effect on average working hour preferences is positive. Then the amount of overemployed women is expected to fall. Thirdly, if both average agreed and preferred hours increase, either the share of under- and overemployed should fall or there may be no shift as the hour increase is driven by at least one of the groups. Fourthly, if neither preferred nor agreed working hours on average change, the group size of unconstrained or under- and overemployed mothers should remain stable.

TABLE 3.1: Child care institutions by providers in Germany

	Total	*of which* Profit (%) organization	Non-profit (%) organization
2010	1,386	164 (11.83)	1,013 (73.09)
2011	1,486	184 (12.38)	1,061 (71.40)
2012	1,631	181 (11.10)	1,185 (72.65)
2013	1,725	185 (10.72)	1,219 (70.67)
2014	1,962	230 (11.72)	1,289 (65.70)
2015	2,029	261 (12.86)	1,348 (66.44)

Notes: Remaining institutions have a public background.
Source: Federal Statistical Office (2010*b*, 2011*c*, 2012*c*, 2013*b*, 2014*c*, 2015*c*). Cut-off date is March 1st.

3.3 Institutional background, estimation strategy, data and descriptive findings

3.3.1 Institutional background and estimation strategy

Institutional background

The German system of child care has several particularities ranging from strong regional variation to the different providers of child care (Kreyenfeld and Hank, 2000). Spatial differences are not only defined between urban and rural areas, but also between the former GDR and the West German states. Still in 2016, child care coverage amounts to 51.8 percent in East Germany in comparison with 28.1 percent in West Germany (Federal Statistical Office, 2016). Child care is usually provided by the communities of which there are more than 11,000 resulting in huge differences not only considering the price but also the availability of child care. A private market is not well-developed as quality regulations and hence market entry are related to high costs. The share of private institutions with a pure profit background amounts to about 11 to 13 percent over the last years (compare Table 3.1). However, there is a variety of non-profit organizations, often with a religious background, that receive public subsidies. About two thirds of all institutions belong to this category.

The expansion of early child care

The expansion of early child care started in 2005 when the German government decided on supplying 230,000 additional child care slots by 2010 (*Tagesbetreuungsausbaugesetz*). Two years later the objective was reinforced by targeting a coverage rate of 35 percent by 2013 (*Krippengipfel*). In 2008, the government decided on a legal claim for a child care slot for children aged one to three years from August 2013 onwards embedded in a law supporting the child's development (*KiFöG*).[1] In line with

[1] The KiFöG came in force in December 2008. Five years later, from August 2013 onwards, the legal claim guaranteed child care provided by a facility or childminder for children aged one to three (§24 SGB VIII). Children younger than one year are also eligible if their parents are employed.

the legal claim for a kindergarten slot introduced in 1996 (children older than two years) the law focuses firstly on the child's education and not on parental employment. The supply of child care is organized on the community level and subsidized by the federal state. Moreover, the federation supports the child care expansion financially. Until 2014, the federation has spent 5.4 billion Euro for improving child care supply and engaged for annual 845 million Euro beginning in 2015 (BMFSFJ, 2015). The allocation of child care on the community level results in strong regional variation that is strengthened by huge disparities between West and East German federal states. In the former German Democratic Republic the education of children was considered to be a public issue translating in a high share of children institutionally cared for until today. In 2011, the coverage rate of children aged up to three years old in subsidized care amounted to 49 percent in East Germany compared to only 20 percent in the rest of the country (Federal Statistical Office, 2011*b*). The reform changed the availability of child care slots dramatically. In 2015, 28.2 percent of children living in West-Germany and 51.9 percent in East Germany were in subsidized care (Federal Statistical Office, 2015*b*).

Although the legal claim was announced five years before it came into force, a shortage of 80,000 to 100,000 slots was predicted in July 2013 for the next month which suggests an almost full take up ratio. In general, the provision of early child care orients on the existing supply of child care slots and not on the actual needs (Kreyenfeld and Hank, 2000; BMFSFJ, 2015). While communities take population growth for the planning process into account, authorities mainly neglect any other factors determining the demand for child care. Table 3.2 shows the take up ratio of child care for several federal states for which official statistics are available. By March 1st, 2013, take up ratios are close to unity in most states. After the introduction of the legal claim in 2014, the ratio gets less tight indicating that the scarcity of child care slots is less severe. Note however, that regional variation on the community level is still high and that in many agglomerated areas child care slots continue being undersupplied.

Home care allowances (HCA)

The reforms of August 2013 included also the introduction of home care allowances (HCA) that were available for children between 15 and 36 months old born after August 2012 and who are not using subsidized child care. Younger children were also eligible if parental leave benefits had exhausted. The subsidy amounted to monthly 100 Euro (150 Euro per month from August 2014 onwards) irrespective of the parents' employment status or income. However an upper bound was set to an annual income of more than 500,000 Euro (married) or 250,000 Euro (singles). Opponents of the allowances feared that they would reinforce traditional employment patterns among couples as they encouraged families to not use subsidized child care. In July 2015, the home care allowances were declared unconstitutional while they normally expired for children already receiving the subsidy.

TABLE 3.2: Take up ratio of child care

	Institution for children aged ... years	2013	2014
Baden-Wuerttemberg	0-3	0.942	0.879
Bavaria	0-3	0.977	0.872
Hamburg	all age groups	0.849	0.802
Hesse	0-3	0.939	0.840
Mecklenburg-Vorpommern	0-3	0.968	0.983
Lower Saxony	0-3	0.895	0.864
North Rhine-Westphalia	0-3	0.946	0.876
Saarland	0-3	0.930	0.882
Saxony-Anhalt	all age groups	0.881	0.880

Notes: The take up rate is defined as actual take up divided by authorized slots.
Source: Own calculations based on the Statistical reports of the Statistical Offices of the Federal States (Statistical Office of Baden-Wuerttemberg, 2013, 2014; Statistical Office of Bavaria, 2013, 2014; Statistical Office of Hamburg and Schleswig-Holstein, 2013, 2014; Statistical Office of Hesse, 2013, 2014; Statistical Office of Mecklenburg-Vorpommern, 2013, 2014; Statistical Office of Lower Saxony, 2013, 2014; Statistical Office of North Rhine-Westphalia, 2013, 2014; Statistical Office of Saarland, 2013, 2014; Statistical Office of Saxony-Anhalt, 2013, 2014). Cut-off date is March 1st.

Although the receipt of these allowances is connected to not using subsidized child care, eligibility criteria for the HCA and subsidized care are not fully opposed to each other. Hence, children could be eligible for child care but not for the allowances in case parental benefits are not completely made use of, i.e., twelve up to 14 months after birth. Furthermore, there is also a small amount (8.4 percent) of eligible families neither requiring subsidies in form of the allowances nor in form of child care possibly because they are unaware or not in need (Alt et al., 2015).

The estimation strategy does not allow to disentangle the reform effect into the impact of the legal claim for a child care slot and the HCA. Theoretically, the allowances are expected to counteract, as they increase the opportunity costs of using a public child care slot and thus, mothers' financial incentive to stay at home is higher. Hence, the resulting estimates are expected to give a lower bound for the effect of the expansion of subsidized child care. Moreover, there are two reasons why I expect negligible effects of the HCA on maternal employment such that the resulting effects are more likely to be solely caused by the child care expansion. Firstly, Gathmann and Sass (2018) show that the introduction of similar HCA in the German federal state Thuringia in 2006 has small and insignificant effects on maternal labor supply. Secondly, the HCA are a potential confounder for the common trend assumption between low- and high-intensity districts. In particular, the HCA violate the assumption if they differently shift the potential employment trend. A possible sensitivity analysis may include to test if the take up of these HCA differs between low- and high-intensity districts. I find that the standardized mean difference[2] of the number of received public subsidies[3] is very low. Hence, I argue that the HCA are

[2]The standardized mean difference is defined as the mean difference divided by the square root of the average variance (see Rubin, 2001).

[3]The data does not contain information on the receipt of HCA. The number of received public subsidies is the closest measure to the HCA.

of minor importance as potential confounder.[4]

Methodological approach

The child care reform of 2013 serves as a quasi-experiment I exploit for DiD estimation. Besides the temporal variation, the expansion of subsidized child care has a spatial dimension that is used to define the treatment and control group. Following the approach of Bauernschuster et al. (2016), Felfe et al. (2015) and Havnes and Mogstad (2011), districts are split at the fourth and sixth percentile of the increase in the child care coverage rate for children aged up to three years old. Hence, treatment definition includes not a change from having no to having child care, but a change from a lower to a higher coverage rate. Furthermore, the resulting effect is an intention-to-treat effect as treatment definition does not inform about actual take up of a child care slot. As from 2005 onwards the Microcensus does not provide information on the attendance of a child care institution, it is not possible to relate the resulting estimates to actual child care take up. However, the resulting estimates clearly state the sign of the reform's impact. One might additionally consider estimating the reform's impact on the coverage rate itself. Any reference to such an analysis on the district level is not meaningful as maternal employment is measured on the individual level.

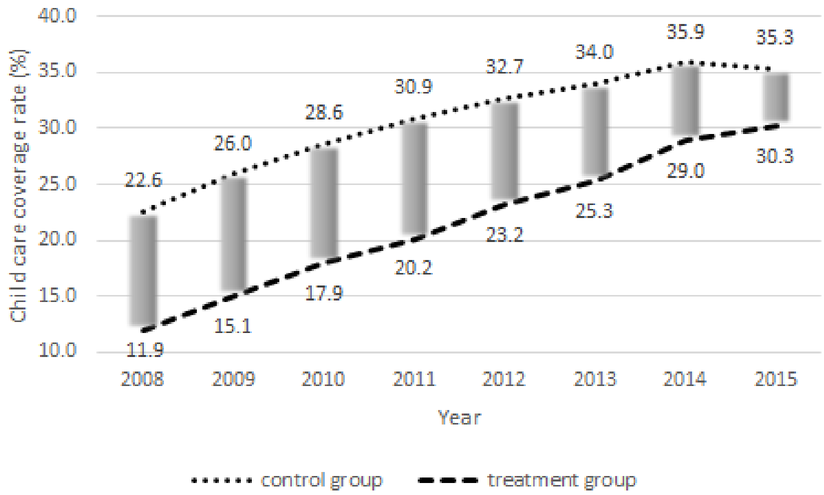
Alternatively, the implementation of the reform using a cut-off date would allow for a regression discontinuity design. A major advantage of DiD estimation, however, is the possibility to take seasonal effects into account which is especially relevant in the given application. Early child care and kindergarten attendance often cannot start at any point in time but orients on the beginning of the school year in August or September. As older children have better chances for a child care slot, mothers with children born shortly before the cut-off date are more likely to take up a job when the school year starts. Empirical studies evaluating German family policies like parental benefit reforms prefer DiD estimation for the same reason albeit the presence of a cut-off date such that cohort effects can be ruled out (Cygan-Rehm, 2016; Cygan-Rehm et al., 2018; Schönberg and Ludsteck, 2014). Furthermore, comparing groups only before and after the cut-off date would not allow for eliminating the impact of the home care allowances for which solely mothers with children born later than August 2012 are eligible. The treatment definition used in this article makes both treated and control group eligible such that the overall effect may only contain the effect of the child care expansion (see also the discussion in the previous section on the home care allowances).

As the reform took place in August 2013, the pre-reform period is measured in 2011 to rule out any anticipation effects.[5] Although the expansion of subsidized child

[4] The number of public subsidies slightly increase for both treated and control group over time. I do not include it as a covariate in the propensity score as the approach of Abadie (2005) used in the article does not allow to use time-variant covariates.

[5] The issue of anticipation is discussed as Assumption 2.

FIGURE 3.1: Child care coverage rates (%) in control and treated districts



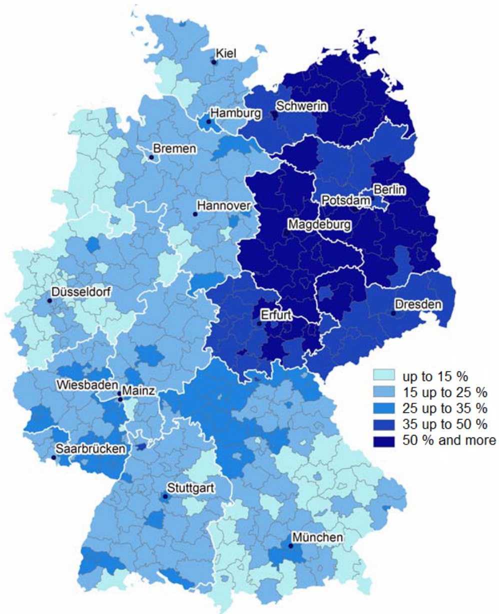
Notes: The child care coverage rate measures the number of children up to three years old in subsidized care in relation to all children in the respective birth cohort.
Source: Own calculations based on numbers from the Federal Statistical Office (2008, 2009, 2010*a*, 2011*b*, 2012*b*, 2013*a*, 2014*b*, 2015*b*).

care has started earlier, the largest increase in child care slots can be observed in the year the legal claim came into force (BMFSFJ, 2015) which additionally supports the use of the chosen survey years in contrast to previous years. From 2015 onwards the increase of the child care coverage rate is significantly smaller. Hence, I set this year as the post-reform period. The sensitivity analysis will provide similar results for the year 2014 as post-reform period. The treatment group comprises mothers whose youngest child is up to three years old and who live in a district in which the coverage rate increased by more than the sixth percentile (8.0 percentage points) between 2011 and 2015.[6] Mothers of children up to three years old living in districts with a lower increase of the coverage rate than the fourth percentile (6.5 percentage points) within these years belong to the control group. Districts within this interval and those undergoing a territorial reform within the considered time span are dropped from the sample resulting in a sample size of 317 districts.[7]
Figure 3.1 shows how the child care coverage rates evolve in control and treated districts. Although the share of institutionally cared for children is higher in low-intensity districts, the lines are almost parallel until the reform has become effective in 2013. From 2014 onwards the difference gets smaller for the first time.
The regional differences can be seen in Figure 3.2 which depicts descriptive statistics of the child care coverage rates on the district level in 2011. It shows that child care

[6]Hence, the pre-(post-)reform period includes mothers with children born between February 2008 (2012) and December 2011 (2015).

[7]Figure B.1 in the Appendix depicts the distribution of the growth of the child care coverage rate between 2011 and 2015. The identification of treatment and control group would be questionable in case of intense concentration around the separation. I find that the distribution is similar to the normal distribution and conclude that the identification strategy does not impose major problems.

FIGURE 3.2: Child care coverage rates for under three-year-olds across districts



Notes: The child care coverage rate measures the number of children up to three years old in subsidized care in relation to all children in the respective birth cohort.
Source: Federal Statistical Office (2011*b*). Cut-off date is March 1st, 2011.

coverage rates are the highest in East Germany while the lowest can be found in the southern and west-northern states.
Moreover, Table 3.3 indicates how the treated districts are spread over the federal states. The majority of northern and western districts belong to the treated group for which the coverage rate increased by more than 8.0 percentage points. In southern states the distinction is less obvious while most districts in East Germany belong to the control group for whom the coverage rate increased to a lesser extent. One may be concerned that most districts of the former GDR belong to the control group. However, a robustness check that drops East German districts will provide similar results compared to the baseline estimates.

Average effects

The idea of the DiD estimator is to compare average outcomes of a group affected by a reform with unaffected individuals before and after the treatment becomes effective. Under the assumptions of 1) parallel trends of control and treated group in the absence of the reform, 2) the absence of anticipation effects and 3) the stable unit treatment value assumption (SUTVA), the average treatment effect on the treated (ATET) can be identified. The assumptions are discussed in the following.

TABLE 3.3: Number of districts by group membership and federal states

Federal state	Control group	Treatment group
West Germany:		
Schleswig-Holstein	0	12
Hamburg	0	1
Lower Saxony	6	31
Bremen	0	1
North Rhine-Westphalia	1	47
Hesse	7	10
Rhineland-Palatinate	21	8
Baden-Wuerttemberg	20	11
Bavaria	50	25
Saarland	1	2
East Germany:		
Berlin	1	0
Brandenburg	13	3
Mecklenburg-Vorpommern	2	0
Saxony	6	3
Saxony-Anhalt	14	0
Thuringia	17	4

Source: Own calculations based on numbers of the Federal Statistical Office (2011*b*, 2015*b*) from 317 districts.

Assumption 1: Parallel trends

Assumption 1) can be expressed as

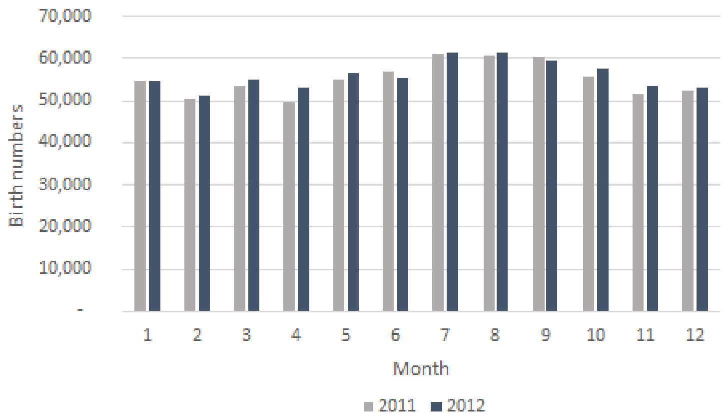
$$\begin{aligned}&\mathbb{E}\left[Y^0(1)|D=1,X\right]-\mathbb{E}\left[Y^0(0)|D=1,X\right]\\=&\mathbb{E}\left[Y^0(1)|D=0,X\right]-\mathbb{E}\left[Y^0(0)|D=0,X\right]\end{aligned}$$

where $Y^0(t)$ denotes the potential outcome in the absence of the treatment at time $T=t$ where $T=0$ is the pre-reform period and $T=1$ the post-reform period. $Y^1(t)$ is its counterpart under the reform. D is the binary treatment status and X depicts some covariates. Controlling for a large set of covariates makes the assumption of parallel trends more likely. I include covariates concerning the mother herself, the household she lives in and also regional dummies to control for the economic background (compare Section 3.3.2). Beyond that, I will run a placebo test by postponing the timing of the reform to 2011. While this kind of sensitivity analysis cannot directly test the common trend assumption, it may give suggestive evidence that it is not violated.

Assumption 2: Absence of anticipation

As the reform was already announced in 2008, anticipation might be relevant in two different forms. Mothers might have tried to postpone firstly, the date of conception or secondly, the date of birth to be eligible for the new regulations (births from August 2012 onwards). Figure 3.3 depicts official birth numbers from the relevant cohort 2012 in comparison with the cohort 2011 and does not show an irregular rise

FIGURE 3.3: Monthly birth numbers



Source: Own representation with numbers from the Official birth registers of the Federal Statistical Office (2011*a*, 2012*a*).

in August 2012. Hence, selection into treatment in the form of anticipation should play a minor role. Additionally, I only use pre-reform observations from 2011 (potential births between February 2008 and December 2011). This definition makes it less plausible that mothers desiring to have a child try to postpone conception longer than half a year such that the subsample of pre-reform mothers would have been selective.

Assumption 3: SUTVA

As further assumption SUTVA rules out interactions between groups. The assumption implies that individuals should not change between groups which might in particular be relevant for families moving from a control district to a treated district or vice versa. Due to the repeated cross sections, I cannot completely exclude these individuals, but I can control for families having moved within the last twelve months. The estimates would also be biased in case of other reforms taking place during the observational period. A major reform on parental leave already came into force in January 2007, incentivizing mothers to return to work at expiration of parental benefits (Bergemann and Riphahn, 2010, 2015; Kluve and Tamm, 2013; Kluve and Schmitz, 2018). However, the regulations were changed in July 2015 to make part-time work during benefit receipt more attractive (*ElterngeldPlus*: see next chapter). Findings suggest that dropping mothers of less than one-year-olds, who are affected by the reform, will turn out to be robust compared to the baseline results.

Under Assumptions 1) to 3) the average treatment effect on the treated (ATET) is identified as

$$\begin{aligned} ATET &= \mathbb{E}\left[Y^1(1) - Y^0(1)|D = 1\right] \\ &= \mathbb{E}\left[\mathbb{E}[Y^1(1) - Y^0(1)|D = 1, X]|D = 1\right] \\ &= \mathbb{E}\left[\mathbb{E}[Y(1) - Y(0)|D = 1, X] - \mathbb{E}[Y(1) - Y(0)|D = 0, X]|D = 1\right] \end{aligned}$$

which implies an outcome model that is usually estimated using OLS. Alternatively, Abadie (2005) shows that the ATET is also identified as

$$ATET = \mathbb{E}\left[\frac{P(D = 1|X)}{P(D = 1)}\rho_0 Y\right]$$

where

$$\rho_0 = \frac{T - \lambda}{\lambda(1 - \lambda)} \frac{D - P(D = 1|X)}{P(D = 1|X)P(D = 0|X)}$$

and λ being the share of post-treatment observations (see Abadie, 2005: for details). This implies a two-step estimation procedure. In a first step the propensity score $P(D = 1|X)$ is estimated by logistic regression. The second step gives the weighted non-parametric mean differences.

This approach has three main advantages. Firstly, it does not require a functional form assumption in the second stage and allows for flexibility which is especially useful for binary outcomes. Linear probability models usually used for parametric DiD estimation cannot satisfy the scale of such outcomes while nonlinear models based on the standard common trend assumption lead to inconsistent estimates (Lechner, 2011). The second advantage concerns the common support between control and treatment group. If an observational unit does not have common support within the other group, it can be dropped leading to higher comparability between treated and control group - a feature that is usually neglected in outcome based models. Finally, the specific form of the estimator allows to infer heterogenous effects (to be discussed in the next section).

Weighting temporal differences in the outcome is additionally relevant, as the reform not only included the expansion of subsidized child care, but also the introduction of home care allowances. Mothers applying for the allowances are supposed to be similar in observed characteristics. Both treated and control group can apply for these benefits and thus, if their outcome dynamics are the same in presence of the home care allowances *given their observed characteristics*, the estimated effect only measures the effect of the child care expansion. Therefore, I rely on defining treatment status based on the increase in child care coverage instead of using mothers of older children as control group (e.g., Bauernschuster and Schlotter, 2015). One might furthermore argue that control districts for which child care increased by a lower amount are characterized by a larger increase in receipt of home care allowances. However, as already mentioned in Section 3.3.1, I find no systematic differences in the take up of public subsidies between high- and low-intensity districts.

Heterogenous effects

To target particular groups, policymakers are often not only interested in average effects for the whole population, but also in an policy's impact for these groups. Hence, previous studies estimate effects for specific subgroups (e.g., Cascio, 2009; Havnes and Mogstad, 2011) - a procedure suffering from the multiple testing problem. The issue aggravates the more heterogeneities are investigated. Abadie (2005) proposes a least squares approximation for the conditional effect

$$\mathbb{E}\left[Y^1(1) - Y^0(1)|D = 1, Z\right]$$

given by $g(Z;\gamma)$ where $Z\epsilon X$:

$$\gamma_0 = argmin_{\gamma\epsilon\Gamma}\mathbb{E}\left[P(D = 1|X)\{\rho_0 Y - g(Z;\gamma)\}^2\right].$$

γ_0 directly indicates how the average effect varies over Z, and joint ordinary least squares inference is given. To my knowledge, this is the first article providing an application for estimating conditional effects following the proposition of Abadie (2005).

3.3.2 Data and descriptive findings

The data is from the German Microcensus,[8] a one percent representative sample of German households. The repeated cross-sections conducted by the Federal Statistical Office contain annual information on the family background, employment and other individual-specific characteristics. A main advantage of the Microcensus is the detailed information on the family composition. Hence, a child's and partner's characteristics can be connected with the observational unit of interest (mothers whose youngest child is aged up to three years old). I restrict the sample to mothers who are between 18 and 45 years old and who live in a private household which corresponds to the main place of residence.

A further particularity of the Microcensus is the availability of individual working hour preferences on top of agreed working hours. In contrast to other surveys like the German Socio-economic Panel (GSOEP) the question on working hours in the Microcensus is filtered. This means that, before stating the amount of preferred working hours, the individual is asked if he/she wants to increase or decrease the agreed weekly working hours conditional on an earnings adjustment[9] (for a methodological comparison of survey data on working hour preferences see Holst and Bringmann, 2016). Thus, there is also a measure for under- (the wish for an increase

[8]For the baseline specification the analysis uses (Microcensus, 2011, 2015: on-site use).

[9]Information on the preference for an hour increase (decrease) is included since 2006 (2008).

of agreed hours) and overemployment (the preference for less weekly hours). Apart from an earnings adjustment, respondents are not supposed to internalize any circumstances preventing them from increasing agreed hours, as follow-up questions explicitly ask for the main reason for not being able to work more hours within the next two weeks. In contrast to the compulsory question on the wish for an hour increase, respondents are free to answer their wish for an hour decrease. Holst and Bringmann (2016) point out that the voluntary indication might imply the underrepresentation of overemployed. The analysis includes only respondents answering the related questions, but I generally expect it to be a minor problem for the subsample of young mothers.[10]

I link the Microcensus data with statistics on the regional child care coverage rate for children aged up to three years old from the German Federal Statistical Office on the district level (Federal Statistical Office, 2010*a*, 2011*b*, 2012*b*, 2013*a*, 2014*b*, 2015*b*). The child care coverage rate is measured on the cut-off date March 1st and includes children in subsidized care not additionally attending another care arrangement and children in other care arrangements apart from subsidized care. The final sample includes 11,640 mothers (of which 3,505 are currently employed) of children not older than three years.

The variables used for estimating the propensity score described in the previous section and their descriptive statistics are listed in Table 3.4: family and individual characteristics, but also information on the interview. These numbers result after trimming observations, i.e., dropping individuals with a propensity score close (< 0.05) to the minimum and maximum value (compare Imbens and Wooldridge, 2009). Trimming excludes 5,192 observations in the whole sample ($N = 348$ in the control group, $N = 4,844$ in the treated group) and 1,710 individuals of the employed sample ($N = 230$ in the control group, $N = 1480$ in the treated group).

A major threat to identification might stem from using repeated cross-sections instead of panel data as individuals could have selected into employment after the reform came effective. Hence, a balancing check looks at the covariate distribution over time. Additional to mean values and standard deviations, Table 3.4 gives the standardized mean difference defined as the mean difference over time divided by the square root of the average variance (see Rubin, 2001). It does not exceed the critical value of 0.25 defined as large suggesting that selection over time depicts a minor problem. The remaining columns show that differences between mothers in high- and low-intensity districts are not large. Not surprisingly, only regional characteristics diverge as treatment is defined upon German districts. Besides, note that the analysis includes federal states instead of a dummy for East Germany to better take regional differences into account.

Table 3.5 shows the means of the child care coverage rate and of the examined outcome variables, their standard deviations and mean differences between treated and

[10] In the group of high-intensity districts only two percent indicate overemployment before the reform.

TABLE 3.4: Descriptive statistics of control variables by group membership

Variable	Pre Mean	Pre sd	Post Mean	Post sd	Post-Pre Std. mean diff.	Control group Mean	Control group sd	Treated group Mean	Treated group sd	Treated-control group Std. mean diff.
Individual age	32.297	5.638	32.395	5.048	0.018	32.270	5.283	32.427	5.426	0.029
Age of youngest child	0.986	0.812	0.969	0.810	-0.020	0.980	0.810	0.975	0.811	-0.007
Number of children	1.943	1.028	1.857	0.993	-0.084	1.877	0.969	1.925	1.056	0.047
Migration background:										
None	0.851	0.357	0.835	0.371	-0.042	0.868	0.339	0.816	0.387	-0.142
From EU country	0.041	0.198	0.057	0.231	0.074	0.041	0.199	0.057	0.231	0.072
Not from EU country	0.109	0.311	0.108	0.310	-0.002	0.091	0.288	0.127	0.333	0.116
Quarter of interview:										
1	0.250	0.433	0.240	0.427	-0.022	0.251	0.434	0.238	0.426	-0.032
2	0.247	0.432	0.239	0.426	-0.020	0.243	0.429	0.244	0.429	0.003
3	0.247	0.431	0.246	0.431	-0.002	0.244	0.429	0.250	0.433	0.015
4	0.256	0.436	0.275	0.446	0.043	0.262	0.440	0.268	0.443	0.014
Interview part										
Head of household	0.726	0.446	0.683	0.465	-0.093	0.712	0.453	0.696	0.460	-0.035
Self-reported	0.189	0.392	0.202	0.402	0.033	0.185	0.388	0.208	0.406	0.058
No information	0.085	0.279	0.114	0.318	0.098	0.141	0.103	0.096	0.295	-0.023
Educational degree:										
Lower secondary school	0.254	0.436	0.225	0.418	-0.069	0.247	0.431	0.232	0.422	-0.033
Middle secondary school	0.353	0.478	0.356	0.479	0.007	0.373	0.484	0.335	0.472	-0.079
High school	0.393	0.488	0.419	0.493	0.053	0.381	0.486	0.433	0.495	0.106
Partner:										
No partner living in household	0.171	0.377	0.124	0.330	-0.132	0.153	0.360	0.142	0.349	-0.031
Activity:										
Inactive	0.047	0.212	0.047	0.211	-0.002	0.042	0.201	0.052	0.221	0.045
Active	0.782	0.413	0.829	0.377	0.119	0.805	0.397	0.806	0.395	0.004
Educational degree:										
Lower secondary school	0.260	0.439	0.243	0.429	-0.040	0.263	0.440	0.239	0.426	-0.056
Middle secondary school	0.225	0.417	0.240	0.427	0.036	0.239	0.426	0.225	0.418	-0.031
High school	0.344	0.475	0.393	0.488	0.101	0.345	0.476	0.394	0.489	0.100
Degree of urbanization:										
Urban	0.373	0.484	0.345	0.475	-0.059	0.266	0.442	0.460	0.498	0.412
Middle	0.459	0.498	0.406	0.491	-0.108	0.476	0.499	0.386	0.487	-0.182
Rural	0.168	0.374	0.249	0.433	0.202	0.259	0.438	0.154	0.361	-0.261
East Germany	0.145	0.352	0.121	0.326	-0.070	0.185	0.389	0.076	0.265	-0.329
N	5,847		5,793			6,052		5,588		

Notes: The sample includes 18 to 45 years old mothers of up to three-year-olds. Instead of a dummy for East Germany, the analysis includes federal states. The standardized mean difference (std. mean diff.) gives the mean difference divided by the square root of the average variance (Rubin, 2001).
Source: Own calculations based on data from the German Microcensus and the Federal Statistical Office (2011*b*, 2015*b*).

TABLE 3.5: Mean outcomes and coverage rate by group membership

Variable	Treated group before refom Mean	sd	*N*	Difference in means Treated-control group Before	After
Coverage rate %	20.16	(8.24)	158	-10.74***	-5.02***
Employed *of which:*	0.348	(0.476)	2,721	-0.042***	-0.009
Agreed hours	25.50	(13.66)	862	0.81	1.16**
Preferred hours	26.96	(13.71)	862	0.95	1.00*
Mismatch	1.46	(6.34)	862	0.14	-0.16
Full-time	0.348	(0.477)	862	0.006	0.011
Part-time	0.470	(0.499)	862	0.021	0.042*

Notes: The sample includes 18 to 45 years old mothers of up to three-year-olds. Agreed and preferred hours are measured on the weekly basis. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.
Source: Own calculations based on data from the German Microcensus and the Federal Statistical Office (2011*b*, 2015*b*).

control group before and after the reform. The average coverage rate shows that less than one quarter of children in high-intensity districts are in subsidized care before the reform came into force. More mothers in low-intensity districts use subsidized care before the reform (negative, statistically significant difference), but high-intensity districts catch up.[11]

As outcomes I examine the extensive and intensive margin, i.e., a dummy for employment, agreed and preferred working hours as well as their mismatch and a binary indicator for working full-time (more than 30 hours per week) or part-time (between 12 and up to 30 hours per week). The Federal Statistical Office measures employment according to the concept of the International Labour Organization (employment for at least one paid hour or self-employment in the week before the interview) which includes employees in maternity protection and parental leave. Hence, I rely on the concept of realized employment and exclude them. About one third of all mothers in high-intensity districts are currently employed with an average of 25.5 hours per week. They prefer to slightly work more, on average one hour per week. However, for the majority working hour preferences and agreed hours match (13.8 percent of treated mothers are underemployed and only two percent overemployed). About 35 percent of them hold a full-time position and almost one half works in part-time. The last two columns of Table 3.5 give the differences in means between treated and control group before and after the reform. Before the reform employment rates in high- and low-intensity districts differ significantly, but the difference vanishes after the reform. Concerning the intensive margin, one cannot detect any strong variation across groups and time for all measures. Only part-time jobs seem to have increased in high-intensity districts. Hence, descriptive findings suggest a positive link between the expansion of subsidized child care and the employment rate, but no or only a weak relation to the intensive margin.

[11]Note that these are aggregated numbers that cannot give information on actual take up of a child care slot on the individual level.

3.4 Estimation results

3.4.1 Main results

Table 3.6 shows the baseline estimation results for the whole sample and different sensitivity checks. Bootstrapped standard errors take the two-step nature of the procedure and clusters on the district level into account.

In general, districts with a large increase of the coverage rate experience a rise of both the employment rate and working hours compared to districts with a lower expansion of child care. The reform effect amounts to an increase of the employment rate of 5.7 percentage points. Agreed and preferred weekly hours increase by 5.1 and 5.3 (20 percent of the pre-reform mean) respectively. Interestingly, these findings suggest an almost equal adjustment of agreed and preferred hours such that the effect on the mismatch size is close to zero. Further estimation results shown in Table A.4 in the appendix demonstrate that the share of under- and overemployed mothers is not significantly affected. These findings imply that the effects on hours are not only driven by involuntarily underemployed mothers who adjust agreed to preferred working hours, but that both distributions change. They suggest (see Figures B.2 and B.3 of agreed and preferred working hours in the appendix) a shift from marginal employment (categorized as up to 12 hours per week) to part-time work (between 12 and up to 30 hours per week). One can also observe a decrease at the upper part of the hour distribution. However, it contributes less to the average effect due to a similar movement in the group of low-intensity districts. Hence, the overall positive effect on working hours is driven by a shift from marginal to part-time employment which also shows up in an unaffected share of full-time employed.

The remaining panels of Table 3.6 contain different robustness checks. Firstly, to investigate the common trend assumption I check whether the time trend before the reform is the same for districts with a high and smaller increase of the coverage rate. I test a specification by introducing a placebo reform with the pre-reform period being 2010 ($T = 0$) and the post-reform period 2011 ($T = 1$). The estimates are close to zero (Panel D). Hence, shortly before the reform treated and control group show a similar time trend.

The next specification (Panel B) uses the median of the increase of the coverage rate for redefining the treatment and control group. The effects are similar to the results in the main specification with the clearer cut. The same holds for changing the post reform year to 2014. While similar in size, the effects for the intensive margin are close to significance on conventional levels.

Other checks deviate from the baseline by changing the sample composition (Panel C). The reform demonstrates to have a similar, but stronger effect when using only

TABLE 3.6: Results of main estimation and sensitivity analysis - Average effects

	Employ-ment	Agreed hours	Preferred hours	Mismatch (hours)	Full time	Part time
Panel A: Baseline						
	0.057**	5.089**	5.303**	0.213	0.063	0.126**
	(0.028)	(2.382)	(2.580)	(0.790)	(0.048)	(0.063)
N	11,640	3,505	3,505	3,505	3,505	3,505
Panel B: Sample composition						
Median	0.069***	3.823**	4.130**	0.307	0.027	0.119**
division	(0.023)	(1.929)	(2.006)	(0.403)	(0.037)	(0.050)
N	16,203	5,113	5,113	5,113	5,113	5,113
$post = 2014$	0.057**	3.263	3.360	0.097	0.037	0.090
	(0.025)	(2.179)	(2.247)	(0.380)	(0.044)	(0.055)
N	15,919	5,142	5,142	5,142	5,142	5,142
Panel C: Sample composition						
West	0.066*	5.316*	6.562**	1.246	0.025	0.184**
Germany	(0.038)	(3.027)	(3.268)	(0.865)	(0.058)	(0.080)
N	10,618	3,196	3,196	3,196	3,196	3,196
Without under	0.110***	6.621***	6.087**	-0.534	0.097*	0.142**
1-year-olds	(0.039)	(2.460)	(2.564)	(0.647)	(0.051)	(0.068)
N	7,695	3,001	3,001	3,001	3,001	3,001
Without	0.052*	4.503*	4.687*	0.184	0.058	0.111*
childminders	(0.029)	(2.491)	(2.700)	(0.767)	(0.051)	(0.065)
N	11,438	3,441	3,441	3,441	3,441	3,441
Without families	0.066**	5.401**	6.186**	0.785	0.054	0.163**
having moved	(0.030)	(2.521)	(2.743)	(0.791)	(0.053)	(0.068)
N	10,330	3,177	3,177	3,177	3,177	3,177
Panel D: Common trend						
Placebo	-0.007	-0.687	0.127	0.814	-0.031	0.008
reform	(0.032)	(2.417)	(2.006)	(0.605)	(0.048)	(0.067)
N	11,307	3,638	3,638	3,638	3,638	3,638

Notes: Standard errors (in columns) are bootstrapped with 1,000 replications considering clusters on the district level. The sample includes 18 to 45 years old mothers of up to three-year-olds. Agreed and preferred hours are measured on the weekly basis. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.
Source: Own calculations based on data from the German Microcensus and the Federal Statistical Office (2011*b*, 2015*b*).

TABLE 3.7: Results of heterogeneity analysis - Effect variation

	Employ-ment	Agreed hours	Preferred hours	Mismatch (hours)	Full time	Part time
Heterogeneities						
Education (Reference: Lower secondary school)						
Middle secondary school	0.044	6.216	5.600	-0.616	0.161	-0.113
	(0.063)	(6.250)	(6.881)	(1.904)	(0.139)	(0.154)
High school	0.122*	10.457*	9.773	-0.684	0.182	0.116
	(0.069)	(6.051)	(6.673)	(1.979)	(0.130)	(0.159)
Number of children	-0.004	1.462	2.550	1.087	0.014	0.055
	(0.025)	(2.629)	(2.912)	(1.059)	(0.053)	(0.074)
Partner (Reference: No partner living in household)						
Partner living in household	0.052	2.641	-0.454	-3.095*	0.027	0.104
	(0.073)	(6.840)	(7.343)	(1.834	(0.140)	(0.160)
N	11,640	3,505	3,505	3,505	3,505	3,505

Notes: Standard errors (in columns) are bootstrapped with 1,000 replications considering clusters on the district level. The sample includes 18 to 45 years old mothers of up to three-year-olds. Agreed and preferred hours are measured on the weekly basis. The estimates give the difference to the reference group for categorical variables or to a one-unit increase in case of continuous variables. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.
Source: Own calculations based on data from the German Microcensus and the Federal Statistical Office (2011*b*, 2015*b*).

West German districts.[12] Employment of mothers living in high-intensity West German districts rises by 6.6 percentage points which is mainly driven by part-time employment. Interestingly, their preferred working hours increase slightly more compared to agreed working hours. These findings show that the overall effect for both East and West Germany turns out to be robust considering any systematic differences between districts of the former GDR and West German districts. Thus, including East Germany in the baseline analysis is not problematic. Dropping mothers of children younger than one year old leads to a slightly larger effect for all outcomes. In particular, full-time employment is positively affected for mothers whose children are older than one year. Moreover, the results show that the parental leave reform of 2015 that affected mothers of less than one-year-olds is not supposed to drive the results. As mothers working in a child care facility might be differently affected by the reform, they are excluded in another specification which only slightly changes the estimates. The same holds when checking for selective migration by excluding those having changed their place of residence within the last twelve months.

3.4.2 Heterogenous effects

Table 3.7 indicates how the effects vary over different subgroups as estimated in Abadie (2005). Note that the estimates give the difference to the reference group for categorical variables or to a one-unit increase in case of continuous variables. E.g., mothers with high school degree show a by twelve percentage points higher employment effect compared to mothers with a degree from lower secondary school.

[12] To estimate the effect for the West German subsample the analysis defines treatment solely based on the quantile increase of the coverage rate in these districts and does not include East Germany.

The impact on the intensive margin is as well higher for better educated women, but the estimates are characterized by a high variance. These findings are in line with Havnes and Mogstad (2011) who also find larger effects for better educated mothers. However, this difference is weaker pronounced as the general reform effect also turns out to be smaller. One explanation of this result could be that external child care costs continue to be too high for mothers with lower educational degree.
While the average effect does not vary for the number of children, further interesting findings concern the presence of a partner. Although the estimates in general do not support deviating adjustment mechanisms for agreed and preferred working hours, cohabiting mothers show a significant higher rise of agreed hours. As the rate of underemployment also decreases for this subgroup, the reform was especially successful for families with a more traditional employment pattern by adapting agreed hours to the desired level.

3.5 Discussion and conclusion

This paper provides empirical evidence for the causal effect of subsidizing early child care on maternal labor market outcomes. It exploits the staggered expansion of early child care provision in Germany culminating in a legal claim for a child care slot introduced in 2013. The presented semi-parametrically estimated intention-to-treat effects suggest a strong impact of 5.7 percentage points on the maternal employment rate and of five on agreed and preferred weekly working hours. Besides, the share of full-time employed women does not significantly change in response to the reform which might result from limited provision of full-time child care slots or the parental preference for part-time care. Although the share of realized full-time slots (defined as more than seven hours per day) almost doubled from 2011 to 2015 in high-intensity districts, only one out of ten children attends full-time care in post-reform years. However, these numbers cannot definitely answer which of the two channels, lack of provision or parental preferences, prevails as they do not give information on the supply of full-time slots.
The main findings are in general in line with previous results for Germany. Bauernschuster and Schlotter (2015) estimate intention-to-treat effects for the eligibility to kindergarten in the range of five to eight percentage points for employment and of 2.5 for weekly hours. Fendel and Jochimsen (2017) find an increase of maternal employment of eight percentage points for the overall reform, i.e., the legal claim for a child care slot and the introduction of the home care allowances. Hence, these findings for Germany turn out to be robust compared to other countries with low maternal labor market participation (Baker et al., 2008; Berlinski and Galiani, 2007; Berlinski et al., 2011; Lefebvre and Merrigan, 2008; Nollenberger and Rodríguez-Planas, 2011; Schlosser, 2005). Another crucial finding concerns the adjustment of agreed and preferred working hours. Both measures change, but in contrast to Reynolds and Johnson (2012) this article finds that agreed and preferred working

hours adjust on average in line with each other. Furthermore, the average effect on the share of under- and overemployed mothers is not significant. These results imply that also the size of the mismatch remains close to zero and that the results are not only driven by involuntarily underemployed mothers adjusting agreed to preferred working hours. On the contrary, the availability of low cost child care has the potential to increase working hour preferences also for other groups represented in an overall shift of the distributions of agreed and preferred working hours. Mothers changing from marginal to part-time work characterize this shift.

Another contribution of this article is the provision of conditional average effects with two interesting findings. Firstly, mothers with high school degree show large positive responses in contrast to women with lower educational degree which can be explained by too high external child care costs for the latter group. Hence, a possible implication is to organize parental contributions for child care slots income-related as many communities already have implemented. Secondly, cohabiting mothers who might have previously provided additional earnings to a partner's main income show a higher adjustment of agreed than of preferred working hours which is reflected in a lower share of underemployed. This finding extends the declarations for ambiguous results in different countries and underlines the possibility for deviating adjustments of preferred and agreed working hours. The effect size can also depend on the degree to which mothers are not satisfied with their actual or agreed working hours. Hence, underlying working hour preferences are relevant to consider when assessing the potential success of a reform targeting female labor supply. These results for cohabiting mothers are supported by a related study for Norway. Andresen and Havnes (2019) find that especially cohabiting mothers respond to child care attendance of two-year-olds by increasing full-time employment in the context of the majority (63 percent) holding a part-time contract before the reform.

Although the reform's overall effect seem to be positive, questions remain. Especially the group of mothers with lower educational degree and singles show small responses. Hence, further research might focus on the channels that drive these results.

Chapter 4

Paid parental leave and maternal reemployment: Do part-time subsidies help or harm?

Joint with Michael Zimmert[c]

Abstract: Employment subsidies can incentivize mothers to shorten employment interruptions after childbirth. We examine a German parental leave reform promoting an early return to work in part-time. Exploiting the exogenous variation in the benefit entitlement length defined by the child's birthday, we apply machine-learning augmented semi-parametric difference-in-difference estimation using administrative data. The reform yields positive average employment effects mainly driven by part-time employment as our dynamic optimization model for mothers on parental leave suggests. Conditional effects show that the policy creates heterogenous incentives depending on the opportunity costs of working part-time.

Keywords: causal machine learning, effect heterogeneity, maternal labor supply, parental leave, Germany

Acknowledgements: The authors are grateful to Michael Lechner, Gesine Stephan and Enzo Weber. The authors benefited also from comments from participants of the joint Econometric Seminar of the Institute for Employment Research (IAB) and the University of Regensburg as well as of the Jour fixe of the Graduate Programme of the IAB and the School of Business and Economics of the University of Erlangen-Nuremberg (FAU).

[c]University of St. Gallen; Dufourstrasse 50, 9000 St Gallen, Switzerland

4.1 Introduction

Motherhood and related employment interruptions are still one of the main causes for the gender wage gap and different labor market prospects for women (Lundborg et al., 2017). While gender roles converge and women catch up in terms of education and job choice (Goldin, 2014), women face statistical discrimination even ex ante in expectation of potential motherhood and its related costs to the firm (Jessen et al., 2019). As long career breaks imply hiring and training costs for a new candidate, employers might anticipate motherhood in their recruiting process. Even if young women find a suitable job, starting a family will fundamentally change their choice set and potentially lead to a reduction of working time with the consequence of worsened employment prospects and lower wage expectations (Goldin, 2014). Hence, labor market interventions promoting an earlier return to employment may be regarded as a suitable tool to cushion these adverse effects. In particular, this article analyzes if subsidized part-time work after child birth shortens employment breaks and affects the working time pattern of mothers sustainably.

There is a growing trend for extending the provision of paid parental leave over the last years (compare Dahl et al., 2016). Especially European countries nowadays offer generous parental leave regulations. In the United States, however, there is no nationwide paid leave period despite some states notably California agreed on a paid protection period (Rossin-Slater et al., 2013). In line with these trends, literature on the effectiveness of maternity protection and (un)paid parental leave policies broadens and analyses use the exogenous variation induced by reforms to investigate maternal labor market outcomes. While previous studies find that short unpaid protection periods like the Federal Maternal Legislation Act (FMLA) in the United States have small effects on maternal employment and wages (Waldfogel, 1999; Baum, 2003), results differ for longer potential parental leave durations. Many authors find that mothers delay their return to work for extended parental leave regulations (Baker and Milligan, 2008; Bergemann and Riphahn, 2015; Dahl et al., 2016; Lalive and Zweimüller, 2009; Kluve and Tamm, 2013; Kluve and Schmitz, 2018; Schönberg and Ludsteck, 2014). Less is known about the employment outcomes, especially working hours, after having been returned to work. This is especially important as length and timing of working hours are considered to be the "last chapter" (Goldin, 2014) for reducing or even closing the gender wage gap. Long working hours signal productivity and employees willing to work long hours are more likely to be promoted (Landers et al., 1996). In this context the question arises if maternity leave improves employment stability enabling mothers to return in less precarious jobs with better career prospects. There exist two articles analyzing working hours after mothers returned to work. Schönberg and Ludsteck (2014) show for Germany that several extensions in paid leave coverage between 1979 and 1992 lead to short-term reductions in full-time employment, but do not have any long-run effects. In contrast, Kluve and Schmitz (2018) find even long-lasting positive effects on full-time employment for mothers from the upper tercile of the income distribution which is the group the most affected by a German parental leave reform in 2007. We also examine Germany as a labor market on which the traditional division of paid and unpaid household work predominates. While the German government enacted several family reforms over the last decades encouraging external child care attendance and a more equal division of unpaid household work, a strict full-time/part-time division of father and mother persists (Wanger, 2015).

We contribute to the literature on the impact of maternal leave on employment in at least three different fields: 1) content-related by focusing on the working time pattern, i.e., the intensive employment margin, 2) theoretically by proposing an illustrative dynamic optimization problem for employees on parental leave and 3) methodologically by providing credible average and subgroup-specific effect estimation using machine learning algorithms and high-quality administrative data.

In detail, we examine the effect of subsidizing part-time on the maternal working time pattern right after the birth of a child. Mothers affected by a new law coming into force for births from July 2015 onwards are encouraged to combine income from part-time work and public subsidies. We develop a heuristic dynamic optimization problem that depicts this mechanism. As part-time work becomes more attractive relative to extending parental leave *and* to working full-time, the overall employment effects are unclear. Even if the effect on the extensive employment margin is positive, the policy might foster the so-called part-time trap. In particular, employees might be unable to increase their agreed working hours to a full-time job at a later point in time. These theoretical findings motivate to empirically assess the effects of an extended part-time subsidy.

The implementation of the reform enables to exploit exogenous variation in the entitlement length and benefit amount of parental leave affecting mothers with children born later than June 2015. We compare those treated mothers with women having children born shortly before the cut-off date. To account for seasonal effects resulting from different patterns for the start of the school year we use difference-in-differences estimation. As different factors such as local differences in the economy or personal characteristics of women may differently shift the employment trends, we include a large list of covariates from administrative data. In particular, we apply a recently proposed semi-parametric difference-in-differences (DiD) estimator (Sant'Anna and Zhao, 2018; Zimmert, 2018). It allows to include covariates in a data-driven way using state of the art machine learning algorithms. We argue that the inclusion of a large set of covariates makes our identifying assumptions more credible. Additionally, we avoid common problems in parametric DiD estimation like arbitrary functional form assumptions and misspecification errors. Moreover, as first shown in Abadie (2005), semi-parametric DiD estimation allows to infer heterogeneous effects that uncover for which subgroups the reform was effective. We give an identification result that implies a new estimator for heterogeneous treatment effects estimation in the DiD setting.

Our results show that women exposed to the reform have on average an about two percentage points higher probability to return to work within the first year which amounts to about 14 percent of the pre-reform level. Like the reform intended, this increase is mainly driven by part-time employment. However, these positive average effects do not continue after the child's first birthday. Although limited to a two-year perspective, these findings cast doubt upon sustainably strengthening female employment prospects. Besides, on average we cannot confirm the existence of a part-time trap for this short time horizon. To some extent, the heterogenous effects show a more refined pattern. Especially mothers with middle income are willing to take up the new part-time subsidy. In turn, mothers with higher income expectations might fear future income losses in case they accept a lower-paid part-time job. The article proceeds as follows. The next two sections describe the institutional setting and the dynamic optimization problem. Section 4.4 and 4.5 explain the estimation strategy and the exploited data before presenting the estimation results and sensitivity checks in Section 4.6. The article ends with a discussion of the results and a conclusion.

4.2 Institutional background

The German system of birth-related legal work interruptions distinguishes two different forms: maternity protection and parental leave. The first concept describes a period of six weeks before and eight weeks after child birth in which mothers are not allowed to work due to health risks. The latter wants to facilitate the employment continuity of parents and especially of mothers by defining a period up to which parents have the right to return to their previous employer. Table 4.1 gives an overview of the two most important parental leave (*Elterngeld* abbreviated *EG*) regulations over the last years. This article will focus on the regulations for births from July 2015 onwards (last column).

4.2.1 Regulations of parental benefits prior to the reform in 2015 (*Elterngeld EG*)

Former regulations (see second column in Table 4.1) for births from January 2007 onwards aimed at facilitating motherhood for working women and engaging fathers in child care. It standardized the maximum benefit receipt duration to 12 months with additional two months if both parents are on leave (so-called daddy months, see Tamm (2019) for their evaluation). Besides, a replacement rate λ of 65 percent (up to 100 percent for parents with low income) was introduced determining the *basic parental benefit amount* based on the average net monthly income measured during the twelve months before child birth denoted by $\bar{y}$. Previously none-working or low-income mothers receive a minimum of 300 Euro per month while the maximum was set to 1800 Euro per month. Part-time work as a share β of a full-time contract and

TABLE 4.1: Parental leave regulations over time

births	$\geq$ 01/2007: *Elterngeld (EG)*	$\geq$ 07/2015: *Elterngeld (EG)* or *ElterngeldPlus (EG+)*
maximum unpaid entitlement length	36 months	36 months
maximum paid entitlement length	12 months	*EG*: 12 months or *EG+*: 24 months
benefit amount Euro/month	*basic parental benefit amount*,: 65 % of average net monthly income measured during 12 months before birth 67 - 100 % if average net monthly income $<$ 1200 Euro minimum 300 maximum 1800	*EG*: full *basic parental benefit amount* *EG+*: up to 0.5 * *basic parental benefit amount*
means testing	no	no
employment during benefit receipt	$\leq$ 30 hours/week	$\leq$ 30 hours/week
benefit deduction in case of part-time work	yes	yes, but *EG+* may imply an equal total benefit amount (compare Table 4.2)
"daddy" months	2 additional months	*EG*: 2 additional months *EG+*: 4 additional months for each parent if both decide for EG+

Source: Own representation according to *Bundeselterngeld- und Elternzeitgesetz* (BEEG).

up to 30 hours per week is also possible, but reduces the benefit amount. For part-time working mothers the difference between former and current net income ($\bar{y} - y$) serves as reference value for the replacement rate τ that also amounts to between 65 and 100 percent (for a graphical representation relating prior with current income see Figure 4.1a):

$$\iota_{EG} = \begin{cases} = 300 \leq \lambda\bar{y} \leq 1800 \text{ if not working, paid for 12 months} \\ = 300 \leq \tau(\bar{y} - \beta y) \leq 1800 \text{ if part-time working, paid for 12 months} \\ = 0 \text{ else} \end{cases}$$

Kluve and Schmitz (2018) show that mothers with income from the upper tercile of the distribution benefit the most from these parental leave regulations having positive effects on full-time employment up to the child's fifth birthday. Moreover, Kluve and Schmitz (2014, 2018) highlight the formation of a social norm to return to work at the end of the maximum entitlement length of 12 months which is challenged by the new regulations coming into force in July 2015.

4.2.2 The reform in July 2015 (*ElterngeldPlus EG+*)

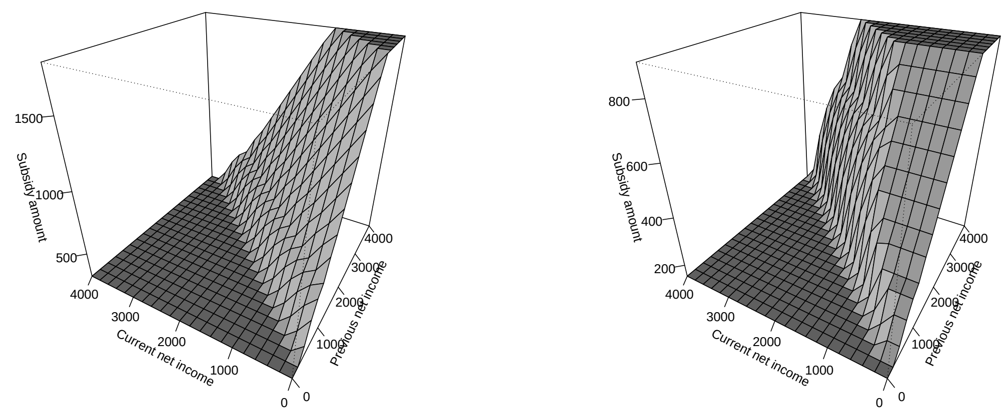
The new regulations coming into effect for births from July 2015 onwards double the maximum entitlement period to 24 months while receiving up to half of the basic benefit amount (compare Figure 4.1b):

$$\iota_{EG+} = \begin{cases} = 150 \leq \frac{\lambda\bar{y}}{2} \leq 900 \text{ if not working, paid for 24 months} \\ = 150 \leq \min(\tau(\bar{y} - \beta y), \frac{\lambda\bar{y}}{2}) \leq 900 \text{ if part-time working, paid for 24 months} \\ = 0 \text{ else} \end{cases}$$

Table 4.2 gives several examples for the calculation of the subsidy under the new regime. Besides, the model presented in Section 4.3 explains the reform mechanisms in detail. The regulations for births until July 2015 discourage mothers to return to work before parental benefits expire as current labor income is taken into account for the calculation of the subsidy. Official statistics show that the majority of female benefit recipients with children born in the third quarter 2015 chooses the *full basic amount* (81 percent) with an average benefit amount of monthly 757 Euro and in total 8,797 Euro (Federal Statistical Office, 2019). The remaining 19 percent decided for the second option (*EG+*) and received on average 492 Euro per month with a slightly higher total sum of 9,130 Euro compared to *EG*. Parents can also share the parental leave period. The two additional daddy months result in four extra months under the new regime. Parents choosing this option are eligible for another four months of benefit receipt resulting in 32 months all together. About ten percent of all male benefit recipients decided to be on leave for at least four months in the relevant birth cohort (Federal Statistical Office, 2019) which amounts to about three percent of all

FIGURE 4.1: Payment schemes before and after cut-off date

(a) Pre-reform (b) Post-reform



Notes: The pre-reform payment scheme is depicted by ι_{EG}, the post-reform payment scheme by ι_{EG+}. The graph gives the benefit amount in dependence from prior and current income.
Source: Own diagram.

births in the third quarter of 2015. Hence, we expect only small effects on paternal labor supply as a channel for maternal employment adjustments. Moreover, Tamm (2019) find that the daddy months established by the previous reform in 2007 do not significantly affect paternal involvement in child care and housework on weekdays for those currently on leave.[1] Hence, paternal leave can rather be considered as shared family time than a promotion of maternal employment.
Unaffected by both reforms, the *unpaid* maximum parental leave duration amounts to 36 months from child birth onwards, i.e., mothers have the right to return to their previous employer until the child's third birthday.

4.2.3 The expansion of subsidized child care

The lack of suitable child care slots may prevent mothers from returning to work. Recent parental leave changes are part of different family policies notably the child care expansion for under three-year-olds starting in 2005 and culminating in a legal claim for a child care slot from August 2013 onwards (see Chapter 3). Before the first parental leave reform in 2007, only 13.6 percent of children younger than three years old attend subsidized child care (Federal Statistical Office, 2006) increasing to 32.9 percent in 2015 (Federal Statistical Office, 2015*b*). Although also under one-year-olds have a claim for a child care slot if both parents are working, jobseeking or in education,[2] only 2.6 percent of this age group attend child care in 2015 (Federal

[1] In turn, Tamm (2019) and also Patnaik (2019) find that paternal leave can strengthen paternal involvement in child care and housework in the longer term, i.e., beyond the leave duration. Unfortunately, the data set does not allow to examine the policy from a comprehensive household context as information on paternal employment and subsidy receipt are unknown.

[2] This regulation is defined by §24 SGB VIII.

TABLE 4.2: Calculation of benefit amount *Elterngeld* and *ElterngeldPlus*

Example	Net monthly income before birth	after birth	Income difference	Parental benefit amount not working	working
I	2,000	0	2,000	$2,000 * 0.65 = 1,300$	-
EG				$12 * 1,300 = 15,600$	$12 * 1,300 = 15,600$
EG+				cap= $1,300/2 = 650$	$24 * 650 = 15,600$
II	2,000	1,200	800	$2,000 * 0.65 = 1,300$	$800 * 0.77 = 616$
EG				$12 * 1,300 = 15,600$	$12 * 616 = 7,392$
EG+				cap= $1,300/2 = 650$	$24 * 616 = 14,784$
III	2,000	500	1,500	$2,000 * 0.65 = 1300$	$1,500 * 0.65 = 975$
EG				$12 * 1,300 = 15,600$	$12 * 975 = 11,700$
EG+				cap= $1,300/2 = 650$	$24 * 650 = 15,600$
IV	2,000	2,000	0	$2,000 * 0.65 = 1300$	300
EG				$12 * 1,300 = 15,600$	$12 * 300 = 3,600$
EG+				cap= $1,300/2 = 650$	$24 * 150 = 3,600$

Notes: Income in Euro. The table gives several examples for the calculation of the benefit amount under the new regime. Mothers can optionally decide for the full basic benefit amount for a period of 12 months (*EG*) which amounts to 1,300 Euro per month or 15,600 Euro in total in the example. The total sum is not affected if the mother decides for half the amount (650 Euro per month) for the longer period of 24 months (15,600 Euro). However, if she has net earnings of 1,200 Euro per month (Example II), she will only receive additional 616 Euro for up to 24 months which is in total less than Example I. Under Example III with a monthly income of 500 Euro she can decide for 975 Euro for one year (in total 11,700 Euro) or 650 Euro for two years (in total 15,600 Euro). In case she has an equal income than before (Example IV) she will receive the minimum amount of 300 or 150 Euro respectively.
Source: Own representation based on BMFSFJ (2018).

Statistical Office, 2015*b*). These numbers tend to be higher in urban areas and in East Germany (4.1 percent) as a legacy of the former GDR. As official statistics lack information on the number of authorized child care slots, we exploit data from the survey FiD (DIW Berlin/SOEP, 2014: wave 2013) to explore the reasons for the low early child care coverage. Similar to official statistics, the survey provides a coverage rate for under one-year-olds of 3.1 percent. It shows that 88 percent of parents not making use of a child care slot consider their child to be too young while five percent indicate the lack of suitable child care slots. Hence, we conclude that institutional restrictions do not determine the low coverage rate but attitudes towards external child care.

4.3 Theoretical effects of part-time subsidies

4.3.1 Model set up

For the sake of illustration we set up a dynamic optimization problem according to the given institutional framework. Mothers in parental leave can generally choose between three options: staying in parental leave (pl), working full-time (f) or working part-time (p). After the end of the maximum parental leave duration in period T mothers can either return to the labour force in part-time or full-time or drop into unemployment (u) where they receive a fixed benefit amount b^u.[3] We neglect the option that mothers have a legal claim to return to their previous employer according

[3]For simplification we do not distinguish between unemployment and non-employment.

to contracted hours and wage until the child's third birthday. However, the model implies that if a mother wants to reduce her working time after child birth but stay with her previous employer, she has to renegotiate her working contract. This kind of simplification will not restrict the main model mechanisms.
Throughout our heuristic model we assume that a decision taken in any period determines the rest of the working life, i.e., mothers choosing part-time will stay in part-time. Even though this might be a strong simplification, it depicts at least partly the German labor market as a legal claim to increase working hours to a full-time job after having worked in part-time only became effective in 2019. We assume that a mother can decline a full-time job offer to work part-time and take only $\beta \times 100$ percent of the income offered. Since parental leave may also only represent a relatively small period compared to the following working life of young women, we approximate the value of unemployment, full- and part-time work by infinite series starting in $T+1$. In particular, let ρ denote the discount rate, y be the income from a full-time job offer in a certain period and l the constant value of leisure when working part-time. We then get the following value functions in $T+1$

$$V^u_{T+1} = b^u \frac{1+\rho}{\rho}, \quad V^f_{T+1} = y_{T+1}\frac{1+\rho}{\rho} \quad \text{and}$$
$$V^p_{T+1} = \beta y_{T+1}\frac{1+\rho}{\rho} + (1-\beta) l \frac{1+\rho}{\rho}. \tag{4.1}$$

During parental leave a mother gets compensation from two different sources.

1. In state pl a mother receives a fixed share λ of her previous labour market income $\bar{y}$ (full-time or part-time).

2. If a mother decides to work part-time, she receives a fixed share of the difference between her previous income and the income from part-time work $\tau \times (\bar{y} - \beta y_t)^+ = \max(0, \tau \times (\bar{y} - \beta y_t))$.

The decision problem of the mother in parental leave is whether to stay out of the labour force or to accept a part-time or full-time job offer, and it can be solved by dynamic optimization. We report the main results in the following sections and give model details in the Appendix C.

4.3.2 Reservation income

We define the income that makes the mother indifferent between working and not working as the reservation income of the extensive margin y^{*EXT} for any time period. Similarly, the reservation income of the intensive margin y^{*INT} is defined as the income that makes the mother indifferent between working part-time and full-time given that she has decided to work.

For the stationary environment after the maximum parental leave duration beginning in $T+1$ we find the reservation income of the extensive margin

$$y_{T+1}^{*EXT} = y_{T+2}^{*EXT} = ... = \begin{cases} b^u & \text{if} \quad y_{T+t} > l \\ \frac{b^u}{\beta} - \frac{1-\beta}{\beta} l & \text{if} \quad y_{T+t} < l \end{cases} \tag{4.2}$$

where the mother decides to work full-time whenever $y_{T+t} > l$ and part-time whenever $y_{T+t} < l$.[4] Hence, in the stationary setting the reservation income at the intensive margin is $y_{T+t}^{*INT} = l$.
Moreover, we can explicitly solve for the value function V_T^{pl} and iterating backwards will give an explicit solution for every V_{T-t}^{pl} in the model. In the non-stationary environment for any period $t \geq 0$ the reservation income will decline compared to the pre period until it reaches the stationary solution in $T+1$ as given in (4.2). We therefore derive an implicit solution for the reservation income in the non-stationary environment. In particular, we obtain

$$y_{T-t}^{*EXT} = \begin{cases} \frac{\rho}{1+\rho} V_{T-t}^{pl} & \text{if} \quad y > \frac{\bar{y}}{\beta} \quad \text{and} \quad y > l \\ \frac{1}{\beta}\left(\frac{\rho}{1+\rho} V_{T-t}^{pl} - (1-\beta) l\right) & \text{if} \quad y > \frac{\bar{y}}{\beta} \quad \text{and} \quad y < l \\ \frac{\rho}{1+\rho} V_{T-t}^{pl} & \text{if} \quad y < \frac{\bar{y}}{\beta} \quad \text{and} \\ & y > l + \frac{1}{1-\beta} D(\rho,t)\tau(\bar{y} - \beta y) \\ \frac{1}{\beta(1-\tau D(\rho,t))} * & \text{if} \quad y < \frac{\bar{y}}{\beta} \quad \text{and} \\ \left(\frac{\rho}{1+\rho} V_{T-t}^{pl} - (1-\beta) l - D(\rho,t)\tau\bar{y}\right) & y < l + \frac{1}{1-\beta} D(\rho,t)\tau(\bar{y} - \beta y) \end{cases} \tag{4.3}$$

with $D(\rho,t) = 1 - \left(\frac{1}{1+\rho}\right)^{t+1}$. The first two cases of (4.3) describe the situation when the mother is not eligible to the part-time subsidy because the offered income is much higher than the previous income and hence $\tau(\bar{y} - \beta y_{T-t})^+ = 0$. Cases 3 and 4 describe a situation when the mother becomes eligible to the part-time subsidy. While the reservation income for working full-time (cases 1 and 3) does not depend on the eligibility of the part-time subsidy, the reservation income for working part-time (cases 2 and 4) is lower when the mother is eligible. Also the decision whether to work full-time or part-time depends on whether the mother is eligible or not. Similarly to previous reasoning we find that

$$y_{T-t}^{*INT} = \begin{cases} l & \text{if} \quad y > \frac{\bar{y}}{\beta} \\ \frac{1}{1-\beta(1-D(\rho,t)\tau)} \left((1-\beta) l + D(\rho,t)\tau\bar{y}\right) & \text{if} \quad y < \frac{\bar{y}}{\beta}. \end{cases} \tag{4.4}$$

Again, the two cases discriminate the reservation income of the intensive margin depending on part-time subsidy eligibility. If not eligible, mothers are indifferent between working part-time and full-time such that they value an additional unit leisure equally to an additional unit of income. If eligible, the offered income has

[4] We get these results by (C.2) of Appendix C.

to equal the utility from leisure plus the time value of the subsidy in order to make mothers indifferent between the two options.

4.3.3 Implications of the reform

The reform changed two parameters simultaneously. First and foremost, it gives mothers the choice to double the maximum parental leave duration if working up to 30 hours per week. Secondly, it optionally halved the replacement rate in case of not working.[5]

Duration effects

We first of all notice that for any period $T-t$ in parental leave an increase in T can be modelled by an increase in t (the end of the maximum duration period is farer away). Since $V^f_{T-t} = V^f_{T+t}$, $V^p_{T-t} > V^p_{T+t}$ and $V^{pl}_{T-t-1} > V^{pl}_{T-t}$ for any $t \geq 0$, we have that $\frac{\partial V^{pl}_{T-t}}{\partial t} > 0$. These results and $\frac{\partial D(\rho,t)}{\partial t} > 0$ directly imply that the reservation income of the extensive margin increases for cases 1-3. For the fourth case we obtain

$$\frac{\partial y^{*EXT}_{T-t}}{\partial t} = \frac{1}{\beta(1-\tau D(\rho,t))}\left(\frac{\rho}{1+\rho}\underbrace{\frac{\partial V^{pl}_{T-t}}{\partial t}}_{>0} - \tau\underbrace{\frac{\partial D(\rho,t)}{\partial t}}_{>0}\underbrace{(\bar{y}-\beta y^{*EXT}_{T-t})}_{>0}\right) \tag{4.5}$$

and so the overall effect is ambiguous. It can be decomposed in the additional value of staying in parental leave and its forgone part-time subsidy.
If eligible to the part-time subsidy, we find for the reservation income at the intensive margin that

$$\frac{\partial y^{*INT}_{T-t}}{\partial t} = \frac{\tau}{1-\beta(1-D(\rho,t)\tau)}\underbrace{\frac{\partial D(\rho,t)}{\partial t}}_{>0}\underbrace{(\bar{y}-\beta y^{*INT}_{T-t})}_{>0} > 0. \tag{4.6}$$

For an increased duration the part-time subsidy becomes more valuable. So, the income that has to be offered in order to make the mother indifferent between working full- or part-time has to increase.

Decreased replacement rate in case of staying on parental leave

Since $\frac{\partial V^{pl}_{T-t}}{\partial \lambda} > 0$, the implications for the reservation income at the extensive margin for cases 1 to 4 of (4.3) follow in a straightforward manner.

[5]The reform also decreased the subsidy schedule τ for $\frac{\lambda\bar{y}}{2} \leq \tau(\bar{y}-\beta y)$ under very special circumstances. For the sake of readability we neglect this feature.

Total effect on reservation incomes

The preceding analysis shows that the total effect on the reservation income at the extensive margin can be summarized for cases 1 to 3 as follows:

$$dy^{*EXT}_{T-t} = \begin{cases} \frac{\rho}{1+\rho}\left(\frac{\partial V^{pl}_{T-t}}{\partial t}dt + \frac{\partial V^{pl}_{T-t}}{\partial \lambda}d\lambda\right) & \text{if} \quad y > \frac{\bar{y}}{\beta} \quad \text{and} \quad y > l \\ \frac{1}{\beta}\frac{\rho}{1+\rho}\left(\frac{\partial V^{pl}_{T-t}}{\partial t}dt + \frac{\partial V^{pl}_{T-t}}{\partial \lambda}d\lambda\right) & \text{if} \quad y > \frac{\bar{y}}{\beta} \quad \text{and} \quad y < l \\ \frac{\rho}{1+\rho}\left(\frac{\partial V^{pl}_{T-t}}{\partial t}dt + \frac{\partial V^{pl}_{T-t}}{\partial \lambda}d\lambda\right) & \text{if} \quad y < \frac{\bar{y}}{\beta} \quad \text{and} \\ & y > l + \frac{1}{1-\beta}D(\rho,t)\tau(\bar{y} - \beta y). \end{cases} \tag{4.7}$$

Since $\frac{\partial V^{pl}_{T-t}}{\partial t}dt > 0$ and $\frac{\partial V^{pl}_{T-t}}{\partial \lambda}d\lambda < 0$, the effect is in principle ambiguous for each case. However, for cases 1 and 2 where mothers are not eligible to the part-time subsidy, we notice that the overall effect should be very small. Likewise for case 3 we postulate that $\frac{\partial V^{pl}_{T-t}}{\partial t}dt + \frac{\partial V^{pl}_{T-t}}{\partial \lambda}d\lambda > 0$ since the duration effect should more than compensate for the decrease in the replacement rate due to the positive effect on the value of the part-time subsidy. For case 4 we have

$$dy^{*EXT}_{T-t} = \frac{1}{\beta(1-\tau D(\rho,t))} * \left(\underbrace{\frac{\rho}{1+\rho}\left(\frac{\partial V^{pl}_{T-t}}{\partial t}dt + \frac{\partial V^{pl}_{T-t}}{\partial \lambda}d\lambda\right)}_{>0} - \underbrace{\tau\frac{\partial D(\rho,t)}{\partial t}(\bar{y} - \beta y^{*EXT}_{T-t})dt}_{>0} \right) \tag{4.8}$$

and hence, the overall effect on the reservation income of the extensive margin is ambiguous. Like the partial effect in (4.5), it includes the additional value of staying in parental leave minus the forgone part-time subsidy. If the additional value of staying in parental leave predominates, mothers will prolong the employment interruption.

However, if eligible to the subsidy, for the reservation income of the intensive margin we find the total reform effect

$$dy^{*INT}_{T-t} = \frac{\partial y^{*INT}_{T-t}}{\partial t}dt > 0. \tag{4.9}$$

Hence, mothers need a higher compensation to work full-time and so the relative attractiveness of working part-time increases.

To conclude, our theoretical model predicts a positive effect of the reform on part-time employment. Since the effects on the extensive margin are ambiguous, though, the question arises if the hypothesized increase in part-time labour supply reduces full-time labour supply or has positive employment effects. We will test these implications empirically in the next section.

4.4 Estimation strategy

4.4.1 Identification

The identification of causal effects typically requires to imagine a counterfactual situation in which the individuals exposed to the reform would not have been exposed. We exploit the exogenous variation induced by the parental leave reform to define an indicator variable $D \in \{0, 1\}$ such that $D = 0$ whenever a mother gave birth shortly before July and $D = 1$ whenever the mother gave birth shortly after that date. A possible identification strategy could be to compare mothers in $D = 0$ with those in $D = 1$ in case many data points are available exactly at the cut-off. Any difference in the employment outcomes of interest Y could now be attributed to the change of the part-time subsidy scheme if nothing else drives a potential difference in the outcomes.[6] However, any estimation strategy, that exploits information away from the cut-off date to increase the sample size, might potentially be exposed to seasonal patterns and time trends. In our setting the starting date of the school year could invalidate such an analysis. Depending on the federal state, the school year usually starts in August or September and child care attendance follows this time plan. Children who are already one year old have better chances to get a child care slot. This implies that children born before the cut-off date who are slightly older than those born after the cut-off date might have a higher probability to attend child care. If employment decisions of mothers systematically differ shortly before and after the cut-off date due to the availability of public child care, then a measured difference in the outcomes can not be purely attributed to the reform. We account for this by comparing the difference in outcomes in 2015 ($T = 1$) to the difference in the previous year ($T = 0$). Similar to prior articles (Cygan-Rehm, 2016; Cygan-Rehm et al., 2018; Schönberg and Ludsteck, 2014), this DiD identification strategy yields an average treatment effect on the treated (ATET) in $T = 1$ under certain assumptions. To clarify things, consider the potential outcomes framework proposed by Rubin (1973). In general, denote variables with capital letters and its realizations with lowercase letters. Define the potential outcome for the two time periods as Y_0^d and Y_1^d such that for every observation in the sample only the potential outcome with $D = d$ of the realized outcome is observed. Further, we observe some covariates which we denote by X. Then, Heckman et al. (1997) show that our parameter of interest $\text{ATET} = \theta = \mathbb{E}\left[Y_1^1 - Y_1^0 | D = 1\right]$ is identified under the following set of assumptions.

Assumption 1 (Common trends):
$\mathbb{E}\left[Y_1^0 - Y_0^0 | X, D = 0\right] = \mathbb{E}\left[Y_1^0 - Y_0^0 | X, D = 1\right]$.
In words, conditional on X the average outcomes for $D = 0$ and $D = 1$ would have followed parallel trends in the absence of the treatment. In our setting this means that the difference in outcomes between mothers giving birth shortly before

[6] This identification strategy was applied by some studies in the context of parental leave implementation using a cut-off date (e.g., Dahl et al., 2016).

and after the cut-off date stays constant between 2014 and 2015 in the absence of the reform. In general, the assumption is empirically not testable and there might be evidence that can raise doubts concerning the validity of the assumption. Following the standard in the literature, in our sensitivity analysis we estimate effects for periods where we would not expect an effect (placebo reform). Crucially, the assumption might only hold conditional on some covariates X. For example, local economic differences or personal characteristics of mothers might affect the trends differently. The inclusion of a rich set of covariates may therefore help to make the assumption more credible. For our analysis we use geographic information as well as personal characteristics like education or the employment history. If the unconditional mean differences drastically differ from an estimator with many included controls, this may at least be interpreted as a non-robustness against the chosen specification. In other words, if specifications with many control variables shift the results, it is very likely that some form of *observed* common trend confounding, that may or may not be fully adjusted for, takes place. Clearly, this argument does not rule out some form of common trend confounding, that is unrelated to the rich set of control variables included.

Assumption 2 (Observational rule): The outcome process follows the observational rule

$$Y_t = \begin{cases} Y_t^0 & \text{if} \quad D_t = 0 \\ Y_t^1 & \text{if} \quad D_t = 1. \end{cases}$$

Hence, we require that reform exposure of one mother does not affect the outcome of another mother. The assumption can be violated if being exposed to the reform has an impact on a colleague's or a friend's reemployment decision. While we cannot completely rule out this kind of peer effect, we argue that the narrow birth interval of four weeks for reform exposure makes the occurrence of peer effects very unlikely.[7]

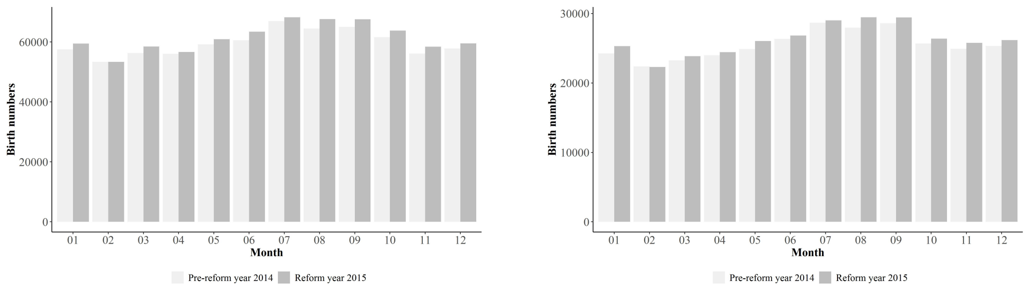
Assumption 3 (No anticipation): $\mathbb{E}\left[Y_0^1 - Y_0^0 | D = 1\right] = 0$.

This assumption requires that being exposed to the reform has no effect prior to the reform and thus rules out anticipation effects. It might be violated if mothers plan to give birth to their child in order to benefit from the new regime. This kind of anticipation can occur in two different forms. Firstly, women considering to become mothers could have tried to plan conception and secondly, they could have tried to postpone the birth shortly before the calculated birthdate. The first type of anticipation is relevant for those mothers knowing about the reform before they are pregnant. However, the German parliament approved the law only in November 2014. Hence, knowledge on the reform becoming definitely effective for births from July 2015 onwards was less than nine months before the cut-off date available when concerned mothers have already been pregnant. To rule out the possibility that mothers might have heard from the draft and waited for another few months, we analyzed

[7] Welteke and Wrohlich (2019) identify peer effects for mothers with births between a much longer period (July 2007 and December 2009).

FIGURE 4.2: Birth numbers in 2014 and 2015

(a) Official birth numbers (b) Birth numbers from *BeH*



Source: Own calculations from *BeH* and the German Federal Statistical Office (2014*a*, 2015*a*).

monthly birth numbers from 2015 in comparison with the previous year. Both statistics from the German Federal Statistical Office and the imputed birth numbers from the *Employment History (BeH)* used in the subsequent analysis show a similar movement in 2014 and 2015 and one cannot detect any sudden increase in July 2015 (see Figure 4.2). We handle the second type of anticipation, trying to postpone the birthdate that might especially relevant for planned Caesarian sections, by dropping individuals with births two weeks around the cut-off date. Obviously, the two weeks rule is arbitrary and we check the sensitivity against it in Section 4.6. Furthermore, the share of mothers wanting a Caesarian section is rather low in Germany (at maximum two to three percent). The concrete definitions of the indicators D and T are summarized in Table 4.3.

Assumption 4 (Common support): $0 < p(X) < 1$ where $p(X) = \mathbb{E}[D|X]$.[8]

It follows that we exclude perfect predictability for belonging to group $D = 0$ or $D = 1$. For our estimation procedure we enforce support by dropping observations with no overlap.

TABLE 4.3: Treatment definition

Cut-off 01/07/2015	Control group	Treated group
$T = 0$	mid May - mid June 2014	mid July - mid August 2014
$T = 1$	mid May - mid June 2015	mid July - mid August 2015

Source: Own representation.

4.4.2 Estimation of average effects

Given these assumptions, different estimands can be shown to identify the ATET. We avoid arbitrary parametric assumptions on the data generating process. We rely instead on results of the semi-parametric DiD literature. For example, Heckman et al. (1997), Abadie (2005) and Lechner (2011) propose different variations of matching and inverse probability weighting type estimators. Recently Sant'Anna and Zhao

[8]Notice that our nonparametric identification of the ATET only requires that $p(X) < 1$. We strengthen this assumption because the estimation strategy proposed in the next section requires the stronger form of common support. For details see (Zimmert, 2018).

(2018) and Zimmert (2018) propose DiD estimators that combine propensity score and outcome estimation (Augmented Inverse Probability Weighting AIPW). In particular, they show that

$$ATET = \theta = \mathbb{E}\left[\frac{1}{\lambda_D}\frac{T-\lambda_T}{\lambda_T(1-\lambda_T)}\frac{D-p(X)}{1-p(X)}\left(Y-\gamma(X,T)\right)\right] \tag{4.10}$$

where $\lambda_D = \mathbb{E}[D]$, $\lambda_T = \mathbb{E}[T]$ and
$\gamma(X,T) = T\mathbb{E}[Y|X,T=1,D=0] + (1-T)\mathbb{E}[Y|X,T=0,D=0]$.
Hence, the propensity score $p(X)$ and the outcome model $\gamma(X,T)$ have to be estimated in a first step. A major advantage of this class of estimands is that they are doubly robust in the sense that when either the outcome model or the propensity score is misspecified, the estimator is still consistent. Misspecification of the propensity score or the outcome model is a particular concern when using parametric models. Depending on the concrete setting, the researcher faces at least two more or less arbitrary decisions regarding the propensity score or the outcome model. Firstly, given a set of potential controls, it is a priori unclear which ones to include in the model. E.g., as argued before, controlling for a large set of regional dummies might improve the credibility of the common trend assumption in our case. However, it remains for example unclear whether we should use dummies at the state or district level. Secondly, the functional form of the covariates (polynomials, interactions) that enter the model has to be manually chosen by the researcher. As employments trends might, e.g., differ by age, controlling for this variable can be necessary. Still, it is unclear if the covariate should enter the model in squared, some higher order polynomial form or interacted with say a regional dummy. So-called supervised machine learning algorithms (for an overview see Hastie et al., 2009) partly avoid these problems and cope with settings where the dimensionality of a model increases with the sample size. In our application a major advantage of using machine learning algorithms compared to standard parametric models is that we can exploit the rich information in the administrative data set more effectively.[9] In particular, we do not rely on a certain specification but choose the covariates and their (implicit) functional form in a data-driven way. Combining machine learning first stages and the nonparametric second stage, we are able to reduce the sensitivity of our results towards functional form assumptions or arbitrary specification choices to a minimum. Building on the double machine learning results of Chernozhukov et al. (2018), Zimmert (2018) shows that the estimator based on the sample analogues of the estimand in (4.10) converges with square-root-N to a normal distribution and has the asymptotic variance

$$\sigma^2 = \mathbb{E}\left[\left(\frac{1}{\lambda_D}\frac{T-\lambda_T}{\lambda_T(1-\lambda_T)}\frac{D-p(X)}{1-p(X)}\left(Y-\gamma(X,T)\right)\right)^2\right] \tag{4.11}$$

[9] Parametric models can be regarded as submodels among the many options the algorithm can choose.

Procedure ATET estimation

Introduce the subsample index $l = 1, 2$ and denote the corresponding information set by $\mathcal{I}_l$ as well as its complement by $\mathcal{I}_l^C$.

1. Randomly split the sample in equally sized subsamples 1 and 2.

2. **for** $l = 1$ **to** 2 **do**

   Estimate the propensity score $p(x)$ and the outcome projections $\gamma(x, 0)$ and $\gamma(x, 1)$ in the sample with $\mathcal{I}_l^C$ using any suitable machine learning method or an ensemble of them.

   Predict $\hat{p}(x)$, $\hat{\gamma}(x, 0)$ and $\hat{\gamma}(x, 1)$ in the sample with $\mathcal{I}_l$.

   end

3. Denote $\hat{p}(x_i) = \hat{p}(x_i)_{l=1,2}$, $\hat{\gamma}(x_i, 0) = \hat{\gamma}(x_i, 0)_{l=1,2}$ and $\hat{\gamma}(x_i, 1) = \hat{\gamma}(x_i, 1)_{l=1,2}$. Then construct the vector with elements $\frac{1}{\lambda_D} \frac{t_i - \lambda_T}{\lambda_T(1-\lambda_T)} \frac{d_i - \hat{p}(x_i)}{1-\hat{p}(x_i)} (y_i - \hat{\gamma}(x_i, t_i))$ for $i = 1, ..., N$ and estimate ATET as

$$\hat{\theta} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\lambda_D} \frac{t_i - \lambda_T}{\lambda_T(1 - \lambda_T)} \frac{d_i - \hat{p}(x_i)}{1 - \hat{p}(x_i)} (y_i - \hat{\gamma}(x_i, t_i)) .$$

as long as the propensity score and the outcome model are consistent and the product of their convergence rates achieves $N^{-\frac{1}{2}}$. These are much lower requirements than for example those needed for parametric models. Importantly, the rate conditions are satisfied for popular machine learning algorithms like Lasso (e.g., Belloni and Chernozhukov, 2013) or Random Forests (Wager and Walther, 2015) under particular forms of sparsity. Hence, the flexibility or dimensionality of the models used can grow with the sample size as long as it grows at a somewhat slower rate. An additional requirement for the validity of the asymptotic results is that training and prediction sample need to be separated. This gives rise to the following ATET estimation procedure as proposed in Zimmert (2018).

The algorithm splits the sample in two different complementary subsamples and estimates the propensity score as well as the outcome projections in one of the samples. Then the values of the propensity score and the outcome projections are predicted in the other sample. Subsequently, this procedure is reverted such that one obtains a vector of propensity score and outcome projection predictions for the whole sample. These first step predictions are then plugged into the sample analogue of the estimand in (4.10). Of course, one could extend this estimation principle and split the sample into much more subsamples. This may increase the small sample efficiency of the estimator because much more information can be used for the estimation of the first step parameters. However, it also drastically increases the computational burden of the procedure. In our application we argue that the sample is large enough

such that estimation on the 50 percent subsample should not decrease efficiency too much.
For the prediction task we use a combinations of Lasso and Random Forests. While the Lasso as a form of penalized regression can be seen as a global nonparametric method, Random Forests are ensembles of regression trees and therefore a local nonparametric method. We merge the predictions from the two methods by choosing out-of-sample mean squared error optimal weights. In this way, we obtain a purely data-driven procedure that assigns a high weight to the machine learner which shows a good predictive performance. This should make our procedure more robust against the tuning parameter choices of the two estimators in the ensemble.[10]

4.4.3 Estimation of heterogenous effects

Yet another advantage of the estimand proposed in (4.10) is its capability to infer subgroup specific average effects. In particular, denote a subset of the observed covariates by $Z \subseteq X$. In our case Z might for example include dummies for income groups or whether the mother worked part-time before parental leave. Then in order to assess how the effect of the reform varies among these subgroups we are interested in the parameter

$$\theta(z) = \mathbb{E}\left[Y_1^1 - Y_1^0 | D = 1, Z = z\right].$$

The parameter represents a so-called conditional average treatment effect on the treated (CATET[11]). Define the propensity score conditional on Z as $\mathbb{E}\left[D|Z\right] = p(Z)$. Then we can show that under the assumptions in Section 4.4.1 and the further assumption that $p(Z) > 0$, the CATET is identified as

$$\theta(z) = \mathbb{E}\left[\frac{1}{p(Z)} \frac{T - \lambda_T}{\lambda_T(1-\lambda_T)} \frac{D - p(X)}{1 - p(X)} (Y - \gamma(X,T)) \Big| Z = z\right]. \tag{4.12}$$

The details for this result are provided in Appendix C. Equation (4.12) suggests to estimate CATET as a projection of the reweighted outcome on Z. A similar strategy was proposed in Abadie (2005) for DiD designs. However, the estimator of Abadie (2005) for the CATET relies on least squares regression weighted by the propensity score $p(X)$. Since we estimate $p(X)$ with our ensemble learner described in Section 4.4.2, inference for this estimator might be very complicated in our setting. We therefore reweight the transformed outcome also used in (4.10) for average effect estimation by $p(Z)$ instead of λ_D to account for the fact that we are interested in a subpopulation that is defined conditional on $D = 1$ *and* $Z = z$. In practice, Z

[10] For the Lasso we choose the penalty term by 5-fold cross-validation and otherwise rely on the default values in the **R**-package glmnet. The Random Forest is estimated using the default values in the **R**-package ranger.

[11] For some recent contributions in other settings see (Abrevaya et al., 2015; Chernozhukov and Semenova, 2017; Fan et al., 2019; Lee et al., 2017; Wager and Athey, 2018; Zimmert and Lechner, 2019)

is low-dimensional and hence $p(Z)$ can for example be estimated using logit regression. The estimand in (4.12) then suggests to simply use ordinary least squares (OLS) regression of the transformed outcome on the independent variables Z. Chernozhukov and Semenova (2017) show that this type of estimation strategy leads to valid inference for the OLS coefficients even when the first stages $p(X)$ and $\gamma(X, T)$ were predicted with sophisticated machine learning algorithms as described in the previous section. In particular, they demonstrate that the first stage estimations have no bearing on the asymptotic behaviour of the estimator. Given the results of Zimmert (2018), we postulate that this also holds for the DiD setting. A rigorous formal argument is, however, beyond the scope of this paper.
In contrast to the standard subgroup analysis, our procedure provides joint OLS inference on the coefficients. The method may therefore be also suitable to avoid the usual multiple testing problem when analyzing heterogeneous effects.

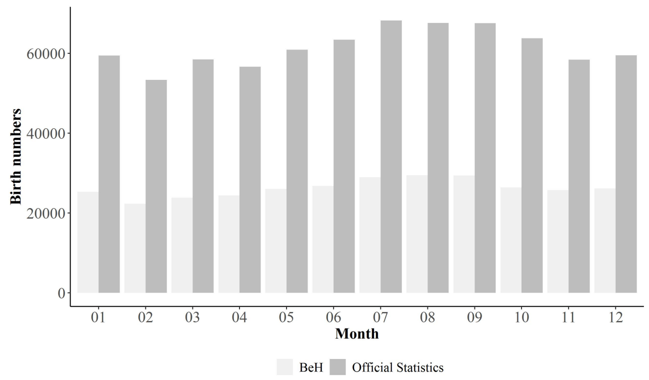
4.5 Data

To empirically test our proposed theoretical considerations we use comprehensive data from the German Federal Employment Agency provided by the research data centre (FDZ) of the Institute for Employment Research (IAB). The exploitation of administrative in contrast to survey data like in previous studies (Bergemann and Riphahn, 2010, 2015; Cygan-Rehm, 2016; Cygan-Rehm et al., 2018; Kluve and Schmitz, 2018) has some major advantages: large sample size, mandatory notification by the employer and detailed longitudinal information on a daily basis (Müller et al., 2017). Still, the fact that the data is collected for the use by the social security system implies that some information normally provided in surveys is not given. Concretely, we do not have exact information on child birth, but rely on a sophisticated imputation by Müller et al. (2017) which is explained in the following section. As a proof of quality, imputed birth numbers will show a movement over the year similar to official statistics. We use the population of mothers employed subject to social security contributions (SSC) before (potential) child birth given in the *Employment History (BeH, version 10.03.00)*. As we focus on the return to work, we neglect mothers previously not working, registered unemployed, in active labor market programs or receiving social assistance. Moreover, the data excludes self-employed and civil servants as they are not subject to SSC.

Child birth defines treatment

The *Employment History* covers all individual employment spells on a daily basis. While employers have to notify authorities at least once a year, notifications are furthermore only recorded if the employment spell ends. The imputation of the day of child birth is based on this information. Employers register when an expectant mother exits her job for the period of maternity protection and receives payment by the statutory health insurance. In general, maternity protection begins six weeks

FIGURE 4.3: Imputed and official birth numbers in 2015



Source: Own calculations from *BeH* and the German Federal Statistical Office (2015*a*).

before the calculated birthday such that Müller et al. (2017) add six weeks of maternity protection to impute child birth. Unfortunately, the exit reason "receiving entitlements from statutory health insurance" can also include long-term sickness ($\geq$ six weeks). However, misspecifications can be minimized by three restrictions. Firstly, the group of young women is more likely to have a child, but less likely to suffer from long-term sickness. Based on official birth statistics, Müller et al. (2017) restrict the childbearing age to 38 years for the first birth and to 40 years for subsequent births. Secondly, mothers are not allowed to work during the 14 weeks lasting period of maternity protection. Any shorter job interruption period is more likely to specify a break due to illness. Thirdly, subsequent births are only possible after a period of about 40 weeks. As the pre-term rate was found to be 9.2 percent in 2010 (March of Dimes et al., 2012), the authors limit the gap to 32 weeks. As we only observe births from mothers previously employed subject to SSC, total numbers are smaller compared to official statistics for whole Germany. However, Figure 4.3 shows that the movement in the considered period is very similar for the official birth numbers and imputed births which highlights the quality of the imputation and the data in general. Additionally, we argue that the exclusion of a certain time window around the cut-off date should mitigate the problem. In section 4.6.1 we will show that our results are insensitive to the specific choice of the window width indicating that the imputation error is empirically a minor concern. We observe between 22,000 and 30,000 births per month while sample size in similar studies using survey data amounts to about 2,000 births for the same period.

Control and outcome variables

TABLE 4.4: Descriptive statistics of covariates by group membership

Variable	Control group Mean	sd	Treated group Mean	sd	Standardized difference
Age	29.36	3.45	29.31	3.47	-0.013
Number of children	1.301	0.522	1.301	0.520	-2.8E-04

Migration background	0.066	0.248	0.066	0.249	0.002
Place of living (Federal state, *baseline* Schleswig-Holstein:)					
Hamburg	0.024	0.152	0.023	0.151	-0.001
Lower Saxony	0.088	0.283	0.088	0.284	0.002
Bremen	0.007	0.082	0.006	0.077	-0.010
North Rhine-Westphalia	0.187	0.390	0.185	0.388	-0.007
Hessen	0.069	0.253	0.072	0.259	0.013
Rhineland-Palatinate	0.045	0.207	0.045	0.208	0.002
Baden-Wuerttemberg	0.133	0.339	0.134	0.340	0.003
Bavaria	0.168	0.374	0.173	0.378	0.013
Saarland	0.012	0.108	0.010	0.100	-0.016
Berlin	0.046	0.209	0.044	0.206	-0.008
Brandenburg	0.036	0.186	0.034	0.182	-0.007
Mecklenburg-Western Pomerania	0.022	0.148	0.023	0.150	0.003
Saxony	0.067	0.251	0.067	0.251	-1.3E-04
Saxony-Anhalt	0.030	0.171	0.030	0.171	1.2E-04
Thuringia	0.033	0.178	0.031	0.174	-0.009
Education (*baseline* Lower/middle secondary school without vocational training):					
Lower/middle secondary school with vocational training	0.494	0.500	0.488	0.500	-0.011
High school without vocational training	0.022	0.148	0.023	0.150	0.004
High school with vocational training	0.199	0.399	0.199	0.399	-0.001
University of applied sciences	0.022	0.147	0.023	0.150	0.007
University	0.179	0.383	0.185	0.388	0.015
Days in					
marginal employment within last five years	122.56	259.24	122.60	260.33	1.5E-04
part-time employment within last five years	327.67	497.69	331.20	500.38	0.007
full-time employment within last five years	979.62	652.75	982.60	655.75	0.005
Previous job:					
Employment pattern (*baseline* Marginal employment):					
Part-time	0.309	0.462	0.309	0.462	0.001
Full-time	0.645	0.478	0.644	0.479	-0.002
Gross monthly income in Euros	2195.92	1241.25	2219.98	1245.32	0.019
Temporary contract	0.225	0.418	0.224	0.417	-0.002
Requirement level (*baseline* Unskilled or semi-skilled activities):					
Specialist activities	0.664	0.472	0.663	0.473	-0.004
Complex specialist activities	0.123	0.329	0.121	0.327	-0.006
Highly complex activities	0.119	0.324	0.122	0.328	0.009
Occupational area (classification system Kldb2010 1-digit, *baseline* Agriculture, forestry, farming, and gardening):					
Production of raw materials and goods and manufacturing	0.061	0.240	0.063	0.243	0.007
Construction, architecture, surveying and technical building services	0.008	0.089	0.007	0.086	-0.006
Natural sciences, geography and informatics	0.021	0.142	0.019	0.138	-0.009
Traffic, logistics, safety and security	0.048	0.213	0.046	0.210	-0.007
Commercial services, trading, sales, the hotel business and tourism	0.187	0.390	0.183	0.387	-0.010
Business organization, accounting, law and administration	0.249	0.432	0.254	0.435	0.012
Health care, the social sector, teaching and education	0.382	0.486	0.383	0.486	0.003

Philology, literature, humanities, social sciences, economics, media, art, culture, and design	0.035	0.184	0.035	0.184	1.5E-04
Establishment:					
Number of female employees	334	999	340	1019	0.006
Total number of employees	695	3049	712	3173	0.005
N	46,263		48,212		

Notes: Instead of using federal states like presented in the table, districts are used for the prediction of the propensity score and the outcome equation. sd = standard deviation.
Source: Own calculations based on employee data from the *Employment History (BeH)*, establishment data from the *Establishment History Panel (BHP)*.

The *Employment History* includes a large set of other individual and job-related characteristics that we use to predict the propensity score and the outcome equation explained in Section 4.4.2. These are measured at the last employment spell, i.e., right before child birth. We include the individual age, the number of children, a binary indicator for having a migration background and the place of residence on the district level. Inserting regional fixed effects is especially important to control for the macroeconomic background or the availability of child care facilities showing large variation over German districts (compare Chapter 3). With the inclusion of 402 districts the number of covariates gets large. While standard parametric models might not converge at these levels of dimensionality, our machine learning approach is able to flexibly include this large list of control variables.
Further information relate to education[12] and occupational characteristics. We include six categories for the educational degree combined with information on the occupation (lower/middle secondary school with(out) vocational training, high school with(out) vocational training, university of applied sciences, university). Other covariates concern the requirement level (unskilled up to highly complex activities) and the occupational code both coming from the German classification of occupations KldB2010.
To control for the individual employment history the days spent in marginal, part- or full-time employment within the last five years and the working time pattern of the previous job (marginal, part- or full-time employment) are considered. The *Employment History* does not contain information on continuously measured working hours. Hence, we use the working time pattern which is provided by the employer as the ninth digit of the classification of occupations and merge additional particularities, i.e., marginal employment as special form of a part-time contract. Marginal employment in Germany, so-called *Mini jobs*, do not exceed earnings of 450 Euro per month and are exempted from income taxation.
The gross monthly income is a generated variable that considers the duration of the

[12]As the variable is characterized by a higher share of missing or inconsistent values compared to other information provided by the employer, we rely on an imputation procedure proposed by Fitzenberger et al. (2005).

employment spell. As employers only have to indicate income up to the SSC assessment ceiling,[13] this variable is right censored. Special payments and misdeclarations can shift the upper ceiling such that we restrict the income range to up to 6,500 Euro per month. The type of working contract (fixed- or long-term) is also controlled for. Additionally, we also include information on the number of (female) employees coming from the *IAB Establishment History Panel (BHP, version 7516 v1)*.
Table 4.4 shows mean values and their standard deviation of the previously described covariates by group status. The last column gives the standardized mean difference (Rubin, 2001) between these two groups and demonstrates that the sample is well balanced as all values are close to zero.
Unfortunately, the data does not contain information on actual receipt of parental subsidies. So, our estimates are intention to treat-effects.
The outcome variables of interest refer to employment after child birth. Since employment spells are available until the end of 2017, we can analyze maternal labor market outcomes up to two years. We measure current employment (in full- or part-time as well as in marginal employment) as binary indicators every three months until the second birthday of the child, i.e., at eight different points in time. Figure 4.4 gives mean outcomes for the treated group before the reform and shows that employment rates are increasing with the child's age. Before the first birthday the employment rate amounts to about 20 percent and is mainly characterized by full-time jobs. Interestingly, the employment rate sharply increases to about 60 percent until the second birthday with the highest share consisting of part-time contracts. It seems that the average pre-reform mother takes the maximum parental leave period of twelve months and returns in a part-time job.
We analyze other variables of job continuity depicted by a binary indicator for staying with the same employer and job quality in terms of earnings accumulated up to the first and second year. The second column of Table 4.5 shows that about 52 percent of treated mothers return to their previous employer before the reform while average earnings amount to 1,744 Euro in the first year and to 13,492 Euro up to the second birthday (including those with zero earnings who have not returned yet).

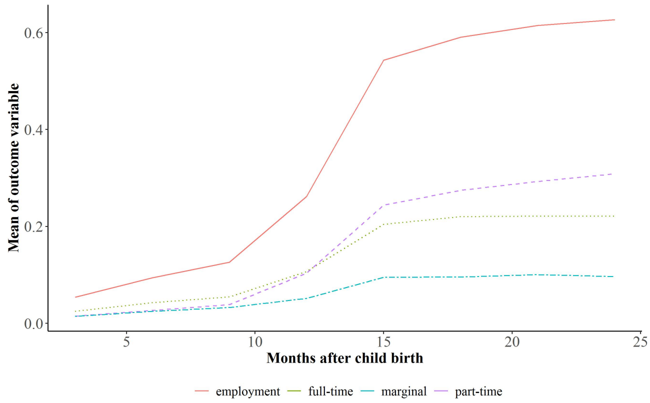
4.6 Results

4.6.1 Estimation results for ATET and sensitivity analysis

We present our main estimation results in graphs where the time in months after child birth is depicted on the horizontal axis and the ATET on the vertical axis. Apart from the machine learning augmented DiD estimator (solid line), we also show results of the unadjusted mean estimator without including any covariates (dashed line).
We start by discussing the overall employment effect in Figure 4.5a. The reform gives

[13] For the statutory pension insurance, the assessment ceiling amounts to 6,050 Euro in 2015. For the health insurance, it was 4,125 Euro in 2015.

FIGURE 4.4: Outcome means of treated mothers before reform



Notes: $T = 0$ in 2014. $N = 23,993$.
Source: Own calculations based on employee data from the *Employment History (BeH)*.

increasing and positive employment effects up to nine months after birth amounting to statistically significant two percentage points at maximum. Although the effect size seems to be small, it accounts for about $\frac{0.020}{0.138} = 14$ percent of the pre-reform mean. Additionally, since take up only amounts to about 19 percent (Federal Statistical Office, 2019), the effect for those actually choosing the new regime should be much higher.[14] The effect vanishes when the child turns one year old indicating that the first birthday remains a reference point for the majority of previously employed mothers. 18 months after child birth the ATET slightly increases again, but does not reach significance on conventional levels.

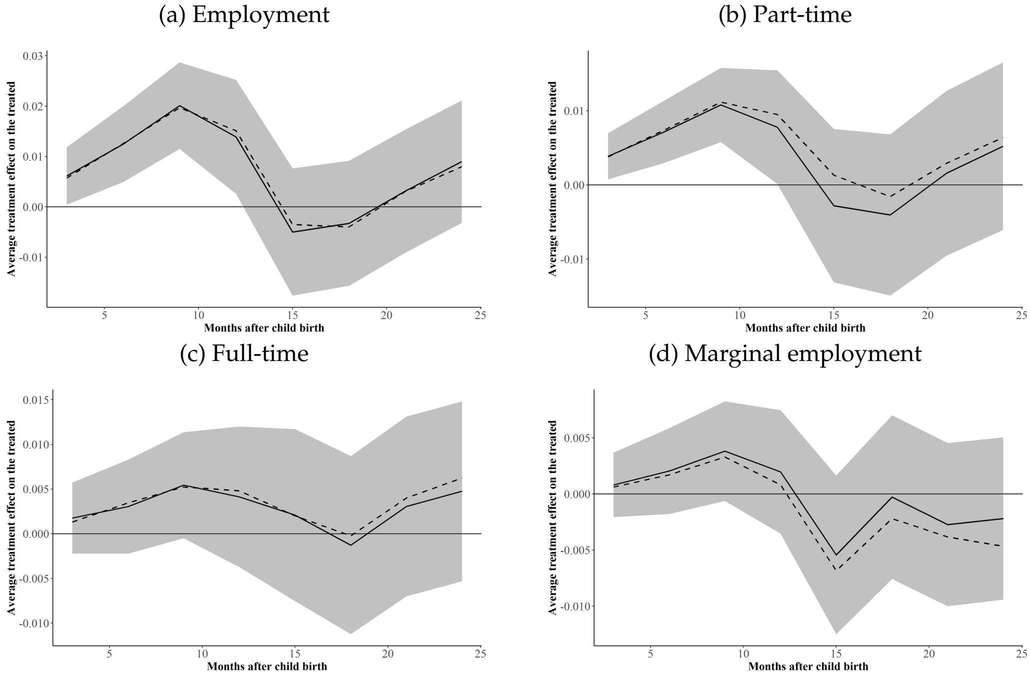
How is this overall positive employment effect in the first year distributed over different employment patterns? Figures 4.5b to 4.5d show that part-time employment mainly drives this finding. At maximum the part-time effect equals about one percentage point which is one half of the overall employment increase. The reform's impact on full-time employment is close to zero. Marginal employment as form of part-time employment is also not significantly affected.

These empirical findings are in line with the proposed model mechanisms of Section 4.3 predicting a decrease of the reservation income for a part-time job relative to a full-time job. Moreover, we find that the theoretically ambiguous effect on the extensive employment margin is empirically positive. Interestingly, the seemingly positive effect on the attractiveness of part-time employment is not associated by a drop in full-time employment. Instead, it is reflected by an increase in overall employment which means that the additional value of further staying in parental leave is dominated by the effect of the forgone part-time subsidy.

We do not identify any persistent employment patterns, i.e., the distribution into full-, part-time or marginal employment is not affected until the child's second birthday. This might suggest that those mothers incentivized to return before the child's

[14] We are cautious when interpreting DiD results in a Wald estimator kind of manner. For example De Chaisemartin and D'Haultfoeuille (2017) show that such an argumentation may only be valid under very restrictive assumptions.

FIGURE 4.5: Baseline estimation results



Notes: Treatment status is measured six weeks around cut-off date excluding the two weeks on each side of the cut-off date. $T = 1$ in 2015, $T = 0$ in 2014. $N = 94,475$. Grey shaded areas depict pointwise 95 % confidence intervals. AIPW DiD is solid line. Unadjusted DiD is dashed line.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

first birthday under a part-time contract would have also returned in part-time under the pre-reform regime after the child's first birthday. In turn, mothers who would have returned in full-time employment under the pre-reform regulations might not be willing to accept reduced working hours supplemented by parental benefits before the child's first birthday as they might fear to get stuck in a part-time contract.

Panel A of Table 4.5 shows that the overall positive employment effects reflect in higher accumulated earnings within the first year (about 273 Euro) and less precisely estimated within the second year (about 314 Euro). Furthermore, the shorter employment break does not affect the probability to return to the previous employer. The validity of our findings is supported by the fact that the unconditional mean differences (dashed lines) are very close to our estimation results using a rich set of covariates. If our setting would be sensitive to confounding with respect to one of the observed covariates, we would expect different results for the simple differences in means estimator and our procedure. Moreover, we examine the plausibility of the common trend assumption between treatment and control group in absence of the reform by postponing the reform year to 2014. While this kind of check cannot directly test the assumption, Figure 4.6 hints at similar employment trends before the reform reflected in ATETs that are precisely measured at around zero. Earnings and job continuity are as well not affected (compare Panel B of Table 4.5).

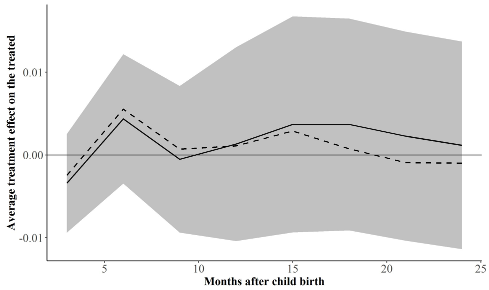
TABLE 4.5: ATETs for job continuity and accumulated earnings

Outcome	$D = 1, T = 0$		Unadjusted DiD		AIPW	
	Mean	sd	ATET	se	ATET	se
Panel A: Baseline						
Same employer	0.522	0.500	0.008	0.006	0.006	0.006
Accumulated earnings						
1st year	1,743.86	5,370.63	282.38***	72.23	272.66***	71.41
2nd year	13,492.12	16,178.52	430.09**	211.53	314.32*	191.22
Notes: Treatment status is measured six weeks around the cut-off date excluding the two weeks on each side of the cut-off date. $T = 1$ in 2015, $T = 0$ in 2014. $N = 94,475$.						
Panel B: Placebo						
Same employer	0.561	0.496	-0.016**	0.007	-0.009	0.006
Accumulated earnings						
1st year	1,656.53	5,343.75	-76.78	73.17	-73.00	71.84
2nd year	12,556.42	15,652.03	-91.02	213.05	44.81	192.52
Notes: Treatment status is measured six weeks around the cut-off date excluding the two weeks on each side of the cut-off date. $T = 1$ in 2014, $T = 0$ in 2013. $N = 89,374$.						
Panel C: Small bandwidth						
Same employer	0.547	0.498	0.015*	0.009	0.016*	0.009
Accumulated earnings						
1st year	1,613.31	5,339.65	237.27**	99.82	242.32**	98.37
2nd year	12,909.01	15,939.58	171.54	294.15	189.89	264.31
Notes: Treatment status is measured four weeks around the cut-off date excluding the two weeks on each side of the cut-off date. $T = 1$ in 2015, $T = 0$ in 2014. $N = 48,544$.						
Panel D: Large bandwidth						
Same employer	0.553	0.497	0.005	0.006	-3.9E-04	0.006
Accumulated earnings						
1st year	1,731.44	5,483.04	262.78***	73.40	268.16***	73.73
2nd year	13,257.52	16,289.88	620.13***	214.52	337.54*	194.84
Notes: Treatment status is measured eight weeks around the cut-off date excluding the four weeks on each side of the cut-off date. $T = 1$ in 2015, $T = 0$ in 2014. $N = 94,493$.						

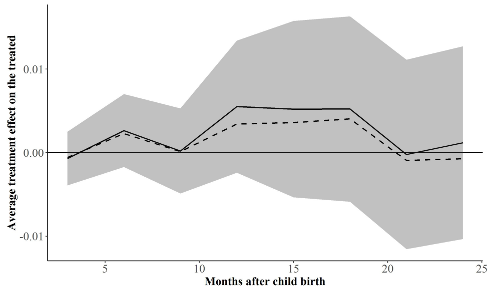
Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

FIGURE 4.6: Estimation results of placebo reform

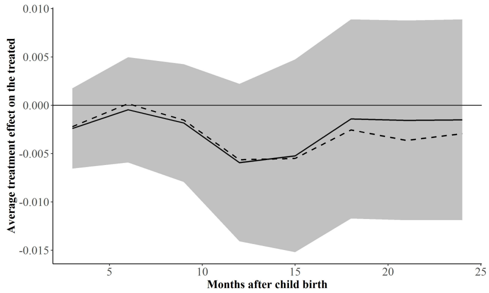
(a) Employment



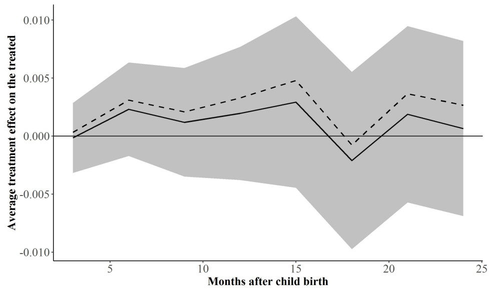
(b) Part-time



(c) Full-time



(d) Marginal employment



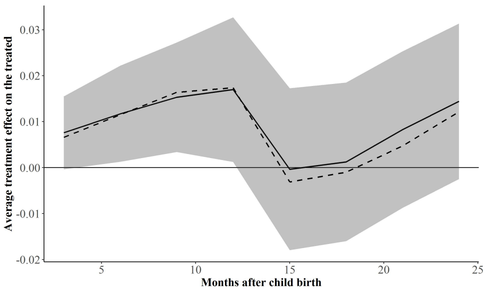
Notes: Treatment status is measured six weeks around cut-off date excluding the two weeks on each side of the cut-off date. $T = 1$ in 2014, $T = 0$ in 2013. $N = 89,374$. Grey shaded areas depict pointwise 95 % confidence intervals. AIPW DiD is solid line. Unadjusted DiD is dashed line.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

A second potential concern in our empirical strategy might be the arbitrary definition of the sampling periods around the cut-off date and the imputation error of the child's birthday (see Section 4.5). We check the sensitivity against these two threats by estimating the effects for different populations. Figure 4.7 shows the employment effects for a smaller bandwidth around the cut-off date of four weeks excluding the two weeks around this date on each side. The employment pattern induced by the reform stays the same compared to the baseline estimates: mothers return earlier in a part-time job. However, we find slightly different effects on accumulated earnings and job continuity (compare Panel C of Table 4.5). The same holds for increasing the bandwidth to eight weeks around the cut-off date with the exclusion of four weeks on each side (compare Figure 4.8). We conclude that our results are robust regarding these potential issues.

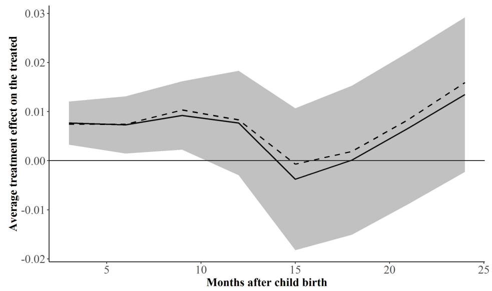
As a possible channel driving the employment outcomes, we look at the effect on subsequent births. E.g., Cygan-Rehm (2016) shows that the parental leave reform of 2007 incentivized mothers to postpone a subsequent pregnancy. Figure 4.9 does not indicate any effect on childbearing within the two year-horizon. However, this finding has to be interpreted with caution as we only observe women with subsequent births who have been employed in the meanwhile.

FIGURE 4.7: Estimation results with small bandwidth

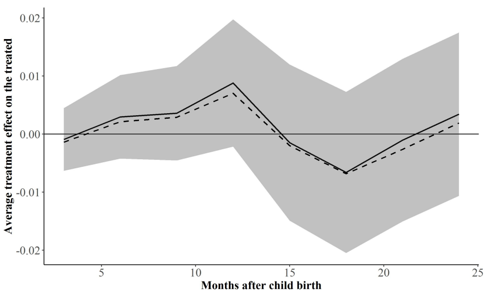
(a) Employment



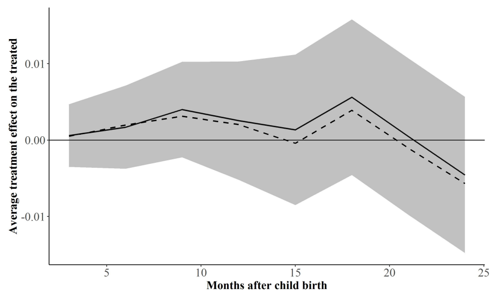
(b) Part-time



(c) Full-time



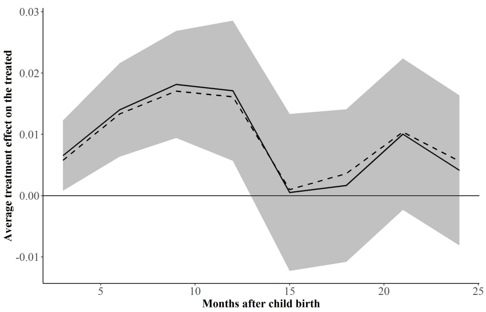
(d) Marginal employment



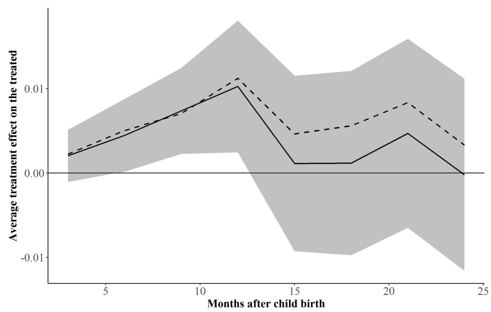
Notes: Treatment status is measured four weeks around cut-off date excluding the two weeks on each side of the cut-off date. $T = 1$ in 2015, $T = 0$ in 2014. $N = 48,544$. Grey shaded areas depict pointwise 95 % confidence intervals. AIPW DiD is solid line. Unadjusted DiD is dashed line.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

FIGURE 4.8: Estimation results with large bandwidth

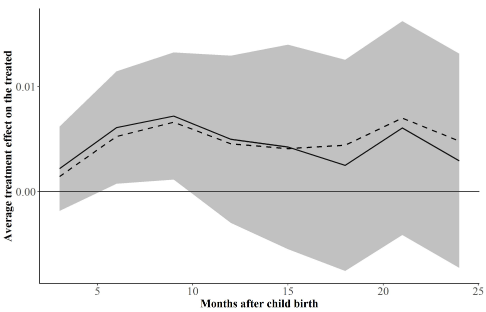
(a) Employment



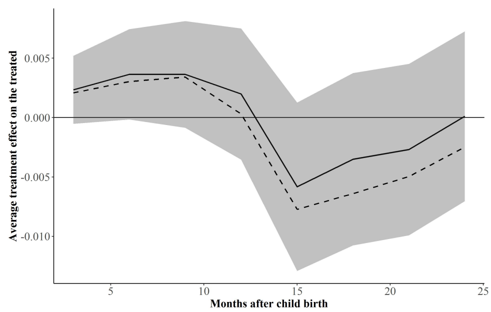
(b) Part-time



(c) Full-time

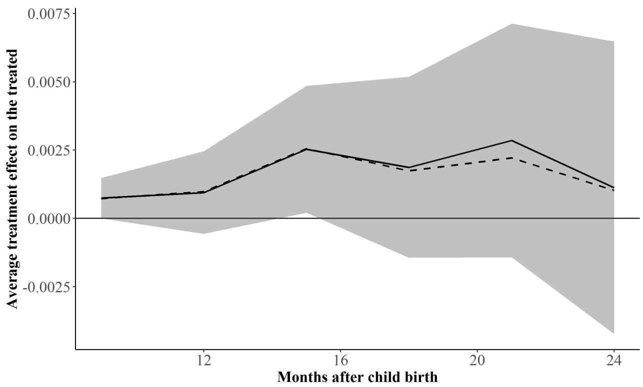


(d) Marginal employment



Notes: Treatment status is measured eight weeks around cut-off date excluding the four weeks on each side of the cut-off date. $T = 1$ in 2015, $T = 0$ in 2014. $N = 94,493$. Grey shaded areas depict pointwise 95 % confidence intervals. AIPW DiD is solid line. Unadjusted DiD is dashed line.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

FIGURE 4.9: Estimation results for subsequent birth within next 24 months



Notes: Treatment status is measured six weeks around cut-off date excluding the two weeks on each side of the cut-off date. $T = 1$ in 2015, $T = 0$ in 2014. $N = 94,475$. Grey shaded areas depict pointwise 95 % confidence intervals. AIPW DiD is solid line. Unadjusted DiD is dashed line.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

4.6.2 Estimation results for conditional effects

To better understand the channels of the reform, we investigate how the treatment effects vary over different pre-specified subgroups.[15] We investigate different income groups as well as heterogeneities concerning the prior working time pattern and the place of living (East and West Germany). The latter might yield interesting results as mothers growing up in the former GDR could have different attitudes towards maternal employment. Hence, in our setting Z contains dummies for the middle and high income groups, whether the mother worked full-time previous to child birth and a dummy for West Germany. In particular, we estimate the following specification using OLS regression

$$\tilde{y} = \beta_0 + \beta z + \epsilon$$

where $\tilde{y}$ represents the sample analogue of $\frac{1}{\hat{p}(Z)} \frac{T-\lambda_T}{\lambda_T(1-\lambda_T)} \frac{D-\hat{p}(X)}{1-\hat{p}(X)}(Y - \hat{\gamma}(X,T))$ with first stages estimated as for the average effects and $\hat{p}(Z)$ by logit regression. The resulting OLS coefficients give the effect variation for the different subgroups. They are depicted on the vertical axis for the eight different periods in Figures 4.10 to 4.13. As discussed in Section 4.4.3 we report the usual OLS standard errors.
Figures 4.10a and 4.11a show that the effect size for employment does not differ with respect to income (low income is chosen as reference group). When we further differentiate the employment effects in part-time and full-time as well as marginal employment, Figure 4.10b reveals that the positive part-time effects are mainly driven by middle income earners. E.g., after nine months the part-time effect for middle income earners is about 1.3 percentage points higher compared with low income

[15]The respective pre-reform outcome means of these subgroups are depicted in Figures B.5 and B.6.

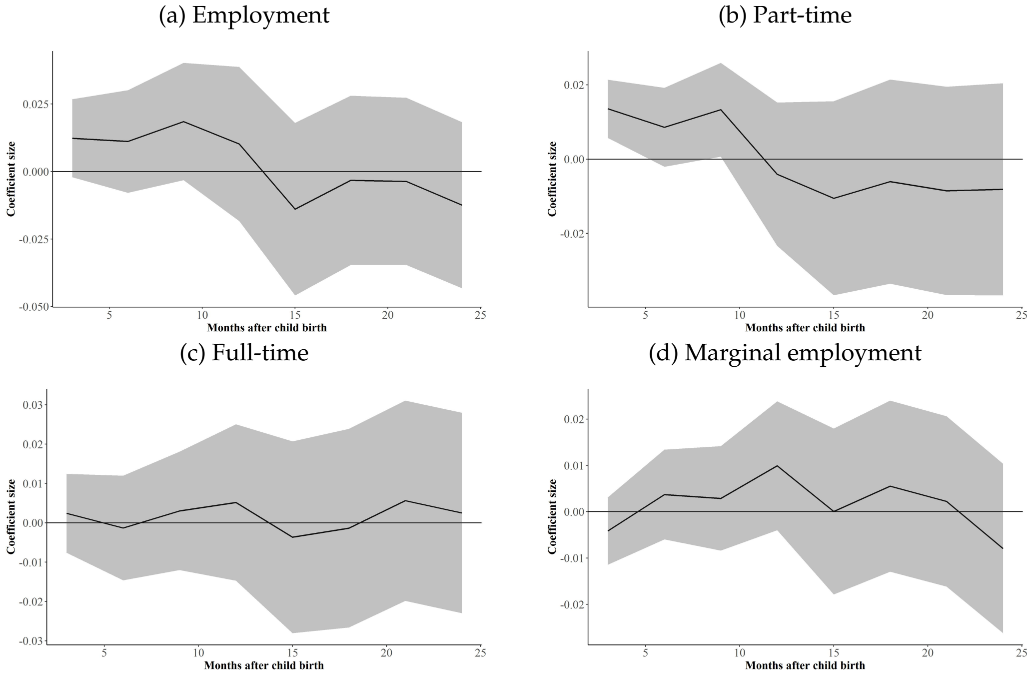
mothers.[16] For high-income mothers, the effects do not significantly differ from those with lower income (see Figure 4.10). Hence, we conclude that high-income and potentially more career-oriented mothers prefer not to return in part-time employment despite the simultaneous provision of parental subsidies since they fear the implications of reducing their working time. Mothers with middle income may be more willing to accept a part-time job because the future potential income loss after expiration of parental benefits is less severe. This argumentation is strongly supported for examining the subgroup of previously full-time employed women. Figure 4.12b demonstrates that they have a lower probability (-1.3 to -3.7 percentage points) to return in part-time employment in the first year after child birth. As their opportunity costs of taking up a part-time job are higher, they are characterized by a weaker response to the reform. We explain the similar effect size of low- and high-income mothers with the amount of the part-time subsidy. Low-income mothers are more likely to receive the minimum amount of 150 Euro such that the incentive to return before the child's first birthday is in general less pronounced. The effects for full-time and marginal employment do not significantly differ over subgroups.

Interestingly, the effect size does not significantly vary with the place of living, i.e., East and West Germany (see Figure 4.13). Bergemann and Riphahn (2015) and Kluve and Schmitz (2014) provide suggestive evidence that the parental leave reform of 2007 defines a social norm to return to work after the child's first birthday. In this regard, the new policy has the potential to further increase cultural acceptance for those mothers preferring a return even before the child turns one year old (Zoch and Hondralis, 2017). As a legacy of the German Democratic Republic, societal acceptance of maternal employment and the reliance on external child care are on general on a higher level in East Germany (e.g., Hanel and Riphahn, 2012). Consequently, a shift of social norms becomes more likely in the West German society where traditional approaches of the household's division of labor predominate. However, we do not observe statistically significant differences until the child's first birthday for mothers living in West Germany compared to East Germany (Figure 4.13a and 4.13b). Thus, we do not find suggestive evidence for a shift of social norms induced by the reform. As the administrative character of the data does not allow to follow up on this suggestion, we interpret it with caution.

Moreover, the financial incentive for part-time work, does not affect maternal employment outcomes after the child's first birthday for almost all subgroups. However, prior full-time working mothers have a lower probability for working part-time of up to 4.1 percentage points shortly before the child gets two years old. Hence, the reform may foster the path dependency of working part-time, at least for the short period of two years.

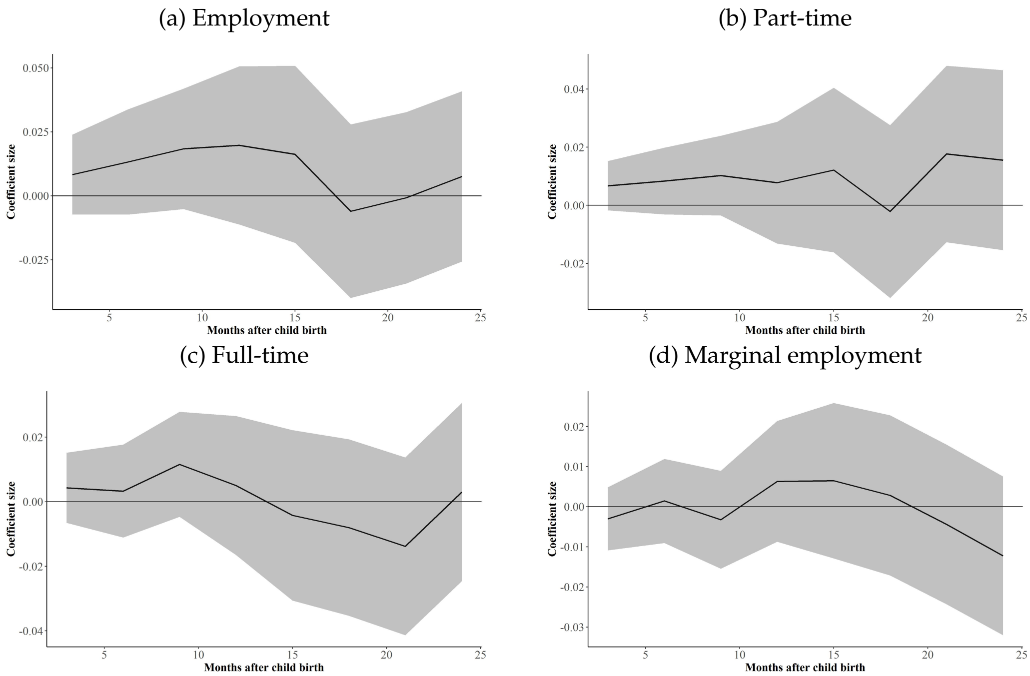
[16] Table A.5 in the appendix shows that these effects reflect in higher accumulated earnings within the first year. In the main analysis we concentrate on the CATETs for the different working time patterns. See Table A.5 for a detailed presentation.

FIGURE 4.10: Estimation results for middle income group



Notes: Low (middle, high) income corresponds to 1st (2nd, 3rd) tercile of gross previous monthly income. $N_{low\_income} = 31,170$, $N_{middle\_income} = 32,139$ and $N_{high\_income} = 31,166$. The coefficients give the effect variation for the respective subgroup. The reference group is low income. Grey shaded areas depict pointwise 95 % confidence intervals.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

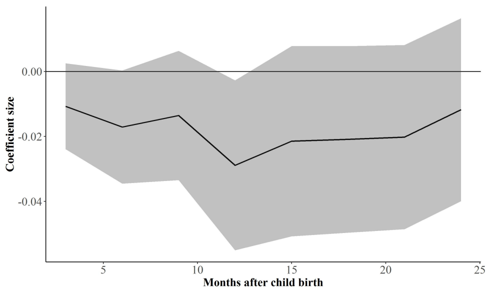
FIGURE 4.11: Estimation results for high income group



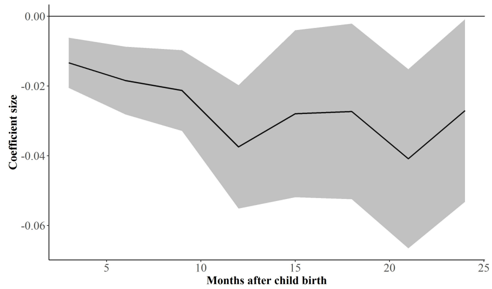
Notes: Low (middle, high) income corresponds to 1st (2nd, 3rd) tercile of gross previous monthly income. $N_{low\_income} = 31,170$, $N_{middle\_income} = 32,139$ and $N_{high\_income} = 31,166$. The coefficients give the effect variation for the respective subgroup. The reference group is low income. Grey shaded areas depict pointwise 95 % confidence intervals.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

FIGURE 4.12: Estimation results for prior full-time

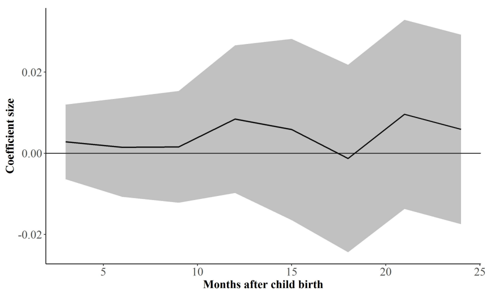
(a) Employment



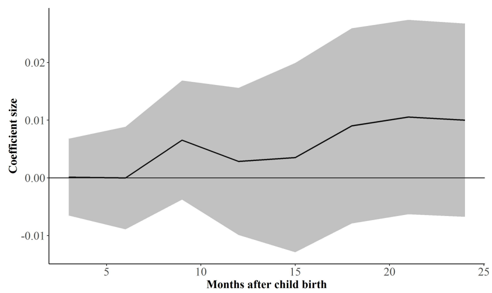
(b) Part-time



(c) Full-time



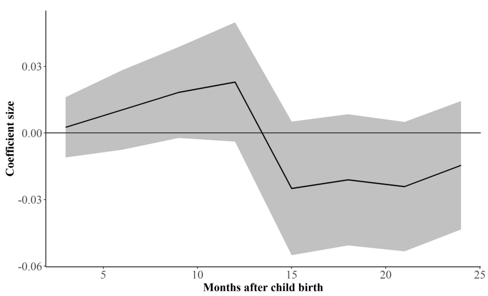
(d) Marginal employment



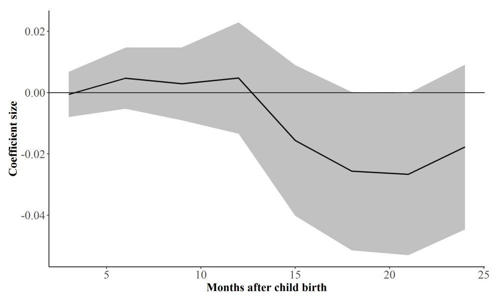
Notes: $N_{part\_time} = 29,214$ and $N_{full\_time} = 60,913$. Coefficient gives effect variation for respective subgroup. The reference group is prior part-time and marginal employment. Grey shaded areas depict pointwise 95 % confidence intervals.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

FIGURE 4.13: Estimation results for West Germany

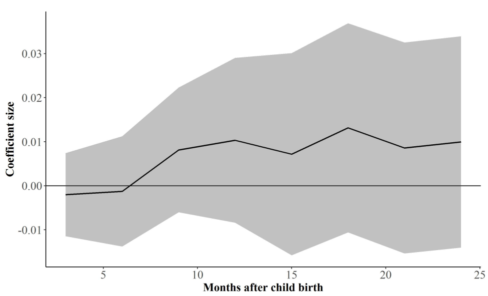
(a) Employment



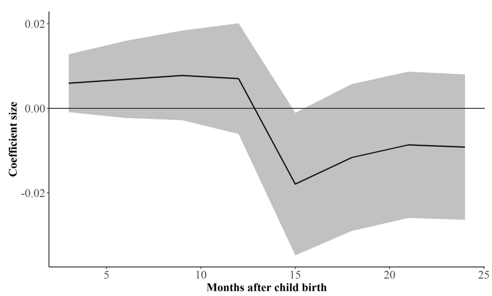
(b) Part-time



(c) Full-time



(d) Marginal employment



Notes: $N_{east} = 21,967$ and $N_{west} = 72,508$. The coefficients give the effect variation for the respective subgroup. The reference group is East Germany. Grey shaded areas depict pointwise 95 % confidence intervals.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

4.7 Discussion

Although the overall employment effects amount to about 14 percent of the pre-reform mean, the new regulations, and consequently a return to work before the child's first birthday are only attractive to about 20 percent of all female benefit recipients (Federal Statistical Office, 2019). In this regard, analyzing individual working hour preferences can be helpful to understand if the remaining 80 percent prefer spending time with the child (working hour preferences are expected to stay close to zero) or if the availability of child care plays a role (working hour preferences are expected to rise). Chapter 3 shows that family policies have the potential to change individual preferences albeit they move on average quite similarly to agreed working hours. Unfortunately, information of working hour preferences is not given in the administrative data we use for the empirical analysis. One also has to keep in mind that our estimates are intention to treat-effects and estimates considering actual receipt of the parental subsidy would be higher. Moreover, it might be possible that only well informed mothers know about the implementation of the reform. As pre-reform regulations are effective for several years and there are many different websites to calculate the benefit amount, we expect it to be a minor issue.

Different countries, notably the United States, discuss an introduction of paid maternity protection or parental leave respectively. Therefore, it might be of interest to investigate the financial expenditure. Given the limited information that we have to calculate the individual tax revenue, we estimate the welfare gain for switching from the old to the new regulations. The "average" mother in our sample has an average monthly gross income of 2,263 Euro before child birth (net: 1,506 Euro[17]) which is in line with official statistics (Federal Statistical Office, 2019). Assume a mother cares for her child until it turns six months old and receives the full basic amount of parental benefits,[18] before she returns in a part-time job with a gross monthly income of 1,528 Euro as in our sample (net: 1,124 Euro) and receives the reduced subsidy until the child's second birthday. The total benefit amount in Euro of this average mother is

$$4 \text{ months} * 1,506 * 0.65 + 18 \text{ months} * (1,506 - 1,124) * 0.67 = 8,532.$$

In case she received the full basic amount until the first birthday, it would be

$$10 \text{ months} * 1,506 * 0.65 = 9,792 \text{ Euro}.$$

[17]The German taxation system is based upon the household. For the tax class we assume an approximately egalitarian household income between partners (class 4).

[18]Note that the first two months are maternity protection during which maternity allowances are paid to previously employed women by the health insurance and the employer.

Additional part-time work generates a tax revenue of about 552 Euro until the child's first birthday. Then, the total public savings of the new regulation amounts to

$$9,792 - 8,532 + 552 = 1,812 \text{ Euro or to } 1,812 * 36,229 = 65,652,600 \text{ Euro}$$

for the 36,229 mothers of the birth cohort of the third quarter in 2015 deciding for the new policy. While the monthly tax revenue cannot compensate the parental leave subsidy, this short-term total welfare gain for switching the parental leave regulations is substantial. Note however, that this simple cost-benefit-analysis does not consider public child care expenditure that might offset the total public savings.

4.8 Conclusion

Improving the labor market prospects of young mothers may imply strong welfare gains. We analyze a German parental leave reform promoting a fast return to part-time work after child birth while receiving parental benefits. Although shorter employment interruptions can improve career prospects, the policy could have pushed mothers to reduce working hours instead of returning to a full-time job when the child is older. Our results from semi-parametric DiD estimation in combination with machine learning algorithms do not provide evidence for such a downside. The reform rather yields additional part-time effects before the child's birthday of up to one percentage point driven by mothers who would have also returned to a part-time job in absence of the reform. Heterogenous effects support this argumentation. We find that mothers with lower opportunity costs of accepting a part-time job (i.e., those with middle previous income and prior part-time workers) show a stronger response to the reform.

Previous regulations established the child's first birthday as a reference point for the parental leave duration reinforced by the legal claim for a child care slot introduced in 2013. Insignificant differences for West and East Germany found in this paper do not hint at the potential to further change societal expectations when to return to work. Our findings have to be interpreted in the context of a labor market in which working mothers with children younger than one year old are a minority and the German tax and health insurance systems additionally promote an inegalitarian household division of paid working hours. This might also explain why the introduction of the new parental benefit system does not indicate better employment prospects in terms of working hours for those women deciding for an early return to work. To further support employees with a temporary preference for a working hour reduction, the German government recently enforced a legal claim to return to a full-time job which might especially be a good instrument for mothers after parental leave. Hence, it would be interesting to learn about long-term effects of the parental leave reform also in combination with the right to return in full-time.

Chapter 5

Conclusion

This thesis examines individual labor supply with a special focus on working hour preferences and maternal employment.
The first article analyzes how household and occupational factors contribute to the evolution of working hour discrepancies defined as the difference between preferred and actual working hours. The results from a discrete duration analysis point out that the occupational autonomy within the job is one of the main driving factors for explaining the creation and resolution of hour discrepancies. Furthermore, there are gender differences concerning mother- and fatherhood, i.e., especially mothers of young children have a lower probability to fulfill the desire for an hour increase. Based on these findings, the following two articles evaluate policies addressing families with children younger than three years old.
The second article examines the public child care expansion culminating in a legal claim for a child care slot from August 2013 onwards. Before the legal claim came into force, child care attendance for under three-year-olds is on a low level in Germany (25.2 percent in 2011, 19.8 percent in West Germany, 49.0 percent in East Germany according to the Federal Statistical Office, 2011*b*). Hence, the availability of low cost subsidized child care has a high potential to increase maternal employment in particular the agreed working hours of underemployed mothers. The empirical analysis uses exogenous variation of the child care expansion for difference-in-differences (DiD) estimation and the findings suggest on average positive effects on the extensive and intensive employment margin. As agreed and preferred working hours change similarly and the shares of under- and overemployed as well as unconstrained mothers are not significantly affected, the article concludes that the availability of child care can tap employment potentials beyond those of currently underemployed.
Available child care is a precondition for an early return to work of young mothers after the birth of a child. Two years after the legal claim for a child care slot became effective, the German government decided for a law further promoting maternal employment. This policy is evaluated in the last article. The reform extended the maximum receipt duration of part-time subsidies and hence, incentivizes for an

early return in part time after child birth. The article exploits the exogenous exposure to the reform for treatment determination and uses machine learning augmented DiD estimation. The policy yields on average positive employment effects until the child's first birthday which are mainly driven by part-time work. Full-time employment is not significantly affected suggesting that mothers do not substitute a full-time contract with reduced hours and additional subsidies. However, these effects do not persist until the child's second birthday.

A major contribution of the second and third article is the provision of conditional average treatment effects (on the treated). Targeted policy implementation requires knowledge on the potential behaviour of specific subgroups. Unlike prior studies that estimate effects for each particular subgroup and potentially suffer from multiple testing, the two articles use CATET estimation as firstly proposed by Abadie (2005). He shows that reweighted outcomes of inverse probability weighting DiD estimation can be used to project them on heterogeneity variables of interest. Resulting coefficients directly give effect variation for the respective variable with valid inference. The last article enhances the approach of Abadie (2005) by using machine learning algorithms to estimate the reweighted outcome.

The subgroup analyses of the second and third article both demonstrate that either mothers with lower school degree or income have smaller employment effects due to the availability of child care slots or the part-time subsidy. The positive average effects are mainly driven by better educated or mothers with middle to higher income. While employees with lower income are more often concerned by old-age poverty, the findings reinforce the question how to mitigate this problem. The Organization for Economic Co-operation and Development (OECD) finds that employment interruptions and short part-time work aggravate the issue of low pension entitlements and are explaining factors for the high gender pension gap in Germany compared to other OECD countries (OECD, 2019). Hence, future research and political action should explicitly concentrate on the employment careers of poorly educated mothers on the lower end of the income distribution.

Bibliography

Abadie, A. (2005), 'Semiparametric difference-in-differences estimators', *The Review of Economic Studies* **72**(1), 1–19.
URL: *https://doi.org/10.1111/0034-6527.00321*

Abrevaya, J., Hsu, Y.-C. and Lieli, R. P. (2015), 'Estimating conditional average treatment effects', *Journal of Business & Economic Statistics* **33**(4), 485–505.
URL: *https://doi.org/10.1080/07350015.2014.975555*

Allison, P. D. (1982), 'Discrete-time methods for the analysis of event histories', *Sociological methodology* **13**, 61–98.
URL: *https://www.jstor.org/stable/270718*

Alt, C., Hubert, S., Jehles, N., Lippert, K., Meiner-Teubner, C., Schilling, C. and Steinberg, H. (2015), *Datenbericht Betreuungsgeld. Auswertung amtlicher Daten und der Kifög-Länderstudien aus den Jahren 2013/2014/2015. Abschlussbericht*, Deutsches Jugendinstitut e. V. München.
URL: *https://www.dji.de/en/media/literature-search/detailansicht/literatur/22672-datenbericht-betreuungsgeld.html*

Andresen, M. E. and Havnes, T. (2019), 'Child care, parental labor supply and tax revenue', *Labour Economics* **61**, Article 101762.
URL: *https://doi.org/10.1016/j.labeco.2019.101762*

Baker, M., Gruber, J. and Milligan, K. (2008), 'Universal child care, maternal labor supply, and family well-being', *Journal of Political Economy* **116**(4), 709–745.
URL: *https://doi.org/10.1086/591908*

Baker, M. and Milligan, K. (2008), 'How does job-protected maternity leave affect mothers' employment?', *Journal of Labor Economics* **26**(4), 655–691.
URL: *https://doi.org/10.1086/591955*

Bauernschuster, S., Hener, T. and Rainer, H. (2016), 'Children of a (policy) revolution: The introduction of universal child care and its effect on fertility', *Journal of the European Economic Association* **14**(4), 975–1005.
URL: *https://doi.org/10.1111/jeea.12158*

Bauernschuster, S. and Schlotter, M. (2015), 'Public child care and mothers' labor supply – Evidence from two quasi-experiments', *Journal of Public Economics* **123**, 1–16.
URL: *https://doi.org/10.1016/j.jpubeco.2014.12.013*

Baum, C. L. (2003), 'The effect of state maternity leave legislation and the 1993 family and medical leave act on employment and wages', *Labour Economics* **10**(5), 573–596.
URL: *https://doi.org/10.1016/S0927-5371(03)00037-X*

Belloni, A. and Chernozhukov, V. (2013), 'Least squares after model selection in high-dimensional sparse models', *Bernoulli* **19**(2), 521–547.
URL: *https://projecteuclid.org/euclid.bj/1363192037*

Bergemann, A. and Riphahn, R. T. (2010), 'Female labour supply and parental leave benefits – the causal effect of paying higher transfers for a shorter period of time', *Applied Economics Letters* **18**(1), 17–20.
URL: *https://doi.org/10.1080/13504850903425173*

Bergemann, A. and Riphahn, R. T. (2015), 'Maternal employment effects of paid parental leave', *IZA Discussion Papers* **9073**.
URL: *https://ssrn.com/abstract=2612325*

Berlinski, S. and Galiani, S. (2007), 'The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment', *Labour Economics* **14**(3), 665–680.
URL: *https://doi.org/10.1016/j.labeco.2007.01.003*

Berlinski, S., Galiani, S. and Mc Ewan, P. J. (2011), 'Preschool and maternal labor market outcomes: Evidence from a regression discontinuity design', *Economic Development and Cultural Change* **59**(2), 313–344.
URL: *https://doi.org/10.1086/657124*

Bettendorf, L. J., Jongen, E. L. and Muller, P. (2015), 'Childcare subsidies and labour supply - evidence from a large Dutch reform', *Labour Economics* **36**, 112–123.
URL: *https://doi.org/10.1016/j.labeco.2015.03.007*

BMFSFJ (2015), *Fünfter Bericht zur Evaluation des Kinderförderungsgesetzes*, Federal Ministry for Family Affairs, Senior Citizens, Women and Youth (BMFSFJ).
URL: *https://www.bmfsfj.de/bmfsfj/service/publikationen/fuenfter-bericht-zur-evaluation-des-kinderfoerderungsgesetzes–kifoeg-bericht-2015-/86370*

BMFSFJ (2018), *Elterngeld, ElterngeldPlus und Elternzeit. Das Bundeselterngeld- und Elternzeitgesetz*, Federal Ministry for Family Affairs, Senior Citizens, Women and Youth (BMFSFJ).
URL: *https://www.bmfsfj.de/bmfsfj/service/publikationen/elterngeld–elterngeldplus-und-elternzeit-/73770*

Bryan, M. L. (2007), 'Free to choose? Differences in the hours determination of constrained and unconstrained workers', *Oxford Economic Papers* **59**(2), 226–252.
URL: *https://doi.org/10.1093/oep/gpl033*

Campbell, I. and van Wanrooy, B. (2013), 'Long working hours and working-time preferences: Between desirability and feasibility', *Human Relations* **66**(8), 1131–1155.
URL: *https://doi.org/10.1177/0018726713478641*

Cascio, E. U. (2009), 'Maternal labor supply and the introduction of kindergartens into American public schools', *Journal of Human Resources* **44**(1), 140–170.
URL: *https://doi.org/10.3368/jhr.44.1.140*

Chamberlain, G. (1984), 'Panel data', *Handbook of econometrics* **2**, 1247–1318.
URL: *https://doi.org/10.1016/S1573-4412(84)02014-6*

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W. and Robins, J. (2018), 'Double/debiased machine learning for treatment and structural parameters', *The Econometrics Journal* **21**(1), C1–C68.
URL: *https://doi.org/10.1111/ectj.12097*

Chernozhukov, V. and Semenova, V. (2017), *Simultaneous inference for Best Linear Predictor of the Conditional Average Treatment Effect and other structural functions*, Unpublished manuscript, Department of Economics, Massachusetts Institute of Technology.
URL: *https://arxiv.org/abs/1702.06240*

Clarkberg, M. and Moen, P. (2001), 'Understanding the time-squeeze married couples' preferred and actual work-hour strategies', *American Behavioral Scientist* **44**(7), 1115–1136.
URL: *https://doi.org/10.1177/00027642010044007005*

Cygan-Rehm, K. (2016), 'Parental leave benefit and differential fertility responses: Evidence from a German reform', *Journal of Population Economics* **29**(1), 73–103.
URL: *https://doi.org/10.1007/s00148-015-0562-z*

Cygan-Rehm, K., Kühnle, D. and Riphahn, R. T. (2018), 'Paid parental leave and families' living arrangements', *Labour Economics* **53**, 182–197.
URL: *https://doi.org/10.1016/j.labeco.2018.05.008*

Dahl, G. B., Løken, K. V., Mogstad, M. and Salvanes, K. V. (2016), 'What is the case for paid maternity leave?', *Review of Economics and Statistics* **98**(4), 655–670.
URL: *https://doi.org/10.1162/REST_a_00602*

De Chaisemartin, C. and D'Haultfoeuille, X. (2017), 'Fuzzy differences-in-differences', *The Review of Economic Studies* **85**(2), 999–1028.
URL: *https://doi.org/10.1093/restud/rdx049*

DIW Berlin/SOEP (2014), *SOEP-RS FiD 2013 - Erhebungsinstrumente 2013 von 'Familien in Deutschland': Elternfragebogen 1: "Für alle Kinder der Jahrgänge 2012/2013"*, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.
URL: *https://www.diw.de/sixcms/detail.php?id=diw_01.c.479306.de*

Drago, R., Tseng, Y.-P. and Wooden, M. (2005), 'Usual and preferred working hours in couple households', *Journal of Family Studies* **11**(1), 46–61.
URL: *https://doi.org/10.5172/jfs.327.11.1.46*

Drago, R., Wooden, M. and Black, D. (2009), 'Who wants flexibility? Changing work hours preferences and life events', *Industrial and Labor Relations Review* **62**(3), 394–414.
URL: *https://doi.org/10.1177/001979390906200308*

Duflo, E. (2001), 'Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment', *American Economic Review* **91**(4), 795–813.
URL: *https://doi.org/10.1257/aer.91.4.795*

Ehing, D. (2014), 'Unter- und Überbeschäftigung in Deutschland: Eine Analyse der Arbeitszeitwünsche von Erwerbstätigen vor dem Hintergrund des demografischen Wandels', *Zeitschrift für Sozialreform* **60**(3), 247–272.
URL: *https://doi.org/10.1515/zsr-2014-0303*

Ellguth, P., Gerner, H.-D. and Zapf, I. (2018), 'Arbeitszeitkonten in Betrieben und Verwaltungen: Flexible Arbeitszeitgestaltung wird immer wichtiger', *IAB-Kurzbericht* **4/2018**.
URL: *http://hdl.handle.net/10419/185850*

Fagan, C. (2001), 'Time, money and the gender order: work orientations and working-time preferences in Britain', *Gender, Work & Organization* **8**(3), 239–266.
URL: *https://doi.org/10.1111/1468-0432.00131*

Fan, Q., Hsu, Y.-C., Lieli, R. P. and Zhang, Y. (2019), *Estimation of Conditional Average Treatment Effects with High-Dimensional Data*, Unpublished manuscript.
URL: *https://arxiv.org/abs/1908.02399*

Federal Statistical Office (2006), *Kindertagesbetreuung regional 2006. Ein Vergleich aller 439 Kreise in Deutschland*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000386*

Federal Statistical Office (2008), *Kindertagesbetreuung regional 2008. Ein Vergleich aller 429 Kreise in Deutschland*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000386*

Federal Statistical Office (2009), *Kindertagesbetreuung regional 2009. Ein Vergleich aller 413 Kreise in Deutschland*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000386*

Federal Statistical Office (2010*a*), *Kindertagesbetreuung regional 2010. Ein Vergleich aller 412 Kreise in Deutschland*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000386*

Federal Statistical Office (2010*b*), *Statistiken der Kinder- und Jugendhilfe. Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01.03.2010*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00002286*

Federal Statistical Office (2011*a*), *Bevölkerung und Erwerbstätigkeit. Natürliche Bevölkerungsbewegung*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000015*

Federal Statistical Office (2011*b*), *Kindertagesbetreuung regional 2011. Ein Vergleich aller 412 Kreise in Deutschland*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000386*

Federal Statistical Office (2011*c*), *Statistiken der Kinder- und Jugendhilfe. Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01.03.2011*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00002286*

Federal Statistical Office (2012*a*), *Bevölkerung und Erwerbstätigkeit. Natürliche Bevölkerungsbewegung*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000015*

Federal Statistical Office (2012*b*), *Kindertagesbetreuung regional 2012. Ein Vergleich aller 402 Kreise in Deutschland*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000386*

Federal Statistical Office (2012*c*), *Statistiken der Kinder- und Jugendhilfe. Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01.03.2012*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00002286*

Federal Statistical Office (2013*a*), *Kindertagesbetreuung regional 2013. Ein Vergleich aller 402 Kreise in Deutschland*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000386*

Federal Statistical Office (2013*b*), *Statistiken der Kinder- und Jugendhilfe. Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01.03.2013*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00002286*

Federal Statistical Office (2014*a*), *Bevölkerung und Erwerbstätigkeit. Natürliche Bevölkerungsbewegung*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000015*

Federal Statistical Office (2014*b*), *Kindertagesbetreuung regional 2014. Ein Vergleich aller 402 Kreise in Deutschland*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000386*

Federal Statistical Office (2014*c*), *Statistiken der Kinder- und Jugendhilfe. Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01.03.2014*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00002286*

Federal Statistical Office (2015*a*), *Bevölkerung und Erwerbstätigkeit. Natürliche Bevölkerungsbewegung*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000015*

Federal Statistical Office (2015*b*), *Kindertagesbetreuung regional 2015. Ein Vergleich aller 402 Kreise in Deutschland*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000386*

Federal Statistical Office (2015*c*), *Statistiken der Kinder- und Jugendhilfe. Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01.03.2015*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00002286*

Federal Statistical Office (2016), *Kindertagesbetreuung regional 2016. Ein Vergleich aller 402 Kreise in Deutschland*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00000386*

Federal Statistical Office (2019), *Öffentliche Sozialleistungen. Statistik zum Elterngeld. Beendete Leistungsbezüge für im 3. Vierteljahr 2015 geborene Kinder*, Federal Statistical Office, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/receive/DESerie_serie_00002282*

Felfe, C. and Lalive, R. (2018), 'Does early child care affect children's development?', *Journal of Public Economics* **159**, 33–53.
URL: *https://doi.org/10.1016/j.jpubeco.2018.01.014*

Felfe, C., Nollenberger, N. and Rodríguez-Planas, N. (2015), 'Can't buy mommy's love? Universal childcare and children's long-term cognitive development', *Journal of Population Economics* **28**(2), 393–422.
URL: *https://doi.org/10.1007/s00148-014-0532-x*

Fendel, T. and Jochimsen, B. (2017), 'Child care reforms and labor participation of migrant and native mothers', *IAB-Discussion Paper* **9/2017**.
URL: *http://hdl.handle.net/10419/162480*

Fitzenberger, B., Osikominu, A. and Völter, R. (2005), 'Imputation rules to improve the education variable in the IAB employment subsample', *ZEW Discussion Paper* **10/2005**.
URL: *https://dx.doi.org/10.2139/ssrn.711044*

Fitzpatrick, M. D. (2012), 'Revising our thinking about the relationship between maternal labor supply and preschool', *Journal of Human Resources* **47**(3), 583–612.
URL: *https://doi.org/10.3368/jhr.47.3.583*

Fuchs, J., Söhnlein, D., Weber, B. and Weber, E. (2016), 'Ein integriertes Modell zur Schätzung von Arbeitskräfteangebot und Bevölkerung', *IAB-Forschungsbericht* **10/2016**.
URL: *http://hdl.handle.net/10419/146108*

Gangl, M. and Ziefle, A. (2015), 'The making of a good woman: Extended parental leave entitlements and mothers' work commitment in Germany', *American Journal of Sociology* **121**(2), 511–563.
URL: *https://doi.org/10.1086/682419*

Gathmann, C. and Sass, B. (2018), 'Taxing childcare: Effects on family labor supply and children', *Journal of Labor Economics* **36**(3), 665–709.
URL: *https://doi.org/10.1086/696143*

Gelbach, J. B. (2002), 'Public schooling for young children and maternal labor supply', *The American Economic Review* **92**(1), 307–322.
URL: *https://doi.org/10.1257/000282802760015748*

Geyer, J., Haan, P. and Wrohlich, K. (2015), 'The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach', *Labour Economics* **36**, 84–98.
URL: *https://doi.org/10.1016/j.labeco.2015.07.001*

Gibbons, R. and Katz, L. F. (1991), 'Layoffs and lemons', *Journal of Labor Economics* **9**(4), 351–380.
URL: *https://doi.org/10.1086/298273*

Gielen, A. C. (2009), 'Working hours flexibility and older workers' labor supply', *Oxford Economic Papers* **61**(2), 240–274.
URL: *https://doi.org/10.1093/oep/gpn035*

Givord, P. and Marbot, C. (2015), 'Does the cost of child care affect female labor market participation? An evaluation of a french reform of childcare subsidies', *Labour Economics* **36**, 99–111.
URL: *https://doi.org/10.1016/j.labeco.2015.07.003*

Goldin, C. (2014), 'A grand gender convergence: Its last chapter', *American Economic Review* **104**(4), 1091–1119.
URL: *http://www.aeaweb.org/articles?id=10.1257/aer.104.4.1091*

Goux, D. and Maurin, E. (2010), 'Public school availability for two-year olds and mothers' labour supply', *Labour Economics* **17**(6), 951–962.
URL: *https://doi.org/10.1016/j.labeco.2010.04.012*

Greenhaus, J. H. and Beutell, N. J. (1985), 'Sources of conflict between work and family roles', *Academy of Management Review* **10**(1), 76–88.
URL: *https://doi.org/10.5465/amr.1985.4277352*

Grözinger, G., Matiaske, W. and Tobsch, V. (2008), 'Arbeitszeitwünsche, Arbeitslosigkeit und Arbeitszeitpolitik', *WSI Mitteilungen* **61**.
URL: *http://hdl.handle.net/10419/150652*

Hanel, B. and Riphahn, R. T. (2012), 'The employment of mothers – recent developments and their determinants in East and West Germany', *Jahrbücher für Nationalökonomie und Statistik* **232**(2), 146–176.
URL: *https://doi.org/10.1515/jbnst-2012-0205*

Hastie, T., Tibshirani, R. and Friedman, J. (2009), *The elements of statistical learning: prediction, inference and data mining*, Springer-Verlag, New York.
URL: *https://link.springer.com/book/10.1007%2F978-0-387-84858-7*

Havnes, T. and Mogstad, M. (2011), 'Money for nothing? Universal child care and maternal employment', *Journal of Public Economics* **95**(11), 1455–1465.
URL: *https://doi.org/10.1016/j.jpubeco.2011.05.016*

Heckman, J. J. and Borjas, G. J. (1980), 'Does unemployment cause future unemployment? definitions, questions and answers from a continuous time model of heterogeneity and state dependence', *Economica* **47**(187), 247.
URL: *https://doi.org/10.2307/2553150*

Heckman, J. J., Ichimura, H. and Todd, P. E. (1997), 'Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme', *The review of economic studies* **64**(4), 605–654.
URL: *https://doi.org/10.2307/2971733*

Hochschild, A. and Machung, A. (1989), *The second shift*, New York: Penguin Books.
URL: *https://www.penguinrandomhouse.com/books/310593/the-second-shift-by-arlie-hochschild-with-anne-machung/*

Holst, E. and Bringmann, J. (2016), 'Arbeitszeitrealitäten und Arbeitszeitwünsche in Deutschland: Methodische Unterschiede ihrer Erfassung im SOEP und Mikrozensus (Desired and actual working hours in Germany: Methodological differences in measurement between SOEP and Mikrozensus)', *SOEPpapers* **859**.
URL: *https://dx.doi.org/10.2139/ssrn.2818955*

Imbens, G. W. and Wooldridge, J. M. (2009), 'Recent developments in the econometrics of program evaluation', *Journal of Economic Literature* **47**(1), 5–86.
URL: *https://doi.org/10.1257/jel.47.1.5*

Jessen, J., Jessen, R. and Kluve, J. (2019), 'Punishing potential mothers? Evidence for statistical employer discrimination from a natural experiment', *Labour Economics* **59**, 164–172.
URL: *https://doi.org/10.1016/j.labeco.2019.04.002*

Kaufman, G. and Uhlenberg, P. (2000), 'The influence of parenthood on the work effort of married men and women', *Social Forces* **78**(3), 931–47.
URL: *https://doi.org/10.1093/sf/78.3.931*

Kluve, J. and Schmitz, S. (2014), 'Social norms and mothers' labor market attachment', *IZA Discussion Paper* **8115**.
URL: *http://dx.doi.org/10.2139/ssrn.2471521*

Kluve, J. and Schmitz, S. (2018), 'Back to work: Parental benefits and mothers' labor market outcomes in the medium run', *Industrial and Labor Relations Review* **71**(1), 143–173.
URL: *https://doi.org/10.1177/0019793917710933*

Kluve, J. and Tamm, M. (2013), 'Parental leave regulations, mothers' labor force attachment and fathers' childcare involvement: Evidence from a natural experiment', *Journal of Population Economics* **26**(3), 983–1005.
URL: *https://doi.org/10.1007/s00148-012-0404-1*

Knaus, M. C. and Otterbach, S. (2019), 'Work hour mismatch and job mobility: Adjustment channels and resolution rates', *Economic Inquiry* **57**(1), 227–242.
URL: *https://doi.org/10.1111/ecin.12586*

Kreyenfeld, M. and Hank, K. (2000), 'Does the availability of child care influence the employment of mothers? Findings from western Germany', *Population Research and Policy Review* **19**(4), 317–337.
URL: *https://doi.org/10.1023/A:1026556309080*

Lalive, R. and Zweimüller, J. (2009), 'How does parental leave affect fertility and return to work? Evidence from two natural experiments', *The Quarterly Journal of Economics* **124**(3), 1363–1402.
URL: *https://doi.org/10.1162/qjec.2009.124.3.1363*

Landers, R., Rebitzer, J. and Taylor, L. (1996), 'Rat race redux: Adverse selection in the determination of work hours in law firms', *The American Economic Review* **86**(3), 329–348.
URL: *https://www.jstor.org/stable/2118200*

Lechner, M. (2011), 'The estimation of causal effects by difference-in-difference methods', *Foundations and Trends in Econometrics* **4**(3), 165–224.
URL: *http://dx.doi.org/10.1561/0800000014*

Lee, S., Okui, R. and Whang, Y.-J. (2017), 'Doubly robust uniform confidence band for the conditional average treatment effect function', *Journal of Applied Econometrics* **32**(7), 1207–1225.
URL: *https://doi.org/10.1002/jae.2574*

Lefebvre, P. and Merrigan, P. (2008), 'Child-care policy and the labor supply of mothers with young children: A natural experiment from Canada', *Journal of Labor Economics* **26**(3), 519–548.
URL: *https://doi.org/10.1086/587760*

Lundborg, P., Plug, E. and Rasmussen, A. W. (2017), 'Can women have children and a career? IV evidence from IVF treatments', *American Economic Review* **107**(6), 1611–37.
URL: *http://www.aeaweb.org/articles?id=10.1257/aer.20141467*

Lundin, D., Mörk, E. and Öckert, B. (2008), 'How far can reduced childcare prices push female labour supply?', *Labour Economics* **15**(4), 647–659.
URL: *https://doi.org/10.1016/j.labeco.2008.04.005*

March of Dimes, PMNCH, Save the Children and WHO (2012), *Born Too Soon: The Global Action Report on Preterm Birth*, World Health Organization, Geneva.
URL: *https://www.who.int/pmnch/media/news/2012/preterm_birth_report/en/index1.html*

Matiaske, W., Schmidt, T., Seifert, H. and Tobsch, V. (2017), 'Arbeitszeitdiskrepanzen mindern Zufriedenheit mit Arbeit und Gesundheit', *WSI Mitteilungen* **4/2017**.
URL: *https://doi.org/10.5771/0342-300x-2017-4-287*

Merz, J. (2002), 'Time and economic well-being - a panel analysis of desired versus actual working hours', *Review of Income and Wealth* **48**(3), 317– 346.
URL: *https://doi.org/10.1111/1475-4991.00057*

Microcensus (2011), *Microcensus*, Federal Statistical Office, Wiesbaden.
URL: *https://www.forschungsdatenzentrum.de/de/10-21242-12211-2011-00-00-1-1-0*

Microcensus (2015), *Microcensus*, Federal Statistical Office, Wiesbaden.
URL: *https://www.forschungsdatenzentrum.de/en/10-21242-12211-2015-00-00-1-2-0*

Mood, C. (2010), 'Logistic regression: Why we cannot do what we think we can do, and what we can do about it', *European Sociological Review* **26**(1), 67–82.
URL: *https://doi.org/10.1093/esr/jcp006*

Morrissey, T. W. (2017), 'Child care and parent labor force participation: A review of the research literature', *Review of Economics of the Household* **15**(1), 1–24.
URL: *https://doi.org/10.1007/s11150-016-9331-3*

Müller, D., Strauch, K. et al. (2017), 'Identifying mothers in administrative data', *FDZ-Methodenreport* **13/2017**.
URL: *http://doku.iab.de/fdz/reporte/2017/MR_13-17_EN.pdf*

Nollenberger, N. and Rodríguez-Planas, N. (2011), 'Child care, maternal employment and persistence: A natural experiment from Spain', *IZA Discussion Paper* **5888**.
URL: *https://ssrn.com/abstract=1906191*

OECD (2017), *OECD Employment Outlook*, Organisation for Economic Co-operation and Development (OECD).
URL: *https://doi.org/10.1787/19991266*

OECD (2019), *OECD Pensions at a Glance*, Organisation for Economic Co-operation and Development (OECD).
URL: *https://doi.org/10.1787/19991363*

Patnaik, A. (2019), 'Reserving time for daddy: The consequences of fathers' quotas', *Journal of Labor Economics* **37**(4), 1009–1059.
URL: *https://doi.org/10.1086/703115*

Pollmann-Schult, M. (2009), 'Arbeitszeitwunsch und -wirklichkeit im Familienkontext. Eine Analyse der Diskrepanzen zwischen präferierter und tatsächlicher Arbeitszeit', *Soziale Welt* **60**(2), 163–178.
URL: *https://doi.org/10.5771/0038-6073-2009-2-163*

Pollmann-Schult, M. and Reynolds, J. (2017), 'The Work and Wishes of Fathers: Actual and Preferred Work Hours among German Fathers', *European Sociological Review* **33**(6), 823–838.
URL: *https://doi.org/10.1093/esr/jcx079*

Potuchek, J. (1997), *Who supports the family?: Gender and breadwinning in dual-earner marriages*, Stanford University Press.
URL: *https://www.sup.org/books/title/?id=1002*

Reynolds, J. (2003), 'You can't always get the hours you want: Mismatches between actual and preferred work hours in the US', *Social Forces* **81**(4), 1171–1199.
URL: *https://doi.org/10.1353/sof.2003.0069*

Reynolds, J. (2004), 'When too much is not enough: Actual and preferred work hours in the United States and abroad', *Sociological Forum* **19**(1), 89–120.
URL: *https://link.springer.com/article/10.1023/B:SOFO.0000019649.59873.08*

Reynolds, J. (2014), 'Prevailing preferences: Actual work hours and work-hour preferences of partners', *Industrial & Labor Relations Review* **67**(3), 1017–1041.
URL: *https://doi.org/10.1177/0019793914537459*

Reynolds, J. and Aletraris, L. (2006), 'Pursuing preferences: The creation and resolution of work hour mismatches', *American Sociological Review* **71**(4), 618–638.
URL: *https://doi.org/10.1177/000312240607100405*

Reynolds, J. and Aletraris, L. (2010), 'Mostly mismatched with a chance of settling: Tracking work hour mismatches in the United States', *Work and Occupations* **37**(4), 476–511.
URL: *https://doi.org/10.1177/0730888410383245*

Reynolds, J. and Johnson, D. R. (2012), 'Don't blame the babies: Work hour mismatches and the role of children', *Social Forces* **91**(1), 131–155.
URL: *https://doi.org/10.1093/sf/sos070*

Rossin-Slater, M., Ruhm, C. J. and Waldfogel, J. (2013), 'The effects of California's paid family leave program on mothers' leave-taking and subsequent labor market outcomes', *Journal of Policy Analysis and Management* **32**(2), 224–245.
URL: *https://doi.org/10.1002/pam.21676*

Rubin, D. B. (1973), 'The use of matched sampling and regression adjustment to remove bias in observational studies', *Biometrics* **29**(1), 185–203.
URL: *https://www.jstor.org/stable/2529685*

Rubin, D. B. (2001), 'Using propensity scores to help design observational studies: Application to the tobacco litigation', *Health Services and Outcomes Research Methodology* **2**(3-4), 169–188.
URL: *https://doi.org/10.1023/A:1020363010465*

Sant'Anna, P. H. and Zhao, J. B. (2018), 'Doubly robust difference-in-differences estimators', *Unpublished manuscript, Department of Economics, Vanderbilt University* .
URL: *http://dx.doi.org/10.2139/ssrn.3293315*

Schlosser, A. (2005), *Public preschool and the labor supply of Arab mothers: Evidence from a natural experiment*, Unpublished manuscript, The Hebrew University of Jerusalem.
URL: *https://doi.org/10.3386/w12179*

Schönberg, U. and Ludsteck, J. (2014), 'Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth', *Journal of Labor Economics* **32**(3), 469–505.
URL: *https://doi.org/10.1086/675078*

Schulz, F. and Blossfeld, H.-P. (2006), 'Wie verändert sich die häusliche Arbeitsteilung im Eheverlauf?', *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie* **58**(1), 23–49.
URL: *https://doi.org/10.1007/s11575-006-0002-0*

Settersten Jr, R. A. and Mayer, K. U. (1997), 'The measurement of age, age structuring, and the life course', *Annual review of sociology* **23**(1), 233–261.
URL: *https://doi.org/10.1146/annurev.soc.23.1.233*

Sousa-Poza, A. and Ziegler, A. (2003), 'Asymmetric information about workers' productivity as a cause for inefficient long working hours', *Labour Economics* **10**(6), 727–747.
URL: *https://doi.org/10.1016/S0927-5371(03)00016-2*

Statistical Office of Baden-Wuerttemberg (2013), *Kinder und tätige Personen in Tageseinrichtungen in Baden-Württemberg am 1. März 2013*, Statistical Office of Baden-Wuerttemberg, Stuttgart.

URL: *https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/BWHeft_derivate_00004647/3833_13001.pdf*

Statistical Office of Baden-Wuerttemberg (2014), *Kinder und tätige Personen in Tageseinrichtungen in Baden-Württemberg am 1. März 2014*, Statistical Office of Baden-Wuerttemberg, Stuttgart.
URL: *https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/BWHeft_derivate_00004649/3833_14001.pdf*

Statistical Office of Bavaria (2013), *Kindertageseinrichtungen und Kindertagespflege in Bayern 2013*, Statistical Office of Bavaria, Munich.
URL: *https://www.statistik.bayern.de/mam/produkte/veroffentlichungen/statistische_berichte/k5300c_201300_16868.pdf*

Statistical Office of Bavaria (2014), *Kindertageseinrichtungen und Kindertagespflege in Bayern 2014*, Statistical Office of Bavaria, Munich.
URL: *https://www.statistik.bayern.de/mam/produkte/veroffentlichungen/statistische_berichte/k5300c_201400_64796.pdf*

Statistical Office of Hamburg and Schleswig-Holstein (2013), *Kinder in Tageseinrichtungen und öffentlich geförderter Kindertagespflege*, Statistical Office of Hamburg and Schleswig-Holstein, Hamburg.
URL: *https://www.statistik-nord.de/fileadmin/Dokumente/Statistische_Berichte/arbeit_und_soziales/K_I_3_j_t/K_I_3_j14_T3_HH.pdf*

Statistical Office of Hamburg and Schleswig-Holstein (2014), *Kinder in Tageseinrichtungen und öffentlich geförderter Kindertagespflege*, Statistical Office of Hamburg and Schleswig-Holstein, Hamburg.
URL: *https://www.statistik-nord.de/fileadmin/Dokumente/Statistische_Berichte/arbeit_und_soziales/K_I_3_j_t/K_I_3_j13_T3_Heft1_HH.pdf*

Statistical Office of Hesse (2013), *Kinder und tätige Personen in Tageseinrichtungen und Kindertagespflege in Hessen am 1. März 2013*, Statistical Office of Hesse, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/HEHeft_derivate_00002536/K%20V%207_J_13.pdf;jsessionid=2D7BA2ACDD9F3AC468EA3BC8280E275B*

Statistical Office of Hesse (2014), *Kinder und tätige Personen in Tageseinrichtungen und in Kindertagespflege am 1. März 2014*, Statistical Office of Hesse, Wiesbaden.
URL: *https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/HEHeft_derivate_00003611/KV7_j14.pdf;jsessionid=07EDA8B6C0D9CE3C45AC986B2C91A7EE*

Statistical Office of Lower Saxony (2013), *Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01. März 2013*, Statistical Office of Lower Saxony, Hannover.
URL: *https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/NIHeft_derivate_00001277/K_I_4_2013_pdfa.pdf*

Statistical Office of Lower Saxony (2014), *Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01. März 2014*, Statistical Office of Lower Saxony, Hannover.
URL: *https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/NIHeft_derivate_00004234/K_I_4_2014_pdfa.pdf*

Statistical Office of Mecklenburg-Vorpommern (2013), *Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege in Mecklenburg-Vorpommern*, Statistical Office of Mecklenburg-Vorpommern, Schwerin.
URL: *https://www.laiv-mv.de/static/LAIV/Statistisches%20Amt/Dateien/Publikationen/K%20V%20Kinder-%20und%20Jugendhilfe/K%20433/K433%202013%2000.pdf*

Statistical Office of Mecklenburg-Vorpommern (2014), *Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege in Mecklenburg-Vorpommern*, Statistical Office of Mecklenburg-Vorpommern, Schwerin.
URL: *https://www.laiv-mv.de/static/LAIV/Statistisches%20Amt/Dateien/Publikationen/K%20V%20Kinder-%20und%20Jugendhilfe/K%20433/K433%202014%2000.pdf*

Statistical Office of North Rhine-Westphalia (2013), *Kindertagesbetreuung in Nordrhein-Westfalen*, Statistical Office of North Rhine-Westphalia, Duesseldorf.
URL: *https://webshop.it.nrw.de/gratis/K239%20201300.pdf*

Statistical Office of North Rhine-Westphalia (2014), *Kindertagesbetreuung in Nordrhein-Westfalen*, Statistical Office of North Rhine-Westphalia, Duesseldorf.
URL: *https://webshop.it.nrw.de/gratis/K239%20201400.pdf*

Statistical Office of Saarland (2013), *Kinder- und Jugendhilfe2013. Teil III: Einrichtungen und tätige Personen*, Statistical Office of Saarland, Saarbruecken.
URL: *https://www.saarland.de/dokumente/thema_statistik/STALA_BER_KV1T3-J-13.pdf*

Statistical Office of Saarland (2014), *Kinder- und Jugendhilfe2014. Teil III: Einrichtungen und tätige Personen*, Statistical Office of Saarland, Saarbruecken.
URL: *https://www.saarland.de/dokumente/thema_statistik/STALA_BER_KV1T3-J-14.pdf*

Statistical Office of Saxony-Anhalt (2013), *Tageseinrichtungen für Kinder und öffentlich geförderte Kindertagespflege*, Statistical Office of Saxony-Anhalt, Halle.
URL: *https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/STHeft_derivate_00001614/6K504_2013.pdf*

Statistical Office of Saxony-Anhalt (2014), *Tageseinrichtungen für Kinder und öffentlich geförderte Kindertagespflege*, Statistical Office of Saxony-Anhalt, Halle.
URL: *https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/STHeft_derivate_00002015/6K504_2014.pdf*

Stewart, M. and Swaffield, J. (1997), 'Constraints on the desired hours of work of British men', *The Economic Journal* **107**(441), 520–535.
URL: *https://doi.org/10.1111/j.0013-0133.1997.175.x*

Stier, H. and Lewin-Epstein, N. (2003), 'Time to work: A comparative analysis of preferences for working hours', *Work and Occupations* **30**(3), 302–326.
URL: *https://doi.org/10.1177/0730888403253897*

Tamm, M. (2019), 'Fathers' parental leave-taking, childcare involvement and labor market participation', *Labour Economics* **59**, 184–197.
URL: *https://doi.org/10.1016/j.labeco.2019.04.007*

Treiman, D. (1976), 'A standard occupational prestige scale for use with historical data', *The Journal of Interdisciplinary History* **7**(2), 283–304.
URL: *https://www.jstor.org/stable/202737*

Wager, S. and Athey, S. (2018), 'Estimation and inference of heterogeneous treatment effects using random forests', *Journal of the American Statistical Association* **113**(523), 1228–1242.
URL: *https://doi.org/10.1080/01621459.2017.1319839*

Wager, S. and Walther, G. (2015), *Adaptive concentration of regression trees, with application to random forests*, Unpublished manuscript, Department of Statistics, Stanford University.
URL: *https://arxiv.org/abs/1503.06388*

Wagner, G., Frick, J. and Schupp, J. (2007), 'The German Socio-Economic Panel study (SOEP) - evolution, scope and enhancements', *Schmollers Jahrbuch* **127**(1), 139–169.
URL: *https://dx.doi.org/10.2139/ssrn.1028709*

Waldfogel, J. (1999), 'The impact of the family and medical leave act', *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management* **18**(2), 281–302.
URL: *https://doi.org/10.1002/(SICI)1520-6688(199921)18:2<281::AID-PAM5>3.0.CO;2-J*

Wanger, S. (2015), 'Frauen und Männer am Arbeitsmarkt: Traditionelle Erwerbs- und Arbeitszeitmuster sind nach wie vor verbreitet', *IAB-Kurzbericht* **18/2015**.
URL: *http://hdl.handle.net/10419/158455*

Welteke, C. and Wrohlich, K. (2019), 'Peer effects in parental leave decisions', *Labour Economics* **57**, 146–163.
URL: *https://doi.org/10.1016/j.labeco.2019.02.008*

Willett, J. B. and Singer, J. D. (1995), 'It's déjà vu all over again: Using multiple-spell discrete-time survival analysis', *Journal of Educational and Behavioral Statistics* **20**(1), 41–67.
URL: *https://doi.org/10.3102/10769986020001041*

Wunder, C. and Heineck, G. (2013), 'Working time preferences, hours mismatch and well-being of couples: Are there spillovers?', *Labour Economics* **24**, 244–252.
URL: *https://doi.org/10.1016/j.labeco.2013.09.002*

Yamaguchi, S., Asai, Y. and Kambayashi, R. (2018), 'Effects of subsidized childcare on mothers' labor supply under a rationing mechanism', *Labour Economics* **55**, 1–17.
URL: *https://doi.org/10.1016/j.labeco.2018.09.002*

Zapf, I. and Weber, E. (2017), 'The role of employer, job and employee characteristics for flexible working time: An empirical analysis of overtime work and flexible working hours' arrangements', *IAB-Discussion Paper* **4/2017**.
URL: *http://hdl.handle.net/10419/162475*

Zimmert, M. (2018), *Efficient Difference-in-Differences Estimation with High-Dimensional Common Trend Confounding*, Unpublished manuscript, Swiss Institute for Empirical Economic Research, University of St. Gallen.
URL: *https://arxiv.org/abs/1809.01643*

Zimmert, M. and Lechner, M. (2019), *Nonparametric estimation of causal heterogeneity under high-dimensional confounding*, Unpublished manuscript, Swiss Institute for Empirical Economic Research, University of St. Gallen.
URL: *https://arxiv.org/abs/1908.08779*

Zoch, G. and Hondralis, I. (2017), 'The expansion of low-cost, state-subsidized childcare availability and mothers' return-to-work behaviour in East and West Germany', *European Sociological Review* **33**(5), 693–707.
URL: *https://doi.org/10.1093/esr/jcx068*

Earlier versions of the articles exist as:

- Chapter 2: Weber, E. and Zimmert, F. (2017), 'The creation and resolution of working hour discrepancies over the life course', *IAB-Discussion Paper* **29/2017**.
- Chapter 3: Zimmert, F. (2019), 'Early child care and maternal employment: empirical evidence from Germany', *IAB-Discussion Paper* **2/2019**.

Appendix A

Tables

A.1 Chapter 2

TABLE A.1: Distribution of individual and partner's characteristics over discrepancy types

	Women			Men		
	nd	ue	oe	nd	ue	oe
N	39,920	16,046	39,839	46,598	9,801	59,228
%	41.67	16.75	41.58	40.30	8.48	51.22
Personal characteristics						
Career stage %						
Learning stage	10.03	8.22	9.10	11.13	17.30	7.03
Career start	7.01	5.93	8.60	6.71	8.86	6.22
Establishing	9.32	8.82	11.62	9.94	10.85	11.23
Middle phase up to 45 years	33.76	39.90	31.90	32.32	30.59	33.63
Middle phase up to 55 years	26.31	27.55	27.35	25.25	20.54	28.03
Pre-retirement	12.44	8.87	10.99	13.05	10.16	13.15
Retirement	1.12	0.71	0.43	1.59	1.70	0.71
Children %						
No children	67.45	53.60	72.75	65.08	66.23	62.11
Children le6	11.68	16.91	9.06	17.06	18.12	18.63
Children 7-10	9.61	13.07	7.83	8.41	7.17	9.18
Children 11-15	11.26	16.42	10.36	9.46	8.48	10.08
Child care facility %						
No facility	39.12	35.07	38.11	50.15	49.67	47.27
Part-time facility	48.60	53.71	45.41	40.99	41.03	43.25
Full-time facility	12.27	11.23	16.48	8.86	9.31	9.48
Daily hours for child care *Mean*	4.47	5.45	3.88	1.59	1.89	1.36
Daily hours for	2.10	2.48	1.72	0.66	0.81	0.60

housekeeping *Mean*						
Education and job %						
No vocational degree	23.21	19.65	15.16	21.22	25.76	12.46
Vocational training	59.73	64.15	57.33	61.45	56.99	59.68
University degree	17.06	16.20	27.50	17.33	17.24	27.86
Occupational autonomy %						
Apprenticeship	5.39	1.98	4.56	5.97	6.10	3.05
Low autonomy (= 1)	18.80	23.82	9.64	19.03	23.90	10.11
2	25.01	34.08	18.91	29.61	28.64	23.28
3	39.20	33.30	42.23	23.28	25.04	22.81
4	9.87	6.09	21.27	18.73	14.35	33.35
High autonomy (= 5)	1.73	0.73	3.39	3.38	1.98	7.41
Gross wage %						
$\leq$ 1st percentile	10.94	13.91	9.57	7.38	12.44	5.42
$\leq$ median	53.45	60.09	43.72	32.69	39.22	29.84
$\leq$ 9th percentile	31.02	21.05	38.82	48.96	38.34	45.77
$>$ 9th percentile	4.59	4.95	7.89	10.97	10.00	18.97
Partner						
Partner %						
No partner	29.80	28.33	30.89	26.99	32.48	21.52
Inactive, currently in education	10.33	8.28	9.28	25.15	24.59	24.21
Active	59.87	63.39	59.83	47.86	42.93	54.27
Occupational autonomy %						
Low autonomy (= 1)	14.51	14.70	10.47	22.61	20.01	13.03
2	25.31	27.14	23.24	25.49	25.81	23.49
3	24.82	25.01	24.96	37.66	39.54	43.35
4	28.96	28.15	33.82	12.19	13.07	17.19
High autonomy (= 5)	6.41	5.00	7.51	2.04	1.57	2.93
Daily hours for child care *Mean*	1.52	1.51	1.69	5.91	6.22	6.10
Daily hours for housekeeping *Mean*	0.77	0.69	0.87	3.08	3.02	2.92

Notes: nd=no discrepancy, ue=underemployed, oe=overemployed.

Source: Own calculations based on GSOEP v33.1, 1985-2016. Pooled analysis.

TABLE A.2: Estimation results for the creation of a discrepancy - Additional covariates

	Women		Men		Women		Men	
	Underemployment				Overemployment			
Institutional child care (*Reference* Full-time slot)								
No attendance	2.575	(0.86)	1.163	(0.33)	1.222	(0.63)	0.995	(-0.02)
Part-time slot	4.731	(1.30)	1.409	(0.68)	1.469	(1.19)	1.434	(1.28)
Children le6 * Period of child care expansion (*Reference* Children le6 * 1985-1995)								
Children le6 * 1996-2005	2.971	(1.41)	2.225	(1.20)	0.823	(-0.43)	0.921	(-0.32)
Children le6 * 2006-2008	*3.412*	(1.55)	2.001	(0.86)	*0.513*	(-1.23)	0.993	(-0.02)
Children le6 * 2009-2016	2.015	(0.94)	1.790	(0.83)	0.627	(-1.08)	0.801	(-0.81)
Unpaid working hours								
Child care	0.961	(-1.25)	1.052	(0.41)	0.879***	(-3.57)	1.033	(0.62)
Housework	0.962	(-0.48)	*1.233*	(1.09)	0.903*	(-1.69)	*0.867*	(-1.59)
More than 20 employees * After introduction of legal claim for a part-time job								
	0.628	(-1.48)	1.054	(0.13)	0.936	(-0.33)	0.993	(-0.05)
N	7,303		5,121		15,884		23,343	
n	1,895		1,335		3,950		5,454	

Notes: Exponentiated coefficients (odds ratios) of fixed effects-logit estimation. Instead of providing marginal effects, odds ratios are indicated as they do not require plugging in a value for the unobserved component. The odds ratio gives the multiplicative value for the odds if the explanatory variable increases by one unit. t-values in parentheses. Standard errors are bootstrapped with 1,000 replications. Other than listed explanatory variables are mentioned in Section 2.3.3. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.
Abbreviations: Children le6 means younger than 7 years old.
Source: Own calculations based on GSOEP v33.1, 1985-2016.

TABLE A.3: Estimation results for the resolution of a discrepancy - Additional covariates

	Women		Men		Women		Men	
	Underemployment				Overemployment			
Institutional child care (*Reference* Full-time slot)								
No attendance	1.356	(0.39)	0.555	(-1.07)	1.183	(0.43)	1.092	(0.33)
Part-time slot	1.377	(0.44)	0.431	(-1.43)	0.767	(-0.66)	1.295	(0.90)
Children le6 * Period of child care expansion (*Reference* Children le6 * 1985-1995)								
Children le6 * 1996-2005	2.453	(1.12)	0.380	(-1.42)	1.555	(1.03)	1.065	(0.24)
Children le6 * 2006-2008	2.824	(1.22)	0.642	(-0.56)	1.343	(0.56)	0.771	(-0.86)
Children le6 * 2009-2016	1.487	(0.50)	0.358	(-1.51)	0.798	(-0.52)	1.028	(0.10)
Unpaid working hours								
Child care	0.963	(-1.06)	1.081	(0.74)	0.934*	(-1.95)	1.005	(0.08)
Housework	0.924	(-1.03)	0.836	(-0.99)	0.882**	(-1.98)	1.057	(0.61)
More than 20 employees * After introduction of legal claim for a part-time job								
	1.131	(0.45)	1.460	(0.99)	1.033	(0.16)	0.882	(-0.81)
N	7,545		5,333		15,964		25,041	
n	2,165		1,702		3,493		4,894	

Notes: Exponentiated coefficients (odds ratios) of fixed effects-logit estimation. Instead of providing marginal effects, odds ratios are indicated as they do not require plugging in a value for the unobserved component. The odds ratio gives the multiplicative value for the odds if the explanatory variable increases by one unit. t-values in parentheses. Standard errors are bootstrapped with 1,000 replications. Other than listed explanatory variables are mentioned in Section 2.3.3. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.
Abbreviations: Children le6 means younger than 7 years old.
Source: Own calculations based on GSOEP v33.1, 1985-2016.

A.2 Chapter 3

TABLE A.4: Results of main estimation - Average effects of additional outcomes

	Baseline	West Germany	Without under 1-year-olds	Without childminders	Without families having moved
Underemployment (binary)					
ATET	0.024	0.097*	-0.002	0.020	0.050
se	(0.042)	(0.051)	(0.041)	(0.041)	(0.042)
Overemployment (binary)					
ATET	0.001	0.010	0.002	2.7E-05	0.010
se	(0.012)	(0.012)	(0.013)	(0.013)	(0.011)
N	3,492	3,183	2,991	3,428	3,177

Standard errors (in columns) are bootstrapped with 1,000 replications considering clusters on the district level. The sample includes 18 to 45 years old mothers of up to three-year-olds. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.
Source: Own calculations based on data from the German Microcensus and the Federal Statistical Office (2011*b*, 2015*b*).

A.3 Chapter 4

TABLE A.5: Conditional effects for job continuity and accumulated earnings

Outcome	$D = 1, T = 0$ Mean	sd	AIPW coefficient	se
Panel A: Middle income				
Same employer	0.426	0.495	0.021	0.016
Accumulated earnings				
1st year	1,049.87	2,942.71	304.47*	179.72
2nd year	6,981.35	8,791.76	244.00	481.68
$N = 32,139$. Middle income corresponds to 2nd tercile of gross previous monthly income. Reference group is low income.				
Panel B: High income				
Same employer	0.563	0.496	0.022	0.017
Accumulated earnings				
1st year	1,382.94	4,359.20	473.79**	194.41
2nd year	11,984.46	12,674.35	342.07	521.10
$N = 31,166$. High income corresponds to 3rd tercile of gross previous monthly income. Reference group is low income.				
Panel C: Prior full-time job				
Same employer	0.563	0.496	-0.012	0.014
Accumulated earnings				
1st year	1,933.89	6,089.10	-271.13*	164.59
2nd year	14,687.05	17,480.42	-680.89	441.14
$N = 60,913$. Reference group is prior part-time or marginal employment.				
Panel D: West Germany				
Same employer	0.534	0.499	-0.031**	0.015
Accumulated earnings				
1st year	1,728.85	5,622.98	128.28	169.09
2nd year	11,834.38	16,158.99	-335.38	453.24
$N = 72,508$. Reference group is East Germany.				

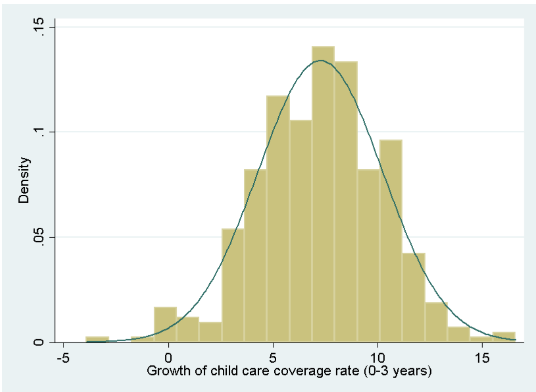
Note: Treatment status is measured six weeks around cut-off date excluding the two weeks on each side of the cut-off date. $T = 1$ in 2015, $T = 0$ in 2014. The coefficients give the effect variation for the respective subgroup. $^{*}p < 0.1$, $^{**}p < 0.05$.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

Appendix B

Figures

B.1 Chapter 3

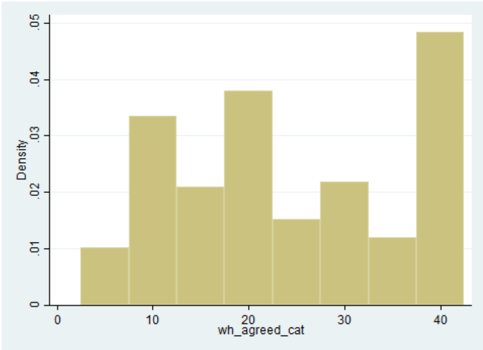
FIGURE B.1: Histogram of child care coverage growth from 2011 to 2015.



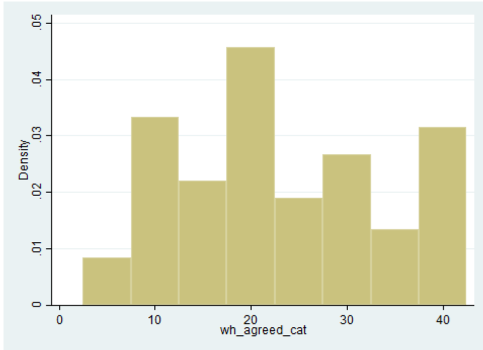
Notes: Unweighted calculations based on 317 districts.
Source: Own calculations based on data from the Federal Statistical Office (2011*b*, 2015*b*).

FIGURE B.2: Distribution of agreed working hours by group status

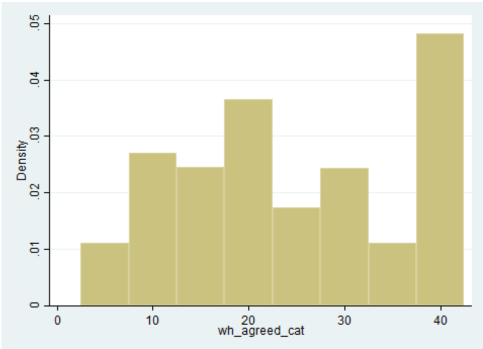
(a) Control group before reform



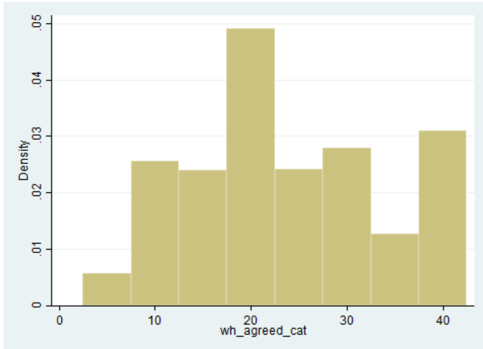
(b) Control group after reform



(c) Treated group before reform



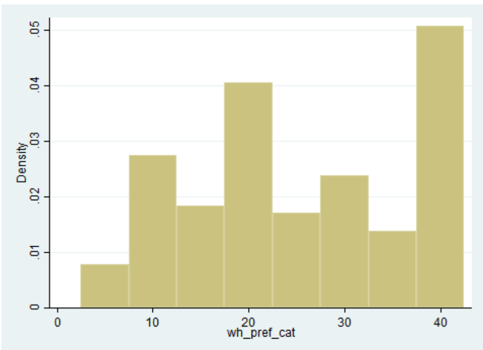
(d) Treated group after reform



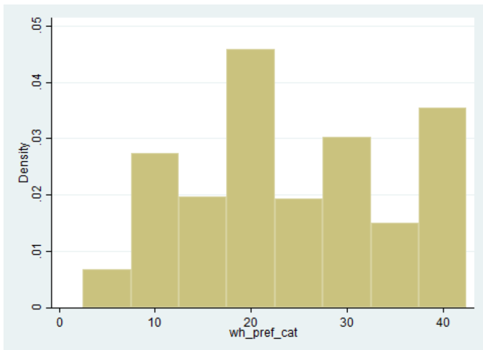
Source: Own calculations based on data from the German Microcensus and the Federal Statistical Office (2011*b*, 2015*b*).

FIGURE B.3: Distribution of preferred working hours by group status

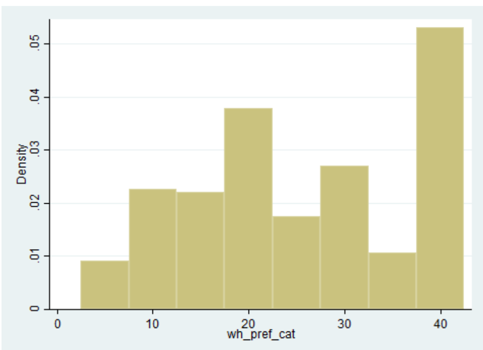
(a) Control group before reform



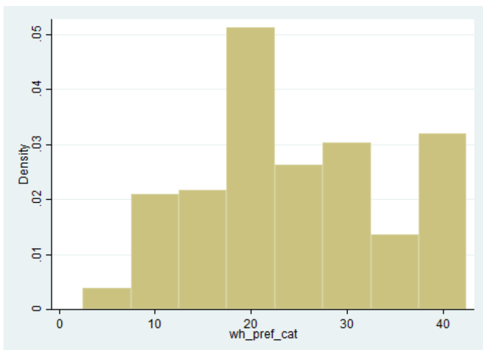
(b) Control group after reform



(c) Treated group before reform



(d) Treated group after reform

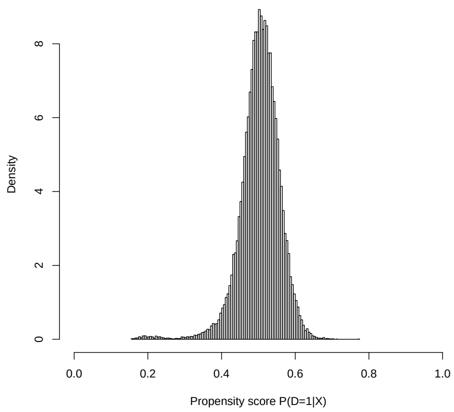


Source: Own calculations based on data from the German Microcensus and the Federal Statistical Office (2011*b*, 2015*b*).

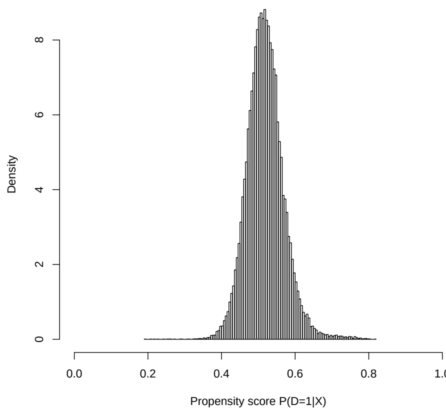
B.2 Chapter 4

FIGURE B.4: Propensity scores by treatment status

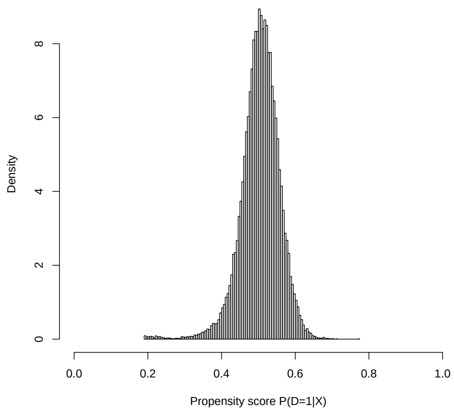
(a) $D = 0$ before trimming



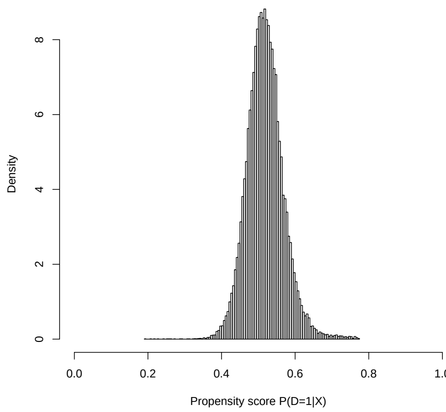
(b) $D = 1$ before trimming



(c) $D = 0$ after trimming



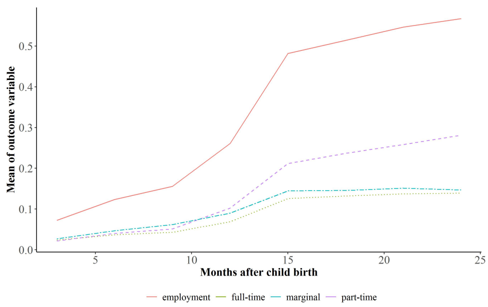
(d) $D = 1$ after trimming



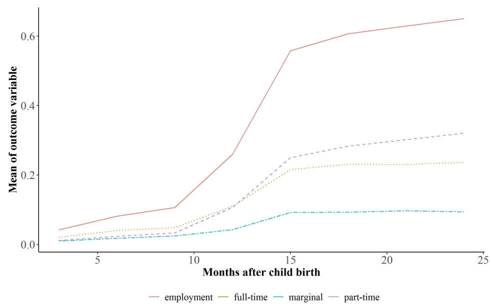
Notes: $N_{D=0} = 46,263$ and $N_{D=1} = 48,212$ before trimming; $N_{D=0} = 46,191$ and $N_{D=1} = 48,184$ after trimming.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

FIGURE B.5: Outcome means of treated mothers before reform (1)

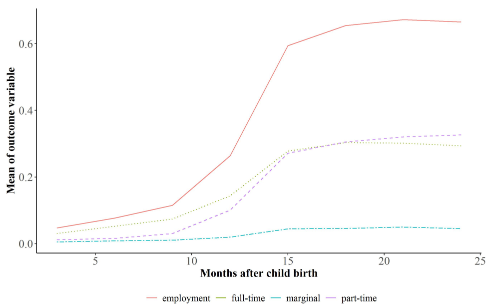
(a) Low income



(b) Middle income

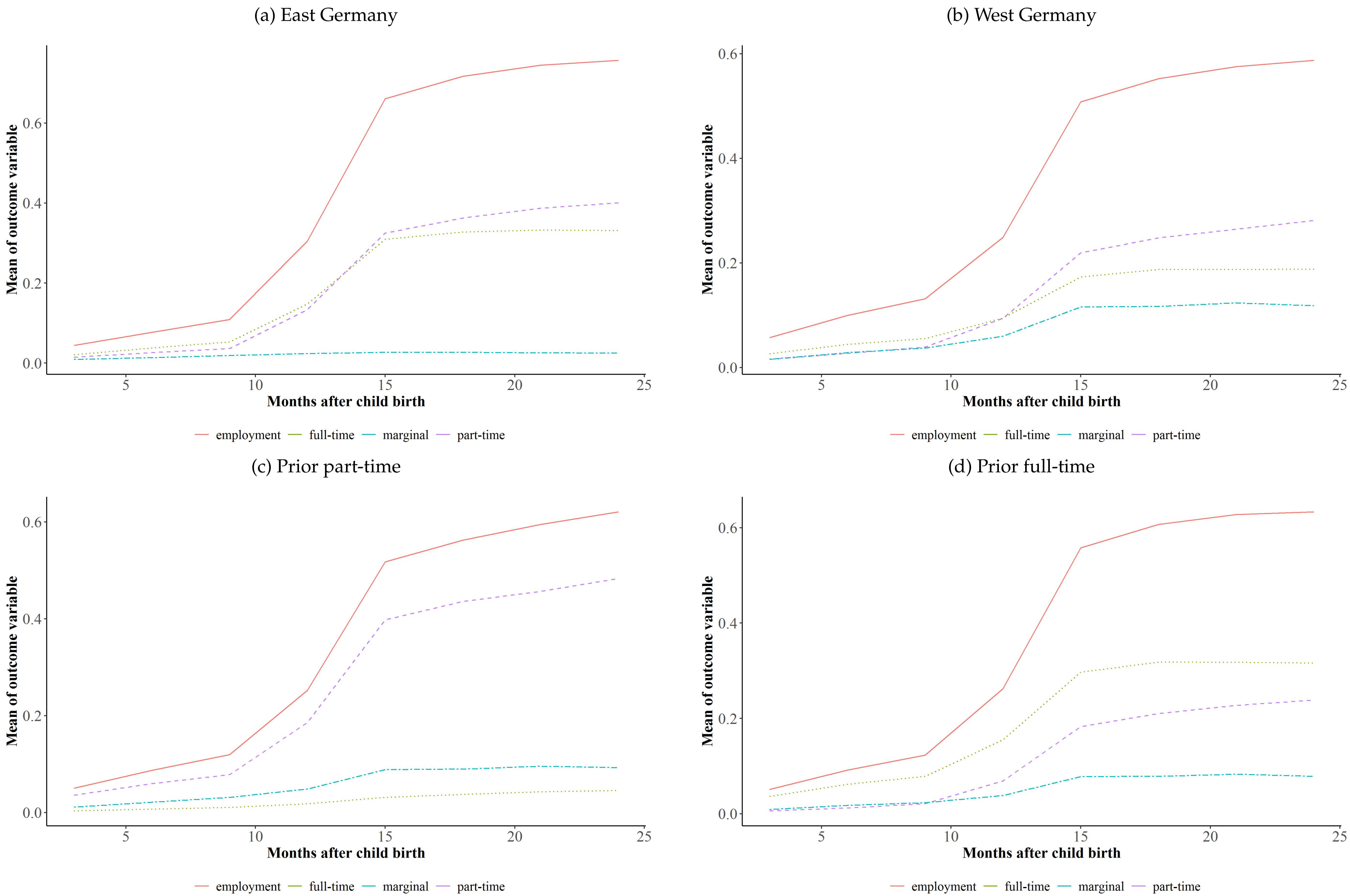


(c) High income



Notes: $T = 0$ in 2014. $N_{low\_income} = 8,198$, $N_{middle\_income} = 8,288$ and $N_{high\_income} = 8,367$.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

FIGURE B.6: Outcome means of treated mothers before reform (2)



Notes:

$T = 0$ in 2014. $N_{east} = 5,524$, $N_{west} = 18,695$, $N_{part} = 7,608$ and $N_{full} = 15,530$.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

Appendix C

Mathemetical appendix

C.1 Chapter 4

C.1.1 Model details

Model set up

The decision problem of the mother in parental leave whether to stay out of the labour force or to accept a part-time or full-time job offer can for any $t > 0$ be fully described by the following Bellman equation:

$$V_{T-t-1}^{pl} = \lambda \bar{y} + l + \frac{1}{1+\rho} \int_0^\infty \max_{pl,f,p}(V_{T-t}^{pl}, V_{T-t}^{f}, V_{T-t}^{p}) dF(y_{T-t}) \tag{C.1}$$

where we are agnostic about the particular form of the cumulative distribution function of the job offer incomes $F(y_{T-t})$.

Reservation income

We first of all notice that for any $t \geq 0$ the value functions for states f and p are given by

$$\begin{aligned} V_{T-t}^{f} &= y_{T-t}\frac{1+\rho}{\rho} \text{ and} \\ V_{T-t}^{p} &= (\beta y_{T-t} + (1-\beta)l + D(\rho,t)\tau(\bar{y} - \beta y_{T-t}))^{+}\,\frac{1+\rho}{\rho} \end{aligned}$$

where $D(\rho, t) = 1 - \left(\frac{1}{1+\rho}\right)^{t+1}$. Second of all, for period T we find

$$\begin{aligned} &V_T^{pl} = \lambda \bar{y} + l + \frac{1}{1+\rho} \int_0^\infty \max(V_{T+1}^{u}, V_{T+1}^{f}, V_{T+1}^{p}) dF(y_{T+1}) \text{ such that} \\ &V_T^{pl} - \lambda \bar{y} - l - \frac{b^u}{\rho} = \\ &\int_0^\infty \max\left(0, \frac{1}{\rho}(y_{T+1} - b^u), \frac{1}{\rho}(\beta y_{T+1} + (1-\beta)l - b^u)\right) dF(y_{T+1}) \end{aligned} \tag{C.2}$$

for inserting the infinite series in Equations 4.1.
Moreover, we can explicitly solve for the value function V_T^{pl} as

$$\begin{aligned}
V_T^{pl} &= \lambda\bar{y} + l + \frac{b^u}{\rho} \\
&+ \frac{1}{\rho}P(y_{T+1} > l, y_{T+1} > b^u)\left(\mathbb{E}(y_{T+1}|y_{T+1} > l, y_{T+1} > b^u) - b^u\right) \\
&+ \frac{1}{\rho}P\left(y_{T+1} < l, y_{T+1} > \frac{b^u}{\beta} - \frac{1-\beta}{\beta}l\right) \\
&* \left(\beta\mathbb{E}\left(y_{T+1}|y_{T+1} < l, y_{T+1} > \frac{b^u}{\beta} - \frac{1-\beta}{\beta}l\right) + (1-\beta)l - b^u\right). \qquad \text{(C.3)}
\end{aligned}$$

Thus, iterating backwards will give an explicit solution for every V_{T-t}^{pl} in the model. In the non-stationary environment for any period $t \geq 0$ the reservation income will decline compared to the pre-period until it reaches the stationary solution in $T+1$ as given above. We therefore derive an implicit solution for the reservation income in the non-stationary environment. In particular, we have

$$\begin{aligned}
&V_{T-t-1}^{pl} - \lambda\bar{y} - l - \frac{1}{1+\rho}V_{T-t}^{pl} = \\
&\int_0^\infty \max\left(0, \frac{y_{T-t}}{\rho} - \frac{1}{1+\rho}V_{T-t}^{pl}, \frac{\beta y_{T-t} + (1-\beta)l}{\rho}\right. \\
&\left. + \frac{D(\rho,t)\tau(\bar{y} - \beta y_{T-t})^+}{\rho} - \frac{1}{1+\rho}V_{T-t}^{pl}\right) dF(y_{T-t}) \qquad \text{(C.4)}
\end{aligned}$$

such that we obtain the results of Equation 4.3.

Duration effects

The derivative of $D(\rho, t)$ with respect to t can be written as

$$\frac{\partial D(\rho,t)}{\partial t} = \ln(1+\rho)\left(\frac{1}{1+\rho}\right)^{t+1} > 0. \qquad \text{(C.5)}$$

C.1.2 Identification of ATET and CATET

Identification of ATET

The following identification result is taken from Zimmert (2018). It is given here for convenience.

We can write

$$
\begin{aligned}
&\mathbb{E}\left[\frac{T-\lambda_T}{\lambda_T(1-\lambda_T)}\frac{D-p(X)}{p(X)(1-p(X))}\left(Y-\gamma(X,T)\right)\Big|X\right] \\
&=\mathbb{E}\left[\mathbb{E}\left[\frac{T-\lambda_T}{\lambda_T(1-\lambda_T)}\frac{D-p(X)}{p(X)(1-p(X))}\left(Y-\gamma(X,T)\right)\Big|X,T\right]\Big|X\right] \\
&=\mathbb{E}\Bigg[\mathbb{E}\left[\frac{T-\lambda_T}{\lambda_T(1-\lambda_T)}\frac{D-p(X)}{p(X)(1-p(X))}\left(Y-\gamma(X,T)\right)\Big|X,T=1\right]P(T=1|X) \\
&+\mathbb{E}\left[\frac{T-\lambda_T}{\lambda_T(1-\lambda_T)}\frac{D-p(X)}{p(X)(1-p(X))}\left(Y-\gamma(X,T)\right)\Big|X,T=0\right](1-P(T=1|X))\Big|X\Bigg] \\
&=\mathbb{E}\left[\frac{D-p(X)}{p(X)(1-p(X))}\left(Y_1-Y_0-\mathbb{E}\left[Y_1-Y_0|X,D=0\right]\right)\Big|X\right] \\
&=\mathbb{E}\left[\frac{D-p(X)}{p(X)(1-p(X))}\left(Y_1-Y_0\right)\Big|X\right]-\mathbb{E}\left[\frac{D-p(X)}{p(X)(1-p(X))}\mathbb{E}\left[Y_1-Y_0|X,D=0\right]\Big|X\right] \\
&=\mathbb{E}\left[\frac{D-p(X)}{p(X)(1-p(X))}\left(Y_1-Y_0\right)\Big|X,D=1\right]p(X) \\
&+\mathbb{E}\left[\frac{D-p(X)}{p(X)(1-p(X))}\left(Y_1-Y_0\right)\Big|X,D=0\right](1-p(X)) \\
&-\mathbb{E}\left[D-p(X)|X\right]\frac{\mathbb{E}\left[Y_1-Y_0|X,D=0\right]}{p(X)(1-p(X))} \\
&=\mathbb{E}\left[Y_1-Y_0|X,D=1\right]-\mathbb{E}\left[Y_1-Y_0|X,D=0\right]
\end{aligned}
$$

where the third equality follows the fact that $P(T=t|X)=\lambda_T$ and by the Observational Rule assumed. The existence of the expectation is guaranteed by the Common Support condition.
Also analogous to the fundamental result of Heckman et al. (1997) we have

$$
\begin{aligned}
\mathbb{E}\left[Y_1^1-Y_1^0|X,D=1\right]&=\mathbb{E}\left[Y_1|X,D=1\right]-\mathbb{E}\left[Y_1^0|X,D=1\right] \\
&=\mathbb{E}\left[Y_1|X,D=1\right]-\mathbb{E}\left[Y_1^0-Y_0^0|X,D=0\right]-\mathbb{E}\left[Y_0^0|X,D=1\right] \\
&=\mathbb{E}\left[Y_1-Y_0|X,D=1\right]-\mathbb{E}\left[Y_1-Y_0|X,D=0\right]
\end{aligned}
$$

which follows by the Observational Rule, the No Anticipation and the Common Trend assumptions. Therefore,

$$
\mathbb{E}\left[Y_1^1-Y_1^0|X,D=1\right]=\mathbb{E}\left[\frac{T-\lambda_T}{\lambda_T(1-\lambda_T)}\frac{D-p(X)}{p(X)(1-p(X))}\left(Y-\gamma(X,T)\right)\Big|X\right].
$$

Denote the conditional density function of X given $D=1$ as $f_{X|D=1}(x,d)$. Then using the previous finding and by the law of iterated expectations similar to Abadie

(2005), it follows that

$$
\begin{aligned}
\text{ATET}(1) &= \mathbb{E}\left[Y_1^1 - Y_1^0 | D = 1\right] \\
&= \int \mathbb{E}\left[Y_1^1 - Y_1^0 | X, D = 1\right] f_{X|D=1}(x, d) dx \\
&= \int \mathbb{E}\left[Y_1^1 - Y_1^0 | X, D = 1\right] \frac{p(X)}{\lambda_D} f_X(x) dx \\
&= \frac{1}{\lambda_D} \mathbb{E}\left[\frac{T - \lambda_T}{\lambda_T(1-\lambda_T)} \frac{D - p(X)}{1 - p(X)} (Y - \gamma(X, T))\right]. \qquad \textit{q.e.d.}
\end{aligned}
$$

Identification of CATET

From the identification of the ATET we know that

$$
\mathbb{E}\left[Y_1^1 - Y_1^0 | X, D = 1\right] = \mathbb{E}\left[\frac{T - \lambda_T}{\lambda_T(1-\lambda_T)} \frac{D - p(X)}{p(X)(1 - p(X))} (Y - \gamma(X, T)) \Big| X\right].
$$

For the CATET we therefore obtain

$$
\begin{aligned}
& \mathbb{E}\left[Y_1^1 - Y_1^0 | Z = z, D = 1\right] \\
=& \mathbb{E}\left[\mathbb{E}\left[Y_1^1 - Y_1^0 | X, Z = z, D = 1\right] | Z = z, D = 1\right] \\
=& \mathbb{E}\left[\mathbb{E}\left[Y_1^1 - Y_1^0 | X, D = 1\right] | Z = z, D = 1\right] \\
=& \mathbb{E}\left[\mathbb{E}\left[\frac{T - \lambda_T}{\lambda_T(1-\lambda_T)} \frac{D - p(X)}{p(X)(1 - p(X))} (Y - \gamma(X, T)) \Big| X\right] \Big| Z = z, D = 1\right] \\
=& \int_X \mathbb{E}\left[\frac{T - \lambda_T}{\lambda_T(1-\lambda_T)} \frac{D - p(X)}{p(X)(1 - p(X))} (Y - \gamma(X, T)) \Big| X\right] f_{X|Z,D=1}(x, z, d) dx \\
=& \int_X \mathbb{E}\left[\frac{T - \lambda_T}{\lambda_T(1-\lambda_T)} \frac{D - p(X)}{p(X)(1 - p(X))} (Y - \gamma(X, T)) \Big| X\right] \frac{f_{X,Z|D=1}(x, z, d)}{f_{Z|D=1}(z, d)} dx \\
=& \int_X \mathbb{E}\left[\frac{T - \lambda_T}{\lambda_T(1-\lambda_T)} \frac{D - p(X)}{p(X)(1 - p(X))} (Y - \gamma(X, T)) \Big| X\right] \frac{f_{D=1|X,Z}(x, z, d) f_{X,Z}(x, z)}{f_{D=1|Z}(z, d) f_Z(z)} dx \\
=& \int_X \mathbb{E}\left[\frac{T - \lambda_T}{\lambda_T(1-\lambda_T)} \frac{D - p(X)}{p(X)(1 - p(X))} (Y - \gamma(X, T)) \Big| X\right] \frac{p(X)}{p(Z)} f_{X|Z}(x, z) dx \\
=& \mathbb{E}\left[\mathbb{E}\left[\frac{T - \lambda_T}{\lambda_T(1-\lambda_T)} \frac{D - p(X)}{1 - p(X)} (Y - \gamma(X, T)) \Big| X\right] \frac{1}{p(Z)} \Big| Z = z\right] \\
=& \mathbb{E}\left[\frac{1}{p(Z)} \frac{T - \lambda_T}{\lambda_T(1-\lambda_T)} \frac{D - p(X)}{1 - p(X)} (Y - \gamma(X, T)) \Big| Z = z\right]
\end{aligned}
$$

which exists under the additional assumption that $p(Z) > 0$. *q.e.d.*