Universität St.Gallen

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

Franziska Zimmert, Michael Zimmert

January 2020 Discussion Paper no. 2020-02

School of Economics and Political Science, Department of Economics

University of St.Gallen

Editor:	Vanessa Pischulti University of St.Gallen School of Economics and Political Science Department of Economics Müller-Friedberg-Strasse 6/8 CH-9000 St.Gallen Phone +41 71 224 23 07 Email seps@unisg.ch
Publisher:	School of Economics and Political Science Department of Economics University of St.Gallen Müller-Friedberg-Strasse 6/8 CH-9000 St.Gallen Phone +41 71 224 23 07
Electronic Publication:	http://www.seps.unisg.ch

Paid parental leave and maternal reemployment:

Do part-time subsidies help or harm?[1]

Franziska Zimmert, Michael Zimmert

Author's address:

Franziska Zimmert
Institute for Employment Research (IAB)
Regensburger Str. 104
DE-90478 Nürnberg
Email franziska.zimmert@iab.de

Michael Zimmert
Swiss Institute for Empirical Economic Research (SEW)
University of St.Gallen
Varnbüelstrasse 14
CH-9000 St.Gallen
Email michael.zimmert@unisg.ch

[1] 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).

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

JEL Classification

J21, J22, C14

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 and 5

explain the estimation strategy and the exploited data before presenting the estimation results and sensitivity checks in Section 6. The article ends with a discussion of the results and a conclusion.

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 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).

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

Former regulations (see second column in Table 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 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 1a):

Table 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 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).

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

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.

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

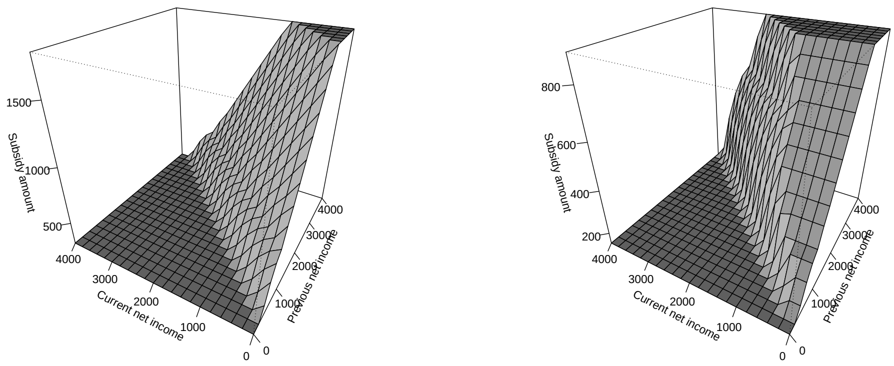
The new regulations coming into effect for births from July 2015 onwards double the maximum entitlement period to 24 months while receiving up to half of the basic benefit amount (compare Figure 1b):

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

Table 2 gives several examples for the calculation of the subsidy under the new regime. Besides, the model presented in Section 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

Figure 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.

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 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.

[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.

Table 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).

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 Zimmert, 2019). 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, 2015b). 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 Statistical Office, 2015b). 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.

3 Theoretical effects of part-time subsidies

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

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

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

a legal claim to return to their previous employer according 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 $T+1$

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

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 A.2.

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

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

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 $y^{*INT}_{T+t} = l$.
Moreover, we can explicitly solve for the value function V^{pl}_T and iterating backwards will give an explicit solution for every V^{pl}_{T-t} 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 (2). We therefore derive an implicit solution for the reservation income in the non-stationary environment. In particular, we obtain

$$y^{*EXT}_{T-t} = \begin{cases} \frac{\rho}{1+\rho} V^{pl}_{T-t} & \text{if} \quad y > \frac{\bar{y}}{\beta} \quad \text{and} \quad y > l \\ \frac{1}{\beta}\left(\frac{\rho}{1+\rho} V^{pl}_{T-t} - (1-\beta) l\right) & \text{if} \quad y > \frac{\bar{y}}{\beta} \quad \text{and} \quad y < l \\ \frac{\rho}{1+\rho} V^{pl}_{T-t} & \text{if} \quad y < \frac{\bar{y}}{\beta} \quad \text{and} \\ & y > l + \frac{1}{1-\beta} D(\rho,t)\tau(\bar{y} - \beta y) \\ \frac{1}{\beta(1-\tau D(\rho,t))} * & \text{if} \quad y < \frac{\bar{y}}{\beta} \quad \text{and} \\ \left(\frac{\rho}{1+\rho} V^{pl}_{T-t} - (1-\beta) l - D(\rho,t)\tau\bar{y}\right) & y < l + \frac{1}{1-\beta} D(\rho,t)\tau(\bar{y} - \beta y) \end{cases} \tag{3}$$

with $D(\rho, t) = 1 - \left(\frac{1}{1+\rho}\right)^{t+1}$. The first two cases of (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 $\tau(\bar{y} - \beta y_{T-t})^+ =$

[4] We get these results by (14) of Appendix A.2.

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

$$y^{*INT}_{T-t} = \begin{cases} l & \text{if} \quad y > \frac{\bar{y}}{\beta} \\ \frac{1}{1-\beta(1-D(\rho,t)\tau)}\left((1-\beta)l + D(\rho,t)\tau\bar{y}\right) & \text{if} \quad y < \frac{\bar{y}}{\beta}. \end{cases} \tag{4}$$

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

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

[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.

1-3. For the fourth case we obtain

$$\frac{\partial y^{*EXT}_{T-t}}{\partial t} = \frac{1}{\beta(1-\tau D(\rho,t))}\left(\frac{\rho}{1+\rho}\underbrace{\frac{\partial V^{pl}_{T-t}}{\partial t}}_{>0} - \tau\underbrace{\frac{\partial D(\rho,t)}{\partial t}}_{>0}\underbrace{(\bar{y}-\beta y^{*EXT}_{T-t})}_{>0}\right) \tag{5}$$

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

$$\frac{\partial y^{*INT}_{T-t}}{\partial t} = \frac{\tau}{1-\beta(1-D(\rho,t)\tau)}\underbrace{\frac{\partial D(\rho,t)}{\partial t}}_{>0}\underbrace{(\bar{y}-\beta y^{*INT}_{T-t})}_{>0} > 0. \tag{6}$$

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 (3) follow in a straightforward manner.

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:

$$dy^{*EXT}_{T-t} = \begin{cases} \frac{\rho}{1+\rho}\left(\frac{\partial V^{pl}_{T-t}}{\partial t}dt + \frac{\partial V^{pl}_{T-t}}{\partial \lambda}d\lambda\right) & \text{if} \quad y > \frac{\bar{y}}{\beta} \quad \text{and} \quad y > l \\ \frac{1}{\beta}\frac{\rho}{1+\rho}\left(\frac{\partial V^{pl}_{T-t}}{\partial t}dt + \frac{\partial V^{pl}_{T-t}}{\partial \lambda}d\lambda\right) & \text{if} \quad y > \frac{\bar{y}}{\beta} \quad \text{and} \quad y < l \\ \frac{\rho}{1+\rho}\left(\frac{\partial V^{pl}_{T-t}}{\partial t}dt + \frac{\partial V^{pl}_{T-t}}{\partial \lambda}d\lambda\right) & \text{if} \quad y < \frac{\bar{y}}{\beta} \quad \text{and} \\ & y > l + \frac{1}{1-\beta}D(\rho,t)\tau(\bar{y}-\beta y). \end{cases} \tag{7}$$

Since $\frac{\partial V_{T-t}^{pl}}{\partial t}dt > 0$ and $\frac{\partial V_{T-t}^{pl}}{\partial \lambda}d\lambda < 0$, 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 $\frac{\partial V_{T-t}^{pl}}{\partial t}dt + \frac{\partial V_{T-t}^{pl}}{\partial \lambda}d\lambda > 0$ 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

$$dy_{T-t}^{*EXT} = \frac{1}{\beta(1-\tau D(\rho,t))} * \left(\underbrace{\frac{\rho}{1+\rho}\left(\frac{\partial V_{T-t}^{pl}}{\partial t}dt + \frac{\partial V_{T-t}^{pl}}{\partial \lambda}d\lambda\right)}_{>0} - \underbrace{\tau\frac{\partial D(\rho,t)}{\partial t}(\bar{y} - \beta y_{T-t}^{*EXT})dt}_{>0} \right) \tag{8}$$

and hence, the overall effect on the reservation income of the extensive margin is ambiguous. Like the partial effect in (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

$$dy_{T-t}^{*INT} = \frac{\partial y_{T-t}^{*INT}}{\partial t}dt > 0. \tag{9}$$

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 Estimation strategy

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

[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).

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 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 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.

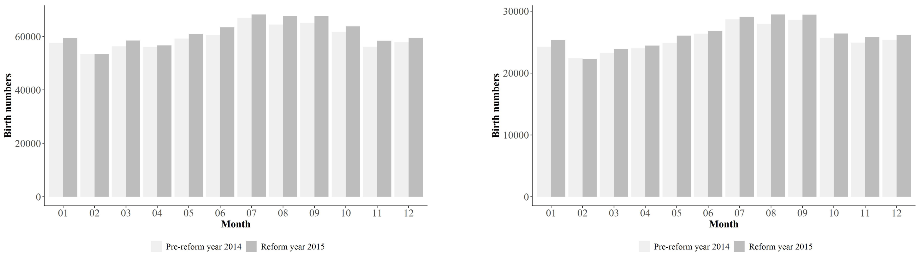
Assumption 2 (Observational rule): The outcome process follows the observational rule

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

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 deci-

Figure 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, 2015a).

sion. 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 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 2). We handle the second type of anticipation, trying to postpone the birthdate that might

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

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 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 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 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.

4.2 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 (2018) and Zimmert (2018) propose DiD estimators that combine propensity score and outcome estimation (Augmented Inverse Probability Weighting AIPW). In particular, they show that

$$ATET = \theta = \mathbb{E}\left[\frac{1}{\lambda_D}\frac{T-\lambda_T}{\lambda_T(1-\lambda_T)}\frac{D-p(X)}{1-p(X)}\left(Y-\gamma(X,T)\right)\right] \tag{10}$$

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

[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).

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 (10) converges with square-root-N to a normal distribution and has the asymptotic variance

$$\sigma^2 = \mathbb{E}\left[\left(\frac{1}{\lambda_D}\frac{T-\lambda_T}{\lambda_T(1-\lambda_T)}\frac{D-p(X)}{1-p(X)}\left(Y-\gamma(X,T)\right)\right)^2\right] \tag{11}$$

[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 (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]

4.3 Estimation of heterogenous effects

Yet another advantage of the estimand proposed in (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

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

[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.

The parameter represents a so-called conditional average treatment effect on the treated (CATET[11]). Define the propensity score conditional on Z as $\mathbb{E}[D|Z] = p(Z)$. Then we can show that under the assumptions in Section 4.1 and the further assumption that $p(Z) > 0$, the CATET is identified as

$$\theta(z) = \mathbb{E}\left[\frac{1}{p(Z)}\frac{T-\lambda_T}{\lambda_T(1-\lambda_T)}\frac{D-p(X)}{1-p(X)}(Y-\gamma(X,T))\Big|Z=z\right]. \tag{12}$$

The details for this result are provided in Appendix A.2. Equation (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 4.2, inference for this estimator might be very complicated in our setting. We therefore reweight the transformed outcome also used in (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 is low-dimensional and hence $p(Z)$ can for example be estimated using logit regression. The estimand in (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.

[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)

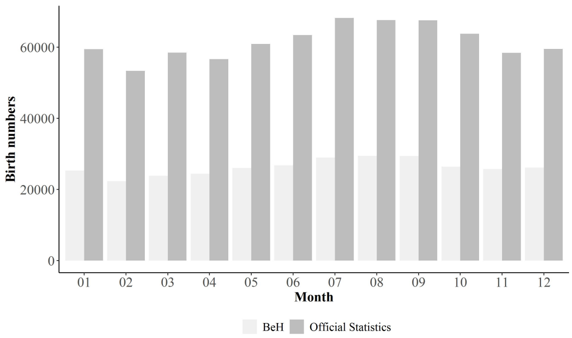
5 Data

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 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

Figure 3: Imputed and official birth numbers in 2015



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

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 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 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

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 4.2. 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 Zimmert, 2019). 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 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

[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).

[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.

include information on the number of (female) employees coming from the *IAB Establishment History Panel (BHP, version 7516 v1).*

Table 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.

Table 4: Descriptive statistics of covariates by group membership

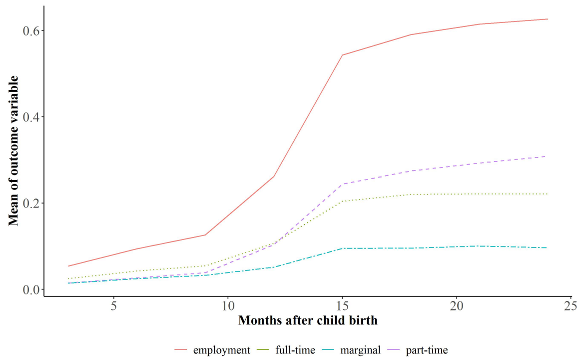
Variable	Control group Mean	sd	Treated group Mean	sd	Standardized difference
Age	29.36	3.45	29.31	3.47	-0.013
Number of children	1.301	0.522	1.301	0.520	-2.8E-04
Migration background	0.066	0.248	0.066	0.249	0.002
Place of living (Federal state, *baseline* Schleswig-Holstein:)					
Hamburg	0.024	0.152	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 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 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

Figure 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)*.

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 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).

6 Results

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 5a. The reform gives 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 5b to 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.

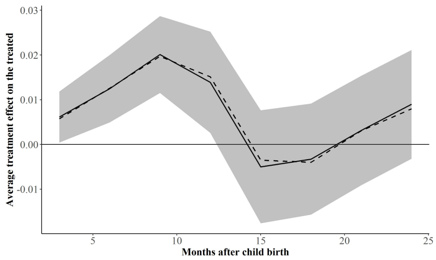
These empirical findings are in line with the proposed model mechanisms of Section 3 predicting a decrease of the reservation income for a part-time job relative to a full-time job. Moreover, we find that the theoretically ambiguous effect on the extensive employment margin is empirically positive. Interestingly, 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 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

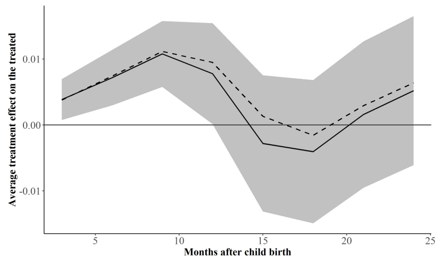
[14] We are cautious when interpreting DiD results in a Wald estimator kind of manner. For example De Chaisemartin and D'Haultfoeuille (2017) show that such an argumentation may only be valid under very restrictive assumptions.

Figure 5: Baseline estimation results

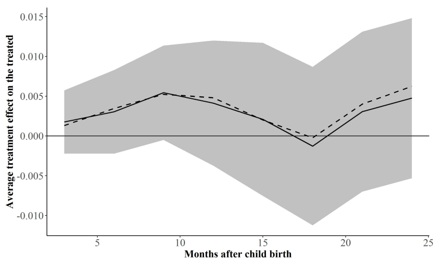
(a) Employment



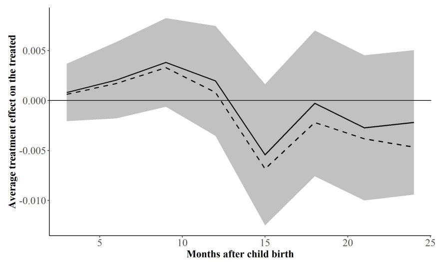
(b) Part-time



(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 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 as they might fear to get stuck in a part-time contract.

Panel A of Table 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 6 hints at similar employment trends before the reform reflected in

Table 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)*.

ATETs that are precisely measured at around zero. Earnings and job continuity are as well not affected (compare Panel B of Table 5).

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 5). We check the sensitivity against these two threats by estimating the effects for different populations. Figure 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 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 8). We conclude that our results are robust regarding these potential issues.

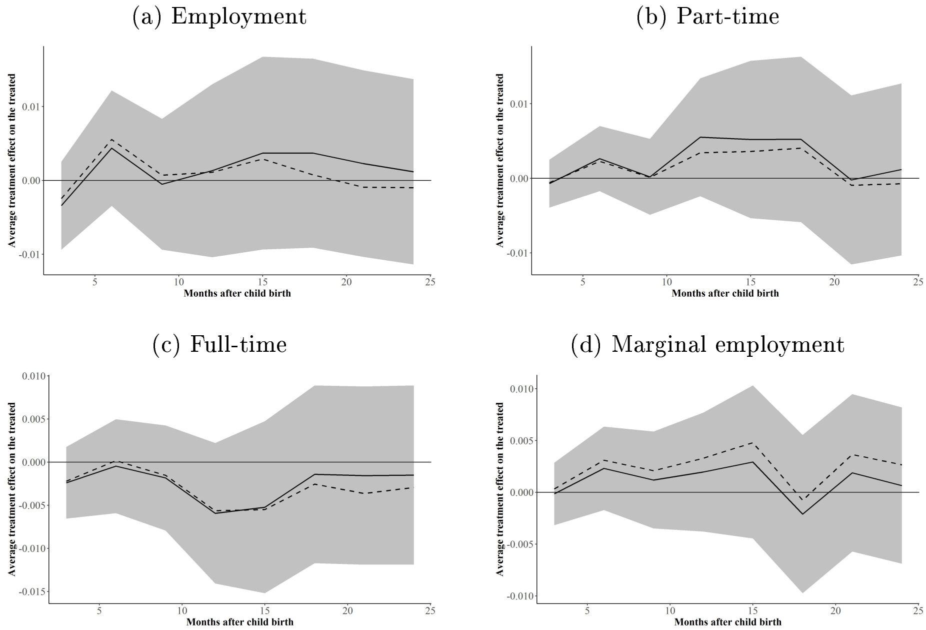
As a possible channel driving the employment outcomes, we look at the effect on subsequent births. E.g., Cygan-Rehm (2016) shows that the parental leave reform of 2007 incentivized mothers to postpone a subsequent pregnancy. Figure 9 does not indicate any effect on childbearing within the two year-horizon. However, this finding has to be interpreted with caution as we only observe women with subsequent births who have been employed in the meanwhile.

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

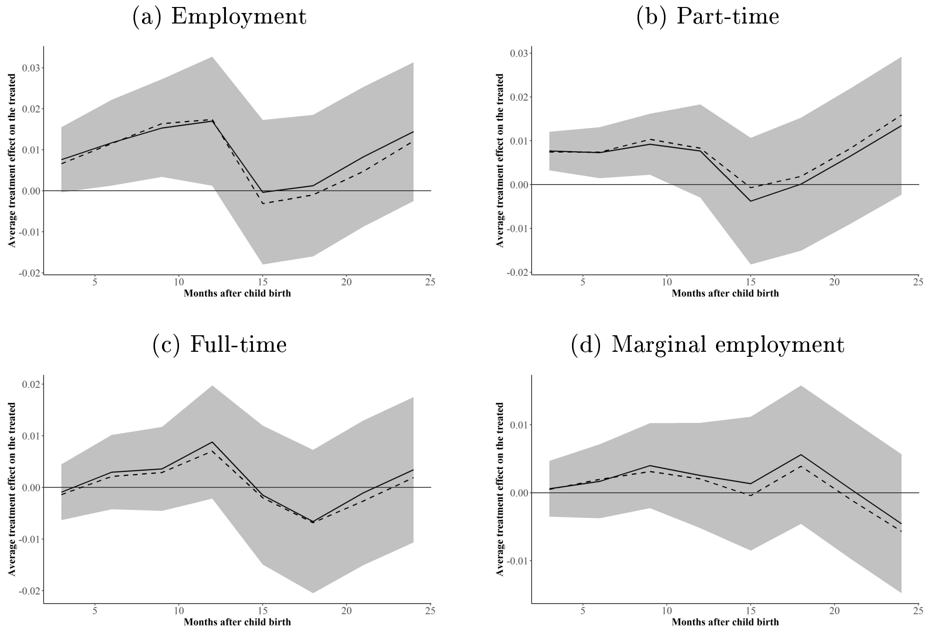
[15] The respective pre-reform outcome means of these subgroups are depicted in Figures A.2 and A.3.

Figure 6: Estimation results of placebo reform



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)*.

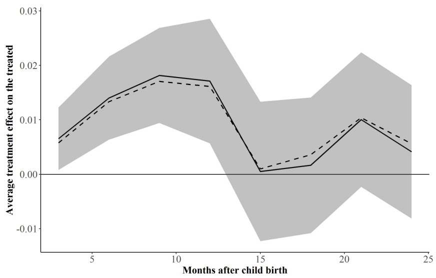
Figure 7: Estimation results with small bandwidth



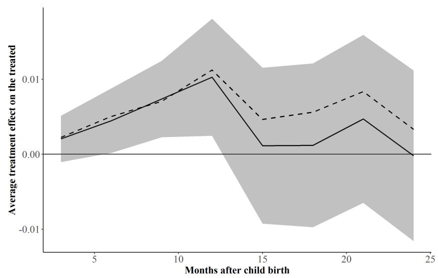
Notes: Treatment status is measured four weeks around cut-off date excluding the two weeks on each side of the cut-off date. $T = 1$ in 2015, $T = 0$ in 2014. $N = 48,544$. Grey shaded areas depict pointwise 95 % confidence intervals. AIPW DiD is solid line. Unadjusted DiD is dashed line.
Source: Own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

Figure 8: Estimation results with large bandwidth

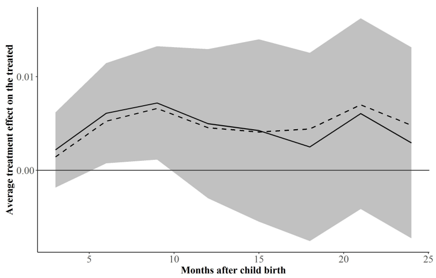
(a) Employment



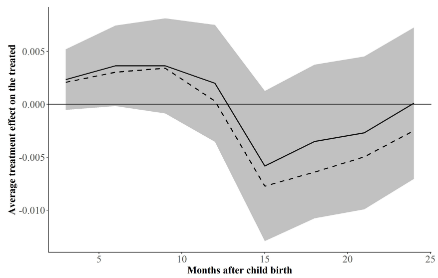
(b) Part-time



(c) Full-time

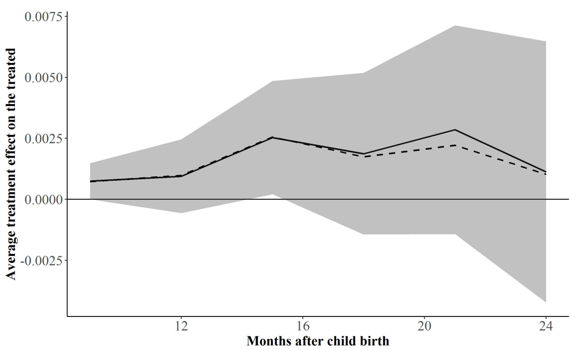


(d) Marginal employment



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 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)*.

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 10 to 13. As discussed in Section 4.3 we report the usual OLS standard errors.

Figures 10a and 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 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 mothers.[16] For high-income mothers, the effects do not significantly differ from those with lower income (see Figure 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 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 sub-

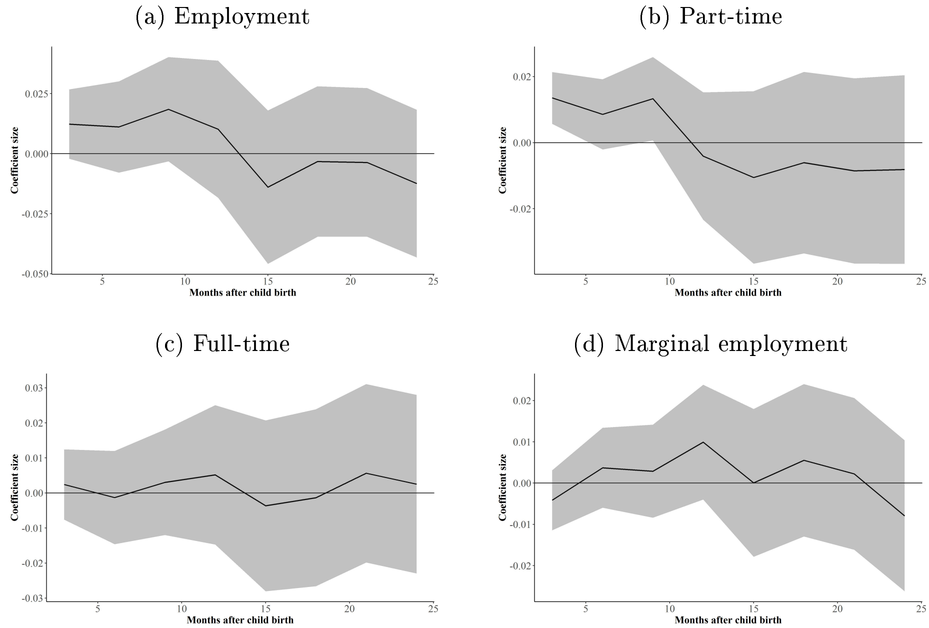
[16]Table A.1 in the appendix shows that these effects reflect in higher accumulated earnings within the first year. In the main analysis we concentrate on the CATETs for the different working time patterns. See Table A.1 for a detailed presentation.

groups.
Interestingly, the effect size does not significantly vary with the place of living, i.e., East and West Germany (see Figure 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 13a and 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.

7 Discussion

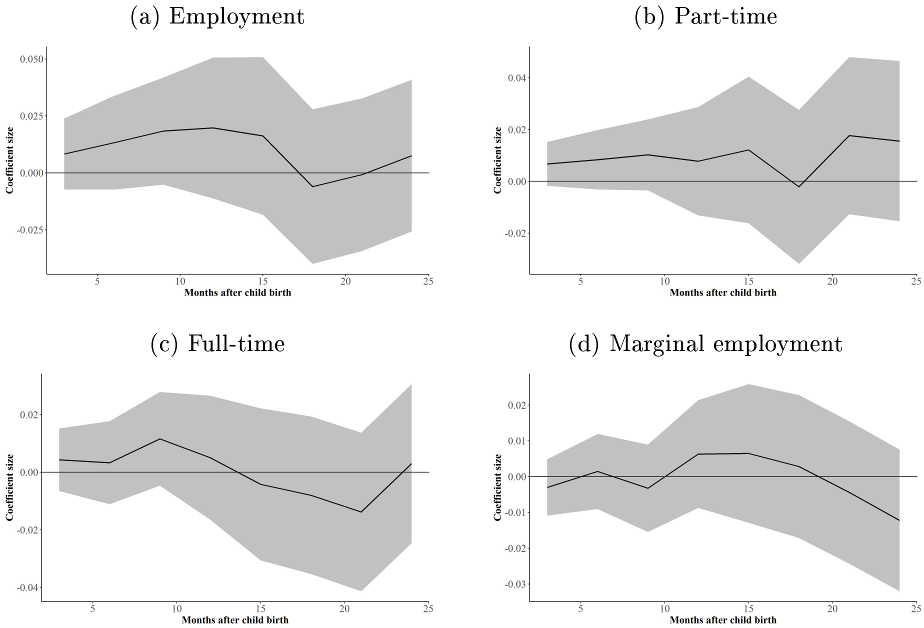
Although the overall employment effects amount to about 14 percent of the pre-reform mean, the new regulations, and consequently a return to work before the child's first birthday are only attractive to about 20 percent of all female benefit recipients (Federal Statistical Office, 2019). In this regard, 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). Zimmert

Figure 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)*.

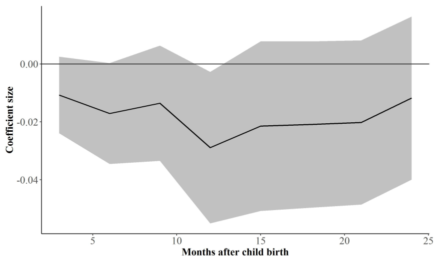
Figure 11: Estimation results for high income group



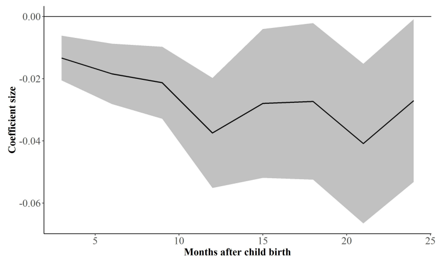
Notes: Low (middle, high) income corresponds to 1st (2nd, 3rd) tercile of gross previous monthly income. $N_{low\_income} = 31,170$, $N_{middle\_income} = 32,139$ and $N_{high\_income} = 31,166$. The coefficients give the effect variation for the respective subgroup. The reference group is low income. 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 12: Estimation results for prior full-time

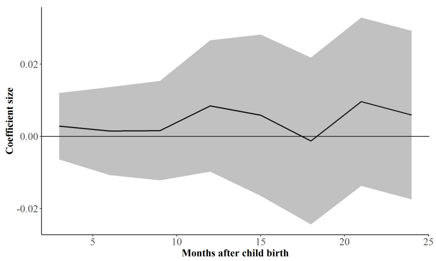
(a) Employment



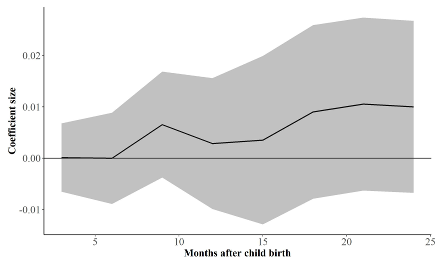
(b) Part-time



(c) Full-time



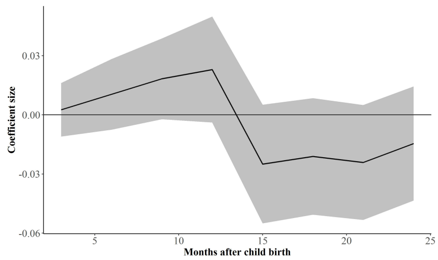
(d) Marginal employment



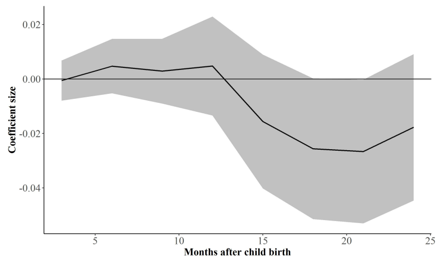
Notes: $N_{part\_time} = 29,214$ and $N_{full\_time} = 60,913$. Coefficient gives effect variation for respective subgroup. The reference group is prior part-time and marginal employment. Grey shaded areas depict pointwise 95 % confidence intervals.
Source: Own calculations based on employee data from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

Figure 13: Estimation results for West Germany

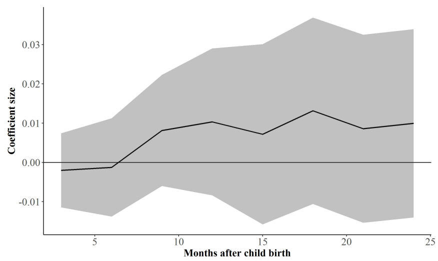
(a) Employment



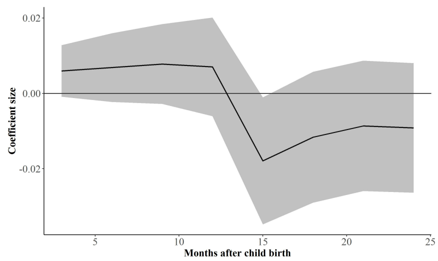
(b) Part-time



(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)*.

(2019) 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

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

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

[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.

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.

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.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies 72*(1), 1–19.

Abrevaya, J., Y.-C. Hsu, and R. P. Lieli (2015). Estimating conditional average treatment effects. *Journal of Business & Economic Statistics 33*(4), 485–505.

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

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.

Belloni, A. and V. Chernozhukov (2013). Least squares after model selection in high-dimensional sparse models. *Bernoulli 19*(2), 521–547.

Bergemann, A. and R. T. Riphahn (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.

Bergemann, A. and R. T. Riphahn (2015). Maternal employment effects of paid parental leave. *IZA Discussion Papers 9073.*

BMFSFJ (2018). *Elterngeld, ElterngeldPlus und Elternzeit. Das Bundeselterngeld- und Elternzeitgesetz.* Federal Ministry for Family Affairs, Senior Citizens, Women and Youth (BMFSFJ).

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

Chernozhukov, V. and V. Semenova (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.

Cygan-Rehm, K. (2016). Parental leave benefit and differential fertility responses: Evidence from a German reform. *Journal of Population Economics 29*(1), 73–103.

Cygan-Rehm, K., D. Kühnle, and R. T. Riphahn (2018). Paid parental leave and families' living arrangements. *Labour Economics 53*, 182–197.

Dahl, G. B., K. V. Løken, M. Mogstad, and K. V. Salvanes (2016). What is the case for paid maternity leave? *Review of Economics and Statistics 98*(4), 655–670.

De Chaisemartin, C. and X. D'Haultfoeuille (2017). Fuzzy differences-in-differences. *The Review of Economic Studies 85*(2), 999–1028.

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.

Fan, Q., Y.-C. Hsu, R. P. Lieli, and Y. Zhang (2019). *Estimation of Conditional Average Treatment Effects with High-Dimensional Data*. Unpublished manuscript.

Federal Statistical Office (2006). *Kindertagesbetreuung regional 2006. Ein Vergleich aller 439 Kreise in Deutschland*. Federal Statistical Office, Wiesbaden.

Federal Statistical Office (2014). *Bevölkerung und Erwerbstätigkeit. Natürliche Bevölkerungsbewegung*. Federal Statistical Office, Wiesbaden.

Federal Statistical Office (2015a). *Bevölkerung und Erwerbstätigkeit. Natürliche Bevölkerungsbewegung*. Federal Statistical Office, Wiesbaden.

Federal Statistical Office (2015b). *Kindertagesbetreuung regional 2015. Ein Vergleich aller 402 Kreise in Deutschland*. Federal Statistical Office, Wiesbaden.

Federal Statistical Office (2019). *Öffentliche Sozialleistungen. Statistik zum Elterngeld. Beendete Leistungsbezüge für im 3. Vierteljahr 2015 geborene Kinder*. Federal Statistical Office, Wiesbaden.

Fitzenberger, B., A. Osikominu, and R. Völter (2005). Imputation rules to improve the education variable in the IAB employment subsample. *ZEW Discussion Paper 10/2005.*

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review 104*(4), 1091–1119.

Hanel, B. and R. T. Riphahn (2012). The employment of mothers – recent developments and their determinants in East and West Germany. *Jahrbücher für Nationalökonomie und Statistik 232*(2), 146–176.

Hastie, T., R. Tibshirani, and J. Friedman (2009). *The elements of statistical learning: prediction, inference and data mining.* Springer-Verlag, New York.

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

Jessen, J., R. Jessen, and J. Kluve (2019). Punishing potential mothers? Evidence for statistical employer discrimination from a natural experiment. *Labour Economics 59*, 164–172.

Kluve, J. and S. Schmitz (2014). Social norms and mothers' labor market attachment. *IZA Discussion Paper 8115.*

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

Kluve, J. and M. Tamm (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.

Lalive, R. and J. Zweimüller (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.

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

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics 4*(3), 165–224.

Lee, S., R. Okui, and Y.-J. Whang (2017). Doubly robust uniform confidence band for the conditional average treatment effect function. *Journal of Applied Econometrics 32*(7), 1207–1225.

Lundborg, P., E. Plug, and A. W. Rasmussen (2017). Can women have children and a career? IV evidence from IVF treatments. *American Economic Review 107*(6), 1611–37.

March of Dimes, PMNCH, Save the Children, and WHO (2012). *Born Too Soon: The Global Action Report on Preterm Birth.* World Health Organization, Geneva.

Müller, D., K. Strauch, et al. (2017). Identifying mothers in administrative data. *FDZ-Methodenreport 13/2017.*

Patnaik, A. (2019). Reserving time for daddy: The consequences of fathers' quotas. *Journal of Labor Economics 37*(4), 1009–1059.

Rossin-Slater, M., C. J. Ruhm, and J. Waldfogel (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.

Rubin, D. B. (1973). The use of matched sampling and regression adjustment to remove bias in observational studies. *Biometrics 29*(1), 185–203.

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.

Sant'Anna, P. H. and J. B. Zhao (2018). Doubly robust difference-in-differences estimators. *Unpublished manuscript, Department of Economics, Vanderbilt University*.

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

Tamm, M. (2019). Fathers' parental leave-taking, childcare involvement and labor market participation. *Labour Economics 59*, 184–197.

Wager, S. and S. Athey (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association 113*(523), 1228–1242.

Wager, S. and G. Walther (2015). *Adaptive concentration of regression trees, with application to random forests.* Unpublished manuscript, Department of Statistics, Stanford University.

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.

Wanger, S. (2015). Frauen und Männer am Arbeitsmarkt: Traditionelle Erwerbs- und Arbeitszeitmuster sind nach wie vor verbreitet. *IAB-Kurzbericht 18/2015.*

Welteke, C. and K. Wrohlich (2019). Peer effects in parental leave decisions. *Labour Economics 57*, 146–163.

Zimmert, F. (2019). Early child care and maternal employment: Empirical evidence from Germany. *IAB-Discussion Paper 2/2019.*

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.

Zimmert, M. and M. Lechner (2019). *Nonparametric estimation of causal heterogeneity under high-dimensional confounding.* Unpublished manuscript, Swiss Institute for Empirical Economic Research, University of St. Gallen.

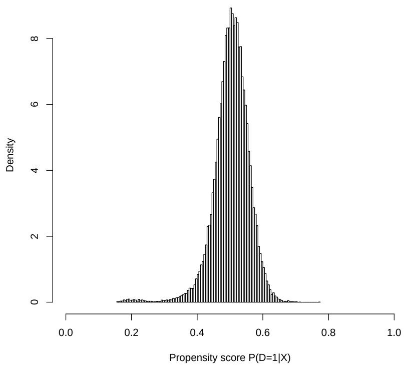
Zoch, G. and I. Hondralis (2017). The expansion of low-cost, state-subsidized childcare availability and mothers' return-to-work behaviour in East and West Germany. *European Sociological Review 33*(5), 693–707.

A Appendix

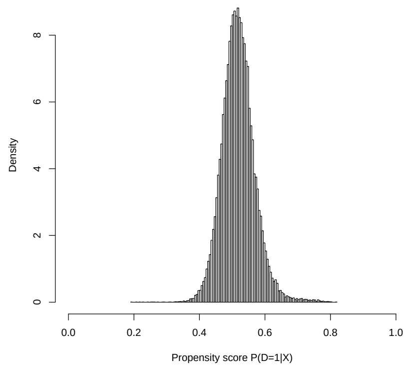
A.1 Figures and Tables

Figure A.1: Propensity scores by treatment status

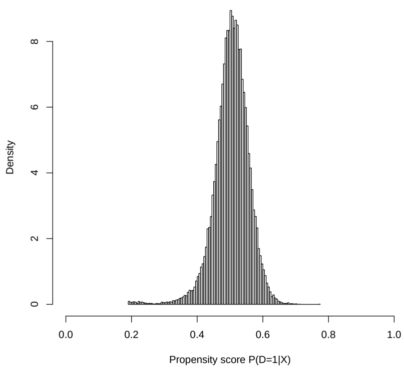
(a) $D = 0$ before trimming



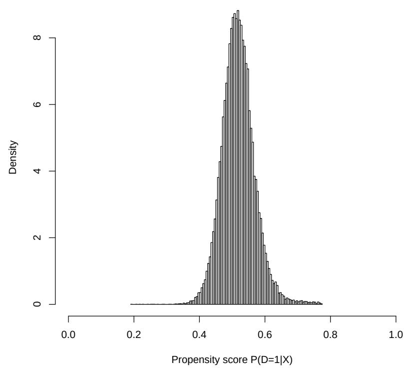
(b) $D = 1$ before trimming



(c) $D = 0$ after trimming



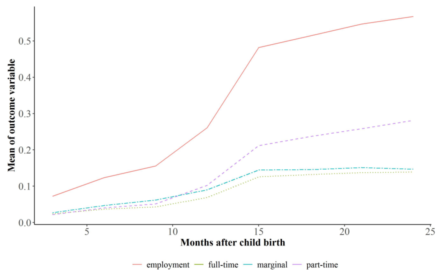
(d) $D = 1$ after trimming



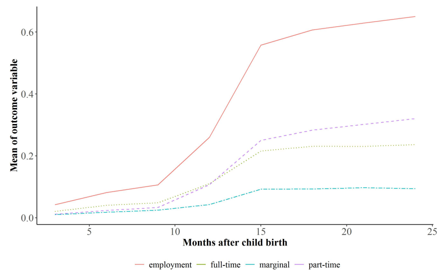
Notes: $N_{D=0} = 46,263$ and $N_{D=1} = 48,212$ before trimming; $N_{D=0} = 46,191$ and $N_{D=1} = 48,184$ after trimming.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

Figure A.2: Outcome means of treated mothers before reform (1)

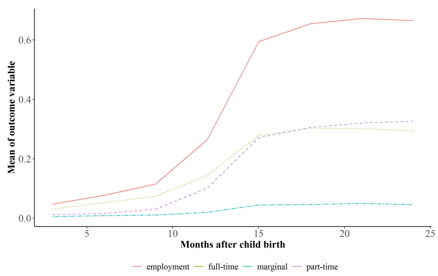
(a) Low income



(b) Middle income



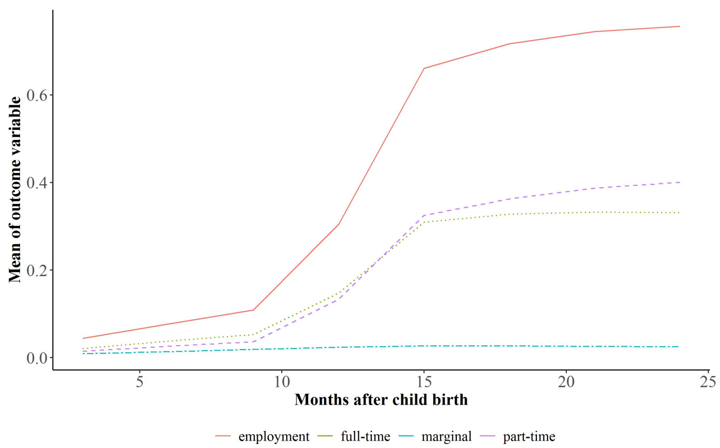
(c) High income



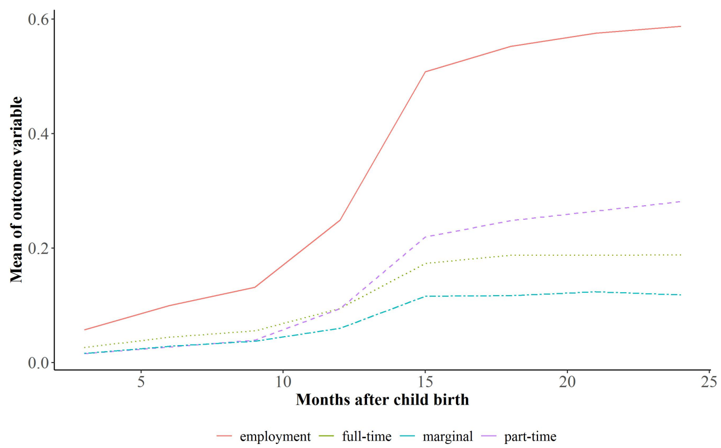
Notes: $T = 0$ in 2014. $N_{low\_income} = 8,198$, $N_{middle\_income} = 8,288$ and $N_{high\_income} = 8,367$.
Source: Own calculations based on employee data from the *Employment History (BeH)* and establishment data from the *Establishment History Panel (BHP)*.

Figure A.3: Outcome means of treated mothers before reform (2)

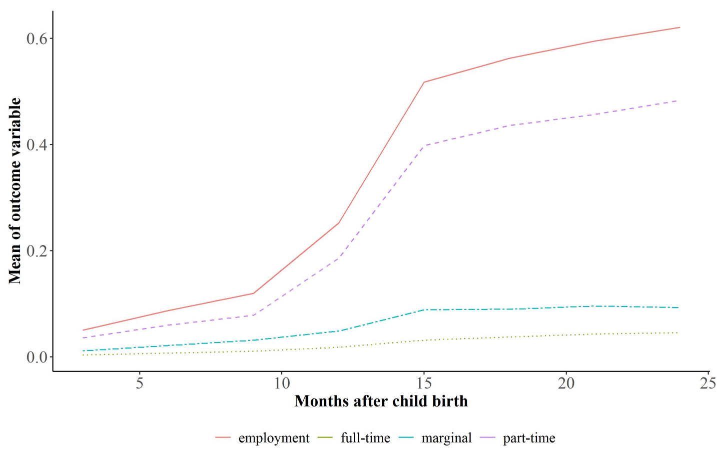
(a) East Germany



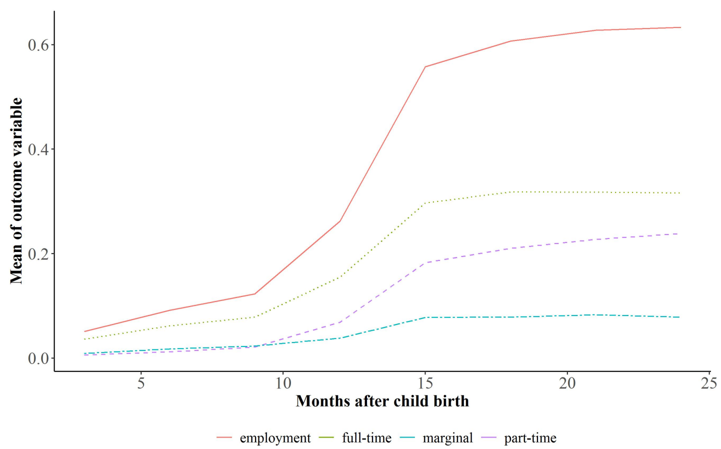
(b) West Germany



(c) Prior part-time



(d) Prior full-time



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)*.

Table A.1: 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)*.

A.2 Mathematical appendix

A.2.1 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:

$$V^{pl}_{T-t-1} = \lambda\bar{y} + l + \frac{1}{1+\rho}\int_0^\infty \max_{pl,f,p}(V^{pl}_{T-t}, V^{f}_{T-t}, V^{p}_{T-t})dF(y_{T-t}) \qquad (13)$$

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

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

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

$$\begin{aligned} &V^{pl}_{T} = \lambda\bar{y} + l + \frac{1}{1+\rho}\int_0^\infty \max(V^{u}_{T+1}, V^{f}_{T+1}, V^{p}_{T+1})dF(y_{T+1}) \text{ such that} \\ &V^{pl}_{T} - \lambda\bar{y} - l - \frac{b^u}{\rho} = \\ &\int_0^\infty \max\left(0, \frac{1}{\rho}(y_{T+1} - b^u), \frac{1}{\rho}(\beta y_{T+1} + (1-\beta)l - b^u)\right) dF(y_{T+1}) \qquad (14) \end{aligned}$$

for inserting the infinite series in Equations 1.
Moreover, we can explicitly solve for the value function V^{pl}_{T} as

$$\begin{aligned} V^{pl}_{T} &= \lambda\bar{y} + l + \frac{b^u}{\rho} \\ &+ \frac{1}{\rho}P(y_{T+1} > l, y_{T+1} > b^u)\left(\mathbb{E}(y_{T+1}|y_{T+1} > l, y_{T+1} > b^u) - b^u\right) \end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{\rho}P\left(y_{T+1}<l, y_{T+1}>\frac{b^u}{\beta}-\frac{1-\beta}{\beta}l\right)\\
&*\left(\beta\mathbb{E}\left(y_{T+1}|y_{T+1}<l, y_{T+1}>\frac{b^u}{\beta}-\frac{1-\beta}{\beta}l\right)+(1-\beta)l-b^u\right). \quad (15)
\end{aligned}
$$

Thus, iterating backwards will give an explicit solution for every V^{pl}_{T-t} 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

$$
\begin{aligned}
&V^{pl}_{T-t-1}-\lambda\bar{y}-l-\frac{1}{1+\rho}V^{pl}_{T-t}=\\
&\int_0^\infty \max\left(0, \frac{y_{T-t}}{\rho}-\frac{1}{1+\rho}V^{pl}_{T-t}, \frac{\beta y_{T-t}+(1-\beta)l}{\rho}\right.\\
&\left.+\frac{D(\rho,t)\tau(\bar{y}-\beta y_{T-t})^+}{\rho}-\frac{1}{1+\rho}V^{pl}_{T-t}\right)dF(y_{T-t}) \qquad (16)
\end{aligned}
$$

such that we obtain the results of Equation 3.

Duration effects

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

$$
\frac{\partial D(\rho,t)}{\partial t}=\ln(1+\rho)\left(\frac{1}{1+\rho}\right)^{t+1}>0. \qquad (17)
$$

A.2.2 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

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

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

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

$$
\begin{aligned}
\text{ATET}(1) &= \mathbb{E}\left[Y_1^1-Y_1^0|D=1\right] \\
&= \int \mathbb{E}\left[Y_1^1-Y_1^0|X,D=1\right]f_{X|D=1}(x,d)dx
\end{aligned}
$$

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

Identification of CATET

From the identification of the ATET we know that

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

For the CATET we therefore obtain

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

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