New Evidence on the 2007 Parental Benefits Reform in Germany: Selection and Reform Effects

Bernd Fitzenberger* Arnim Seidlitz†

April 5, 2022

Preliminary; please do not cite or circulate without authors' permission.

Abstract

This paper studies the causal effect of giving first birth and how that effect changes by the 2007 reform in parental benefits in Germany. We further investigate how the reform affects the selection of women into motherhood and how that selection effect affects labor market outcomes. A large novel data set merging data from the pension insurance with administrative labor market data provides information on all births. We apply a dynamic treatment effect approach which differs from other strategies used so far in most of the literature to estimate the causal effect of motherhood and to evaluate the 2007 reform. The reform has positive medium-run effects on earnings and employment. There are no effects on second-order fertility and on full-time employment. While the reform slightly changes the selection of mothers, this has little impact on the reform effect for the causal effect of motherhood.

JEL-Classification: J08, J13, J16, J22.

Keywords: Parental leave reform, causal effect of motherhood, earnings, employment dynamic treatment effect.

*IAB, FAU Erlangen-Nürnberg, IFS, CESifo, IZA and ROA

†Humboldt University Berlin and Berlin School of Economics

1 Introduction

The birth of the first child is still a turning point in the career of many mothers and the so-called "child penalty" due to motherhood is viewed as a key reason for the persistent gender gap in the labor market (Angelov et al., 2016; Kleven et al., 2019b,a). Employment and earnings drops to (almost) zero for a few months immediately before and after child birth. The process of reentering employment is on average sluggish. After returning to employment first-time mothers often work less hours and often soon have their second child (Sigle-Rushton and Waldfogel, 2007; Fitzenberger et al., 2013). Most governments provide paid parental benefits and/or unpaid parental leave (henceforth job protection) to support parents after child birth and to enhance the compatability between work and having children. All OECD countries except the U.S. pay nationwide parental benefits, and in the U.S. some states like California have such programs (OECD, 2019).

Olivetti and Petrongolo (2017) and Rossin-Slater (2018) review different systems of family policies around the world and the economic research literature on this topic. There is evidence that job protection helps mothers to return to their previous employers. However, medium-run effects on labor market outcomes of extending an already existing job protection are often insignificant, and there is concern that longer periods of paid or unpaid parental leave are detrimental for the post-birth career of mothers. For Germany, Schönberg and Ludsteck (2014) investigate various reforms extending job protection from two to 36 months between 1979 and 1992. The study finds sizeable negative short-run effects on post-birth employment of mothers but only small negative long-run effects. Angelov et al. (2016), Kleven et al. (2019b), and Kleven et al. (2019a) find large negative long-run effects of child birth on mothers' labor market outcomes in various countries with long unpaid parental leave regulations.

Paid parental benefit systems have also gained interests within the research community. Rossin-Slater et al. (2013) evaluate the introduction of benefits in California offering transfers for up to six weeks. This relatively short duration seems to have positive effects on maternal employment until the third year after giving birth. The international evidence on extending benefits duration include studies on Austria (Lalive et al., 2014) and Norway (Dahl et al., 2016). Both countries had compared to California already a more generous parental benefits system. Increasing the benefit duration further had in both countries no significant medium run effects on employment outcomes. The results on the German reform in 2007 are presented in the following section.

Our study contributes to this literature by our approach of identifying first the causal effect of giving birth and in a second step, the reform effect from the effects of giving birth in pre- and post-reform periods. For estimating the effect of giving birth, we use a dynamic treatment approach combining the advantages of a dynamic matching approach and an event study (Fitzenberger et al., 2013; Sianesi, 2008). This strategy involves a control group which is the most important difference to all other papers on the 2007-reform and to most papers on the effect of giving birth (for example Angelov et al. (2016); Kleven et al. (2019a,b)). It allows us further to assess the change in the selection of mothers due to the reform and to distinguish this selection effect from the reform effect on the "treated" e.g. the mothers.

As one main advantage of our project, we have access to a novel administrative data set which is a merge of data from the pension insurance on fertility with labor market data. It provides information on all births, independent of the pre-birth employment status. Likewise, second births can be observed. From the labor market data, we construct a monthly panel on earnings, employment and full-time status which allows us to display these outcomes with monthly precision.

We find that the reform has positive medium-run effects on earnings and employment while it does not affect significantly full-time employment. There is further no effects on second-order fertility. To our surprise, the reform had no large effect on the selection of mothers.

The remainder is organized as follows. Section 2 discuss the reform and the literature on it, section 3 contains a description of the data, section 4 describes the estimation approach for the causal effect of giving birth, section 5 shows a straightforward approach how these effects can be used in an RDD-setting, section 6 continues the econometric approach, section 7 presents our main results, section 8 conducts robustness checks, section 9 compares our approach to alternative estimating techniques and section 10 concludes.

2 The 2007 parental leave reform

The 2007-reform made Germany to one of the most generous countries concerning parental benefits worldwide. All mothers of children born from January 2007 on may receive the new "parental allowance" (Elterngeld) until the 12 month after giving birth. The reform was implemented with a clear cutoff. Mothers of children born until December 2006 faced a different benefit system which paid substantially less benefits to most mothers.

The German system contains generally different institutions. From six weeks before the expected date of delivery until eights weeks after giving birth, the mothers are in so-called "maternity protection" (Mutterschutz). During that period, it is for them forbidden to maintain a paid work but they receive the "maternity allowance" (Mutterschaftsgeld) which covers 100 percent of the pre-birth earnings[1]. Health insurances

[1]To be precise, the pre-birth earnings are calculated as average of the three calender month before

and employers provide the funding of this "maternity allowance". It is not seen as part of the family policy but to insure the health of the mother and her child. Most importantly, it was not reformed in 2007 such that pre- and post-reform mothers are treated equally (BMFSFJ, 2020b).

Under the old system, the eligibility for further benefits depend on the mother's and potentially her partner's income (henceforth for convenience household income[2]). Mothers of high income households received after the end of maternity protection – from the third month after giving birth on – no further financial benefit. Medium and low income households were paid 300 Euro monthly so-called "child raising allowance" (Erziehungsgeld). The former group was eligible for this benefit until the sixth month and the latter group until the 24th month after giving birth. Low income households alternatively had the possibility to choose 450 Euro monthly until the end of the first year instead of receiving 300 Euro monthly until the end of the second year (BMFSFJ, 2004). That second option for low-eaners was however only chosen by a minority of mothers (Kluve and Tamm, 2013).

From 2007 on, eligibility is universal and the amount of monthly benefits is earning dependent. Mothers receive 65 percent[3] of their monthy earnings in the 12 months before entering "maternity protection" but at least 300 and at most 1800 Euro. This benefit is paid until the end of the first year after giving birth. Single-mothers have the possibilities to receive two additional months of "parental allowance"[4] (Ehlert, 2008).

It is also important to note what was not touched by the reform. The "maternity protection" and the "maternity allowance" directly before and after giving birth remained unchanged. Further, mothers loose eligibility for both, the old "child-raising allowance" and the new "parental allowance" when working full-time, working until at most 30 hours weekly is allowed[5]. Mothers also enjoy job protection (Elternzeit) for three years, unchanged since 1992. Within that period mothers may return to their previous jobs (BMFSFJ, 2020a).

Figure 1 shows the pre- and post-reform situation with respect to household income. It is quite easy to distinguish reform-winners from reform-losers. Medium and especially high income household receive unambiguously longer and higher benefits while low-

the mother enters pregnancy.

[2]The relevant quantity was the yearly net-income of the mother if she raised her child as single-mother and the combined income of her and her partner if they lived in one household. For both groups of mothers different income thresholds were used.

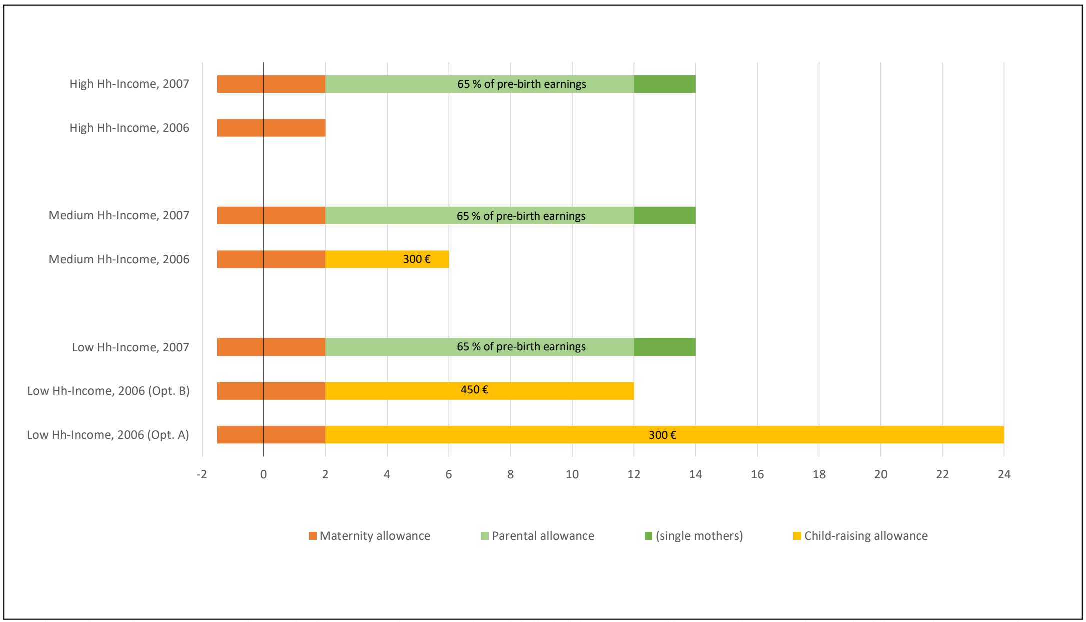
[3]The replacement rate decreases in earnings. The differences are moderate for earnings above 1000 Euro montly (between 65 and 67 percent), below 1000 Euro it increases to 100 percent for 440 Euro monthly earnings (BMFSFJ, 2020a).

[4]Fathers are also targeted by the reform. From 2007 it is possible that both parents receive benefits. However, effects on them cannot be covered in this projects. The same holds true for adjustments in 2015 which added more flexibility for the receivers of "parental allowance" (BMFSFJ, 2020a).

[5]If mothers work during receivance of "parental allowance", their benefits will decrease to 65 percent of the difference between pre- and post birth earnings (BMFSFJ, 2020a)

income household may fare worse compared to the old system. To break down all the complex regulations to a simple message, it is the case that Germany changed from a system paying higher benefits to low-earners to one paying higher benefits to high-earners.

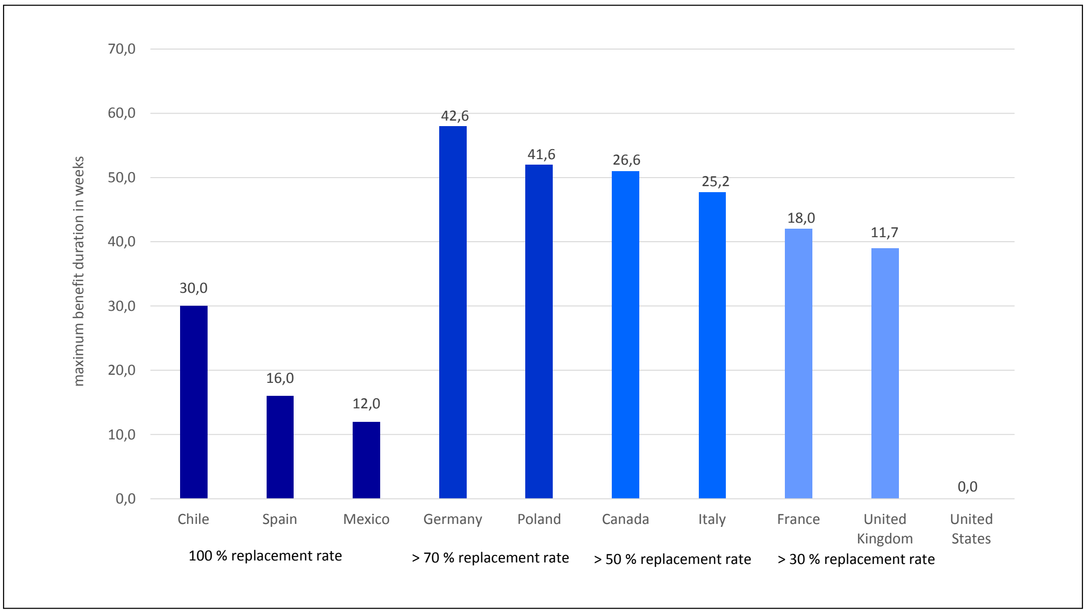
Figure 1: Benefits under new and old regimes by household income



High and medium income households according to the old "child-raising allowance" were households with an yearly net-income above 30000 and 22000 (23000 and 16000 for single-mothers), respectively.
"Parental allowance" is bound to be at least 300 and at most 1800 Euro monthly.
Receivers of both, "child-raising allowance" and "parental allowance" are restricted to work at most 30 hours per week.

The "parental allowance" is internationally one of the most generous systems. Figure 2 compares the system of some of the larger OECD-countries. The countries are grouped by the average replacement rate in terms of the previous income. The height of the bars gives the maximum duration of benefits. The Spanish speaking countries have institutions comparable to "maternity allowance" in Germany, a replacement of 100 percent but only for a short period. The other European countries and Canada pay benefits for a longer duration but do not replace the earnings entirely. Outstandingly, the US do not have parental benefits on national level. Poland and Germany are arguably the most generous countries of the displayed countries. An indication for this claim is the number on top of the bars which gives the product of replacement rate and benefits duration, the theoretical amount of weeks with 100 percent replacement. The investigation of the German reform is thus interesting for most other countries which would have the room for an increment in transfers paid to mothers.

Figure 2: Comparison of different parental benefit systems



Source: OECD Family Database (OECD, 2019).
Countries are grouped by the average level of replacement rate. The number at the top of the bar is the product of the replacement rate and the maximum duration in weeks. It can be interpreted as the theoretical number of weeks with 100 percent replacement.
For Germany: 14 weeks x 100 % (maternity allowance)+ 44 weeks x 65 % (parental allowance)= 42,6

2.1 Literature on the German reform

Shortly after the reform was implemented and data became available, researcher started to investigate its short-term effects. Bergemann and Riphahn (2011) found a positive reform effect on employment for the second year while Kluve and Tamm (2013) found a negative reform effect on employment for the first year after giving birth. Geyer et al. (2015) confirms both – the positive effect in the second year at least for low-earners. These short-term effects are not contradictory. Given the increase of benefits in the first year and the decrease in the second year (figure 1), they are rather expectable.
Meanwhile, there are also some papers studying the medium term outcomes. Bergemann and Riphahn (2020) show that the temporary gains in employment probability do not persist in the medium run. Frodermann et al. (2020) concentrate in their study on earnings and find a significant positive reform effect two years after giving birth. For low-earners, that is the only significant effect, while for high-earners the positive effects stays significant but diminishes over the following years. After eight years it is insignificant. The working hours do not seem to be significantly affected. This is widely in line with the results of Kluve and Schmitz (2018) who find significant effects on employment until five years after giving birth which are driven by medium and high earners. They show further heterogenous effects for full-time employment. The probability to work full-time after five years is reduced for low-earners but increased for high earners due to the reform. On average, it is insignificant. It seems thus that the effect in the medium run (after five years) are less striking than the short run effects. There is

at least evidence for a positive reform effect on outcomes as earnings and participation which are significant for high-earners.
Two studies work on fertility effects of the reform. Cygan-Rehm (2016) finds temporarily a huge decrease in higher-order fertility. In the medium run this negative effect becomes insignificant. It is driven by the reform-losers. Raute (2019) shows that the propensity to become mother (first and higher order fertility) for high earning women increases compared to low earning women as response to the reform.
This results implies that it is important to investigate the effect of the selection induced by the reform on labor market outcomes. That is something which is so far missing in the literature. Frodermann et al. (2020); Kluve and Schmitz (2018) use a diff-in-diff and regression discontinuity design relying only on observations directly before and after the reform. They need to assume that differences in the selection of mothers do not drive their results and that is for their setting convincing as the reform was announced only six months in advance such that there was no time to adjust fertility behavior. Bergemann and Riphahn (2020), on the other hand, work with mothers of the year 2005 to 2008, two years before and after the reform. Changes in the selection of mothers might be a bigger threat to them although they show that their result do not change much when they restrict themselves to a narrower time window.
However, to our best knowledge, we are the first who try to quantify the effect of the selection induced by the reform on labor market outcomes.

3 Data

We have the great chance to work with the merge of two different high quality administrative data sources from Germany. These are data which originate on one hand from social insurance and on the other hand from pension insurance.

3.1 Social insurance data

The data of the social insurance are the so-called Integrated Employment Biographies (IEB). They are administrated by the Institute for Employment Research (IAB) and contain data on all employees in Germany but not on civil servants and self-employed. For them, there are precise information on employment spells including start and end date, gross-earnings, a part-time status, education level and the county of employment[6]. From these spell data, we form a monthly panel and construct an average daily earnings rate for each month. Therefore, we add potentially the earnings of three different

[6]For a detailed description of the data which include many more features, we are not using in this project, we would recommend the documentation of the SIAB. The SIAB is a two percent sample of the IEB (Antoni et al., 2019).

jobs but we do not regard marginally employments with less than 14 days duration or earnings below 300 Euros in one month.
One potential issue of the IEB-data is the top coding of earnings. However, for the population of women only the top four percent of employment spells are distorted (Fitzenberger and Seidlitz, 2020). We hence decided to neglect this issue. The earnings are further deflated by the annual consumer price index of the Federal Statistical Office (Statistisches Bundesamt, 2019).

3.2 Pension insurance data and data merge

The data from the pension insurance are called "Versichertenkontostichprobe (VSKT)". These are also spell data on employment outcomes (FDZ-RV, 2021). Most important for us, they contain for mothers the concrete month of giving birth for each of her children. This information is extremely valuable as it is not included in the data of the social insurance.
The data merge was conducted within the project "Custom Shaped Administrative Data for the Analysis of Labour Market (CADAL)"[7] by researchers at the IAB. As both sources provided their data anonymized, the union of the two data sources is based on a probabilistic matching. This procedure is possible as the data of the pension insurance also contains information on employment spells although not as detailed as the social insurance. The colleagues at the IAB designed for the matching an algorithm which is mainly based on the ratio of employment spells which coincides in terms of start and end date, earnings and status of employment (covered by social insurance, marginal or in apprenticeship) in both data bases and required that both data bases reported the same month and year of birth.
This algorithm was applied to roughly 306000 female observation of the pension insurance. For two third of them a unique match in the social insurance was found. The rest of the data where the algorithm identified either no or more than one candidate match were not used.
For our analysis, we introduced a few additional minimum criteria on the working history and exclude all observations with a working history in the former East German GDR, to decrease the risk of mismatches[8]. Of course this approach contains the danger

[7]This project was part of the SPP "The German Labor Market in a Globalized World: Challenges through Trade, Technology, and Demographics" (SPP 1764) which was funded by the German Research Foundation (DFG).

[8]Given our probabilistic matching strategy, it is clear that the accuracy increases with the length of the working history. We therefore require our observations to have at least worked for two years within the social insurance and to report at least ten spells in total and five working spells covered by the social insurance. The reporting quality of spells in marginal employment is further a potential issue. Therefore, we require the observations to have earned on average more than the marginal employment threshold in her working spells. Lastly, the algorithm may produces mismatches if an observation has periods which are covered in pension insurances but not in social insurance. Consequently, we exclude

that our sample has an above average labor market attachment and is less representative for the entire population. However, we want to highlight that our selection is arguable much less restrictive than the selection which arises when maternity is directly identified from social insurance data. Müller and Strauch (2017) described a way how mothers can be detected from the information on working spells in the IEB data. This approach is very helpful and widely used (for example by Frodermann et al. (2020)) but contains a strong restriction. It requires mothers to work directly before entering motherhood. Our strategy basically requires the observations, treated and control, to have worked for some time in their life at any time before or after giving birth. Importantly, there is no reason to suspect spurious differences in the selection between mothers and control women.

3.3 Descriptives

To evaluate the effect of the 2007 reform, we regard the mothers who have her first child within three years before or after the reform eg. the first time-mothers of the six years from 2004 to 2009. This means naturally that we exclude all women who have her first child before 2004 from the matched sample. Further, we restrict the sample to a core working and fertility age group of women aged 21 to 40. All observations are required to be in this age group in at least one year between 2004 and 2009. This is equivalent with stating that we only evaluate the population born between 1964 and 1988. For the control group, we keep all women who are childless until 2010, regardless whether they become later mothers.
This restrictions leave us 50000 observation, a bit more than 10000 of them become her first child in the period of interest. Table 1 shows some descriptives for our sample.
We learn that the sample of mothers is a bit better educated. 88 percent of them hold at least a secondary educational certificate compared to 81 percent of the childless women. The differences in age and region of residence are small. The share of East Germans in the sample around four percent is very low which we explain by the selection induced by the matching.
Earnings, participation and full-time status are the main outcome variables of our analysis. Table 1 contains the dramatic decline in means of these variables if the pre-birth period is compared to the post-birth period for mothers. Mean earnings are reduced to less than half of the pre-birth earnings, employment decreases by 30, full-time employment even more than 40 percentage points. Regarding full-time, we use the correction, we have proposed in Fitzenberger and Seidlitz (2020)[9].

observation who have started their working career before 1975 when the social insurance records began (given our observational period that is no severe restriction) and who have a working history in the former GDR which is generally not included in social insurance data.

[9]After a reform in the data collection procedure in 2011, it became clear that the social insurance

Table 1: Average values for outcome and control variables

	childless (2004-2009)	mothers (first child, 2004 -2009) before giving birth	after giving birth
labor markot outcomes:			
daily earnings	55.89	62.97	27.24
participation	.743	.813	.518
(corrected) full-time	.482	.584	.123
(raw full-time)	.493	.603	.136
controls:			
year of birth	1977.77	1976.28	
former East Germany	.042	.033	
medium education	.666	.714	
high education	.139	.161	
N	40281	10702	

The left hand part of the table contains the averages for those women who are childless in our period of interests. The right hand part of the table contains the averages for those women who have their first child in our period of interest, three years before and after the reform (2004 - 2009).
The averages are based on a monthly panel including the period 2001 - 2014 as we regard the time from three years before until five years after giving birth.
High and medium education refer to having a tertiary and secondary certificate, respectively.

We further see that the group of mothers before giving birth shows in all three observed dimensions on average better outcomes than childless women. They seemed to be positively selected in both, labor market outcomes and educational achievement. For any empirical approach using a control group to evaluate the causal effect of giving birth, this results shows the relevance to account for these pre-birth differences. In our approach, we reweight the control group to match the group of mothers by an inverse probability weighting.

4 Econometric Approach

Firstly, we estimate the causal effect of motherhood on various post-birth outcomes before and after the 2007 parental leave reform using a dynamic treatment effects approach as in Fitzenberger et al. (2013). Secondly, we decompose the difference in the

data overreport full-time in the years until 2010. Figure A.1 shows the share of full-time employed over time. The raw full-time values drop dramatically after the implementation of the new procedure, e.g. from 2010 to 2012. Neglecting this data issue may lead in our case to an overestimation of post-birth full-time employment for the pre-reform sample compared to the post-reform sample. The corrected version of the full-time variable adjusts the share of full-time employed downwards while it incorporates the original trend in the 2000er years. Due to the lower intital level, it does not show the large decline form 2010 to 2012.

post- and pre-reform effects of motherhood into the causal reform effect and the effect of changing selection of mothers (section 4.2).

4.1 The causal effect of giving birth

Econometrically, we estimate the average treatment effect for the treated (ATT) on post-birth outcomes based on discrete time data. The treatment is 'first child birth at a certain point in time' against the alternative of waiting (i.e. postponing the birth of the first child to the future, possibly ad infinitum). The alternative of waiting entails hence both, the possibility of never having a child and the possibility of having a child at a later date. Thus, for a certain point in time, we estimate the ATT of having a first child at this point of time versus not yet having a child.[10] This definition of the control group avoids conditioning on future outcomes. We condition however on not having a child yet, both for treated mothers and nontreated women in the time-varying control group (see subsection 4.1.1).

This section draws on Fitzenberger et al. (2013) with the difference that Fitzenberger et al. (2013) align treated mothers and nontreated women by age at first birth whereas we align treated and nontreated women by time at first birth, which allows us to clearly separate treated and nontreated women before and after the reform. There is no fundamental difference between the two approaches because either way the second dimension (age or time, respectively) is controlled for when aligning treated and nontreated women through inverse probability reweighting.

The key identification assumption for our analyis is a dynamic conditional independence assumption (Fitzenberger et al., 2013). It states that – conditional on the variables controlled for – until a certain time period the assignment to treatment in this time period is random, i.e. independent of the potential outcomes. Specifically, our dynamic conditional independence assumption stipulates that given the duration of childlessness and given the covariates, having a first birth within the next year is random.

The dynamic conditional independence assumption can be motivated as follows. Our rich administrative data allow to control for a number of socio-economic characteristics and labor market history. One year before birth, the treated women are not likely to differ systematically from those women who stay childless until shortly after the birth of the child. The exact timing of birth cannot be planned with certainty and may depend upon random circumstances not reflected in long-run labor market choices. It is highly implausible that women plan the exact month of first birth more than a year ahead. At the same time, women differ in their probability to have a child within the next

[10]The treatment effect we estimate is an example for the dynamic treatment approach applied in the context of program evaluation of active labor market policies by Sianesi (2004) for Sweden or by Biewen et al. (2014) for Germany.

year and this is likely to be reflected by the characteristics that are being controlled for. In our analysis, we match treated mothers and nontreated women by observable characteristics twelve months before birth, which is about the time when the decision for becoming pregnant is made.

In our application, for estimating the effects of giving first birth, we do not exclude the alternative of giving first birth at a later date. This corresponds to the fact that fertility decisions are taken jointly with career decisions and that women repeatedly make choices regarding fertility throughout life, such that as a stopping problem each month a women decides to remain childless until the month the decision is made to have her first child. This results in a dynamic selection of first time mothers at a certain date. Thus, using solely a control group of women who do not give first birth until a much later date or who will never have a child would bias the control group due to further dynamic selection towards women with a low propensity of having a child. This bias is likely to be correlated with labor market outcomes (e.g., women with a strong unobserved career orientation may be more likely to exhibit a higher labor market attachment and to never have a child).

Our approach assumes that women giving birth to their first child are comparable before the gestation lag, i.e. at the time before pregnancy, to women who do not give birth at this date. Econometrically, the treatment occurs at the beginning of the gestation period, i.e. the treatment effects already materialize before actual child birth. This approach assumes that women do not know the exact timing of first birth before the gestation period. But they may know the probability of having a first birth now versus later and they may act upon the determinants of this probability (Abbring and Van den Berg, 2003). Assuming a no-anticipation condition with respect to the precise date of pregnancy before the gestation period allows us to match treated and nontreated women at this date. To account for potential short-run anticipation effects at the time of conception, we take as treatment time 12 months before the actual child birth assuming that the gestation period starts 11 months before the actual child birth.

The treatment group in our analysis consists of women who have their first child between the age of 21 and 40. The control group varying with calendar time of actual birth consists of women who are still childless when the 'treated group' gives birth. We measure the treatment effect at a monthly frequency. Our analysis uses an estimate of the average counterfactual outcome for each treated woman based on the individual-specific comparison group of individuals not treated yet.

4.1.1 Temporal alignment

Our evaluation approach requires a temporal alignment between treated mothers and control observations. For the treatment group the treatment time is given by the month (time period) at giving first birth, i.e. the treatment time is defined as the calendar time relative to the month of birth. Women in the control group are aligned to the treated mothers based on not giving first birth in the calendar month of birth of the mother. We do not impute a random placebo treatment time for control observations (as discussed, e.g., in Kleven et al. (2019a)).[11]

As classification window for our benchmark analysis, we consider as suitable control observation a woman who is childless and continues to be so for the following 11 months. When this classification window is reduced, the control group also includes women who are pregnant or who already have concrete plans to become mothers at the date of birth of the treated mother. When this classification window is extended, the control group would be childless for a longer period which conditions on the future thus increasing the dynamic selection of the control group. As robustness checks, we also investigate the sensitivity of our core findings to reducing the classification window to 1 month, i.e. the control group includes all women giving first births after the birth of the treated mother, and to extending the classification window to 24 months, i.e. women giving first births during the seond year after the birth of the treated mother are excluded from the control group.

We consider each treatment month from January 2004 to December 2009 separately. For each of these 72 months, we pool treated women, who have her first child in this month, and all potential control women, those who are childless for at least the following 11 months for the benchmark analysis (the classification window is 1 month or 24 months for robustness checks, respectively). Further, we require all observations to be between 21 and 40 years of age at the treatment time (date of birth).

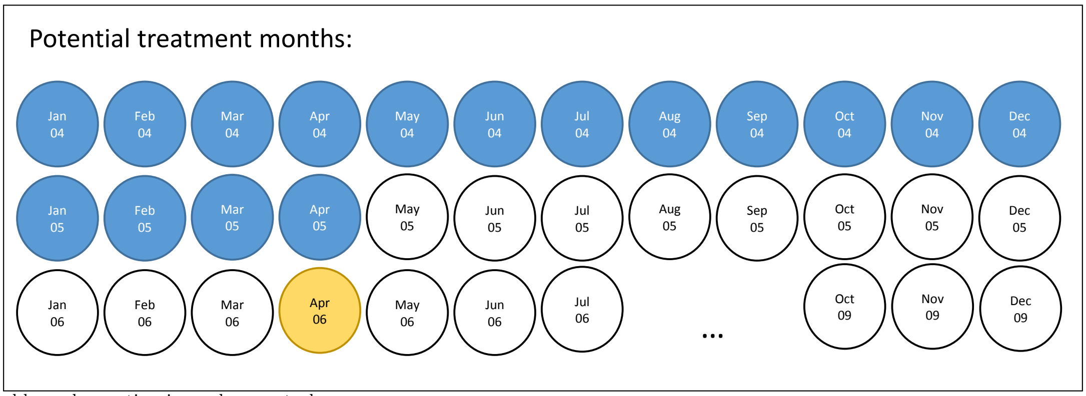
This way, we generate a $person \times treatment-calendar-months\,[2004,1 to 2009,12] \times month-relative-to-treatment-time\,[-35\,to\,59]$ data set used for our analysis. For *month-relative-to-treatment-time*, 0 denotes the month of birth, -35 means 35 months before birth, and 59 means 59 months after birth. Effectively, for the panel data set used for our empirical analysis, women are thus duplicated for each (potential) treatment month $t \in [2004{,}1 \quad \text{to} \quad 2009{,}12]$. A woman may be used in one treatment period as treated, in the month where she becomes mother and potentially multiple times as control observation, in all months which are at least one year before she has her child. $t \in [2004{,}1 \quad \text{to} \quad 2006{,}12]$ represents the pre-reform period and $t \in [2007{,}1 \quad \text{to} \quad 2009{,}12]$ the post-reform period.

Take as an example a woman who gives first birth in April 2006. In the months until April 2005, those which are marked blue in figure 3, she serves as control observation. She is not used in May 2005 as then her own child has less than one year to come. So,

[11] This is because we do not want to restrict the control group to women who do not give birth during the entire observation window. Only for the latter group, it would be plausible to simulate placebo treatment times.

we do not regard her as suitable control observation anymore. The same argumentation holds until March 2006. In April of that year, she give birth to her first child and enters the sample as treated. Afterwards, for all treatment months until December 2009, she is not used, as she is not part of the childless sample anymore.

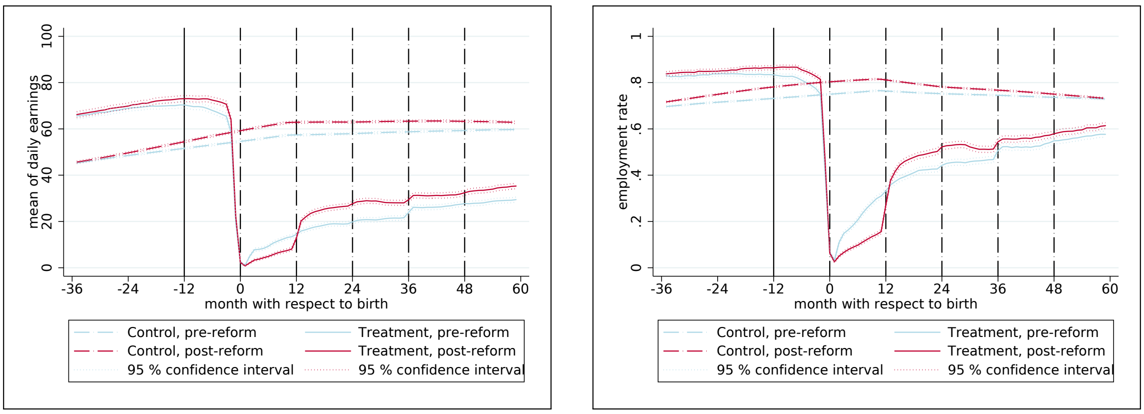
Figure 3: Duplication of observations for temporal alignment



blue: observation is used as control
yellow: observation is used as "treated"
white: observation is not used

Using this approach results in a very large data set. Our sample includes about 50,000 women (see table 1), yielding 10,695 treatment and about 2.6 Million control observations, the latter number is so large because the not-yet-treated women are duplicated for each month observed. The main advantage of the temporal alignment is that the outcomes for both treated and control observation can be now analyzed with respect to the timing of treatment. It is for example possible to plot average daily earnings and labor market participation for treatment and control group and separated by time periods (pre-reform: treatment years 2004 to 2006 and post-reform: treatment years 2007 to 2009).

Figure 4: Raw data after temporal alignment



(a) average of daily earnings

(b) employment rate

We determine the average earnings and the employment rate for the treated and non-treated by month with respect to birth (treatment time) after temporal alignment. The average outcomes are calculated for the pre-reform and the post-reform period as averages of the pooled data set. As to be expected, Figure 4 show that the labor market outcomes for mothers drop sharply to virtually zero at child birth. The recovery afterwards seems to be influenced by the reform as the lines for pre- and post-reform differ significantly. We see further that treatment and control group differ already strongly in the pre-treatment period. Mothers earn more and have a higher probability to be employed in the third and second year before giving birth. As there is yet no child neither present nor anticipated, this difference is a clear evidence for the positive selection of mothers. The graphs are hence not yet informative for conclusions on the causal effect of giving birth. However, we will come back to them in section 5.1 which investigates whether the reform changed the selection of mothers in terms of pre-birth labor market outcomes.

4.1.2 Inverse Probability Weighting (IPW)

To control the selection of treated women, we use monthly propensity scores for giving birth at a certain date. We estimate separate models for the pre-reform and the post-reform period while pooling across treatment months within each subperiod. Specifically, we model the probability of having the first child twelve months from now as a function of socio-demographic characteristics and labor market history. Controlling for past labor market career is crucial for a successful matching.

Under the unconfoundedness of the treatment and perfect overlap in the propensity score, Busso et al. (2014) find that in small samples with unknown propensity score, a modified inverse probability weighting estimator (IPW) performs well. The crucial modification of the IPW estimator involves the normalization of weights for the nontreated women.

Our analysis has to account for the fact that the group of eligible comparison women changes by month of birth, see section 4.1.1. Correspondingly, the alignment between treated and nontreated observations changes as well by age and month of birth, respectively. Recall, that we estimate the average treatment effect on the treated (ATT), i.e. the average treatment effect for the actual mothers. This treatment effect is the effect of giving birth in different months t on labor market outcomes at different *months-relative-to-treatment-time* which can be estimated by IPW in two steps. Firstly, we estimate the treatment effect separately for each treatment month t, (1) and secondly, we weight these treatment effects of the treatment month by the number of treated observations (2):

$$\hat{\theta}_t = \frac{\sum_{i=1}^{n} T_{i,t} Y_{i,t}}{\sum_{i=1}^{n} T_{i,t}} - \frac{\sum_{i=1}^{n} (1 - T_{i,t}) \hat{W}_{i,t} Y_{i,t}}{\sum_{i=1}^{n} (1 - T_{i,t}) \hat{W}_{i,t}} \tag{1}$$

$$\hat{\theta} = \frac{\sum_{i=1}^{n} T_{i,t}\hat{\theta}_t}{\sum_{t=1}^{T}\sum_{i=1}^{n} T_{i,t}} \tag{2}$$

n is the total number of women in the data set. $T_{i,t}$ denote the treatment dummy variable for individuals i in treatment month t (treated and non-treated), respectively. The first fraction in equation (1) is the (unweighted) average of the treatment group and the second fraction is the average of the control group with weights $\hat{W}_{i,t}$:

$$\hat{W}_{i,t} = E_{i,t} \times \frac{\hat{p}_t(X_{i,t})}{1 - \hat{p}_t(X_{i,t})}. \tag{3}$$

$E_{i,t}$ is a dummy variable for eligibility as nontreated observation which takes the value of one if woman i can be used as a control observation for treatment month t. Otherwise, $E_{i,t}$ is set to zero giving this woman zero weight in month t. Thus, $E_{i,t}$ is a dynamic non-treatment dummy. In our benchmark analysis sets the classification window to 12 months. $E_{i,t}$ will be one if i is childless for at least 11 months after the respective time period t.

$\hat{p}_t(X_{i,t})$ denotes the estimated probability to have a first child in time period t as a function of covariates $X_{i,t}$ which can vary over time. The application of the estimated weights $\hat{W}_{i,t}$ leads to a reweighting of the nontreated women according to the odds–ratio of having a child within the next year.

The IPW reweighting estimator has the advantage of not relying on a tuning parameter. Moreover, it is easy to implement and standard errors are readily obtained by bootstrapping. The probability to give first birth in month t given the characteristics $X_{i,t}$ is estimated by a probit regression based on the observations in the aforementioned duplicated data set at $month - relative - to - treatment - time = -12$ months, i.e. one year before giving birth in the duplicated data set. The characteristics we use to determine the chance to be a mother are an indicator for working in former East-Germany ($east_{it}$) and dummies for medium and high education (low education serves as reference category). Further, we include fixed effects for the calender month, years and age at treatment. Arguably, most important are the controls for the employment history. We include three times 24 variables for the earnings, employment and full-time status of the second and third year before giving birth (month -35 to -12 with repect to giving birth). This regression is conducted separately for four age groups (women aged 21 to 25, 26 to 30, 31 to 35 and 36 to 40) and separately for the pre-reform and the post-reform period. We use the following specification for treatment month t:

$$P(mother_{i,t} = 1|X_{i,t}) = \Phi\Bigg(\beta_0 + \beta_1 east_{it} + \sum_{l=22}^{40} \beta_l I(age_{it} = l) + \sum_{j=1}^{3} \gamma_j I(edu_{it} = j) \\ + \sum_{k=-35}^{-12} (\delta_k earn(k)_{it} + \alpha_k empl(k)_{it} + \theta_k ft(k)_{it}) + \lambda_m + \mu_y\Bigg) \quad (4)$$

with $\Phi(.)$ representing the standard normal distribution function and $X_{i,t}$ comprising the aforementioned partly time-varying covariates considered in eq. (4). The fitted value of $P(mother_{it} = 1|X_i)$ are used as $\hat{p}_{t(i)}$ in eq. (1).

4.1.3 From effect of motherhood to reform effects

The estimates $\hat{\theta}_{pre}$ and $\hat{\theta}_{post}$ gives the approximation to the causal effect of motherhood for pre- and post-reform periods. Taking the difference would be a very simple approach to obtain an estimate for the reform effect:

$$RE = \hat{\theta}_{post} - \hat{\theta}_{pre} \quad (5)$$

This approach neglects potential differences in selection and should be only used in a regression discontinuity design with a narrow time frame around the reform. Frodermann et al. (2020); Kluve and Schmitz (2018) restrict their projects on the 2007-reform to first time mothers between October 2006 and March 2007. They argue that the mothers of this three months period around the reform were unaware at the time of conception and that therefore selection into motherhood is no threat to them. In section 6.1, we present results calculated like in (5) to compare our ATT-based approach to the two papers of Frodermann et al. (2020); Kluve and Schmitz (2018).
However, the RDD estimate may miss possible selection effects of the reform - recall that the reform was intended to increase fertility. The reform can manifest itself in the change in the selection of mothers and in the change in the effect of motherhood. By construction, the local view of RDD reveals only the latter. The selection effect on the other hand has not gained much attention. For a comprehensive assessment of the reform, these selection effects are important and therefore, we want to outline in the following section an approach for estimating both, the causal reform effect on mothers and the effect caused by differences in the selection of women into motherhood.

4.2 Selection and Reform effects

A simple comparison of the pre- and post-birth motherhood effects on post-birth outcomes may confound reform and selection effects. In the following, we describe our

approach to separate the two effects. Accounting for the selection effect of the reform, we then determine the pure reform effect on the causal effect of motherhood.
In section 5.1, we present regression results for a regression of the probability of becoming mother on pre-birth labor market history and its interaction with the reform implementation. We show that women with higher pre-birth earnings have an increased probability to become mother in the post-reform period. This differences in pre-birth outcomes can be thought of the first dimension of selection effects. It may drive differences in post-birth outcome (the causal effect of motherhood) between pre- and post-reform periods, the second dimension of selection effects.
To assess this second dimension of the reform effect on the selection of mothers with regard to the causal effect of motherhood on post-birth outcomes, we use a second IPW to reweight the sample of pre-reform observations to the sample of post-reform mothers. This reweighting allows to estimate the counterfactual causal effect of motherhood - the ATT - on post-birth outcomes that would have applied for a sample of mothers as observed in the post-reform period but facing the pre-reform regulations.
Estimating this counterfactual ATT is implemented in two steps:

1. We take the post-reform mothers and separately the sample of pre-reform mothers and control women and run two Probit regressions to estimate the probability that a woman gives birth in the post-reform period as a function of pre-birth characteristics and history as observed 12 months before birth. This probability $P(post_{i,t} = 1, mother_{i,t} = 1|X_i)$ is specified as a function of age, education, region (east, west), and labor market history (earnings, employment, full-time employment) during the second and third year before birth. The labor market outcomes are averaged for the four half-year periods $[-35, -30]$, $[-29, -24]$, $[-23, -18]$, and $[-17, -12]$ (months-relative-to-treatment). This aggregation of the labor market history and the absence of year fixed effects are the only differences to the regression of (4).

2. Then, we estimate the counterfactual ATT of motherhood by reweighting pre-reform observations with

$$\hat{G}_{i,t} = E_{i,t} \times \frac{\hat{P}(post_{i,t} = 1, mother_{i,t} = 1|X_{i,t})}{1 - \hat{P}(post_{i,t} = 1, mother_{i,t} = 1|X_{i,t})},$$

where $\hat{P}(post_{i,t} = 1, tr_{i,t} = 1|X_i)$ are the fitted probabilities for the Probit regression in step 1. The counterfactual ATT is then given by

$$\hat{\theta}(\text{post-sample,pre-effect})_t = \frac{\sum_{i=1}^{n} T_{i,t}\hat{G}_{i,t}Y_{i,t}}{\sum_{i=1}^{n} T_{i,t}\hat{G}_{i,t}} - \frac{\sum_{i=1}^{n}(1 - T_{i,t})\hat{G}_{i,t}Y_{i,t}}{\sum_{i=1}^{n}(1 - T_{i,t})\hat{G}_{i,t}} \qquad (6)$$

$$\hat{\theta}(\text{post-sample,pre-effect}) = \frac{\sum_{i=1}^{n} T_{i,t}\hat{\theta}(\text{post-sample,pre-effect})_t}{\sum_{t=1}^{T}\sum_{i=1}^{n} T_{i,t}} \tag{7}$$

The approach is implemented analogously to eq. (1). The treatment effect for month t is given by the difference of treatment and control average, both averages are now reweighted with weights $\hat{G}$, the estimated odds-ratio to appear as treated mother. The pre-reform observations, treatment periods from January 2004 to December 2006, receive a weight $\hat{G}_{i,t}$ which is high for those who are comparable to the mothers of the year from 2007 to 2009. Here, $i = 1, ..., n$ denotes the sample of women observed for the pre-reform period.

Let the factual ATTs of motherhood in the pre- and post-reform period - as given by eq. (1) - be denoted by $\hat{\theta}$(pre-sample,pre-effect) and $\hat{\theta}$(post-sample,post-effect), respectively. The raw difference between pre-reform and post-reform period is given by the differences of these ATTs (the equivalent of equation (5) for the RDD setting), i.e. by

$$\text{raw\_diff} = \hat{\theta}(\text{post-sample,post-effect}) - \hat{\theta}(\text{pre-sample,pre-effect})\,. \tag{8}$$

However, this is not the causal reform effect because the treated mothers before and after the reform may differ systematically in a socio-economic characteristics and labor market history. These difference in the selection of mother may itself result in a change in the causal effect of motherhood. The effect on the change of the selection of mother (the second dimension of the selection effect) on the ATT can be estimated by

$$\text{sel\_eff} = \hat{\theta}(\text{post-sample,pre-effect}) - \hat{\theta}(\text{pre-sample,pre-effect})\,. \tag{9}$$

This selection effect quantifies how the counterfactual effect of motherhood for the post-reform sample of mother would have differed in the pre-reform period from the effect of motherhood as observed for the pre-reform sample. This difference arises because of the change in the sample of mothers between the pre-reform sample and the post-reform sample.

The causal reform effect of the ATT of motherhood for the post-reform sample of mothers is then given by

$$\begin{aligned}\text{reform-eff} &= \text{raw\_diff} - \text{sel\_-eff}\\ &= \hat{\theta}(\text{post-sample,post-effect}) - \hat{\theta}(\text{post-sample,pre-effect})\end{aligned} \tag{10}$$

This version of the reform effect measures for the post-reform sample of mothers how the causal effect of motherhood changes due to the reform. Section 5.2 presents the results on the labor market outcome. It will illustrate the different effects in equation (10).

5 Baseline Results

As outlined, we start this section by showing how the reform changed the selection of women into motherhood in terms of their pre-birth labor market history. Afterwards, we present the effects of the 2007 reform on post-birth outcomes. We regard earnings, employment and full-time employment as labor market outcomes. Additionally, the effects on second order fertility are evaluated. We also show how the effects differ for different age groups of women. The results of the RDD-setting conclude this section.

5.1 Selection effects of the reform

One of the goals of the parental leave benefits reform was to increase the low birth rate in Germany, especially for high-earning women who have a high attachment to the labor market. We first investigate how the reform changes fertility. Then, we investigate how the selection of mothers changes with regard to pre-birth labor market outcomes. As part of the analysis of the reform effect on the causal effect of first birth in the subsequent section, we then estimate how the reform changes the selection of mothers with regard to post-birth outcomes.

To estimate the effect of the reform on fertility, we follow Raute (2019) and estimate the following difference-in-differences regression

$$child_{iy} = \alpha lm\_outcome_{iy} + \beta lm\_outcome_{iy} \times postreform_y + \gamma X_{iy} + \lambda_y + \epsilon_{iy} \qquad (11)$$

where $child_{iy}$ is a dummy for having a first child in year y, $lm\_outcome_{iy}$ is the labor market outcome of interest (log-earnings or the identifier to have gross-earnings above the threshold of 8550 €), $postreform$ is the postreform dummy, which is equal to one for 2007 onward, X_{iy} are other covariates, and λ_y is a fixed effect for year y. This regression measures differences in the reform effect on fertility by labor market outcomes relative to the general time trend.

The right panel of Table 2 shows the time trend in the birth rate among women aged 21 to 40 in our sample, who have not had a child until the year before. There is a positive trend in the birth rate increasing steadily from 4.2% in 2004 to 4.8% in 2009. There is no disproportionate increase in the birth rate in 2007 (+ .1 percentage points compared to 2006) but a disproportionate increase in 2008 (+ .3 percentage points compared to 2007). Though this should not be interpreted as causal, it is suggestive of a positive reform effect considering the delay of 9-10 months between conception and actual birth.

The left panel of Table 2 involves two regression specifications. The upper specification controls for non-employment and log earnings, both as observed in the second calendar year before birth. The birth rate increases due to the reform significantly for employed

Table 2: Share of women entering motherhood and selection into motherhood

	Difference-in-differences regression		Descriptive share of women entering motherhood		
	(1)	(2)	year	share	N
log earnings	0.0145***	0.0103***	2004	.042	40831
	(28.60)	(19.61)	2005	.043	40726
log earnings	0.0016**	0.0021***	2006	.044	40261
× postreform	(2.21)	(2.90)	2007	.045	39661
non-employment	0.1140***	0.0756***	2008	.048	38723
	(24.04)	(15.47)	2009	.048	37375
non-employment	0.0158**	0.0204***	total	.045	237577
× postreform	(2.27)	(2.95)			
earnings > 8.550 €	0.0328***	0.0257***			
	(30.12)	(22.69)			
(earnings > 8.550 €)	0.0024	0.0029*			
× postreform	(1.50)	(1.86)			
year FE	yes	yes			
controls		yes			
N	237577	237577			

t statistics in parentheses

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

Note: The left panel (first two columns) shows estimates for two difference-in-differences regressions of the probability of giving birth on the labor market history (11). The earnings variable measures log cumulated earnings during the second calender year before giving birth (zero earnings are coded as zero and as one for the non-employment identifier). For the bottom panel, the labor market outcome evaluated is an identifier whether the cumulated earnings exceed 8550 Euro in the second year before giving birth. The covariates used in column (1) are the respective labor market outcome, the interaction term and year fixed-effects. Column (2) includes additionally an identifier for former East-Germany, controls for medium and high education and age fixed-effects. The standard error are clustered at the individual level. The right panel shows the share of women who give first birth in the year of observation.

women with higher earnings. For employed women earnings above (below) annual gross-earnings of 16500 €, the increase in fertility is larger (smaller) than for non-employed women. Following Raute (2019), the lower specification controls for whether earnings lie above or below 8.550 €, which is the threshold generating a group of reform-winners and a group whose benefit remained on average unchanged, see section 2. Women with earnings below 8.550 €, who were in principle eligible for the means-tested benefits before the reform, received lower parental benefits after the reform. The findings reveals a weakly significant additional increase in fertility for women whose earnings were above the threshold. Altogether there is evidence for a particularly strong increase in fertility among women with higher earnings. According to the first specification, the increase in fertility for women at the 75th percentile (gross-earnings 30000 €) of

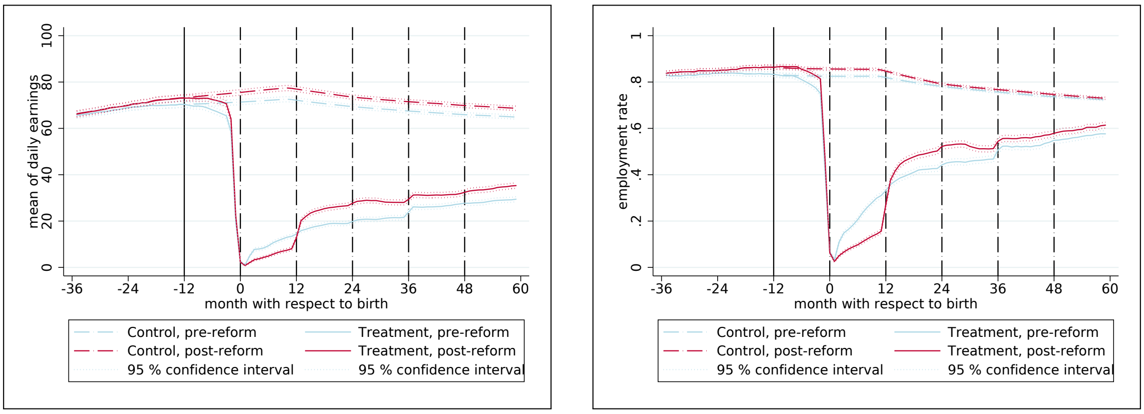
the earnings distribution is 1.9 percentage points higher than for women earning 8.550 €(38th percentile).
Next, we investigate how pre-birth labor market outcomes changed after the reform based on the evidence in Figure 4. In the second and third year before birth, earnings and employment among mothers increases after the reform and the increase grows steadily from month -36 to month -12. However, labor market outcomes also improve for the control group by a similar magnitude, and even more strongly for the control group. Thus, the selection of mothers improves with regard to pre-birth labor market outcomes in the second and third year before birth. However, the same applies to the outcomes for the control group.

5.2 Causal effects of first birth and reform effects

5.2.1 Findings for labor market outcomes

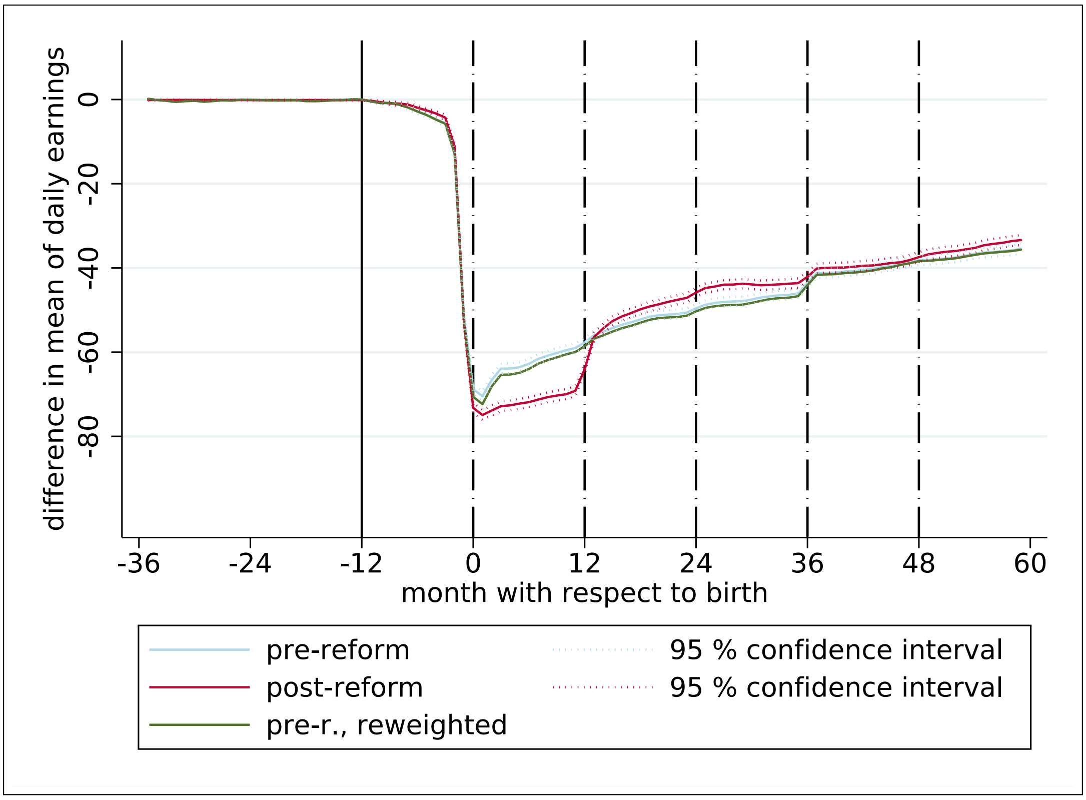
Applying the IPW-strategy described above results in the earnings and employment profiles when aligning the control group to the treatment group with regard to the covariates considered and labor market history up to month -12, as depicted in figure 5. Compared to figure 4, showing the profiles after temporal alignment, the two lines for the control group are shifted (dotted light [pre-reform] and dotted dark [post-reform]). In the pre-treatment period until one year before giving birth, the profiles controls results match very precisely the treatment results, both for the pre-reform and the post-reform sample. This makes us confident that our econometric specfication for the propensity score is sufficiently flexible to achieve comparability between treated mothers and non-treated women.
The same holds true for the reweighted pre-reform sample using the weights of (6) (results not shown here to save space but available upon request). After reweighting, this group matches the labor market history of the post-reform treatment group.
The differences between the profiles for the treatment and the control group after IPW provides the estimated causal effect of giving first birth, i.e. the ATT of giving birth now against the alternative of waiting. The ATT for earnings shown in figure 6 is depicted post-reform, pre-reform, and for the reweighted pre-reform sample, respectively. These baseline effects follow for all three groups similar pattern, until one year before giving birth there is no effect at all, shortly before giving birth, there is a dramatic decline in earnings and than a slow recovery. There are yet also significant differences between pre- and post-reform groups, which entails the estimated reform effect. Note that all effects are precisely estimated as the confidence intervals are quite small.
The reform has its largest effects in the short run – the first three years after giving birth. Replicating key findings results in the literature, there a strong fall in earnings

Figure 5: Treatment and control group after IPW



(a) average of daily earnings (b) labor market participation

Figure 6: Effects of giving birth on daily earnings



during the first year and a moderate increase in the second and third year. The general job protection for mothers ends after three years and a sizeable number of mothers returns to work after month 36, thus reducing the positive but still significant earnings effect in the medium run – years four and five after giving birth. We estimate that the reform increased daily earnings in the fifth year after giving birth by 1.83 Euro (an increase of more than 650 Euro in annual earnings)(see the estimates in the top panel of table A.1 in the appendix). Furthermore, there is a significant increase in earnings in the first year before giving birth. To our knowledge, this effect during the year before birth has not been noted so far in the literature. After the reform, mothers have an incentive to maintain employment until the beginning of the job protection period because parental benefis depend upon previous earnings. This "entitlement effect" is significant at the 92 percent level. The effect is visible for the month -9 to -4, amounting to 0.64 Euro in daily earnings (230 Euro annually).

We do not detect a significant difference between the unweighted and reweighted pre-reform sample with regard to outcomes from month -11 onward. Such a difference would be the selection effect of equation (10) as the latter group provides the counterfactual situation when the reform would not have changed the selection of mothers with regard to the causal effect of first birth. This finding complements our findings in section 5.1, as the selection effects found there are rather small. Note that by construction the differences in figure 6 during months -36 to -12 are insignificant. While the reform is likely to change the composition of mothers towards higher fertility among high-earning women and there is a small positive entitlement during the year before birth, there is no noticeable selection effect concerning the reform effect on the causal effect of first birth. Formally, the reform-eff(ect) in equation (10) basically coincides with the raw differences in the post- and pre-reform ATT's of first birth.

The reform effects for employment are generally very similar what we have seen for employment as shown in figure 7 and the top panel of table A.2 (the same holds for the patterns of the causal effects of first birth, pre- and post-reform, the graphs are available upon request). There is a negative reform effect during the first year after birth, positive reform effects during the second and third year, and smaller positive reform effects in the medium run. In the post-reform period, mothers are up to 10 percentage points more likely to be employed in the second and third year and only 2.5 percentage points more likely to be employed afterwards compared to mothers in the pre-reform period. Graphical inspection of figure 7 suggests that the reform reduced the employment interruption for some mothers from 36 months in the pre-reform period to less than 15 months in the post-reform period - recall that the length of the job protection period is 36 months and remained unchanged with the reform. There is also evidence for a small entitlement effect during the first year before giving birth, which amounts to one percentage point.

The reform effects for full-time employment are shown in figure 8 and table A.3. The reform effects during the first three years go in the same direction as for overall employment but the effects are much smaller in size. The medium run effects in the forth and fifth year are negative and insignificant. The entitlement effect before giving birth is of similar size as the effect on employment indicating that the jobs maintained until maternity protection are predominately full-time employments.

Hence, in the medium run the reform increased earnings and employment without

Figure 7: Effects of giving birth on employment rate

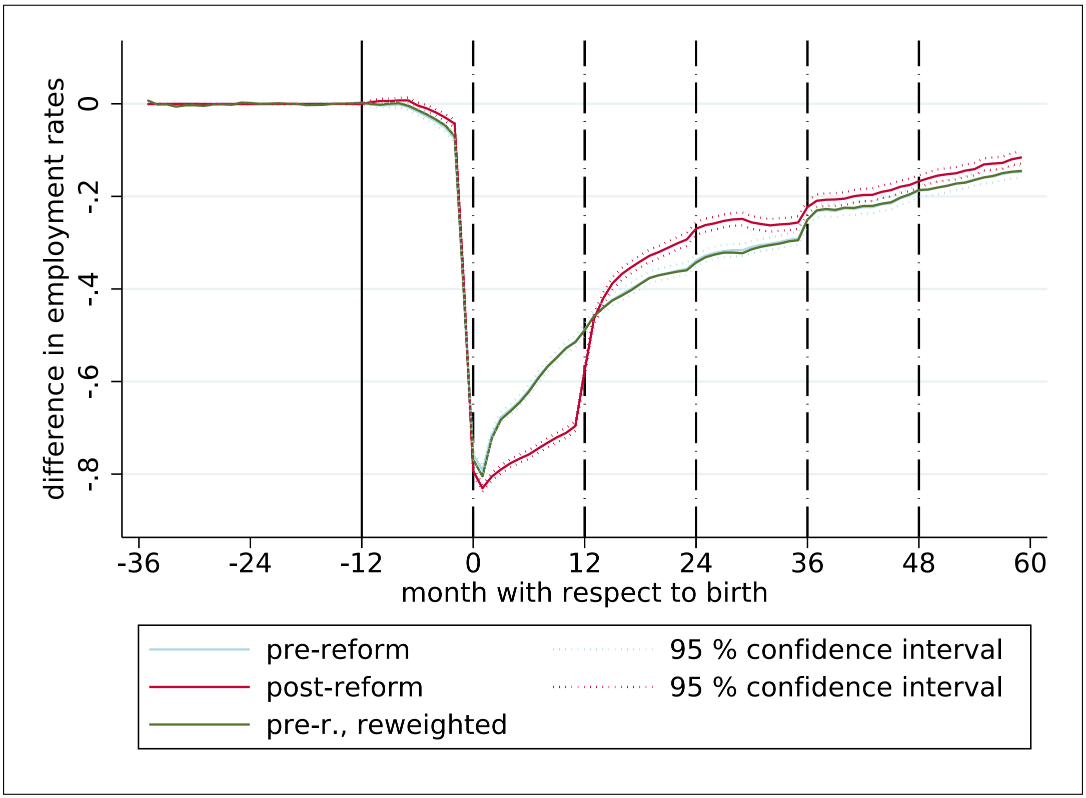
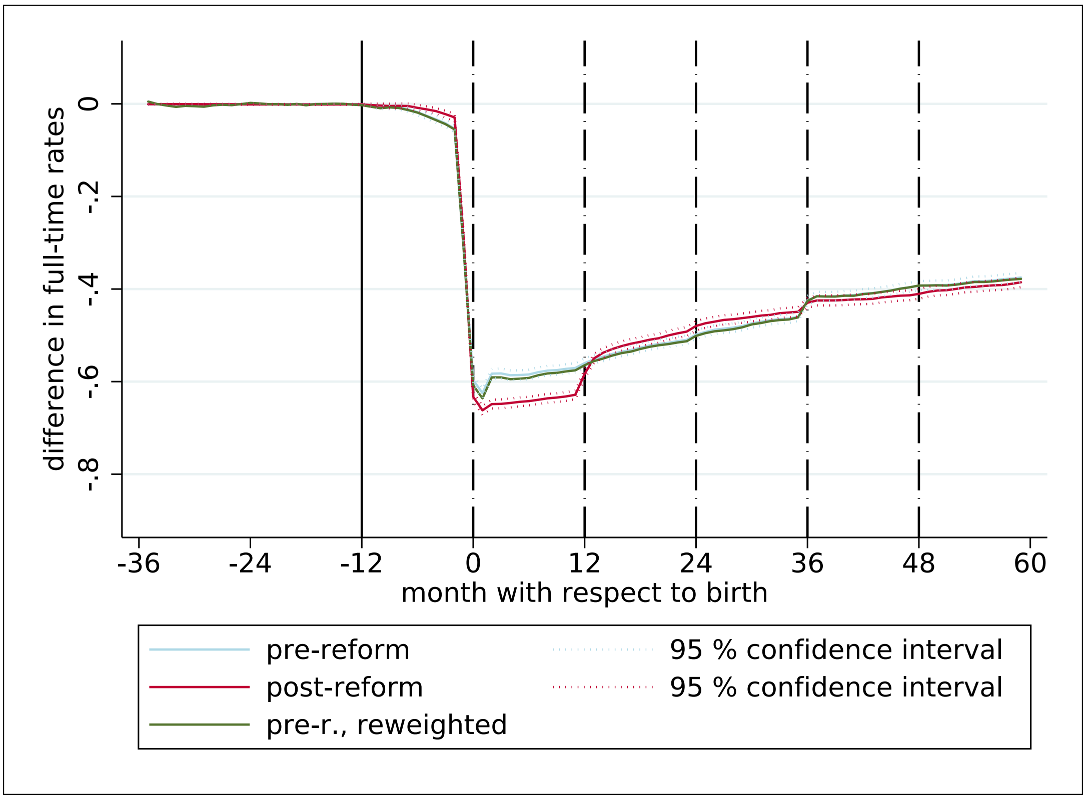


Figure 8: Effects of giving birth on full-time rate



increasing full-time employment, suggesting the positive medium run effects on earnings and employment stem from part-time jobs. We further conclude that the reform did not change the selection of women into motherhood with regard to post-birth outcomes (detailed results confirming this for employment are available upon request) and there is a small positive entitlement effect during the 12 months before birth. Accounting for these selection effects does not affect the estimated reform effects.

5.2.2 Heterogenous effects by age of mothers and pre-birth earnings

In this section, we investigate how the effects differ for different groups according to the age of mothers and the amount of pre-birth earnings.
Results on how the reform effects differ by the age of the mother are shown in figures A.2 to A.4 and the upper panels of tables A.1 to A.3. Generally, the medium run reform effects on employment and earnings are positive for mothers above age 26, which drive in fact the overall positive effect. The medium-aged group, age 31 to 36, seems to benefit most. For full-time employment, there are negative effects for all except this group. The group 26 to 30 year of age, has the interesting result of significant positive effects on overall employment but negative effects on full-time employment. Further, the youngest age group shows negative but insignificant effects for all three outcomes.
Concerning pre-birth earnings, we split the sample in three groups according to their earnings in the third and second year before giving birth (months -35 to -12) when the women are unaware of becoming mother or at least the concrete timing. We base the subsampling on the results of section 5.1. The results (shown in table 2) suggest that two groups of women have a strongly increased propensity to become mother after the reform. These groups are women in non-employment (with zero earnings) and those with earnings above 16500 Euro yearly gross earnings. The group of "low earners" – in between with positive earnings but below 16500 Euro – have a decreased likelihood to become mother after the reform compared to non-employed. Consequently, we show the results for women with "zero earnings", "low earnings" (positve earnings but below 33000 Euro in the two years period) and "high earnings" (earnings above 33000 Euro in the two years period), see figures A.2 to A.4 and the lower panels of tables A.1 to A.3.
On one hand, we find the high earners to benefit above average from the reform. Their gains amount to three Euro in daily earnings in the fifth year. Further, their earning increase in the year before giving birth, the "entitlement effect" is above average. On the other hand, the low earner show no significant positive results. On the contrary, they are estimated to have a decreased likelihood to be full-time employed in the medium run by two percentage points.
The results on the zero earners are very interesting as that is a small group which is

not detectable in the framework of Frodermann et al. (2020) but which shows quite different results as the low earners. Since the group is by far the smallest one, it has the largest confidence interval and a lack of statistical power might be an issue of our analysis. However, their results indicate a very strong positive entitlement and medium-run effects on earnings. The point estimates are in both cases twice as much as the one for the sample average. It seems convincing that the increased incentives to stay employed before giving birth (as the benefits depend on the earnings in that time) keep the becoming mothers in employment and thereby lead to a lasting labor market attachment.

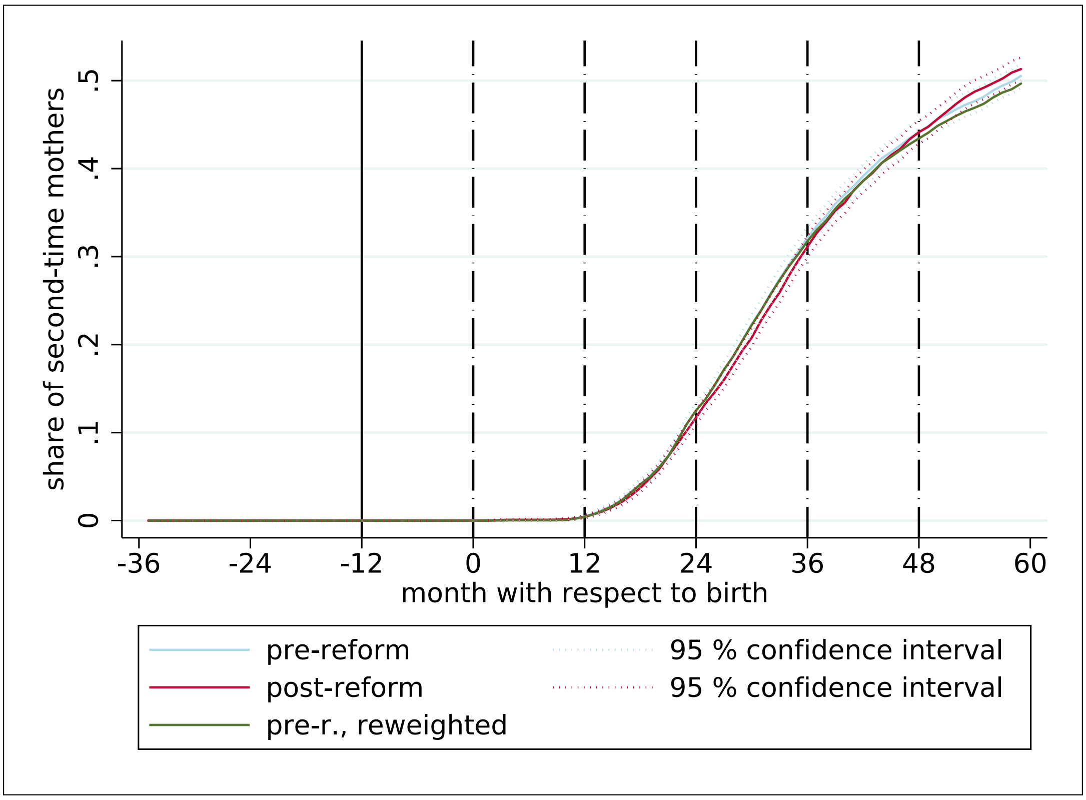
The findings on low and high earners are similar to those reported in the existing literature separating results by pre-birth earnings (reform winners and losers). Bergemann and Riphahn (2011), Frodermann et al. (2020), and Kluve and Schmitz (2018) find that the reform winners (with high pre-birth earnings) benefit above average while the reform losers do not gain in the medium run.

The results on the age groups have not been shown explicitly so far. But given that medium aged mothers may have on average higher pre-birth earnings than young mothers, they are well in line to the existing and our own results. The investigation of zero earners is also an interesting contribution to the literature.

5.2.3 Findings for second order fertility

One of the goals of the parental benefits reform was to increase fertility. Here, we investigate how the reform affected second order fertility, i.e. the propensity to have a second child, using the incidence of giving birth to child a second time as outcome variable. There are no significant effects, As shown in figure 9 and table A.4, there are only a small and insignificant differences in the cumulative incidence of a second birth pre- and post-reform. This finding holds both for the entire age group and for different five-year age intervals. These results mostly confirm Cygan-Rehm (2016), who finds temporary but no lasting reform effects on second order fertility.

Figure 9: Effects of giving birth on second order fertility



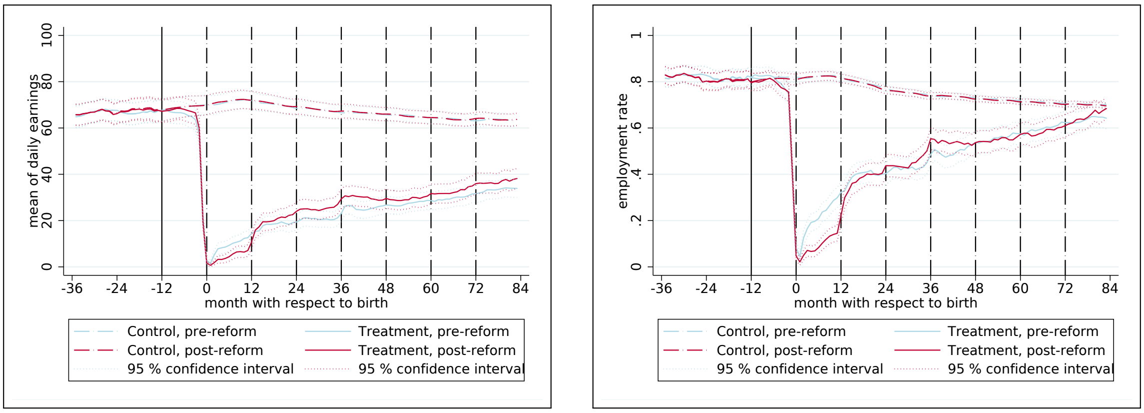
6 Alternative estimation strategies

6.1 Regression Discontinuity Design around reform date

To contrast our baseline findings with RDD estimates as in Frodermann et al. (2020); Kluve and Schmitz (2018), we now use the same narrow time range of three months before and after the reform (October 2006 to March 2007) as described in section 4.1.3. The RDD studies argue that these mothers were unaware of the reform at the time of conception such that the reform did not change the selection of mothers in this sample. Thus, while the RDD approach estimates the reform effect for the sample of mothers at the time of the reform, this is a local effect based on a small sample of 840 mothers pre- and post-reform (our baseline analysis is based on a sample of 10.695 mothers) and does not allow to estimate the reform effect on fertility and the overall selection of mothers. The RDD estimates in Frodermann et al. (2020); Kluve and Schmitz (2018) are based on the sample of mothers only. However, the fact that we use a control group of non-mothers before and after the reform should not make a difference under the RDD identification assumptions.

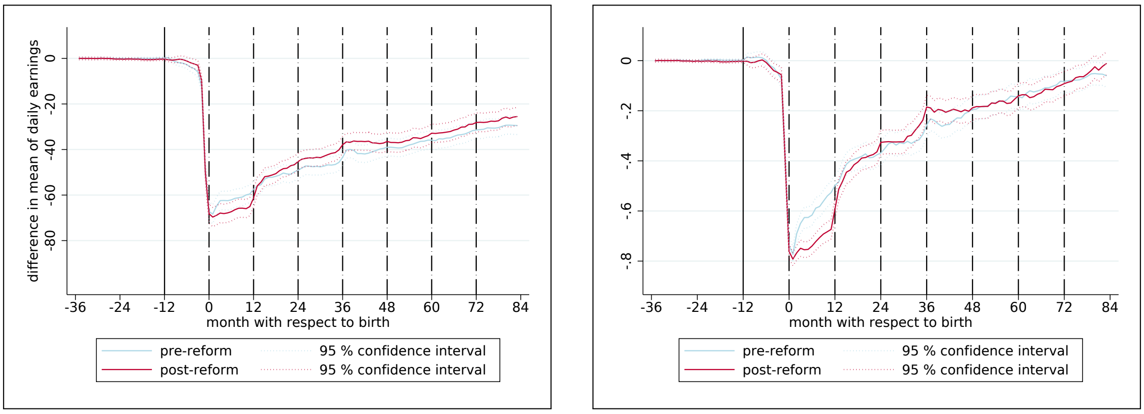
Figure 10 depicts earnings and employment for the mothers pre- and post-reform. The evidence confirms that the selection of mothers did not change with regard to labor market history before conception because pre- and post-reform mothers show the same outcomes during the second and third year before birth. The controls are reweighted using the weights in equation (3). Because selection is unchanged, no reweighting of pre-reform mothers is needed to mimic the post-reform sample of mothers.

Figure 10: Treatment and reweighted control group, RDD-sample



(a) earnings (b) employment

Figure 11: Causal effect of giving birth, RDD-sample



(a) earnings (b) employment

The difference between treatment and control group gives the causal effect of giving birth. Figure 11 shows these effects for the pre- and post-reform period. The difference between the two lines provides the estimated causal reform effect. In line with the literature and our baseline results above, there is a strong negative reform effect during the first year after giving birth. For earnings, this effects turn positive shortly after. In the third and fourth year after giving birth the positive effect is sifnificant on 95 percent level. Afterwards the effects on earnings loose significance but they show until the seventh year roughly the same size which would imply that there are small but lasting gains. For employment, the changes through the reform from the second year onward seem quite small and vary in sign. Hence, the positive average effects on earnings do not stem from an increase in employment.

The RDD-sample allows us to investigate a longer post-birth period. It is remarkable that both for the pre- and post-reform period the causal effect of birth on employment is close to zero (figure 11), i.e. employment rates of mothers have converged to those of the controls seven years after birth. This is partly driven by a decline of about 10 percentage points in the employment rate among the controls (not-yet-mothers), which can be related to the birth of the first child (fig 10). However, most of the convergence is due to the increase in employment among mothers again accelerating in years 6 and 7 after birth. In contrast, a large gap in earnings remains, reflecting the high level of part-time employment among mothers.

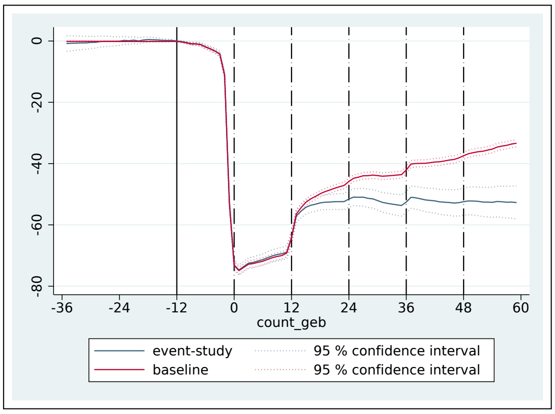
6.2 Event study approach without control group

In two widely cited studies, (Kleven et al., 2019a,b) use an event study approach without control group to estimate the causal effect of birth on post-birth labor market outcomes for mothers and how these post-birth outcomes changes under different family policy scenarios. We implement this approach in our setting estimating regressions of the type

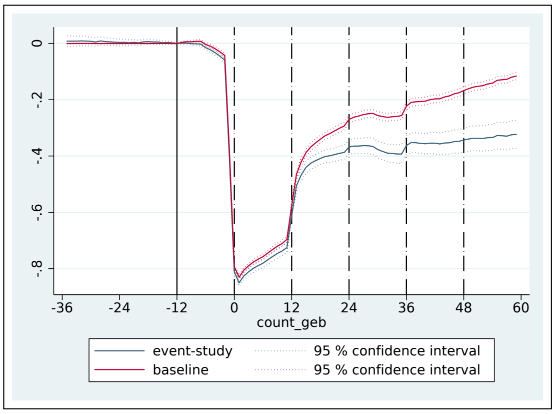
$$\begin{aligned} Y_{iym} = & \sum_{\substack{j=-35 \\ j\neq -12}}^{59} \alpha_j I(ym - t_i = j) + \sum_{k=-35}^{59} \beta_k I(ym - t_i = k) postreform_i \\ & + \pi east_{iym} + \sum_{j=1}^{3} \theta_j I(edu_{ism} = j) + \sum_{l=19}^{45} \gamma_l I(age = l) + \lambda_y + \delta_m + \epsilon_{iym} \ , \qquad (12) \end{aligned}$$

for the sample of mothers only. Here, Y_{iym} is the labor market outcome of interest of mother i in year y and calendar month m (January,...,December) [ym is an integer counter for the observation month] and t_i is a counter of the month of birth. Hence, $ym - t_i (= j)$ measure time (in months) to birth, α_j ($j = -35, ..., 59; j \neq -12$) measures how the outcome variable varies by time to birth relative to month -12 in the pre-reform period, and β_j how this changes in the post-reform period, i.e. the reform effect. Kleven et al. (2019a,b) denote these effects after birth (α_j pre-reform and $\alpha_j + \beta_j$ post-reform, for $j \geq 0$) the child penalty of birth. This measure of the (causal) effect of birth takes the difference of post-birth outcomes to the pre-birth outcome at -12, i.e. the latter is taken as the comparison level for the non-birth outcome after controlling for general time and age effects and other time-invariant individual characteristics. While this approach may work well to account for the strong immediate effects after birth, which are likely to be much larger than the changes for the counterfactual of not having a child during the same time period. However, as Kleven et al. (2019a) suggest themselves, the event study estimates for the medium run post-birth outcomes may be biased because the counterfactual may entail larger changes as well, which may not be captured by the general time and age effect. Put differently, the pre-birth outcomes for mothers, corrected for the general time and age effects, may not an appropriate estimate for the counterfactual outcome of not having a child at this point of time.

Figure 12: Effects of giving birth, baseline compared to event study estimates without control groups



(a) earnings



(b) employment

Figure 12 contrasts the estimated causal effects of first birth for earnings and employment using our baseline estimates in figures 6 and 7 and the event study estimates for α_j. Both for employment and earnings, the event study estimate shows basically zero effects during the second and third year before birth and also matches well the baseline estimates one year pre- and post-birth. However, the event study estimates diverge strongly from the baseline estimates from the second post-birth year onward. Hence, the event study estimates imply a persistent and much larger child penalty compared to our baseline estimates. This findings is driven by the fact that the pre-birth outcomes of mothers provide a higher comparison level compared to the control group used by our baseline estimates. And the gap increases over time on the one hand due to the general upward trend in employment and earnings among women pre-birth and on the other hand due to the growing share of women in the control group having a child (recall that the control group of the baseline estimates excludes women who give birth during the first year after birth for the treated mother).

To investigate this further, we explore two alternative control groups. The larger control group also includes all childless women at the time of birth of the treated women who give afterwards from month 1 onward. The smaller control group only includes future mothers who give birth at least two years after the birth for the treated mother. Figure A.5 in the appendix shows the differences in the effects of giving birth in the pre-reform period for alternative estimation strategies compared to the baseline estimates. The findings show only minor differences between the three control group approaches from the third year onward. The RDD approach yields noisier effect estimates which differ noticeably from the baseline during the first year after birth for earnings and during the first three years for employment, but which are very similar to the control group estimates in the medium run. The event study estimates show the strongest gap to the baseline estimates. This gap increases over time and it becomes much larger than the gap between the different control group estimates from the third year onward.

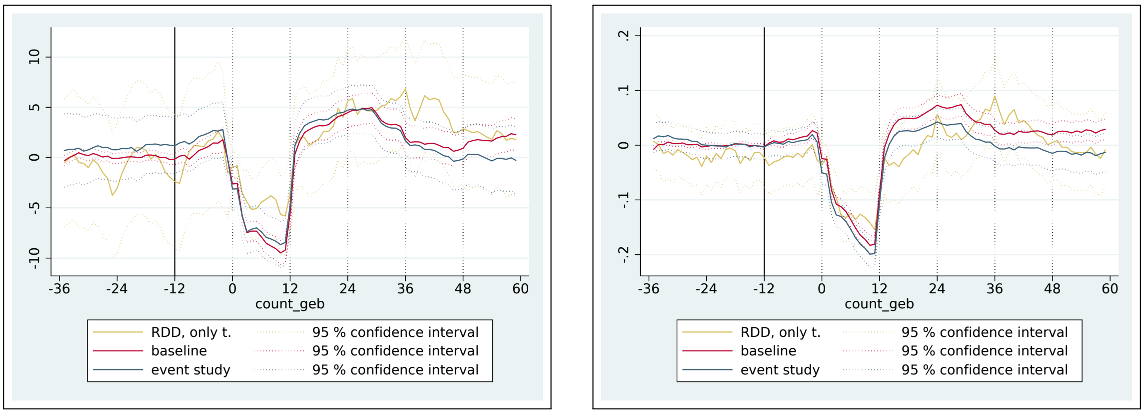
6.3 Comparison of estimated reform effects

Figure 13 shows estimated reform effects based on our baseline control group estimates, the RDD sample without control group (RDD with a control as presented in section 6.1 would show almost the same results), and the event study approach without control groups. The most noteworthy finding is that the baseline results are very precisely estimated and often differ significantly from the point estimates of the two other approaches, when ignoring the much larger statistical variation of the two other approaches. At the same time, the baseline estimates lie within the much larger confidence intervals of the two other approaches. This means that the alternative estimates are unlikely to differ significantly from the baseline estimates.

Regarding the estimates, the RDD reform effect shows a very noisy pattern, entailing the largest confidence intervals. The reform effect is a bit smaller than the baseline during the first year after birth and then moves around the baseline estimates afterwards. The event study estimates are estimated more precisely than the RDD estimates but still much more imprecisely than the baseline. The event study effect suggests positive earnings effect before birth already in the second and third year before birth and a stronger effect than the baseline during the first year before birth. The event study effect basically coincides with the baseline effects during the first year after birth for employment and for the first three years for earnings, but it is considerably smaller later on. In fact, the event study estimates would not entail a significantly positive medium reform effect whereas the baseline effects prove significantly positive for almost all months from the second post-birth year onward. The two alternative control groups, we used in the previous section, do not differ significantly from the baseline.

In sum there are few important methodological conclusions. The RDD approach is too noisy to detect significant reform effects. The event study approach also yield less precise estimates than the baseline estimates. It tends to find stronger positive reform effects before birth than baseline, possibly confusing the reform effect with the general upward trend in earnings and employment, thus overestimating the entitlement effect. And it underestimates the positive reform effects on employment and earnings after the first post-birth year.

Figure 13: Reform effects with event-study, baseline compared to two alternative control groups, RDD without and with control group and event-study



(a) earnings (b) employment

7 Conclusions

This paper estimates the causal effect of first-time motherhood on various post-birth outcomes in Germany and then investigates the selection of women into motherhood and the effect of the 2007 parental leave reform. Mothers are positively selected in terms of their pre-birth labor market outcomes. The probability to become first-time mother increases in log-earnings. The reform had some impact on these selection patterns. Higher earning as well as non-employed women have an increased chance to enter motherhood.

Despite significant negative effects of the reform on labor market outcome during the first year, during which the new benefits are paid, the medium-run effects on earnings and employment are significantly positive. We estimate that the reform increased yearly gross-earnings by almost 650 Euros and the employment rate by 2.5 percentage points in the fifth year after giving birth. The most positive effects are found for medium aged mothers (31 to 35 years of age) while the youngest age group (21 to 25 years of age) shows worse labor market outcomes after the reform. The reform effect on full-time and on second-order fertility is insignificant. It is hence likely that part-time employments drive the positive effects on earnings and (overall) employment.

Further, we find a positive effect during the year before giving birth. This "entitlement effect" is significant for all three labor market outcomes. This finding is plausible because the parental leave benefits in the post-reform period depend on earnings of mothers immediately before birth. To our knowledge, this is the first study to establish this effect.

This evidence on the reform effects on post-birth outcomes fits quite well to the existing literature which finds positive medium-run effects on earnings and employment per se but not for full-time employment (Frodermann et al., 2020; Kluve and Schmitz, 2018). The positive reform effects stem from mothers with better pre-birth labor market outcomes. These "reform winners" are women with higher earnings (Bergemann and Riphahn, 2020; Frodermann et al., 2020; Kluve and Schmitz, 2018). Our results separated by age groups show that the post-birth labor market outcomes for mothers giving birth in their 30s increase after the reform while we do not find such positive reform effects for mothers giving birth in their early 20s. Both pre-birth employment and earnings are considerably higher among the former group compared to the latter group. Consequently, the share of "reform winner" is much larger for the older age group.

The econometric strategy of our project differs substantially from most of the literature on estimating the causal effect of motherhood on post-birth outcomes and on assessing the impact of institutional changes in this context. For short run post-birth outcomes, the methodological differences do not matter much because findings are driven by the strong dip in employment and earnings for mothers immediately after birth. However, the results on the effect of giving birth start to diverge from the second year after giving birth onward compared to a event study approach as in Kleven et al. (2019b). Our findings using a control group approach in a dynamic treatment setting shows better medium-run effects of motherhood compared to the event study approach. Even though there is some ambiguity regarding the choice of the appropriate control group, the differences between alternative definitions of the control group do not fundamentally change the results in the medium run. Regarding the reform effect, the differences between our approach and other approaches are also important. While the effects on earnings and employment are insignificant using an event study or RDD, our control group approach implies positively significant reform effects on post-birth earnings and employment in the medium run. This means that the choice of the econometric strategy may matter for assessing the effects of an institutional reform. For the 2007 parental leave reform in Germany, it does.

References

Abbring, J. H. and Van den Berg, G. J. (2003). The nonparametric identification of treatment effects in duration models. *Econometrica*, 71(5):1491–1517.

Angelov, N., Johansson, P., and Lindahl, E. (2016). Parenthood and the Gender Gap in Pay. *Journal of Labor Economics*, 34(3):545–579.

Antoni, M., Ganzer, A., and vom Berge, P. (2019). Sample of Integrated Labour Market Biographies Regional File (SIAB-R) 1975 - 2017. FDZ-Datenreport 04/2019 (en), Nuremberg.

Bergemann, A. and Riphahn, R. T. (2011). 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 Riphahn, R. T. (2020). Maternal employment effects of paid parental leave. Working Paper Series 2020:6, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Biewen, M., Fitzenberger, B., Osikominu, A., and Paul, M. (2014). The effectiveness of public-sponsored training revisited: The importance of data and methodological choices. *Journal of Labor Economics*, 32(4):837–897.

BMFSFJ (2004). Erziehungsgeld, Elternzeit. Das Bundeserziehungsgeldgesetzt. Technical report, Bundesministerium für Familie, Senioren, Frauen und Jugend.

BMFSFJ (2020a). Elterngeld, ElterngeldPlus und Elternzeit. Das Bundeselterngeld- und Elternzeitgesetz, 23. Auflage. Technical report, Bundesministerium für Familie, Senioren, Frauen und Jugend.

BMFSFJ (2020b). Leitfaden zum Mutterschutz, 16. Auflage. Technical report, Bundesministerium für Familie, Senioren, Frauen und Jugend.

Busso, M., DiNardo, J., and McCrary, J. (2014). New evidence on the finite sample properties of propensity score reweighting and matching estimators. *Review of Economics and Statistics*, 96(5):885–897.

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

Dahl, G. B., Løken, K. V., Mogstad, M., and Salvanes, K. V. (2016). What Is the Case for Paid Maternity Leave? *The Review of Economics and Statistics*, 98(4):655–670.

Ehlert, N. (2008). Elterngeld als Teil nachhaltiger Familienpolitik, 3. Auflage. Technical report, Bundesministerium für Familie, Senioren, Frauen und Jugend.

FDZ-RV (2021). Codeplan FDZ-Biografiedatensatz - VSKT 2018. Technical report, Deutsche Rentenversicherung.

Fitzenberger, B. and Seidlitz, A. (2020). The 2011 break in the part-time indicator and the evolution of wage inequality in Germany. *Journal for Labour Market Research*, 54(1).

Fitzenberger, B., Sommerfeld, K., and Steffes, S. (2013). Causal effects on employment after first birth - A dynamic treatment approach. *Labour Economics*, 25(C):49–62.

Frodermann, C., Wrohlich, K., and Zucco, A. (2020). Parental Leave Reform and Long-Run Earnings of Mothers. Discussion Papers of DIW Berlin 1847, DIW Berlin, German Institute for Economic Research.

Geyer, J., Haan, P., and Wrohlich, K. (2015). The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach. *Labour Economics*, 36:84 – 98.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., and Zweimüller, J. (2019a). Child penalties across countries: Evidence and explanations. In *AEA Papers and Proceedings*, volume 109, pages 122–26.

Kleven, H., Landais, C., and Søgaard, J. E. (2019b). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4).

Kluve, J. and Schmitz, S. (2018). Back to work: Parental Benefits and Mothers' Labor Market Outcomes in the Medium Run. *ILR Review*, 71(1):143–173.

Kluve, J. and Tamm, M. (2013). Parental leave regulations, mothers' labor force attachment and fathers' childcare involvement: evidence from a natural experiment. *Journal of Population Economics*, 26(3):983–1005.

Lalive, R., Schlosser, A., Steinhauer, A., and Zweimüller, J. (2014). Parental Leave and Mothers' Careers: The Relative Importance of job Protection and Cash Benefits. *Review of Economic Studies*, 81(1):219–265.

Müller, D. and Strauch, K. (2017). Identifying mothers in administrative data. Fdz methodenreport, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany].

OECD (2019). Family Database. Technical report, Organisation for Economic Cooperation and Development.

Olivetti, C. and Petrongolo, B. (2017). The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries. *Journal of Economic Perspectives*, 31(1):205–30.

Raute, A. (2019). Can financial incentives reduce the baby gap? Evidence from a reform in maternity leave benefits. *Journal of Public Economics*, 169:203 – 222.

Rossin-Slater, M. (2018). Maternity and Family Leave Policy. In Susan L. Averett, L. M. A. and Hoffman, S. D., editors, *Oxford Handbook of Women and the Economy*. Oxford University Press.

Rossin-Slater, M., Ruhm, C. J., and Waldfogel, J. (2013). The Effects of California's Paid Family Leave Program on Mothers' Leave-Taking and subsequent Labor Market Outcomes. *Journal of Policy Analysis and Management*, 32(2):224–245.

Schönberg, U. and Ludsteck, J. (2014). Expansions in Maternity Leave Coverage and Mothers' Labor Market Outcomes after Childbirth. *Journal of Labor Economics*, 32(3):469–505.

Sianesi, B. (2004). An evaluation of the swedish system of active labor market programs in the 1990s. *Review of Economics and statistics*, 86(1):133–155.

Sianesi, B. (2008). Differential effects of active labour market programs for the unemployed. *Labour Economics*, 15(3):370–399.

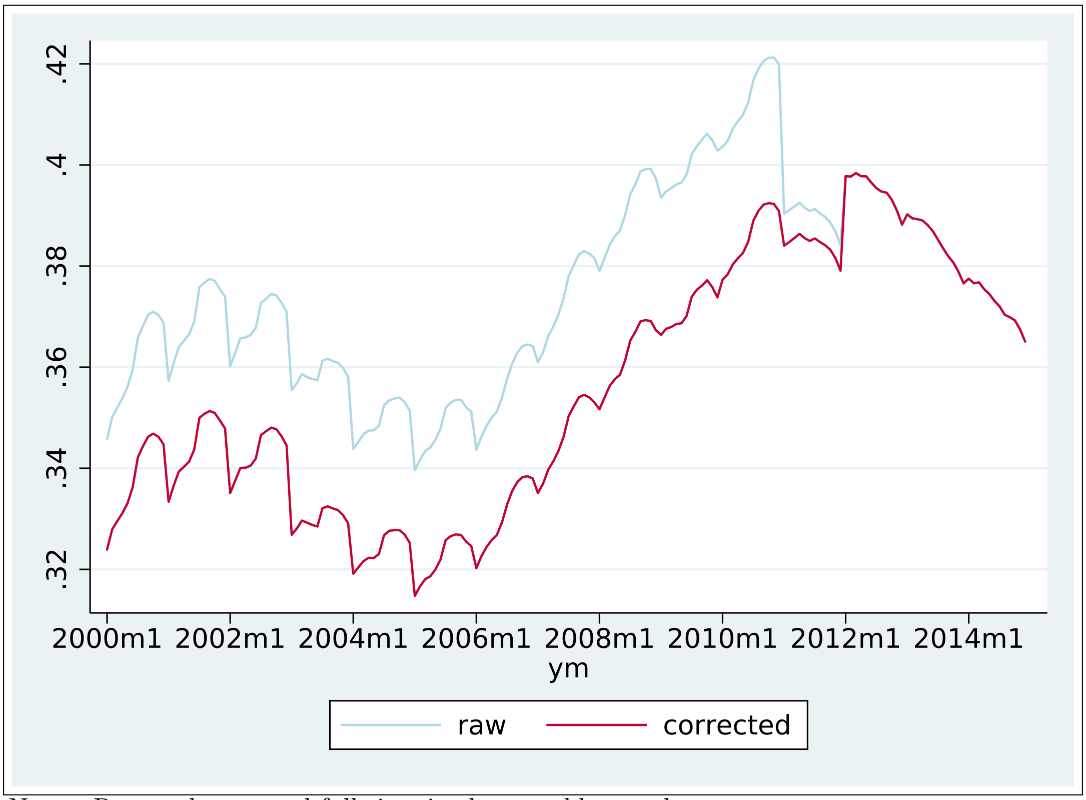
Sigle-Rushton, W. and Waldfogel, J. (2007). Motherhood and women's earnings in Anglo-American, Continental European, and nordic Countries. *Feminist Economics*, 13(2):55–91.

Statistisches Bundesamt (2019). Verbraucherpreisindizes für Deutschland. *Jahresbericht 2018*.

Appendix

Full-time correction

Figure A.1: Share of full-time employed over time



Notes: Raw and corrected full-time in the monthly panel.

For the correction, we used the approach proposed in Fitzenberger and Seidlitz (2020).

Tables

Table A.1: Causal reform effects on daily earnings

	before giving birth	after giving birth				
	1st year	1st year	2nd year	3rd year	4th year	5th year
entire sample	0.63*	-7.16***	2.20***	4.15***	1.28*	1.82**
	(.079)	(.000)	(.001)	(.000)	(.089)	(.016)
women age 21-25	-0.67	-6.09***	-2.04*	-0.98	-2.26*	-2.28
	(.340)	(.000)	(.050)	(.408)	(.084)	(.110)
women age 26-30	0.85*	-6.88***	0.95	1.57	-0.12	0.98
	(.068)	(.000)	(.260)	(.103)	(.909)	(.372)
women age 31-35	0.27	-7.50***	3.62***	6.40***	1.91	3.06**
	(.656)	(.000)	(.003)	(.000)	(.173)	(.032)
women age 36-40	2.35	-8.00***	5.52**	9.50***	6.72**	4.37
	(.126)	(.000)	(.028)	(.000)	(.013)	(.114)

Average causal reform effect for the respective year.
P-values in parentheses refer to a t-test on significance of the average effect for the respective years.
***,** and * indicate significance on 99, 95 and 90 percent level.

Table A.2: Causal reform effects on employment rate

	before giving birth	after giving birth				
	1st year	1st year	2nd year	3rd year	4th year	5th year
entire sample	.011***	-.122***	.033***	.058***	.023***	.025***
	(.005)	(.000)	(.000)	(.000)	(.009)	(.004)
women age 21-25	.013	-.114***	.014	.040	.028	-.013
	(.250)	(.000)	(.556)	(.100)	(.236)	(.602)
women age 26-30	.013**	-.130***	.032**	.054***	.019	.032**
	(.041)	(.000)	(.026)	(.000)	(.177)	(.029)
women age 31-35	.009	-.114***	.045***	.064***	.019	.031**
	(.119)	(.000)	(.002)	(.000)	(.185)	(.037)
women age 36-40	.012	-.130***	.017	.067**	.041	.023
	(.379)	(.000)	(.535)	(.017)	(.123)	(.397)

Average causal reform effect for the respective year.
P-values in parentheses refer to a t-test on significance of the average effect for the respective years.
***,** and * indicate significance on 99, 95 and 90 percent level.

Table A.3: Causal reform effects on full-time rate

	before giving birth	after giving birth				
	1st year	1st year	2nd year	3rd year	4th year	5th year
entire	.012***	-.049***	.013**	.018***	-.011	-.010
sample	(.003)	(.000)	(.019)	(.003)	(.106)	(.111)
women	.003	-.050***	-.009	-.007	-.023	-.023
age 21-25	(.814)	(.000)	(.443)	(.600)	(.132)	(.143)
women	.015**	-.052***	.004	.004	-.015	-.017*
age 26-30	(.022)	(.000)	(.657)	(.720)	(.137)	(.099)
women	.011*	-.041***	.032***	.043***	.006	.014
age 31-35	(.072)	(.000)	(.001)	(.000)	(.587)	(.206)
women	.017	-.061***	.000	.006	-.037*	-.054**
age 36-40	(.205)	(.000)	(.991)	(.764)	(.093)	(.015)

Average causal reform effect for the respective year.
P-values in parentheses refer to a t-test on significance of the average effect for the respective years.
***,** and * indicate significance on 99, 95 and 90 percent level.

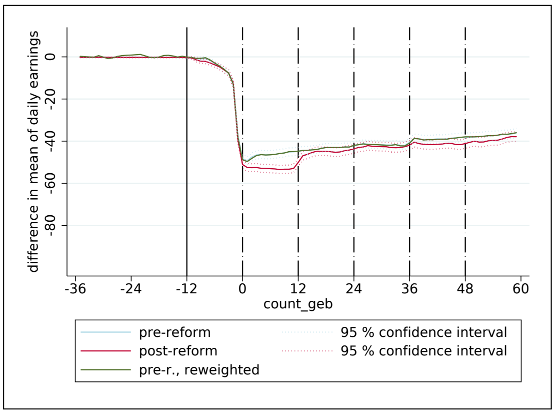
Table A.4: Average reform effects on second order fertility

	1st year	2nd year	3rd year	4th year	5th year
entire	.000	-.002	-.010	-.001	.014
sample	(.328)	(.487)	(.145)	(.927)	(.137)
women	.002	-.004	-.021	-.023	-.001
age 21-25	(.154)	(.710)	(.273)	(.355)	(.986)
women	.000	-.005	-.017	-.010	.008
age 26-30	(.856)	(.279)	(.157)	(.479)	(.611)
women	.001	.000	-.001	.015	.022
age 31-35	(.142)	(.998)	(.951)	(.329)	(.161)
women	-.001	.002	-.010	.001	.020
age 36-40	(.325)	(.824)	(.600)	(.958)	(.443)

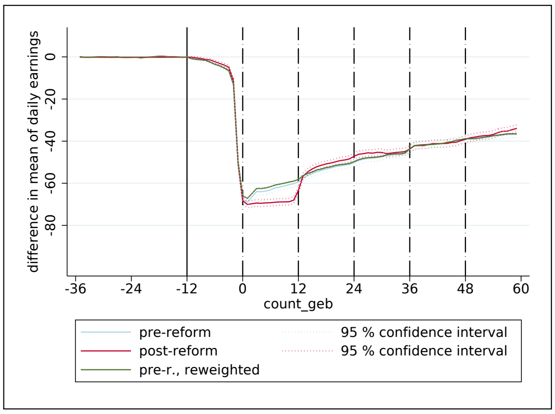
Average effects as difference between the post-reform and pre-reform (reweighted) results for the reform effect (RE) and between post-reform and pre-reform results for the "full reform effect (full RE)". The former excludes and the latter includes potential reform effects on the selection of mothers.
P-values in parentheses refer to a t-test on significance of the average effect for the respective years.
***,** and * indicate significance on 99, 95 and 90 percent level.

Graphs on heterogenous effects by age groups

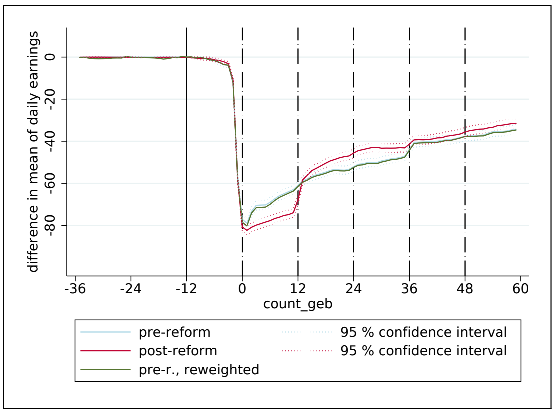
Figure A.2: Effects of giving birth on daily earnings according to age groups



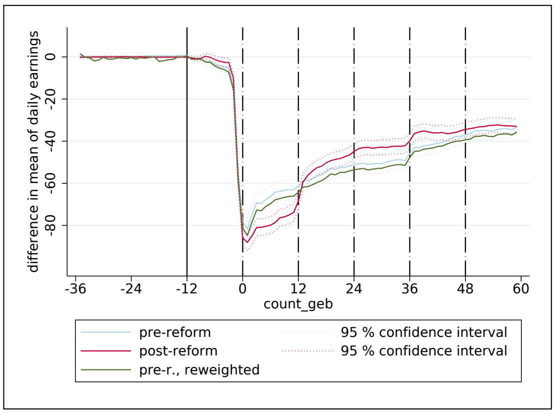
(a) estimates for mothers aged 21 to 25 years



(b) estimates for mothers aged 26 to 30 years

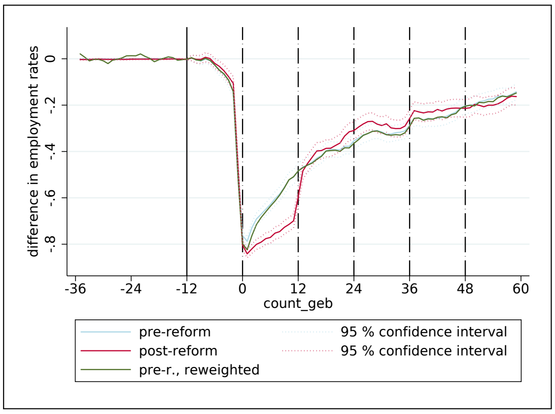


(c) estimates for mothers aged 31 to 35 years

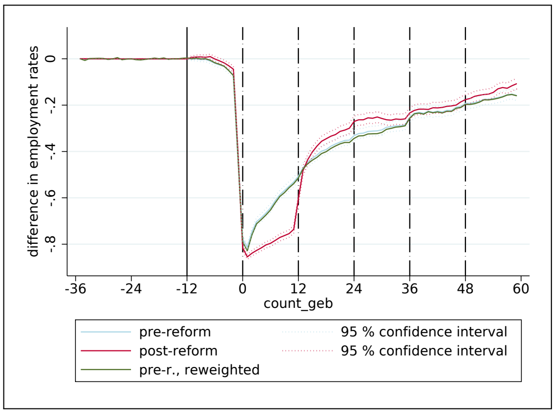


(d) estimates for mothers aged 36 to 40 years

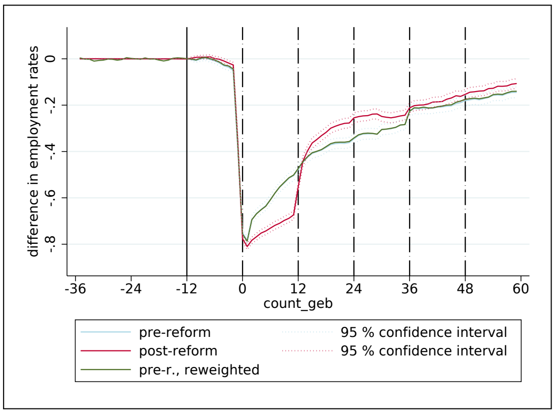
Figure A.3: Effects of giving birth on employment rate by age group



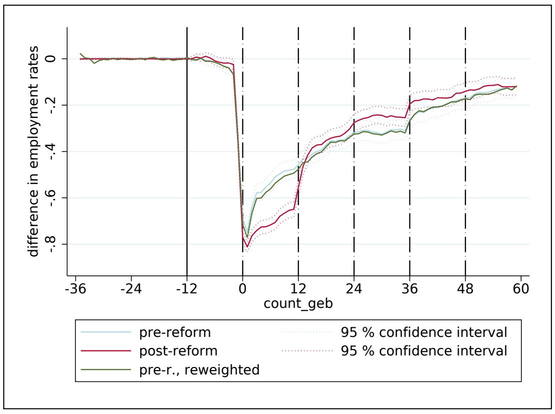
(a) estimates for mothers aged 21 to 25 years



(b) estimates for mothers aged 26 to 30 years

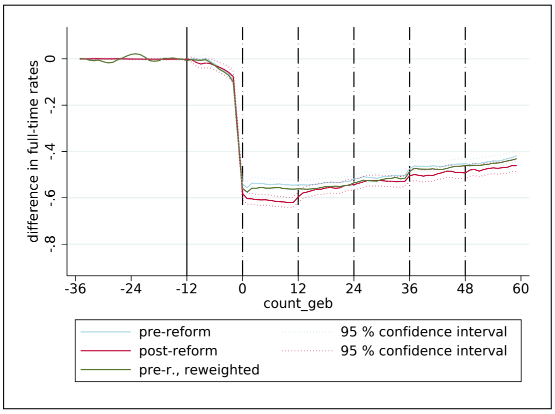


(c) estimates for mothers aged 31 to 35 years

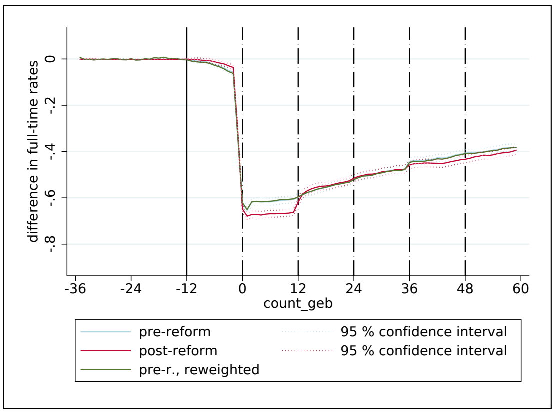


(d) estimates for mothers aged 36 to 40 years

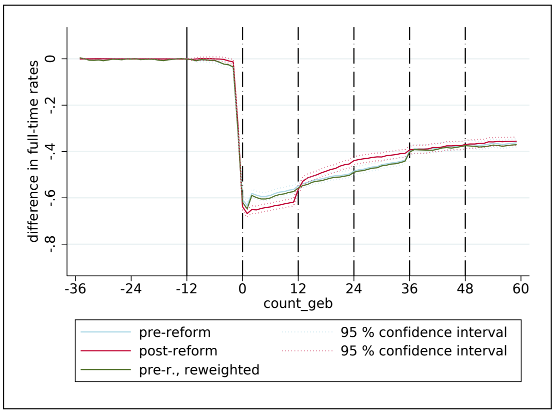
Figure A.4: Effects of giving birth on full-time rate according to age groups



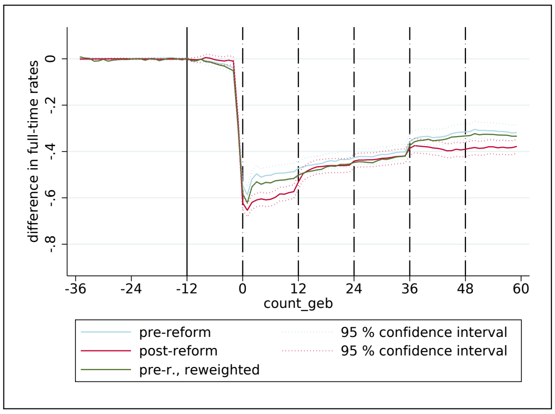
(a) estimates for mothers aged 21 to 25 years



(b) estimates for mothers aged 26 to 30 years



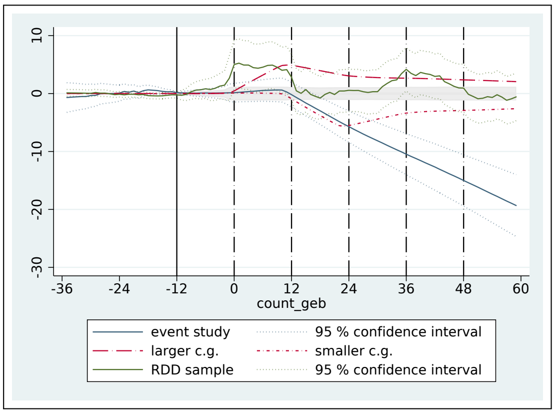
(c) estimates for mothers aged 31 to 35 years



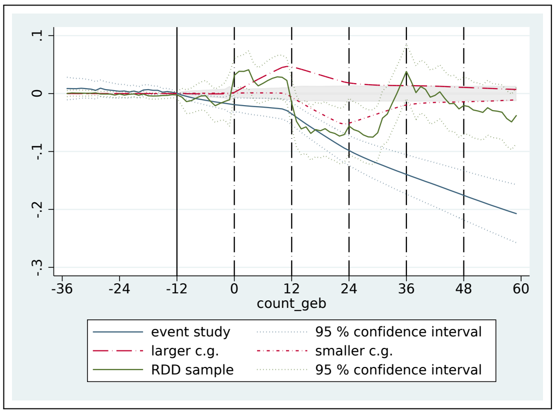
(d) estimates for mothers aged 36 to 40 years

Difference between baseline and event-study without control group

Figure A.5: Differences in the estimates for the effects of giving birth compared to baseline



(a) earnings



(b) employment

Notes: The graphs show the differences between baseline estimates and the estimates for event-study and variations of our baseline (larger control group: include all childless women, smaller control group: restrict control group to women who remain childless for two years after treatment group enter motherhood, RDD sample: uses only treatment groups of 10/2006 to 03/2007 and their control groups).
The gray area shows the 95-confidence interval for the baseline estimates.