DIW BERLIN

2016

Discussion Papers

Deutsches Institut für Wirtschaftsforschung 2022

Parental Leave Benefits and Child Penalties

Sevrin Waights

IMPRESSUM

DIW Berlin, 2022

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
https://www.diw.de

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
https://www.diw.de/discussionpapers

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
https://ideas.repec.org/s/diw/diwwpp.html
https://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html

Parental Leave Benefits and Child Penalties

Sevrin Waights*
HU Berlin and DIW Berlin

This version: August 2022

Abstract

I use the universe of tax returns in Germany and a regression kink design to estimate the impact of the benefit amount available to high-earning women after their first childbirth on subsequent within-couple earnings inequality. Lower benefit amounts result in a reduced earnings gap that persists beyond the benefit period for at least nine years after the birth. The longer-term impacts are driven by couples where the mother earned more than the father pre-birth. Simulations suggest it would take a 50% reduction in the benefit amount to completely eliminate long-run child penalties for sample couples.

Keywords: Child penalties, gender inequality in earnings, social norms, parental leave policy, regression kink design
JEL: D63, H31, J13, J16, K31, M52, Z13

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 752053. For helpful comments and suggestions, I thank Abigail Adams-Prassl, Sebastian Calonico, Henning Finseraas, Jonas Jessen, Ariel Kalil, Claudia Olivetti, Steven Rivkin, Ulf Zölitz, and participants of the Counterfactual Methods for Policy Impact Evaluation conference (2018), the German Economic Association annual Congress (2019), and the Potsdam Research Seminar in Economics (2020), as well as many colleagues at internal seminars and discussions. I also thank Stefanie Setzer, Maurice Brandt, Hannes Fauser, and others at the Federal Statistical Office for their assistance in accessing the data. Many thanks also to Zhuan Pei for providing code for the FG bandwidths. Email: swaights@diw.de.

I. Introduction

The persistence of gender inequality in earnings is largely the result of a child penalty that is likely driven more by social norms than by biological differences between men and women (Lundborg et al., 2017; Kleven et al., 2019a,b; Andresen and Nix, 2021; Kleven et al., 2021). Policymakers see parental leave as a potential tool to foster maternal labour market attachment and increase paternal involvement with childcare.[1] Indeed, evidence shows parental leave policy can have significant impacts, especially aspects of policy design to do with benefit payments.[2] For example, reserving some months of paid leave for fathers may be beneficial for mothers' outcomes (e.g. Patnaik, 2019; Druedahl et al., 2019), whereas extending paid leave beyond the period of job-protection may be detrimental (e.g. Schönberg and Ludsteck, 2014). Given that leave payments appear important, there is a surprising lack of evidence on the impacts of the benefit amount itself.

In this paper I ask: can the benefit amount available to mothers affect within-couple gender inequality in earnings after childbirth? The benefit amount is likely an important factor in household decisions regarding the relative take-up and duration of leave of either parent and could, therefore, affect subsequent earnings trajectories. The benefit amount might also affect labour market decisions independently of the length of leave through an income effect (Wingender and LaLumia, 2017). However, the impacts of the benefit amount are typically difficult to identify since payments tend to be a function of previous earnings and, therefore, highly endogenous to post-birth earnings. As such, only a few existing studies estimate causal effects (Asai, 2015; Bana et al., 2020; Ginja et al., 2020).

I estimate the causal impact of the mothers' benefit amount on the earnings of each parent and the within-couple gender gap up to nine years after the birth of their first child. I do so using data on the universe of tax returns filed in Germany over the 2002–2016 period. My main specifications focus on joint tax returns filed by married couples

[1] Tackling the gender pay gap is an important item on the policy agenda in many European countries (Plantenga et al., 2006).

[2] Job protection is also an important aspect for women's labour market outcomes (e.g. Stearns, 2018; Baker and Milligan, 2008; Waldfogel, 1998) but most countries already have job-protection in place and adjustment to the length of the job-protection period appears to have limited medium-term consequences affecting only the timing of the return to work (Dahl et al., 2016; Lalive et al., 2014; Lalive and Zweimüller, 2009). An important exception is for very long job-protected leaves (Mullerova, 2017). For broader summaries of the effects of the leave policy see Rossin-Slater (2017) and Olivetti and Petrongolo (2017).

since these allow me to observe both parents together. Births to married couples comprise around two thirds of births in Germany. However, in alternative specifications, I also examine the impacts for unmarried women. In order to estimate a causal impact, I make use of exogenous variation due to a cap in benefits. The German parental leave system offers new parents up to 14 months of payments (or twelve if only one parent takes paid leave) compensated at a rate of approximately 65% of the leave-taking parent's pre-birth net earnings up to a maximum monthly benefit amount of €1,800. This benefits cap allows me to estimate the impacts of the mothers' benefit amount using a regression kink (RK) design (see Card et al., 2015).

My main finding is that a lower mothers' benefit amount reduces the gender gap in the leave period and that this effect persists into the post-leave period up to nine years after the birth. The reduced earnings gap in the leave period is driven primarily through increased mothers' earnings, whereas the post-leave impacts comprise both increased mothers' earnings and decreased fathers' earnings. The estimates are robust to bias-correction and different bandwidth specifications and they stand up to several placebo tests. I also find that a lower benefit amount decreases the likelihood of having further children in the longer run and has no effect on split from partner.

These results represent the first causal evidence on the impacts of the benefit amount on couple earnings (after the first birth) and on other family outcomes. Bana et al. (2020) for California and Asai (2015) for Japan assess the impact of the benefit amount on mothers' labour market outcomes but only in the short run and without observing partners. Neither paper finds evidence that the benefit amount affects the labour market outcomes of mothers, perhaps due to the specific institutional and cultural backgrounds in each case.[3] The only study to look at longer-term impacts and to examine partner outcomes is Ginja et al. (2020), for Sweden, but does so for second births, specifically. They find that a lower mothers' benefit increases her earnings only in the short run and without impacts on spousal earnings. A possible explanation for the lack of long-run or spousal impacts is that a gendered intra-household specialisation may already be somewhat entrenched by the second birth.[4] My results also provide the first causal impacts of benefits on broader family outcomes. Studies

[3]The Californian scheme is much shorter, at only six weeks, than the 14 months available in Germany, leaving less scope for large labour market impacts. The only detected effects are a small increase in job continuity and an increase in future claims. In Japan the low compensation rate of 40% (changed from 25%) may have been too small to overcome the strong cultural norm for mothers to stay at home.

[4]The evidence on child penalties suggests that it is the moment of entering parenthood, i.e. the first birth, that is most crucial for intra-household specialisation.

show a decrease in fertility (Farré and González, 2019) and mixed impacts for marital stability (Olafsson and Steingrimsdottir, 2020; Avdic and Karimi, 2018) as a result of paternity quotas, but there is a lack of evidence on the benefit amount.[5]

An examination of heterogeneity reveals that the longer-term (post-leave) impacts on the earnings gap are driven by couples where the mother earned more than the father in the year before the birth (representing approximately half the couples in the main sample). In couples where the mother did not previously earn more than her spouse, the positive impact on mothers' earnings lasts only in the short run, dissipating once the benefit payments elapse, and there are no effects on fathers' earnings. A possible explanation for these heterogeneous effects is if financial incentives, as a tool to alter within-household specialisation, are less effective in the presence of strong gender identities (Ichino et al., 2019; Bertrand et al., 2015). In support of this idea, I present data from a representative household survey showing that restrictive gender role attitudes are more common among parents where the mother did not earn more than her partner pre-birth.

I investigate leave-taking behaviour as a potential mechanism finding that a lower benefit amount reduces mothers' take up of paid leave. Furthermore, in couples where the mother earned more than her partner pre-birth, the reduction in leave take-up by the mother is larger and there is an increase in the length of the father's leave. In the remaining couples, the reduction in leave take-up by mothers is smaller and fathers actually take shorter leaves, on average. These findings present a plausible mechanism for the main results and for the effect heterogeneity. Longer-term reductions in the gender gap appear to occur only where there are positive impacts on the length of paternity leave, underlining the potentially important role for reserved partner months in addressing gender inequality.[6]

I provide insight into the generalisability of my main findings beyond high-earning married women in two further analyses. Firstly, I make use of an alternative method, difference-in-differences (DD), that is not reliant on using a small bandwidth around the threshold. I compare outcomes for couples with births before and after the 2007 reform with higher vs. lower predicted benefit amounts. The DD results for a sam-

[5]One exception is (Raute, 2019) who finds a positive relationship between the benefit amount and further children in Germany, consistent with my results. However, given the methodology used it is unclear if the paper measures an impact of the benefit amount or the benefit duration.

[6]Some studies show that reserved months are beneficial for women's outcomes (Patnaik, 2019; Druedahl et al., 2019; Tamm, 2019) but others show no effects (Ekberg et al., 2013; Cools et al., 2015).

ple similar to that used in the RK design are broadly supportive of the main results. However, when switching to using a more general sample, including lower-earning married women, a lower benefit amount has only small negative impacts on the leave-period earning gap and it no longer reduces the earnings gap in the post-leave period. Secondly, I make use of tax returns for unmarried women to estimate the impacts on mothers' earnings by marital status. Here, too, I find a lack of generalisability: for unmarried women there are no impacts on earnings in the leave period and there is even a decrease in the post-leave period. To help understand this finding, I present evidence from a household survey that shows that unmarried mothers are far less likely to have a partner in the household who contributes to childcare than married women. Thus, overall, my main findings appear to apply only for higher-earning women who are also married in the pre-birth year.

My main approach is a reduced form RK design. As such, I provide estimates of the effect of the predicted benefit amount rather than the actual benefit amount. I do not estimate a fuzzy RK principally because I have data on annual but not monthly benefits received. Nevertheless, there are some clear advantages of the using predicted amounts. Firstly, the predicted benefit amount is perhaps of more policy relevance since it is the value that is under the direct control of policymakers. Secondly, and more importantly, using predicted amounts does not require a selected sample of policy takers as can be the case for analyses using datasets of claimants that record amounts received. Since I observe predicted amounts for all eligible births, regardless of take-up, I mitigate any concern regarding whether the benefit amount affects take up potentially violating the assumptions of the RK design. Instead, I am able to specifically examine policy take-up as an outcome which is of heightened policy interest and a key piece of understanding for the mechanisms of impacts.

Finally, I make use of an event study design to examine how the earnings impacts might affect the actual child penalties for married couples. To do so, I predict earnings in post-birth years following a 20% or 50% decrease in the benefit amount based on the dynamic RK estimates. I then estimate long-run child penalties based on the earnings predicted after the counterfactual benefit reductions as well as actual observed earnings. This analysis suggests that the child penalty could be completely eliminated in the long run for married couples in the main estimation sample following a 50% reduction in benefits. These findings show there may be some potential for parental leave policy to address child penalties. At the same time, the fact that my results do not seem to generalise to broader samples is consistent with the current understanding that these policies are not generally associated with improved gender equality in earnings (Kleven et al., 2020).

II. German parental leave policy

Parental leave policy in Germany was subject to a major reform in 2007. Before the reform, new parents were entitled to 36 months of job-protected leave. A flat-rate benefit of €300 per month was paid for the first 24 months, with an option to instead get €450 per month for twelve months. These benefits, however, were means-tested and available only to parents with a net household income below €30,000 for the first six months, and below €16,470 from the seventh month of benefits. These thresholds compare to a median net household income in Germany in 2006 of €15,646 (EU-SILC, 2022). The system was intended to be gender-neutral with both mothers and fathers entitled to the same length of leave and benefits. However, in reality the long leave period and low, flat-rate means-tested benefit were considered only to have reinforced the German male breadwinner model, with few fathers taking parental leave (Waldfogel, 1998). The 2007 reform overhauled this system, shortening the length of job-protected leave to 24 months and introducing twelve months of universal benefits that were paid as a proportion of previous earnings, as well as an extra two months of benefits for the partner.[7]

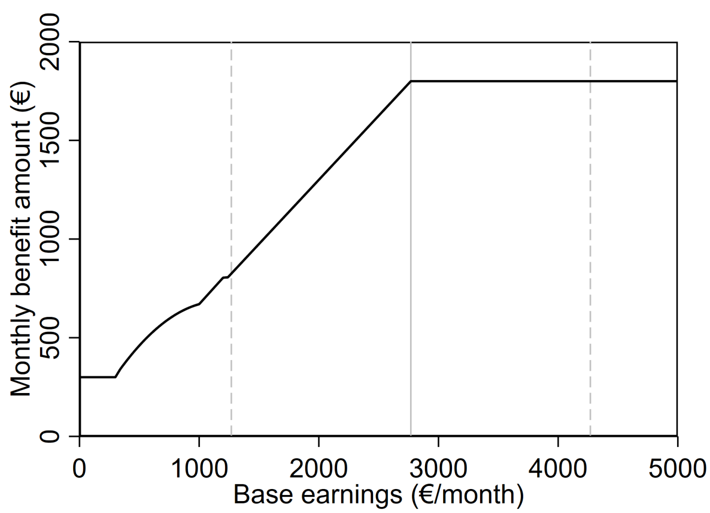
A number of studies examine the impacts of the 2007 German parental leave reform. Given eligibility for the new benefits system were based on a sharp cut-off date on January 1, 2007, and that the announcement was too short notice to affect conception, the reform lends itself neatly to regression discontinuity designs. However, given the reform represented a bundle of changes, it is difficult to assign measured effects to any particular policy design aspect. Indeed, both Frodermann et al. (2020) and Kluve and Schmitz (2018) find that the reform increases earnings for mothers with births in 2007 or later but provide differing reasons for the effect. Kluve and Schmitz (2018) suggest that the higher earnings might reflect the value that employers place

[7] Additionally, there is a 'mother protection' (*Mutterschutz*) period that mandates six weeks before the birth and eight weeks after to be taken off at full pay. This scheme was is place in both the new and old parental leave systems. An implication of this law is that the first two months of paid parental leave are compulsory for mothers, although these months are paid at full pay rather than using the standard benefit schedule.

on having greater certainty regarding the timing of return to work.[8] Frodermann et al. (2020), however, suggest that the two reserved months for partners may be the reason for improved labour market outcomes for mothers.

My paper instead focuses specifically on estimating the impact of the benefit amount making use of the kinked benefit schedule. Leave payments are computed as a function of previous net earnings. Roughly speaking, payments are 65% of net earnings (up to the cap of €1,800) for earnings above €1,200, 67% for earnings between €1,000 and €1,200, and increasing from 100% to 67% for earnings between €300 and €1,000. The minimum payment is €300. Figure 1 illustrates the kinked parental leave payments schedule against net earnings in the year before birth. The main kink in the schedule falls at a net earnings of €2,769, which is the earnings level that relates to the €1,800 benefits cap. There are some further non-linearities at the bottom of the schedule, in particular at the €300 minimum payment. To ensure these do not interfere with estimation around the main kink an observation window of €2,769 ± €1,500 is used.

FIGURE 1: PARENTAL LEAVE BENEFITS SCHEDULE



Notes: Benefits schedule shows monthly parental leave payments as a function of monthly net earnings in the year before the child is born. The solid vertical line indicates the locations of the (upper) benefits cap at a pre-birth net earnings of €2,769 and dashed lines indicate the sample window of €2,769 ± 1,500.

[8]The new twelve month benefit payments meant that most mothers returned after exactly one year.

III. Data

I use administrative data on the universe of tax returns filed in Germany between 2002 and 2016, provided by the German Federal Statistical Office (Destatis, 2018). Around 27 million of Germany's 39 million tax payers file an annual income tax return and, therefore, appear in the data. Additionally, from 2013 onwards the data also include cases for non-filers that are automatically generated based on social security records. The tax returns include information on earnings, transfer benefits, age, gender, marital status, federal state of residence, and birth dates of all children. I focus on married couples since they file joint tax returns, allowing me to see mothers and fathers together.[9] The data comprises only opposite-sex partnerships since same-sex marriages were allowed only from 2017 in Germany. The data has a panel structure whereby the same married couple is linked across all the years in which they file a tax return. Splits are identifiable since in these cases the husband files a single tax return under the same case number as previously when married. Using birth dates of children reported on tax return I identify around 150,000 first births to married couples per year. While these correspond to around 70% of the official count for this type of birth in Germany,[10] I am able to use the automatically generated returns from the later waves to confirm that coverage is nearly complete (98%-99%) for the group of married couples where mothers earnings are close to the benefits cap who are much more likely to file.[11] As such, my estimation sample is effectively the full population of married couples around the threshold. I also make use of tax returns from unmarried and lower-earnings mothers to assess generalisability of the main analysis.

My main outcome variable is the gender gap in gross annual earnings, which is computed as fathers' minus mothers' gross annual earnings in each tax year relative to the birth year of the first child.[12] As additional outcomes, I use mothers' earnings and

[9] Married couples in Germany have the option to file a joint tax return and it is unusual for them not to do so. Opting to file individually means they would lose out on certain tax benefits available to wedded pairs. As a result, married couples typically opt to file individually only in rare cases such as when the couple is about to divorce.

[10] In Table A1 in the appendix, I provide a comparison between the official statistics on births and the number of births identified by the dates of birth of children reported on the tax returns. The overall coverage rate of births in the tax data of about 70% roughly coincides with the share of tax-payers that file a tax return.

[11] Moreover, not filing a tax return in a later wave is also very small and not impacted by the treatment. See the last paragraph of the robustness checks in section V for further details.

[12] I sum employment earnings and self-employment earnings, which enter the data individually.

fathers' earnings, separately. Following the child penalties literature, I use earnings in levels in order to keep instances of zero earnings (Kleven et al., 2019b). As further outcomes, I examine whether the pair splits (if the husband file alone subsequently, as mentioned above) and whether they have further children. I examine (annual) transfer benefits received to demonstrate the existence of a first stage relationship. In order to examine leave-taking as a mechanism, I compute an indicator of leave take-up reflecting whether the annual benefits received by a parent divided by twelve is €300 or greater, which is the minimum benefit payment for one month. I also compute a proxy for the number of months of leave as the total benefit amount received in the first two years after birth divided by the predicted monthly amount. I compute the benefit assignment variable as net earnings in the year before the birth year of the first child (the base year).[13]

The data include 1.1 million couples who have their first child in 2007 or after and file a joint tax return in the year before the birth. I focus on the 56% of these couples where the mothers' base earnings fall within €1,500 of the benefits cap, i.e. €2,769 $\pm$ €1,500. I make the further restrictions that the mother has no self-employment earnings in the base year (97.8%), is not a civil servant in the base year (93.4%), and that the couple has a valid tax code combination (93.1%), since these are required for correct calculation of net earnings for the assignment variable. In the end, I am left with around half a million couples, corresponding to 6.4 million couple-year observations when using the longitudinal format of the data. As discussed, these observations reflect the near-universe of such couples. Table 1 provides means of the outcome variables and covariates for all 1.1 million eligible joint-filers as well as the estimation sample with restrictions as described and a bandwidths of €1,500, €1,000 and €500.[14] The samples closer to the cap are noticeably different to the full population: Mothers have higher earnings, are older, and are more likely to earn more than their partner the closer one gets to the threshold. As such, my analysis pays careful attention to issues of generalisability.

[13]Net earnings are created by deducting taxation and social security from mothers' gross earnings using publicly available formulas.

[14]Table A2 in the appendix provides more summary statistics over the panel dimension of the data.

Table 1: Descriptive statistics in base year

	Joint-filers with first birth≥ 2007	Samples by bandwidth		
		€1500	€1000	€500
Earnings gap (€000 pa)	11.8	2.47	-1.59	-5.37
	(39.9)	(34.8)	(33)	(34.5)
Mothers' earnings (€000 pa)	26.7	38.3	45.2	52.8
	(20)	(12.8)	(10.9)	(7.82)
Fathers' earnings (€000 pa)	38.5	40.8	43.6	47.5
	(38.4)	(36)	(34.7)	(36.4)
Mother is the bigger earner	.308	.452	.545	.606
	(.462)	(.498)	(.498)	(.489)
Mother is capped	.0568	.0957	.154	.299
	(.231)	(.294)	(.361)	(.458)
Father is capped	.18	.185	.226	.294
	(.384)	(.388)	(.418)	(.456)
Age of mother (years)	28.7	29.6	30.4	31.2
	(4.52)	(3.9)	(3.71)	(3.52)
Age of father (years)	31.7	32.1	32.8	33.5
	(5.36)	(4.81)	(4.71)	(4.6)
Age gap (years)	2.93	2.53	2.4	2.29
	(4.1)	(3.7)	(3.68)	(3.65)
Lives in former East	.0995	.086	.0777	.0751
	(.299)	(.28)	(.268)	(.264)
Religious household	.67	.732	.73	.709
	(.47)	(.443)	(.444)	(.454)
Observations	1,097,386	502,175	273,548	98,000

Notes: Table reports the mean and standard deviation (in parentheses) for key variables from the joint tax returns filed by married couples with first birth in 2007 or later, when the benefits system was in place. Variables correspond to the year before the birth year. The first column includes all eligible married couples and the remaining columns show estimation samples restricted to couples where the mothers' earnings fall in the specified bandwidth around the benefits cap plus the additional restrictions described in the text.

IV. Empirical approach

A naive approach might be to regress couple outcomes on the monthly benefit amount available to the mother, plus controls. However, if one found that, say, mothers with a higher benefit amount have higher post-birth earnings than mothers with a lower benefit amount, such an estimate could not be interpreted as a causal effect. The most obvious reason why not is that that mothers with a higher benefit amount necessarily had higher pre-birth earnings as illustrated by the sloped section of the schedule in in Figure 1. These mothers, by way of having higher pre-birth earnings, are likely on average to have a higher education level, more experience, higher skills, to hold more egalitarian gender attitudes as a couple, and to have a more highly paid job to return to. Thus, higher post-birth earnings for these mothers could reflect any of those factors rather than the higher benefit amount. Similar logic holds for paternal earnings or for the earnings gap as outcomes.

In order to estimate the causal effect of the benefit amount, I estimate a regression kink (RK) design (see e.g. Card et al., 2015). The approach requires the benefit amount b to be a kinked function of base earnings v. In my case, the benefit schedule as illustrated in Figure 1 may be written as $b(v) = \min[0.65v, 1800]$, where €1,800 is the maximum benefit amount. This function implies a sudden change in slope from 0.65 to 0 at the threshold earnings v_0 of €2,679. The RK approach assumes that the outcome of interest y depends smoothly on the benefit amount and relevant covariates that may be observable or unobservable, such as education, experience, and gender norms: $y = y(b(v), u(v))$. Identification requires the assumption that the covariates u are a smooth function of base earnings. The causal effect of the benefit amount is then given by the change in slope in the outcome divided by the change in slope of the benefits function at the threshold: i.e. $\frac{\lim_{v_0 \to 0^+} \frac{\partial y | v=v_0}{\partial v} - \lim_{v_0 \to 0^-} \frac{\partial y | v=v_0}{\partial v}}{-0.65}$.

I order to implement the approach econometrically, I estimate the following RK regression model using OLS:

$$y_{it\tau} = \alpha + \beta(v_i - v_0)\mathbf{1}[v_i > v_0] + \gamma(v_i - v_0) + \delta X_{it\tau} + \epsilon_{it\tau}, \text{ if } |v_i - v_0| \leq h \tag{1}$$

where $y_{it\tau}$ is the outcome for couple i, observed in year t, and event time τ (which is years since the birth year s, i.e. $\tau = t - s$), v_i is the mother's base year earnings, v_0 is the threshold earnings of €2,679, $\mathbf{1}[v_i > v_0]$ is an indicator variable for whether the mother is subject to the benefits cap, and X_i is a vector of covariates for the mother's age, the mother's age squared, the age difference between mother and father, the base year earnings gap, whether the spouse is subject to the benefits cap, whether the couple lives in the former East, whether the couple is a member of a church, as well as indicator variables for the tax year t and event time τ.[15] The estimate $\hat{\beta}$ gives the change in slope of the outcome at the benefit cap. Subject to the identifying assumptions of the RK design, $\hat{\beta}/-0.65$ provides a causal estimate of the impacts of the predicted benefit amount on the outcome.

Using too large a bandwidth h runs the risk that the estimate captures a smooth change in slope, i.e. a curve rather than a kink, which could be the result of nuisance unobservables. However, using too small a bandwidth may reduce the number of observations available to estimate the slopes either side of the threshold, resulting in imprecision. To demonstrate robustness to bandwidth choice, I estimate RK models with bandwidths from €300 to €1,500 in €100 increments. I also carry out checks for sensitivity to RK options by using a triangular instead of uniform kernel and the bias-corrected, robust estimate of Calonico et al. (2014) instead of conventional estimates. Furthermore, I present results using the optimal bandwidth selection procedures of Calonico et al. (2014), Imbens and Kalyanaraman (2012), and Fan and Gijbels (1995), denoted hereafter as CCT, IK, and FG, respectively. As discussed by Card et al. (2015), automated selection procedures may choose undesirable bandwidths. As such, I choose a fixed bandwidth of €600 for my main specifications. This bandwidth is chosen in a way to be small enough that there is only smooth variation in (observed) characteristics but wide enough to obtain precise estimation. A further advantage of the researcher-chosen bandwidth is that estimates are based on the same sample aiding comparison across specifications.

In addition to the main specification, I estimate dynamic effects using the following OLS regression:

$$\begin{aligned} y_{it\tau} = &\sum_{j\neq-1} \alpha_j \mathbf{1}[\tau=j] + \sum_{j\neq-1} \beta_j \mathbf{1}[\tau=j](v_i - v_0)\mathbf{1}[v_i > v_0] + \sum_{j\neq-1} \gamma_j \mathbf{1}[\tau=j](v_i - v_0) \\ &+\alpha_{-1} + \beta_{-1}(v_i - v_0)\mathbf{1}[v_i > v_0] + \gamma_{-1}(v_i - v_0) + \zeta X_{it\tau} + \varepsilon_{it\tau}, \text{ if } |v_i - v_0| \leq h \end{aligned} \tag{2}$$

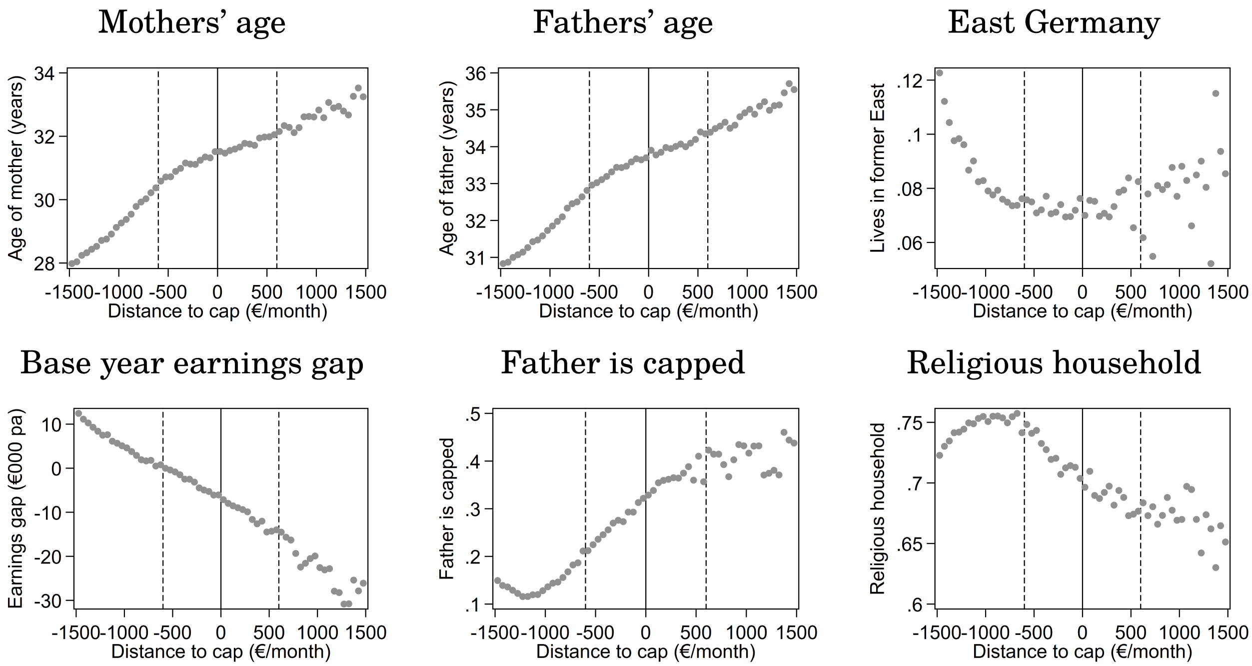
[15] Members of churches pay an extra tax making them identifiable on the tax returns.

where $\sum_{j \neq -1} \mathbf{1}[\tau = j]$ are a set of event time indicator variables starting with seven years before the birth ($j = -7$) and ending seven years after ($j = 7$), but omitting the base year ($j = -1$). In practice, in order to improve precision of the dynamic estimates, I pool two event time years in each case, e.g. ($j = -7, -6$). The estimates $\hat{\beta}_j$, therefore, represent the kinks at event time $\tau = j$ relative to the kink estimated in the base year $\tau = -1$. Such an event study type approach removes any potential pre-existing non-linearities in the outcome variable, strengthening the sub-group estimates.

Evidence on identifying assumptions. The RK approach requires that (1) the nuisance relationship between base earnings and the outcome is smooth across the threshold, and that (2) the frequency distribution is smooth across the threshold. A potential threat to these assumptions is if individuals manipulate their base earnings to cross the threshold. Such manipulation may be less likely in kink settings where small changes cannot lead to large changes in the benefit amount (as in discontinuity settings). Another common problem in RK studies is if the benefit amount affects selection into the programme and if the data include only claimants. As discussed, however, my data include both claimants and non-claimants avoiding this issue. The standard tests are to plot the frequency distribution and covariates around the benefits cap. Figure 2 shows that covariates vary smoothly over the threshold when looking within the chosen €600 bandwidth, and Figure 3 shows a smooth frequency distribution across the threshold.

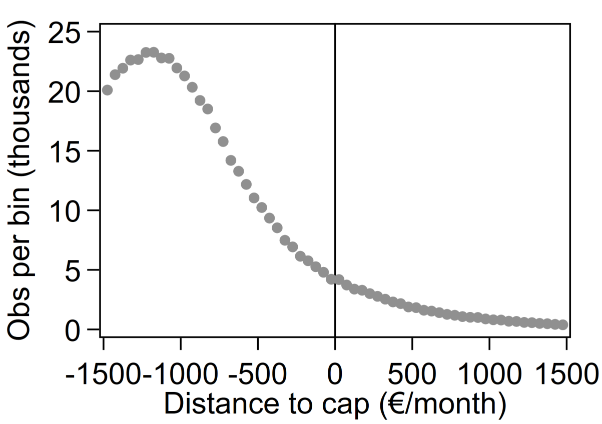
In addition to these two main assumptions, the fuzzy version of the design also requires a strong first stage relationship between the predicted benefit amount and the actual benefit amount. My reduced form approach also implies the existence of a first stage relationship. Since I do not see the monthly benefit amount, I instead plot the annual transfer income received in €50 bins around the threshold in Figure 4. The plot shows a clear kink in transfer income reported on tax returns during the leave period. This is either evidence of a drop in the monthly benefit amount or a reduction in the duration of leave taken, with the latter unlikely to occur in the absence of a first stage relationship.

Figure 2: Variation in covariates across the threshold



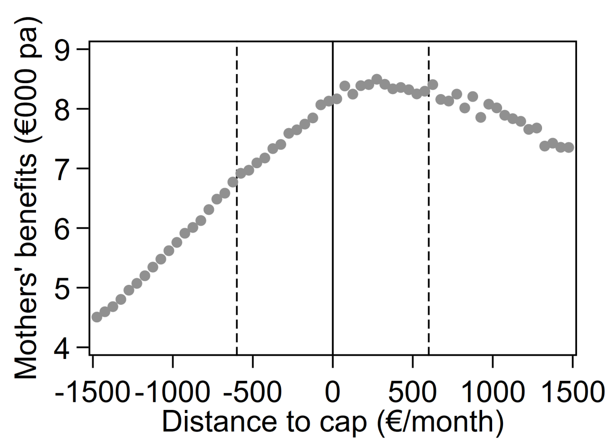
Notes: Plots show covariates in €50 bins of mothers' normalised base year earnings. Covariates refer to the base year ($\tau = -1$).

Figure 3: Frequency distribution



Notes: Plot shows number of observations (couples observed in the base year) in €50 bins of mothers' normalised base year earnings.

Figure 4: Evidence of first stage



Notes: Plot shows mothers' annual transfer income pooled in $\tau = 0, 1$ in €50 bins of mothers' normalised base year earnings.

V. Impacts of the benefit cap

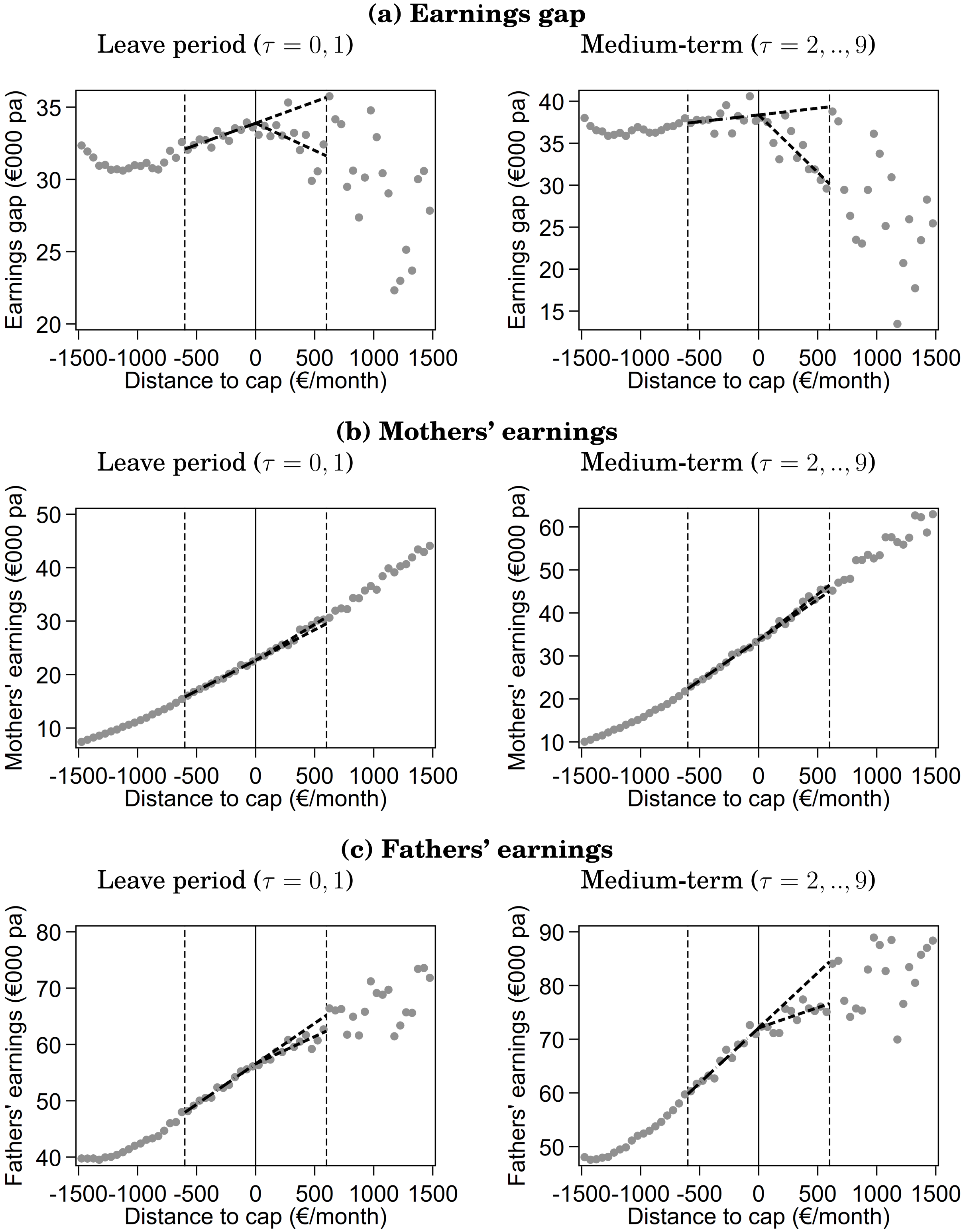
RK plots. Figure 5 presents RK plots for the gender gap and for mothers' and fathers' earnings, separately. The plots show the average for the outcome variable in €50 bins of mothers' normalised base year net earnings, i.e. earnings minus the threshold (€2,679). The leave period ($\tau = 0, 1$) and the post-leave period ($\tau = 2, ..., 9$) are plotted separately to assess immediate and persistent effects. The figures show strong visual evidence of a negative trend shift in the gender gap after the threshold. As such, a reduction in the benefit amount appears to reduce the gender gap. This effect is visible in both the leave period as well as the post-leave period. The graphs also suggest that a reduction in benefits increases mothers' earnings and decrease fathers' earnings. While the increase in mothers' earnings emerges immediately in the leave period, the decrease in fathers' earnings appears small at first and grows over time.

RK estimates. Table 2 reports results the results of the estimation of equation (1). Panel A includes the full set of covariates for the couple as listed in the table notes. Panel B shows the results without covariates, controlling only for the indicators for tax return year and child age. The key finding is that the benefits cap reduces the gender gap in the leave period and that this effect not only persists into the post-leave period but actually grows in size. In the leave period, the reduced gender gap comes mostly from increased mothers' earnings, whereas in the post-leave period, it comes from both increased mothers' earnings and decreased fathers' earnings.

In order to provide an idea of the magnitude of the impacts, I report the effect of a ten percent decrease in the mothers' benefit amount.[16] In the post leave-period, the estimates correspond to a €2,200 decrease in the annual earnings gap, a €600 increase in mothers' earnings and a €1,600 decrease in fathers' earnings. The impact on fathers' earnings is larger in absolute terms (although in the opposite direction) than the impact on mothers earnings but the effects as a proportion of the mean of the outcome variable are comparable. This fact is illustrated by the elasticities, which in the post-leave period are 0.30 for mother's earnings, -0.31 for fathers' earnings, and 0.68 for the earnings gap. Effects of equal proportions could, for example, represent a shift in paid work hours from the father to the mother.

[16] The calculation of the effect of the ten percent benefits decrease and the elasticities is described in the table notes.

Figure 5: RK plots



Notes: Figures shows mean parental earnings in €50 bins of the running variable. Black lines indicate the trend shift estimated within a bandwidth of €600 around the cap.

Table 2: RK estimates of the impact of mothers' predicted benefit amount

Outcome (€ 000s):	Earnings gap		Mothers' earnings		Fathers' earnings	
Period:	Leave	Post-leave	Leave	Post-leave	Leave	Post-leave
	(1)	(2)	(3)	(4)	(5)	(6)
			A. With covariates			
Kink estimate ($\hat{\beta}$)	-0.0034**	-0.0090***	0.0021***	0.0025***	-0.0012	-0.0065***
	(0.0014)	(0.0019)	(0.0006)	(0.0007)	(0.0013)	(0.0017)
Effect of -10% benefits	-0.8	-2.2	0.5	0.6	-0.3	-1.6
Mean of outcome	32.9	32.9	20.5	20.5	53.4	53.4
Elasticity	0.26	0.68	-0.25	-0.30	0.06	0.31
Observations	211,978	289,303	212,547	289,702	212,273	289,718
			B. Without covariates			
Kink estimate ($\hat{\beta}$)	-0.0069***	-0.0151***	0.0022***	0.0025***	-0.0047***	-0.0127***
	(0.0019)	(0.0024)	(0.0006)	(0.0007)	(0.0018)	(0.0022)
Effect of -10% benefits	-1.7	-3.8	0.5	0.6	-1.2	-3.2
Mean of outcome	32.9	32.9	20.5	20.5	53.4	53.4
Elasticity	0.53	1.15	-0.27	-0.31	0.22	0.60
Observations	212,079	289,812	212,648	290,211	212,374	290,227

Notes: Table presents estimates from regressions of equation (1) using the tax returns data. A €600 bandwidth is used in all regressions, with a uniform kernel. Panel A includes covariates for the mother's age, the mother's age squared, the age difference with spouse, the base year earnings gap, whether the spouse is subject to the benefits cap, whether the couple lives in the former East, whether the couple is religious, and indicators for the tax return year and the event time year. Panel B shows the results without covariates, controlling only for the indicators for tax return year and event time. Leave period results pool $\tau = 0, 1$ and post-leave period results pool $\tau = 3, .., 9$. The effect of a ten percent decrease in the benefit amount is calculated as $E[-0.1b(v_i) \cdot \hat{\beta}/-0.65]$ where $b(v_i)$ is predicted benefit amount according to the schedule in Figure 1 and v_i is the mothers' base earnings. The elasticity is calculated as the effect of a 1% increase in predicted benefits divided by 1% of the mean of the outcome. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Robustness checks for RK estimates of impact on earnings gap

Model:	Baseline	Triangular kernel	Bias corrected	CCT bandwidth	IK bandwidth	FG bandwidth
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Leave period					
Kink estimate ($\hat{\beta}$)	-0.0034***	-0.0025*	-0.0027	0.0027	-0.0008	-0.0010
	(0.0013)	(0.0015)	(0.0018)	(0.0052)	(0.0007)	(0.0008)
Bandwidth	600	600	600	211	1113	979
Observations	211,978	211,978	211,978	61,644	564,987	455,503
	B. Post-leave period					
Kink estimate ($\hat{\beta}$)	-0.0087***	-0.0076***	-0.0096***	-0.0213**	-0.0087***	-0.0060***
	(0.0013)	(0.0017)	(0.0018)	(0.0104)	(0.0008)	(0.0021)
Bandwidth	600	600	600	140	900	427
Observations	289,303	289,303	289,303	53,263	570,920	183,821

Notes: Table presents estimates from regressions of equation (1) using different RK options. Column (2) uses a triangular rather than uniform kernel. Column (3) presents bias corrected robust estimates of Calonico et al. (2014) and uses a pilot bandwidth of €1500. Columns (4), (5) and (6) use the bandwidth selection procedures of Calonico et al. (2014), Imbens and Kalyanaraman (2012), and Fan and Gijbels (1995), respectively. The specifications are otherwise equivalent to those in Table 2, Panel A. See the notes from that table for the control variables included and definition of leave and post periods. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Robustness. Figure 6 and Table 3 examine the sensitivity of the RK estimates in the post-leave period to bandwidth choice and RK specification. Figure 6 reports estimates for the earnings gap, mothers' earnings, and fathers' earnings that are relatively stable across bandwidth choice between € 300 and € 1,500. Columns (1)–(3) in Table 3 show that the earning gap estimates are similar when using triangular instead of uniform kernel and a bias corrected robust estimate instead of conventional estimates.[17] Columns (4)–(6) show that the post-period results are robust to using the IK and FG optimal bandwidth selection procedures, but significantly larger when using CCT. For the leave period, the results are more sensitive to bandwidth choice. The concerns of Card et al. (2015) regarding the selection procedures appear relevant here, especially for CCT which chooses small bandwidths and delivers imprecise and unstable estimates.

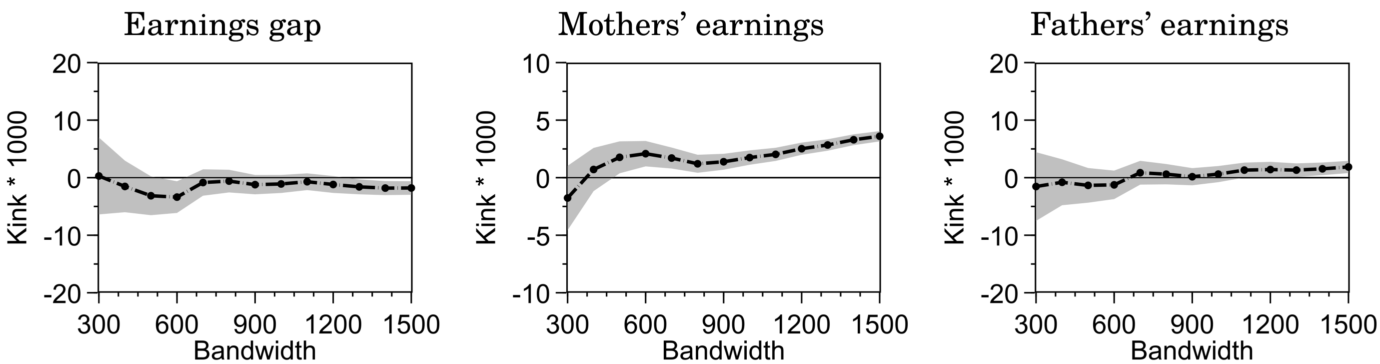
Table 4 performs placebo tests for the RK estimates. Panel A presents the placebo estimates obtained by estimating the kinks for couples observed in the pre-birth period. Here the estimates are small and mostly statistically insignificant. However, one exception is mothers' earnings, which is significant only at the 10% level and much smaller than the kinks estimated in the leave and post-leave periods. Furthermore, Panel B shows estimates of kinks in the post-leave period but for births pre-2007 that were not eligible for the benefits. Here the estimates are more imprecise due to a smaller sample size, but none are statistically significant. Altogether, the placebo results suggest that the estimates reported in the main table do not reflect non-linearities that existed either in the pre-birth period or for births that occurred in the pre-policy period.

In the appendix, I carry out further tests addressing possible concerns of sample section and attrition due to non-filing of tax returns. Table B1 uses data from 2013 onwards that includes non-filers automatically. It shows estimates are not sensitive to including non-filers focusing on births in the year 2014, for which non-filers are available in the pre-birth year, the leave period, and one post period. The analysis also shows that non-filers represent only 1-2% of the sample for my main estimations. Furthermore, Table B2 shows that non-filing in a later year is not impacted by the benefit amount, using this as an outcome in the RK design. As such, it rules out attrition as a source of bias in the main estimates.

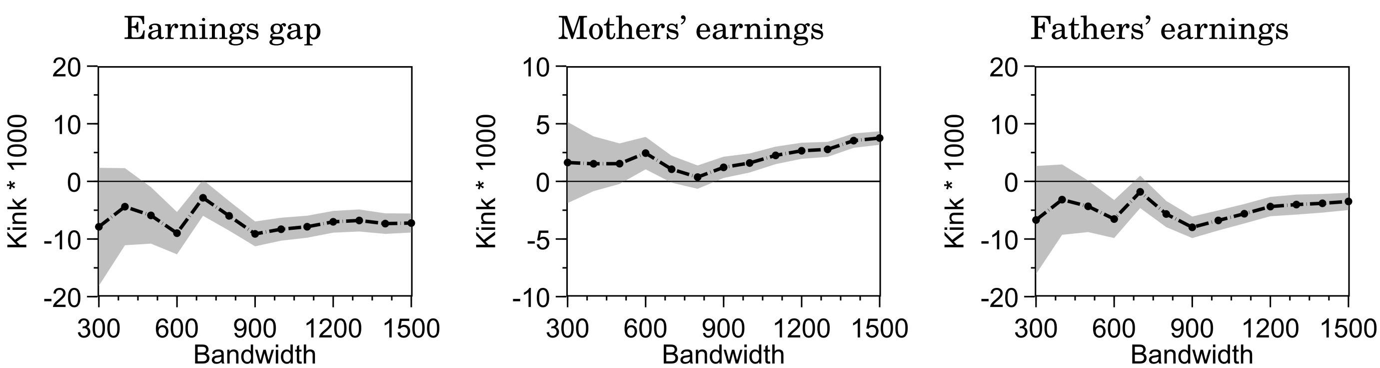
[17] Table B3 in the appendix shows comparable robustness findings using mothers' earnings and fathers' earnings individually as outcomes.

Figure 6: RK estimates in post-leave period by bandwidth choice

(a) Leave period



(b) Post-leave period



Notes: Figure shows RK estimates as black dots connected by a dashed line, from regressions of equation (1) using the bandwidth specified on the x-axis. The specifications are otherwise equivalent to those with covariates in Table 2, Panel A. See the notes from that table for the control variables included and definition of leave and post periods. The grey areas represent the 5% confidence intervals using robust standard errors.

Table 4: Placebo RK estimates

Outcome:	Earnings gap (1)	Mothers' earnings (2)	Fathers' earnings (3)
	A. Placebo outcomes: before birth year		
Kink estimate ($\hat{\beta}$)	-0.0010	0.0008*	-0.0003
	(0.0008)	(0.0004)	(0.0008)
Obervations	227,349	228,232	227,868
	B. Placebo births: before policy period		
Kink estimate ($\hat{\beta}$)	-0.0028	0.0009	-0.0025
	(0.0037)	(0.0015)	(0.0033)
Obervations	116,250	116,793	116,701

Notes: Table presents estimates from regressions of equation (1) using placebo outcomes and placebo births. Panel A uses outcomes measured in couples before the birth of their first child, pooling $\tau = -2, -3, -4$. Panel B uses post-birth outcomes, pooling $\tau = 2, ..., 9$, but for couples who have their first child before the benefits system is in place, i.e. in 2006 or earlier. The specifications are otherwise equivalent to those in Table 2, Panel A. See the notes from that table for the control variables included, bandwidth used, and definition of leave and post periods. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

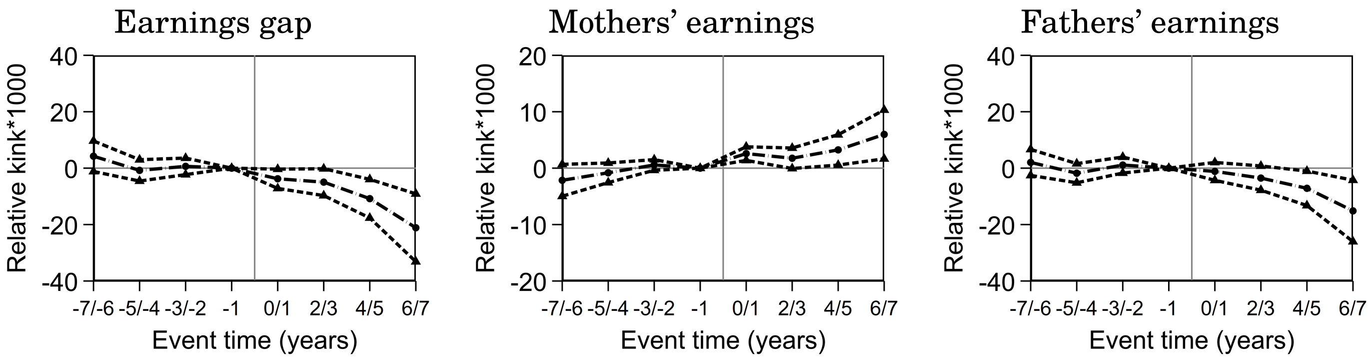
Dynamic effects and heterogeneity. Figure 7 plots estimates from regressions of the dynamic model of equation (2). The estimates represent the kinks relative to the base year at event times from seven years before the first birth until seven years after, in each case pooling two years. The estimates are scaled up by 1,000 for display purposes and, as such, dividing them by four gives approximately the impact of a ten percent decrease in benefits. The figure shows effect dynamics for the main sample, but also split by whether or not the mother earned more than the father in the base year. The pattern for the full-sample (panel a) confirms the main results but shows more precisely when the positive impact on mothers' earnings begins to grow larger and when the negative impact on fathers' earnings emerges. In both cases, it is around four years after the birth of the first child, which aligns with, and could potentially be related to, the slow down in having further children observed at this point in the data.[18] The pre-period estimates going back until seven years before the birth are all insignificant providing support for the parallel trends assumption of the event study design.

The sample split in panels (b) and (c) show that the leave-period impacts are sim-

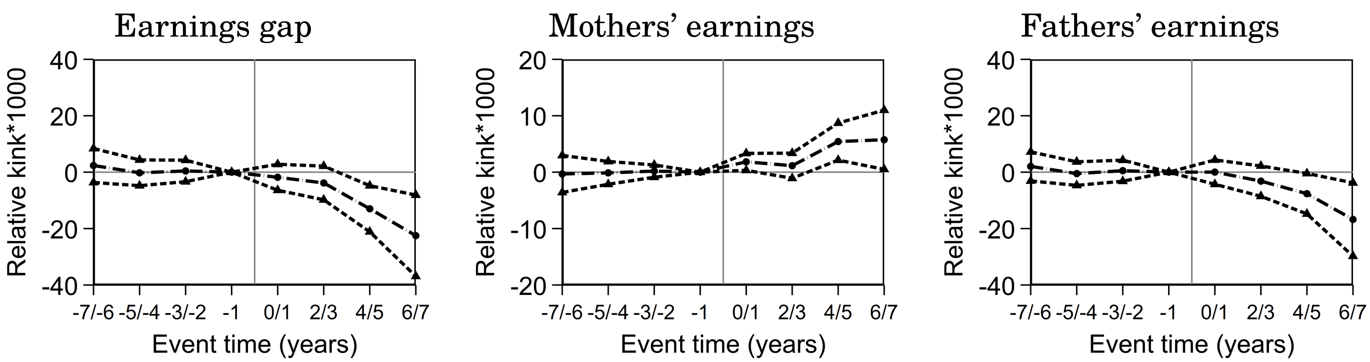
[18]This feature is observed in the further children variable in Table A2.

FIGURE 7: DYNAMIC RK ESTIMATES

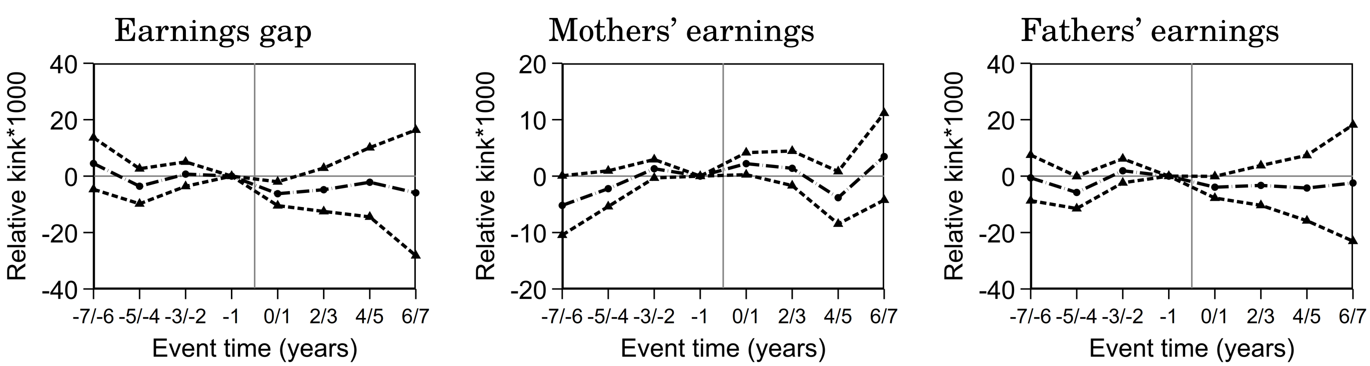
(a) Full sample



(b) Mother earned more than father pre-birth



(c) Mother did not earn more than father pre-birth



Notes: The black dots, connected by a dashed line, indicate separate estimates from regressions of equation (2) looking at the outcome at event times specified on the x-axis. The specifications are otherwise equivalent to those with covariates in Table 2, Panel A. See the notes from that table for the control variables included and bandwidth used. The triangles illustrate the 95% confidence intervals using robust standard errors.

ilar for both couple types but that there are important differences in the post-leave period. Only if the mother was previously the higher earner does the positive impact on mothers' earnings persist into the post-leave period, and only in those couples is there a negative impact on fathers' earnings. For couples where the mother was not the bigger earner in the base year, the positive impact on her earnings dissipates once the benefit period elapses and fathers do not reduce their earnings. A possible explanation for this heterogeneity is if couples where the mother did not earn more than the father hold stronger gender role attitudes and, as such, are less responsive to financial benefits that might promote equality (Bertrand et al., 2015). Table 5 presents evidence from the nationally representative German Socioeconomic-Panel that supports this explanation. Couples where the mother earned more than her partner in the base year are less likely to believe that a child under three suffers if his/her mother works. This particular gender role attitude has been demonstrated to strongly predict child penalties across countries (see Kleven et al., 2019a).

Mechanisms. Table 6 uses the RK design to examine the impact of the benefit amount on the taking of paid leave by mothers and fathers.[19] A key result visible in the full sample (panel A) is that the lower benefit amount reduces the probability that the mother takes any leave by around 1.6 percentage points. This impact is slightly larger in households where the mother previously earned more than the father, where it is also matched by an increase in the number of months of leave taken by the father. These findings present not just a plausible mechanism for the positive impacts on mothers earnings in the leave period but also a basis for long-run impacts on household specialisation in paid work consistent with the observed post-leave effects. However, the table also shows an *increase* in months of leave by the mother, albeit a small relative increase at 2% of the mean, and a decrease in take-up and duration of leave by the father in the full sample.

A possible interpretation of these results is based on conceptualising that there are two types of couple. First, in the gender-traditional type, the mother takes a long parental leave (with full certainly) whereas a short parental leave for the father, of perhaps 2-3 months, is considered optional. This type may be more greatly represented in Panel C, in line with evidence presented in Table 5. In such households, a lower mothers' benefit amount will not reduce the length of the mothers' leave but

[19] Respective RK plots are presented in Figure B1 in the appendix

Table 5: Parents' gender role attitudes by whether mother earned more than partner

Responses to question: To what extent do you agree with the statement "A child under three years suffers if their mother works"?						
Respondents:	Mothers			Their partners		
Couple type:	Mother earned more			Mother earned more		
	No	Yes	Diff	No	Yes	Diff
Score = 1 (Do not	0.22	0.40	0.175***	0.16	0.35	0.189***
agree at all)	(0.01)	(0.05)	(0.045)	(0.01)	(0.05)	(0.042)
Score = 2	0.19	0.25	0.056	0.16	0.19	0.030
	(0.01)	(0.04)	(0.042)	(0.01)	(0.04)	(0.041)
Score = 3	0.11	0.09	-0.027	0.11	0.11	0.003
	(0.01)	(0.03)	(0.034)	(0.01)	(0.03)	(0.035)
Score = 4	0.10	0.04	-0.061*	0.11	0.09	-0.018
	(0.01)	(0.02)	(0.032)	(0.01)	(0.03)	(0.034)
Score = 5	0.11	0.08	-0.034	0.12	0.08	-0.041
	(0.01)	(0.03)	(0.033)	(0.01)	(0.03)	(0.036)
Score = 6	0.10	0.05	-0.050	0.14	0.08	-0.060
	(0.01)	(0.02)	(0.032)	(0.01)	(0.03)	(0.038)
Score = 7 (Agree	0.16	0.10	-0.059	0.19	0.09	-0.103**
completely)	(0.01)	(0.03)	(0.038)	(0.01)	(0.03)	(0.043)
Score (average)	3.62	2.70	-0.922***	4.05	2.95	-1.100***
	(0.05)	(0.21)	(0.230)	(0.06)	(0.22)	(0.237)
N	1613	93	1706	1320	88	1408

Notes: Gender role attitudes question appears in the Families in Germany (FiD) module in the 2018 wave of the German Socioeconomic-Panel (G-SOEP). 'Mother earned more' is a binary indicator variable for whether the mother had higher gross annual earnings than her partner in the year before the birth year of her first child. Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: G-SOEP v35.

instead reduce household income during the leave period. Consequently, fathers in these households may decide to forgo their short leave in order to maintain their family's finances. The existence of such households could explain the negative impacts on fathers' leave take-up in column (3) and their varying impacts by panel. Furthermore, if the father forgoes his short leave, it may free up some months that can be transferred to mother, explaining the small increase in column (2).

Secondly, in the gender-egalitarian type, the mother is not necessarily tied to a long leave and the father is likely to take leave, and in some cases will take more than 2-3 months, depending on the circumstances. This type may be more greatly represented in Panel B. In these households, a lower mothers' benefit amount may induce the mother to not take any parental leave or to take a shorter one, whereby the father takes up some of the remaining months. Thus, the presence of this household type can explain the negative impacts in column (1) and their varying magnitudes in each panel. It can also explain why, in panel B, the father increases his months of leave. Note, that this average increase in duration comes despite a negative impact on fathers' take up. If the reduction in take up comes from the traditional couples, then the increase in fathers duration must be much larger than 0.245 months in gender-egalitarian couples, which is already quite large at around 17% of the mean.

Further outcomes. Table 7 shows the impact of the benefit amount on further outcomes.[20] Lower benefits reduce the chances of having further children in the post-leave period, thus reversing a temporary increase in this probability seen in the leave period. The negative impact on having children is stronger in households where the mother did not earn more than the father before birth. In these households there was no improvement in labour market equality in the long-run and a smaller reduction in leave taking, thus the major impact of the lower benefit amount would have been reduced family income during the early years of the first child's life. This could perhaps have changed the experience of child-rearing in a way that dissuaded couples from having a further child in the observation window. Furthermore, there appears an increase in the likelihood of splitting from partner for these couples. Nevertheless there is no significant impact on split from partner when looking across all couples.

[20]Respective RK plots are presented in Figure B1 in the appendix.

TABLE 6: RK ESTIMATES OF IMPACT OF MOTHERS' BENEFITS ON LEAVE TAKING

	Mother		Father	
Outcome:	Takes leave (yes/no)	Duration (months)	Takes leave (yes/no)	Duration (months)
	(1)	(2)	(3)	(4)
	A. Full sample			
Kink estimate ($\hat{\beta}$) $\times 1000$	-0.016***	0.229***	-0.102***	-0.159***
	(0.004)	(0.076)	(0.011)	(0.061)
Mean of outcome	0.98	10.52	0.59	1.40
Observations	207,923	207,923	207,923	188,176
	B. Mother earned more than father			
Kink estimate ($\hat{\beta}$) $\times 1000$	-0.017***	0.227**	-0.074***	0.245***
	(0.006)	(0.103)	(0.014)	(0.089)
Mean of outcome	0.97	10.25	0.59	1.44
Observations	108,178	108,178	108,178	87,645
	C. Mother did not earn more than father			
Kink estimate ($\hat{\beta}$) $\times 1000$	-0.010*	0.301***	-0.170***	-0.434***
	(0.006)	(0.115)	(0.020)	(0.076)
Mean of outcome	0.99	10.80	0.59	1.37
Observations	99,745	99,745	99,745	100,531

Notes: Table presents estimates from regressions of equation (1) using leave-taking outcomes for the mother and the father. The kink estimates are scaled by 1,000 for display purposes. The bandwidth used for these regressions € 1,000 which is larger than for other outcomes in order to make up for the loss of precision when missing the panel dimension. Take-up and duration of leave are proxy variables and their creation is described in the data section. The sample split is by whether the mother earned more than father in base year. The specifications are otherwise equivalent to those in Table 2, Panel A. See the notes from that table for the control variables included and definition of leave and post periods. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: RK estimates for further outcomes

Outcome:	Has further children		Split from partner	
Period:	Leave	Post-leave	Leave	Post-leave
	(1)	(2)	(3)	(4)
	A. Full sample			
Kink estimate ($\hat{\beta} \times 1000$)	0.0123*	-0.0249**	0.0032	0.0061
	(0.0067)	(0.0117)	(0.0068)	(0.0058)
Mean of outcome	0.06	0.54	0.05	0.05
Observations	212,857	290,153	227,880	314,860
	B. Mother earned more than father			
Kink estimate ($\hat{\beta} \times 1000$)	0.0207**	-0.0360**	0.0080	0.0023
	(0.0083)	(0.0144)	(0.0098)	(0.0075)
Mean of outcome	0.05	0.52	0.08	0.06
Observations	123,574	178,400	136,192	196,576
	C. Mother did not earn more than father			
Kink estimate ($\hat{\beta} \times 1000$)	-0.0025	-0.0450**	0.0213***	0.0320***
	(0.0119)	(0.0208)	(0.0069)	(0.0092)
Mean of outcome	0.06	0.57	0.02	0.03
Observations	89,283	111,753	91,688	118,284

Notes: Table presents estimates from regressions of equation (1) using further outcomes. The kink estimates are scaled by 1,000 for display purposes. A €600 bandwidth is used in all regressions. The creation of further kids and split from partner is described in the data section. The sample split is by whether the mother earned more than father in base year. The specifications are otherwise equivalent to those in Table 2, Panel A. See the notes from that table for the control variables included and definition of leave and post periods. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

VI. Child penalties

In this section, I aim to assess the extent to which the impacts of changes to the monthly benefit amount could affect long-run child penalties. My main aim is not the causal estimation of child penalties, but to illustrate the potential interplay between parental leave policy and gender inequality after child birth. I begin by estimating event study models separately for married men and women using the maximum span of the tax returns data:[21]

$$y^g_{it\tau} = \sum_{j \neq -1} \eta^g_j \mathbf{1}[\tau = j] + \sum_p \lambda^g_p \mathbf{1}[age_{it} = p] + \sum_q \mu^g_q \mathbf{1}[t = q] + \xi_{it\tau} \quad (3)$$

where $y^g_{it\tau}$ is the annual earnings of the partner of gender g from couple i at event time τ and tax year t, $\mathbf{1}[\tau = j]$ are event time indicators from five years before the birth ($\tau = -5$) until ten years after ($\tau = 10$) omitting the base year, $\mathbf{1}[age_{it} = p]$ are age indicators, and $\mathbf{1}[t = q]$ are tax year indicators. I follow the child penalties literature (e.g. Kleven et al., 2019b) in using earnings in levels, converting effects into percentages by calculating $P^g_\tau = \hat{\eta}^g_\tau / E[\tilde{y}^g_{i\tau t}|\tau]$, where $\tilde{y}^g_{i\tau t}$ is the predicted earnings omitting the impact of parenthood, i.e. $\tilde{y}^g_{i\tau t} = \sum_p \hat{\lambda}^g_p \mathbf{1}[age_{it} = p] + \sum_q \hat{\mu}^g_q \mathbf{1}[t = q]$. Child penalties are given by $P_\tau = \dfrac{\hat{\eta}^m_\tau - \hat{\eta}^w_\tau}{E[\tilde{y}^w_{i\tau t}|\tau]}$ and I take the tenth and last event year in my model as the long-run child penalty, i.e. P_{10}.

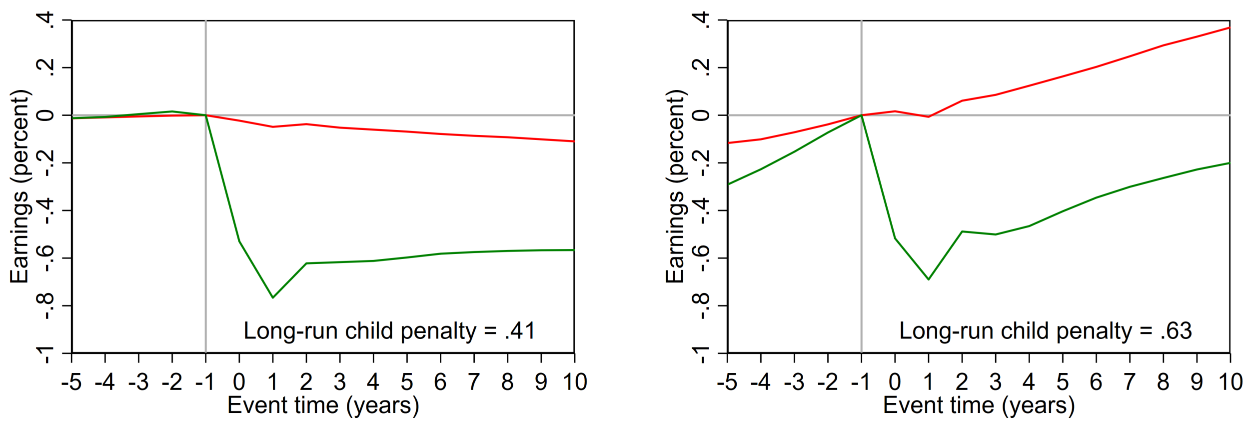
Figure 8 plots the results of event study estimations of equation (3) using all married couples (panel a) and the main RK sample with bandwidth €600 (panel b). The results using either of these samples are in line with the literature that shows a relatively unchanged earnings trajectory for men next to a dramatic drop in earnings for women at event time zero, which persists until the end of the data frame. The long run child penalty for all married women is 41% and for the RK sample it is larger still at 63%. These very large penalties are consistent with the long-run child penalty of 61% estimated by Kleven et al. (2019a) using representative data for Germany going back until 1984. Note that there is a deviation in parallel trends in panel (b) and the previously faster earnings growth for women implies the child penalty could be larger

[21]The data span of 14 years—2002 through 2016—is too short to create a balanced panel. Estimating the model with an unbalanced panel requires having at least one pre-birth year so I use all married men and women with births from 2003 until 2016.

still. The fact that the child penalty is larger in the RK sample where the mothers are higher earners is consistent with evidence suggesting child penalties are actually larger after controlling for previous income differences between partners (Andresen and Nix, 2021).

FIGURE 8: CHILD PENALTIES FOR MARRIED PARENTS

(a) All married couples (b) Main RK sample (bw=€600)



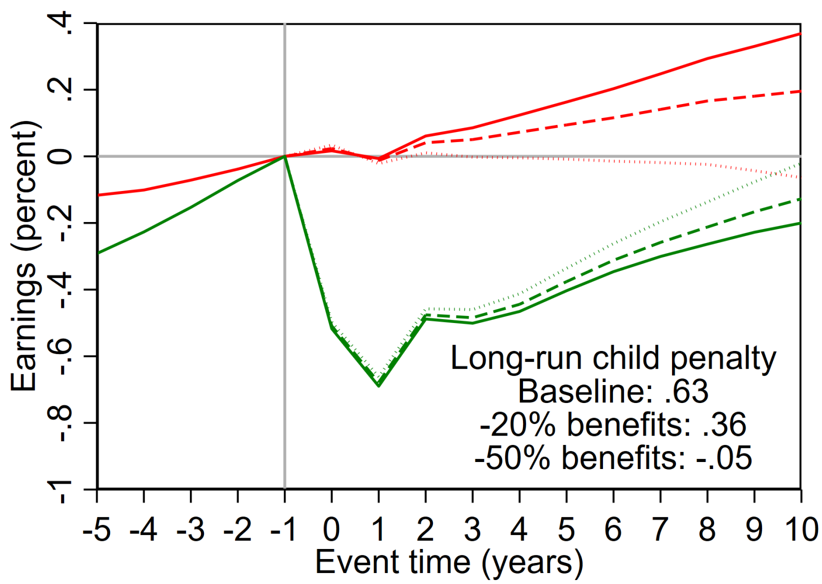
Notes: Figures plot estimates from regressions of equation (3). The red lines are male earnings trajectories captured by $\hat{\eta}^m_j$ and the green lines are female earnings trajectories $\hat{\eta}^f_j$ from five year before until ten years are the birth of the first child, relative to earnings in the base year ($j = -1$). Confidence intervals are not plotted but are so small as to be virtually indistinguishable from the coefficient lines, due to the large sample size.

Figure 9 shows results of event study estimations for the RK sample that use both observed earnings (solid line) as well as simulated earnings under the scenarios of a 20% (dashed) and a 50% (dotted) drop in the monthly benefit amount. Earnings are simulated making use of estimates from the dynamic RK design. Specifically, the simulated earnings in event time $\tau = j$ for a 20% decrease in benefits is calculated as $E[y^g_{itj} + -0.2b(v_i) \cdot (\hat{\beta}^g_j / -0.65)]$, where $\hat{\beta}^g_j$ are from the regressions of equation (2) that are presented in Figure 7.[22] The simulations suggest that a 50% reduction in the benefit amount would eliminate the long-run child penalty for couples within the €600 bandwidth. The full effect takes until ten years post birth to manifest, while a short-run child penalty remains in earlier event years. While a 50% cut in benefits is hardly a viable policy suggestion, the findings demonstrate that differences in finan-

[22]For the purpose of these simulations I estimate the dynamic models as far as the data allow until nine years after the first birth. I then use a linear regression to smooth the bin estimates that pool two event time years and extrapolate (slightly) until the tenth year.

cial incentives at a crucial point in time can dramatically alter the long-run paths in a way that improve gender equality in married couples where the mother is a high earner.

FIGURE 9: CHILD PENALTIES FOR MAIN RK SAMPLE (BW=€ 600) WITH SIMULATED BENEFITS CHANGE



Notes: Figures plot estimates from regressions of equation (3) using observed and simulated earnings from a 20% and 50% decrease in the benefit amount (solid, dashed, and dotted lines, respectively). The red lines are male earnings trajectories captured by $\hat{\eta}_j^m$ and the green lines are female earnings trajectories $\hat{\eta}_j^f$ from five year before until ten years are the birth of the first child, relative to earnings in the base year ($j = -1$).

VII. Generalisability of findings

In order to obtain internally valid estimates, my analysis up until now has focussed on a specific sample of married couples with high-earning mothers. In this section I aim to regain some generalisability by making use of an alternative approach and an alternative sample to examine the extent to which the findings hold for lower earning and unmarried women.

Effects for lower-earning women. To examine impacts for couples where the mother is not necessarily a higher-earner, I make use of a difference-in-differences (DD) approach that does not require a tight bandwidth around the benefit cap. I look at the difference in outcomes before and after the policy reform for couples where the mother has a high predicted benefit amount against couples where the mother has a low predicted benefit. Specifically, I estimate the following DD regression:

$$Y_{is\tau} = \pi + \rho b(v_i)\mathbf{1}[s \geq 2007] + \sigma b(v_i) + \phi\mathbf{1}[s \geq 2007] + \psi X_{is\tau} + e_{is\tau} \quad \textbf{(4)}$$

where $Y_{is\tau}$ is the outcome of couple i, with first birth in year s, observed in event time τ, $b(v_i)$ is the predicted benefit amount, which is a function of base earnings as depicted in Figure 1, $\mathbf{1}[s \geq 2007]$ is an indicator for the birth year s being in the post-reform period when the benefit schedule is in place, and $X_{is\tau}$ are the same controls used previously, including tax year and event time indicators. In the main specification, b is a continuous variable and is divided by -180 to approximate the impact of a 10% decrease in benefits for comparison with the RK estimates in Table 2. Under the parallel trends assumption, the DD estimate, $\hat{\rho}$, identifies the causal effect of the benefit amount.[23]

Table 8 presents the DD results and Figure 10 presents plots of parallel trends. In panel A, I use a sample that is €600 below the cap, broadly in line with the main RK specification.[24] The results of the DD using this sample are broadly consistent with those using the RK approach from Table 2. A 10% decrease in the mothers' benefit amount decreases the earnings gap in both the leave and post-leave periods, doing so through decreased mothers' earnings and increased fathers' earnings. Some of the magnitudes are also in line, such as the post-leave earnings gap decrease of around €2,000. There are some key differences, such as the positive impact on mothers' earnings being larger and the negative impact on fathers' earnings being a little smaller.

In panel B, I extend the sample to €1,500 below the cap to be more representative of the greater population. I do not use those couples below that bandwidth partly in order to avoid a negative treatment, whereby the previous means-tested benefits were

[23] Since my setting involves a single treatment date, I do not need to assume homogeneous treatment effects as highlighted in the DD literature (e.g. Goodman-Bacon, 2021).

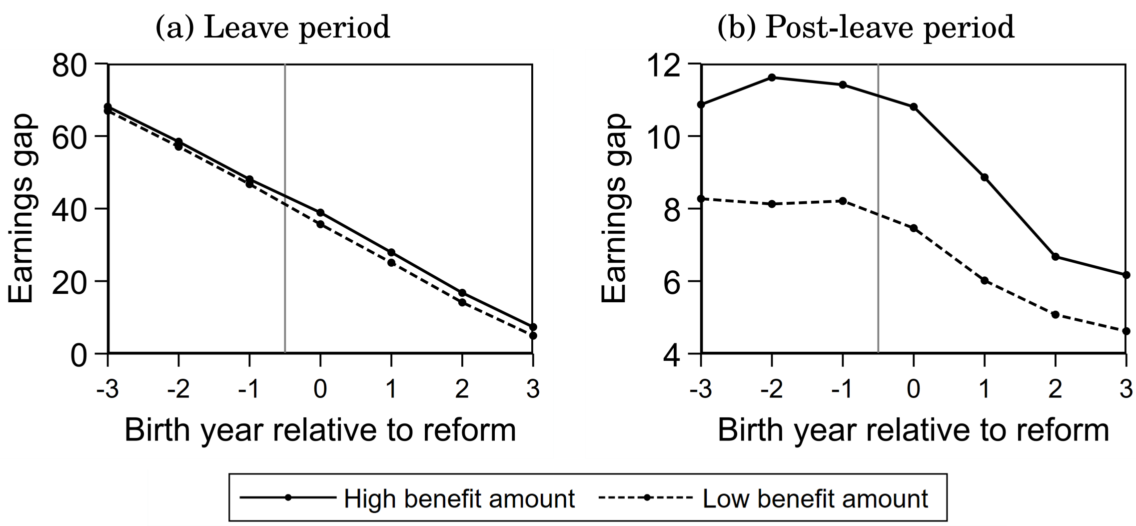
[24] I do not use the relatively small proportion of couples above the cap in either sample since they do not vary in predicted benefit amounts.

Table 8: Difference-in-differences estimates

Outcome:	Earnings gap		Mothers' earnings		Fathers' earnings	
	Leave	Post-leave	Leave	Post-leave	Leave	Post-leave
	(1)	(2)	(3)	(4)	(5)	(6)
	A. RK main sample (bw=600)					
DD estimate ($\hat{\rho}$)	-1.94***	-1.94***	1.33***	1.39***	-0.60	-0.72*
	(0.45)	(0.44)	(0.21)	(0.18)	(0.38)	(0.38)
Mean of outcome	30.9	36.3	19.2	28.3	50.0	64.4
Obervations	221,160	370,765	221,903	372,685	223,520	380,368
	B. Full sample window (bw=1500)					
DD estimate ($\hat{\rho}$)	-0.46***	0.65***	1.02***	0.99***	0.56***	1.61***
	(0.16)	(0.10)	(0.03)	(0.02)	(0.15)	(0.10)
Mean of outcome	32.2	38.2	8.5	11.8	40.7	49.8
Obervations	1,236,479	2,878,110	1,240,090	2,891,821	1,247,603	2,964,966

Notes: Table presents estimate from regression of equation (4). The DD estimate give the estimate on the interaction between an indicator for births post-2007 with a measure of the predicted benefit amount divided by -180 to give the impact of an approximate 10% decrease in the benefit amount. All specification control for mother's age, the mother's age squared, the age difference with spouse, the base year earnings gap, whether the spouse is subject to the benefits cap, whether the couple lives in the former East, and whether the couple is religious. Leave period results pool $\tau = 0, 1$ and post-leave period results pool $\tau = 3, .., 9$. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 10: Parallel trends plot



Notes: Plots show the evolution of the earnings gap before and after the birth year in couples in the full sample window where the mother receives an above median benefit amount (solid line) and a below median benefit amount (dashed line).

abolished as part of the 2007 reform. The results deviate in several ways when using this more representative sample. The negative impact on the gender gap is much smaller in the leave period due to a smaller positive impact on mothers' earnings but also a positive, rather than negative, impact on fathers' earnings. In the post-leave period, the impact on the gender gap is actually positive, mainly due to a large increase in fathers' earnings. This result could be explained if a lower mothers' benefit amount had the effect of decreasing fathers' leave taking and increasing mothers' leave taking, a possibility highlighted in the mechanisms section that may be even more applicable in the general population. While the post period findings find less support in parallel trends (see figure 10), this result does open up the possibility that the main findings are specific to that sample.

Effects for unmarried women. Table 9 reverts back to the RK approach but makes use of individual tax returns filed by women. For individual tax returns I cannot see partner outcomes or compute a gender gap, but am able to show impacts on mothers' earnings. Column (1) presents results for all mothers, column (2) for mothers filing an individual return in the base year (unmarried), and column (3) for mothers filing jointly (married).[25] For unmarried women, there is no significant impact of the benefit amount on earnings in the leave period and a negative impact on earnings in the post-leave period. A possible explanation for the lack of leave period effect is if unmarried mothers are less likely to have partners in the household that contribute to childcare. Table 10 presents evidence in support of this explanation from the G-SOEP dataset. The partners of unmarried women indeed contribute less to childcare in the birth year than partners of married women. Furthermore, unmarried women are much more likely to have no partner in the household at all. Altogether, unmarried women are nearly three times more likely to have no contributing partner in the household in the birth year. The negative effect on mothers' earnings in the post-leave period could be explained if non-contributing partners increase their employment work hours as a result of the lower mothers' benefit amount, with longer-run consequences for the gender split of childcare and paid work.

[25]Column (3) is a near identical specification and sample to the main results for mothers' earnings, with the only difference being the non-inclusion of controls for the age gap and pre-birth earning gap in order to remain consistent across specifications where partners are not be present.

Table 9: RK estimates of the impact on mothers' earnings by marital status

	All (1)	Unmarried (2)	Married (3)
	A. Leave period		
Kink estimate ($\hat{\beta}$)	0.0015***	0.0004	0.0021***
	(0.0005)	(0.0009)	(0.0006)
Obervations	326,073	113,493	212,580
	B. Post-leave period		
Kink estimate ($\hat{\beta}$)	0.0010	-0.0036**	0.0028***
	(0.0007)	(0.0016)	(0.0007)
Obervations	392,649	102,890	289,759

Notes: Table presents estimates from regressions of equation (1) for mothers' earnings including both married and unmarried women. The specifications are otherwise equivalent to those in Table 2, Panel A. See the notes from that table for the control variables included, bandwidth used, and definition of leave and post periods. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 10: Presence of contributing partner by marital status

	Unmarried	Married	Difference	Obs
Partner childcare (hours/day)	1.33	2.05	0.725***	3,060
	(0.19)	(0.05)	(0.186)	
Partner no childcare	0.65	0.28	-0.379***	3,060
	(0.03)	(0.01)	(0.030)	
No partner in household	0.34	0.02	-0.323***	3,349
	(0.02)	(0.00)	(0.011)	
No partner in household or partner no childcare	0.74	0.28	-0.462***	3,349
	(0.02)	(0.01)	(0.024)	

Notes: Sample restricted to mothers of first children born in 2007 or later. Separated and divorced mothers are excluded. Variables refer to birth year. Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: The German Socioeconomic Panel (G-SOEP) v35.

VIII. Conclusion

Using an RK design, I estimate the impact of the mothers' parental leave benefit amount on the gender gap in earnings in married couples after the birth of their first child. A reduced benefit amount decreases the gender gap in the leave period, an effect persisting into the post-leave period. The reduction in the gender gap in the leave period comes mainly through an increase in mothers earnings. In the post-leave period it comes through an increase in mothers' earnings and a reduction in fathers' earnings of similar relative magnitudes—but the post leave effects are driven entirely by couples where the mother previously out-earned the father. These heterogeneous findings suggest that large differences in financial incentives may shift the intra-family constellation only in cases where traditional gender norms are already weakened.

The estimated dynamic impacts of the benefit amount combined with simulations and an event study approach show that a 50% decrease in the parental leave benefit amount could eliminate the child penalty for a group of high-earning mothers. The results of this paper do not infer that benefit amounts *should* be reduced since they pertain to a particular group and there may be a range of other potentially negative effects to consider such as on fertility, household income, and child welfare. However, the results do suggest a promising role of family policy in addressing gender inequality in the labour market and they are supportive of policy design that changes the split of leave such as reserved partner months.

References

ANDRESEN, M. AND E. NIX (2021): "What causes the child penalty? Evidence from adopting and same-sex couples," *Journal of Labor Economics*.

ASAI, Y. (2015): "Parental leave reforms and the employment of new mothers: Quasi-experimental evidence from Japan," *Labour Economics*, 36, 72–83.

AVDIC, D. AND A. KARIMI (2018): "Modern family? Paternity leave and marital stability," *American Economic Journal: Applied Economics*, 10, 283–307.

BAKER, M. AND K. MILLIGAN (2008): "How does job-protected maternity leave affect mothersâ employment?" *Journal of Labor Economics*, 26, 655–691.

BANA, S. H., K. BEDARD, AND M. ROSSIN-SLATER (2020): "The impacts of paid family leave benefits: regression kink evidence from California administrative data," *Journal of Policy Analysis and Management*, 39, 888–929.

BERTRAND, M., E. KAMENICA, AND J. PAN (2015): "Gender identity and relative income within households," *The Quarterly Journal of Economics*, 130, 571–614.

CALONICO, S., M. D. CATTANEO, AND R. TITIUNIK (2014): "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 82, 2295–2326.

CARD, D., D. S. LEE, Z. PEI, AND A. WEBER (2015): "Inference on causal effects in a generalized regression kink design," *Econometrica*, 83, 2453–2483.

COOLS, S., J. H. FIVA, AND L. J. KIRKEBØEN (2015): "Causal effects of paternity leave on children and parents," *The Scandinavian Journal of Economics*, 117, 801–828.

DAHL, G. B., K. V. LØKEN, M. MOGSTAD, AND K. V. SALVANES (2016): "What is the case for paid maternity leave?" *Review of Economics and Statistics*, 98, 655–670.

DESTATIS (2018): "Taxpayer-Panel 2001-2016, GWAP, Version 0." Dataset of the Research Data Centres of the Federal Statistical Office (Destatis) and the statistical offices of the federal states. https://doi.org/10.21242/73111.2016.00.01.2.1.0.

DRUEDAHL, J., M. EJRNÆS, AND T. H. JØRGENSEN (2019): "Earmarked paternity leave and the relative income within couples," *Economics Letters*, 180, 85–88.

EKBERG, J., R. ERIKSSON, AND G. FRIEBEL (2013): "Parental leave–A policy evaluation of the Swedish "Daddy-Month" reform," *Journal of Public Economics*, 97, 131–143.

EU-SILC (2022): "Mean and median income by household type," *European Union Statistics on Income and Living Conditions*.

Fan, J. and I. Gijbels (1995): "Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation," *Journal of the Royal Statistical Society: Series B (Methodological)*, 57, 371–394.

Farré, L. and L. González (2019): "Does paternity leave reduce fertility?" *Journal of Public Economics*, 172, 52–66.

Frodermann, C., K. Wrohlich, A. Zucco, et al. (2020): "Parental Leave Reform and Long-Run Earnings of Mothers," Tech. rep., DIW Berlin, German Institute for Economic Research.

Ginja, R., J. Jans, and A. Karimi (2020): "Parental leave benefits, household labor supply, and childrenâs long-run outcomes," *Journal of Labor Economics*, 38, 261–320.

Goodman-Bacon, A. (2021): "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*, 225, 254–277.

Ichino, A., M. Olsson, B. Petrongolo, and P. S. Thoursie (2019): "Economic Incentives, Home Production and Gender Identity Norms," *IZA Discussion Paper Series*.

Imbens, G. and K. Kalyanaraman (2012): "Optimal bandwidth choice for the regression discontinuity estimator," *The Review of economic studies*, 79, 933–959.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimuller (2019a): "Child penalties across countries: Evidence and explanations," 109, 122–26.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2020): "Do family policies reduce gender inequality? Evidence from 60 years of policy experimentation," .

Kleven, H., C. Landais, and J. E. Søgaard (2019b): "Children and gender inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, 11, 181–209.

——— (2021): "Does biology drive child penalties? evidence from biological and adoptive families," *American Economic Review: Insights*, 3, 183–98.

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

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

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

LUNDBORG, P., E. PLUG, AND A. W. RASMUSSEN (2017): "Can women have children and a career? IV evidence from IVF treatments," *American Economic Review*, 107, 1611–37.

MULLEROVA, A. (2017): "Family policy and maternal employment in the Czech transition: a natural experiment," *Journal of Population Economics*, 30, 1185–1210.

OLAFSSON, A. AND H. STEINGRIMSDOTTIR (2020): "How does daddy at home affect marital stability?" *The Economic Journal*, 130, 1471–1500.

OLIVETTI, C. AND B. PETRONGOLO (2017): "The economic consequences of family policies: lessons from a century of legislation in high-income countries," *Journal of Economic Perspectives*, 31, 205–30.

PATNAIK, A. (2019): "Reserving time for daddy: The consequences of fathersâ quotas," *Journal of Labor Economics*, 37, 1009–1059.

PLANTENGA, J., C. REMERY, ET AL. (2006): "The gender pay gap. Origins and policy responses. A comparative review of thirty European countries," *The co-ordinators' synthesis report prepared for the Equality Unit, European Commission*.

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. (2017): "Maternity and family leave policy," Tech. rep., National Bureau of Economic Research.

SCHÖNBERG, U. AND J. LUDSTECK (2014): "Expansions in maternity leave coverage and mothersâ labor market outcomes after childbirth," *Journal of Labor Economics*, 32, 469–505.

STEARNS, J. (2018): "The long-run effects of wage replacement and job protection: Evidence from two maternity leave reforms in Great Britain," *Available at SSRN 3030808*.

TAMM, M. (2019): "Fathers' parental leave-taking, childcare involvement and labor market participation," *Labour Economics*.

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

WINGENDER, P. AND S. LALUMIA (2017): "Income effects on maternal labor supply: Evidence from child-related tax benefits," *National Tax Journal*, 70, 11–52.

APPENDIX (FOR ONLINE PUBLICATION)

A. Additional information on the data

Comparison with official births. Table A1 compares the number of first births to married couples seen in the tax returns data to the official count in years 2007 until 2015. It also provides the corresponding coverage rate in each year. Roughly 70% of all births of this type are seen in the tax returns data. The difference is likely accounted for by the fact that 27 million of Germany's 39 million tax payers (also about 70%) file an annual income tax return and, therefore, appear in the data. Table B1 makes use of 2013 and later waves that include non-filers to show that coverage is much higher for the relevant sub-sample.

Panel descriptives. Table A2 provides more summary statistics over the panel dimension of the data. The growth in the earnings gap from event time zero illustrates the child penalty that is widely documented in the literature. The transfer income variable is close to zero pre-birth showing that this field is mostly reserved for parental leave payments, especially in the mother's case. The number of observation drops at later years after birth where the end of the data is reached – however, the number of married couples that no longer file jointly, either because they are missing in the data or split is relatively small. The probability of having a further child rises rapidly from the second and third years after first birth.

TABLE A1: NUMBER OF FIRST BIRTHS TO MARRIED COUPLES

Year	Official data	Tax returns data	Coverage rate
2007	209,041	147,248	0.70
2008	208,371	142,013	0.68
2009	205,797	138,019	0.67
2010	208,245	142,262	0.68
2011	201,585	142,610	0.71
2012	201,346	144,395	0.72
2013	203,601	145,749	0.72
2014	213,668	151,833	0.71
2015	219,801	156,695	0.71

Notes: Table compares official birth count each year with the number of births seen in the tax returns data in each case focussing on first births to married couple. Source: Statistisches Bundesamt (Destatis).

Table A2: Key variables over event time

Variables	Event time (τ)													
	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9
Earnings gap	2.7	1.8	0.1	-2.5	29	38	33.1	37.4	39.4	39.6	40.7	42.7	45.0	46.4
Mothers' earnings	36.6	40.5	45.9	52.8	23.6	15.7	26.8	25.8	27.9	31.5	34.6	37.6	40.1	42.5
Fathers' earnings	39.3	42.3	46	50.3	52.6	53.7	59.9	63.2	67.3	71.1	75.3	80.3	85.1	88.9
Mothers' benefits	0.11	0.11	0.09	0.07	7.02	7.69	2.24	2.96	2.23	1.41	1.01	0.76	0.58	0.45
Fathers' benefits	0.17	0.18	0.17	0.16	0.89	1.52	0.53	0.75	0.6	0.45	0.35	0.33	0.28	0.31
Joint tax return	0.47	0.66	0.84	1	0.92	0.95	0.94	0.93	0.92	0.91	0.90	0.88	0.87	0.85
Missing	0.42	0.24	0.09	0	0.01	0.01	0.02	0.02	0.02	0.03	0.04	0.05	0.06	0.07
Split	0.11	0.11	0.07	0	0.07	0.04	0.04	0.05	0.05	0.06	0.07	0.07	0.07	0.07
Further children	0	0	0	0	0.03	0.08	0.31	0.52	0.62	0.65	0.67	0.66	0.66	0.66
Observations	58,552	64,644	78,661	98,000	97,438	81,521	66,750	53,867	42,023	31,783	22,525	15,343	9,565	3,938

Notes: Table shows the mean value for each outcome variable as well as the observation count at given event time, i.e. year relative to birth year of first child in a couple. Statistics refer to the sample with a bandwidth of €500 around the threshold. Earnings and benefits variables are in € thousands, and the remaining outcomes are indicator variables.

B. Further impacts of mothers' benefit cap

Further RK plots. Figure B1 shows RK plots for the further outcomes of fertility and marital stability, as well as the mechanism of leave-taking behaviour of each parent.

Non-filers. Table B1 shows results looking at both filers and non-filers with first births in 2014. The table makes use of the fact that, from 2013, taxpayers not filing a tax return were included in the dataset as automatically generated cases based on social security records. Selecting only births in 2014 means that I see both filers and non-filers in the pre-birth year, 2013, and can track their outcomes in the leave period (2014 and 2015) as well as one post-birth period (2016). The regressions focus only on mothers' earnings since a match with the partner is not possible using the automatically generated cases. The sample size is much smaller than the main specifications since births are restricted to one year only. Nevertheless, the filers results in column (2) are not too far off the main estimates, albeit less precise. Including non-filers makes almost no change to the estimates is column (1). This is because the number of non-filers is very small, corresponding to only about 1% of all observations. Therefore, while the number of non-filers in Germany is around 30% of taxpayers, it is a much smaller share for those married parents where the mother is a higher earner, within €600 of the threshold for the benefits cap. As such, the main estimates are likely to closely reflect the effect for the population of such couples.

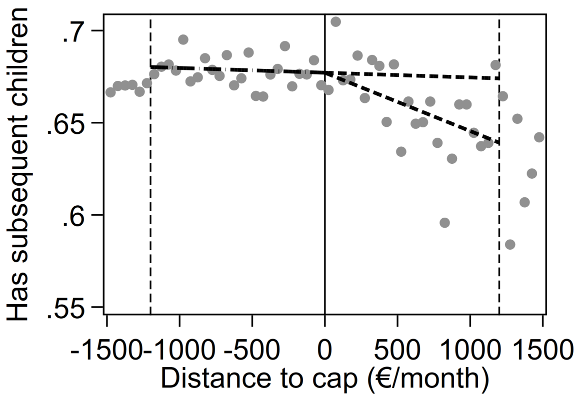
Data attrition. Table B2 shows results looking at couples not filing a tax return later in the panel data-set as an outcome variable. There is no significant impact for the full sample or either sub-sample, suggesting that dropping out of the sample of tax-return filing tax-payers is not influenced by the benefit amount, ruling this out as a source of bias in the main estimates.

Robustness for parental earnings. Table B3 reports RK robustness checks of the main effects for mothers' and fathers' earnings individually. It is analogous to the main paper Table 3 using earnings gap as the outcome and is consistent with those results.

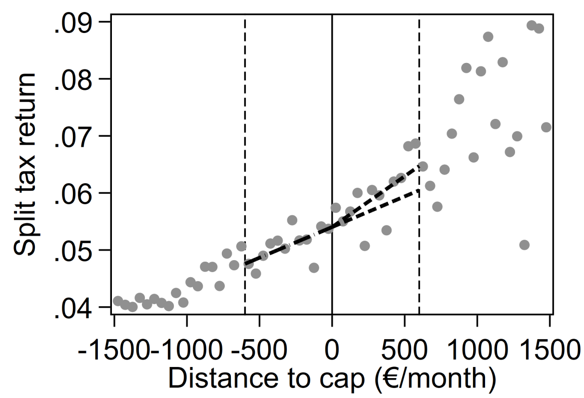
Figure B1: RK plots for further outcomes

(a) Further outcomes

Has further children ($\tau = 5$)

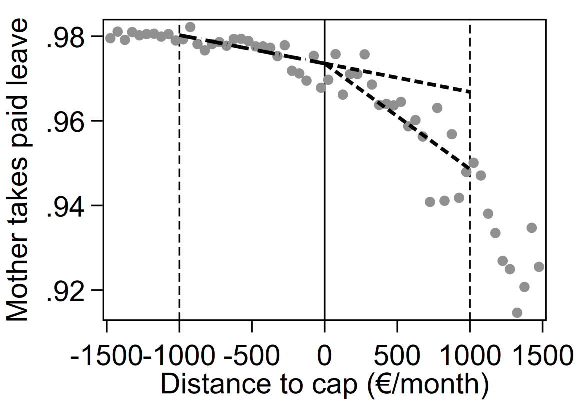


Split from partner ($\tau = 2, .., 9$)

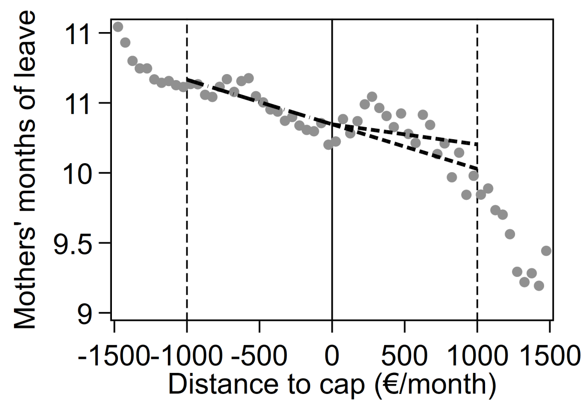


(b) Mothers' leave

Takes leave

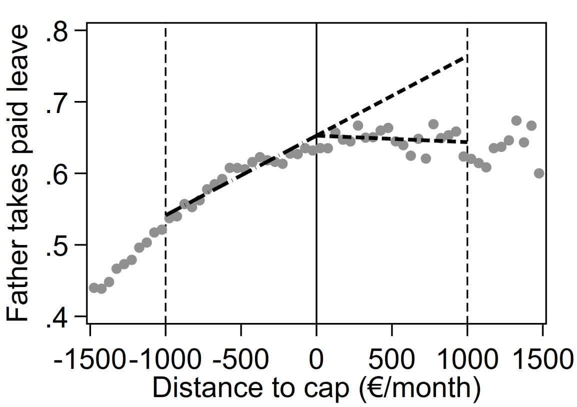


Duration (months)

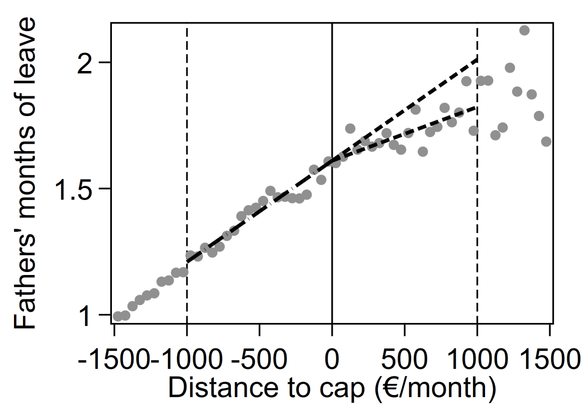


(c) Fathers' leave

Takes leave



Duration (months)



Notes: Figures shows mean parental earnings in €50 bins of the running variable. Black lines indicate the trend shift estimated within a bandwidth of €600 or €1000 around the cap.

Table B1: RK estimates for filers and nonfilers with births in 2014

	All (1)	Filers (2)	Non-filers (3)
	A. Leave period		
Kink estimate ($\hat{\beta}$)	0.0016	0.0016	0.0046
	(0.0012)	(0.0012)	(0.0106)
Obervations	50,487	49,514	973
	B. Post-leave period		
Kink estimate ($\hat{\beta}$)	0.0027	0.0029	-0.0023
	(0.0021)	(0.0021)	(0.0251)
Obervations	23,346	23,070	276

Notes: Table presents estimates from regressions of equation (1) for mothers' earnings including both filers and non-filers using data for 2013 and later where earnings for non-filers are included in the data based on social security records. The sample is restricted to couples with first births in 2014. The leave period is the same as in the main specification ($\tau = 1, 2$) but there is only one post leave period ($\tau = 3$), i.e. 2016. The specifications are otherwise equivalent to those in Table 2, Panel A. See the notes from that table for the control variables included and bandwidth used. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B2: RK estimates using non-filing as an outcome

	Leave (1)	Post-leave (2)
	A. Full sample	
Kink estimate ($\hat{\beta} \times 1000$)	0.0026	0.0011
	(0.0030)	(0.0040)
Mean of outcome	0.01	0.03
Observations	227,880	314,860
	B. Mother earned more	
Kink estimate ($\hat{\beta} \times 1000$)	0.0025	-0.0020
	(0.0040)	(0.0049)
Mean of outcome	0.01	0.03
Observations	136,192	196,576
	C. Mother did not earn more	
Kink estimate ($\hat{\beta} \times 1000$)	0.0038	0.0074
	(0.0045)	(0.0074)
Mean of outcome	0.01	0.02
Observations	91,688	118,284

Notes: Table presents estimates from regressions of equation (1) using an indicator for not entering the data (non-filer) in later periods as an outcome. The sample split is by whether the mother earned more than father in base year. The specifications are otherwise equivalent to those in Table 2, Panel A. See the notes from that table for the control variables included, bandwidth used, and definition of leave and post periods. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B3: Robustness checks for RK estimates of impact on mothers' earnings and fathers' earnings

Model:	Baseline	Triangular kernel	Bias corrected	CCT bandwidth	IK bandwidth	FG bandwidth
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Mothers' earnings, leave period					
Kink estimate ($\hat{\beta}$)	0.0021***	0.0013*	0.0001	-0.0015	0.0021***	0.0005
	(0.0006)	(0.0007)	(0.0008)	(0.0013)	(0.0006)	(0.0010)
Bandwidth	600	600	600	329	591	381
Observations	212,547	212,547	212,547	100,296	208,164	118,948
	B. Mothers' earnings, post-leave period					
Kink estimate ($\hat{\beta}$)	0.0027***	0.0025***	0.0008	0.0081***	0.0031***	0.0053***
	(0.0005)	(0.0007)	(0.0008)	(0.0023)	(0.0011)	(0.0020)
Bandwidth	600	600	600	212	364	230
Observations	289,702	289,702	289,702	83,005	151,839	90,603
	C. Fathers' earnings, leave period					
Kink estimate ($\hat{\beta}$)	-0.0012	-0.0012	-0.0026	0.0004	-0.0006	-0.0001
	(0.0011)	(0.0013)	(0.0016)	(0.0037)	(0.0017)	(0.0006)
Bandwidth	600	600	600	239	441	923
Observations	212,273	212,273	212,273	70,358	141,898	413,777
	D. Fathers' earnings, post-leave period					
Kink estimate ($\hat{\beta}$)	-0.0061***	-0.0053***	-0.0088***	-0.0011	-0.0058***	-0.0040***
	(0.0012)	(0.0015)	(0.0017)	(0.0038)	(0.0007)	(0.0011)
Bandwidth	600	600	600	267	841	631
Observations	289,718	289,718	289,718	106,926	503,828	311,636

Notes: Table presents estimates from regressions of equation (1) with mothers' and fathers' earnings as outcomes using different RK options. Column (2) uses a triangular rather than uniform kernel. Column (3) presents bias corrected robust estimates of Calonico et al. (2014) and uses a pilot bandwidth of €1500. Columns (4), (5) and (6) use the bandwidth selection procedures of Calonico et al. (2014), Imbens and Kalyanaraman (2012), and Fan and Gijbels (1995), respectively. The specifications are otherwise equivalent to those in Table 2, Panel A. See the notes from that table for the control variables included and definition of leave and post periods. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.