OXFORD

Do Parental Leaves Make the Motherhood Wage Penalty Worse?

Gabriele Mari [ID] [1,*] and Giorgio Cutuli[2]

[1]Department of Public Administration and Sociology, Erasmus University Rotterdam. 3062 PA, Rotterdam, the Netherlands and [2]Department of Sociology and Social Research, University of Trento, 38122, Trento, Italy

*Corresponding author. Email: mari@essb.eur.nl

Abstract

We assess if and how motherhood wage penalties change in response to the design of parental leave regulations. Focusing on Germany, we compare sweeps of reforms inspired by opposite principles. One allowed for longer periods out of paid work in the 1990s, the other prompted quicker re-entry in the labour market in the late 2000s. These reforms may have first exacerbated and later mitigated wage losses for new mothers, albeit each component of leave schemes may trigger separate, and at times zero-sum, mechanisms. We rely on Socio-Economic Panel (SOEP) data and a difference-in-differences design. Focusing on first-time mothers, we find that motherhood wage penalties were substantial (around 20–30 per cent of pre-birth wages) and also changed little during the 1990s. As parental leave reform triggered longer time spent on leave coupled with better tenure accumulation, wage losses for mothers remained stable in this first period. Following parental leave reform in the late 2000s, instead, the wage prospects of first-time mothers improved, thanks in part to shorter work interruptions and increased work hours. We suggest that the nuts and bolts of leave schemes can be fine-tuned to reduce child penalties and, thus, gender wage disparities.

Introduction

Mothers typically face wage losses when reprising paid work after having a child. Such a motherhood wage penalty has become a key component of gender pay gaps in labour markets, but its size and roots are still debated (for a review, Ponthieux and Meurs, 2015). Women's labour supply patterns, particularly work interruptions for family-related reasons, account for at least part of the wage dip (Albrecht *et al.*, 1999; Gupta and Smith, 2002; Gangl and Ziefle, 2009; Adda, Dustmann and Stevens, 2017). Work interruptions are in turn influenced by leave policies and reforms to those policies over time (Gregg, Gutíerrez-Domènech and Waldfogel,

2007; Lalive and Zweimüller, 2009; Ziefle and Gangl, 2014; Baum and Ruhm, 2016). Few studies, however, have questioned if and how motherhood wage penalties may change in relation to leave reform.

We address these questions here, focusing on two decades of German reforms (e.g. Schönberg and Ludsteck, 2014; Ziefle and Gangl, 2014; Kluve and Schmitz, 2018). Germany offers an exemplary case study for it once encompassed a maternalist leave scheme among the longest and most generous in high-income countries, and has now shifted to provisions similar to those of Scandinavian countries (Ray, Gornick and Schmitt, 2010). Germany also features one

of the largest motherhood penalties and gender pay gaps in international comparison (e.g. Olivetti and Petrongolo, 2008; Gangl and Ziefle, 2009), with far-reaching consequences on intra-household equilibria (Musick, Bea and Gonalons-Pons, 2019), on the opportunity-cost of children amidst Germany's fertility decline (Buhr and Huinink, 2015), and on gender economic disparities along the lifecycle (OECD, 2017: p. 173; see also, Lersch, Jacob and Hank, 2017).

Relying on high-quality panel data from the Socio-Economic Panel (SOEP 1985–2014), we make two main contributions. First, we show that wage penalties might be larger than what existing estimates suggest (e.g. Cukrowska-Torzewska and Matysiak, 2020; de Linde Leonard and Stanley, 2020), up to 20–30 per cent of pre-birth wages. Previous studies largely averaged wage losses across all periods after the birth of a child. We distinguish these periods in an event-study framework and suggest that this assumption might be unwarranted, with implications for the size and, hence, for the societal relevance of motherhood wage penalties.

Second, we show that child penalties can change over time, exploiting two major parental leave reforms. In doing so, we argue that the 'nuts and bolts' of leave policy—duration, job protection rights, sharing rights between partners, etc.—might trigger distinct, and at times zero-sum, mechanisms. We find that German reform in 1992 prolonged women's work interruptions, hardly changed their working hours after re-entry, but improved their accumulation of job tenure thanks to job-guarantee rights. This resulted in large yet stable wage losses, that substantially improved only after the latest reform in 2007. In fact, we can no longer detect (large) penalties for mothers after 2007, a finding we impute to shorter career breaks and longer hours upon re-entry—both plausibly aided by new leave-sharing rights for partners.

Previous studies have either inferred the importance of institutions by comparing the magnitude of wage loss across countries (Davies and Pierre, 2005; Gangl and Ziefle, 2009) or by investigating change over time in a single country in relation to changing bundles of work–family policies (Petersen, Penner and Høgsnes, 2014). Others have highlighted robust associations between child penalties and leave policies, but only focusing on leave length and only across countries (Budig, Misra and Boeckmann, 2016). Our study complements these, suggesting that leave policy can shape motherhood wage penalties, over time and within a single country, albeit much depends on policy design. In line with few others in the field (Andersen, 2018), we suggest that in-depth consideration of policies and their effects can further advance our understanding of the roots of child penalties. Our findings inform existing debates on the consequences of work–family policies (e.g. Ruhm, 1998; Mandel and Semyonov, 2006; Arulampalam, Booth and Bryan, 2007; Mandel, 2012; Aisenbrey and Fasang, 2017) and on German parental leave in particular (e.g. Schönberg and Ludsteck, 2014; Ziefle and Gangl, 2014; Cygan-Rehm, 2016; Raute, 2019), both of which have overlooked motherhood wage penalties.

Background

Parental Leave Reforms and Maternal Labour Supply

Maternity leave in Germany has long covered a period of 14 weeks with full income replacement and job protection. Parental leave, on the other hand, has been subject to two sweeps of reforms, one during the 1980s–1990s and culminating in 1992, and another during the 2000s and culminating in 2007. Reforms variously redesigned the scheme with respect to its duration, job-guarantee rights (i.e. to return to a comparable job with the pre-birth employer after taking leave), payments, and sharing rights between partners.

Prior to 1986, employed (West) German mothers could access up to 6 months of paid and job-guaranteed parental leave. A number of reforms progressively increased the duration of the job guarantee. Most strikingly, such duration was doubled in 1992, from 18 to 36 months. In the meantime, benefits also changed, switching from earnings-related to a mix of flat-rate and means-tested payments. Payments became available to all mothers regardless of pre-birth employment status (1986), for up to 12 months since 1992 and 24 months since 1993. In short, early reform granted long spells of paid leave, with extensive job protection rights.

In contrast, reforms in the 2000s sought to make shorter work interruptions more financially appealing, with the aim of fostering, at once, paid work and family formation. In 2001, mothers returning to paid work after 12 months of leave, rather than 24, could have access to a larger payment (flat-rate, 450 EUR). Parents, maintaining their eligibility intact, could also work up to 30 hours a week while on leave—in contrast to the 19-hour limit in place prior to 2001. The 2007 reform (*Elternzeitgesetz*) went further, limiting paid leave to 12 months, or 14 when each parent takes at least 2 months. Benefits are now earnings-related once again, with a replacement rate of around 67 per cent of the pre-birth net labour earnings and a cap at 1,800 euros a month. Parents who are not in employment in the year prior to childbirth are entitled to a minimum of 300

euros a month, similar to the pre-2007 regime. Quite importantly, the job-guarantee period remained unchanged throughout. In short, reforms in the 1980s/1990s broadened benefit receipt and extended the job guarantee to allow mothers a prolonged absence from paid work. Reforms in the 2000s squeezed more generous benefits in shorter periods of benefit receipt with the explicit aim of maintaining mothers in employment (see also Kluve and Tamm, 2013; Schönberg and Ludsteck, 2014; Ziefle and Gangl, 2014).

According to previous studies, the introduction or expansion of paid leaves may lead to sharp reductions in women's employment probability right after childbirth (e.g. Gregg, Gutíerrez-Domènech and Waldfogel, 2007; Lalive and Zweimüller, 2009). In the long run, however, paid leaves may exert positive effects on job continuity and increase the share of women returning to the labour market when entitlements expire (e.g. Baker and Milligan, 2008; Schönberg and Ludsteck, 2014; Baum and Ruhm, 2016). In Germany, parental leave reform has been found to neatly shape the short-term labour supply of mothers (Schönberg and Ludsteck, 2014; Ziefle and Gangl, 2014). After 1992, the median length of leave periods in the West rose to 27 months and returns started peaking also at 36 months, coinciding with the exhaustion of job-protection rights. In the East, mothers responded similarly to policy change, albeit to a lesser extent given historically superior female labour force participation. Early reforms particularly depressed mothers' short-run chances of working full-time (Schönberg and Ludsteck, 2014; Arntz, Dlugosz and Wilke, 2017). Part-time employment conversely became the norm for mothers returning to paid work, particularly in former West Germany and, to a lesser extent but increasingly over time, in the former East too (Trappe, Pollmann-Schult and Schmitt, 2015; Dieckhoff *et al.*, 2016). Although around half of returners maintained their pre-birth employer in the aftermath of 1992, the share of returners with a new employer and that of those having a second child out of inactivity also increased (Arntz, Dlugosz and Wilke, 2017). Yet, in the long run, the impact of parental leave expansion on mothers' accumulated labour market experience was modest (Schönberg and Ludsteck, 2014; see also Lalive and Zweimüller, 2009 for similar conclusions on leave expansion in Austria). Reforms seemingly accentuated positive selection into employment, as employment (and job) continuity after motherhood became even more skewed in favour of highly educated women (Drasch, 2012; Arntz, Dlugosz and Wilke, 2017).

In the midst of and contributing to rising female labour market participation, parental leave reforms in the 2000s lead to a reversal in maternal labour supply in the first years after childbirth. While evidence is mixed on the role of the 2001 reform (cf. Fitzenberger, Sommerfeld and Steffes, 2013; Ziefle and Gangl, 2014), the 2007 reform seems to have fuelled strong labour supply responses. Consistent evidence in the literature indeed points to *increased* time spent off work in the first year after childbirth—that is, during *paid* leave—coupled though with higher re-employment chances and longer working hours after the 12-month mark, when payments expire (Kluve and Tamm, 2013; Ziefle and Gangl, 2014; Bergemann and Riphahn, 2017; Kluve and Schmitz, 2018). Still, long-run maternal labour force participation was largely unaffected by the new regime and women's higher propensity to reprise working part-time rather than full-time, as well as the 'high-skill skew' among employed mothers, have persisted (Drasch, 2012; Bergemann and Riphahn, 2017; Kluve and Schmitz, 2018).

Overall, parental leave reform triggered longer short-term dips in the labour supply of mothers, and these dips significantly shortened only in the late 2000s. Throughout, selective return and part-time work remained common among new mothers. The extension of the job guarantee in 1992 and the reform of the benefit scheme in 2007 have been identified as the two watershed reforms with respect to the labour supply behaviour of German mothers (Kluve and Tamm, 2013; Ziefle and Gangl, 2014; Bergemann and Riphahn, 2017; see also Fitzenberger, Sommerfeld and Steffes, 2013; Gangl and Ziefle, 2015). In the remainder, we will thus focus on these two reforms. For a summary, we portray the reforms' key features, in terms of duration, benefit receipt, and job protection, in Figure 1.

Wage Responses to Parental Leave Reform

Parental leave regulations may affect wages via several mechanisms. From a human capital perspective, leaves, and wage attainment are linked both at the extensive and at the intensive margin of labour supply. At the extensive margin, expanding the duration of parental leaves might exacerbate motherhood wage penalties. Prolonged leave uptake can lead to human capital loss and skill atrophy. This may result in lower wage offers for mothers returning to paid work after taking leave (Gupta and Smith, 2002; Anderson, Binder and Krause, 2003; Adda, Dustmann and Stevens, 2017). Conversely, if leaves grant job-protected work interruptions, job tenure (and, thus, firm-specific human capital) can be preserved and motherhood wage penalties could be offset (Waldfogel, 1998; Baker and Milligan, 2008; Zhang,
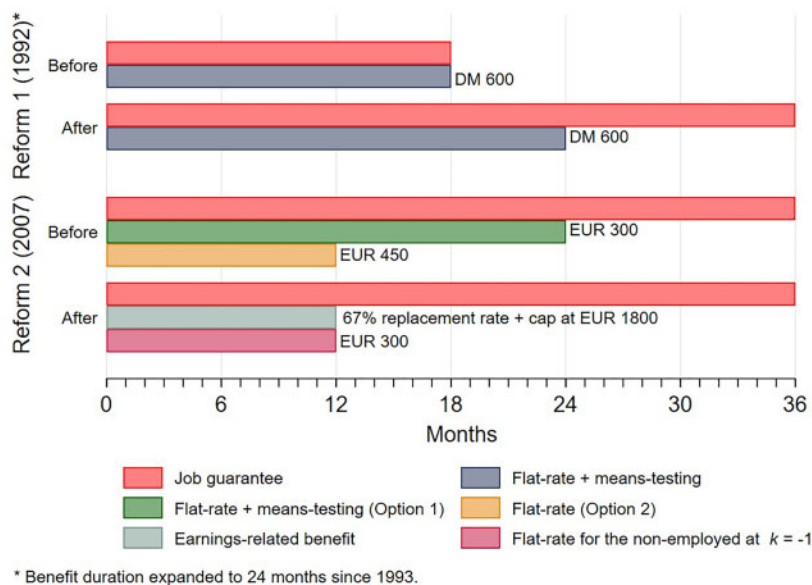
Figure 1. Summary of the relevant changes to benefit receipt and job guarantee periods under different parental leave regulations

Note: the year $k = 1$ refers to the year prior to the birth of a new child for a given woman.

2010; Fernández-Kranz, Lacuesta and Rodrìguez-Planas, 2013).

Hence, it is unclear whether the 1992 reform made wage penalties worse. On the one hand, it doubled leave duration with the risk of intensifying human capital loss. On the other hand, though, the reform extended job-guarantee rights to 36 months, and this may have helped mothers maintain their tenure with their pre-birth employer. The impact of the 2007 reform is more clear-cut. Shorter benefit duration incentivized mothers to concentrate leave uptake in the first 12 months after childbirth, and then resume employment, thereby mitigating wage losses.

At the intensive margin, parental leaves can affect wages via working hours. Part-time employment has long been the modal arrangement after taking leave in (West) Germany (e.g. Trappe, Pollmann-Schult and Schmitt, 2015; Dieckhoff *et al.*, 2016). This can depress hourly wages immediately upon re-entry, but comparative research has shown that part-time jobs are relatively well-paid in Germany (Bardasi and Gornick, 2008) and are associated with lower wage rates only for very short schedules (5–15 hours, Paul, 2016). Nevertheless, wage returns to experience are typically inferior in part-time (vs full-time) employment, regardless of the 'intensity' of part-time and to the detriment of wage attainment in the long run (Fernández-Kranz, Paul and Rodrìguez-Planas , 2015; Paul, 2016).

Mothers subject to the 1992 regulations have been shown to display stronger preferences for family commitments over paid work (Gangl and Ziefle, 2015). This could have translated in effort re-allocation, manifesting in shorter work hours and low wages. In contrast, the 2007 reform spurred re-entries with longer hours, albeit still predominantly within part-time schedules (e.g. Kluve and Schmitz, 2018). New mandates for fathers may have played a role in this respect (Bünning, 2015), freeing time for mothers to work longer hours and thereby improving their wage prospects (see e.g. Andersen, 2018), at least in the short run.

Alternative to a human capital perspective, it could be that employers interpret changes in maternal behaviours and preferences as market signals. Given that leave *uptake* is widespread among mothers, leave *length* may serve as a signal (Albrecht *et al.*, 1999; Albrecht, Thoursie and Vroman, 2015; Evertsson, 2016). The price of taking family leave may spike if women stay out more than what is statutorily granted or if—given the choice—they spread their leave period rather than exhaust it all at once. These might be perceived as signals of low commitment to the job or employment in general. Coherently, research has highlighted a significant jump in the wage penalty for leaves exceeding the job-guaranteed period of 36 months in Germany in the years 1994–2005 (Buligescu *et al.*, 2009). Similar 'threshold

effects' have been shown to hold also in other countries (Albrecht, Thoursie and Vroman, 2015; Evertsson, 2016). Signalling may thus generate heterogeneity in the effects of each reform. In the midst of a general drift towards longer career interruptions after 1992, mothers taking shorter leaves could have positively signalled themselves to employers, perhaps avoiding a (more) negative wage shock. In contrast, women could have sent adverse signals by not complying with the new 12-month interruption norm after 2007 (e.g. Bergemann and Riphahn, 2017), resulting in wage penalties also under the new leave regime.

In sum, human capital accounts provide only mixed indications for the effects of the 1992 reform, due to the combination of long work interruptions, extensive job protection, and short hours. The 2007 reform, on the other hand, may have improved the wage prospects of mothers in the short run, thanks to more concentrated leave breaks and longer hours. Extensive part-time work, if sustained, might still lead to wage deterioration though. For both reforms, signalling predicts instead discontinuities in the wage penalty, depending on the timing of mothers' return to work vis-à-vis leave 'thresholds'. Time spent on leave (human capital/signalling), working hours (human capital), and job tenure (human capital) are thus the proxies (mechanisms) we will investigate to make sense of the reforms' wage effects, if any.

Empirical Approach

Data and Samples

We use longitudinal data from the German SOEP (v. 31.1, German Institute for Economic Research (DIW), Berlin, 2016), a multipurpose household panel survey carried out annually since 1984 (Goebel *et al.*, 2018). We rely on samples A–K (with the exception of the Innovation Sample I), that is, all original samples for both West and East Germany as well as refreshment and boost samples added up to 2012.

We ran separate analyses for the two reform periods. To evaluate the extension of the job guarantee in 1992 (Reform 1, hereafter), we focus on the period between 1985 and 1998. To examine the change in the benefit scheme carried out in 2007 (Reform 2, hereafter), we select the subsequent time window between 1999 and 2014. The choice of these thresholds allows us to have periods of equal length prior and after each reform (see e.g. Francesconi and Van der Klaauw, 2007). The two midpoints, 1992 and 2007, respectively, fit precisely with the culmination points of each sweep of reforms, as

previously discussed. The extensiveness of the time window is also motivated by the specificity of the treatment effects we are investigating (e.g. Ziefle and Gangl, 2014). Since women may take up to 3 years of parental leave from 1992 onwards, and since we necessarily measure their post-birth wages only once they returned to the labour market anyway, we need to allow (enough) women in our sample to make such re-entry in paid work.

Following conventional practices in the literature, our sample is restricted to women aged 16–45, working as dependent employees, with at least two valid observations for any of our outcome variables, and with non-missing information on all other variables involved in the analysis. To fully reconstruct women's fertility biographies, we take advantage of data from the Biography and Life History module of SOEP (Goebel, 2017). We are thus able to build on info on the timing of childbirth events, precise to the month and available up to the 15th parity.

We define the group *treated* by the policy change as those women who become mothers for the first time between 1992 and 1998 for Reform 1 and between 2007 and 2014 for Reform 2. The control group[1] in each case also comprises first-time mothers, who have given birth for the first time between 1985 and 1991 for Reform 1, and between 1999 and 2006 for Reform 2. To avoid overlaps between the two treatment arms, mothers belonging to each control group did not give birth to a child in the respective post-reform periods and are thus unaffected by parental leave reforms[2] We also drop all person-year records past the fifth year after childbirth due to the small sample size for $k \geq 6$. Finally, women who do not give birth in either observation period are not part of the analyses. This choice is, nonetheless, largely inconsequential for our estimates (see Supplementary Section A5).

We are thus left with 706 women (of which 462 are treated) for the evaluation of Reform 1 and 1,017 women (of which 494 are treated) for Reform 2, followed for an average of 5.6 and 6.2 waves, respectively. These numbers are comparable to those of previous research on parental leave effects on labour supply and earnings (e.g. Joseph *et al.*, 2013; Bergemann and Riphahn, 2017).

Estimating the Effects of Parental Leave Reforms

We implement a difference-in-differences (DiD) design, augmented by entropy balancing (Hainmueller, 2012) and individual fixed effects (FE) (as in, e.g. Francesconi and Van der Klaauw, 2007; Gangl and Ziefle, 2015).

These add-ons have several advantages. First, weighting our estimates via entropy balancing improves the comparability of mothers in each 'treatment-control' pair, at least on the basis of a set of observable characteristics.

The latter are variables tapping into women's work histories and household characteristics; we present the procedure in full in Supplementary Section A2. By including individual FEs, we remove bias due to compositional differences across groups due to time-invariant unobserved characteristics, as long as these have constant wage returns (implied by the 'fixed' effect).

These choices arguably enhance internal validity, yet several assumptions are still needed to grant a causal interpretation to our findings. Broadly speaking, these concern if and to what extent our estimates are shielded from confounding bias and sample selection bias. We discuss these at length in Supplementary Sections A2 and A4, and provide evidence supporting the robustness of our approach. With this in mind, estimation involves the following linear model:

$$y_{it} = \alpha + \sum_{k=0}^{k=5} \beta_k \cdot 1[t = k] + \sum_{k=0}^{k=5} \gamma_k \cdot 1[t = k] \cdot D_i + X_{it}\theta + \phi_i + \varepsilon_{it}$$

$$(1)$$

where the dependent variable y_{it} is the log of real hourly wages for an individual i in calendar year t. Hourly wages are derived from gross monthly labour income, as reported by SOEP respondents who are currently employed with reference to the month prior to the interview. This is then divided by the amount of actual weekly working hours multiplied by 4.35 (the approximate number of weeks in a month). If actual working time is not available, we substitute for it by taking the sum of contractual working hours and overtime (Kühhirt and Ludwig, 2012). Wages are then logged and indexed at 2014 consumer prices. Person-year observations were dropped when real hourly wage values were smaller than 1 or bigger than 100, to reduce the influence of outliers on our estimates. As per Equation 1, we draw on an event-study specification to study the effect of motherhood on wages before and after a given reform (e.g. Borusyak and Jaravel, 2016; Kleven, Landais and Søgaard, 2018; Kuziemko *et al.*, 2018). In a nutshell, what we aim to retrieve is the average wage loss for a given woman, in each year after giving birth as compared to her average wage prior to giving birth. These are estimates of the *total* motherhood wage penalty after first birth. We want to further contrast these estimates

of the motherhood wage penalty across groups of women who became mothers prior to and after parental leave reform.

Each year of interview t after first childbirth is named k and a corresponding 'event-time' dummy is included in the model (i.e. $1[t = k]$). These dummies cover the post-birth period, from the first interview after the child was born (Year 0) to the sixth interview in its aftermath (Year 5). We first estimate the motherhood penalty for each year after the birth of the first child for the control group (β_k). We then take the product of these event-time dummies and our treatment dummy D_i, the latter separating women in the treated group (coded as 1) from women in the control group (the reference, coded 0). This product gives us the difference in the motherhood penalty (γ_k) between treated and controls, that is, the average effect of parental leave reform on the treated. Such effect γ_k should be regarded as an intention-to-treat estimate, given that we do not explicitly model parental leave receipt. This holds particularly for Reform 1, as benefit receipt became rather universal in the aftermath of Reform 2 (e.g. Schönberg and Ludsteck, 2014; Raute, 2019).

To grant an easier first approach to our results, we sum these latter coefficients γ_k to the respective β_k for each year after first childbirth. This sum gives us the estimated motherhood penalty for the treated, and we can readily contrast this to the motherhood penalty for the controls. Full estimates can be found in Supplementary Section A3. Two aspects of our design are worth stressing. First, we omit the 'main effect' of the treatment dummy D_i. Belonging to one or the other treatment group is a time-constant feature for each individual, and as such this feature is already encompassed by individual FEs ϕ_i. Yet product terms that involve time-constant variables and time-varying ones can be estimated. We can therefore retrieve our estimates of interest, the differences in the motherhood effect across treatment groups γ_k.

Second, year-specific sample sizes in years 0–5 vary from a minimum of 30 women to a maximum of 279 women (Supplementary Table SA1). This may raise concerns on whether our design is under-powered and with what consequences for the credibility of our estimates (e.g. Gelman and Carlin, 2014), an issue we explore further in Supplementary Section A8. Even if underpowered, we chose an event-study specification for one main reason. We follow both studies on the wages of German mothers (Ejrnæs and Kunze, 2013) and on the motherhood wage penalty more broadly (e.g. Loughran and Zissimopoulos, 2009; Fernández-Kranz, Lacuesta

and Rodrìguez-Planas, 2013; Kleven, Landais and Søgaard, 2018), who commonly suggest to disentangle short, medium, and (if possible) long run effects of motherhood on wages. As years go by since the event, estimates may vary in magnitude as well as in the amount of uncertainty that surrounds them. Simpler approaches - e.g. a single dummy for before-after first childbirth - would assume this heterogeneity away and may provide severely biased estimates of the effect of interest (for recent appraisals, see Borusyak and Jaravel, 2016; Imai and Kim, 2017; de Chaisemartin and D'Haultfœuille, 2020).

Among variables in the vector $\mathbf{X}_{it}$, we include a quadratic for age, to net out pure lifecycle effects, and dummies for region of residence. Individual FEs ϕ_i and an idiosyncratic error term ε_{it} complete our preferred specification. Robust standard errors are estimated to account for the possibility of serial correlation in the disturbance term. Finally, to investigate mechanisms, we ran the same model displayed in Equation 1 on three auxiliary outcomes. First, we track a woman's uptake of leave provisions by determining the share of months she spent on leave in the year preceding the current interview (e.g. Buligescu *et al.*, 2009). Leave shares range from 0, indicating that no time was spent on leave in the year prior to the current interview, to 1, indicating that a mother spent on leave all 12 calendar months in the previous year. Importantly, we kept focusing on first births only, meaning that leave uptake is set to 0 following a higher-order parity. We look at leave share to assess changing family-related career interruptions prior and after each reform, for both the purpose of validating our design with respect to the ample previous literature on the topic and to look into the role of human capital and signalling mechanisms in shaping motherhood wage penalties. Also, given that leave share can be observed regardless of whether a woman works or not, we use all available person-year records belonging to the women part of our sample (6,585 total person-year records for Reform 1, 9,756 total person-year records for Reform 2). This further allows the comparison of our estimates to those of prior research on the labour supply of mothers (e.g. Schönberg and Ludsteck, 2014; Bergemann and Riphahn, 2017).

For the intensive margin of labour supply, we focus on changes in women's weekly working hours. As for job continuity, implicated by the job guarantee built in parental leave provisions, our dependent variable is tenure with the current employer. For these two latter outcomes, which are observed only if a woman is in paid work at a given point in time t, we stick to person-year

records in which women were employed similar to the main models for wages.

Findings

Wage Responses

We start by examining the effects of motherhood on wages before and after Reform 1, that is, parental leave expansion in 1992. Figure 2 plots our estimates separately for those exposed to Reform 1 and their comparison group. The overarching pattern is that of a sizeable motherhood penalty, stable over time. Women in the control group, who gave birth to their first child prior to 1992, confront wage losses in the range of 14–30 per cent of their pre-birth wages in all years except Year 1. Similar, women treated by Reform 1 face a wage penalty in the range of 15–23 per cent of their pre-birth wages. Little differences are found, overall, between the two groups of women, with the exception of Year 0 and, more markedly, Year 1. In Year 0, control-group women face a penalty of around 29 per cent, while the penalty for treated women stops at around 15 per cent. Women in the treated group could be positively signalling themselves, as they forgo a larger job-guaranteed leave compared to control-group women in the same year. A similar pattern should hold in Year 1, yet we find little evidence of a penalty for women in the control group, whereas women exposed to Reform 1 report a substantial wage loss of around 23 per cent of their pre-birth wages. Evidence for positive signalling among early returners is therefore mixed, in the aftermath of Reform 1.

Figure 3 plots the wage effects of motherhood for women affected by Reform 2, the latest major reform of German parental leave (2007). We compare these estimates to those for a comparison group of women who became mothers prior to Reform 2. For this latter group, wage penalties are only detected in the medium term, in Year 2 after the birth of a child and onwards. Such wage effects for control-group women range from around 7 per cent of pre-birth wages in Year 2 to around 21 per cent in Year 4. For women giving birth for the first time after Reform 2, on the other hand, wage losses cannot be detected, also in the long run. In Years 4 and 5 after first birth, in particular, we detect an improvement of around 15–17 points in the wage effects of motherhood in favour of the treated.

Our main findings thus point, first, to persistence in the aftermath of Reform 1. Such persistence may suggest that the negative wage effects of human capital loss and the positive wage effects of job
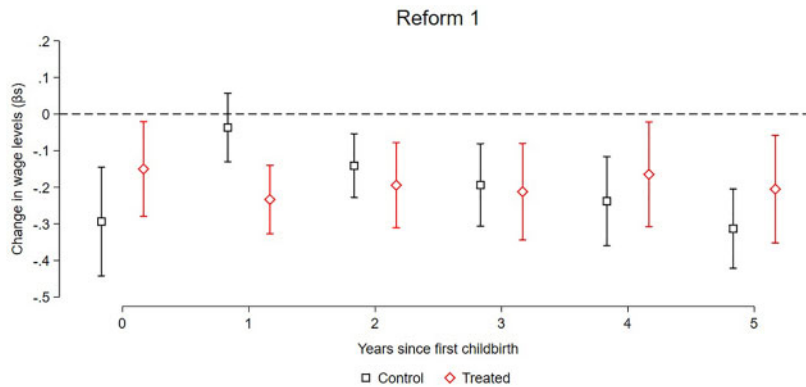
Figure 2. FE estimates (95 per cent confidence intervals) of the motherhood wage penalty before and after Reform 1, across treatment groups, and relative to event time (SOEP 1985–1998)
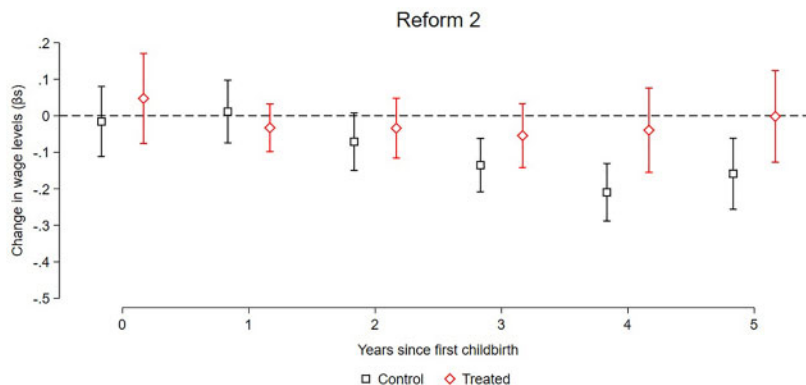

Figure 3. FE estimates (95 per cent confidence intervals) of the motherhood wage penalty before and after Reform 2, across treatment groups, and relative to event time (SOEP 1999–2014)

continuity, both spurred by the expanded job guarantee introduced by Reform 1, cancel each other out on average. Evidence for Years 0 and 1 provides conflicting indications on 'positive' signalling instead. Mothers contributing to those estimates are 'early returners', and even more so among the treated, given the 36-month job guarantee instituted in 1992. Yet, early returners among the treated face smaller penalties than their counterparts in Year 0 ($p = 0.10$) and harsher penalties in Year 1 ($p < 0.01$), offering mixed evidence, at best, for positive signalling.

As for Reform 2, we can no longer detect a large motherhood wage penalty for women under the new parental leave benefit instituted in 2007. Nevertheless, estimates across treated and controls do not statistically differ from each other in the first years after childbirth, but only later on. Coherent with a human capital argument, it could be that treated women reap the benefits

of shorter career breaks per effect of Reform 2. Differently, in line with negative signalling arguments, we would have expected a larger penalty in the years long after first childbirth, as these latter comprise 'late returners' that now defect the norm of reprising work after 12 months of benefit receipt.[3]

Labour Supply Outcomes

To shed light on the wage effects of parental leave reform, we examine time spent on live, weekly working hours, and job tenure (Figure 4, see also Supplementary Section A3). We estimate the same model specification used for hourly wages.

Results in the upper panel of Figure 4 are largely in line with our expectations. We find that the expansion of the job guarantee in 1992 lead to opposite effects on leave uptake and tenure. The share of months spent on
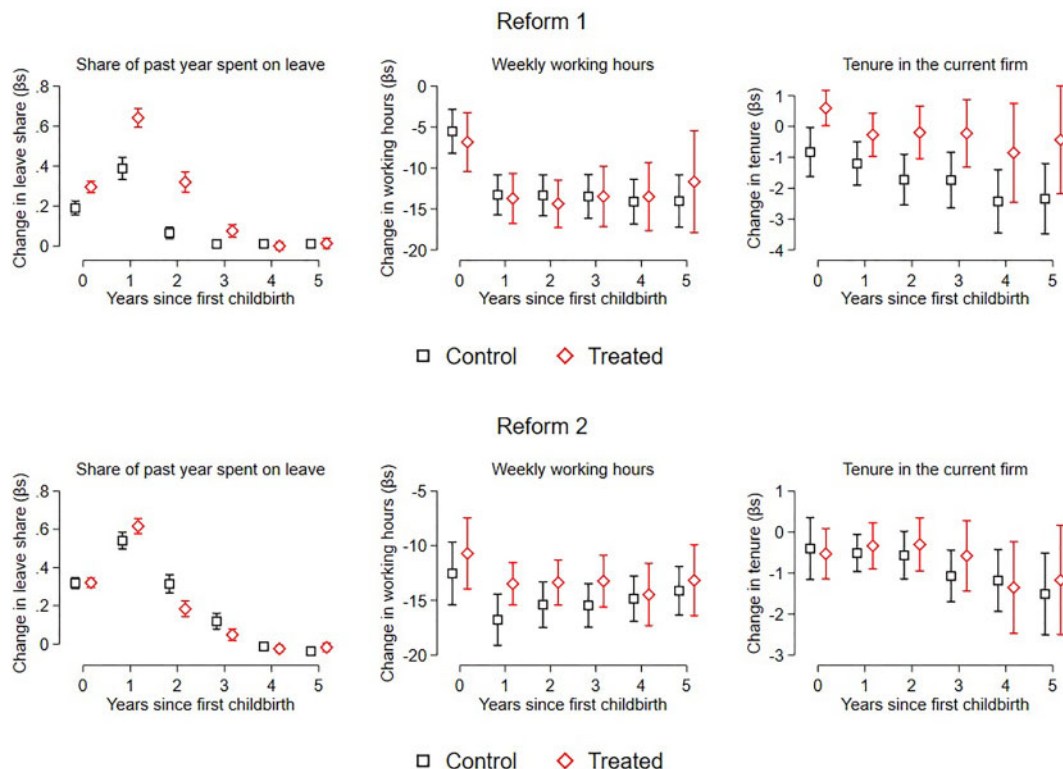
Figure 4. FE estimates (95 per cent confidence intervals) of motherhood ffects on (i) leave share, (ii) working hours, and (iii) tenure (SOEP 1985–2014)

leave, as displayed in the top-left corner, increased for women treated by Reform 1 as compared to the control group. The increase for the treated is remarkable, with women more than doubling their share of time spent on leave in some years (e.g. Year 2; full estimates in Supplementary Table SA3). These effects point to a more prolonged use of parental leave after Reform 1, in line with what has been widely documented in previous studies (Schönberg and Ludsteck, 2014; Ziefle and Gangl, 2014).

At the same time, looking at the top-right corner of Figure 4, mothers subject to the expanded leave post-1992 appear to be better off in terms of job tenure with their current firm. Prior to 1992, in fact, mothers faced a 'tenure penalty', possibly contributing to their wage losses. If these women faced losses of around 1–2 years of tenure per year after childbirth, we do not detect statistically similar penalties for women treated by Reform 1. The effect of the reform are thus relatively large when it comes to tenure (see Supplementary Table SA5), whereas for working hours we find evidence of overall stability over time. As per Figure 4, we find that

mothers, both exposed and unexposed to Reform 1, reduce their working hours of around 5–6 hours if returning already in Year 0, and of around 11–14 hours in the following years (see Supplementary Table SA4).

Our findings at the bottom of Figure 4 pertain to Reform 2. For leave share, in the bottom-left corner, changes across groups are mixed. We find little evidence of change in Year 0 and a slight increase in the share of time spent on leave for the treated in Year 1. In Years 2 and 3, though, women affected by Reform 2 are found to *reduce* their time spent on leave. This is consistent with previous research, highlighting a fuller use of leave in the first 12 months covered by the new parental leave benefit and a higher chance of return to paid work after its expiration (Bergemann and Riphahn, 2017; Kluve and Schmitz, 2018).

Opposite to Reform 1, we further find somewhat of a relative improvement in the working hours of mothers, but not in their post-birth tenure after Reform 2 (see also Supplementary Tables SA8 and SA9). Consistent with past evidence (Kluve and Schmitz, 2018), we find that mothers affected by the 2007 leave benefit log more hours of work

than their counterparts, especially in Year 1. Still, employed mothers reduce their working hours substantially even under the new scheme, with no change across treatment groups detected in the medium term.

Other than resemblance to previous evidence on maternal labour supply, our findings in this section shed light on the wage responses to parental leave reform. First, we suggest that prolonged time on leave combined with job-protection rights ultimately did not alter the wage penalty experienced by German women in the 1990s (Gangl and Ziefle, 2009; Schönberg and Ludsteck, 2014). Wage responses to long work interruptions are typically negative, whilst preserving job tenure might mitigate wage loss. The two forces might have thus cancelled each other out, resulting in the persistence of wage penalties for mothers in the aftermath of Reform 1. Second, the increase in time spent on leave in the immediate aftermath of childbirth might further explain why substantial wage penalties were found for the treated by the 1992 reform also in Year 1. Conversely, while time spent on leave increased also long after childbirth, due to the expanded job guarantee, wage responses are not consistent with negative signalling further dampening down the wage of mothers in the medium term. Substantial yet unchanging reductions in working hours, last, might have underpinned wage penalties both before and after Reform 1. We interpret this as evidence that a change in the relative importance of work versus family, spurred by Reform 1 (Gangl and Ziefle, 2015), did not translate into changes in work hours among employed women. Effort, as proxied by work hours, had little consequences on the *relative* wage prospects of women treated by Reform 1.

Moving on to Reform 2, evidence in Figure 4 might account for the improvement in mothers' wage rates via shorter time spent on leave in the medium run and slight increases in working hours. The wage effects of these labour supply changes are more consistent with human capital accounts than with signalling. If in line with signalling, indeed, we would have expected stronger improvements in the wage penalty for those returning early, and more evidence of a wage penalty when including 'late returners' later on.

Robustness and Sensitivity Analyses

We performed a number of additional analyses to examine the robustness of our estimates with respect to model assumptions, as well as to probe the sensitivity of our estimates to specification and sample selection choices. These analyses are reported and discussed more at length in our Supplementary.

A first assumption that might be violated in our setting is the absence of contemporaneous shocks to the wages and labour supply of mothers. Relying on relatively small sample numbers, we had to select long time periods before and after a given reform, thereby increasing the chance of co-occurring policy changes or period effects. In Supplementary Section A4, we focus on concomitant childcare reforms. Childcare expansion in 1995, for one, might have improved the employment rates of women exposed to Reform 1 in our study (Bauernschuster and Schlotter, 2015). Restricting our definition of treated and comparison groups to women unlikely affected by this childcare reform does not change our substantial conclusions with respect to Reform 1. As for later efforts to expand childcare coverage, in 2005 and 2008, we could not perform similar sample restrictions and retrieve statistically reliable estimates. Recent appraisals of this new wave of childcare expansion found larger positive labour supply effects for the birth of a second child rather than the first (Zoch and Hondralis, 2017). Our estimates, also in consideration of previous research, centre around first parities (see also Supplementary Section A6).

We also probed our models to the inclusion of year FEs and of regional (East-West) unemployment rates to account for shocks that may similarly hit the wages of treated and control-group women. Two-way fixed effects, with individual and year fixed effects, are increasingly understood as biased when it comes to recovering the treatment effect (of an event) of interest, especially in the absence of a control group that never experiences the event/treatment (e.g. Borusyak and Jaravel, 2016; Imai and Kim, 2017; de Chaisemartin and D'Haultfœuille, 2020). In the specification reported in Supplementary Section A5, however, we augment our samples by including childless women. This group, while not experiencing the birth of a child by definition, may aid the identification of other variables in our models, including period effects. Our main findings are largely unchanged.

In Supplementary Sections A6–A8 we explore whether our findings are sensitive, respectively, to (i) censoring observations after higher-order parities rather than selecting only control-group women who had children prior to a given reform, (ii) aggregating multiple event-time periods rather than using year-by-year event-time dummies, and (iii) adding a boost sample available for the evaluation of Reform 2. All these additional analyses aim at increasing statistical power with respect to our main analyses, possibly accepting more bias in exchange for more variance. Results in these sections reinforce our main findings, pointing to stability after

Reform 1 and a relative improvement in the wage effects of motherhood following Reform 2. In this latter respect, the lower bounds of our confidence intervals are typically consistent with penalties of around 5–10 per cent for those women treated by Reform 2.[4] Hence, even if wage estimates for treated women do not reach conventional levels of statistical significance, we do *not* suggest that the motherhood penalty has 'vanished' in contemporary Germany.

Discussion and Conclusions

We have examined how motherhood wage penalties may have responded to two distinct parental leave reforms in Germany, with a focus on first-time mothers. From 1992 onwards, a maternalist leave scheme combined long periods of benefit receipt and even longer periods protected by job-guarantee rights. These provisions further delayed German women's re-entry into the labour market. Yet we find that longer time spent on leave coupled with improved tenure with the current employer resulted in an overall stable penalty, before and after 1992. Penalties are larger than suggested in previous studies, up to 20–30 per cent of pre-birth wages. In contrast, since 2007, German parental leave has featured a shorter, earnings-related benefit, intact job-guarantee rights, and a bonus of two months of benefit receipt usually taken up by fathers. Per effect of this reform, we can no longer detect (large) wage penalties for new mothers, who now concentrate their leave taking in their first years after childbirth and also work longer hours upon re-entry.

Motherhood wage penalties can be large, justifying intense academic and societal interest in their roots. We suggest that an event-study design might provide more faithful estimates of the size of such penalties, as compared to studies focusing on 'one-shot' losses per child (cf. Cukrowska-Torzewska and Matysiak, 2020; de Linde Leonard and Stanley, 2020). A first limit of our analyses, though, is that we rely on relatively small sample sizes. As a result, our analyses only examine medium-term penalties and also assume homogeneity in maternal wage effects *within* each group of 'control' or 'treated' women. Future studies adopting an event-study framework could shed light on long-term penalties and on further sources of heterogeneity across women. With respect to the German context, more specifically, studies should particularly assess heterogeneity depending on education level/income. Previous research has shown that highly educated/high-income women were the 'winners' out of Reform 2 (Kluve and Schmitz, 2018;

Frodermann, Wrohlich and Zucco, 2020), yet we could not go beyond 'average' penalties in this study.

Nevertheless, we show that child penalties can change over time, and parental leave policies may play a role in this regard. If granting our estimates a causal interpretation, the right mix to reduce or even undo motherhood wage penalties may consist of a relatively short leave, with job-protection rights, and incentives to share leave uptake so that women can work longer hours upon re-entry. Our findings thus echo and update previous literature on best practices when designing leave schemes (e.g. Waldfogel, 1998; Baker and Milligan, 2008). Results also inform the larger debate on the consequences of work–family policies, often found to foster women's employment chances at the expense of their wages and career attainment (Ruhm, 1998; Mandel and Semyonov, 2006; Arulampalam, Booth and Bryan, 2007; Mandel, 2012; Aisenbrey and Fasang, 2017). In line with recent studies (Mun and Jung, 2018), we suggest that perverse effects largely depend on policy design. Examining the 'nuts and bolts' of any given policy in a single country (e.g. Andersen, 2018) could nicely complement comparative research on work–family bundles.

This in-depth focus can shed light on mechanisms triggered by different components of a given policy, mechanisms that can have zero-sum consequences (e.g. Reform 1). By and large, wages responses in our study are compatible with a human capital account. We did not detect sharp discontinuities as predicted by signalling theory, but future studies could improve on our design in this respect building on previous research (e.g. Evertsson, 2016). We also overlooked demand-sided mechanisms, yet parental leave provisions may contribute to statistical discrimination. Risk-averse employers may fear the costs of maternal work interruptions, and pull down (up) the wages of all women in contexts with more (less) extensive leave regulations (e.g. Jessen, Jessen and Kluve, 2019). These are fruitful directions, we believe, for further inquiry.

In conclusion, large motherhood wage penalties in Germany persisted in recent decades, up until a substantial improvement in the late 2000s. Both early stability and later improvement are here credited to the design of parental leave policy and to the behavioural responses parental leave provisions might have triggered. Our contribution suggests that institutions may critically shape gender economic inequality, at times maintaining it and at times reducing it.

Supplementary Data

Supplementary data are available at *ESR* online.

Notes

1 The use of terms such as 'control group' in observational settings like ours is well-established in the relevant literature (e.g. Shadish, Cook and Campbell, 2002; Francesconi and Van der Klaauw, 2007; Gangl and Ziefle, 2015). In such observational studies, however, 'treated' and 'control' conditions are not randomly assigned by the researcher, but rather identified by the researcher as one in which members of a population have/have not been exposed to some policy or intervention. As such, we will use 'control group' and 'comparison group' interchangeably, to emphasise the absence of random assignment.

2 Also, using these sample selection criteria and due to SOEP complex design (Goebel *et al.*, 2018), East German women make up around 12 per cent of person-year records for women treated by Reform 1 but only 3.5 per cent from the respective control group. This imbalance is, however, inconsequential for our main findings.

3 Notably though, in our estimates we simply disentangle wage responses separately by each year since first childbirth, not depending on the timing of a woman's re-entry more specifically. This means that, while in Year 0 our estimation sample comprises only women who have returned to work by that time, in Year 1 women in the estimation sample will comprise returners in Year 0 and 1, in Year 2 returners in Years 0, 1, and 2, and so forth. In other words, we are bound to mix women with different timings of re-entry for our wage estimates in the medium and long run. We do not conduct separate analyses depending on the timing of a mother's re-entry for two reasons, namely sample size considerations and due to the endogeneity of such timing to the design of parental leave policy itself.

4 We computed 90 per cent confidence intervals for our wage estimates. Checking whether 90 per cent confidence intervals lie entirely above (below) a certain value would lead us to reject the hypothesis of a meaningful effect, or smallest effect of interest (Rainey, 2014; Bernardi, Chakhaia and Leopold, 2017; Lakens, Scheel and Isager, 2018). Considering the wage effects of motherhood for those women treated by Reform 2, we typically cannot exclude penalties of around 5 per cent in the first years after childbirth, and 10 per cent later on. We consider penalties in this range to be of non-negligible size, and our estimates are thus compatible with such meaningful effects.

Acknowledgements

References

Adda, J., Dustmann, C. and Stevens, K. (2017). The career costs of children. *Journal of Political Economy*, **125**, 293–337.

Aisenbrey, S. and Fasang, A. (2017). The interplay of work and family trajectories over the life course: Germany and the United States in comparison. *American Journal of Sociology*, **122**, 1448–1484.

Albrecht, J., Thoursie, P. S. and Vroman, S. (2015). Parental leave and the glass ceiling in Sweden. *Research in Labor Economics*, **41**, 89–114.

Albrecht, J. W. *et al.* (1999). Career interruptions and subsequent earnings: a reexamination using Swedish data. *Journal of Human Resources*, **34**, 294–311.

Andersen, S. H. (2018). Paternity leave and the motherhood penalty: new causal evidence. *Journal of Marriage and Family*, **80**, 1125–1143.

Anderson, D. J., Binder, M. and Krause, K. (2003). The motherhood wage penalty revisited: experience, heterogeneity, work effort, and work-schedule flexibility. *Industrial & Labor Relations Review*, **56**, 273–294.

Arntz, M., Dlugosz, S. and Wilke, R. A. (2017). The sorting of female careers after first birth: a competing risks analysis of maternity leave duration. *Oxford Bulletin of Economics and Statistics*, **79**, 689–716.

Arulampalam, W., Booth, A. L. and Bryan, M. L. (2007). Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *ILR Review*, **60**, 163–186.

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

Bardasi, E. and Gornick, J. C. (2008). Working for less? Women's part-time wage penalties across countries. *Feminist Economics*, **14**, 37–72.

Bauernschuster, S. and Schlotter, M. (2015). Public child care and mothers' labor supply—evidence from two quasi-experiments. *Journal of Public Economics*, **123**, 1–16.

Baum, C. L. and Ruhm, C. J. (2016). The effects of paid family leave in California on labor market outcomes. *Journal of Policy Analysis and Management*, **35**, 333–356.

Bergemann, A. and Riphahn, R. T. (2017). Maternal employment effects of paid parental leave. *SOEPpapers on Multidisciplinary Panel Data Research*, **900–2017**, 1–83.

Bernardi, F., Chakhaia, L. and Leopold, L. (2017). 'Sing me a song with social significance': the (mis) use of statistical significance testing in European sociological research. *European Sociological Review*, **33**, 1–15.

Borusyak, K. and Jaravel, X. (2016). *Revisiting Event Study Designs*. Working Paper, available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2826228 [accessed 10 October 2020].

Budig, M. J., Misra, J. and Boeckmann, I. (2016). Work–family policy trade-offs for mothers? Unpacking the cross-national variation in motherhood earnings penalties. *Work and Occupations*, 43, 119–177.

Buhr, P. and Huinink, J. (2015). The German low fertility: how we got there and what we can expect for the future. *European Sociological Review*, 31, 197–210.

Buligescu, B. *et al.* (2009). Panel estimates of the wage penalty for maternal leave. *Oxford Economic Papers*, 61, i35–55.

Bünning, M. (2015). What happens after the 'Daddy Months'? Fathers' involvement in paid work, childcare, and housework after taking parental leave in Germany. *European Sociological Review*, 31, 738–748.

Cukrowska-Torzewska, E. and Matysiak, A. (2020). The motherhood wage penalty: a meta-analysis. *Social Science Research*, 88–89, 102416

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

Davies, R. and Pierre, G. (2005). The family gap in pay in Europe: a cross-country study. *Labour Economics*, 12, 469–486.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110, 2964–2996.

de Linde Leonard, M. and Stanley, T. (2020). The wages of mothers' labor: a meta-regression analysis. *Journal of Marriage and Family*, 1–19 10.1111/jomf.12693

Dieckhoff, M. *et al.* (2016). A stalled revolution? What can we learn from women's drop-out to part-time jobs: a comparative analysis of Germany and the UK. *Research in Social Stratification and Mobility*, 46, 129–140.

Drasch, K. (2012). Educational attainment and family-related employment interruptions in Germany: do changing institutional settings matter? *European Sociological Review*, 29, 981–995.

Ejrnæs, M. and Kunze, A. (2013). Work and wage dynamics around childbirth. *The Scandinavian Journal of Economics*, 115, 856–877.

Evertsson, M. (2016). Parental leave and careers: women's and men's wages after parental leave in Sweden. *Advances in Life Course Research*, 29, 26–40.

Fernández-Kranz, D., Lacuesta, A. and Rodrìguez-Planas, N. (2013). The motherhood earnings dip: evidence from administrative records. *Journal of Human Resources*, 48, 169–197.

Fernández-Kranz, D., Paul, M. and Rodrìguez-Planas, N. (2015). Part-time work, fixed-term contracts, and the returns to experience. *Oxford Bulletin of Economics and Statistics*, 77, 512–541.

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

Francesconi, M. and Van der Klaauw, W. (2007). The socioeconomic consequences of "in-work" benefit reform for British lone mothers. *Journal of Human Resources*, 42, 1–31.

Frodermann, C., Wrohlich, K. and Zucco, A. (2020). *Parental Leave Reform and Long-run Earnings of Mothers*. IZA Discussion Paper No. 12935. Bonn, Germany: IZA.

Gangl, M. and Ziefle, A. (2009). Motherhood, labor force behavior, and women's careers: an empirical assessment of the wage penalty for motherhood in Britain, Germany, and the United States. *Demography*, 46, 341–369.

Gangl, M. and Ziefle, A. (2015). The making of a good woman: extended parental leave entitlements and mothers' work commitment in Germany. *American Journal of Sociology*, 121, 511–563.

Gelman, A. and Carlin, J. (2014). Beyond power calculations: assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science*, 9, 641–651.

German Institute for Economic Research (DIW), Berlin. (2016). *Socio-Economic Panel (SOEP), Data for Years 1984–2014, Version 31.1*. SOEP, doi:10.5684/soep.v31. German Institute for Economic Research (DIW), Berlin.

Goebel, J. (2017). *SOEP-Core v32—Documentation on Biography and Life History Data*. SOEP Survey Papers No. 418.

Goebel, J. *et al.* (2018). The german socio-economic panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik*, 239, 2, 345–360..

Gregg, P., Gutíerrez-Domènech, M. and Waldfogel, J. (2007). The employment of married mothers in Great Britain, 1974–2000. *Economica*, 74, 842–864.

Gupta, N. D. and Smith, N. (2002). Children and career interruptions: the family gap in Denmark. *Economica*, 69, 609–629.

Hainmueller, J. (2012). Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20, 25–46.

Imai, K. and Kim, I. S. (2017). *When Should We Use Fixed Effects Regression Models for Causal Inference with Longitudinal Data?* Working Paper, available from: https://imai.princeton.edu/research/files/FEmatch.pdf.

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

Joseph, O. *et al.* (2013). The economic impact of taking short parental leave: evaluation of a French reform. *Labour Economics*, 25, 63–75.

Kleven, H., Landais, C. and Søgaard, J. E. (2018). *Children and Gender Inequality: Evidence from Denmark*. NBER Working Paper No. 24219.

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

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

Kühhirt, M. and Ludwig, V. (2012). Domestic work and the wage penalty for motherhood in West Germany. *Journal of Marriage and Family*, 74, 186–200.

Kuziemko, I. *et al.* (2018). *The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?* NBER Working Paper No. 24740.

Lakens, D., Scheel, A. M. and Isager, P. M. (2018). Equivalence testing for psychological research: a tutorial. *Advances in Methods and Practices in Psychological Science*, **1**, 259–269.

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

Lersch, P. M., Jacob, M. and Hank, K. (2017). Parenthood, gender, and personal wealth. *European Sociological Review*, **33**, 410.

Loughran, D. S. and Zissimopoulos, J. M. (2009). Why wait? The effect of marriage and childbearing on the wages of men and women. *Journal of Human Resources*, **44**, 326–349.

Mandel, H. (2012). Winners and losers: the consequences of welfare state policies for gender wage inequality. *European Sociological Review*, **28**, 241–262.

Mandel, H. and Semyonov, M. (2006). A welfare state paradox: state Interventions and women's employment opportunities in 22 countries. *American Journal of Sociology*, **111**, 1910–1949.

Mun, E. and Jung, J. (2018). Policy generosity, employer heterogeneity, and women's employment opportunities: the welfare state paradox reexamined. *American Sociological Review*, **83**, 508–535.

Musick, K., Bea, M. D. and Gonalons-Pons, P. (2019). His and her earnings following parenthood in the United States, Germany, and the United Kingdom. *American Sociological Review*, **85**, 639–674.

OECD. (2017). *The Pursuit of Gender Equality: An Uphill Battle*. Paris: OECD Publishing.

Olivetti, C. and Petrongolo, B. (2008). Unequal pay or unequal employment? A crosscountry analysis of gender gaps. *Journal of Labor Economics*, **26**, 621–654.

Paul, M. (2016). Is there a causal effect of working part-time on current and future wages? *The Scandinavian Journal of Economics*, **118**, 494–523.

Petersen, T., Penner, A. M. and Høgsnes, G. (2014). From motherhood penalties to husband premia: the new challenge for gender equality and family policy, lessons from Norway. *American Journal of Sociology*, **119**, 1434–1472.

Ponthieux, S. and Meurs, D. (2015). Gender inequality. In Atkinson A. B. and Bourdignon F. (Eds.), *Handbook of Income Distribution*. Vol. 2A. North Holland: Elsevier.

Rainey, C. (2014). Arguing for a negligible effect. *American Journal of Political Science*, **58**, 1083–1091.

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.

Ray, R., Gornick, J. C. and Schmitt, J. (2010). Who cares? Assessing generosity and gender equality in parental leave policy designs in 21 countries. *Journal of European Social Policy*, **20**, 196–216.

Ruhm, C. J. (1998). The economic consequences of parental leave mandates: lessons from Europe. *The Quarterly Journal of Economics*, **113**, 285–317.

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

Shadish, W. R., Cook, T. D. and Campbell, D. T. (2002). *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.

Tamm, M. (2018). *Fathers' Parental Leave-taking, Childcare Involvement and Mothers' Labor Market Participation*. IZA Discussion Paper No. 11873. Bonn, Germany: IZA.

Trappe, H., Pollmann-Schult, M. and Schmitt, C. (2015). The rise and decline of the male breadwinner model: institutional underpinnings and future expectations. *European Sociological Review*, **31**, 230–242.

Waldfogel, J. (1998). The family gap for young women in the United States and Britain: can maternity leave make a difference? *Journal of Labor Economics*, **16**, 505–545.

Zeldow, B. and Hatfield, L. A. (2019). Confounding and Regression adjustment in difference-in-differences. arXiv preprint: 1911.12185.

Zhang, X. (2010). Can motherhood earnings losses be ever regained? Evidence from Canada. *Journal of Family Issues*, **31**, 1671–1688.

Ziefle, A. and Gangl, M. (2014). Do women respond to changes in family policy? A quasi-experimental study of the duration of mothers' employment interruptions in Germany. *European Sociological Review*, **30**, 562–581.

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

Gabriele Mari, PhD, is a postdoctoral researcher at Erasmus University Rotterdam. His research interests include social policy, families, and causal inference. His work has been published in peer-reviewed journals such as Journal of Marriage and Family and Research in Social Stratification and Mobility.

Giorgio Cutuli, PhD, is currently a research fellow at the Department of Sociology and Social Research, University of Trento, Italy. His main research interests include welfare and labour market dynamics, socioeconomic inequality, and quantitative longitudinal analysis. His work has been published in peer-reviewed journals such as Socio-Economic Review, European Sociological Review, and Research in Social Stratification and Mobility.