

SOEPpapers

on Multidisciplinary Panel Data Research

SOEP – The German Socio-Economic Panel Study at DIW Berlin

1025-2019

Do Parental Leaves Make the Motherhood Wage Penalty Worse? Assessing Two Decades of German Reforms

Gabriele Mari and Giorgio Cutuli

SOEPPapers on Multidisciplinary Panel Data Research at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPPapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPPapers are available at
<http://www.diw.de/soeppapers>

Editors:

Jan **Goebel** (Spatial Economics)
Stefan **Liebig** (Sociology)
David **Richter** (Psychology)
Carsten **Schröder** (Public Economics)
Jürgen **Schupp** (Sociology)

Conchita **D'Ambrosio** (Public Economics, DIW Research Fellow)
Denis **Gerstorff** (Psychology, DIW Research Fellow)
Elke **Holst** (Gender Studies, DIW Research Director)
Martin **Kroh** (Political Science, Survey Methodology)
Jörg-Peter **Schräpler** (Survey Methodology, DIW Research Fellow)
Thomas **Siedler** (Empirical Economics, DIW Research Fellow)
C. Katharina **Spieß** (Education and Family Economics)
Gert G. **Wagner** (Social Sciences)

ISSN: 1864-6689 (online)

German Socio-Economic Panel (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: soeppapers@diw.de



Do Parental Leaves Make the Motherhood Wage Penalty Worse? Assessing Two Decades of German Reforms

Gabriele Mari* Giorgio Cutuli†

Abstract

Women-friendly policies may have perverse effects on the wages of employed women and mothers in particular. Yet few have addressed the causal impact of such policies and the mechanisms they might trigger at the individual level to produce such wage responses. We assess if and how two decades of reforms of parental leave schemes in Germany have shaped changes in the motherhood wage penalty over time. We compare two sweeps of reforms inspired by opposite principles, one allowing for longer periods out of paid work, the other prompting quicker re-entry in the labour market. We deploy panel data (SOEP 1985-2014) and a within-person difference-in-differences design.

Motherhood wage penalties were found to be harsher than previously assessed in 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. Conversely, we can no longer detect motherhood wage penalties for women affected by the later reform. Shorter career breaks and increased work hours may have benefited new mothers in the late 2000s, leading to a substantial improvement in their wage prospects.

Keywords: parental leave, motherhood wage penalty, difference-in-difference, gender inequality

JEL: D13, J13, J16, J31

*mari@essb.eur.nl, Erasmus University Rotterdam.

†g.cutuli@unitn.it, University of Trento.

Note: Access to the SOEP data was kindly granted by the German Institute for Economic Research (DIW), Berlin, 2016. The data collector bears no responsibility for the analyses and interpretations provided in this paper. The authors wish to thank Lynn Prince Cooke, as well as participants at the Understanding Society Scientific Conference (University of Essex, 2017), the 2017 ECSR conference (Bocconi University, Milan), and in internal seminars at the University of Trento and at the University of Bath, for their useful comments on previous versions of this paper. All remaining errors are the authors' own.

Employed mothers typically face wage losses when returning to paid work after childbirth, to the point that such a motherhood wage penalty has become a key component of the gender pay gap in labour markets (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 et al., 2017). Work interruptions are in turn influenced by family-leave policies and reforms to those policies over time (Gregg et al., 2007; Lalive and Zweimüller, 2009; Ziefle and Gangl, 2014; Baum and Ruhm, 2016), yet few studies have looked into the corresponding changes, if any, in the motherhood wage penalty.

Prolonged absence from work granted by statutory leaves may spur human capital losses (Gupta and Smith, 2002), effort re-allocation (Gangl and Ziefle, 2015), or signal low commitment to current and future employers (Albrecht et al., 1999; Evertsson, 2016), all to the detriment of women’s wages. At the same time, family leaves may offset the motherhood wage penalty by granting job protection and thereby job continuity (Waldfogel, 1998; Zhang, 2010) or by encouraging a more equal division of childcare within couples (Petersen et al., 2014; Andersen, 2018).

Leave arrangements, with their duration, job protection rights, and so forth, may thus contribute to shape the motherhood wage penalty. In this paper, we aim to establish a link between reforms of parental leave mandates and changes to the motherhood wage penalty over time. We focus on Germany and rely on a within-person difference-in-differences (DiD) design and on long-running panel data (SOEP 1985-2014). Parental leaves in Germany have undergone substantial change in the last decades. A first round of reforms, culminated in 1992-1993, gradually extended the duration of both benefits and job protection under the parental leave scheme. A second sweep in the 2000s prompted mothers to quickly return to the workforce, while also introducing a take-it-or-leave-it quota to encourage leave uptake among fathers. Germany thus offers an exemplary case study for it once encompassed a maternalist leave scheme, *de facto* reserved to women and among the longest and most generous in high-income countries, and has now shifted to flexible and shared provisions similar to those of Scandinavian countries (Ray et al., 2010).

Our contribution is threefold. First, by showing how family policies shape the motherhood

wage penalty in a single country, our findings complement previous comparative work. Studies have mainly inferred the importance of the institutional setting by comparing the magnitude of the penalty across countries (e.g. Davies and Pierre, 2005; Gangl and Ziefle, 2009). Differently, we examine how parental leave mandates have causally affected the motherhood wage penalty over time in Germany, and which mechanisms have been triggered at the individual level to produce such wage responses.

Our second aim is to contribute to the broader debate on women-friendly family policies. The consensus in the literature reads that such policies may help women maintain their footing in paid work at the cost of lower wages and inferior career attainment (Ruhm, 1998; Albrecht et al., 2003; Mandel and Semyonov, 2005, 2006; Arulampalam et al., 2007; Mandel, 2012; Aisenbrey and Fasang, 2017). Most of these studies have featured cross-national comparisons, often using cross-sectional data or, when taking advantage of panel data, without explicitly testing for the impact of family policies. In the following, we thus aim to contribute with a within-country, longitudinal account of how the design of family policies may influence the motherhood wage penalty.

Third and last, we add here to the rich body of evidence on two decades of parental leave reform in Germany. The effects of changing parental leave arrangements on the labour supply of German women have been extensively examined (Schönberg and Ludsteck, 2014; Ziefle and Gangl, 2014; Fitzenberger et al., 2013; Bergemann and Riphahn, 2017; Kluve and Schmitz, 2018). Beyond labour supply, the impact of parental leave reform has been assessed with respect to fertility behaviour (Cygan-Rehm, 2016), breastfeeding practices (Kottwitz et al., 2016), child development (Huebener et al., 2018), and father's involvement in childcare and housework (Kluve and Tamm, 2013; Tamm, 2018). Evidence on the wage responses to parental leave reform is relatively under-developed instead (cf. Schönberg and Ludsteck, 2014). This omission is particularly relevant when considering Germany, a country where both wage penalties for mothers and the gender wage gap more broadly are among the harshest in international comparison (e.g. Gangl and Ziefle, 2009; Olivetti and Petrongolo, 2008). These gender gaps arising in the labour market arguably raise the opportunity costs of parenthood amid Germany's fertility decline (Buhr and Huinink, 2015) and hamper women's accumulation of personal wealth and pension income

later in life (Lersch et al., 2017; OECD, 2017: 173). The consequences of parental leave reform on the motherhood wage penalty may thus be far-reaching, making Germany a compelling avenue for the study of gender economic inequality more at large.

1. Background

1.1. Parental leave reforms and women's return to paid work

Maternity leave in Germany has long covered a period of 14 weeks, six weeks before and eight weeks after delivery, with full income replacement and job protection. Parental leave, on the other hand, has been redesigned by multiple reforms touching both benefits and the job guarantee, that is, the right to return to a comparable job with the pre-birth employer.

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, the parental leave benefit also changed, switching from earnings-related to a mix of flat-rate and means-tested payments. It became available to all mothers regardless of pre-birth employment status (1986), for up to 18 months in 1992, then raised to 24 months the following year.

In 2001 a monetary incentive for shorter leaves granted a more generous benefit to mothers returning to paid work after 12 rather than 24 months of benefit. 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 if each parent takes at least two months. Benefits are now earnings-related once again, with a replacement rate of 67% 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/90s broadened benefit receipt and extended the job guarantee period 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).

Introducing or expanding paid leaves leads to sharp reductions in women’s employment probability right after childbirth (e.g. Gregg et al., 2007; Lalive and Zweimüller, 2009). In the long run, however, paid leaves 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 neatly shaped 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 the job guarantee. 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 et al., 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 et al., 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 et al., 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 et al., 2017).

In the midst of and contributing to rising female labour market participation, parental leave reforms in the 2000s lead to a reversal in mothers’ behaviour in the first years after childbirth. While evidence is mixed on the role of the 2001 reform (cfr. Fitzenberger et al., 2013; Ziefle and Gangl, 2014), the more radical overhaul of the benefit in 2007 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 the

entitlement period – coupled though with higher re-employment chances and longer working hours after the 12-month mark, when the benefit expires (Kluve and Tamm, 2013; Ziefle and Gangl, 2014; Bergemann and Riphahn, 2017; Kluve and Schmitz, 2018). Still, long-run maternal labour force participation was little affected 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).

1.2. Wage responses to parental leave reform

Overall, parental leave reform triggered longer short-term dips in the labour supply of mothers, and these dips decisively reduced only in the late 2000s. Selective return and part-time working remained common, among new mothers, all along. The extension of the job guarantee period 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 (Ziefle and Gangl, 2014; Kluve and Tamm, 2013; Bergemann and Riphahn, 2017; see also Fitzenberger et al., 2013; Gangl and Ziefle, 2015). In the remainder, we will focus on these two reform moments. The effects of these policies spilled over into the wages of mothers through three main channels: human capital loss during leave periods, the re-allocation of effort/commitment from paid work to the household, and adverse signalling to employers.

Following human capital theory, first, prolonged time off may result in forgone experience¹ and tenure, missed opportunities for training and promotion, and the depreciation of one’s extant stock of skills. This kind of skill atrophy will be reflected in lower wage offers for mothers returning to paid work after childbirth (Gupta and Smith, 2002; Anderson et al., 2003; Adda et al., 2017). Expanding the duration of parental leaves might exacerbate motherhood wage penalties by extending short-term career breaks and, thereby, leading to human capital loss. At the same time, job guarantee rights also boost job continuity (Baker and Milligan, 2008; Arntz et al., 2017). If preserving their pre-birth employer and thus firm-specific human capital, mothers could effectively offset wage losses (Waldfogel, 1998; Zhang, 2010; Fernández-Kranz et al., 2013). Whether the 1992 reform propelled human capital loss is therefore unclear, since it doubled the time on leave available to

mothers, but it did so while also providing job guarantee rights for the whole leave period of 36 months. Differently, after 2007, opting for a shorter benefit duration should have mitigated skill atrophy and its costs by prompting mothers to reprise paid work more quickly.

Second, mothers could re-orient their work-family preferences during *long* work interruptions, to the detriment of work commitment (Evertsson, 2013; Gangl and Ziefle, 2015). The expansion of job-guaranteed leave in 1992 has indeed lead to a sizeable slump in mothers' work commitment, i.e. the relative importance individuals assign to having a career over having a family for life satisfaction. Such a change, in turn, played a part in mothers' labour supply responses to the reform (Gangl and Ziefle, 2015). That German mothers have been pushed to divert effort/commitment from paid to unpaid work is also coherent with mothers' heightened chances of a higher-order birth after parental leave expansion (Arntz et al., 2017).

Hence, a worsening of the motherhood wage dip after 1992 would be consistent with a shift in effort, job guarantee notwithstanding. Conversely, the shorter paid leave introduced in 2007 may have spurred a re-allocation of effort to the market, to the economic benefit of mothers. This is in line with the observed increase in mothers' labour supply at the intensive margin after the 2007 reform (e.g. Kluge and Schmitz, 2018). New mandates for fathers may play a role in this regard: parental leave uptake by fathers gradually increased from around 3% before 2007 to around 30% in recent years, with leave periods typically averaging two months (Bünning, 2015; Tamm, 2018). The heightened involvement of fathers in childcare and housework, also after the leave period, may in turn free up time for mothers to work longer hours (Tamm, 2018) and mitigate the motherhood wage penalty (Andersen, 2018; Petersen et al., 2014).

Third and last, it could be that employers pick up changes in the behaviours and preferences of mothers as market signals. Since leave *uptake* is widespread among mothers, leave *length* may serve as a signal (Albrecht et al., 1999, 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, thus signalling low commitment to the job or to employment in general. Coherently, re-

search has highlighted a significant jump in the wage penalty for leaves exceeding the job-guaranteed arc of 36 months in Germany in the period 1994-2005 (Buligescu et al., 2009). Similar ‘threshold effects’ have been shown to hold also in other countries (Albrecht et al., 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. At the opposite, 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 accrued motherhood wage penalties also under the new leave regime.

Table 1 sums up our expectations regarding the wage responses mothers may have confronted as a result of parental leave reforms in 1992 and 2007. If skill atrophy during work interruptions is key, then the 1992 reform could have exacerbated the motherhood wage penalty. Yet, if job protections shield mothers from the loss of firm-specific human capital, extending the job guarantee in 1992 may have generated relatively smaller penalties for those who return to the same employer rather than changing. Differently, effort-based arguments point, unambiguously, to a worsening of the wage penalty for mothers in the aftermath of 1992, regardless of job continuity. With respect to both human capital loss and effort re-allocation, our hypotheses for the causal impact of the 2007 reform are largely symmetrical. We predict a shrinkage in the motherhood wage penalty as a by-product of a quicker re-entry in the labour market and longer working hours. Over and beyond, signalling dynamics may have generated heterogeneity in the impact of each policy shift, depending on the timing of mothers’ return to work.

2. Empirical approach

2.1. Data, sample, and design

We use longitudinal data from the German Socio-Economic Panel (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 to K, that is, all original samples for both West and East Germany as well as refreshment and booster samples added up to 2012.

We ran separate analyses for two 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, a requirement of our preferred modelling strategy (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 three 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 to 45, working as dependent employees, with at least two valid wage observations, and with non-missing information on all 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 fifteenth 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 in each case is similarly made up of 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 never give birth to a child in the respective post-reform period and are thus unaffected by parental leave reforms. We are thus left with 709 women (of which 456 are treated) for the evaluation of Reform 1 and 1,040 women (of which 490 are treated) for Reform 2, followed for an average of around 6 and 7 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).

2.2. Estimating the effects of parental leave reforms

To identify the causal effect of parental leave reforms on the motherhood wage penalty, we implement a person-level difference-in-differences (DiD) design relying on the fixed-effects (FE) estimator (Francesconi and Van der Klaauw, 2007; Gangl and Ziefle, 2015). Rather than focusing on a group-level comparison, as in classic DiD, our strategy singles out within-individual variation and accounts for time-invariant unobserved heterogeneity. Via the inclusion of individual fixed effects, indeed, we manage to net out, first, compositional differences across treatment groups and, second, endogenous changes to women’s fertility and labour market participation choices, as long as time-constant factors are at the root of both these sources of bias (see the discussion in the next paragraph). Additionally, the choice of the FE estimator makes our design more comparable to the bulk of the literature on the motherhood wage penalty, also deploying such specification strategy (e.g. Anderson et al., 2003; Gangl and Ziefle, 2009; Kühhirt and Ludwig, 2012; Fernández-Kranz et al., 2013; Harkness, 2016).

The core of our model boils down to the product between event-time dummies capturing the effects of motherhood on log-hourly wages and a dummy for treatment status D_i . More specifically, our model takes the form:

$$y_{it} = \sum_{k=0}^{k=K} \beta_k \mathbb{1}(k = t - e^i) \times D_i + \mathbf{X}_{it}'\gamma + \theta_i + \epsilon_{it} \quad (1)$$

where the dependent variable is the log of real hourly wages for an individual i in year t . Hourly wages are obtained by dividing gross monthly earnings by the amount of actual weekly working hours multiplied by 4.35 (the approximate number of weeks in a month). If the actual working time is not available, we substitute for it adding up the amount of contractual working hours and the reported hours of overtime (Kühhirt and Ludwig, 2012). Wages are then logged and indexed at 2014 consumer prices. We also trim observations whose wage values were smaller than 1 or bigger than 100 to reduce the influence of outliers on our estimates.

To study the effect of motherhood on wages before and after a given reform, we use an

event-study specification (e.g. Kleven et al., 2018; Kuziemko et al., 2018). Let e^i denote the year of the event of interest, the birth of a first child for a woman i , and t be the calendar year. Our specification includes event-time dummies k for each year since the event, starting from $k = 0$, i.e. the first interview year after a child’s birth, to year K , the most recent interview year observable for a given woman after the birth of her first child. Since the post-reform period for Reform 1 runs from 1992 to 1998, a woman giving birth in 1992 would possibly be observed up to $k = 6$, one giving birth in 1993 up to $k = 5$, and so forth. Similar, for the period after Reform 2 running from 2007 to 2014, we are able to track treated women up to year $k = 7$ after a child’s birth. Hence, we will compare wage changes for treated and controls for $k = 0, 1, 2, \dots, 6$ for Reform 1 and $k = 0, 1, 2, \dots, 7$ for Reform 2. Control-group women will be observed for longer periods of event time and dummies for such later years k will also be included in the model.

Sample sizes in each cell k in which we have both treated and control-group women are reported in Table 1A in the Appendix. Given the extremely small number of treated women in $k = 6$ for Reform 1 and $k = 7$ for Reform 2, we will report but not focus on the corresponding estimates in our regression models. Year-specific sample sizes in the remaining years vary from a minimum of 30 women to a maximum of 274 women. 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 will return to in our Robustness section. Even if under-powered, 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 et al., 2013; Kleven et al., 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).

Multiplying our event-time dummies by the indicator variable D_i ($= 1$ if treated by a given

reform), we obtain an estimate of the effects motherhood has on wages as years go by after first childbirth and separately for women belonging to each treatment arm. Since we do not explicitly model benefit receipt, the differences between the effects motherhood has on the wages of treated vis-à-vis control-group women have to be regarded as intention-to-treat (ITT) effects of each parental leave reform. Also, when attributing such effects to parental leave reforms, we are assuming that no contemporaneous shock could have also induced changes to the motherhood wage penalty. In other words, there should not be any contextual policy change explicitly targeting our treated group and so much so to produce wage effects for this group relative to the controls. Recent federal expansions of state-subsidised childcare for toddlers, however, fit this profile. These reforms, being carried out in 2005 and 2008, surround the 2007 parental leave reform and similarly target new mothers. The subsequent increase of childcare availability has been associated with a reduction in the length of work interruptions for (West) German women and an increase in their probability of returning to paid work (Zoch and Hondralis, 2017). Yet, these associations were limited to women experiencing second-order parities, a finding that further motivates our choice to focus on first-time mothers instead.

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 residence². Individual fixed effects θ_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. Regression coefficients in our wage equations are in the log scale. They can be considered accurate approximations of changes in wage levels in the percent-change scale insofar as they lie roughly in the $(-.25, .25)$ interval. For the purpose of substantial interpretation, we will map coefficients that exceeds these thresholds back to the percent-change scale (using the transformation $e^{\beta_k} - 1$).

Finally, to investigate mechanisms, we ran the same model displayed in Equation 1 on three auxiliary outcomes. First, we track a woman's take-up 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 twelve calendar months in the previous year. Importantly, we kept focusing on first births only, meaning that leave uptake following a higher-order parity does not contribute to a woman’s leave share for the purposes of our analysis. 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, since leave share can be observed regardless of whether a woman works or not, for this outcome we include in the estimation sample all available person-year records for both the 709 women on which we evaluate the impact of Reform 1 (7,476 total person-year records) and for the 1,040 women on which we evaluate Reform 2 (11,789 total person-year records).

For the other two outcomes, 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 as in the main models for wages³. Specifically, to track work effort, we complement previous evidence (e.g. Gangl and Ziefle, 2015) by focusing on changes in women’s actual working hours. As for job continuity, implicated by the job guarantee built in parental leave provisions, our dependent variable becomes tenure with the current employer.

2.3. Endogenous fertility: weighting by means of IPTW

Endogenous fertility behaviour and sample selection bias are well-recognised threats in the study of motherhood and wages (e.g. Elwert and Winship, 2014) and our design is no exception. Specifically, fertility behaviour might be endogenous to the reform process itself. After all, treatment assignment is here conditional on giving birth to a child after a given reform threshold (1992 or 2007). Bias could derive then from differences in the characteristics of women selecting into parenthood before and after the reform (e.g. Tamm, 2013).

Fixed-effects estimation takes care of time-invariant factors affecting fertility and, thereby, treatment assignment. We can thus exclude compositional differences in the “type” of women becoming first-time mothers before and after a given reform, as long as these differences lie in time-invariant characteristics. Also, we focus on first-time mothers considering that parental leave reforms in Germany may have affected second- and higher-order

births. Birth spacing has seemingly declined following the expansion of parental leave in 1992 (Arntz et al., 2017) and conversely appears to have extended after the retrenchment of the benefit scheme in 2007 (Cygan-Rehm, 2016).

Nevertheless, to improve the credibility of our design, we weight our models by means of inverse probability of treatment weighting (IPTW, e.g. Morgan and Winship, 2007; Hernan and Robins, forthcoming; for recent applications in sociology, see e.g. Breen and Ermisch, 2017; Biegert and Kühhirt, 2018). Weights are first derived by running a logistic regression for the probability of belonging to the treated group ($D_i = 1$) rather than the control group ($D_i = 0$) for a given reform. We model this probability as a function of covariates Z_i measured in the wave prior to the first child’s birth for each woman. Covariates tap both women’s work history up to that point and the characteristics of their household. Specifically, we include: years spent in full-time employment; years spent in part-time employment; years spent in unemployment; employment status (dummy, 1 if employed); weekly working hours, the hourly wage, and tenure with the current employer (all three set to 0 if not employed); marital status (dummy, 1 if married); household income and household income squared (both divided by 100 and excluding a woman’s own labour income); the years spent in unemployment by a woman’s partner, plus an indicator for whether this information is missing⁴.

We deploy stabilised treatment weights stw_i defined as the ratio between the unconditional probability for an individual i to belong to the her treatment arm and the same probability conditional on covariates Z_i ,

$$stw_{D_i=1} = \frac{P(D_i = 1)}{P(D_i = 1|Z_i)} \quad stw_{D_i=0} = \frac{1 - P(D_i = 1)}{1 - P(D_i = 1|Z_i)}$$

Weighting by means of IPTW creates a pseudopopulation in which treatment assignment is independent from observable confounders Z_i . In other words, observations with covariate values that are over-represented among one treatment arm are down-weighted, while observations with covariate values that are under-represented in one treatment arm are up-weighted. Stabilised treatment weights, in particular, typically take less extreme values than conventional IPT weights (e.g. $tw_{D_i=1} = \frac{1}{P(D_i=1|Z_i)}$) for those units with either very

low or very high probabilities of treatment, thus limiting the influence of such outliers on the final weighted estimates⁵. To further counteract this risk, we also bottom code our weights at the 5th percentile and top code them at the 95th percentile, following previous empirical applications in the literature (Biegert and Kühhirt, 2018).

Covariate balance obtained by deploying stabilised treatment weights is assessed in Table 2, in which we display the weighted average of each variable in the treated and control groups. Normalised differences⁶ between means in each group are all below the threshold of $\pm .25$, signalling a satisfying balance between treatment arms is achieved thanks to our weighting strategy (Imbens and Wooldridge, 2009). To strengthen our claims on time-varying confounding, Figures 1A and 2A plot the the main variables included in Z_i , respectively for Reform 1 and Reform 2. We compute the mean of each variable and weight it by stw_i , separately for control and treatment groups and over event time. To validate our design, covariate balance across treated and controls should hold in the years prior to first childbirth. Visual inspection suggests that a good balance on observable characteristics holds in the re-weighted treatment groups for both reforms, with the only possible exception of the unbalance in household income across Reform 1 groups. Overall balance also extends to years prior to the one immediately preceding first childbirth ($k = -1$), i.e. the year in which variables Z_i are measured to compute our weights. We are thus confident that IPTW eliminates or at least reduces observable group differences in the type of women becoming first-time mothers before and after each reform⁷.

2.4. Selection into employment and fixed-effects estimation

A second source of concern is the extent to which our findings might be tainted by sample selection bias. At any given point in time t , women take part in the estimation sample only if observed in gainful employment. If such women systematically differ from their non-employed counterparts, estimates based only on the former group might be biased. More specifically, since both motherhood and the magnitude of wage offers influence a woman's chances to accept paid work, conditioning our estimates to the sub-sample of employed women may invalidate the claim of a *causal* motherhood wage penalty. Accepting paid work becomes a collider on the causal path between motherhood and wages, as formalised by Elwert and Winship (2014).

Studies focused more broadly on the gender wage gap suggest that German women are, on average, positively selected into employment (Olivetti and Petrongolo, 2008). Women with better earnings potentials will thus be disproportionately represented in our estimation sample. Such positive selection may be particularly pronounced in the aftermath of motherhood and even more so for mothers subjected to subsequent parental leave regimes. Indeed, highly-educated women are more likely to return (faster) to paid work after childbirth (Gutiérrez-Domènech, 2005; Fitzenberger et al., 2013) and have been so more clearly after parental leave reform in the early 1990s (Drasch, 2012; Arntz et al., 2017), with no indication of a reversal in the 2000s (Stahl and Schober, 2017; Kluve and Schmitz, 2018). Hence, our estimates of both the motherhood penalty *per se* and of the effects of parental leave reform could be regarded as conservative. Yet, fixed-effects estimation should take care of the changing composition of women after motherhood and after policy reform, insofar as these changes in the composition of women can be traced back to time-invariant variables such as a woman’s level of education (see Francesconi and Van der Klaauw, 2007).

3. Findings

3.1. Wage responses

Our main findings are depicted in Figure 1 and full estimates are available in Tables 2A and 3A in the Appendix. On the left panel of Figure 1, we contrast motherhood wage effects for control-group and treated women, respectively prior and after Reform 1. Control-group women, who gave birth prior to parental leave expansion in 1992, experience a sizeable motherhood wage penalty in the years following birth. Such wage losses oscillate between roughly 13 log points in Year 2 after first childbirth to 35 log points ($e^{(-.354)} = -29$ percent) in Year 6. As for treated women, giving birth after parental leave expansion in 1992, the motherhood penalty mostly hangs between 10 (Years 0 and 2 to 5) and 16 log points (Year 1), with the exception of Year 6 after first birth (around 37 log points, $e^{(-.377)} = -31$ percent). Overall, differences in the two sets of estimates suggest little change in the motherhood penalty per Reform 1 and, if anything, an improvement felt by women treated by parental leave expansion. As displayed in the third column of Table 2A, differences across treated and controls are typically quite noisy, with the

exception of Years 4 ($p = .083$) and 5 ($p = .015$).

Turning to the right panel of Figure 1, control-group women, who had their first child prior to 2007, experience a wage penalty from Year 2 onwards. Wage losses range from 8 to little in excess of 20 log points, from Year 2 onwards. Differently, for women benefiting from the new parental leave benefit in 2007, wage effects in the aftermath of motherhood are closer to zero and noisier, as 95% confidence interval include both negative and positive values. At the same time, we cannot rule out nil *differences* between treated and controls with the exception of Years 4 ($p = .001$) and 5 ($p < .001$).

Overall, both reforms seem to have brought about modest changes to the motherhood wage penalty in Germany. For Reform 1, estimates in the first years after childbirth are, for the most part, statistically indistinguishable between treated and controls. This suggests limited scope for the kind of positive signalling that one might have expected after Reform 1. Mothers contributing to those estimates are “early returners” with respect to the norm in Germany at the time, especially considering the 36-months job guarantee period installed in 1992. Yet, also such early returners face wage losses. The persistence of the penalty in the following years, albeit somewhat improved for women treated by Reform 1, may suggest that the negative wage effects of human capital loss and the positive wage effects of job continuity, both spurred by the expanded job guarantee, cancel each other out on average.

As for Reform 2, we can no longer detect a motherhood wage penalty for women treated by the overhaul of the parental leave benefit in 2007. At the same time, as displayed more fully in the last column of Table 3A, 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 harsher, not lighter, penalty in the years long after first childbirth, as “late returners” now deflect the new norm of reprising work after the 12 months of benefit receipt⁸. To shed light on the possible mechanisms underlying wage responses to both reforms we then turn to our auxiliary outcomes.

3.2. Mechanisms

Figure 2 contrasts treated and control-group women for Reform 1, first, in terms of their leave share following first childbirths. Complementing previous evidence on the labour supply responses to parental leave expansion (Schönberg and Ludsteck, 2014; Ziefle and Gangl, 2014), we find evidence of sizeable increases in the share of time treated mothers spend on leave in Years 0 to 3 after first childbirth. Differences in leave share between treated and control-group women range from 9 percentage points in Year 3 to 37 and 35 in Years 1 and 2 respectively ($p < .001$ for each difference between treated and controls in Years 0, 1, 2, and 3). Especially in later years (2 and 3), such increases thus testify to the use new mothers make of the longer job guarantee period installed in 1992.

As for working hours, we do not find any evidence of change in the aftermath of Reform 1. In each year following childbirth, mothers in the treated and control group similarly reduce their weekly working hours by around 5-6 hours in Year 0 and up to around 13 in the following years. As much as working hours can proxy for work effort, we thus do not identify a shift in effort in our sample of employed mothers after Reform 1 (cf. Gangl and Ziefle, 2015). Differently, the third panel of Figure 2 suggests substantial improvements for the treated in terms of tenure with the current employer, especially in Years 0 to 3. In line with existing research (Baker and Milligan, 2008), extending job guarantee rights may have thus improved German mothers' job continuity. At the same time, the effects of motherhood on tenure remain negative also for the treated, suggesting that switching employers after motherhood was not at all infrequent despite the longer job-guaranteed period installed in 1992 (e.g. Arntz et al., 2017).

On balance, the combination of longer leave uptake, stable reductions in working hours, and somewhat improved job continuity may account for the overall stability in the motherhood wage penalty we have observed comparing prior and after Reform 1. Our conclusion is that the costs attached to extended career breaks, on the one hand, and the benefits of maintaining one's pre-birth employer, on the other, largely cancel each other out on average.

Change in the same auxiliary outcomes is also investigated in Figure 3, this time comparing treated and controls for Reform 2. In the first panel, evidence supports the idea that

the 2007 Reform reduced the share of time spent on leave in the medium run. Specifically, we find reductions in leave share for the treated vis-à-vis the controls in Years 2 and 3 after first childbirth. While control-group women experienced an increase in leave share of around 36 and 12 percentage points in Years 2 and 3 ($p < .001$), treated women experienced increases of around 23 and 5 percentage points in the same years ($p < .001$). Differences between the two sets of estimates for treated and controls in Years 2 and 3 could also be detected ($p < .001$ and $p = .002$ respectively). Conversely, in Years 0 and 1 treated and controls increase their leave share similarly, even if with a slight increase for the treated in Year 1 ($\beta_{Controls, Year1} = .63$, $\beta_{Treated, Year1} = .67$, p for the difference between the two = .140). We take this to be in line with studies highlighting how German mothers, after 2007, expanded their time spent on leave in the period of benefit receipt and started to return more often right after its expiration (Ziefle and Gangl, 2014; Bergemann and Riphahn, 2017; Kluve and Schmitz, 2018).

We can also detect modest increases in working hours for the treated as compared to control-group women, in the years immediately following childbirth. As depicted in the central panel of Figure 3, while both treated and controls substantially reduce working hours after first childbirth, mothers after 2007 register some increase in working hours vis-à-vis the controls, notably of around 3.1 hours in Year 0 ($p = .137$), 3.8 hours in Year 1 ($p = .008$), 2.6 hours in Year 2 ($p = .052$), and 2.3 hours in Year 3 ($p = .088$). This pattern is in line with previous evidence pointing to an increase in working hours – among part-timers – spurred by the 2007 reform (Kluve and Schmitz, 2018). Differently, we cannot find any evidence of changes in tenure accumulated with the current employer on average, when comparing treated and controls in the third panel of Figure 3 (cf. *ibidem*).

All in all, if the wage prospects of German mothers improved after 2007, evidence in Figure 3 points to shorter time spent on leave and slight increases in working hours, providing some evidence for human capital and effort mechanisms as motors behind the wage responses to Reform 2.

4. Robustness

Having examined motherhood wage penalties before and after two parental leave reforms in Germany, we come to two main findings. As a result of longer work interruptions covered though by job protection rights, wage penalties for German women persisted after Reform 1, yet the size of such penalties reduced especially in the medium term. Reform 2 in 2007 also mitigated wage penalties, again detected primarily in the medium term. We attribute this to human capital and effort channels, as German mothers cut their time on leave after the very first years after childbirth and slightly increased their weekly working hours. We find limited scope for signalling accounts throughout.

Netting out time-constant unobserved heterogeneity via individual fixed effects, our findings critically hinge on a *selection on time-varying observables* assumption. We use IPT weights to address imbalance across treatment groups on a set of time-varying variables we were able to observe in the year prior to first childbirth. As noted in the previous sections (and per Table 2 and Figures 1A and 2A), weighted average characteristics across treatment groups are similar enough, with few exceptions. For example, treated women in 1992 have higher levels of household income, on average, than control-group women in the years prior to birth. Improvements in the wage penalty for treated women after Reform 1 may thus still reflect this composition imbalance, rather than a causal effect of the reform itself.

To address this, we repeated our analysis using an even stronger balancing technique, namely entropy balancing (Hainmueller, 2012). Entropy balancing is a non-parametric algorithm that estimates weights such that, once re-weighted, the distribution of a set of variables in the control group matches that of the treated. The algorithm matches covariates across treatment groups directly on sample moments (mean, variance, skewness). As displayed in Figure 4A in the Appendix, covariates after balancing are matched exactly across treatment groups, and differences in means between treated and controls reduce to 0.

We apply entropy-balancing weights and re-estimate our main models for the motherhood wage penalty across parental leave reforms. As portrayed in Figure 4, estimates for treated and controls now closely resemble one another in the case of Reform 1. For both

groups, we find a motherhood wage penalty of around 20 log points or more from Year 1 onwards. Hence, our diagnosis of persistence in the motherhood penalty before and after Reform 1 comes out reinforced. Such findings are in line with previous literature similarly documenting persisting motherhood wage penalties across cohorts of German mothers in the same period (Gangl and Ziefle, 2009). Our event-study approach, though, provides larger estimates of motherhood wage losses with respect to previous studies, especially when it comes to older cohorts of women such as control-group mothers before Reform 1 (cf. *ibidem*; Kühhirt and Ludwig, 2012).

Differently, we can no longer detect motherhood wage penalties after Reform 2. The right panel of Figure 4 shows that such conclusions for Reform 2 are substantially unchanged when applying entropy balancing rather than IPT weights. Beside the assumption of selection on observables, though, small cell size may limit the credibility of our conclusions due to low statistical power. For Reform 2 only, we can partially address this issue by augmenting our sample size, adding the SOEP sub-sample L1. L1 is a booster sample of a specific family type, that of households having at least one newborn between 2007 and 2009. Among L1 sample members we thus select women aged 16-45 with their first childbirth occurring in 2007 or afterwards, i.e. women that qualify to be in the treatment group for the 2007 reform. We added them up to our main Reform 2 sample and ended up with an augmented sample of 1,747 women (707 more than in the main analyses) and 9,444 person-year records. For such augmented sample, however, we cannot re-weight estimates by means of IPT or entropy-balancing weights since women in the L1 sub-sample were not interviewed the year prior to first childbirth.

Nevertheless, we repeated our main analyses for Reform 2 on such bigger sample to increase statistical power. Figure 5 juxtaposes our main estimates for Reform 2 on the left panel and their replication adding the booster sample on the right. Our substantial conclusions on the effect of Reform 2 hold in that, particularly in the augmented sample, we see a neat improvement in the motherhood wage penalty for the treated vis-à-vis the controls. Differences between the two groups can be ascertained already starting in Year 2 and, for the treated, we can no longer detect a motherhood wage penalty in any of the years following childbirth. To be sure, our conclusion is then that Reform

2 lead to an improvement in the wage effects of motherhood. Both our main estimates for 2007 (Figure 1) and the augmented ones in Figure 5 are consistent with the absence of a penalty. Yet, lower bounds of our 95% confidence intervals are also consistent with penalties of non-negligible size for those women treated by Reform 2 (e.g. Bernardi et al., 2017).

Discussion and conclusions

We have examined the causal effects that parental leave legislation may have had on motherhood wage penalties in Germany. From 1992 onwards, a maternalist leave scheme combined long periods of benefit receipt and even longer periods protected with 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. Conversely, since 2007, German parental leave features a shorter, earnings-related benefit, intact job guarantee rights, and a take-it-or-leave-it quota usually taken up by fathers. Per effect of this reform, we can no longer detect a wage penalty for new mothers, who now concentrate their leave taking in their first years after childbirth and also work longer hours in the same years.

Our findings complement previous comparative work who inferred the importance of labour market and welfare institutions indirectly, comparing motherhood wage penalties across countries. The design we pursued here traces wage penalties back to how parental leave legislation changes, sometimes drastically, over time within a single country. Such approach also sheds light on how “women/family-friendly” policies may have or not perverse effects on the careers of women depending on context and policy design (e.g. Mun and Jung, 2018).

We sought to identify the effects of a single policy program as it changes over time. Following a vast literature on German women’s labour supply behaviour, we focused on parental leave reforms in 1992 and 2007. Nevertheless, policy changes are often gradual and concomitant. Future research could tease out the wage effects, if any, of policy transformation beyond parental leave mandates, such as those related to the availability of childcare service or to the flexibilisation of working-time arrangements.

Additionally, with a small sample size, our conclusions regard only first-time mothers and are only concerned with the “average” wage penalty. A second limit of our analyses is thus that we cannot ascertain heterogeneity in women’s wage responses. In particular, previous research has shown that the labour supply benefits of Reform 2 were only felt by highly-educated/high-income mothers (Bergemann and Riphahn, 2017; Kluve and Schmitz, 2018). Our findings suggest that such changes to labour supply behaviour improved the motherhood wage penalty. It is therefore worth to ask if wage gains were also a prerogative of “advantaged” women and with what consequences for class-and-gender inequality (cf. Mandel, 2012).

Related to effect heterogeneity across social groups is the quest for mechanisms behind such effects. We cannot claim that the wage responses we have highlighted are uniquely ascribable to human capital, effort, or signalling. With respect to employers’ role, for example, a prominent demand-sided explanation hinges upon theories of statistical discrimination. Risk-averse employers might pay women less than (equally productive) men, believing that women will eventually take time out from work, have their working hours reduced or leave their job altogether. In the eyes of employers, long and generous leaves raise the costs and uncertainties surrounding female labour turnover and work effort. Employers might thereby resort to statistical discrimination to insure themselves against employing female workers they deem more ‘risky’.

Although commonplace in the literature (e.g. Gangl and Ziefle, 2009), we suggest that the causal chain linking parental leaves to the motherhood wage penalty through statistical discrimination may need refinement. It is unclear, first, if such discriminatory practices would apply to mothers, thus contributing to their wage penalty, or more generally to women of childbearing age who might have kids and take leaves in the future (e.g. Gupta and Smith, 2002; Ejrnæs and Kunze, 2013).

Second, even if applying to all women of childbearing age, statistical discrimination hardly applies to all women equally. Facing uncertainty regarding which women will take advantage of leave schemes, employers could end up paying unfairly low wages particularly to those women who will eventually remain childless or that will manage to commit to a career even after motherhood, thus defying employers’ ‘group’ expectations. Employer

discrimination in this framework amounts to overshooting, hurting the most those women ‘advantaged’ in terms of human capital, skill, occupational attainment, career aspirations, and so forth (Mandel, 2012).

Hence, we believe future research could tackle whether Reform 1, or similar policy schemes, depressed the wages of childbearing-age women more generally and highly-educated women in particular, providing some grounds for the model of statistical discrimination we sketched here. Conversely, it could be that Reform 2 mitigated employers’ expectation of long motherhood breaks, with positive spillover effects on the wages of (highly-educated) women of childbearing age more broadly. We leave these hypotheses to further research.

To wrap up, motherhood wage penalties in Germany persisted in the last 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 the nuts and bolts of specific institutions may thus critically shape gender economic inequality, at times maintaining it and at times reducing it.

Notes

¹Given the high usage of part-time work at labour market re-entry, encouraged by parental leave schemes themselves (Schönberg and Ludsteck, 2014; Bergemann and Riphahn, 2017), (West) German mothers could have experienced wage losses due to the lower returns to work experience for part-timers (e.g. Fernández-Kranz et al., 2015). Part-time employment is relatively well-paid in Germany though, both with respect to full-time equivalent jobs and in international comparison (Bardasi and Gornick, 2008), and previous research has suggested that maternal part-time working has little to do with the motherhood wage penalty in this context (Gangl and Ziefle, 2009; Kühhirt and Ludwig, 2012).

²We also probed our models to the inclusion of year fixed effects to account for exogenous shocks that may similarly hit the wages of treated and control-group women. Our main results are substantially unchanged (output available upon request). Year fixed effects, nonetheless, may introduce multicollinearity with respect to event time dummies and therefore we deem it safer to exclude them (e.g. Borusyak and Jaravel, 2016; Imai and Kim, 2017).

³With respect to the main analyses on wages, due to missing data on tenure with a woman’s current employer, we lose a relatively small number of person-year records for Reform 1 (36 observations) and for Reform 2 (3 observations).

⁴We build these two variables such that, for instance, if the information on partner’s unemployment is missing (e.g. because a woman has no partner), “years spent in unemployment by the partner” is set to 0 while the indicator for “missing information on partner” is set to 1.

⁵In our analytical samples, weights $stw_{i,Reform1}$ have a mean equal to 0.92 ($SD = 0.27$), while weights $stw_{i,Reform2}$ have a mean equal to 0.96 ($SD = 0.24$). Weights whose mean value is close to 1 and whose standard deviation is relatively small suggest that the positivity assumption holds, i.e. there is a non-zero probability of belonging to each treatment group D_i for any combination of the values of covariates Z_i (e.g. Hernan and Robins, forthcoming; Biegert and Kühhirt, 2018).

⁶Normalised differences are given by the difference in means of a given x between treated and controls, scaled by the square root of the sum of the variances of x for treated and controls (Imbens and Wooldridge, 2009).

⁷To validate our design, we also looked at whether re-weighted control and treatment groups exhibit parallel trends in the development of our dependent variable prior to first childbirth. Figure 3A in the Appendix plots the within-transformed log of hourly wages separately for each reform period, treatment group, and year prior to first childbirth (up to five years prior). Visual inspection suggests no substantial violation of the parallel trend assumption for each reform pair.

⁸Notably 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 Year 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.

References

- Adda, J., C. Dustmann, and K. Stevens (2017). The career costs of children. *Journal of Political Economy* 125(2), 293–337.
- Aisenbrey, S. and A. Fasang (2017). The Interplay of Work and Family Trajectories over the Life Course: Germany and the United States in Comparison. *American Journal of Sociology* 122(5), 1448–1484.
- Albrecht, J., A. Björklund, and S. Vroman (2003). Is there a glass ceiling in Sweden? *Journal of Labor Economics* 21(1), 145–177.
- Albrecht, J., P. S. Thoursie, and S. Vroman (2015). Parental leave and the glass ceiling in Sweden. *Research in Labor Economics* 41, 89–114.
- Albrecht, J. W., P.-A. Edin, M. Sundström, and S. B. Vroman (1999). Career Interruptions and Subsequent Earnings: a Reexamination using Swedish Data. *Journal of Human Resources* 34(2), 294–311.
- Andersen, S. H. (2018). Paternity Leave and the Motherhood Penalty: New Causal Evidence. *Journal of Marriage and Family*, DOI:10.1111/jomf.12507.
- Anderson, D. J., M. Binder, and K. Krause (2003). The motherhood wage penalty revisited: Experience, heterogeneity, work effort, and work-schedule flexibility. *Industrial & Labor Relations Review* 56(2), 273–294.
- Arntz, M., S. Dlugosz, and R. A. Wilke (2017). The sorting of female careers after first birth: a competing risks analysis of maternity leave duration. *Oxford Bulletin of Economics and Statistics* 79(5), 689–716.
- Arulampalam, W., A. L. Booth, and M. L. Bryan (2007). Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *ILR Review* 60(2), 163–186.
- Baker, M. and K. Milligan (2008). How does job-protected maternity leave affect mothers’ employment? *Journal of Labor Economics* 26(4), 655–691.
- Bardasi, E. and J. C. Gornick (2008). Working for less? Women’s part-time wage penalties across countries. *Feminist Economics* 14(1), 37–72.
- Baum, C. L. and C. J. Ruhm (2016). The effects of paid family leave in California on labor market outcomes. *Journal of Policy Analysis and Management* 35(2), 333–356.
- Bergemann, A. and R. T. Riphahn (2017). Maternal Employment Effects of Paid Parental Leave. *SOEPpapers on Multidisciplinary Panel Data Research* 900–2017.

- Bernardi, F., L. Chakhaia, and L. Leopold (2017). ‘Sing me a song with social significance’: The (mis) use of statistical significance testing in European sociological research. *European Sociological Review* 33(1), 1–15.
- Biegert, T. and M. Kühhirt (2018). Taking lemons for a trial run: Does type of job exit affect the risk of entering fixed-term employment in Germany? *European Sociological Review* 34(2), 184–197.
- Borusyak, K. and X. Jaravel (2016). Revisiting Event Study Designs. *Working Paper*, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2826228.
- Breen, R. and J. Ermisch (2017). Educational reproduction in Great Britain: a prospective approach. *European Sociological Review* 33(4), 590–603.
- Buhr, P. and J. Huinink (2015). The German Low Fertility: How We Got There and What We Can Expect for the Future. *European Sociological Review* 31(2), 197–210.
- Buligescu, B., D. De Crombrugghe, G. Mentesoglu, and R. Montizaan (2009). Panel estimates of the wage penalty for maternal leave. *Oxford Economic Papers* 61, i35–i55.
- 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(6), 738–748.
- Cygan-Rehm, K. (2016). Parental leave benefit and differential fertility responses: Evidence from a German reform. *Journal of Population Economics* 29(1), 73–103.
- Davies, R. and G. Pierre (2005). The family gap in pay in Europe: a cross-country study. *Labour Economics* 12(4), 469–486.
- Dieckhoff, M., V. Gash, A. Mertens, and L. R. Gordo (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(5), 981–995.
- Ejrnæs, M. and A. Kunze (2013). Work and wage dynamics around childbirth. *The Scandinavian Journal of Economics* 115(3), 856–877.
- Elwert, F. and C. Winship (2014). Endogenous selection bias: The problem of conditioning on a collider variable. *Annual Review of Sociology* 40, 31–53.
- Evertsson, M. (2013). The importance of work: Changing work commitment following the transition to motherhood. *Acta Sociologica* 56(2), 139–153.

- 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., A. Lacuesta, and N. Rodríguez-Planas (2013). The motherhood earnings dip: Evidence from administrative records. *Journal of Human Resources* 48(1), 169–197.
- Fernández-Kranz, D., M. Paul, and N. Rodríguez-Planas (2015). Part-time work, fixed-term contracts, and the returns to experience. *Oxford Bulletin of Economics and Statistics* 77(4), 512–541.
- Fitzenberger, B., K. Sommerfeld, and S. Steffes (2013). Causal effects on employment after first birth — A dynamic treatment approach. *Labour Economics* 25, 49–62.
- Francesconi, M. and W. Van der Klaauw (2007). The socioeconomic consequences of “in-work” benefit reform for British lone mothers. *Journal of Human Resources* 42(1), 1–31.
- Gangl, M. and A. Ziefle (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(2), 341–369.
- Gangl, M. and A. Ziefle (2015). The Making of a Good Woman: Extended Parental Leave Entitlements and Mothers’ Work Commitment in Germany. *American Journal of Sociology* 121(2), 511–563.
- Gelman, A. and J. Carlin (2014). Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science* 9(6), 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.
- Goebel, J. (2017). SOEP-Core v32 - Documentation on biography and life history data. *SOEP Survey Papers*, No. 418.
- Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2018). The german socio-economic panel (soep). *Jahrbücher für Nationalökonomie und Statistik*, <https://doi.org/10.1515/jbnst--2018--0022>.
- Gregg, P., M. Gutiérrez-Domènech, and J. Waldfogel (2007). The employment of married mothers in Great Britain, 1974–2000. *Economica* 74(296), 842–864.
- Gupta, N. D. and N. Smith (2002). Children and career interruptions: the family gap in Denmark. *Economica* 69(276), 609–629.

- Gutiérrez-Domènech, M. (2005). Employment after motherhood: a European comparison. *Labour Economics* 12(1), 99–123.
- Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis* 20(1), 25–46.
- Harkness, S. E. (2016). The Effect of Motherhood and Lone Motherhood on the Employment and Earnings of British Women: A Lifecycle Approach. *European Sociological Review* 32(6), 850–863.
- Hernan, M. A. and J. M. Robins (forthcoming). *Causal Inference*. Boca Raton: Chapman & Hall/CRC.
- Huebener, M., D. Kühnle, C. K. Spieß, et al. (2018). Parental Leave Policies and Socio-Economic Gaps in Child Development: Evidence from a Substantial Benefit Reform Using Administrative Data. *IZA Discussion Paper No. 11794*.
- Imai, K. and I. S. Kim (2017). When Should We Use Fixed Effects Regression Models for Causal Inference with Longitudinal Data? *Working paper*, <https://imai.princeton.edu/research/files/FEmatch.pdf>.
- Imbens, G. W. and J. M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* 47(1), 5–86.
- Joseph, O., A. Pailhé, I. Recotillet, and A. Solaz (2013). The economic impact of taking short parental leave: Evaluation of a French reform. *Labour Economics* 25, 63–75.
- Kleven, H., C. Landaïs, and J. E. Søgaaard (2018). Children and gender inequality: Evidence from Denmark. *NBER Working Paper 24219*.
- Kluve, J. and S. Schmitz (2018). Back to Work: Parental Benefits and Mothers’ Labor Market Outcomes in the Medium Run. *ILR Review* 71(1), 143–173.
- Kluve, J. and M. Tamm (2013). Parental leave regulations, mothers’ labor force attachment and fathers’ childcare involvement: evidence from a natural experiment. *Journal of Population Economics* 26(3), 983–1005.
- Kottwitz, A., A. Oppermann, and C. K. Spiess (2016). Parental leave benefits and breastfeeding in Germany: Effects of the 2007 reform. *Review of Economics of the Household* 14(4), 859–890.
- Kühhirt, M. and V. Ludwig (2012). Domestic work and the wage penalty for motherhood in West Germany. *Journal of Marriage and Family* 74(1), 186–200.

- Kuziemko, I., J. Pan, J. Shen, and E. Washington (2018). The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood? *NBER Working Paper 24740*.
- Lalive, R. and J. Zweimüller (2009). How does parental leave affect fertility and return to work? Evidence from two natural experiments. *The Quarterly Journal of Economics* 124(3), 1363–1402.
- Lersch, P. M., M. Jacob, and K. Hank (2017). Parenthood, Gender, and Personal Wealth. *European Sociological Review* 33(3), 410.
- Loughran, D. S. and J. M. Zissimopoulos (2009). Why wait? The effect of marriage and childbearing on the wages of men and women. *Journal of Human Resources* 44(2), 326–349.
- Mandel, H. (2012). Winners and losers: The consequences of welfare state policies for gender wage inequality. *European Sociological Review* 28(2), 241–262.
- Mandel, H. and M. Semyonov (2005). Family policies, wage structures, and gender gaps: Sources of earnings inequality in 20 countries. *American Sociological Review* 70(6), 949–967.
- Mandel, H. and M. Semyonov (2006). A Welfare State Paradox: State Interventions and Women’s Employment Opportunities in 22 Countries. *American Journal of Sociology* 111(6), 1910–1949.
- Morgan, S. L. and C. Winship (2007). *Counterfactuals and causal analysis: Methods and principles for social research*. Cambridge: Harvard University Press.
- Mun, E. and J. Jung (2018). Policy Generosity, Employer Heterogeneity, and Women’s Employment Opportunities: The Welfare State Paradox Reexamined. *American Sociological Review* 83(3), 508–535.
- OECD (2017). *The Pursuit of Gender Equality: An Uphill Battle*. Paris: OECD Publishing.
- Olivetti, C. and B. Petrongolo (2008). Unequal pay or unequal employment? A cross-country analysis of gender gaps. *Journal of Labor Economics* 26(4), 621–654.
- Petersen, T., A. M. Penner, and G. Høgsnes (2014). From Motherhood Penalties to Husband Premia: The New Challenge for Gender Equality and Family Policy, Lessons from Norway. *American Journal of Sociology* 119(5), 1434–1472.
- Ponthieux, S. and D. Meurs (2015). Gender Inequality. In A. B. Atkinson and F. Bourdignon (Eds.), *Handbook of Income Distribution. Volume 2A*. North Holland: Elsevier.

- Ray, R., J. C. Gornick, and J. Schmitt (2010). Who cares? Assessing generosity and gender equality in parental leave policy designs in 21 countries. *Journal of European Social Policy* 20(3), 196–216.
- Ruhm, C. J. (1998). The economic consequences of parental leave mandates: Lessons from Europe. *The Quarterly Journal of Economics* 113(1), 285–317.
- Schönberg, U. and J. Ludsteck (2014). Expansions in maternity leave coverage and mothers’ labor market outcomes after childbirth. *Journal of Labor Economics* 32(3), 469–505.
- Stahl, J. S. and P. S. Schober (2017). Convergence or divergence? Educational discrepancies in work-care arrangements of mothers with young children in Germany. *Work, Employment & Society* (<http://dx.doi.org/10.1177/0950017017692503>).
- Tamm, M. (2013). The impact of a large parental leave benefit reform on the timing of birth around the day of implementation. *Oxford Bulletin of Economics and Statistics* 75(4), 585–601.
- Tamm, M. (2018). Fathers’ Parental Leave-Taking, Childcare Involvement and Mothers’ Labor Market Participation. *IZA Discussion Paper No. 11873*.
- Trappe, H., M. Pollmann-Schult, and C. Schmitt (2015). The rise and decline of the male breadwinner model: institutional underpinnings and future expectations. *European Sociological Review* 31(2), 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(3), 505–545.
- Zhang, X. (2010). Can motherhood earnings losses be ever regained? Evidence from Canada. *Journal of Family Issues* 31(12), 1671–1688.
- Ziefle, A. and M. Gangl (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(5), 562–581.
- Zoch, G. and I. Hondralis (2017). The Expansion of Low-Cost, State-Subsidized Child-care Availability and Mothers’ Return-to-Work Behaviour in East and West Germany. *European Sociological Review* 33(5), 693–707.

Tables & graphs

Table 1: Summary of the expected changes to the motherhood wage penalty as a consequence of each reform of parental leave in Germany.

PL Reforms	Labour supply	Mechanisms	Wage responses
1992 Reform	Drift to longer career breaks	Human capital	+/-
	Job continuity		
	Maternal PT work	Effort	-
	Selective return to paid work	Signalling	+/-
2007 Reform	Career breaks scaled back	Human capital	+
	Maternal PT work		
	Selective return to paid work	Effort	+
		Signalling	+/-

Note: PL = Parental leave; PT = Part-time.

Table 2: Summary of selected variables for treatment and control groups, weighted by means of IPTW.

	Reform 1			Reform 2		
	Means		Normalized differences	Means		Normalised differences
	Controls	Treated		Controls	Treated	
Years in full-time (FT)	5.37	5.61	0.04	5.97	5.98	0.00
Years in part-time (PT)	0.36	0.44	0.05	1.00	1.23	0.08
Years unemployed	0.26	0.26	0.00	0.40	0.39	0.00
Employed	0.79	0.80	0.01	0.82	0.83	0.01
Hourly wage	4.61	5.09	0.07	8.26	8.80	0.06
Working hours	29.85	30.67	0.03	32.13	32.33	0.01
Tenure	3.80	3.91	0.02	4.02	4.06	0.01
Married	0.65	0.66	0.01	0.47	0.47	0.00
Household income/100	145.85	179.88	0.17	269.81	285.32	0.05
Household income ² /100	32216.87	60607.67	0.10	125147.10	145714.20	0.05
Years unemployed (partner)	0.24	0.22	−0.02	0.33	0.33	−0.01
Missing “Years unemployed (partner)”	0.30	0.27	−0.05	0.30	0.31	0.01
Number of individuals	253	456		550	490	

Notes: All variables are measured in the interview year occurring prior to the first childbirth event.

Source: SOEP 1985-2014.

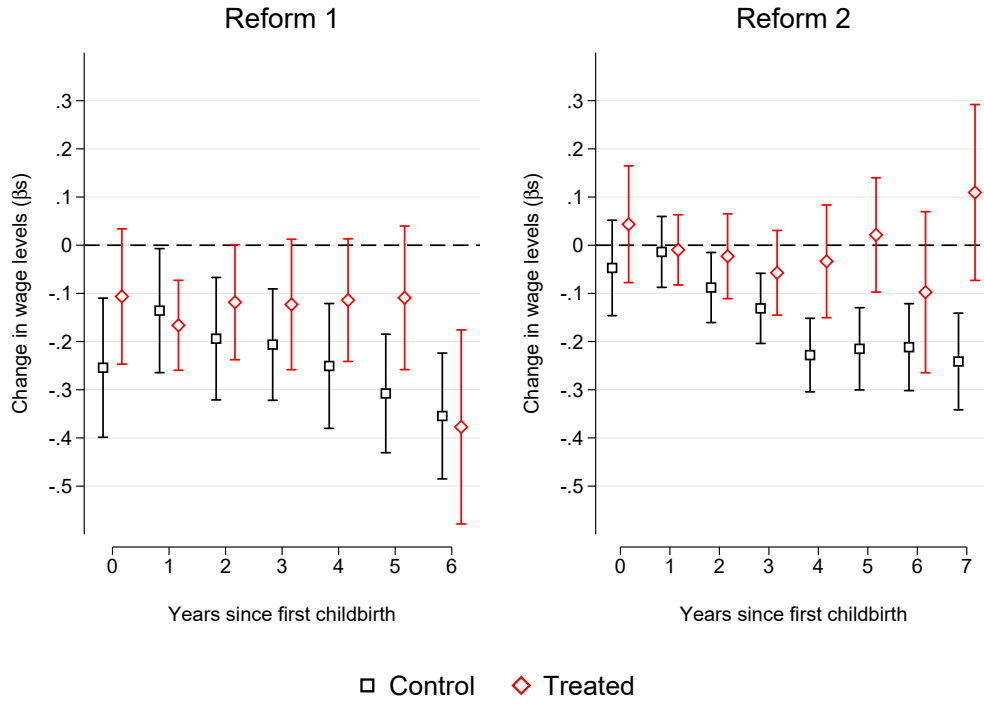


Figure 1: FE estimates (95% confidence intervals) of the motherhood wage penalty across reforms, treatment groups, and relative to event time. Estimates are weighted by IPTW, as detailed in the main text (SOEP 1985-2014).

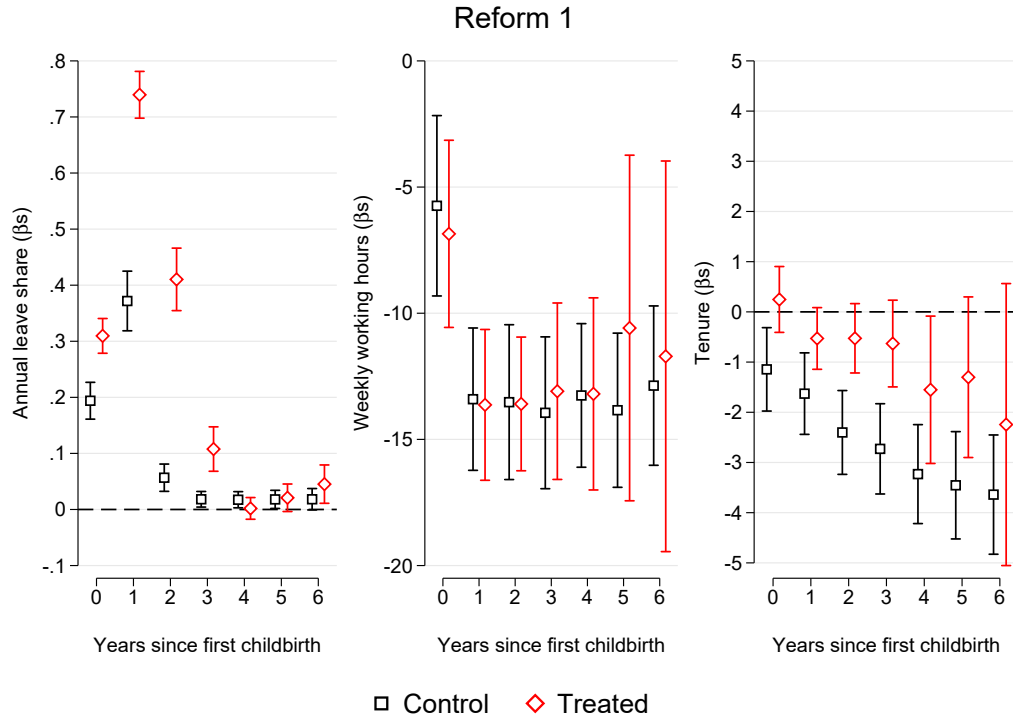


Figure 2: FE estimates (95% confidence intervals) of motherhood effects on (i) leave share, (ii) working hours, and (iii) tenure. Estimates refer to Reform 1, across treatment groups and relative to event time. Estimates are weighted by IPTW, as detailed in the main text (SOEP 1985-1998).

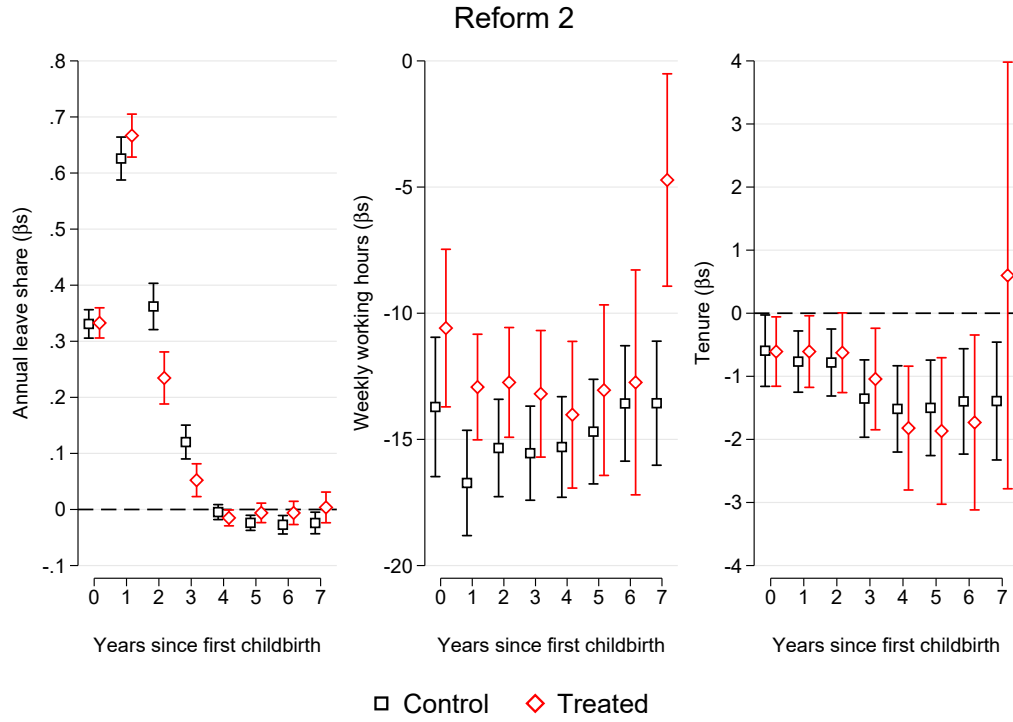


Figure 3: FE estimates (95% confidence intervals) of motherhood effects on (i) leave share, (ii) working hours, and (iii) tenure. Estimates refer to Reform 2, across treatment groups and relative to event time. Estimates are weighted by IPTW, as detailed in the main text (SOEP 1999-2014).

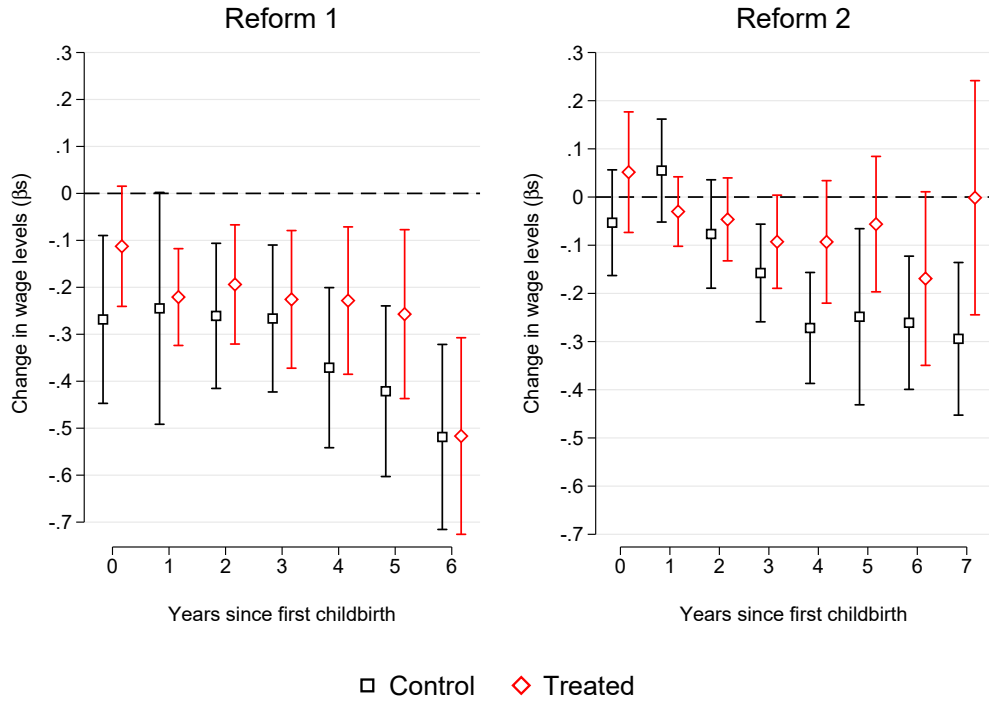


Figure 4: FE estimates (95% confidence intervals) of the motherhood wage penalty across reforms, treatment groups, and relative to event time. Estimates are weighted by means of entropy-balancing weights, as detailed in the main text (SOEP 1985-2014).

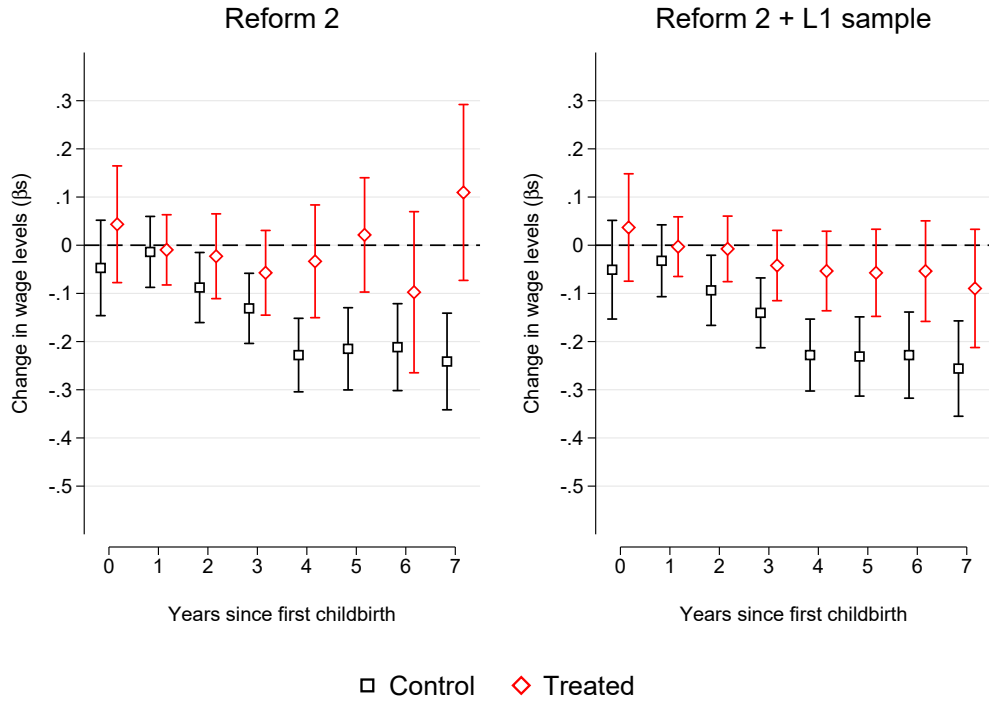


Figure 5: FE estimates (95% confidence intervals) of the motherhood wage penalty across treatment groups and relative to event time. Estimates on the left panel are weighted by means of IPTW, as detailed in the main text (SOEP 1985-2014).

A. Appendix

Table 1A: Sample size for each treatment group by year since first childbirth (companion to the estimates in Figure 1, Tables 2A and 3A). Unweighted counts are raw sample counts, weighted counts are sample counts weighted by IPTW.

	Reform 1				Reform 2			
	Controls		Treated		Controls		Treated	
	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
Year 0: first birth	61	53.8	41	38.1	124	122.3	62	63.6
Year 1 after first birth	86	72.3	86	84.6	180	178.7	203	199.5
Year 2 after first birth	94	78.3	94	95.9	219	207.1	177	179.0
Year 3 after first birth	98	83.0	79	80.2	244	232.9	117	118.6
Year 4 after first birth	109	90.6	59	60.2	261	252.6	86	89.7
Year 5 after first birth	111	93.9	30	31.8	274	267.2	57	59.1
Year 6 after first birth	115	99.7	7	7.7	253	240.8	36	40.5
Year 7 after first birth	.	.	.	.	251	235.7	4	3.5

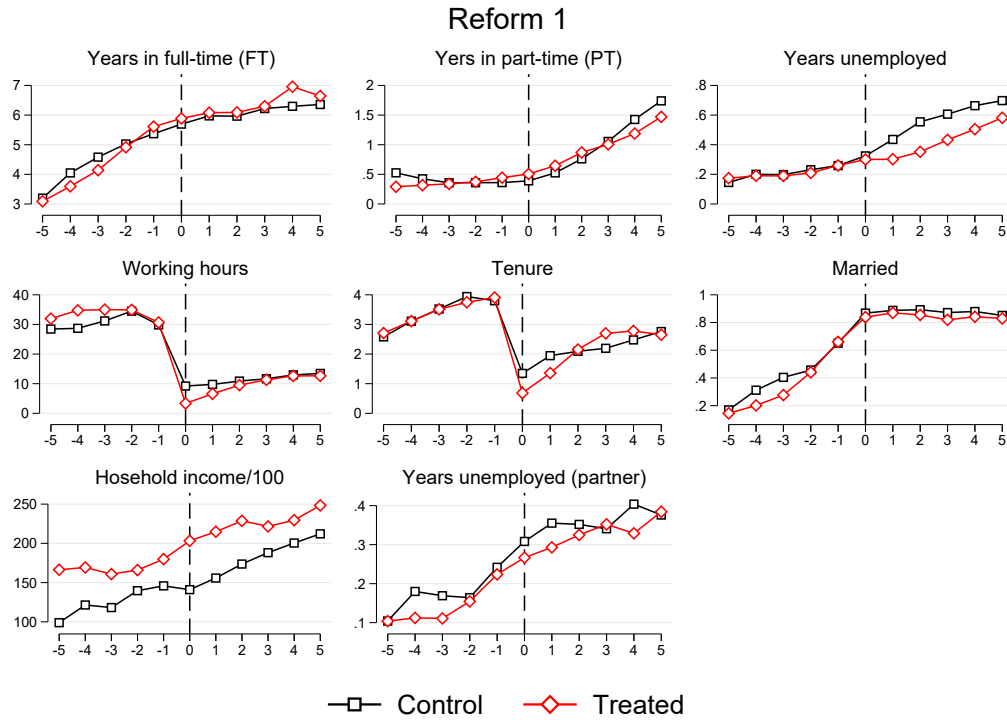


Figure 1A: Means of selected variables for treated and control units (Reform 1, SOEP 1985-1998). Means are computed for each year up to the fifth year prior (after) first childbirth. All means are weighted by IPTW as detailed in the main text.

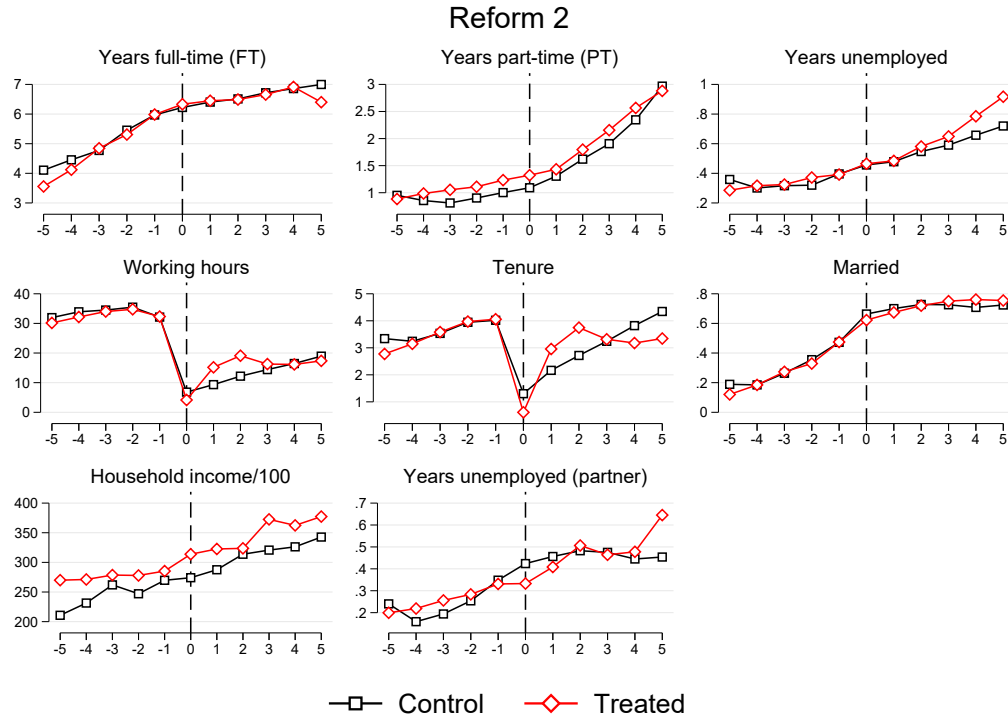


Figure 2A: Means of selected variables for treated and control units (Reform 2, SOEP 1999-2014). Means are computed for each year up to the fifth year prior (after) first childbirth (x axis). All means are weighted by IPTW as detailed in the main text.

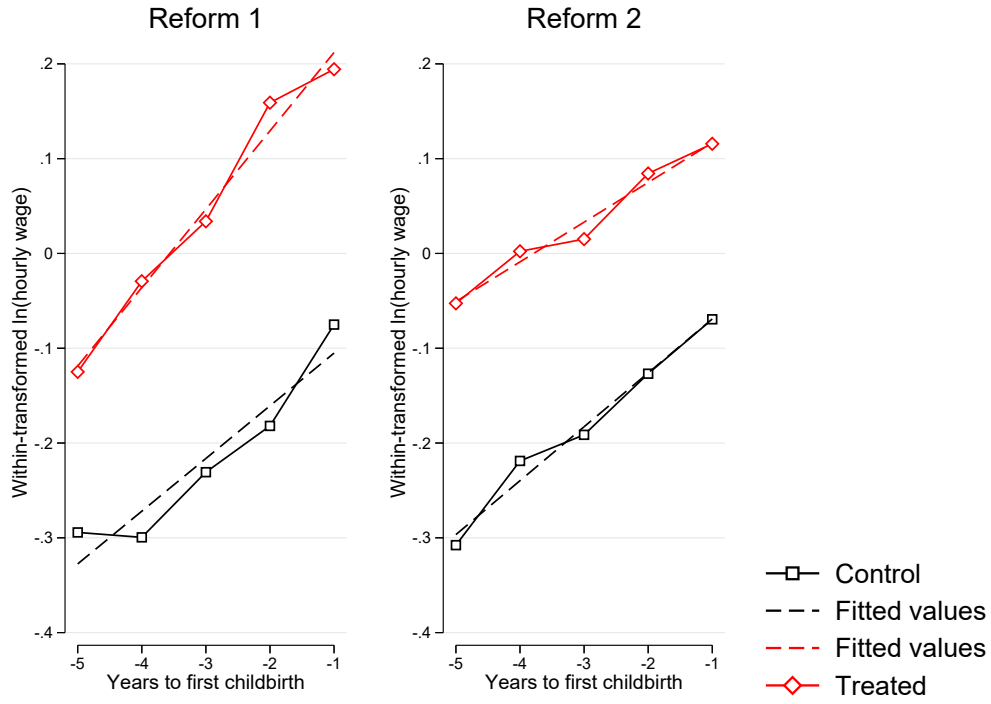


Figure 3A: Within-transformed log of real hourly wages plotted over years to first childbirth, for treated and control units across reform periods. Dashed lines are obtained fitting a linear trend for each treatment-reform group using Stata's `lfit` (SOEP 1985-2014).

Table 2A: FE estimates for the motherhood wage penalty, by treatment group and year since first childbirth (Reform 1).

	Reform 1		
	(1)	(2)	(3)
	Controls	Treated	Difference
	β (SE)	β (SE)	β (SE)
Year 0: first birth	-0.254*** (0.074)	-0.106 (0.072)	0.148 (0.104)
Year 1 after first birth	-0.136** (0.066)	-0.166*** (0.048)	-0.030 (0.078)
Year 2 after first birth	-0.194*** (0.065)	-0.118* (0.061)	0.076 (0.084)
Year 3 after first birth	-0.206*** (0.059)	-0.123* (0.069)	0.083 (0.081)
Year 4 after first birth	-0.251*** (0.066)	-0.114* (0.065)	0.137* (0.079)
Year 5 after first birth	-0.308*** (0.063)	-0.109 (0.076)	0.198** (0.081)
Year 6 after first birth	-0.354*** (0.067)	-0.377*** (0.103)	-0.023 (0.104)
Number of individuals	709	709	709
Number of person-years	4,300	4,300	4,300

* $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: All models include individual fixed effects, as well as dummies for each additional year after childbirth (control group only), a quadratic for age, and dummies for region of residence. Standard errors are clustered at the individual level.

Source: SOEP 1985-1998.

Table 3A: FE estimates for the motherhood wage penalty, by treatment group and year since first childbirth (Reform 2).

	Reform 2		
	(1) Controls	(2) Treated	(3) Difference
	β (SE)	β (SE)	β (SE)
Year 0: first birth	-0.047 (0.051)	0.044 (0.062)	0.091 (0.079)
Year 1 after first birth	-0.014 (0.038)	-0.010 (0.037)	0.004 (0.049)
Year 2 after first birth	-0.088** (0.037)	-0.023 (0.045)	0.065 (0.052)
Year 3 after first birth	-0.131*** (0.037)	-0.057 (0.045)	0.074 (0.048)
Year 4 after first birth	-0.228*** (0.039)	-0.033 (0.060)	0.195*** (0.058)
Year 5 after first birth	-0.215*** (0.044)	0.021 (0.061)	0.236*** (0.057)
Year 6 after first birth	-0.212*** (0.046)	-0.098 (0.085)	0.114 (0.079)
Year 7 after first birth	-0.241*** (0.051)	0.110 (0.093)	0.351*** (0.081)
Number of individuals	1,040	1,040	1,040
Number of person-years	7,668	7,668	7,668

* $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: All models include individual fixed effects, as well as dummies for each additional year after childbirth (control group only), a quadratic for age, and dummies for region of residence. Standard errors are clustered at the individual level.

Source: SOEP 1999-2014.

Table 4A: FE estimates for leave share, by treatment group and year since first childbirth (Reform 1).

	Reform 1		
	(1)	(2)	(3)
	Controls	Treated	Difference
	β (SE)	β (SE)	β (SE)
Year 0: first birth	0.194*** (0.017)	0.310*** (0.016)	0.116*** (0.023)
Year 1 after first birth	0.372*** (0.027)	0.740*** (0.021)	0.368*** (0.035)
Year 2 after first birth	0.057*** (0.012)	0.410*** (0.028)	0.354*** (0.031)
Year 3 after first birth	0.018** (0.007)	0.108*** (0.020)	0.090*** (0.021)
Year 4 after first birth	0.017** (0.007)	0.002 (0.010)	-0.015* (0.009)
Year 5 after first birth	0.018** (0.008)	0.021* (0.013)	0.003 (0.011)
Year 6 after first birth	0.018* (0.010)	0.045*** (0.017)	0.027 (0.016)
Number of individuals	709	709	709
Number of person-years	7,476	7,476	7,476

* $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: All models include individual fixed effects, as well as dummies for each additional year after childbirth (control group only), a quadratic for age, and dummies for region of residence. Standard errors are clustered at the individual level.

Source: SOEP 1985-1998.

Table 5A: FE estimates for weekly working hours, by treatment group and year since first childbirth (Reform 1).

	Reform 1		
	(1)	(2)	(3)
	Controls β (SE)	Treated β (SE)	Difference β (SE)
Year 0: first birth	-5.742*** (1.823)	-6.854*** (1.892)	-1.350 (2.567)
Year 1 after first birth	-13.406*** (1.439)	-13.632*** (1.525)	-0.030 (2.042)
Year 2 after first birth	-13.524*** (1.565)	-13.594*** (1.352)	-0.586 (1.928)
Year 3 after first birth	-13.946*** (1.535)	-13.091*** (1.784)	0.360 (2.184)
Year 4 after first birth	-13.258*** (1.452)	-13.197*** (1.942)	0.182 (2.141)
Year 5 after first birth	-13.844*** (1.559)	-10.585*** (3.494)	3.045 (3.465)
Year 6 after first birth	-12.869*** (1.611)	-11.707*** (3.949)	0.754 (3.759)
Number of individuals	709	709	709
Number of person-years	4,300	4,300	4,300

* $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: All models include individual fixed effects, as well as dummies for each additional year after childbirth (control group only), a quadratic for age, and dummies for region of residence. Standard errors are clustered at the individual level.

Source: SOEP 1985-1998.

Table 6A: FE estimates for tenure with current employer, by treatment group and year since first childbirth (Reform 1).

	Reform 1		
	(1)	(2)	(3)
	Controls	Treated	Difference
	β (SE)	β (SE)	β (SE)
Year 0: first birth	-1.145*** (0.424)	0.248 (0.335)	1.392** (0.549)
Year 1 after first birth	-1.628*** (0.414)	-0.529* (0.314)	1.099** (0.515)
Year 2 after first birth	-2.402*** (0.426)	-0.526 (0.352)	1.875*** (0.530)
Year 3 after first birth	-2.728*** (0.459)	-0.630 (0.441)	2.098*** (0.591)
Year 4 after first birth	-3.231*** (0.503)	-1.552** (0.749)	1.680** (0.854)
Year 5 after first birth	-3.454*** (0.545)	-1.301 (0.816)	2.153** (0.884)
Year 6 after first birth	-3.639*** (0.606)	-2.244 (1.433)	1.395 (1.452)
Number of individuals	709	709	709
Number of person-years	4,264	4,264	4,264

* $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: All models include individual fixed effects, as well as dummies for each additional year after childbirth (control group only), a quadratic for age, and dummies for region of residence. Standard errors are clustered at the individual level.

Source: SOEP 1985-1998.

Table 7A: FE estimates for leave share, by treatment group and year since first childbirth (Reform 2).

	Reform 2		
	(1) Controls	(2) Treated	(3) Difference
	β (SE)	β (SE)	β (SE)
Year 0: first birth	0.331*** (0.013)	0.333*** (0.014)	0.002 (0.019)
Year 1 after first birth	0.626*** (0.020)	0.667*** (0.020)	0.041 (0.028)
Year 2 after first birth	0.362*** (0.021)	0.235*** (0.024)	-0.128*** (0.032)
Year 3 after first birth	0.120*** (0.015)	0.052*** (0.015)	-0.068*** (0.022)
Year 4 after first birth	-0.005 (0.007)	-0.015** (0.007)	-0.010 (0.009)
Year 5 after first birth	-0.024*** (0.007)	-0.006 (0.009)	0.018** (0.009)
Year 6 after first birth	-0.027*** (0.008)	-0.006 (0.011)	0.021* (0.011)
Year 7 after first birth	-0.024** (0.010)	0.004 (0.014)	0.028** (0.014)
Number of individuals	1,040	1,040	1,040
Number of person-years	11,789	11,789	11,789

* $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: All models include individual fixed effects, as well as dummies for each additional year after childbirth (control group only), a quadratic for age, and dummies for region of residence. Standard errors are clustered at the individual level.

Source: SOEP 1999-2014.

Table 8A: FE estimates for weekly working hours, by treatment group and year since first childbirth (Reform 2).

	Reform 2		
	(1) Controls	(2) Treated	(3) Difference
	β (SE)	β (SE)	β (SE)
Year 0: first birth	-13.714*** (1.409)	-10.588*** (1.593)	3.126 (2.098)
Year 1 after first birth	-16.726*** (1.065)	-12.926*** (1.068)	3.800*** (1.431)
Year 2 after first birth	-15.341*** (0.984)	-12.742*** (1.110)	2.599* (1.334)
Year 3 after first birth	-15.547*** (0.951)	-13.194*** (1.279)	2.353* (1.378)
Year 4 after first birth	-15.301*** (1.017)	-14.023*** (1.483)	1.278 (1.522)
Year 5 after first birth	-14.692*** (1.057)	-13.049*** (1.724)	1.643 (1.696)
Year 6 after first birth	-13.575*** (1.167)	-12.742*** (2.274)	0.833 (2.158)
Year 7 after first birth	-13.566*** (1.254)	-4.721** (2.147)	8.845*** (1.819)
Number of individuals	1,040	1,040	1,040
Number of person-years	7,668	7,668	7,668

* $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: All models include individual fixed effects, as well as dummies for each additional year after childbirth (control group only), a quadratic for age, and dummies for region of residence. Standard errors are clustered at the individual level.

Source: SOEP 1999-2014.

Table 9A: FE estimates for tenure with current employer, by treatment group and year since first childbirth (Reform 2).

	Reform 2		
	(1)	(2)	(3)
	Controls	Treated	Difference
	β (SE)	β (SE)	β (SE)
Year 0: first birth	-0.594** (0.289)	-0.608** (0.281)	-0.014 (0.396)
Year 1 after first birth	-0.767*** (0.248)	-0.608** (0.290)	0.159 (0.363)
Year 2 after first birth	-0.781*** (0.271)	-0.626* (0.322)	0.155 (0.384)
Year 3 after first birth	-1.353*** (0.313)	-1.043** (0.411)	0.310 (0.467)
Year 4 after first birth	-1.517*** (0.349)	-1.821*** (0.501)	-0.304 (0.562)
Year 5 after first birth	-1.500*** (0.386)	-1.867*** (0.592)	-0.367 (0.655)
Year 6 after first birth	-1.398*** (0.426)	-1.732** (0.707)	-0.334 (0.751)
Year 7 after first birth	-1.392*** (0.476)	0.599 (1.725)	1.991 (1.718)
Number of individuals	1,040	1,040	1,040
Number of person-years	7,665	7,665	7,665

* $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: All models include individual fixed effects, as well as dummies for each additional year after childbirth (control group only), a quadratic for age, and dummies for region of residence. Standard errors are clustered at the individual level.

Source: SOEP 1999-2014.

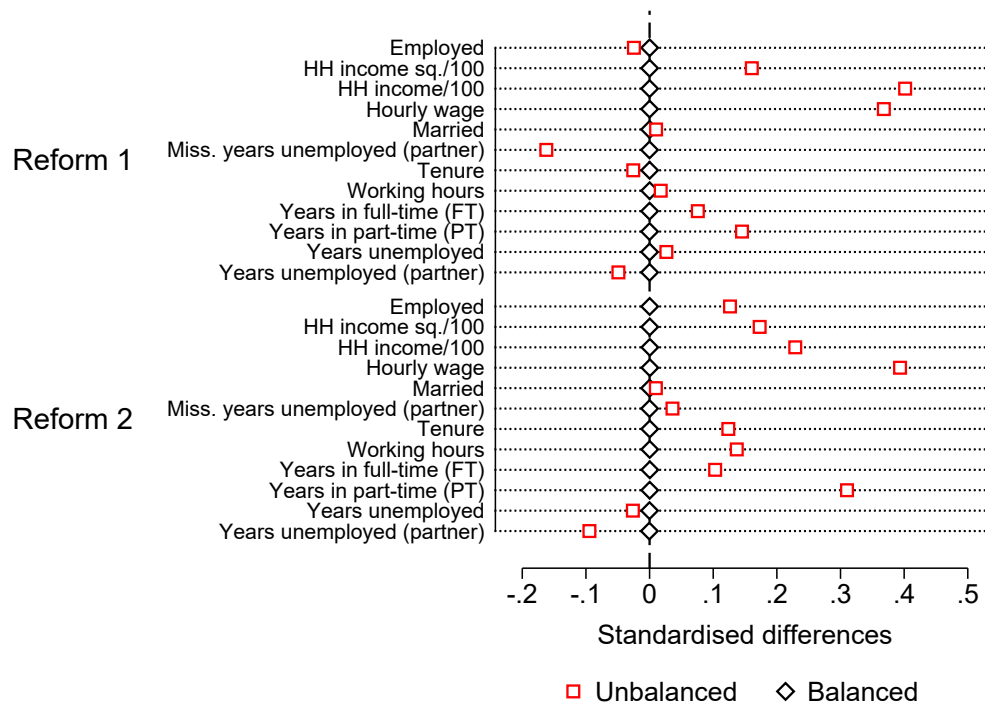


Figure 4A: Standardised differences between the means for treated and control-group women, by reform group. “Balanced” differences are obtained re-weighting via entropy balancing (SOEP 1985-2014).