

WOMEN WITH CHILDREN FIRST? PARENTHOOD, POLICIES, AND GENDER
GAPS IN THREE EUROPEAN LABOUR MARKETS

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan Tilburg University op gezag van prof. dr. G.M. Duijsters, als tijdelijk waarnemer van de functie rector magnificus en uit dien hoofde vervangend voorzitter van het college voor promoties, en Universiteit van Trento op gezag van de rector, prof. dr. P. Collini, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de Aula van de Universiteit op dinsdag 25 juni 2019 om 16.00 uur door

Gabriele Mari
geboren te Mirandola, Italië

PROMOTIECOMMISSIE

Promotores prof. dr. R. J. A. Muffels
 prof. dr. P. Barbieri

Copromotor dr. A. R. C. M. Luijkx

Overige Leden prof. dr. M. Gangl
 prof. dr. S. Harkness
 prof. dr. R. Keizer
 dr. C. L. Dewilde
 dr. G. A. Veltri

Contents

	Page
Introduction	6
Chapter 1 – Is There a Fatherhood Bonus?	51
Chapter 2 – Do Parental Leaves Make the Motherhood Wage Penalty Worse?	93
Chapter 3 – Policy, Compensating Differentials, and Gender Career Gaps	139
Chapter 4 – Gender, Parenthood, and Hiring Decisions in Sex-Typical Jobs	168
References	200
Summary	232
Acknowledgements	235
About the author	236

To those who care

Introduction

In August 2017, Jacinda Ardern, then Labour candidate for prime minister of New Zealand, faced this question on a radio show¹:

If you are the employer of a company you need to know that type of thing from the woman that you are employing because legally, you have to give them maternity leave. So therefore the question is, is it OK for a PM to take maternity leave while in office?

One year later, prime minister Ardern proved it OK returning to work six weeks after giving birth to her first child². Yet, to this day, parenthood puts women at a crossroads between career and family in high-income countries and beyond. Once confined to ‘home-making’, women have made strides in labour markets, taking up paid work at higher rates across cohorts and more continuously along the lifecycle. Gender gaps in pay have shrank and women and men distribute more evenly across jobs, albeit only up to a point. And while family formation, on the other hand, remains a key life-course transition for many, parenthood is increasingly postponed when not forgone altogether. More and more, childbearing and child rearing are carried out independently from marriage ties, by two parents engaged in both paid and unpaid work or in single-headed households where one parent does it all.

Amid these secular trends in labour markets and family life, parenthood divides today the careers of women and men. A *family gap* features contemporary labour markets, as women pay economic and career prices for motherhood while men’s career progression marches on come fatherhood. Gender inequality in paid work persists despite institutional change aimed at mitigating it or curbing it altogether. Labour market and welfare institutions have variously departed from the *family wage* model once supporting male breadwinning through secure, well-paid employment, surrounded by social protections. In particular, the United Kingdom, Germany, and the Netherlands drifted away from this family wage model in recent decades. Two main institutional changes have marked this transition, namely the expansion of family leave rights and the flexibilisation of employment relationships. Beyond their common features, however, policy trajectories have diverged in the three countries and so have their consequences for the family gap and gender inequality more broadly.

Hence, throughout this dissertation I ask how the family gap has shaped in the midst of akin and yet distinct changes in the labour market and welfare institutions formerly devoted to the family wage principle in the UK, Germany, and the Netherlands. By highlighting progress and stall in the ways these three countries came to modify their male breadwinner order, my main tenet is that policies aimed at women and families are not, by default, women- or family-friendly. The family gap, I will argue, is often the unintended or perverse by-product of gradual and selective institutional change.

In this Introduction I develop my argument in four steps. In a first section I define the main contours of the family gap in the context of changing labour markets and changing families. I then ask “What has parenthood got to do with your career?”, framing the question in causal terms and bringing to the fore the main empirical challenges that arise when answering such a question. In a third step I devote my attention to those institutions that have departed from the family wage model in the three countries under consideration. Fourth and last, I discuss the mechanisms producing the family gap with an emphasis on their ties to the aforementioned institutional change. Throughout, I highlight the goals and contributions of each chapter, re-affirming them together with directions for future research in the concluding sections.

1. Gender, parenthood, and labour markets

1.1. Gender inequality in labour markets: progress and stall

In high-income countries women and men work for pay at rates more similar today than ever before (Ahn and Mira, 2002; Charles, 2011). Figure 1 plots trends over the last 40 years (1977-2017). Even if at the tail end of a secular trend, the OECD average female labour force participation (FLMP) rate is now 14 percentage points higher than it was in 1977. Of the three countries of interest here, the Netherlands features the most dramatic rise in FLMP, from 32.5% in 1977 to 75.2% in 2017. Germany displays a steady growth in FLMP³ as well, with a 24 percentage-point difference in the period considered. FLMP in the UK has risen to similar levels despite a higher starting point back in the 1980s. Notably, convergence between the three countries under study is evident since the late 2000s.

Beyond overall rates, female labour supply has become more continuous over the life-

cycle across cohorts (Fouarge et al., 2010; Goldin and Mitchell, 2017). Compared to older cohorts of women the typical life-cycle pattern comprises today later labour market entry due to prolonged education, an evident yet smoother dip during childbearing years, and sustained participation in later life. Both rising participation rates and continuity in paid work have been aided by the availability of part-time work, especially in the countries being considered (Gregg et al., 2007; Euwals et al., 2011; Trappe et al., 2015). Germany, the UK, and the Netherlands indeed share a pattern of maternal part-time work, with women working part-time disproportionately after the birth of a child (e.g. Anxo et al., 2007). Gaps of around 20 percentage points exist between the part-time employment rate of mothers of young children and the same rate for childless women in all three countries (OECD, 2013a: 163).

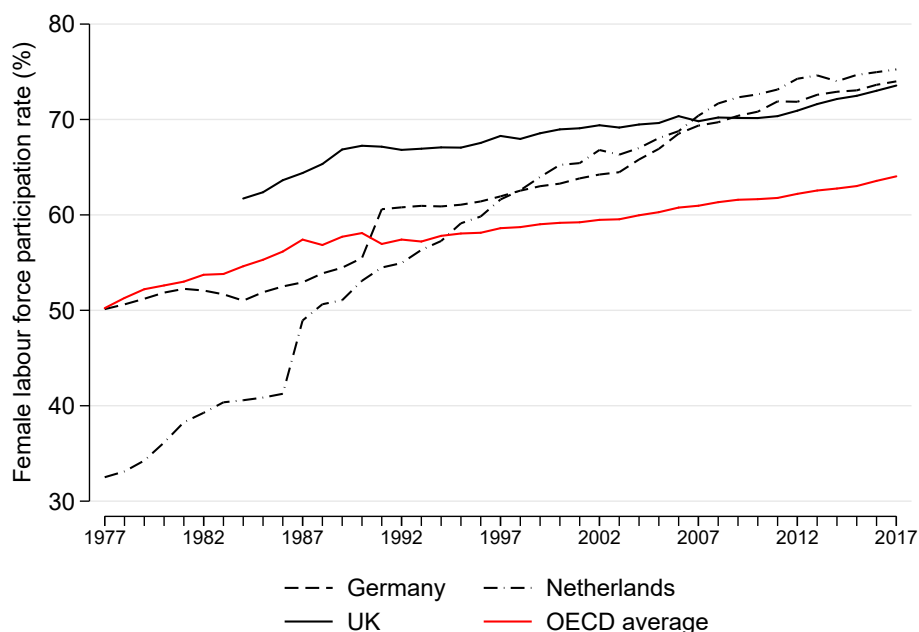


Figure 1: Female labour force participation rates in selected countries and in the OECD area. Prior to 1989, data for Germany refers to West Germany. *Source:* OECD Statistics, <https://stats.oecd.org/>.

Moving on from labour supply to earnings, progress has stalled in recent decades across OECD countries. In the US, for example, rising FLMP coincided with a fast-pace shrinkage in the pay gap in the 1970s and 1980s, followed by a slowdown in the 1990s and 2000s, with the gap halting at around 17% in 2014⁴ (Blau and Kahn, 2000, 2006, 2017).

For a comparison, Figure 2 plots available OECD data for gender gaps in gross earnings at the median in the UK, Germany, and the Netherlands. In the UK, the gender gap decreased by roughly 11 percentage points in the 1970s, from around 47% in 1970 to 36% in 1980. By contrast, in the last decade or so, the gap reduced from around 21% in 2006 to 17% in 2016, a 4 percentage-point reduction. Similar gap levels and slow-pace change feature both contemporary Germany and the Netherlands, albeit shorter time-series are available especially for the latter country.

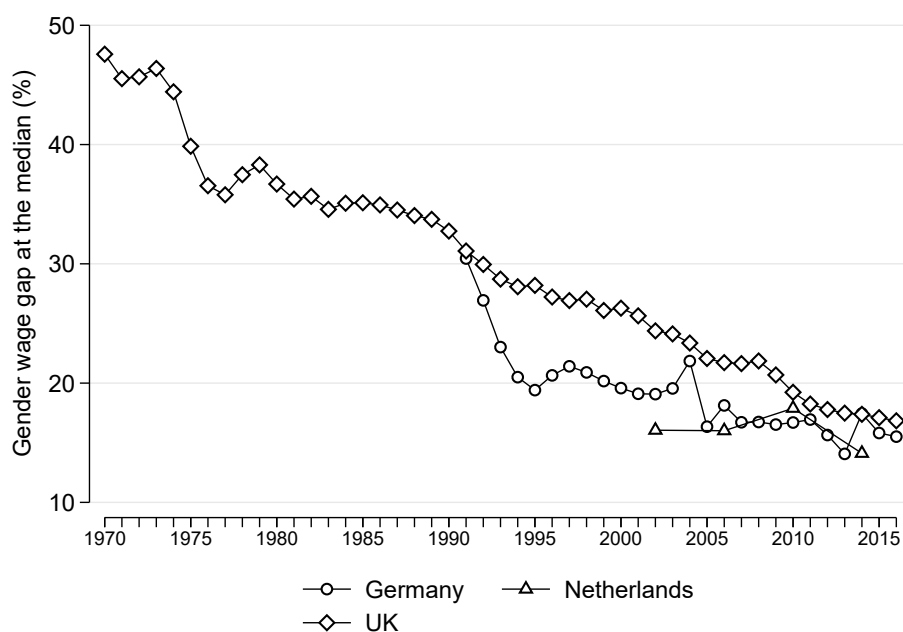


Figure 2: Unadjusted gender wage gaps at the median in selected countries. Data refer to full-time employees and to the self-employed. Prior to 1989, data for Germany refers to West Germany. *Source*: OECD Statistics, <https://stats.oecd.org/>.

Notably, turning from variation over calendar time to variation over the lifecycle, studies have reliably shown that the gender wage gap widens particularly through the early career, and specifically around the time of first childbirth (cf. Loprest, 1992; Kunze, 2005; Manning and Swaffield, 2008; Napari, 2009; Bertrand et al., 2010; Del Bono and Vuri, 2011; Goldin et al., 2017; Adda et al., 2017; Francesconi and Parey, 2018).

Underlying the raw gap and its evolution over both calendar time and the lifecycle are multiple factors (for reviews, Blau and Kahn, 2017; Ponthieux and Meurs, 2015; Kunze, 2018). For one, as the gender gap in educational attainment reversed in recent decades

(e.g. Goldin, 2006), gender differences in educational level between women and men now contribute to the gap to the *benefit* of women. More continuous labour supply during the lifecycle has also translated in a reduction of the portion of the gap explained by gender differences in work experience (see also O'Neill and Polachek, 1993; Bar-Haim et al., 2018).

What lies behind the gap in high-income countries today then? Labour market participation remains selective among women and, on average, those with better earnings potential disproportionately fill the ranks of the workforce. Such positive selection leads to the underestimation of gender wage gaps and in some high-income countries (take Italy, for one) severely so (Olivetti and Petrongolo, 2008). Keeping this caveat in mind, an inevitably partial list of causes includes: (i) segregation by field of study, and consequently, by occupation (e.g. Murphy and Oesch, 2015; Adda et al., 2017; Francesconi and Parey, 2018); (ii) differences in working-time arrangements (Triventi, 2013; Cha and Weeden, 2014) and in how working time is rewarded across different occupations (Goldin, 2014); (iii) differential sorting across firms and, thus, within-occupation differentials (e.g. Bayard et al., 2003; Card et al., 2016); (iv) related, gender differences in job search, job mobility, and their returns (e.g. Hirsch et al., 2010; Kunze and Troske, 2012); (v) gender (sex) differences in psychological traits (for a review, Blau and Kahn, 2017); (vi) employer discrimination (cf., for example, Hellerstein et al., 1999; Gayle and Golan, 2012; Lesner, 2018; Charles et al., 2018).

Adjudicating the relative weight of each of these mechanisms is beyond the scope of this dissertation. Rather, my aim is to highlight how some of these mechanisms are triggered by one life event, the transition to parenthood.

1.2. Changing families, changing parenthood

The transition to parenthood remains a key life-course stage for many. Yet in the midst of secular fertility decline and rising childlessness (e.g. Ahn and Mira, 2002; Kreyenfeld and Konietzka, 2017), families are changing in high-income countries. I deem two of these changes crucial for my purposes in this dissertation: the postponement of parenthood and the increasingly complex entanglement of parenthood and couple formation.

When not forgone altogether, parenthood is more and more experienced later in life in

OECD countries. With few exceptions, such as the US, the transition to parenthood has been postponed by an average of 4 years between 1970 and the late 2000s (Gustafsson, 2001; Mills et al., 2011). Table 1 focuses on the UK, Germany, and the Netherlands. Postponement is evident in all three countries, with mean age at first birth now surpassing age 29 in former West Germany and the Netherlands, and age 27 in the UK⁵ and former East Germany. Postponement is thus spreading and unequally so, as tertiary educated women and men are by and large leading the way (Nicoletti and Tanturri, 2008).

Table 1: Mean age at first birth for women in selected countries and periods. *Sources:* Gustafsson, 2001; Mills et al., 2011, OECD Family Database, and Human Fertility Database.

	1970	1980	1990	2000	2010	$\Delta_{2010-1970}$
United Kingdom	23.5	24.5	25.5	27.1	27.8	4.3
Germany (West)	23.8	25.0	26.3	-	29.0	5.2
Germany (East)	22.5	22.2	22.7	-	27.2	4.7
Netherlands	24.7	25.7	27.4	28.6	29.2	4.5

Later life means later career stage. Parenthood, now more frequently than ever, may thus come at a time when labour market careers are consolidating via human capital accumulation (e.g. Lagakos et al., 2018) and voluntary mobility to better-paying jobs (e.g. Schmelzer and Ramos, 2015). It follows, first, that career attainment might predate parenthood as much as it may follow it. An average man, for example, might already be climbing the wage ladder prior to the transition to fatherhood and, as such, it is an empirical question whether fatherhood itself boosts a man’s wages or rather the other way round. Second and related, individuals who postpone parenthood may gain a relative advantage in the labour market as compared to those who do not postpone or do not as much. As mentioned, cohort and skill are key dimensions in stratifying parenthood postponement. Highly-educated women and men, especially in younger cohorts, have children later, while early parenthood is often a marker of social disadvantage (McMunn et al., 2015; Struffolino et al., 2016). Heterogeneity in the career effects of parenthood, consistent with these patterns, can be expected (see, for men, **Chapter 1**; for women, e.g. Miller, 2011).

Parenthood is not just happening later, but it is also more likely to intertwine with differentiated union histories. Although marriage remains the modal port of entry into family formation (Holland, 2017), non-marital transitions to parenthood have become increasingly common in recent decades (Perelli-Harris et al., 2010, 2012). At the same time, especially for couples in their 20s, marriages occurring right after conception but prior to first birth have remained a substantial portion (from ≈ 10 to more than 20%) of all marriages in the UK, Germany, and the Netherlands (Holland, 2017). Although less frequent, marriages occurring right after first birth have also remained quite stable across cohorts, being much more common in Germany and the UK rather than in the Netherlands (*ibidem*). It is not far-fetched, therefore, to assume that the transition to parenthood still leads to the transition to marriage, at least for a portion of the population.

Turning from union formation to its possible dissolution, life-time marriages have become more infrequent, while divorce and re-partnering are commonplace (e.g. Elzinga and Liefbroer, 2007). The share of single-headed, and overwhelmingly female-headed, households with children has been rising in recent decades, up to around 20% in Germany and 15% in the Netherlands, while stable at around 25% in the UK (Nieuwenhuis and Maldonado, 2018). What is more, the mere presence of children, as well as the couple dynamics children can trigger (for instance, in terms of a couple's division of labour), have been linked to couple's chances of divorce and to individual chances of re-partnering (for a review, Lyngstad and Jalovaara, 2010; see also, e.g. Avdic and Karimi, 2018; Di Nallo, 2018).

This multifaceted “endogeneity of family status”, as Lundberg put it (2005), makes the inquiry into the labour market effects of parenthood prone to bias. For example, consider a scenario in which couples with children are more likely to split if women do *not* specialise in unpaid work (e.g. Kalmijn et al., 2007). At any point in time, couples that “survive” vice versa happen to be, disproportionately, those who have specialised along traditional gendered lines (Becker, 1981). If my interest lies, say, in how fatherhood affects wages, and I restrict my analyses to married men, these men may also be disproportionately part of the surviving couples who have specialised. I would therefore risk to overstate the impact of specialisation and, as a result, overestimate the effect fatherhood *per se* may exert on wages (cf. Killewald and Gough, 2013).

All in all, from this brief consideration of demographic change, I derive two guidelines for

the purposes of this dissertation. One is to pay as much attention to the (labour market) processes that lead to the transition to parenthood as to what follows it. The second guideline is not to condition the study of parenthood to couple-level dynamics (e.g. the transition to marriage), for they are cyclically intertwined with parenthood itself.

1.3. Babies and careers: a (quick) review of the family gap

A wealth of research in fact points to parenthood as a key factor behind persisting gender (wage) gaps in labour markets (Angelov et al., 2016; Wilner, 2016; Kleven et al., 2017; Adda et al., 2017; Butikofer et al., 2018). In a series of seminal articles, economist Jane Waldfogel (1995; 1997; 1998b; 1998a) first introduced the expression “family gap” to indicate the observed wage losses attached to the transition to motherhood. Mothers earn lower wages with respect to what they did prior to giving birth to their offspring (a *within* component if you will, e.g. Gangl and Ziefle, 2009). Largely as a result, mothers mature an earning disadvantage compared to women who delay or forgo motherhood entirely, as well as with respect to men (a *between* component, e.g. Sigle-Rushton and Waldfogel, 2007).

Focusing on the *within* component, motherhood wage penalties have been assessed all around high-income countries and beyond (see e.g. for Germany: Kühhirt and Ludwig, 2012; the UK: Harkness, 2016; the Netherlands: De Hoon et al., 2017; the US: Budig and England, 2001; Canada: Fuller and Hirsh, 2018; Australia: Livermore et al., 2011; Denmark: Kleven et al., 2017; Sweden: Angelov et al., 2016; Norway: Cools et al., 2017; France: Lucifora et al., 2017; Switzerland: Oesch et al., 2017; Italy: Martino, 2017; Spain: Fernández-Kranz et al., 2013; China: Yu and Xie, 2018). These economic losses vary in size, ranging from modest dips of around 3% up to and exceeding 20%, in comparison to women’s earnings preceding childbirth. The wide range of the wage penalty for an average woman can be traced back to variation among countries. In the small group of countries under consideration here, for instance, harsher motherhood wage penalties are typically found in Germany vis-à-vis the UK and the Netherlands (Davies and Pierre, 2005; Gangl and Ziefle, 2009; Kühhirt and Ludwig, 2012; Harkness, 2016).

Yet differences in study design hinder credible rankings among countries. In particular, early studies may have underestimated the magnitude of the penalty by neglecting non-

linearities, i.e. that penalties could be harsher in the first years after the event and be mitigated by some wage recovery, however partial, later on. Alternatively, penalties could be long-lasting, with no rebounding as years go by since the event. Event-study designs, in which wage changes are investigated in each year relative to the timing of the transition to parenthood, have started to shed light on such dynamics (Angelov et al., 2016; Cheng, 2016; Kleven et al., 2017; **Chapter 1** and **2** in this dissertation). The case of Denmark is illustrative: studies have found small or even negligible wage penalties for Danish mothers in the short term (Simonsen and Skipper, 2006, 2012). Differently, the event-study design of Kleven and colleagues (2017) provides evidence for a sizeable and long-lasting motherhood penalty in Denmark, pulling down mothers' wages by roughly 20 percent (see also Lundborg et al., 2017). Similar discrepancies can be found for other countries and, as a result, policy implications starkly differ across studies. In a frequently cited comparative paper, Davies and Pierre (2005: 485) trace their inability to detect a wage penalty for French mothers back to the role of institutions such as “publicly funded crèches, day care institutions and after school facilities developed to enable mothers to work full-time”. In contrast, Lucifora and colleagues (2017) find a long-lasting wage drop of around 10 percent for French mothers, a result they attribute to the impact of reduction in working hours and absenteeism – thus calling out market dynamics incompatible with the presence of small children in the household.

Despite the ongoing debate on its size and dynamics, the effect motherhood exerts on the wages of women is nonetheless well established, especially as compared to other links between parenthood and labour market outcomes. First, extending the scope of Waldfogel's original definition, studies have contended that the *family gap* may also comprise a fatherhood wage premium. Opposite to the wage penalty for mothers, parenthood may further divide the careers of women and men by boosting men's wages. Fathers earn more after the transition to parenthood than they did prior to it and fathers typically out-earn childless men (e.g. Lundberg and Rose, 2000; Petersen et al., 2011; Killewald, 2013; Cooke and Fuller, 2018). Overall though, the size of such premiums rarely exceeds 1-2% and, similar to motherhood, the dynamics of the fatherhood effect (if any) have not been scrutinized in the literature yet. Also, recent studies on a closely related subject, the wage premium seemingly attached to marriage (male marital premium), have

cast doubts on whether family formation indeed causally affects men's wages. Rather, the observed premiums might be more of a statistical artifact (Killewald and Lundberg, 2017; Ludwig and Brüderl, 2018; see also Ioannidis et al., 2017). I will explore this issue further, focusing on fatherhood, in **Chapter 1**.

Closely related to wages, the family gap has also been assessed in terms of the type of jobs parents of both sexes end up doing. Motherhood penalties in particular have been assessed with respect to access to supervisory and managerial jobs (Bygren and Gähler, 2012; Kleven et al., 2017; Lucifora et al., 2017), promotion chances (Kunze, 2015), as well as occupational status more broadly (Aisenbrey et al., 2009; Abendroth et al., 2014). Becoming a father, differently, seems to spark little change in the chances men have of climbing the job ladder, although results are more of a mixed bag here (cf. Bygren and Gähler, 2012; Kleven et al., 2017; Lucifora et al., 2017).

A third arena in which a family gap may manifest is hiring chances. A seminal study combining a lab and a field experiment in the US found mothers to be the least-preferred candidate for hire in marketing and business jobs (Correll et al., 2007). Field experiments in the financial sector in France (Petit, 2007) and across a wide array of jobs in Sweden (Bygren et al., 2017) have not replicated this finding, while a vignette study in Switzerland found evidence of a motherhood penalty for women applying for a HR assistant position (Oesch et al., 2017). As for men, field experiments could not detect any employer preference for fathers over women or childless men (Correll et al., 2007; Bygren et al., 2017). Far less established than wage responses, the literature on parenthood and hiring is thus quite inconclusive and **Chapter 4** in this dissertation will provide a contribution to it.

As in Waldfogel's original proposal, hence, motherhood wage penalties still constitute the bulk of the family gap both in terms of size and reliability of the statistical finding. Additionally, a recent wave of studies suggests that the effect motherhood has on wages has strong dynamics, meaning that we are better off thinking about the motherhood wage penalty as something distributed over a woman's lifecycle rather than a simple one-off change in a woman's wage rate. Broadening the family gap literature to career outcomes other than wages, as well as to men's outcomes after parenthood, is a relatively under-developed research endeavor and consensus in many areas is still lacking.

2. What has parenthood got to do with your career? A causal question

One question predates that of why (how, where, when, or for whom) parenthood may affect the careers of women and men: does parenthood affect the careers of women and men? Investigating the causal effect of parenthood on job rewards – such as wages, access to top jobs, and so forth – requires addressing two main processes of selection by virtue of which parenthood and career opportunities may very well intertwine, yet not causally so. These two processes are selection into parenthood (or the problem of *endogenous fertility*) and selection into employment (or the problem of *endogenous sample selection bias*).

In Figure 1, I present a unified graphical representation of these theoretical hurdles by means of a Directed Acyclic Graph (DAG⁶, see e.g. Pearl, 1995; Greenland et al., 1999; Hernán et al., 2004; Elwert and Winship, 2014). In brief, nodes in the network represent observed and unobserved random variables (the latter in between brackets or denoted by the letter U), each identified by a letter. If a letter is surrounded by a box, that variable is conditioned on in the analysis from the get-go – meaning that it is either adjusted for or the analysis is carried out only among units with a particular value of that variable. Pointed arrows express direct links from causes to effects. Dashed arrows originating in U indicate that U comprises multiple variables whose interrelations are not shown in the graph (e.g. Greenland et al., 1999: 39). Error terms are typically absent in a DAG, as their inputs into each variable are assumed to be marginally independent⁷. Finally, the example is formalized with respect to wages (W_O) as the outcome variable, but could be easily transposed to other career outcomes.

2.1. Selection into parenthood

Positing a causal effect of parenthood (P) on wages (W_O) is first threatened by common causes of both the transition to parenthood and wage determination. In the graph, I include a vector U standing for such characteristics. A *back-door* path passing through U leads from the outcome W_O back to the treatment P, qualifying U as comprising common causes or confounders. Focusing for now on the left-hand side of the graph, valid causal inference thus depends on blocking this path by adjusting for U. Otherwise, causal inference may be biased by antecedent factors that select individuals into parenthood

while also influencing wages. For instance, if low-paid women are, by the same token, more likely to have children, one risks overestimating the motherhood wage penalty. If, vice versa, high-earning men are more likely to become fathers, fatherhood wage premiums may be overestimated.

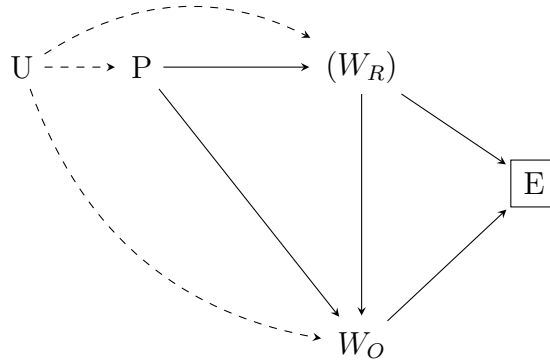


Figure 3: DAG for the causal path between parenthood P and wage offers/observed wages W_O . W_R stands for the (unobserved) reservation wage; E for employment; U for confounders in the path between P and W_O . Adopted and modified from Elwert and Winship (2014).

By and large, research in the field has developed and is still grounded on the working assumption that common causes of P and W_O are time-constant and can thereby be netted out by adjusting for individual fixed effects in panel data analysis (e.g. Budig and England, 2001; Gangl and Ziefle, 2009; Killewald, 2013). The latter operation has typically lead to the claim that negative selection operates for both women (e.g. Budig and England, 2001; Davies and Pierre, 2005; Gangl and Ziefle, 2009) and men (e.g. Lundberg and Rose, 2000; Hodges and Budig, 2010), meaning that unobserved and time-invariant factors boost the chances of becoming a parent while also depressing wages. For example, preferences morphed by socialization (e.g. Hakim, 2002; Polavieja and Platt, 2014) or education choices accounting for future earnings (e.g. Polachek, 1981) are examples of antecedents that may underlie women’s combination of family formation and low-paying (but perhaps “family-friendly”) jobs in adulthood.

Time-varying common causes of parenthood and wages, also subsumed in U , may still represent a threat to identification. Focusing on women for illustrative purposes, accounting for the time-invariant component of U takes care of the possibility that women who even-

tually have kids may typically earn lower (higher) wages than women who will eventually remaining childless. Yet, it could still be that women sort into motherhood depending not just on time-constant factors subsumed in their wage levels, but also on time-varying factors leading to their wage growth. Put differently, one should account both for *a*) the fact that parents(-to-be) may have a different earning potential than non-parents and *b*) the fact that parents(-to-be) may be on a different wage growth path than non-parents. For motherhood, a wealth of research has dealt with the possibility of time-varying endogenous fertility, either by accounting for selection on wage growth rather than on wage levels only (Loughran and Zissimopoulos, 2009; Livermore et al., 2011) or by deploying a variety of instrumental variables for the transition to motherhood⁸ (see e.g. Hotz et al., 1997; Miller, 2011; Kleven et al., 2017; Lundborg et al., 2017; Farbmacher et al., 2018). Unanimously, these studies support a causal story for the motherhood wage penalty. No comparable evidence exists for men, although in two US-based studies fatherhood wage premiums could not be detected when accounting for the differential wage growth of men eventually becoming fathers (Loughran and Zissimopoulos, 2009; Ludwig and Brüderl, 2018). **Chapter 1** applies this line of reasoning and adds new evidence on fatherhood and wages.

2.2. Selection into employment

Selection bias may also come in the form of selection into employment. At any given time point, wage offers W_O are only observed for those individuals who have accepted a wage and therefore participate in paid work (therefore $W_O \rightarrow E$). It is well established though that the transition to motherhood negatively affects the chances of accepting paid work in the market (e.g. Gutiérrez-Domènech, 2005; Fouarge et al., 2010; Fitzenberger et al., 2013; Kleven et al., 2017; Kuziemko et al., 2018). Being a common effect of W_O and P, employment is a collider in the path between P and W_O . As a result, limiting the analysis to a sample of employed individuals risks biasing estimates of the effect of P on W_O , to the point that one may retrieve a statistically reliable “effect” even in the absence of a “real” causal path leading from P to W_O (Elwert and Winship, 2014: 41).

What can be done? Motherhood is supposed here to adversely affect employment by raising a woman’s reservation wage W_R (e.g. Gronau, 1974), i.e. the minimum wage offer

that would offset the utility⁹ she enjoys outside of the labour market. It follows that conditioning on W_R would eliminate the bias¹⁰, by blocking the path $P \rightarrow E$ ¹¹.

A first strategy aims to adjust for some “unobserved propensity to work”, conceivable as a proxy of the reservation wage W_R (Heckman, 1979). In a nutshell, one should run a statistical model with E as the dependent variable, estimating women’s probability of being employed rather than non-employed. A non-selection hazard (the so called Inverse Mills Ratio) is then derived and plugged in the outcome model, i.e. the one with W_O as the dependent variable (for panel data, see e.g. Wooldridge, 1995; Semykina and Wooldridge, 2010; see also extensions to polytomous selection in Ermisch and Wright, 1993; Matteazzi et al., 2014). The outcome model is still ran on the subsample of employed women (thus conditioning on $E = 1$), but adjusting for the non-selection hazard blocks the non-causal path opened by the collider E .

Gangl and Ziefle (2009) and Livermore and colleagues (2011) have taken this route when it comes to the motherhood wage penalty. They highlight that positive selection into employment may bias upward the estimates of the motherhood wage penalty. At any time point in time, women with better earnings potentials are disproportionately represented in the sample. Failing to account for this, “conventional estimators will underestimate the motherhood wage penalty” (Gangl and Ziefle, 2009: 364; Elwert and Winship, 2014). This approach has two main pitfalls. First, it typically assumes a unique selection rule, meaning that women are assumed to be, on average, either positively selected or negatively selected into paid work. Yet, the selection of women into paid work may change over time (Mulligan and Rubinstein, 2008; Ejrnæs and Kunze, 2013) and, on top of that, multiple selection rules may co-exist at a given moment in the labour market, as shown by Neal (2004) documenting negative selection for white women and positive selection for black women in the US (see also Machado, 2017). Second, the strategy heavily relies on the use of exclusion restrictions, i.e. a set of variables that directly influence E , but not W_O . Identification in the absence of such restrictions is questionable (Puhani, 2000) and theoretical and/or statistical justifications for a given exclusion restriction are hard to come by.

Moving then to a second class of approaches to this particular selection problem requires a shift in focus. Rather than aiming at conditioning their analyses on W_R or a proxy for

it, researchers can extend their analyses to the non-employed ($E = 0$) by imputing values for the missing W_O . Sample selection bias can thus be framed as a missing data problem. Three strategies can be briefly outlined here. One is to arbitrarily assign to non-employed women a value of 0 for their missing W_O . In doing so, statistical analyses can be performed without stratifying on the sub-sample of employed women, thus purportedly solving the collider problem of Figure 1. Despite its promise, statistical estimates of the effect of $P \rightarrow W_O$ will then mix up the effects of parenthood on women’s wages and the effects of parenthood on women’s labour supply (see, e.g. Schönberg and Ludsteck, 2014; Baum and Ruhm, 2016). Besides, assigning the lowest possible wage value (0) to non-employed women clashes with the theoretical assumption, on the other hand, that women who do not participate in paid work have higher reservation wages on average than those who do participate.

A second imputation strategy is available when relying on panel data. Missing values of W_O at a given point time point t for a woman i can be substituted with valid values of W_O observed, for the same woman i , in the previous period(s) $t - k$ (for a recent application, Jee et al., 2018). Hot deck imputations of this kind have proven useful to correct year-by-year estimates of the gender wage gap (Blau and Kahn, 2006; Olivetti and Petrongolo, 2008). Yet, this strategy might only trade selection bias for another kind of bias in the case of the motherhood penalty. Since motherhood negatively affects employment, particularly in the years immediately following a child’s birth, researchers would impute for women – especially for those less prone to return to the market – their pre-pregnancy wages. As a result, for these women, there would be no within-individual variation in wages come parenthood, by design. This would automatically give rise to attenuation bias, moving (within-individual) estimates of the overall motherhood wage penalty closer to 0.

The third approach to imputation I survey here is based on matching. The key idea is that wages W_O of non-employed women (recipients) can be imputed using the observed wages W_O of “observationally equivalent” employed women (donors). Two additional assumptions need to be met: (1) assignment to the donor or recipient group should depend on a set of observable characteristics (*selection on observables*); (2) enough women with similar characteristics (collapsed in the propensity score) should exist in the two groups

(*common support*). Resting on these assumptions, the strategy would then require to “donate” the observed wages of employed women to non-employed women with similar values of the propensity score. Neuberger and colleagues (2011) apply nearest-neighbour matching to adjust estimates of the gender wage gap in Britain, finding evidence of positive selection for British women coherently with previous studies (Blundell et al., 2007; Olivetti and Petrongolo, 2008). While this imputation strategy arguably overcomes the limits of the previous two and does not rely on exclusion restrictions à la Heckman, it has yet to find application in the literature on motherhood wage penalties.

In the absence of a well-established strategy to address selection into employment, one can point out that, once again, netting out individual-level, time-invariant confounding gets us at least half the way. Per the DAG of Figure 1, selection bias can be muted by conditioning on W_R or on so called parents of W_R contained in U (e.g. Greenland et al., 1999). It follows that, at least for its time-invariant component, selection into employment is accounted for by conditioning our analysis on U (for example, by including individual fixed effects in panel data analysis). This will be my working assumption in the relevant chapters (2 and 3). I will return on this issue, particularly on the uncharted waters of selection into employment on the basis of time-varying variables, in my concluding remarks.

3. The family gap unpacked (1): institutions

The overarching question “What has parenthood got to do with your career?”, aptly dressed up in causal clothing, can now be explored further. I ask under which institutional conditions and through which mechanisms parenthood may affect the labour market careers of women and men. I discuss, first, institutional change in the former family wage models of the UK, Germany, and the Netherlands. Second, I present individual-level mechanisms that may produce the family gap and highlight how they are linked to institutional change.

Often overlooked in the family gap literature (cf. Gough and Noonan, 2013; Ponthieux and Meurs, 2015), in fact, is how individual-level mechanisms depend on the welfare and labour market institutions in which employed parents are embedded. This dependence may manifest in a twofold manner (DiPrete, 2002). First, institutions may *trigger*

life-course events such as the transition to parenthood and related mechanisms at the individual level. Second, institutions may mitigate, leave unchallenged, or rather exacerbate the *consequences* of life-course events and related mechanisms, with distinct implications for the family gap and for gender (social) inequality.

In the three countries under consideration¹², labour market and welfare institutions have long supported and enmeshed with the cultural and empirical norm of male breadwinning (Lewis, 1992; Crompton, 1999; Hobson, 2002; Gornick and Meyers, 2003; Iversen and Rosenbluth, 2010). Although never a monolith, male breadwinning had its cornerstone in the family wage ideal, prescribing that “the male head of the household would be paid a family wage, sufficient to support children and a wife and mother, who performed domestic labor without pay” (Fraser, 1994: 591). Several institutions certified or actively supported this gender order: (i) marriage bars, that is, laws and regulations prohibiting the employment or forcing the dismissal of women upon marriage, in place in specific sectors, occupations, or firms up to the second half of the 20th century (e.g. Smith, 1986; Kolinsky, 1989; see also, Goldin, 1990; Brinton et al., 1995); (ii) joint taxation of a family’s income, typically discouraging women’s participation in paid work (Cooke, 2011); (iii) social programs, such as unemployment benefits or sickness absence, tailored to “male” employment-related risks (Esping-Andersen, 1990); (iv) policies promoting, on the opposite, “female” full-time homemaking such as leaves reserved *de jure* or *de facto* to mothers (see Table 1A in the Appendix).

It is precisely this latter pillar of the family wage ideal, that of female full-time homemaking, that I will consider here for it has eroded in recent decades in what have become the “modified” male breadwinner societies of the UK, Germany, and the Netherlands (Gregg et al., 2007; Trappe et al., 2015; Begall and Grunow, 2015). Institutional transformations, pertaining to leave rights and to employment relationships, have contributed to this erosion, albeit not homogeneously across the three countries.

3.1. Leave rights

Leave rights grant individuals the opportunity to take time off from paid work and prepare or recover from childbirth, take care of newborns and infants, or assist ill family members. Leave programs vary widely in terms of eligibility criteria, duration, financing, and benefit

levels and structure, as well as in the take-up behaviour of recipients. Statutory national paid maternity leaves are now widespread all over the world, the US and Papua New Guinea being the only exceptions. Paternity, parental, and family leave provisions are common in a smaller yet growing number of countries, although overwhelmingly high-income ones (Rossin-Slater, 2017).

Table 1A in the Appendix offers an overview of how maternity, paternity, and parental leaves have changed in the UK, Germany, and the Netherlands. In all three countries, paid maternity provisions date back to the 1960s and 1970s, replacing ‘work restrictions’ for childbearing women put in place already at the turn of the 20th century (CESifo database, <https://bit.ly/2zTwwvV>). Following the example, however varied, of Nordic countries (Eydal and Rostgaard, 2016) and the input of EU directives (e.g. Lewis, 2002), leave rights have also been extended to fathers in the 1990s, and more decisively in the 2000s.

The UK and the Netherlands have taken a similar route in many respects. For one, eligibility for maternity leave in the UK, and for parental leave in both countries, has been tied to parents’ involvement in paid work, in terms of tenure and working hours (Burgess et al., 2008; Begall and Grunow, 2015). With respect to generosity, benefits for both programs rank in the lower tier among OECD countries (Ray et al., 2010). Parental leave, in particular, is currently unpaid in the Netherlands and paid only in part in the UK under the new Shared Parental Leave scheme (2015). Parental leave uptake is far from universal among women and rather low for men in both countries, and higher among high-income/highly educated parents anyway (Huerta et al., 2014; Blum et al., 2018). By contrast, more than 80% of new fathers in both countries use paternity leaves nowadays (*ibidem*). Introduced in the 2000s, paternity leave in the Netherlands has recently being expanded to five days at full wage replacement (up from the two granted since 2001), while in the UK it spans two weeks at a flat-rate payment. Notably, while paternity leave was introduced in the UK, maternity leave was simultaneously extended, perhaps signalling contradictory commitments to changing gendered work-family reconciliation (Baird and O’Brien, 2015).

Both the UK and the Netherlands have thus taken the approach of a “cost-efficient minimum standard of family policy” (Begall and Grunow, 2015: 698; Baird and O’Brien,

2015) when it comes to leave rights, particularly those of fathers¹³. Germany, on the other hand, has reinforced and then challenged the female home-making principle the most via its reforms to leave provisions over time. German parental leave, first instituted in 1984 in the former West, has long been among the longest and most generous in comparative perspective (Ray et al., 2010). After a series of reforms culminated in 1992-1993, parents could access up to 24 months of benefit, paid irrespective of employment status, and 36 months of job guaranteed leave (e.g. Ziefle and Gangl, 2014). German mothers, especially in the former West, were taking the longest career breaks among the countries of interest here (Fouarge et al., 2010), largely as a result of the design of parental leave itself (Schönberg and Ludsteck, 2014; Ziefle and Gangl, 2014). Formally entitled, men were instead exploiting leave provisions only in single-digit shares for much of the 1990s and early 2000s (Bünning, 2015).

Things changed in the 2000s. With the double aim of facilitating women's quicker return to paid work after childbirth and of contrasting Germany's fertility decline, a new parental leave benefit came into effect in 2007. Now earnings-related, and thus more advantageous for employed parents, the new benefit spans only 12 months, or 14 if each parent takes at least two months. Women's leave interruptions shortened, falling in line with a new 12 months norm (e.g. Ziefle and Gangl, 2014; **Chapter 2**), and fathers started chipping in at higher rates. Today, over 30% of German fathers go on parental leave, for an average of two months (e.g. Bünning, 2015). For a comparison, both figures resemble those of parental leave uptake among Swedish fathers (Albrecht et al., 2015).

Overall, institutional change in the realm of leave rights has challenged the female full-time home-making ideal in all three countries. Maternity and parental leave mandates, when of moderate length, help women stay in paid work (e.g. Ruhm, 1998; Rønsen and Sundström, 2002; Gregg et al., 2007; Baker and Milligan, 2008; Lalive and Zweimüller, 2009; Kluge and Schmitz, 2018). The extension of leave rights to men has been heterogeneous across countries (see also Ray et al., 2010; OECD, 2017), but may have increased fathers' involvement in childcare and further freed up women to participate in paid work (Tanaka and Waldfogel, 2007; Huerta et al., 2014; Tamm, 2018). How these changes can in turn impact the family gap, starting from the motherhood wage penalty, is much more ambiguous (Waldfogel, 1998a; Schönberg and Ludsteck, 2014; Baum and Ruhm, 2016;

Andersen, 2018). I contribute in this respect by highlighting, in this Introduction, which relevant mechanisms parental leave regulations may trigger, to the benefit or detriment of the earnings of mothers (fathers). In **Chapter 2**, I tackle the issue empirically and exploit the ‘German experience’ as an example of how leave rights can shape the family gap.

3.2. Flexible employment

Female full-time home-making has also eroded thanks to the flexibilization of employment relationships. Temporal flexibility in particular has been on the rise in labour markets, as flexible working-time arrangements have become common alternatives to the standard full-time work schedule. Women’s lifetime labour market participation has been aided crucially by the availability and regulation of part-time work in all three countries (Gregg et al., 2007; Euwals et al., 2011; Trappe et al., 2015). Other flexible working-time schedules, such as flexitime or working from home, are also in high demand among employed parents (e.g. Felfe, 2012b; Bryan and Sevilla, 2017; OECD, 2017).

Nevertheless, working-time flexibility has been subject to different regulations in the UK, Germany, and the Netherlands, and the career consequences attached to holding flexible jobs have diverged. The Netherlands set up the most extensive legal entitlements surrounding working-time flexibility through the 1980s and 1990s, largely via agreements between the state, trade unions, and employer organisations (not to mention impulse from the EU, Visser, 2002). Over time, part-time work has become not just the prerogative of married women with children, but also increasingly common among women without children (Bosch et al., 2010) and more widespread among men than in all other OECD countries (OECD, 2013b). What is more, in the Netherlands, women in part-time jobs have similar chances to receive firm-sponsored training (Picchio and van Ours, 2016) and experience comparatively small wage penalties (Fouarge and Muffels, 2009) vis-à-vis their full-time counterparts, while also reporting high levels of job satisfaction (Booth and van Ours, 2013).

In Germany, part-time work has long been the modal port of re-entry into the labour market for women after childbirth, especially in the former West (Trappe et al., 2015; Dieckhoff et al., 2016). Parental leave reforms over the years have fostered moves to

part-time jobs, but importantly did so by encouraging mobility within rather than across employers (Schönberg and Ludsteck, 2014; Kluge and Schmitz, 2018). Since 2001, for example, mothers (parents) have been granted the right to work up to 30 hours a week while on the aforementioned job-protected parental leave. At the same time though, labour market reforms ('Hartz' reforms, 2003-2005) promoted work arrangements combining temporal flexibility and lax employment and social protections, with the aim of integrating women and other outsiders into the labour market (Palier and Thelen, 2010; Biegert, 2014). Part-time and marginal employment are more and more synonym of low-wage employment, contributing to rising in-work poverty and wage inequality in Germany (e.g. Brülle et al., 2018).

For women in particular, the career costs of part-time work are more ambiguous in Germany than in the Netherlands. Working part-time generates lower returns to experience and, thereby, lower wage growth (Paul, 2016). German women working part-time suffer a wage penalty when compared to their full-time counterparts, but this gap is relatively modest in international comparison (Bardasi and Gornick, 2008). Previous research has also suggested that switches to part-time jobs after childbirth are not responsible for the motherhood wage penalty experienced by German women (Gangl and Ziefle, 2009; Kühhirt and Ludwig, 2012; cf. **Chapter 2**).

Women work disproportionately part-time after childbirth in the UK as well (e.g. Paull, 2008). For the most part, a *laissez-faire* principle has driven Britain's approach to the regulation of part-time work (e.g. Rubery, 2011), and the economic and career costs of part-time are more clear-cut in the UK than elsewhere. Part-timers receive less training and lower pay than comparable full-timers (Arulampalam and Booth, 1998; Manning and Petrongolo, 2008), and are both horizontally and vertically segregated (Connolly and Gregory, 2008; Matteazzi et al., 2014). Transitions to part-time jobs during the career-cycle have also been found to explain rising downward class mobility among British women (Bukodi et al., 2017). And yet, British women employed part-time are usually more satisfied with their jobs than do full-timers, and relatively more satisfied than their counterparts in European countries (Booth and Van Ours, 2008; Gallie et al., 2016).

Divergent paths aside, one common piece of institutional change in all countries has been the introduction of "right-to-request" legislation in the early 2000s. These laws pertain the

right for employees to ask changes to their current working schedule and formalise how employers should manage, accept, or refuse such requests (Hegewisch, 2005; Lewis and Campbell, 2007). While Germany and the Netherlands granted such right to all employees in firms above a certain size threshold, the UK first limited the right only to parents of small children. Such selectivity may have deepened further the negative toll working-time flexibility has on British mothers, an issue I will explore further in **Chapter 3**.

In all three countries, in sum, mothers' work is part-time work and part-time work is mothers' work. Drifting away from pure female full-time homemaking, these three countries have molded in a "one-and-a-half" arrangement where women are part-time earners and part-time carers (e.g. Crompton, 1999, 2006). The Netherlands stands out, however, in terms of regulation and equal treatment of part-time work, the UK being the polar opposite in both respects, and Germany falling somewhere in between.

3.3. 'Women-friendly' changes? Putting it all together

Having discussed relevant institutional change, I situate here my contributions with respect to the few examples in previous research that have also sought to 'unpack' how the family gap differs across institutional settings.

Comparative research has highlighted the importance of institutions by showing how the magnitude, drivers, or consequences of the family gap differ across countries belonging to different welfare regimes, gender orders, or work-family packages (e.g. Davies and Pierre, 2005; Gangl and Ziefle, 2009; Dotti Sani, 2015). Regardless of the label, a merit of this research agenda is to elaborate on and provide empirical evidence for how individual life courses are shaped in contexts featuring distinct institutional complementarities (e.g. **Chapter 1**). I move three critiques to such approach nonetheless. First, the importance of institutional arrangements is gauged only indirectly, without explicit measurement and modelling of macro-level variables. Even when measured and part of the empirical models (Abendroth et al., 2014; Budig et al., 2016), second, institutions are rarely considered as they change over time due to policy reform. Third, while work-family policies undoubtedly come in bundles and interact with each other to influence individual outcomes, the effects of single policies – as they change over time – are difficult to disentangle following this comparative approach (cfr. Aisenbrey and Fasang, 2017; Kluge and Schmitz, 2018).

In line with this critical appraisal, I pursue a complementary design aimed at assessing how a given policy in a given country evolved over time and with what consequences for the family gap (esp. **Chapter 2** and **Chapter 3**). In doing so, my goal is also to contribute to the debate on the ‘perverse effects’ of work-family policies. Well-established in both sociology and economics is the finding that work-family policies help women maintain their footing in paid work at the expense of persistent gender gaps in pay and labour market career (e.g. Ruhm, 1998; Arulampalam et al., 2007; Mandel, 2012; Blau and Kahn, 2013; Aisenbrey and Fasang, 2017). This has been variously called a ‘trade-off’ (Pettit and Hook, 2009), a ‘boomerang effect’ (Gupta et al., 2008), a ‘welfare paradox’ (Mandel and Semyonov, 2006). Similar to research on the institutional contours of the family gap, the bulk of this broader literature has focused on Nordic countries or cross-national comparisons, often resorting to cross-sectional data. My aim, by contrast, is to assess whether and how changes to policies commonly deemed women- and family-friendly have affected the family gap, within countries other than Nordic ones and over time. Examining how leave rights and the promotion of flexible employment have changed, and with what consequences, in the UK, Germany, and the Netherlands, one question leads the investigation: are all-encompassing narratives about the ‘perverse’ effects of family-friendly policies satisfactory in all contexts or, looking at the nuts and bolts of each single policy, are we bound to conclude that, in fact, “it depends” (e.g. Mun and Jung, 2018)?

4. The family gap unpacked (2): mechanisms

Keeping this question in mind, here I survey individual-level mechanisms that may underlie the family gap in labour markets. Mechanisms are here lined up on a continuum, from those whose explanatory power mainly rests on labour supply factors to the ones who insist, conversely, on the role of labour demand. Arguments based on human capital and effort considerations, presented first, mainly point to the labour supply behaviours of employed parents. Relevant behaviours, such as taking time off or adjusting one’s working hours, are expected to affect individual productivity, to which wage setting in the labour market is assumed to be fine-tuned. Signalling mechanisms, on the other hand, posit that what matters is how the labour supply behaviours of parents are read (re-

warded or sanctioned) by employers. Similarly integrating labour demand into the mix, the theory of compensating differentials shifts the focus to how family gaps may arise as a consequences of what preferences parents have in terms of job features and of employers' leverage over those job features. Finally, employer discrimination theories grant labour demand the most prominent influence, as employers are hypothesized to treat differently otherwise equal workers depending on their sex category and parental status.

Whenever pertinent, I emphasise how each of these mechanisms relate to the aforementioned institutional change.

4.1. Human capital: loss, depreciation, and returns

As per human capital theory (Mincer and Polachek, 1974; Becker, 1981), motherhood penalties in labour markets may stem from career interruptions around the time of a child's birth. Human capital loss may harm women's career as taking time off halts the accumulation of experience and job tenure, possibly resulting in the loss of training and promotion opportunities. Secondly, women's current stock of human capital may depreciate, meaning that the longer they stay off work the higher the chances their skills may become obsolete and thus less remunerative when reprising their former or a new job¹⁴. Third, if returning to work on a part-time basis, women may face the economic costs associated with the lower (returns to) human capital accumulation granted by working short hours (Fernández-Kranz et al., 2015; Paul, 2016).

A number of studies on the motherhood wage penalty highlights how losses and depreciations of, as well as lower returns to, human capital account for at least part of women's wage losses (Gupta and Smith, 2002; Anderson et al., 2002; Gangl and Ziefle, 2009; Adda et al., 2017). Also consistent with the implications of human capital theory, mothers staying with their pre-pregnancy employer – and thus retaining firm-specific skills – experience smaller wage penalties than those who switch employers (Waldfogel, 1997; Zhang, 2010; Felfe, 2012a; Fuller, 2017; see also Looze, 2014).

Career interruptions and their costs do not come about in a vacuum though. Particularly, work interruptions are influenced by family-leave policies and reforms to those policies over time (Ruhm, 1998; Rønsen and Sundström, 2002; Gregg et al., 2007; Schönberg and Ludsteck, 2014; Lalive et al., 2013; Baum and Ruhm, 2016; Kluge and Schmitz, 2018).

Yet, the link between leave policies, career interruptions, and motherhood wage penalties has been seldom teased out. Leave mandates mechanically trigger work interruptions. The length, replacement rates, and rights (to work part-time during leave, to return to one's pre-pregnancy employer) regulated by leave policies shape the consequences that career interruptions may have for motherhood penalties. The analyses of such macro-micro link are carried out in **Chapter 2**, focusing on German parental leave reforms.

4.2. Effort

A different strand of human capital theory (Becker, 1965, 1985) points to mothers' allocation of time and effort. Childless women can allocate more of their non-market time to leisure activities as compared to mothers. Heightened housework hours and childcare duties add up instead to mothers' involvement in the household (Kühhirt and Ludwig, 2012; Schober, 2013; Cooke and Baxter, 2010), consuming their energies and depressing their work effort much more than leisure would do. Wage losses in the aftermath of parenthood may thus reflect a reduction in work effort.

One key issue here is that of measurement. Scholars have commonly referred to effort as proxied by the age of the youngest child in the household (Anderson et al., 2003), by the amount of hours spent doing housework (Kühhirt and Ludwig, 2012), or measured in terms of work-life primacy, that is, the relative importance a person assigns to work over family (Evertsson, 2013; Gangl and Ziefle, 2015; Bielby and Bielby, 1984). Evidence is mixed on whether mothers' labour market outcomes rebound as children grow older (cf. Anderson et al., 2003; Kahn et al., 2014; Abendroth et al., 2014; Kunze, 2015) and household effort, consequentially, declines (e.g. Vargha et al., 2017). Housework hours have been relatively overlooked, yet seem to mediate part of the motherhood wage penalty particularly for mothers of young children, consistent with theory (Kühhirt and Ludwig, 2012).

Differently, subjective evaluations of work-life primacy have not been explicitly linked to the motherhood penalty. It is this latter measure of work effort, however, that has been shown to vary in response to institutional constraints, namely to the design of parental leave policies. Long parental leaves indeed seem to depress mothers' work effort (Evertsson, 2013; Gangl and Ziefle, 2015), leading to a re-orientation of preferences from

market to household production. Whether this macro-to-micro link has spillover effects into mothers' wages is an open question. In **Chapter 2**, I build on previous studies (Gangl and Ziefle, 2015) to ask whether German parental leave reforms in the early 1990s, by depressing work commitment, have hurt mothers' wage attainment.

What about men? Fatherhood may act as a motivating force and push men to increase their effort on the job (Townsend, 2002; Percheski and Wildeman, 2008; Petersen et al., 2011), either by working longer hours or being more productive. For the US, Killewald (2013) provides indirect support for this hypothesis when she finds that men get a fatherhood wage premium only when married and co-residing with their biological children. Stronger ties to their children and partner, she argues, may motivate men to commit more fully to breadwinning, yet the study lacks direct measures of work effort to further support this conclusion.

Turning to working hours as a proxy for effort, however, findings in the extant literature rarely corroborate the fatherhood-effort nexus. Indeed, average working hours actually decline after the transition to fatherhood in most European countries (Bünning and Pollmann-Schult, 2016). Studies for Germany, for example, have shown that becoming a father prompted an increase in working hours for men born prior to 1960, but a *decrease* in working hours for men belonging to younger cohorts, and both changes are modest in size (1 h of paid work at most, Pollmann-Schult and Reynolds, 2017). For Britain, previous studies provide little evidence that the presence of children affects men's working hours at all (Bryan, 2007; Paull, 2008; Schober, 2013). Further and differently from the US (Lundberg and Rose, 2000), men's allocation of time seems hardly affected by fatherhood in European countries, even in couples where the female partner reduces her working hours and devotes more time to housework and childcare (Schober, 2013; Kühhirt, 2012; Grunow et al., 2012). In **Chapter 1**, therefore, I will further evaluate this stance asking if fatherhood, absent any increase in work effort, indeed leads to wage premiums for men in Germany and Britain.

4.3. Signalling

In labour markets with imperfect information, employers may use employees' behaviours to better infer individual productivity (Spence, 1973). For instance, workers may signal

their ability or motivation by participating in training or higher education. In contrast, taking time off may result in adverse signaling, suggesting low commitment to the current job or to paid work in general.

Evidence in support of the signalling value of family-related career interruptions, vis-à-vis an interpretation resting on human capital, comes in three flavours. First, studies focused on human capital depreciation neglect whether or not depreciation follows a linear trend (Görlich and De Grip, 2009). Picking up on this, research has shown that the career costs of women’s work interruptions rather develop in non-monotonic fashion over interruption time (Aisenbrey et al., 2009; Buligescu et al., 2009; Evertsson and Duvander, 2011). Notably such patterns can be traced back to the design of family leave mandates, that is, taking leaves longer than the maximum duration granted by the law – or longer than the ‘norm’ (e.g. Bergemann and Riphahn, 2017) – may acquire negative signalling value and thus trigger economic costs.

A second critique moved to human capital arguments is that, in principle, depreciation should result out of a work interruption regardless of the reason behind it. Career breaks of the same length due to family reasons or unemployment should therefore give rise to comparable career costs, yet studies have found otherwise (Albrecht et al., 1999; Evertsson et al., 2016). Third and last, signalling has emerged as a powerful explanation for the wage costs of taking time out among *fathers*. Across Europe, paternity and parental leaves allow men to leave work for relatively short periods of time (Karu and Tremblay, 2018). Even in countries where men’s uptake is the highest like Sweden or present-day Germany, fathers’ time out stops at around two months on average (Bünning, 2015; Albrecht et al., 2015). With such short breaks human capital depreciation is hardly triggered, but wage losses for men after parental leave are well documented nonetheless (Albrecht et al., 1999, 2015; Evertsson, 2016). The consensus, hence, is that women who choose to stay out more than some threshold and men who choose to stay out *per se* experience career penalties as a result of adverse signalling¹⁵.

Taken all together, explanations resting on human capital accumulation, effort, and signalling are pitted against each other in **Chapter 2** as possible accounts of the wage responses to parental leave reform in Germany over the last two decades.

4.4. Compensating differentials

Focusing on the type of jobs mothers hold, the family career gap may arise because of compensating differentials. In this framework, jobs are bundles of monetary and non-monetary features such that losses in one domain are compensated by gains in another. In the parlance of the theory, jobs combine amenities (what makes a job a “good” job) and disamenities (what makes a job a “bad” job).

Originally conceived to suggest that jobs involving disamenities (e.g. physical hazards) should pay a wage premium to attract enough workers to fill them (Smith, 1776; Rosen, 1986), compensating differentials theory also highlights how workers may have a “willingness to pay”, vice versa, for job amenities. While some have argued for the “pervasive absence of compensating differentials” in labour markets when considering men and women together (Bonhomme and Jolivet, 2009), a growing body of literature highlights gender differences. Women, and not (or much more than) men, value job features like schedule flexibility and regard some working-time arrangements such as long hours as disamenities instead (Flabbi and Moro, 2012; Mas and Pallais, 2017; Wiswall and Zafar, 2017). Mothers in particular may opt¹⁶ for amenities such as flexible schedules at the cost of disamenities such as lower wages or scant chances of promotion (e.g. Filer, 1985; Felfe, 2012b). Motherhood penalties are thus manifestations of a trade-off between job features coveted by mothers themselves, like working short hours, and the career costs that may combine with such features in labour markets, such as the lower pay often associated with working short hours.

Most commonly, compensating differentials are invoked to explain women’s career hurdles in Scandinavian countries. Whether it is the glass ceiling in pay (Albrecht et al., 2003), the wage penalty upon motherhood (Simonsen and Skipper, 2006; Kleven et al., 2017), or the chances of climbing to top jobs (Kunze, 2015; Hardoy et al., 2017), scholars associate women’s career setbacks with Scandinavia’s family-friendly jobs and workplaces, yielding indirect support for the compensating differentials story.

Explicit tests of the argument only provide mixed evidence though. Notably, Felfe has shown how, within the mandates of German parental leave granting mothers continuity with their pre-pregnancy employer, German mothers may trade off pay cuts for some flex-

ible schedules (e.g. rotating shifts and working during the evenings) and not others (Felfe, 2012b). And yet, motherhood penalties, especially for mothers who switch employers, are not fully accounted by changes in work schedules (Felfe, 2012a). For Sweden, Hotz and colleagues (2017) first develop a composite index of workplace-level family-friendliness and then find the transition to motherhood to increase the chances of switching to family-friendly jobs at the expense of skill and career progression. Differently, for the US and Canada, studies have found that family-friendly job features (mainly pertaining to working-time arrangements) do not “explain away” motherhood wage penalties (Budig and England, 2001; Glauber, 2012; Fuller, 2017) suggesting limited scope for the trade-off between monetary and non-monetary job features¹⁷.

In line with the key tenets of my argument, I consider the job amenity value of flexible schedules to be context-dependent. As a *litmus test* of this proposition, in **Chapter 3** I ask if compensating differentials arise for mothers (and women more broadly) in a context such as that of the UK, where working-time flexibility and particularly part-time work is associated with dismal career prospects (Manning and Petrongolo, 2008; Connolly and Gregory, 2008; Matteazzi et al., 2014). More specifically, I ask if policies aimed at easing access to flexible schedules have helped install compensating differentials for British mothers, deepening both family- and gender- wage gaps while simultaneously increasing transitions to part-time jobs and satisfaction with these arrangements.

4.5. Employer discrimination

Two theories of employer discrimination are often mentioned and seldom tested in the literature on the family gap: statistical discrimination theories and status-characteristic theory (Correll et al., 2007; Gangl and Ziefle, 2009; Bygren et al., 2017). In reviewing them, I develop two arguments. First, statistical discrimination theories *do not* account specifically for mothers’ disadvantage, but rather apply to all women of childbearing age, while status-characteristic theory predicts motherhood penalties more specifically. Second, the incentives to discriminate statistically depend on labor market institutions and welfare institutions, as well as on occupational features. The latter are also relevant for status-characteristic theory and predictions from the two models are, by and large, complementary. As such, they will be put to test in a survey experiment on employer

discrimination in **Chapter 4**.

4.5. Statistical discrimination

Statistical discrimination theories (Phelps, 1972; Aigner and Cain, 1977) hold that 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. Under statistical discrimination, employers are rational actors aiming at maximizing expected profits in labour markets characterized by imperfect information (for a review, Fang and Moro, 2011). Specifically, employers may find it difficult or too expensive to access precise information on the individual productivity of job applicants. Group markers such as a candidate's sex are instead easily accessible – on CVs, at job interviews, etc. – and employers may thereby determine individual productivity by combining the expected productivity of a given female (male) candidate with the group-level productivity they estimate for women (men). Even if employers believe women and men to be equally productive on average, the productivity signals of a female applicant might be deemed more noisy (Aigner and Cain, 1977). Employers' underlying assumption is that, despite equal educational credentials or accumulated work experience, female employees might be more likely than men to take career breaks, reduce their working hours, or leave their job altogether for family-related reasons.

To compensate for this “risk”, employers become more reluctant to hire women and offer them lower wages than those of men, all else equal. Economic discrimination arises then as a result of overshooting, as female employees deciding against motherhood or whose productivity is not affected by the presence of children will be discriminated against. In this framework, employers are thus forward-looking and their concern lies with women potentially becoming parents and altering their labour supply in ways harmful to productivity. Women of childbearing age, irrespective of current parental status, could be discriminated against with respect to men (Gupta and Smith, 2002; Petit, 2007; Yip and Wong, 2014; Biewen and Seifert, 2016).

In a nutshell, employers insure themselves against the risk they attach to female employees of childbearing age. The estimation of such risk though might be moderated by institutional factors, such as employment protection legislation (e.g. Dieckhoff et al., 2015).

Particularly, when open-ended contracts are surrounded by stringent protections against dismissal, temporary contracts may serve as extended probationary periods to solve asymmetric information with respect to job candidates' productivity (e.g. Wang and Weiss, 1998). This may become especially valuable when deciding upon the hiring of women of childbearing age. Statistically discriminating employers can offer temporary contracts to insure themselves against the costs they expect to bear come motherhood. Additionally, waiting for a temporary contract to expire helps circumvent anti-discriminatory laws that may prohibit the dismissal of employees because of pregnancy. The key prediction is, therefore, that employers have more incentives to discriminate statistically if the job post being offered is permanent rather than temporary, as previously found for France by Petit (2007).

In **Chapter 4**, I investigate this hypothesis looking at the Dutch context. Recent labour market reforms have relaxed the level of employment protection attached to temporary contracts but left largely untouched that of permanent contracts (Mooi-Reci and Dekker, 2015), which remains above the European average in the Netherlands (OECD, 2014). I thereby expect Dutch employers to be reluctant to hire women or offer them equal starting salaries, vis-à-vis men with identical CVs, for permanent job posts.

4.5. Status-characteristic theory

According to status-characteristic theory, no matter information asymmetries and risk aversion, status beliefs may bias employers' evaluations against women and, particularly, mothers (Correll et al., 2007; Ridgeway, 2011). Status beliefs are conceived as a particular class of stereotypes (e.g. Fiske et al., 2002) by virtue of which individuals categorize members of social groups on the basis of perceived competence. Any nominal characteristic that groups together individuals in a social setting – sex, ethnicity, sexual orientation, and so forth – may become a status characteristic if actors share beliefs regarding that group's competence, further conceptualized as the sum of ability and commitment (Berger et al., 1977; Ridgeway and Correll, 2006; Mark et al., 2009). Seeking coordination with one another, individuals may activate performance expectations regarding how capable and committed others will be with respect to the task at hand, depending on the salient social memberships. Performance expectations may then drive the distribution of rewards such

that, in the hiring setting, low-status actors end up penalized in terms of hiring chances or salary offers (Correll et al., 2007; Pedulla, 2016).

Motherhood is such a status characteristic insofar as it amplifies status beliefs morphed along gender lines (Ridgeway and Correll, 2004; Correll et al., 2007). Specifically, women, especially if mothers, are perceived to be less capable than men in workplace settings (Cuddy et al., 2004; Correll et al., 2007; Thébaud, 2015). Superior ability may be granted to mothers only for tasks that involve nurturance and care (Ridgeway and Correll, 2004), in line with stereotypes broadly associating women with communion, i.e. being selfless and concerned with others, rather than with agency, i.e. being assertive and motivated to master a task (for a review, Ellemers, 2018). As for commitment, women and mothers in particular are not expected to prioritize work over family obligations nor to make sacrifices to build a career as much as men would (Correll et al., 2007).

While all of this points to motherhood penalties, in female-typical jobs where nurturing and caring skills are needed (teaching, nursing, social work etc.), proponents of status-characteristic theory suggest women and mothers might actually be highly regarded in terms of competence (Ridgeway and Correll, 2004). This proposition may help explain the large body of evidence on gender discrimination in favour of women in female-typical lines of work. In reviewing such findings, Neumark notes that they pose “a bit of a puzzle for labor economists, since our models of discrimination do not naturally predict this pattern. That said, perspectives on discrimination from other fields, such as those emphasizing norms regarding who does which job, might fit these facts better” (Neumark, 2016: 77). I test whether status-characteristic theory can account for differential treatment of mothers (women) across occupations in **Chapter 4**.

5. Content of the dissertation

In this dissertation I put together four essays on the family gap in the UK, Germany, and the Netherlands. Each chapter descends from the general framework traced in the previous sections and attempts to provide a little piece of the puzzle. Starting from **Chapter 1**, I address whether fatherhood causally affects the wages of men in the UK and Germany, relying on long-running household panel data and on new advances in linear fixed effects modelling. Relatively little evidence exists on whether fatherhood grants men

a wage premium or superior wages spur the transition to fatherhood instead. Also, most longitudinal studies have been US-based and neglected a comparative perspective that may unravel the contextual underpinnings of fatherhood wage premiums, if any. I carry out here a comparative and longitudinal analysis of how fatherhood may affect men's wages. Micro-level mechanisms supporting the idea of a wage premium - changes in work effort, couple specialization, and employer discrimination - are discussed in light of stability and changes in the institutional settings of the two modified male-breadwinner societies. Empirical evidence in this first chapter, however, cannot support the idea of a causal premium for men, even in such contexts. Rather, I highlight the role of previously neglected sources of selection into fatherhood, particularly on the basis of prior wage growth.

The main findings of this first chapter indirectly call out on the family gap literature to renew its focus on motherhood penalties, for they might be the prime and only manifestation of how parenthood deepens the labour market divide between women and men. Exploiting a quasi-experimental design and panel data, **Chapter 2** explores the causal effect of family-friendly policies on the motherhood wage penalty. Together with co-author Giorgio Cutuli, 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. 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.

Chapter 3 also zooms in on the effects of specific policy reforms, this time turning from leave rights to the flexibilisation of the employment relationship. A commonly held account of the family gap in labour markets is that mothers favour flexible working-time arrangements over career attainment, yet little is known on whether this compensating differential is shaped directly by public policy. Relying on panel data and a difference-

in-difference design, I examine the introduction of a ‘right to request’ flexible schedules for parents of small children in Britain. Fitting the theory of compensating differentials, mothers experience wage cuts combined with a reduction in working hours and accrued satisfaction with working-time arrangements. This is especially the case for mothers of children aged 0-2 at the time of the reform. Evidence also suggests that the negative economic impact of the reform might have deepened gender gaps in the British labour market. For mothers of older children though, I can only detect wage losses and not reductions in working hours or increases in satisfaction with working time.

Finally, in **Chapter 4** I turn to the causal effects parenthood may have on hiring decisions, “bringing employers back in”. Together with co-author Ruud Luijkx, we ran two survey experiments with Dutch employers to investigate hiring discrimination in sex-typical jobs. We ask if women are especially discriminated against when they have children, whether discrimination applies similarly in different occupations, and whether statistical discrimination or status-characteristic theories best account for discriminatory practices. Informed by these theories, we set up our experimental study having employers rate fictitious candidates for either a female-typical job (primary school teacher) or a male-typical job (software engineer). Employers display a slight preference for female candidates all else equal when filling a teacher post, although such bias is less strong for female applicants with children. No such ranking is found for a software engineer vacancy, nor we find different salary offers across candidates and across vacancies. Employers do not appear to favour men over women for positions likely to be on the career track, as predicted by statistical discrimination theories, nor expect women to be less capable than men, as status-characteristic theory suggests. If mothers, however, female candidates are expected to be less committed to their job and work fewer hours, especially in the teacher experiment. Differently from US-based research, such expectations seem to have small consequences for the hiring decisions and salary offers Dutch employers make in our study.

6. Concluding remarks

6.1. Mind the gap: contributions

Parenthood is key to the persistence of gender inequality in labour markets. Decades of societal and policy changes have eroded the family wage principle once underpinning the male breadwinner order across European countries. Yet it appears a new equilibrium has been achieved, one in which parenthood deepens (economic) inequalities between men and the women newly participating in labour markets. Hopeful to inform a future research programme, the contributions of this dissertation are mainly two.

First, the family gap does not come about in a vacuum. The family wage has been enforced at the intersection of state, market, and family: we meet the family gap at the same crossroads. The allocation of jobs and job rewards such as wages and promotions in labour markets is intertwined with family dynamics, and so should be their study as masterfully argued by Lundberg (2005). What I have attempted to add here though is that parents and employers are also embedded in contexts featuring distinct sets of labour market and welfare institutions. Their impact on the family gap makes such phenomenon “context-dependent” (e.g. Gangl and Ziefle, 2009).

In making this claim, however, I have proposed here an alternative and complementary approach to two broad narratives. One narrative groups contexts in more or less coherent clusters whose features and consequences for individuals along the life course are to be contrasted with one another (e.g. Esping-Andersen, 1999, 2009; Mandel and Semyonov, 2005, 2006; Mandel and Shalev, 2009; Aisenbrey and Fasang, 2017). In choosing the UK, Germany, and the Netherlands, I have considered countries that could all be grouped as ‘modified male breadwinner’ or ‘one-and-a-half’ arrangements. Yet, among them, my aim has been to highlight difference despite similarity, and similarity despite difference. Take part-time work, for one, commonly underlying the labour market integration of women in the UK and the Netherlands. Evidence in **Chapter 3** suggests a causal link between policies promoting part-time work and the labour market disadvantage experienced by British mothers. By contrast, **Chapter 4** points out that employers in the Netherlands, even if associating motherhood with part-time work and low commitment, do not impose substantial penalties on mothers at the hiring stage. And as for similarity despite

difference, while the UK and Germany pursued different policy strategies on fatherhood, I cannot detect a fatherhood wage premium in both countries (**Chapter 1**) - not unlike previous research for Nordic countries (e.g. Cools et al., 2017; Kleven et al., 2017).

A second broad narrative points to the unintended and often perverse effects of women- and family-friendly policies (e.g. Mandel and Semyonov, 2006; Gupta et al., 2008; Pettit and Hook, 2009). Fostering women's employment, these policies may end up creating labour market 'shelters' made of low paid, part-time, female-typical jobs with little career opportunities. **Chapter 2** puts to test this claim, highlighting the radical overhaul of parental leave legislation in Germany, and how family-friendly policy can at times mitigate rather than exacerbate forms of gender inequality. In line with recent contributions in sociology (Mun and Jung, 2018) and a much longer tradition in economics (e.g. Heckman et al., 1999), I thus take issue with the evaluation of a single policy as it changes over time and provide evidence for how, decomposed to its nuts and bolts, this policy variously affected the motherhood penalty.

In short, I advocate for a piecemeal evaluation approach to the study of what is, not infrequently, piecemeal reform. The trade-off is clear, the synthesis granted by a broad-scope research agenda for the variation that "single-issue" studies may unveil. In favouring the latter, my second contribution has been to bring to the fore a causal perspective with a twofold target. One is that of policy evaluation. The other centres on the causal effects parenthood may have on the labour market outcomes of women and men. My point of departure is that, as Lundberg put it, "No single econometric technique or set of techniques can 'solve' the family-work simultaneity problem" (2005: 592). Modern causal inference may help though, being a tool to "draw assumptions" rather than a set of techniques to arrive at a conclusion (e.g. Elwert and Winship, 2014). And so, for example, **Chapter 1** has wrestled with the assumption that men who eventually become fathers are on earning trajectories parallel to those of men who will remain childless. In **Chapter 2** and **Chapter 3** valid inference rests on the assumption that time-varying confounding is muted when one re-weights groups of women exposed and not exposed to a given reform. Carrying out experimental manipulation in **Chapter 4**, finally, our conclusions stem from the assumption that randomisation is effective in cancelling out any confounding behind the effect that signalling parental status may have on hiring

chances (cf. Deaton and Cartwright, 2018).

All of these assumptions, and the path to ‘solve’ them, have been comprised in the causal model purported in this Introduction and, hopefully, future research can build on it and on its previous iterations (Elwert and Winship, 2014). To be sure, transparent assumptions do not automatically make for valid conclusions¹⁸, yet they make of causal inference less of a “dirty deed” (Hernán, 2018).

6.2. Mend the gap: limits and ways forward

In this respect, a first limit of this dissertation is that, coming to its end, two causal problems - namely “the endogeneity of family status” and selection into employment - remain hard problems. Some of the chapters discuss at length these issues and provide some partial solutions. As proposed in **Chapter 1**, the attention of further research could go to still neglected processes of selection into parenthood, such as the demography of fatherhood (cf. Balbo et al., 2013). Techniques like inverse probability of treatment weighting may then hold promise to mute what spurs the transition to fatherhood and identify, on the other hand, the causal effect of fatherhood on labour market outcomes and beyond (e.g. for marriage, Mincy et al., 2009). As for selection into employment, examined more in depth in **Chapter 2**, it remains particularly elusive when it changes over life-course time. As women’s participation in paid work becomes more continuous over the lifecycle, bias deriving from non-random selection into employment may decline. Falling male activity rates in some countries (like the US, Krueger, 2017) may motivate an analysis of men’s changing selection into employment in turn. Measuring reservation wages more comprehensively, covering all employed and non-employed people, and accounting for this measure in our empirical models could solve the problem as portrayed in Figure 1. Alternatively, techniques based on matching, briefly surveyed in this Introduction, have yet to find application.

A second limit of my inquiry is that, for all the fuss made about the causal effect of parenthood on labour market outcomes, I have not looked into its heterogeneity as much. Research on how the family gap is not one but many is burgeoning. Heterogeneity has been explored not just across national contexts, but also depending on race, class, skill, marriage ties, job title, the timing of parenthood, and so forth. Here, particularly in

Chapter 1, I have argued that questioning causal assumptions underlying the effect of fatherhood (motherhood) on the labour market outcomes of men (women) necessarily predates addressing whether such causal effect, if any, varies across social groupings. Small sample sizes hindered, throughout, this latter line of research. I have focused nonetheless on how the family gap changed over time, before and after policy reform (**Chapters 2 and 3**), as well as across birth cohorts and skill groups (**Chapter 1**). In **Chapter 4**, I have examined whether signalling parental status affects hiring chances across male- and female-typical occupations, but compared only one of each kind (cf. Bygren et al., 2017). Tackling heterogeneity, studies on the family gap may in particular contribute to the broader debate on gender-and-class divides in high-income countries. Some scholars have argued that class and gender inequality add up to a seemingly zero-sum game, as some institutional settings magnify one and not the other (Cooke, 2011; Aisenbrey and Fasang, 2017). Some others have suggested that institutional change of the kind examined in these pages has benefited middle classes the most, for they are the ones who have demanded and obtained work-family reconciliation (Moen, 2011; Warren, 2015). Others even have posited, on the opposite, that family-friendly policies hurt the hardest the career attainment of women in advantageous class positions (Mandel and Semyonov, 2006; Mandel, 2012). For future contributions to this debate, the blueprint provided here is one that focuses on single policies, prior and complementary to painting the big picture. Asking if and how family gaps contribute to gender-and-class divides may fit smoothly within the framework followed here (e.g. Kluge and Schmitz, 2018).

Finally, I have contended that institutional transformations pertaining to leave rights and to employment relationships are factors not just in the transition out of the family wage, but in the one “into” the family gap as well. Not only institutional change has been heterogeneous among countries formerly devoted to the family wage, but it has been contradictory too and further research could highlight the consequences of such contradictions. For one, changes to parental leave legislation, as well as to the provision of public childcare (Zoch and Hondralis, 2017), have put Germany in a better position to sustain universal breadwinning. Perhaps surprisingly then, Germany also introduced a cash-for-care program in 2013, subsidizing childcare for those parents (mothers) that choose not to enroll children to childcare facilities. Such a policy, inspired by a “caregiver

parity” principle (Fraser, 1994), follows the footsteps of Nordic countries too, inheriting their dilemmas in the process: while it appears to be in high demand among parents, it may depress the labour supply of mothers especially in low-income households (Sipila et al., 2010).

On a similar note, will the expansion of leave rights to fathers bring about fatherhood wage penalties, Nordic-style too (Rege and Solli, 2013; Albrecht et al., 2015)? Will parental leave combining job protection with generous benefits of moderate length help curb the motherhood penalty (**Chapter 2**)? What about a new class of policies, austerity measures on the one hand and policies mandating gender equality on the other (board quotas, pay equality, etc.), how will they impact the family gap? For these and more questions, I hope this thesis has provided useful starting points for future research design. Evidence amassed in this thesis and this body of literature dissects the contours of the family gap as a market failure. As such, what to do about it is largely up to governments (Atkinson, 2015). Mind the gap, and mend it too.

Notes

¹*The Guardian*, 19 January 2018 (<https://bit.ly/20mglws>).

²*The Guardian*, 2 August 2018 (<https://bit.ly/20X0x4n>).

³Part of the increase displayed in Figure 1 is artificially due to German re-unification. A discontinuity in the trend is in fact visible between 1989 and 1990 and likely due to the much higher FLMP in East Germany now taking part in the estimates. More recently, the rise in FLMP is rather due to changing patterns of female and, especially, maternal labour supply in West Germany (Trappe et al., 2015).

⁴This figure refers to the gap between the average weekly earnings of men and women working full-time in 2014 (Blau and Kahn, 2017).

⁵Data from the latest available time-point (2016) suggests a further acceleration for the UK. Mean age at first birth in the UK now reaches 28.9 years (OECD Family Database).

⁶DAGs are graphical models encompassing all the assumptions underlying the causal effect of a treatment X on an outcome Y . Now a standard item in the toolkit of causal inference, DAGs do not just help in formalizing a causal question but also provide theory and rules to address the “three horsemen” of causal relationships, namely confounders, mediators, and colliders (see Pearl, 1995; Pearl et al., 2016; Morgan and Winship, 2007; Hernan and Robins, forthcoming).

⁷In their DAG for the effect of motherhood on wages, Elwert and Winship (2014) suggest that determinants of wages, exogenous to the transition to parenthood and subsumed in an error term ϵ , may bias the estimation of the motherhood effect regardless of whether W_R is measured and controlled for. The authors, however, do not provide examples of what variables might be omitted here and that would thus end up in the error term. For simplicity, I omit this element from the DAG in Figure 2.

⁸Usual disclaimers on the availability and credibility of instruments apply (see Wilde et al., 2010; Farbmacher et al., 2018; Bhalotra and Clarke, 2018 for discussions of the case at hand).

⁹Such utility comprises the amenity value women may derive from children, leisure, or from other sources of income, such as the partner’s income.

¹⁰It is useful here to clarify that, following DAG theory, conditioning on a confounder may *close* non-causal paths between treatment and outcome, whereas conditioning on a collider may *open* non-causal paths between treatment and outcome (see e.g. Elwert and Winship, 2014).

¹¹This conclusion derived from the causal graph depicted in Figure 1 is exactly the same as Heckman’s econometric framing of selection bias as a “specification error” (Heckman, 1979), i.e. sample selection (collider) bias stems from omitted variable bias! Unfortunately, reservation wages often remain unobserved. In two of the main data sources used in this dissertation (BHPS and G-SOEP), for instance, respondents are only asked to estimate their (gross) reservation wage if also reporting to be ready to accept paid work. This arguably leaves out individuals with the lowest propensity to participate in paid work, that is, precisely those that might differ the most from currently employed individuals.

¹²This is not the case, of course, for former East Germany whose institutions have long encouraged

women's roles as workers *and* mothers (e.g. Trappe et al., 2015). Since my main concern is with institutional change after German re-unification, I will not systematically delve into the East-West divide.

¹³Although not surveyed here, similarities can also be traced with respect to the provision of child-care services in both countries (e.g. Yerkes and Javornik, 2018) and looking at tax policies directed to parents, as both countries have implemented earned income tax credits with the goal of stimulating the participation in paid work of (lone) mothers (with mixed results, Francesconi and Van der Klaauw, 2007; Francesconi et al., 2009; Bettendorf et al., 2014).

¹⁴An implication of this is that women who anticipate family formation may choose occupations in which human capital depreciates at slow(er) rates and career interruptions are thus less costly (Polachek, 1981; Görlich and De Grip, 2009; Polavieja, 2012). From this lifecycle perspective, the 'career costs of children' would then comprise the initial occupational sorting operated by women. In other words, the fact that women typically sort into low-paid occupations (e.g. Murphy and Oesch, 2015) should be considered as an indirect component of the family pay gap. Throughout this dissertation, however, I focus on the costs attached to the *transition* to parenthood, i.e. how career attainment deviates as a result of parenthood. Even Adda and colleagues (2017), who provided perhaps the most comprehensive treatment and support for Polachek's model, show that part of the income costs of motherhood arise *following* the birth of a child, in addition to the costs (forgone earnings) attached to human capital investments anticipating motherhood (see, by contrast, Kuziemko et al., 2018 for a competing model and evidence supporting a scenario in which women make education choices under *uncertainty*, and not perfect information, and the employment costs of motherhood are by and large *unanticipated* by young women).

¹⁵In a sense, signalling also rests on the concept of work effort/commitment. What may trigger adverse signalling though is employers' expectation of lower commitment on the part of working parents, regardless of whether such reductions in work commitment are factual or not. A similar argument is put forward within the framework of status-characteristic theory (Ridgeway and Correll, 2004; Correll et al., 2007).

¹⁶Importantly, compensating differentials arise come motherhood. In other words, differently from Polachek's model (1981), women sort into family-friendly jobs after becoming a mother rather than picking a family-friendly line of work for their future when making their initial human capital investment (say, when choosing a field of study).

Compensating differentials are closer to Hakim's concept of *adaptiveness* (Hakim, 2000, 2002), the idea that some women may adapt their preferences during their lifecycle and end up combining work and family rather than committing fully to either a career or to home-making. Differently from preference theory though, compensating differentials explicitly deal with the career costs that come with opting for family-friendly jobs. Also, the focus is on the measurement of job features (conceivable as "revealed preferences", e.g. Hotz et al., 2017) rather than that of women's preferences. The former are easily available across the waves of panel data used in this dissertation and my operationalization choices can

build on an internally consistent body of research (e.g. Smith, 1979; Rosen, 1986; Bonhomme and Jolivet, 2009). Measures of women’s preferences, instead, are less readily constructed, especially considering that concepts such as preferences, attitudes, commitment, expectations, or work orientations are sometimes used interchangeably and sometimes analytically distinguished in the extant literature (e.g. Hakim, 2002; Kan, 2007; Kahn et al., 2014; García-Manglano, 2015; Gangl and Ziefle, 2015). For its superior scope and precision, the theory of compensating differentials is here preferred to Hakim’s framework.

¹⁷Two recent studies, one for the US (Yu and Kuo, 2017) and one for Canada (Fuller and Hirsh, 2018), pursue a different route to the assessment of compensating differentials. They look at whether the extent of the motherhood wage penalty differs depending on mothers’ job characteristics. In both studies, penalties are smaller or not detectable in jobs featuring family-friendly schedules, a piece of evidence interpreted as inconsistent with compensating differentials and rather pointing to a “work-life facilitation” function of flexible working-time arrangements.

There might be a number of inter-related conceptual and empirical problems with such a strategy. Job features become a moderator rather than a mediator in the analysis of the penalty, i.e. a variable conditioned on to examine the heterogeneity of motherhood effects rather than one that lies in the causal path between motherhood and wages. Conceptually though, compensating differentials theory frames working-time flexibility as a *mediator* of motherhood wage penalties. One expects mothers to pay a price for switching to flexible yet less lucrative job posts. Further, if mothers indeed switch jobs to achieve their preferred balance between monetary and non-monetary features, it becomes tricky to assess when, relative to the transition to motherhood, one should measure those job features then deployed as moderators. Related, the fact that women stay or switch across jobs with different attributes raises the threat of selection bias, seldom addressed by the literature at large. For instance, women with higher earnings potential may be better poised to “purchase” family-friendly working-time arrangements in the firm or in the market, creating a positive, and yet spurious, association between earnings and job amenities (see Heywood et al., 2007; **Chapter 3** in this dissertation).

¹⁸No doubt, internal validity was favoured throughout this dissertation (for a more unified approach to validity, Westreich et al., 2018). The reasons are threefold (Athey and Imbens, 2017). First, causal inference that privileges internal validity already incorporates concerns over external validity when it investigates treatment effect heterogeneity (e.g. **Chapter 1**) and if it aims at replication in different settings (e.g. Nosek et al., 2015). Second, generalising findings that lack internal validity risks to resolve into a ‘garbage-in-garbage-out’ model of evidence accumulation (King et al., 1994; Shadish et al., 2002). Third and last, random samples from multiple contexts are no guarantee for external validity given how common modelling techniques end up circumscribing estimates to smaller, often non-representative and perniciously weighted, ‘effective samples’ (Aronow and Samii, 2016).

Table 1A: Chronology of main changes to family-leave arrangements in the UK, Germany (DE), and the Netherlands (NL), from the 1980s to the present day.

	1980s	1990s	2000s - present day
UK	<p><i>Maternity leave:</i> job guarantee up to 29 weeks after delivery (40 in total); only entitled if</p> <p>a) ≥ 2 years of tenure, ≥ 16 h/week or</p> <p>b) ≥ 5 years of tenure, between 8 and 16 h/week; mix of earnings-related (first 6 weeks) and flat-rate payments (since 1975-1977)</p>	<p><i>Maternity leave</i></p> <p>1994: 14 weeks of job guarantee for all women (statutory); additional period (up to 28th week) if ≥ 2 years of tenure; improved flat-rate, but payments only if ≥ 2 years of tenure</p> <p><i>Parental leave</i></p> <p>1999: up to 13 weeks, unpaid</p>	<p><i>Maternity leave:</i></p> <p>2003: extended to 52 weeks, unpaid in the second 26 weeks</p> <p>2007: flat-rate payments up to 33 weeks (from 20)</p> <p><i>Paternity leave</i></p> <p>2003: two weeks, flat rate</p> <p>2009: option to take unused maternity leave for fathers</p> <p><i>Parental leave:</i></p> <p>2015: Shared Parental Leave (SPL), 50 weeks (37 paid); eligibility based on tenure and h/week; no mandate for fathers</p>
DE	<p><i>Maternity leave:</i> 14 weeks, job guarantee, full income replacement (since 1968)</p> <p><i>Parental leave</i></p> <p>1984: 6 months job guarantee, earnings-related benefit</p> <p>1986-1989: duration gradually up to 15 months, mix of flat-rate and means-testing benefit for 12 months; access to the benefit granted to non-employed mothers</p>	<p><i>Parental leave</i></p> <p>1990: job guarantee up to 18 months</p> <p>1992: job guarantee up to 36 months</p> <p>1993/1995: payments up to the 24th month</p>	<p><i>Parental leave</i></p> <p>2001: increased benefit if leave of 12 months only; 12 months of leave can be taken between a child's 2nd and 8th birthday; parents allowed to work 30h/week (up from 19h limit)</p> <p>2007: earnings-related benefit, 67% replacement rate (capped at 1,800 EUR/month); "daddy months"</p>

Table 1A continued from previous page

1980s	1990s	2000s - present day
		<i>Maternity leave</i> 2015: six weeks after delivery, remaining weeks can be spread over a maximum of 30 weeks
	<i>Maternity leave</i> 1990: duration up to 16 weeks	
	<i>Parental leave</i> 1991: 13 times weekly working hours, eligible if ≥ 20 h/week and ≥ 1 year of tenure; unpaid in the private sector 1997: extension to workers < 20 h/week; flexibility within a child's 8th birthday; option to reduce hours instead of leave period	<i>Parental leave</i> 2009-2014: expansion up to 26 times the weekly working hours and tax credit of 50 per cent of minimum wage; extended if < 1 year of tenure; added flexibility and job protection 2015: leave is now unpaid
NL	<i>Maternity leave:</i> 12 weeks, job guarantee, full income replacement (since 1969)	<i>Paternity leave</i> 2001: two days (fully paid) 2019: extended to five days (gov't plans to expand it to 5 weeks, 70% wage replacement, from 2020)

Sources: Burgess et al. (2008); Ziefle and Gangl (2014); Begall and Grunow (2015); Blum et al. (2018); CESifo database, <https://bit.ly/2zTwwwV>.

Chapter 1

Is There a Fatherhood Wage Premium? A Re-Assessment in Societies with Strong Male-Breadwinner Legacies

Is There a Fatherhood Wage Premium? A Re-Assessment in Societies with Strong Male-Breadwinner Legacies*

Abstract

This study asks whether fatherhood sparks the wage attainment of men or rather entry into fatherhood is simply more typical for high-earning men and at times of wage growth during the career-cycle. Fatherhood premiums may contribute to gender economic inequalities, particularly in countries with strong male breadwinner legacies such as Germany and the UK. And yet, as male breadwinner norms have waned and policies have started fostering men's role as carers, wage premiums could be a thing of the past.

I use long-running panel data for both countries. Pitfalls and benefits of three regression-based approaches (pooled OLS, fixed effects estimation, and fixed effects individual-slope estimation) are highlighted. Overall, fatherhood wage bonuses cannot be detected, on average as well as across birth cohorts. At best, estimates are compatible with premiums of a few percentage points among the older cohorts examined. Positive selection on both prior wage levels and wage growth is largely responsible for the apparent wage boost. The contribution of selection on prior wage levels though is fading across cohorts, meaning that men select into fatherhood less and less on the basis of time-invariant characteristics positively related to both wages and the chance of becoming a father. Similar to what has been suggested for marriage, the link between fatherhood and wages appears to be more of a selection story than a causal one, even in contexts with strong male-breadwinner legacies.

*Data from the BHPS and UKHLS were made available through the UK Data Archive (University of Essex, Institute for Social and Economic Research, 2018), while data from the G-SOEP were made available by the German Institute for Economic Research (DIW), Berlin, 2016. Neither the original collectors of the data nor the archive bear any responsibility for the analyses or interpretations presented here. I wish to thank Lynn Prince Cooke, Renske Keizer, Volker Ludwig, and Rossella Icardi for useful comments on previous versions of this paper. This paper has previously been presented at the RC28 Spring Meeting 2019 (University of Frankfurt). A slightly different version of this paper currently under review at an international peer-review journal.

Wage penalties after motherhood are now key drivers of gender wage gaps in high-income countries (for a review, Ponthieux and Meurs, 2015). Counterpart to wage losses for mothers, fathers may receive a more modest wage premium (Lundberg and Rose, 2000; Petersen et al., 2011; Killewald, 2013). Yet, while the question of whether motherhood causally affects wages has come under intense scrutiny (e.g. Elwert and Winship, 2014), few have asked the same about fatherhood (for an exception, Loughran and Zissimopoulos, 2009).

I reconsider, first, the possibility that fathers earn more due to how men select into fatherhood (e.g. Kravdal and Rindfuss, 2008; Trimarchi and Van Bavel, 2017). Previous research has carefully accounted for the fact that high-earning men are more likely to become fathers. Selection though may also operate through the superior wage growth of fathers-to-be, rather than just through their wage levels (Ludwig and Brüderl, 2018). Accounting for both, this study provides a more severe test for the causal effect of fatherhood on wages.

Second, adding to a literature mainly based on the US, my re-assessment is conducted comparing Germany and the UK for a number of reasons. These two countries have long supported male breadwinning, both through policy and culturally: The inability to support a causal story for fatherhood and wages would thus be particularly meaningful in such contexts. As male breadwinner norms have waned (Knight and Brinton, 2017) and policies in both countries have extended family leave rights to men (Blum et al., 2018), rich longitudinal data allows me to investigate heterogeneity across cohorts. The latter has been overlooked in previous studies, perhaps partly because of the invariance of the policy context of fatherhood in the US (at least at the federal level, e.g. Baum and Ruhm, 2016).

If present in the past, wage premiums may have declined and particularly in Germany, for comprehensive evidence has pointed to a (small) shift in effort from the market to the household for German men in recent cohorts (Pollmann-Schult and Reynolds, 2017; Leopold et al., 2018; Tamm, 2018). Notably, German fathers have increasingly accessed a generous and relatively long parental leave provision put in place since 2007 (e.g. Bünning, 2015). In the UK, evidence points to similar changes in men's contribution to household

production (e.g. Huerta et al., 2014; Altintas and Sullivan, 2016), but policy change regarding fathers has been less extensive (Lewis, 2002; Lewis and Campbell, 2007; Tanaka and Waldfogel, 2007). This further motivates a cross-country, cross-cohort comparative perspective: If fatherhood wage premiums prove persisting, they may further deepen gender wage gaps or make it such that, even if the mommy penalty will narrow down, gender wage gaps will endure (Petersen et al., 2014).

Considering whether fatherhood, to this day, causally affects men's wages is thus paramount to disentangle the sources of gender economic inequality. In short, I ask here (i) if fatherhood premiums are causal or rather a by-product of the process by which men select into fatherhood, (ii) whether premiums are found in contexts with strong, yet fading, male-breadwinner legacies, and (iii) whether premiums for fathers have faded too or have rather persisted across cohorts.

1. Background

1.1. Why there could be a fatherhood premium: reviewing previous evidence and mechanisms

Considering previous evidence, the fatherhood wage premium seems generally modest in size and is not always detected across contexts. Both findings are in stark contrast with research on motherhood wage penalties. Estimates of the gross premium for all fathers have been the largest in North America, up to a range of 6-13% in Canada and the US (Lundberg and Rose, 2000; Hodges and Budig, 2010; Cooke and Fuller, 2018). In Norway and Denmark, differently, premiums have not been ascertained (Cools et al., 2017; Kleven et al., 2018) or have at best amounted to 1-2% (Petersen et al., 2011). Few studies have examined the premium in Germany and the UK, finding for the former a 2-3% wage boost for higher-order parities (Pollmann-Schult, 2011). For a comparison, gross motherhood wage losses have typically been estimated in excess of 15% in Germany (Beblo et al., 2008; Gangl and Ziefle, 2009; Kühhirt and Ludwig, 2012) and of around 12% in the UK (Gangl and Ziefle, 2009; see also Harkness, 2016). Beyond Germany and the UK, substantial economic losses for mothers seem rather universal across contexts (e.g. Kleven et al., 2019).

When detected, wage premiums for fathers have been traced back to individual changes

in work effort, couple specialisation, and employer discrimination. Based on previous evidence and theoretical considerations though, the explanatory power of each of these mechanisms seems dubious. For one, fatherhood may elicit an increase in men's work effort (e.g. Eggebeen and Knoester, 2001). Yet studies based on household data have not been able to assess this due to the lack of precise measures of productivity, relying instead on proxies of effort/productivity such as working hours. As previously acknowledged (Killewald, 2013), working hours may be poor proxies of effort or only affect wages in the long run, as hard-working employees signal themselves to employers and secure thereby better-paid positions in internal labor markets (Gibbons and Waldman, 1999). In the short run, working longer hours may increase total earnings rather than wage rates, but studies of the fatherhood premium have typically been concerned with the latter (for an exception, Cooke, 2014).

Further, fatherhood may propel a man to increase his work effort only conditional on whether his earning potential exceeds that of his partner (Becker, 1981). The implication of such a model of couple specialization is that fatherhood wage premiums should be observed in particular for those men whose partner reduces working hours or leaves paid work come parenthood. Indeed US-based studies find that married fathers, and among them especially those whose wives interrupt employment or cut back work hours, add extra working hours of their own and gain a wage premium (Lundberg and Rose, 2000; Killewald and Gough, 2013; Killewald, 2013).

Assessing the impact of couple-level specialization on the wages of fathers, however, presents additional complications given the "endogeneity of family" (Lundberg, 2005). Couple formation is often a transitory event, considering that splitting and re-partnering are commonplace (Elzinga and Liefbroer, 2007). If specialization affects couple formation and stability, conditioning the analyses of fatherhood premiums on the presence and characteristics of a partner may induce selection bias (Elwert and Winship, 2014). For example, if couples in which women do not specialize in household production are more likely to split (e.g. Kalmijn et al., 2007; Lepinteur et al., 2016; for a review, Cooke and Baxter 2010), "surviving" couples at any given point in time may be the ones more likely to (have) specialize(d). The role of couple specialization on the wages of fathers may thus

be exaggerated, yet studies have neglected this point so far (e.g. Killewald and Gough, 2013).

Moving on from the behaviors of fathers and couples, wage premiums for fathers may also precipitate from employer discrimination. Employers may have a preference for fathers, rooted in the perception that fathers will be more productive, competent and/or committed than their childless counterparts (cf. Phelps, 1972; Correll et al., 2007). Despite the lack of factual productivity differences between male employees, men could be differentially treated in the workplace depending on parental status. In a lab experiment with US undergraduate students, Correll and colleagues (2007) indeed find that fathers are evaluated as more committed, would be hired more often, and would be offered higher starting salaries than childless men, holding job applicants' features equal. Yet, in the companion field experiment, differences in call-back rates between "equivalent" fathers and childless men were not detected. The same inconclusive evidence has recently emerged from a large field experiment across multiple job titles in Sweden (Bygren et al., 2017).

Hence, a review of the size and generative mechanisms of the fatherhood bonus motivates asking whether expecting a causal effect of fatherhood on wages is warranted in the first place.

1.2. Contextual underpinnings of the wage trajectories of fathers in Germany and the UK

While addressing wage determination at the individual level, contextual factors that may shape the fatherhood premium (and its drivers) have been overlooked in the literature (cf. Cooke, 2014). One might expect that a strong male-breadwinner norm, both culturally and institutionally enforced, may foster shifts in market effort after fatherhood, traditional specialization patterns within couples, or employer bias in favor of fathers.

Looking at the UK and Germany, I compare two countries with strong yet drifting male breadwinner legacies (e.g. Crompton, 1999). Culturally, male-breadwinner norms have been particularly strong in former West Germany rather than in the UK (Knight and Brinton, 2017; Trappe et al., 2015). Unfavorable attitudes towards mothers' employment – whether full-time or in general during a child's pre-school years – persist in both countries (O'Reilly et al., 2014; Dechant and Rinklake, 2016). Still, attitudes in both countries have

shifted away from traditionalist views that assigned to men the role of (sole) breadwinners and to women that of full-time carers and home-makers (Knight and Brinton, 2017).

As for policy, fatherhood has long been synonym with “providing”, emphasised, for one, by mandated cash transfers from fathers to mothers in case of couple dissolution (Hobson, 2002). Lately though, and to a somewhat greater extent in Germany, policy reform started targeting fathers for their *care* obligations rather than for *cash* provision (Adler and Lenz, 2017). Germany introduced two bonus months of paid parental leave if both parents take some leave at all (2007), and succeeded in increasing father involvement in the household (e.g. Tamm, 2018) in times of a broader, if slow, gender convergence in the division of labor in Germany (Pollmann-Schult and Reynolds, 2017; Leopold et al., 2018). The UK opted for a paid statutory paternity leave (2003), with high uptake rates but lasting only two weeks, and a parental leave scheme (2015), only partly paid and so far largely unutilised by new parents (Blum et al., 2018). Change in the cultural and policy context surrounding fatherhood thus further enriches the chance to understand the contextual underpinnings of fatherhood premiums (if any) over time in the two countries.

First, fatherhood may act as a transforming event and spur men’s work effort particularly in contexts where male breadwinning is culturally reinforced (Townsend, 2002). Yet the transition to fatherhood does little to change men’s attitudes towards work and family, as men seem not to become more (or less) traditional after the birth of a child (Grinza et al., 2017; Kuziemko et al., 2018). Coherently, findings regarding fathers’ working hours as a proxy for effort suggest that, even in former male-breadwinner regimes, men do not commit more fully to breadwinning after the birth of a child. Indeed, average working hours actually decline after the transition to fatherhood in most European countries (Bünning and Pollmann-Schult, 2016). In Germany, becoming a father prompts an increase in working hours for men born prior to 1960, but a decrease in working hours for men belonging to younger cohorts, and both changes are modest in size (1 h of paid work at most, Pollmann-Schult and Reynolds, 2017). For Britain, previous studies provide little evidence that the presence of children affects men’s working hours at all (Bryan, 2007; Paull, 2008; Schober, 2013).

Nevertheless and second, men could increase their work effort conditional on their part-

ner's investment in the household, in line with a traditional mode of couple specialization. Tax policies, for example, may provide incentives for particular arrangements of paid and unpaid work between partners. Pooling the income of both partners to determine personal income tax, Germany discourages paid employment among the secondary earners, typically women (Smith et al., 2003; Bick and Fuchs-Schündeln, 2017), and may thereby foster traditional couple specialization. Differently, Britain switched from joint to individual taxation in 1990, further pursuing tax credit policies since the end of the 1990s with the aim of encouraging maternal labor supply (Francesconi and Van der Klaauw, 2007; Francesconi et al., 2009). Perhaps surprisingly, though, couple specialization in both the UK and Germany deviates from the Beckerian model when it comes to the transition to parenthood. While mothers indeed trade off employment hours with time spent in housework and childcare, fathers' allocation of time to either paid or unpaid work is hardly affected by parenthood regardless of their partner's behavior (e.g. Schober, 2013; Kühhirt, 2012).

Employer bias in favor of fathers may still rest on the assumption, however justified, that men will maintain or even increase their commitment to paid work after the birth of a child. In recent years, such an assumption may have eroded because of fathers' use of paternity and parental leaves, which may signal a parallel and potentially conflicting commitment to the family sphere. In Britain, around 80% of fathers now take time off around the birth of a child, although mostly in the form of the two-week paternity leave introduced in 2003 (Blum et al., 2018). In Germany, fathers have been entitled to paid parental leave provisions since the end of the 1980s, but fathers' uptake became substantial only after the aforementioned 2007 reform. According to the latest figures, more than 30% of German fathers now use parental leave provisions, typically for the statutory minimum of two months (Bünning, 2015; Kluge and Tamm, 2013). Notably, these figures for German fathers approach those of their counterparts in Sweden. Swedish fathers taking parental leave have been found to experience modest wage penalties after returning to work, a finding scholars have interpreted as evidence of adverse signalling (Albrecht et al., 2015; Evertsson et al., 2016). Also, wage penalties for fathers taking leave may stem from a re-orientation of effort from the market to the household (Rege and Solli,

2013). It is at best unclear, therefore, if employers may still assume a fuller commitment to work from fathers and discriminate in their favor, particularly in modern-day Germany. More broadly, as male breadwinner norms waned and policies also shifted emphasis from “cash” to “care”, it could be that fatherhood premiums are at best a thing of the past in both countries, and possibly even more in Germany rather than the UK. Other than investigating the average causal effect of fatherhood on wages, I will look into its possible heterogeneity across cohorts.

1.3. Why there might not be a fatherhood premium after all: the role of selection

Mixed support for the mechanisms seemingly generating wage premiums for fathers, even when considering contexts in which we might expect them at work the most, prompts taking a step back. It is natural to ask whether apparent wage boosts come fatherhood are causal or rather driven by selection into fatherhood. I consider here two sources of selection: selection on prior wage levels and selection on prior wage growth.

Similar to selection into marriage (for a review, Ludwig and Brüderl, 2018), selection on prior wage levels entails that high-earning men are, by the same token, more likely to become fathers. As Cooke and Fuller put it (2018: 783), “positive selection might account for the gross (wage) premium if the men who become fathers have unmeasured characteristics such as loyalty and commitment valued similarly by employers and potential partners”. If such positive selection holds, ignoring it would lead to an overestimation of fatherhood wage premiums.

Research on what kind of men eventually become fathers, and what kind does not, is relatively under-developed (Balbo et al., 2013; Kreyenfeld and Konietzka, 2017). Across countries, highly educated men have better chances of becoming a father than do low-educated men, yet much is due to selection into union (e.g. Trimarchi and Van Bavel, 2017). If fatherhood and union formation are a compound, then one might expect the type of selection into fatherhood to overlap with that into marriage, with both of them being positive (for a review, Ludwig and Brüderl, 2018). As long as such drivers of positive selection are unobserved, cross-sectional estimates of the bonus (e.g. Cooke, 2014; Petersen et al., 2011, 2014; Cooke and Fuller, 2018) might thus suffer from selection bias due to omitted variables in the regression equation. Panel estimates, differently,

can be augmented by adding individual fixed effects to curb estimates from selection on such time-invariant unobservables. Comparing the latter to cross-sectional estimates has highlighted negative selection may be at play, as panel estimate of the daddy bonus are typically bigger than their cross-sectional counterparts (Lundberg and Rose, 2000; Hodges and Budig, 2010). Yet, once again, these studies have only focused on the US. They could thus speak to the specificity of the American context where fatherhood – or at least early fatherhood – may go hand in hand with markers of earning and life-course disadvantage such as dropping out of high school or incarceration (e.g. Dariotis et al., 2011). Hence, also the direction of selection on prior wage levels is at best ambiguous, much like the effectiveness of mechanisms purportedly leading to a causal bonus discussed in the previous sections.

Yet selection could also operate through a different path. Men may select into fatherhood, depending on their wage growth rather than simply on their wage levels (Ludwig and Brüderl, 2018). The transition to parenthood, much like that to marriage (Killewald and Lundberg, 2017), may simply occur at times of fast wage growth in the career-cycle. For one, as men who become fathers are disproportionately better educated (Kravdal and Rindfuss, 2008; Trimarchi and Van Bavel, 2017), they do not simply enjoy high wages but also steep wage growth paths (e.g. Lagakos et al., 2018) – steeper, possibly, than that of relatively less educated men who are more likely to remain childless. This might also be more pronounced for men belonging to younger cohorts, as they typically have children later in life after considerable accumulation of experience and wages in the market (McMunn et al., 2015; Struffolino et al., 2016). Once again, the US might not provide the best context to assess this selection dynamic, as mean age at first birth has not increased much across cohorts, and the transition to parenthood is prominent in the early 20s when labor market careers are not yet consolidated (e.g. Mills et al., 2011).

Previous comparative studies for European countries, including the UK and Germany, have indeed found that the wages of men are growing already in the period prior to fatherhood (Smith Koslowski, 2011). Yet such literature has not drawn one important implication out of this finding. If the wage spikes prior to fatherhood are comparable to the wage spikes observed after fatherhood, then speaking of a bonus sparked by fatherhood is

unwarranted given the observed data pattern (e.g. Killewald and Lundberg, 2017; Ludwig and Brüderl, 2018). The tests I devise in this study, therefore, will try to detect a spike in wages occurring in the aftermath of fatherhood, once men’s selection into fatherhood – based on both wage levels and on wage growth – is accounted for.

2. Methods

2.1. Data and samples

I employ long-running household panel data, namely the German Socio-Economic Panel (SOEP v31 1984-2014, [doi:10.5684/soep.v33](https://doi.org/10.5684/soep.v33)) and the British Household Panel (1991-2016, [doi:10.5255/UKDA-SN-6614-11](https://doi.org/10.5255/UKDA-SN-6614-11)). Both are multipurpose household surveys following the lives of a representative sample of each country’s residents (Goebel et al., 2018; Taylor et al., 2010; Buck and McFall, 2011). Both datasets are augmented by fertility history files (see, respectively, Goebel, 2017; Pronzato, 2011) to recover information on the transition to fatherhood.

For the UK, I rely on all BHPS sample members. While the BHPS was temporarily discontinued in 2009, its sample started being interviewed again in 2010-11, within the framework of the UK Household Longitudinal Study (UKHLS). My analyses thus rely on the full data available for the BHPS sample, covering, despite the gap, the period 1991-2016. For Germany I employ samples A to H, as well as refreshment samples J and K. In the main analyses, I focus on men aged 20 to 50, working as dependent employees, with non-missing information on their fertility history as well as on their current wage. Further, given my focus on the transition to fatherhood, I restrict my analyses to men that, when first observed in the panel, had no children. During the observation period then, part of this initial pool will experience the transition to fatherhood (“treated group”), while the rest will remain childless (“control group”). For some of the men in this latter group, family histories may be truncated though: they might become fathers after the last datapoint in each panel or drop out of the panels prior to their transition to fatherhood. If arguments on selection hold, this subgroup of men among the controls might actually be more similar to fathers(-to-be) than to childless men, and this would in turn attenuate estimates of the fatherhood premium. Hence, to ensure that the control group does not include prospective fathers, who might be on a similar wage trajectory to that of the

treated, I further limit my analysis to men who have been observed at least until age 40, thereby selecting cohorts of men born not after 1974 for Germany and not after 1976 for the UK. The cut-point at 40 is assumed to be indicative of completed fertility for men, as less than 10% of the transitions to fatherhood occurred past age forty in the final samples for both countries (see also Kleven et al., 2018). Applying or not this sample restriction, however, does not alter the substantial conclusions of this paper (for sensitivity checks and a discussion, see Appendix H).

Finally, due the requirements of one of the statistical models I will use (FEIS model, see below), I further limit my analyses to men who have been observed for at least three waves in the panel (Brüderl and Ludwig, 2015; Ludwig and Brüderl, 2018). All together, these restrictions result in a sample of 2,709 men (1,012 of which will become first-time fathers) and 34,879 person-year records for Germany, and a sample of 1,251 men (572 of which will become first-time fathers) and 15,454 person-year records for the UK. In both samples, men are followed for an average of roughly 12 waves. After (prior) the birth of a child, in particular, first-time fathers are followed for an average of roughly 12 (5) waves in the German sample and 10 (5) waves in the British sample.

Table 1 sums up all sample restrictions and relative sample sizes. Notably, restricting the sample to individuals observed for at least three waves has a minor impact on the final sample counts. My main findings, coherently, are unaltered by this choice (see Figure 1A, Appendix B).

With the exception of the selection of suitable treatment and control groups, the most substantial drop in sample size is observed because of the use of fertility history files in SOEP. For Germany, SOEP started collecting men's fertility histories only from 2000 onwards (Goebel, 2017). Respondents that dropped from the panel prior to that date are not included in the analysis. Nevertheless, replicating the analyses for Germany using fertility info from the core file ("number of children in the household") does not alter any of the conclusions on the impact of fatherhood on wages for German men (see Appendix F).

2.2. Measures

The outcome variable in this study is the log of real hourly wages. Hourly wages are computed dividing gross monthly pay by the amount of weekly working hours multiplied by 4.35, the approximate number of weeks in a month. For the BHPS arm of my analysis, I sum weekly working hours and hours of overtime (Bryan and Sevilla-Sanz, 2011). For SOEP, I use actual working hours or, when missing, the sum of contractual working hours and of overtime (Kühhirt and Ludwig, 2012). As per Table 1, listwise deletion on missing wages only lead to marginal sample losses, of 0.3% and 0.4% of the potential sample in SOEP and BHPS respectively.

Wages are then indexed at 2014 prices (2016 for the UK) and values below 1 or above 100 are trimmed, following standard practices in the literature (e.g. Kühhirt and Ludwig, 2012). Taking the natural logarithm of real wages then enables the interpretation of coefficients in terms of percentage effects on wage levels, since the log scale well approximates the percentage-point scale as long as coefficients lie in the $(-.25, .25)$ interval.

I deploy two measures to single out the effects of the transition to fatherhood. The first is a simple dummy ($Child_{it}$) that equals 0 prior to the birth of the first child and 1 in its aftermath. In a second model specification, I consider the effect of the transition to fatherhood as distributed along the life course of men (e.g. Dougherty, 2006). Instead of a single dummy switching from 0 to 1 after the transition to fatherhood, I operationalize fatherhood via a series of dummies k for each year t since the first child's birth. Whenever relevant data was available (thus, with the exclusion of the UKHLS arm), both measures are corrected for the month in which the child is born. If the child's birth occurred in the interview year but months prior to the interview month the appropriate dummy is set to 0; only if the interview occurred in the same month of birth or after, the appropriate dummy is set to 1.

Focusing on the transition to fatherhood, and thus on first childbirth events, is consistent with part of the literature on family events and wages (e.g. Kleven et al., 2018; Ludwig and Brüderl, 2018). Many studies in the field have rather employed dummies for different parities (e.g. Pollmann-Schult, 2011; Petersen et al., 2011) or a single counter for the number of children (e.g. Gangl and Ziefle, 2009; Cooke, 2014). My operationalization

choice was then driven by two types of considerations. One is that arguments being developed regarding both mechanisms and selection dynamics deal with the transition to fatherhood and not with the effect of specific (higher-order) parities. It is posited, for example, that employers treat differently fathers and childless men, rather than discriminate among fathers depending on their number of children. As for selection dynamics, the core contrast is between childless men who will eventually become fathers and men who will not.

Differently it could be argued that the need for increasing work effort and for couple specialization might be heightened after higher-order births, as additional children demand higher income. Even if of sure interest, the effects of different parities might not be easily disentangled in the regression framework though, considering that higher-order births might be endogenous to previous births. Notably, one birth may causally affect the chances of a subsequent birth via its effect on a parent's labor market standing. If men, say, receive a wage premium after the birth of a first child, this might increase their chances of having a second one because they can afford a bigger family size. If, in turn, the second birth propels men to specialise in paid work even more and get even higher wages, part of the effect of a first birth on wages will "work through" the second birth. Focusing on the transition to fatherhood, I therefore look for the total effect of fatherhood on wages: Such total effect *includes* wage responses to higher-order parities, yet avoids the empirical hurdles of disentangling separate effects for each endogenous childbirth event.

Models throughout also include age and age squared to net out pure lifecycle effects. Results are unchanged when opting for different polynomial forms (quadratic, cubic, and quartic) or for a full set of dummies for age, potential labor market experience, and (for SOEP only, due to data availability) actual labor market experience. Period dummies are also included, grouping years in 3-year bans (with only two broader residual categories: 2011-14 for SOEP, 2012-2016 for BHPS). In SOEP, interviews are carried out annually: when applying the within-individual transformation in fixed effects (FE) models, age and interview year thus increase of one unit each year creating collinearity between the two variables in the FE regression model. Grouping period dummies circumvents this issue.

For consistency, I deploy 3-year dummies for the UK arm of the analysis as well, although BHPS interviews naturally span over multiple years.

For cross-cohort comparisons, I identify three birth cohorts, distinguishing men born between 1950 and 1959, 1960 and 1967, and between 1968 and 1976 (1974 for Germany). Cut-points derive, first, from the need to assure enough cell size in each group. Second, I follow previous literature on Germany, suggesting that the relationship between fatherhood and labor market participation (and thus, perhaps, wages) changed starting from men born in the 1960s (Pollmann-Schult and Reynolds, 2017). For the UK, I will not report estimates for the 1950-59 British cohort since they would be based on relatively older men (32 or older) experiencing the transition to fatherhood.

2.3. Fatherhood and wages: model specifications

I start from a simple OLS specification of the wage equation,

$$y_{it} = \alpha + \beta Child_{it} + \gamma_1 Age_{it} + \gamma_2 Age_{it}^2 + \phi_t + \epsilon_{it} \quad (1)$$

where y_{it} is the log of real hourly wages. Apart from age (Age_{it} and Age_{it}^2) and period (ϕ_t), no other variables are adjusted for on the right-hand side of the equation (e.g. Loughran and Zissimopoulos, 2009; Killewald and Lundberg, 2017). My goal here is to test whether data are compatible with a causal fatherhood wage bonus. Parsimony in model specification allows, first, to estimate such gross bonus, if any. Second, parsimony shields estimates from the risk of “overcontrol bias” (Elwert and Winship, 2014), that is, of muting the effect fatherhood has on wages by accounting for the channels through which the effect manifests in the first place. Adjusting for working hours, for example, could already account for part of the bonus in accordance with the work effort mechanism. I refrain, therefore, from including this and similar variables in the analyses, as the question of whether data supports a causal fatherhood wage bonus predates asking what mediates the bonus if present.

Still, pooling together all observations as if they belonged to different units, the OLS model does not distinguish person-year records belonging to the same person from person-year records belonging to a different one. Estimates of β in Equation 1, therefore, simply

contrast records in which the transition to fatherhood has occurred ($Child_{it} = 1$) to records in which the transition has not occurred ($Child_{it} = 0$). The latter group of observations includes both fathers-to-be prior to first childbirth and men belonging to the control group of childless men.

Differently, to focus on within-individual change, and thus on the transition to fatherhood, I contrast Equation 1 to the following fixed effects (FE) specification:

$$y_{it} = \beta Child_{it} + \gamma_1 Age_{it} + \gamma_2 Age_{it}^2 + \phi_t + \theta_i + \epsilon_{it} \quad (2)$$

The estimation of β associated with $Child_{it}$ in Equation 2 rests only on within-unit variance, thanks to the inclusion of individual fixed effects θ_i . It can be interpreted as the average one-off shift in wages that men experience when becoming fathers (for details on how this average is computed, see Borusyak and Jaravel, 2016; Imai and Kim, 2017). Further, individual fixed effects curb estimates of β from time-invariant sources of selection into fatherhood. The role and direction of such type of selection can be assessed in two ways. First, I contrast OLS and FE estimates of the β coefficient in both equations. Negative selection on time-invariant individual characteristics should be signalled by an increase in the magnitude of β in the FE model; positive selection, vice versa, by a decrease. Second, I look at the correlation coefficient $r(\theta_i, Child_{it})$, expressing the sign and magnitude of selection into fatherhood in terms of the correlation between wage-relevant time-constant unobservables θ_i and the variable for fatherhood, $Child_{it}$ (e.g. Gangl and Ziefle, 2009).

While the model in Equation 2 nets out time-invariant sources of selection, it still makes a number of assumptions. One is that the change in wage levels come fatherhood can be expressed as a one-off change, summarized by a single coefficient β . Research on male marital wage premiums and motherhood wage penalties alike has now well established that family events impact wages in a dynamic fashion (e.g. Korenman and Neumark, 1991; Loughran and Zissimopoulos, 2009; Kleven et al., 2018). In other words, one is better off modelling parenthood effects year by year after the birth of a child, as short-, medium-, and long-run shifts in wage levels may differ. This first assumption is relaxed by deploying an event-study design (e.g. Kleven et al., 2018), adding leads k of $Child_{it}$

as follows (see also, Imai and Kim, 2017):

$$y_{it} = \sum_{k=0}^{k=10} \beta_k \mathbb{1}(k = t - Child^i) + \gamma_1 Age_{it} + \gamma_2 Age_{it}^2 + \phi_t + \theta_i + \epsilon_{it} \quad (3)$$

With $Child^i$, I now indicate the year in which the child’s birth occurs for an individual i . Calendar time keeps being signalled by the subscript t . $\mathbb{1}(k = \dots)$ is the indicator function whose argument can be either 1 or 0, i.e. a (set of) dummy variable(s). Equation 3 then includes event-time dummies k for each year t since the year of first childbirth and up to the tenth after. I cap the last indicator variable (for $k = 10$), coding as 1 also all years after the tenth since first childbirth. Results out of Equation 3 are unaltered by this choice (see also Appendix C; for cell sizes for each dummy k see Table 1A).

With Equation 3, I am thus able to capture the dynamic evolution of wages (if any) after fatherhood. Yet, prior to fatherhood, fathers-to-be and childless men could be on different wage growth paths. This would violate another assumption behind the FE estimator, namely the strict exogeneity assumption for a treatment (here, $Child_{it}$). Equations 2 and 3 indeed require that outcomes evolve in parallel for controls and treated prior to the treatment (e.g. Wooldridge, 2010; Brüderl and Ludwig, 2015). In my case, childless men and fathers-to-be (prior to fatherhood onset) should experience similar wage growth. This might not be the case considering, for example, that much of selection into fatherhood operates through selection into a union (e.g. Trimarchi and Van Bavel, 2017) and men who enter a union are on steeper wage growth profiles compared to men who will remain single (e.g. Ludwig and Brüderl, 2018). Since entry into a union has little causal effect *per se* on men’s wages (*ibidem*: Loughran and Zissimopoulos, 2009; Killewald and Lundberg, 2017), there is little risk of mistaking for fatherhood wage bonuses what are actually marital wage bonuses. Adjusting estimates for men’s entry into union is likely not necessary to achieve credible causal inference on fatherhood and wages (cf. Ludwig and Brüderl, 2018; see also Figure 4A in the Appendices). Rather, a process of selection on the basis of wage growth might underlie family formation (union entry and fatherhood) and failing to account for it may lead to incorrectly claim the existence of a positive wage boost brought about by fatherhood.

Hence, to compound evidence obtained via the FE estimators in Equations 2 and 3, one

can ask if any effect of fatherhood on wages can be detected once any individual trend in wage growth is netted out (Ludwig and Brüderl, 2018). A more restrictive specification of this kind fully accounts for the chance that fathers-to-be and childless men may be on different wage-growth paths. It could permit wage growth to depend on any (time-invariant) individual characteristic, since each man in the sample would basically have his own wage slope. A fixed effects individual-slope (FEIS) model (e.g. Brüderl and Ludwig, 2015; Ludwig, 2019) is thus fitted in the following steps: first, one should estimate an OLS regression for each individual i , regressing log-wages y_{it} on a constant and a linear term for age; second, get the predicted values and subtract them from y_{it} , thus obtaining wage values for each individual i that are both de-measured (constant term in step 1) and de-trended (linear term for age in step 1); third, repeat step 1 for all independent variables to de-mean and de-trend them as well; finally, run an OLS on the transformed data. All these steps are automated in the STATA routine `xtfeis` (Ludwig, 2019) which I deploy. A more compact formulation of the model is as follows (Wooldridge, 2010: 377):

$$y_{it} = \beta Child_{it} + \gamma_2 Age_{it}^2 + \phi_t + \mathbf{W}'_{it} \theta_i + \epsilon_{it} \quad (4)$$

With respect to the fixed effects specification in Equation 2, I allow now for the product of individual fixed effects θ_i and some observable variable, namely the linear term for age, contained in the vector \mathbf{W}'_{it} . I ran such FEIS estimation both using the single dummy variable $Child_{it}$ and, separately, the event-study approach of Equation 3. Notably, since FEIS estimation is based on a OLS model for each single individual (step 1) including two parameters (the constant term plus a linear slope for age), at least three observations per person are needed, hence the aforementioned sample restriction (Ludwig and Brüderl, 2018).

Throughout, standard errors are handled via the Huber-White estimator to account for the clustering of observations within each individual (via the option `robust` in STATA 14.0).

3. Findings

Pooling data from all the available waves, Figure 1 depicts average wage levels from age 20 onwards and separately for men who will eventually become fathers and for men who will not. Trends are similar in Germany and Britain. Wages at the mean grow at a slightly faster pace for fathers relative to non-fathers up until their late 20s/early 30s. After that, the wages of men who will remain childless stagnate, while men who did or will experience the transition to parenthood earn even higher wages than in their youth. Such a pattern could be consistent with the existence of a fatherhood wage bonus, as wage trajectories diverge right around the age of first-time fatherhood in both countries (McMunn et al., 2015; Struffolino et al., 2016). The picture cannot be conclusive, however, on whether men experience a wage premium come parenthood or some antecedent factors boost the wages of fathers-to-be.

Turning to statistical models, OLS estimates in the first column of Table 1A and Table 3 reflect the patterns displayed in Figure 1. Fatherhood is associated with an average wage gain of about 14 percent in Germany ($p < .001$) and 17 percent in Britain ($p < .001$). In terms of magnitude, these estimates are compatible, if not slightly higher, with those highlighted for the US in previous studies (Lundberg and Rose, 2000; Hodges and Budig, 2010). Yet OLS regression pools all observations in the panel and does not distinguish records belonging to the same individual from records belonging to a different individual. As a result, OLS estimates in the first column of both tables simply contrast person-year records of men once they have become fathers with person-year records of men who are not (or not yet) fathers. I thus quantify the differential already shown in Figure 1, but cannot address the potential bias stemming from selection into fatherhood.

A step forward in this direction comes from FE estimates in the second column of Tables 1A and 3. Such estimates focus on within-individual variation and therefore can be interpreted as the one-off shift in wage levels brought about, on average, by the transition to fatherhood. Netting time-invariant differences between men by means of individual fixed effects, estimates are reduced. The shift in wages brought about by fatherhood stops at around 4.2 percent in Germany ($p < .001$). For British men, differently, the fatherhood wage premium halts at around 2.6 percent ($p = .059$). The decrease in size compared to

OLS estimates is opposite to what previous studies have found for the US. Selection into fatherhood in both Germany and Britain seems to be positive, on average: high-earning men in both countries are more likely to become fathers. At the bottom of both tables, correlations $r(\theta_i, Child_{it})$ further support this conclusion. For both countries, I found moderate and positive correlations between time-invariant unobserved heterogeneity θ_i and my indicator variable for fatherhood. Unobserved, time-invariant factors relevant for wage determination are thus positively correlated with the transition to fatherhood.

While netting out time-invariant heterogeneity between individuals substantially reduces the bonus, FE estimates could still be biased if the wages of men who eventually become fathers grow at a faster pace than those of their childless counterparts. In the third column of Tables 1A and 3, fixed effects individual-slope (FEIS) estimates address this by letting the wage-age profile vary between men. Now, including individual slopes, I cannot reject the hypothesis that the effect fatherhood has on wages is actually nil for German men ($\beta = -0.007$, $p = .505$). The same can be said for British men looking at FEIS estimates in Table 3 ($\beta = -0.003$, $p = .803$). Hence, once the possibility of divergent individual wage trajectories is accounted for, the evidence does not support a causal story for the fatherhood wage bonus (as in Ludwig and Brüderl, 2018 for US men).

So far, I have assumed that fatherhood may bring about a one-off shift in men’s wage levels. Such assumption could be unwarranted though. For example, men could increase their work effort or specialize in paid work particularly in the first years following a child’s birth, often compensating for mothers’ work interruptions and subsequent income loss in that period. As many applied and methodological contributions have shown (e.g. Korenman and Neumark, 1991; Borusyak and Jaravel, 2016), a simple dummy for “before-after” the event of childbirth will not help retrieving more complex, dynamic effects. Figure 2 thus displays estimates from an event-study approach, in which the wage response to fatherhood is singled out for each year after that of a child’s birth – as per Equation 3. For German men, the figure supports the presence of a fatherhood wage bonus, especially in the short term. FE estimates in the first five years are all positive and around the size previously assessed with the single-dummy approach (≈ 5 percent). For British men, point estimates are somewhat smaller in the aftermath of fatherhood, compatible with

wage bonuses of around 2 to 4 percentage points.

Yet, turning to FEIS estimates and thus accounting for idiosyncratic wage growth, the alleged premium reduces further. For Germany, in particular, estimates turn negative already in years 0 ($p = .478$), 1 ($p = .068$), and 2 ($p = .144$). In later years, estimates are not just negative in sign but substantial in size. Yet, as the number of fathers observed in such later years shrinks, the precision of FEIS estimates in the long run becomes questionable – with respect to both the sign of the point estimates and to the width of confidence intervals (see also Ludwig and Brüderl, 2018; Appendix C for sensitivity checks). It is therefore unwarranted to consider such estimates as evidence of some substantial long-run fatherhood wage penalty in Germany.

FEIS estimates for the UK are close to 0 in the first years after childbirth, and quite noisy thereafter. For the aforementioned sensitivity of such long-run estimates, caution should be applied this time to the idea that fatherhood spurs substantial wage premiums for British fathers in the long run.

Hence, even if the wage trajectories of fathers and non-fathers depicted in Figure 2 differ markedly in both countries, there is no strong evidence to conclude that fatherhood boosts men’s wages in Germany and the UK. Rather, the apparent wage premium can be traced back, by and large, to positive selection both on the basis of wage levels (static) and on the basis of wage growth (dynamic). Already prior to their child’s birth, that is, fathers-to-be earn relatively higher wages and are on superior wage growth paths than their childless counterparts.

3.1. Heterogeneous effects? A comparison across cohorts

On balance, the evidence presented in the previous section cannot support the idea of a causal fatherhood wage premium. Regardless of the estimation method though, all my analyses have tested for the presence of a premium for all men, on average. In the remainder, I examine cohort differences: After all, fatherhood wage premiums might simply be “a thing of the past”, as the UK and Germany shifted away from a traditional male-breadwinner model in recent decades.

In Figure 3, I repeat the analyses for Germany separately by birth cohort. OLS estimates suggest that fatherhood is associated with higher wages for men, especially for

those belonging to the older cohorts (up to around 24% for the 1950-59 cohort). After the inclusion of individual fixed effects, estimates are substantially reduced in the older cohorts, while they remain small but stationary for the youngest. Finally, looking at the FEIS specification, estimates further reduce and are no longer compatible with the idea of a wage premium particularly in the youngest cohorts (for the 1950-59 cohort, somewhat differently, $\beta Child_{it} = .02, p = .272$).

Results for Britain in Figure 4 exhibit similar patterns across model specifications. In the FEIS specification, in particular, fatherhood brings about a wage gain of around 2 percent for the 1960-67 cohort ($p = .266$) while the point estimate for the youngest cohort is even negative ($-0.01, p = .496$).

Given the partitioning of the sample in different cohorts, I refrain from presenting here the event-study part of my analyses for each separate group due to the small cell size. Nonetheless, event-study results (see Figures 5A and 6A in the Appendices), with both FE and FEIS specifications, give no strong indication either of a fatherhood bonus across cohorts in the UK and Germany.

At best, small causal fatherhood premiums are indeed a “thing of the past” in Germany and the UK. Notably, cross-cohort analyses also reveal that selection into fatherhood on the basis of wage-relevant, time-invariant characteristics may have become less positive across subsequent cohorts. This is evidenced by the smaller and smaller contribution that including individual fixed-effects makes to estimates of the premium across cohorts in Figures 3 and 4.

4. Discussion and conclusion

In this study I ask whether men get a wage premium when they become fathers in countries that have long supported male breadwinning. I propose that observed premiums could be a statistical artifact if men who become fathers have a higher earning potential and experience steeper wage growth than men that will eventually remain childless. By deploying several model specifications that variously account for such selection dynamics, I cannot reject the null of no effect of fatherhood on men’s wages. What is more, across cohorts, both the size of premiums and the importance of selection into fatherhood appear to be fading in Germany and the UK. Over time, fatherhood may have thus become less

of a marker of breadwinning in the labor market, and traits associated with breadwinning may have become less important for the transition to fatherhood in the marriage market. Such findings stand in contrast with previous studies in a twofold sense. First, pertaining to selection into fatherhood, I find that British and German men positively select into parenthood on the basis of time-invariant unobserved factors. In the US, such selection dynamic appears to be negative instead (Lundberg and Rose, 2000; Hodges and Budig, 2010; Killewald and Gough, 2013). The portion of the premium due to positive selection, also, has been decreasing across cohorts in both Germany and the UK. This differs from findings from cross-cohort analyses carried out in Nordic countries, typically finding persistent or even increasing positive selection of men into parenthood on the basis of wage levels or on antecedents of earning power such as educational attainment (Kravdal and Rindfuss, 2008; Hart, 2015; Jalovaara et al., 2018). As the question of “which” men become fathers has broad implications – not limited to men’s own wage attainment, but extending to income inequality across households or the intergenerational transmission of (dis)advantage (e.g. McCall and Percheski, 2010; Huerta et al., 2014) – this study motivates research on the demography of fatherhood and whether and how it is changing over time.

Second, differently from the US-based literature, considering selection on both time-invariant unobserved factors and on prior wage growth, I cannot conclude that fatherhood sparks the wage attainment of men in the UK and Germany. As a consequence, highlighting the crucial role played by selection dynamics casts doubts on the credibility of estimates of fatherhood premiums derived by cross-sectional data, lest they will account for selection into fatherhood as well. So far largely confined to the role of “unobservables”, results in this study prompt further research on those features (personality, genetic variants, non-cognitive skills, beauty, etc., see e.g. Bowles et al., 2001) that may matter for both wage attainment and the transition to fatherhood, to be considered in a comparative and cross-cohort perspective.

Surely, the inability to reject the null in this study could reflect the fact that fatherhood indeed has a negligible (causal) effect on men’s wages, but it may also stem from *a*) lack of statistical power, *b*) measurement error, and *c*) deficiencies in the study design. Sample

sizes in this study are highly comparable to previous longitudinal studies on fatherhood and wages (e.g. Smith Koslowski, 2011; Killewald and Gough, 2013). Yet, if true effects are small or very small I may have failed to detect them due to lack of statistical power. Across both my main and sub-group analyses, at best fatherhood premiums amounted to 4-5%. If even smaller premiums exist and could not be detected here, however, it is worth to question what their substantial significance would be, especially when contrasted, on the other hand, with the order of magnitude of motherhood wage penalties (e.g. Gangl and Ziefle, 2009; Harkness, 2016; Kleven et al., 2018; Cools et al., 2017).

Findings in this paper may also be invalidated by measurement errors. Measurement error could affect the computation of log hourly wages as well as the construction of the key independent variable operationalizing the transition to fatherhood. For the former, I relied on well-established practices in the literature (e.g. Gangl and Ziefle, 2009; Bryan and Sevilla-Sanz, 2011; Kühhirt and Ludwig, 2012) and I can only note that my conclusions on fatherhood and wages are in line with those of studies using perhaps more precise register data, albeit for different countries (Kleven et al., 2018; Cools et al., 2017). Further, using fertility history files has enabled me to detect the precise timing of the transition to fatherhood, a crucial requirement for the event-study part of the analyses presented here. Finally, while the study design of this paper has sought to curb estimates of the fatherhood bonus from multiple sources of selection bias, evidence from alternative causal designs should complement my findings. Examples of alternative designs could include quasi-experimental studies matching childless men and fathers on time-constant and, especially, time-varying confounders (e.g. Mincy et al., 2009 for the male marital wage premium), or approaches based on instrumenting the transition to parenthood (e.g. Cools et al., 2017; Kleven et al., 2018). Experimental studies, so far failing to detect a preference for fathers in terms of hiring chances in field settings (Correll et al., 2007; Bygren et al., 2017), could also complement the evidence of this study by investigating employers' wage offers to prospective male employees, depending on parental status. Additionally, while this study has focused on the transition to fatherhood and possible heterogeneities by cohort, future research could deploy the same stringent tests to assess how men's wages respond to higher-order parities and across other social groupings. Parity- and group-

specific mechanisms should of course motivate these analyses.

All in all, the evidence amassed in this study cannot support the idea of a causal wage premium for fathers in Germany and the UK. One consequence is that, as much as gender wage gaps in the two countries are driven by the transition to parenthood, the penalty for mothers rather than the premium for fathers really drives the gap in labor markets. The wage response to motherhood, however, has yet to be put to test using FEIS, accounting thereby for heterogeneous individual wage growth (but see Loughran and Zissimopoulos, 2009; Livermore et al., 2011). If, on average, mothers-to-be are on worse (better) wage growth paths than women who will not become mothers, motherhood penalties could have actually been overestimated (underestimated) in the literature so far. This study motivates future research into the reciprocal relationship between wages and motherhood too.

Absent wage boosts for fathers, nonetheless, it appears wage losses for mothers will not be compensated for within the confines of households alone. Evidence on men's wages in this paper thus indirectly calls for continued research on those policies that may mitigate the motherhood wage penalty and reduce gender economic inequalities more at large. Notably, my findings for Germany and the UK are in line with those of similar event-study designs for Scandinavian countries (for Denmark, Kleven et al., 2018; for Norway, Cools et al., 2017), countries that have a much longer tradition of support for men's role as carers and women's role as earners. As increasingly inclusive care policies are implemented in former male-breadwinner societies too, especially in the form of parental leaves, future research could evaluate their impact on men's labor market outcomes (e.g. for Scandinavia, Rege and Solli, 2013; Albrecht et al., 2015; Evertsson et al., 2016). Evidence in this study cannot support a causal story linking wages and fatherhood *per se*, after all.

Tables & graphs

Table 1: Summary of sample restrictions and relative sample sizes.

	SOEP		BHPS-UKHLS	
	<i>N</i> person-years (<i>N</i> individuals)	Percent	<i>N</i> person-years (<i>N</i> individuals)	Percent
All men aged 20-50	146,334 (20,109)	100	74,462 (10,883)	100
In dependent employment	106,054 (16,452)	72.5 (81.8)	53,550 (8,544)	71.9 (78.5)
Complete fertility history	88,088 (11,879)	60.2 (59.1)	52,917 (8,227)	71.1 (75.6)
First-time fathers and childless men	58,625 (7,591)	40.1 (37.7)	33,902 (5,395)	45.5 (49.6)
Observed at least until age 40	37,076 (3,236)	25.3 (16.1)	17,437 (1,649)	23.4 (15.1)
Non-missing on current wage	35,879 (3,156)	24.5 (15.6)	16,197 (1,536)	21.7 (14.1)
Observed for at least three waves	35,297 (2,743)	24.1 (13.6)	15,730 (1,283)	21.1 (11.8)
No child when first observed	34,879 (2,709)	23.8 (13.5)	15,454 (1,251)	20.7 (11.5)

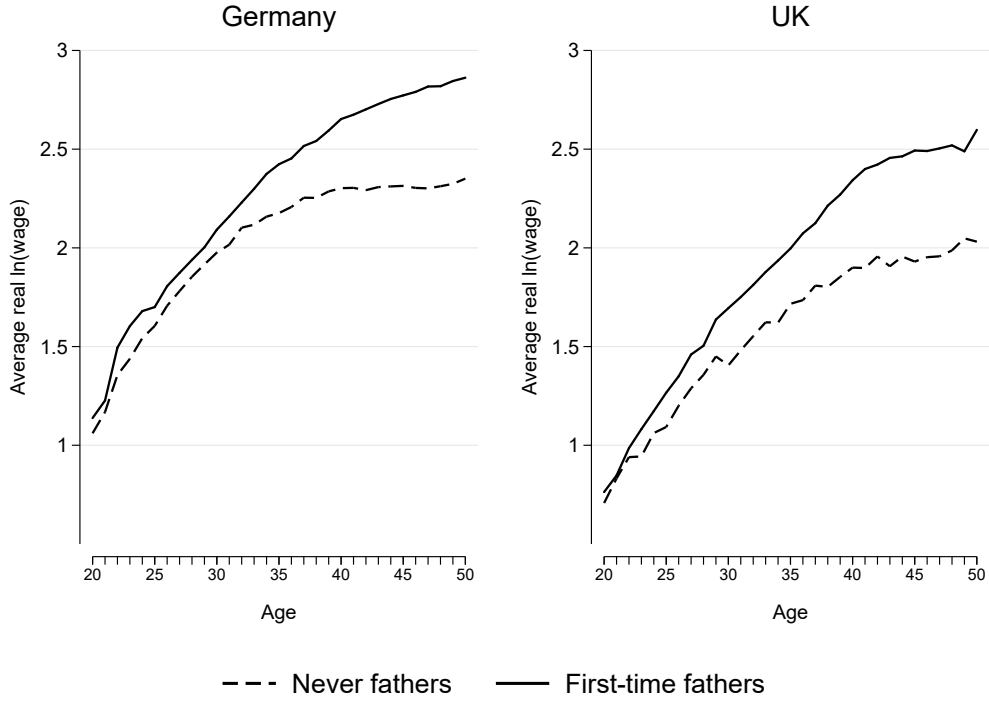


Figure 1: Average log hourly wages by age of the respondent, separately for childless men and fathers. Sources: SOEP 1984-2014, BHPS 1991-2016.

Table 2: OLS, FE, and FEIS for the log of real hourly wages. German men (SOEP 1984-2014).

	(1) OLS β (SE)	(2) FE β (SE)	(3) FEIS β (SE)
First-time father (<i>ref.</i> childless)	0.146*** (0.005)	0.042*** (0.011)	-0.007 (0.010)
$r(\theta_i, Child_{it})$		0.23	
R^2 (within R^2 in Columns 2 and 3)	0.46	0.61	0.10
Number of individuals	2,743	2,743	2,743
Number of person-years	34,879	34,879	34,879

Notes: All models include period dummies and a quadratic for age. The latter is allowed to vary across individuals in the FEIS specification. Standard errors are clustered at the individual level.

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

Table 3: OLS, FE, and FEIS for the log of real hourly wages. British men (BHPS 1991-2016).

	(1)	(2)	(3)
	OLS	FE	FEIS
	β (SE)	β (SE)	β (SE)
First-time father (<i>ref.</i> childless)	0.177*** (0.008)	0.026 (0.014)	-0.003 (0.013)
$r(\theta_i, Child_{it})$		0.23	
R^2 (within R^2 in Columns 2 and 3)	0.51	0.74	0.05
Number of individuals	1,251	1,251	1,251
Number of person-years	15,454	15,454	15,454

Notes: All models include period dummies and a quadratic for age. The latter is allowed to vary across individuals in the FEIS specification. Standard errors are clustered at the individual level.

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

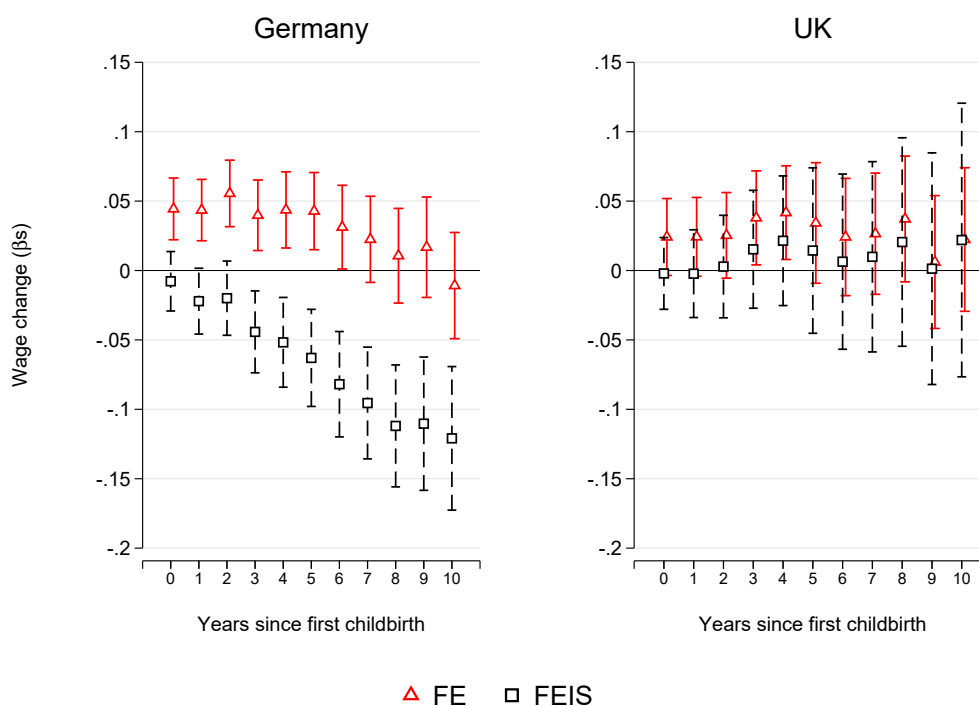


Figure 2: Point estimates and 95% CI for the event study of fatherhood and wages. fixed effects (FE) estimates and fixed effects individual-slope (FEIS) estimates on display. Models are detailed in the main text. *Sources:* SOEP 1984-2014, BHPS 1991-2016.

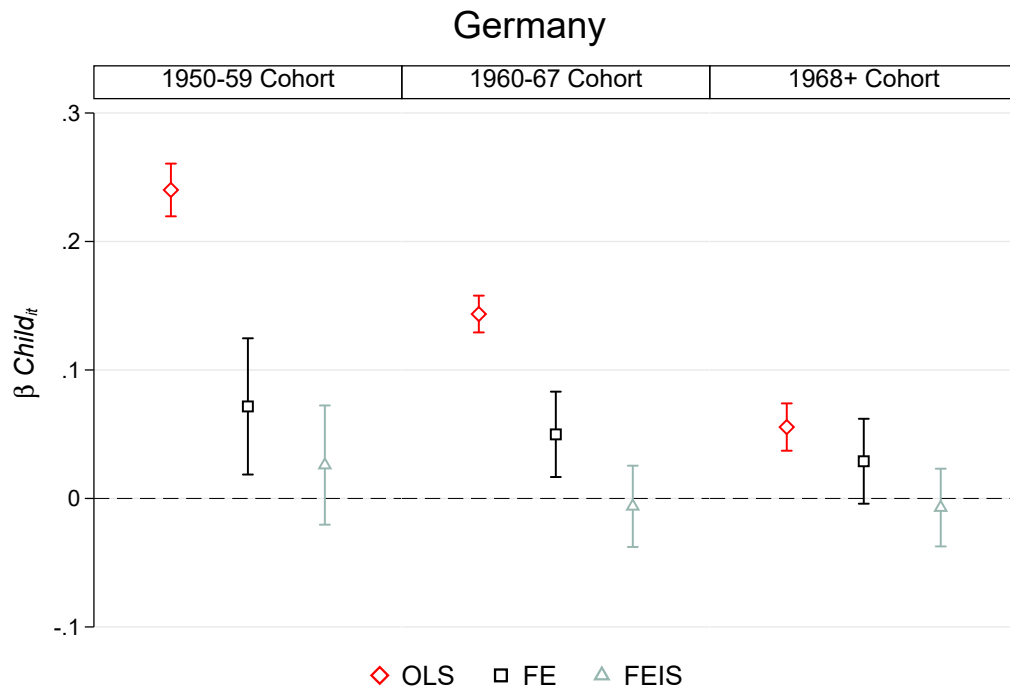


Figure 3: Point estimates and 95% CI for the coefficient of first-time fatherhood on log wages. Separate models for different birth cohorts. Sources: SOEP 1984-2014.

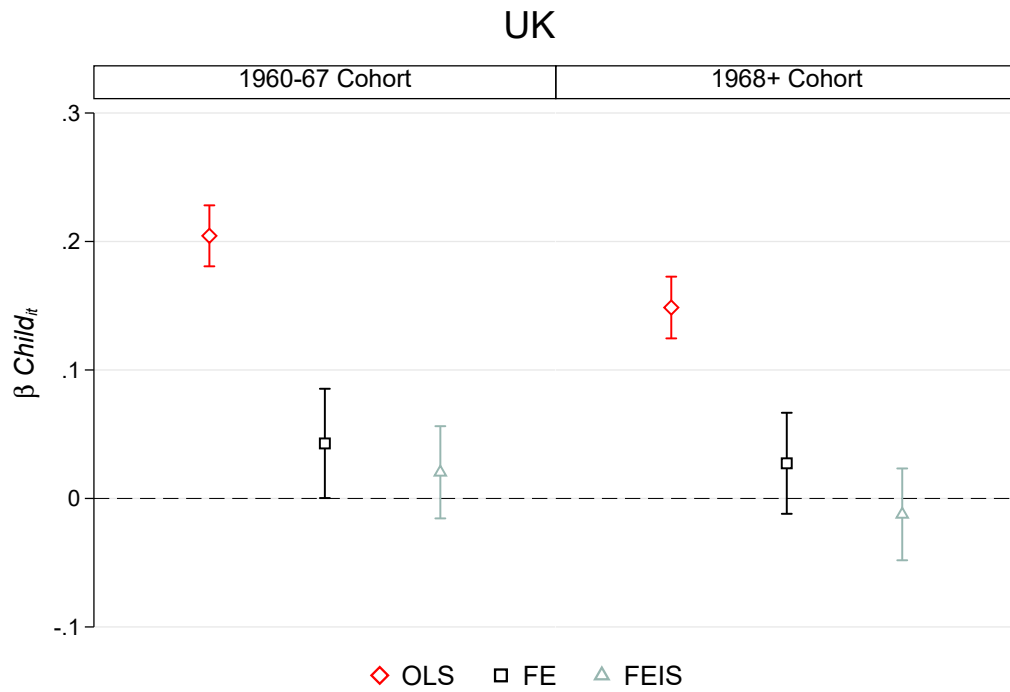


Figure 4: Point estimates and 95% CI for the coefficient of first-time fatherhood on log wages. Separate models for different birth cohorts. Source: BHPS 1991-2016.

Appendices

A. Cell size for each dummy k after first childbirth

Table 1A: Number of first-time fathers (treated units) for each dummy k after first childbirth (Equations 3 and 4 in the main text).

	SOEP	BHPS-UKHLS
<i>Years since first childbirth</i>	N first-time fathers	N first-time fathers
Year 0	906	496
Year 1	909	495
Year 2	901	487
Year 3	891	469
Year 4	850	451
Year 5	841	431
Year 6	808	400
Year 7	762	380
Year 8	740	370
Year 9	686	332
Year 10+	4,400	1,791

B. With and without restrictions on the number of waves

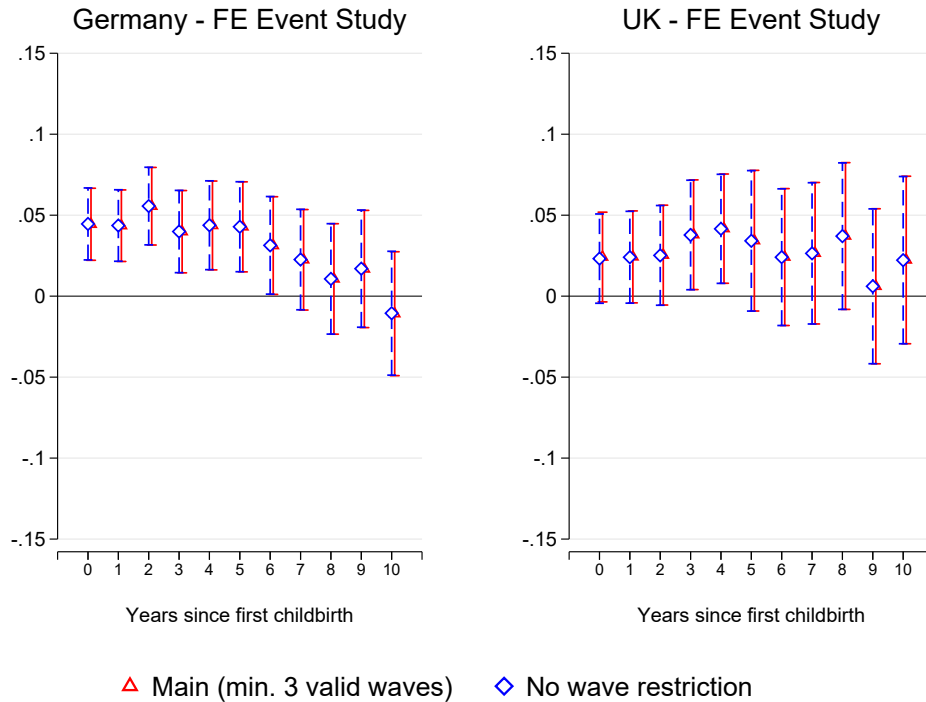


Figure 1A: Point estimates and 95% CIs. FE model specifications (Equation 3, e.g. Figure 2 in the main text) with and without restrictions on the number of waves. Sources: SOEP 1984-2014, BHPS 1991-2016.

C. Event-study dummies: investigating model dependence

In the following I replicate my event-study analyses displayed in Figure 2 in the main text. Estimates corresponding to the main analyses are henceforth labeled “Main”. I contrast here three alternative ways of dummy-coding the years k since the birth of the first child:

1. I include dummies for $k = 0, 1, \dots, 10$, but, differently from “Main”, I do not set $k_{10} = 1$ if $k > 10$ (“No cap at $k = 10$ ”);
2. I exclude observations for fathers after $k = 10$ (“Trim if $k > 10$ ”);
3. Last (“Full set of dummies”), I include all dummies k for which at least 140 first-time fathers were observed in the panel (capping at $k = 20$ for SOEP, $k = 15$ for BHPS).

Estimates in FE specifications are substantially independent from such specification choices. Estimates for FEIS specifications vary to a greater extent, but mainly with respect to the uncertainty surrounding point estimates in later years after childbirth.

Regardless, no specification choice for FEIS is compatible with a substantial, causal fatherhood wage bonus in the years immediately following the transition to fatherhood. Also, uncertainty surrounding estimates may stem from the fewer number of units (fathers) that are observed in later years after childbirth. The FEIS procedure requires to “detrrend” both the outcome and each predictor, and so each of the dummies k is detrended via individual-specific OLS regression models. The predictions of these OLS models, then plugged in for the estimation of FEIS coefficients, are likely to be more and more imprecise the less the variation in k , i.e. the lower the number of fathers that are observed in that particular year k after first childbirth.

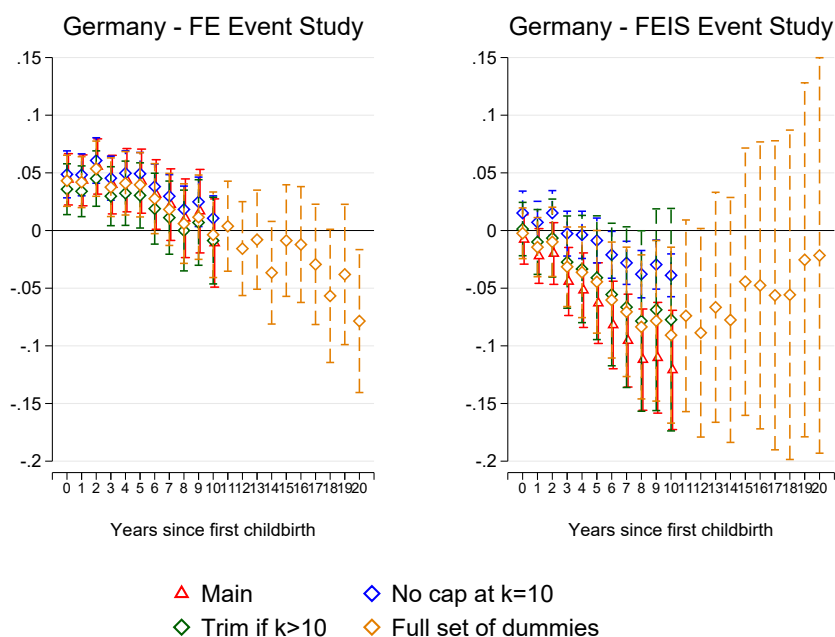


Figure 2A: Point estimates and 95% CIs. Source: SOEP 1984-2014.

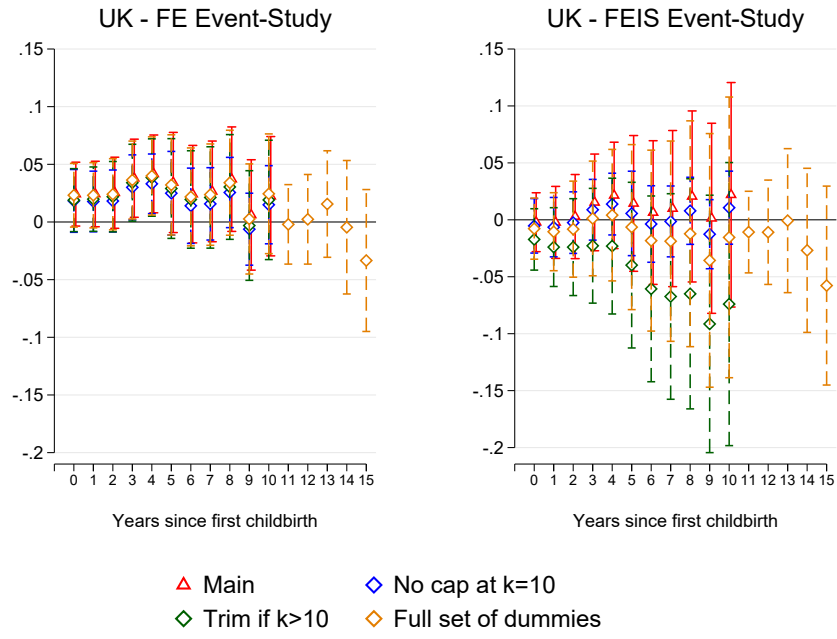


Figure 3A: Point estimates and 95% CIs. Source: BHPS 1991-2016.

D. Adding a dummy for marriage

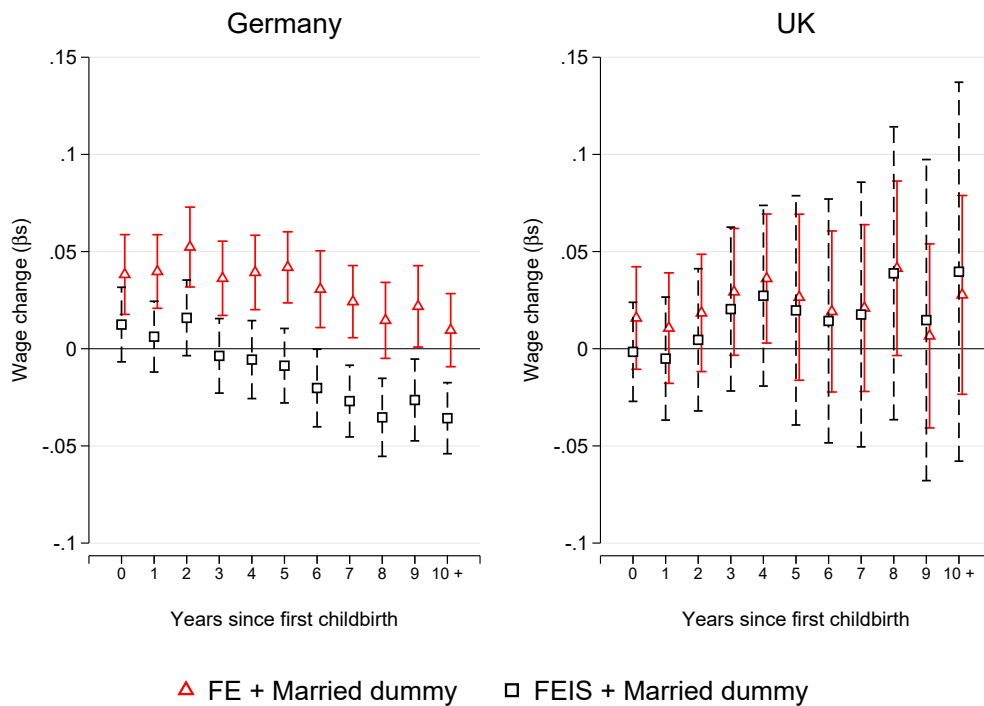


Figure 4A: Point estimates and 95% CIs. FE and FEIS model specifications (Equations 3 and 4, respectively; e.g. Figure 2 in the main text) adjusting for marriage (1 if in a union, 0 otherwise). Sources: SOEP 1984-2014, BHPS 1991-2016

E. Cross-cohort event-study design

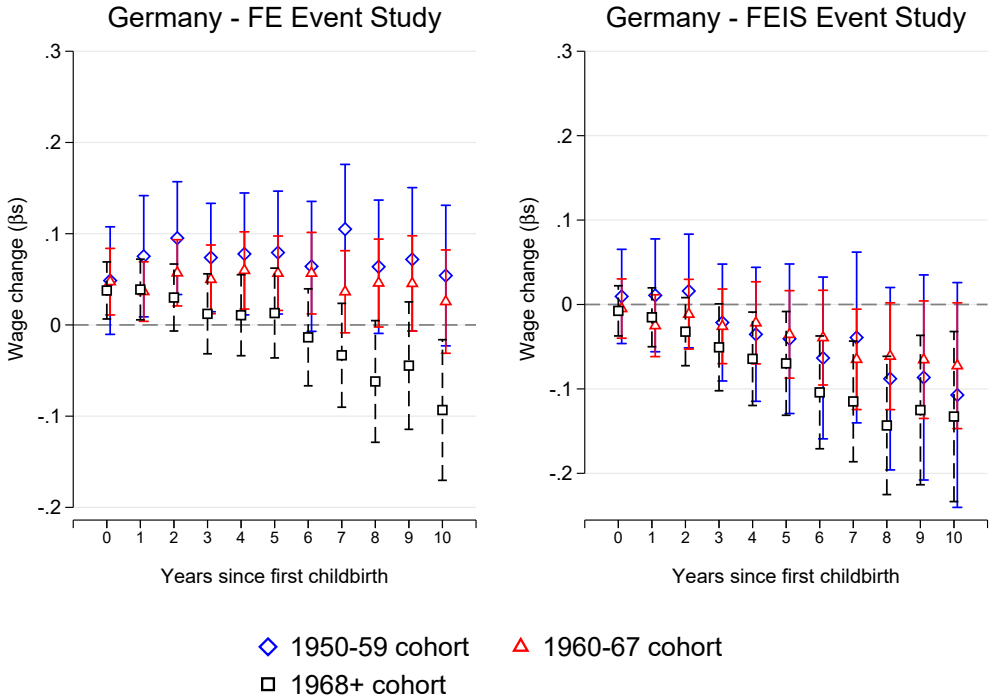


Figure 5A: Point estimates and 95% CIs for the event-study design by men’s birth cohort. Source: SOEP 1984-2014.

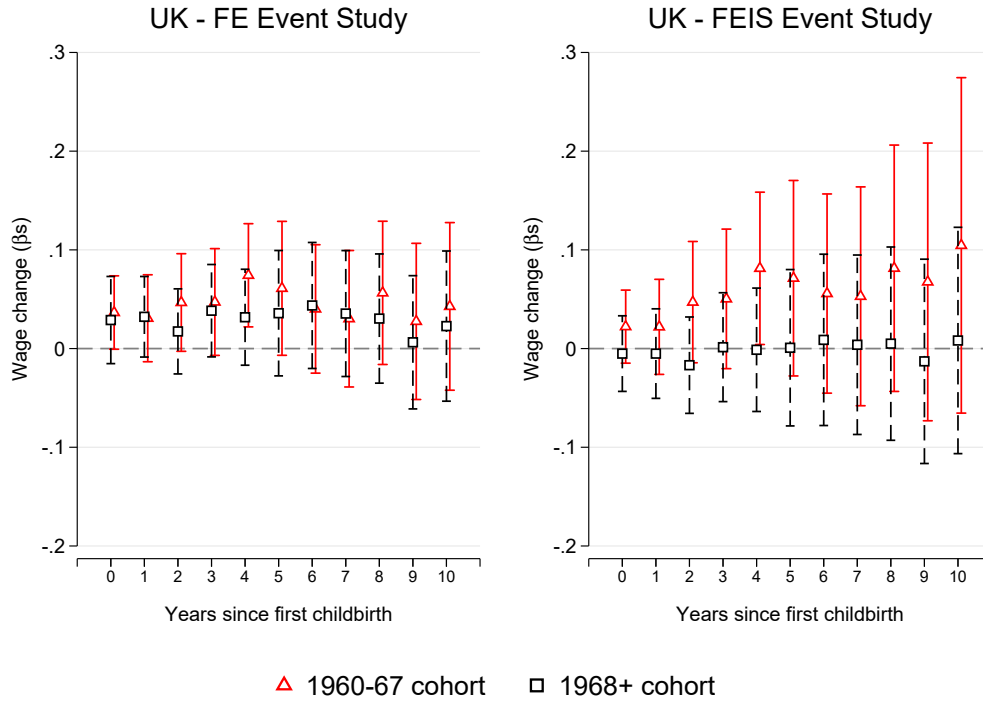


Figure 6A: Point estimates and 95% CIs for the event-study design by men’s birth cohort. Source: BHPS 1991-2016.

F. Missing fertility histories in SOEP

Only men who have been interviewed in 2000 or later completed fertility history questionnaires for SOEP. Men who dropped out of the panel prior to 2000 cannot therefore participate to the estimates displayed in the main text. Table 1 in the main text reports a substantial loss of around 12.5% of the potential sample when the restriction on the availability of fertility info is applied.

Hence, I have replicated here my analyses exploiting info contained in the main questionnaire. Relying on the variable for the “the number of children in the household”, I code it 1 after first childbirth and 0 otherwise. Applying the same sample restrictions utilised for the main analyses, with the exception of the one pertaining fertility histories, I am able to retain 5,223 individuals (and 41,170 person-year records), of which 1,794 will become first-time fathers in the observation period.

Figure 7 portrays FE and FEIS specifications in the augmented German sample. FE estimates are similar in magnitude and uncertainty to those reported in the main text. FEIS estimates, differently, are positive in the years following childbirth, although the pattern is not clear-cut (e.g. β for year 0 = .01, $p = .487$, while β for year 1 = .02, $p = .017$).

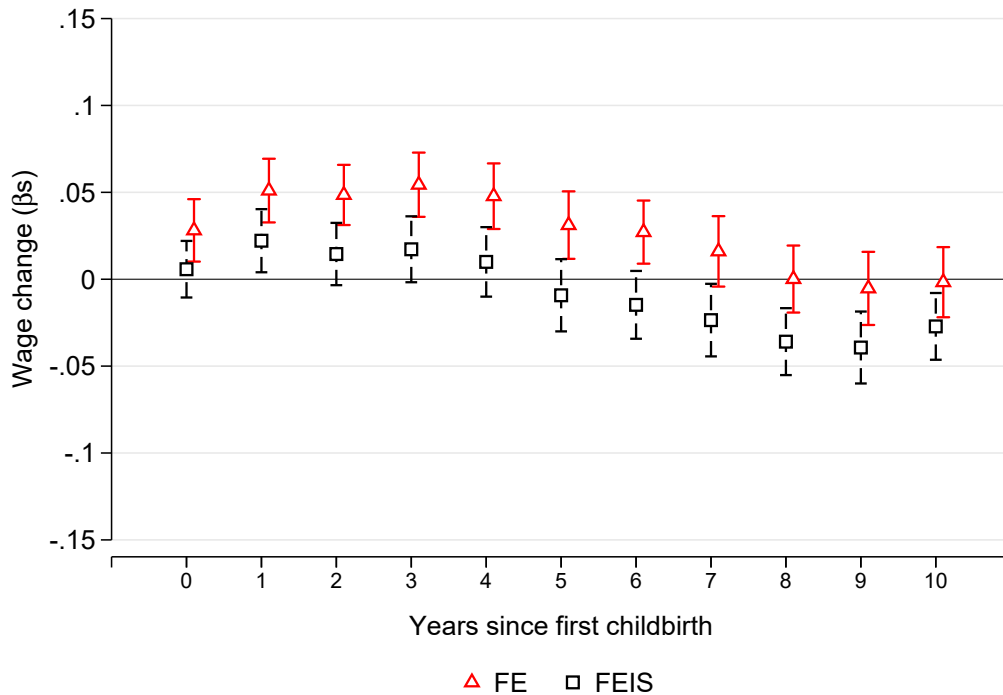


Figure 7A: Point estimates and 95% CIs for men in the augmented German sample. Source: SOEP 1984-2014.

This difference could be due to the composition of the new sample, now encompassing a higher portion of men belonging to older cohorts as compared to the estimation sample in the main text. To investigate this intuition, I have replicated the event-study design, split by birth cohort, using the new augmented sample. Results for the FEIS specification in Figure 8 support the idea that small causal premiums (at best, $\approx 4\%$) may resist adjusting for selection on prior wage growth, only among older cohorts. For men in the youngest cohort, estimates are instead indistinguishable from the average estimates reported in the main test. As reported in the main text then, one can conclude that fatherhood wage premiums are at best a “thing of the past” in Germany.

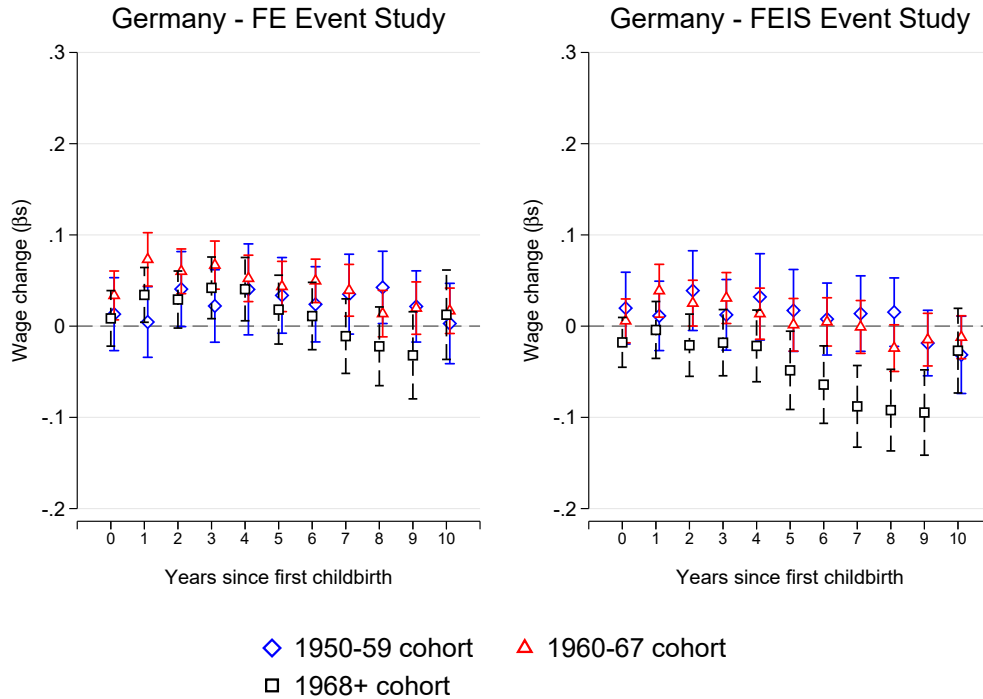


Figure 8A: Point estimates and 95% CIs for men in the augmented German sample, by men's birth cohort. Source: SOEP 1984-2014.

G. Adding dummies for the years preceding fatherhood

Event-study designs are typically specified by including dummies for years since the event of interest as well as prior to it. Dummies for the years prior to fatherhood, however, are not included in the paper as they would be superfluous in FEIS specifications. Concentrating on the years after the transition to fatherhood only, FE and FEIS specifications are as similar as possible in the paper, and this helps comparing them.

Nonetheless, I here contrast my “Main” estimates (FE specification, Figure 2 in the text) with the ones that can be obtained running a fully dynamic event-study specification. In line with common practices in the literature, I include observations for the treated group for each year up to the fifth prior to first childbirth and up to the tenth after. Person-year records in which the treated were observed earlier or later than those moments are dropped. The year preceding childbirth ($k = -1$) serves as the reference category.

To a somewhat lesser extent in the German sample, estimates in the two specifications closely resemble one another (Figure 9). Once again, evidence from FE specifications is at best compatible with small wage premiums for fathers (particularly in Germany).

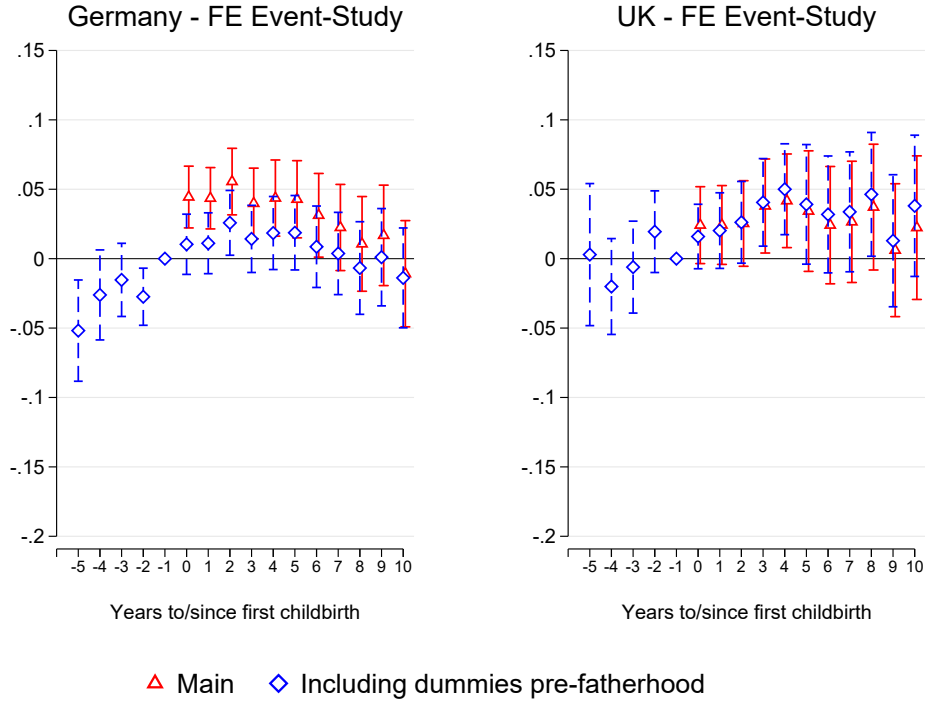


Figure 9A: Point estimates and 95% CIs for the main model (Figure 2 in the main text) and a FE model including dummies for up to the fifth year prior to the transition to fatherhood ($k = -1$ is the reference category). Sources: SOEP 1984-2014, BHPS 1991-2016.

H. Age restrictions and suitable control groups

My final analytical sample only includes men who have been observed at least until age 40. This sample restriction is necessary to define the preferred control group consisting of childless men. Yet it is not without consequences, first and foremost bringing about a substantial sample size loss (see Table 1, main text). Further, it may introduce bias by what could be described as “conditioning on future outcomes”. If the argument about selection holds, men who are high-earning and/or on steep wage growth paths are more likely to become fathers. In a sense then, the control group I chose disproportionately consists of men who will never attain wages high enough (or a wage growth steep enough) to make them “suitable” for the transition to fatherhood. I am thus conditioning the definition of the control group, being childless at age 40, to men’s wage levels and trajectory, their “future outcomes” I can expect under a scenario of positive selection into fatherhood. In a way, I may thus be contrasting “lucky” men (fathers-to-be) to “unlucky” men (never fathers), luck being men’s wage attainment. It follows that this may, by design, lead to overestimate wage premiums related to fatherhood. As a sensitivity check, I re-estimate all the main models dropping the sample requirement to observe men at least until age 40. After applying all the other sample restrictions as per Table 1, I can rely on a much

bigger analytical sample: 53,277 person-year records (5,343 men) for SOEP and 27,518 person-year records (2,790 men) for BHPS.

All the main analyses are here replicated, in Figure 10 for Germany and in Figure 11 for the UK. Note that they now refer to a different contrast between groups, as the wage attainment of first-time fathers is compared to men who will never become fathers *or* may become fathers in the future. The latter may simply experience the transition to fatherhood at a later date than when last observed in the panel. From a causal inference perspective, this contrast provides an estimate of the effect of becoming a father v. not having become a father yet. All the panels in both figures closely resemble the main analyses, highlighting indeed, if anything, some overestimation of the wage premium in FE models for the German sample.

Ultimately, this alternative design strengthens the paper's main conclusions, that there is no strong evidence pointing to a causal effect of fatherhood on the wages of German and British men.

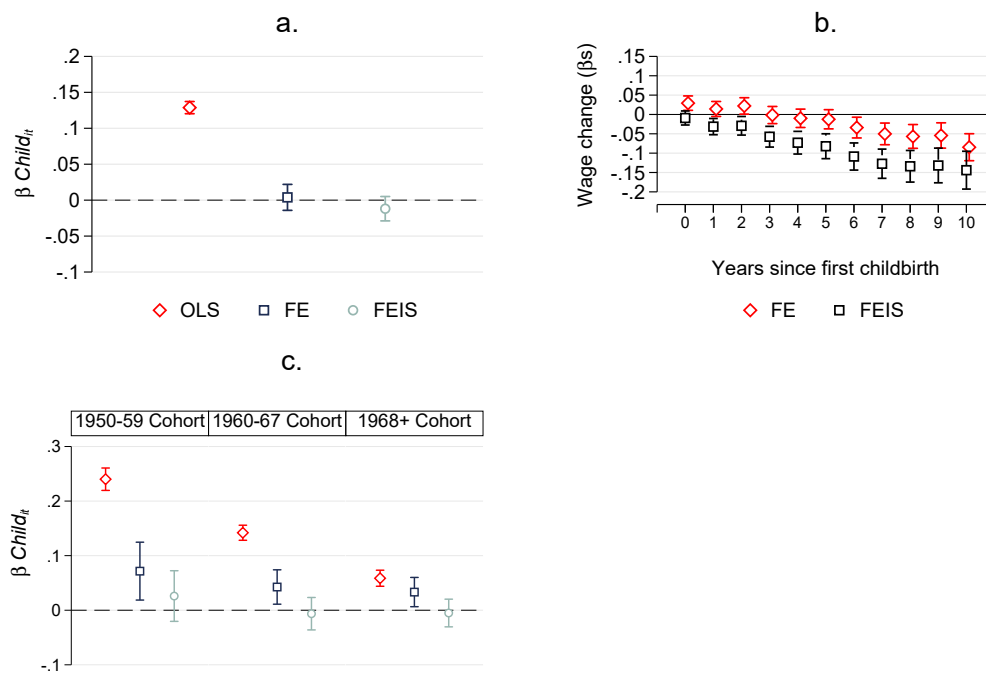


Figure 10A: Point estimates and 95% CIs for the effect of fatherhood on wages in the German sample, dropping age restrictions used to define the analytical sample in the main text. Panel a. replicates the analyses displayed in Table 2 in the main text. Panel b. replicates the event-study estimates displayed in Figure 2 in the main text. Panel c. replicates the cross-cohort analyses of Figure 3 in the main text.

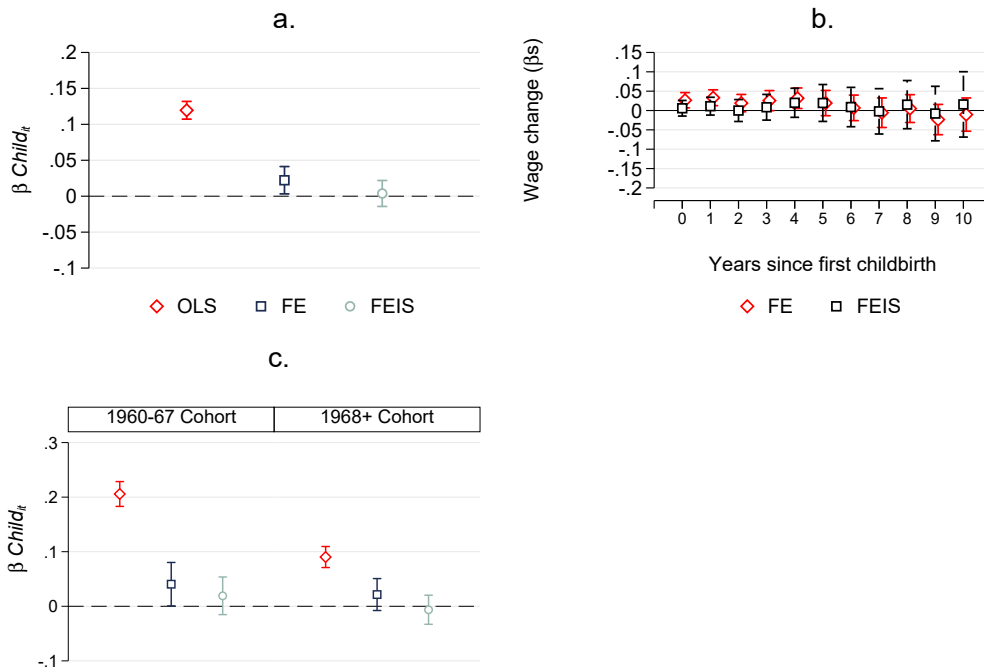


Figure 11A: Point estimates and 95% CIs for the effect of fatherhood on wages in the British sample, dropping age restrictions used to define the analytical sample in the main text. Panel a. replicates the analyses displayed in Table 3 in the main text. Panel b. replicates the event-study estimates displayed in Figure 2 in the main text. Panel c. replicates the cross-cohort analyses of Figure 4 in the main text.

Chapter 2

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

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

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.

*This paper is co-authored with Giorgio Cutuli. 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. A slightly different version of this paper is currently under review at an international peer-reviewed journal. The paper also appeared as part of the SOEPPapers collection (dp1025, 2019)

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; Kluge 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 (Kluge 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 Kluge 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 (Kluge and Tamm, 2013; Ziefle and Gangl, 2014; Bergemann and Riphahn, 2017; Kluge 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; Kluge 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; Kluge 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. $tw_{D_i=1} = \frac{1}{P(D_i|Z_i)}$) 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; Kluge 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; Kluge 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 (Kluge 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; Kluge 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.

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.

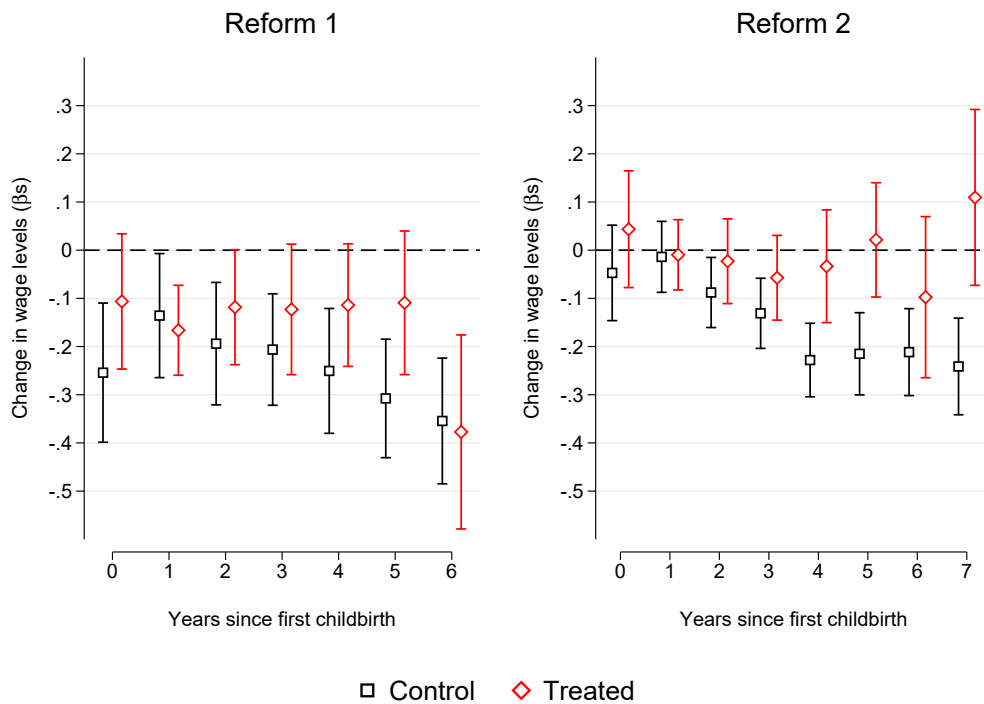


Figure 1: FE estimates (95% confidence intervals) of the motherhood wage penalty across reforms, treatment groups, and relative to event time. Estimates are weighted by IPTW, as detailed in the main text (SOEP 1985-2014).

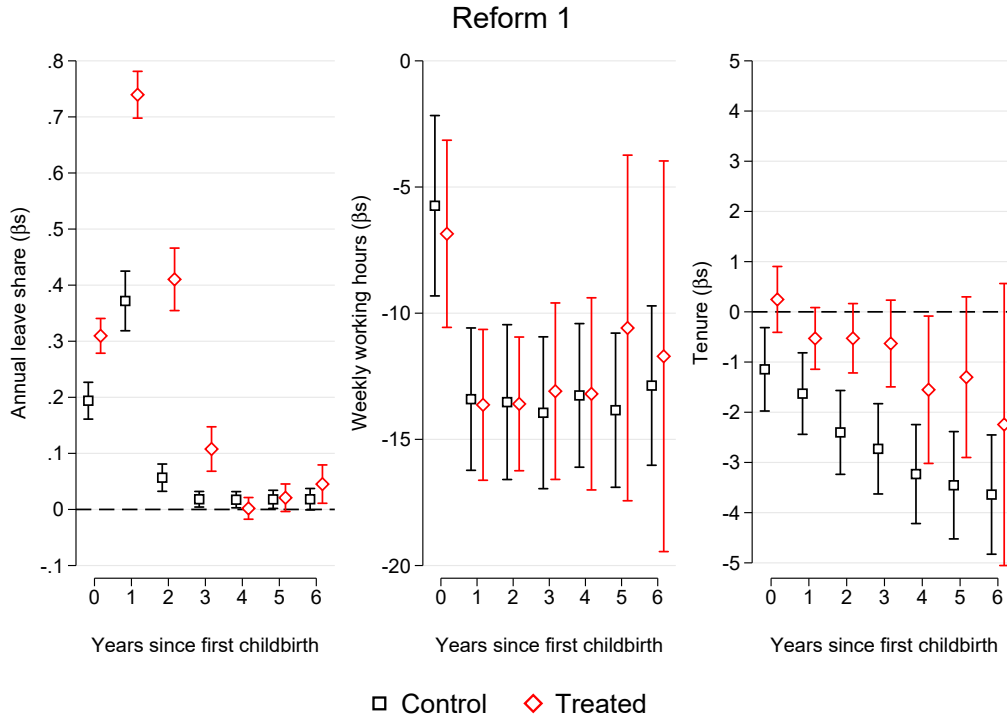


Figure 2: FE estimates (95% confidence intervals) of motherhood effects on (i) leave share, (ii) working hours, and (iii) tenure. Estimates refer to Reform 1, across treatment groups and relative to event time. Estimates are weighted by IPTW, as detailed in the main text (SOEP 1985-1998).

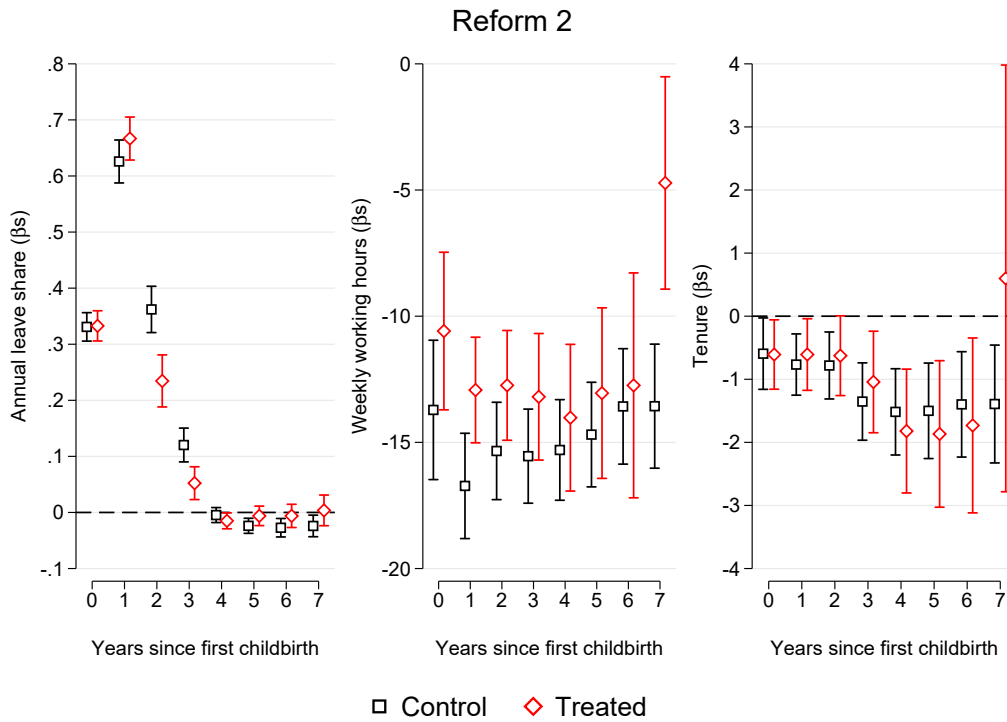


Figure 3: FE estimates (95% confidence intervals) of motherhood effects on (i) leave share, (ii) working hours, and (iii) tenure. Estimates refer to Reform 2, across treatment groups and relative to event time. Estimates are weighted by IPTW, as detailed in the main text (SOEP 1999-2014).

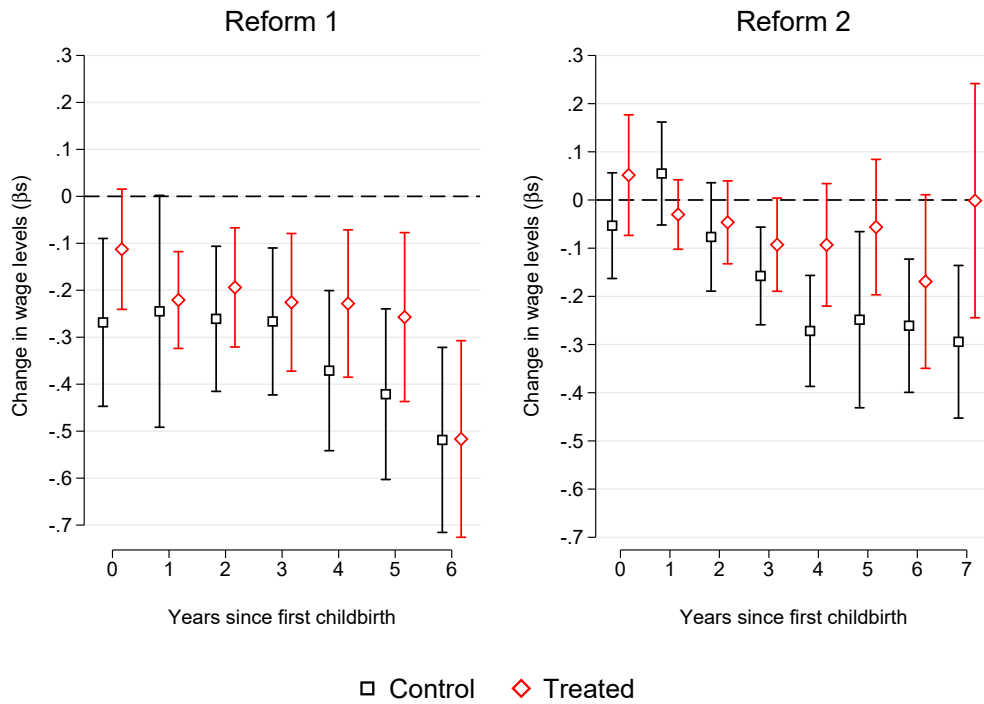


Figure 4: FE estimates (95% confidence intervals) of the motherhood wage penalty across reforms, treatment groups, and relative to event time. Estimates are weighted by means of entropy-balancing weights, as detailed in the main text (SOEP 1985-2014).

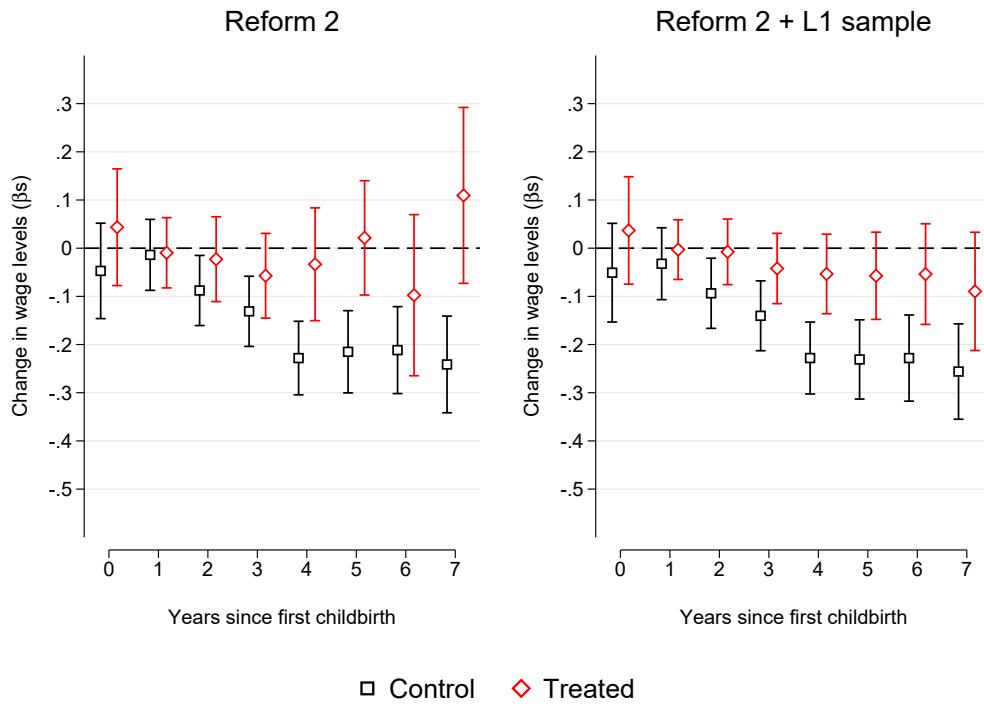


Figure 5: FE estimates (95% confidence intervals) of the motherhood wage penalty across treatment groups and relative to event time. Estimates on the left panel are weighted by means of IPTW, as detailed in the main text (SOEP 1985-2014).

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

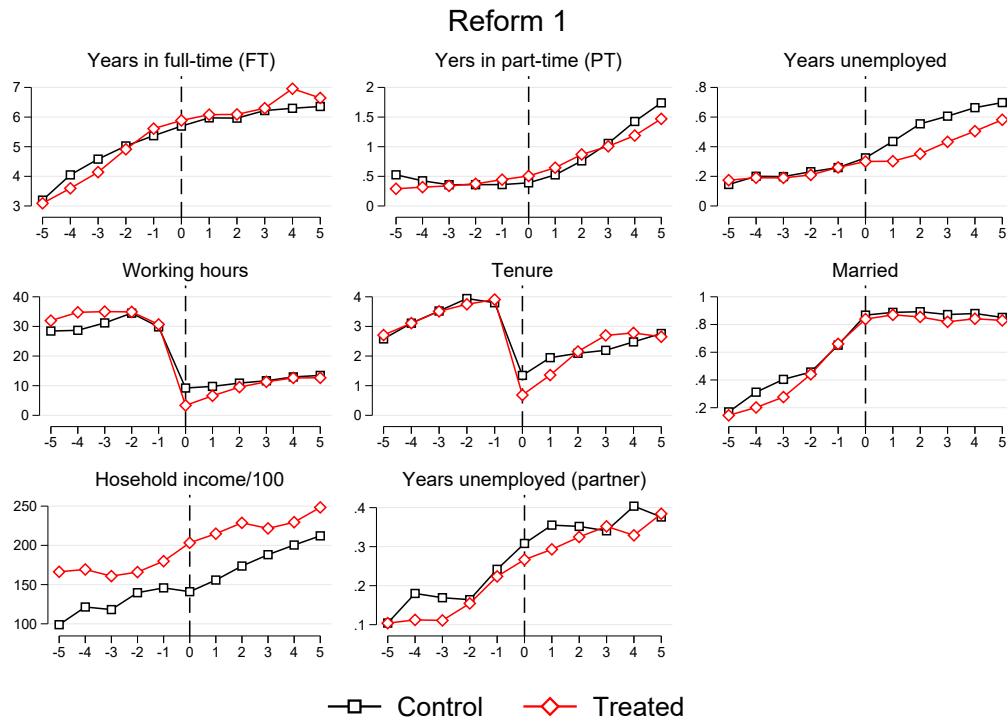


Figure 1A: Means of selected variables for treated and control units (Reform 1, SOEP 1985-1998). Means are computed for each year up to the fifth year prior (after) first childbirth. All means are weighted by IPTW as detailed in the main text.

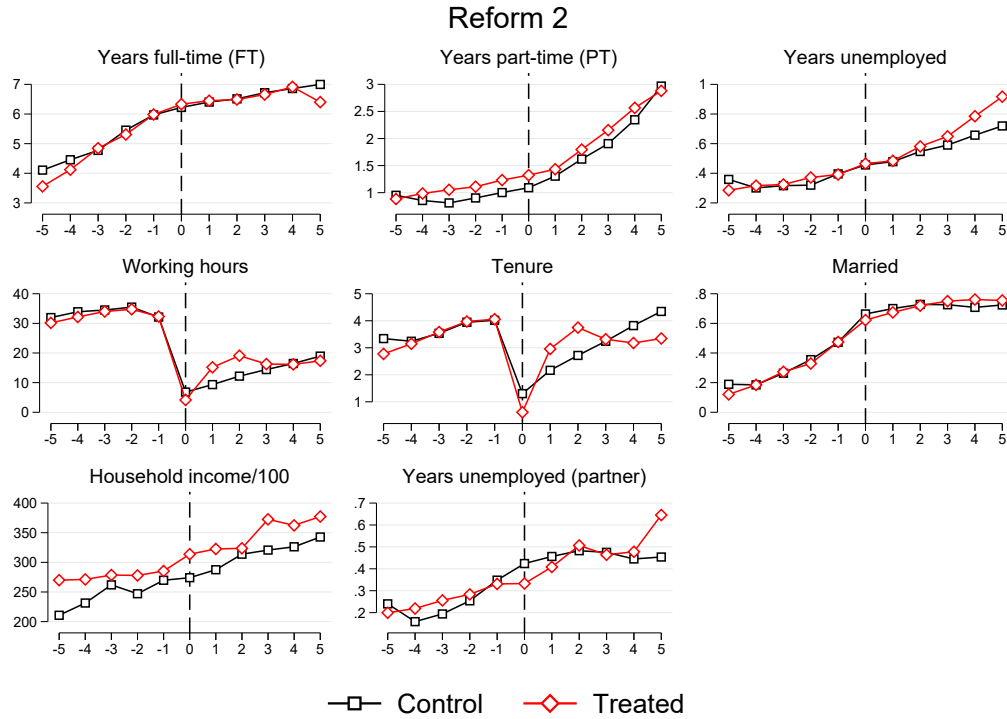


Figure 2A: Means of selected variables for treated and control units (Reform 2, SOEP 1999-2014). Means are computed for each year up to the fifth year prior (after) first childbirth (x axis). All means are weighted by IPTW as detailed in the main text.

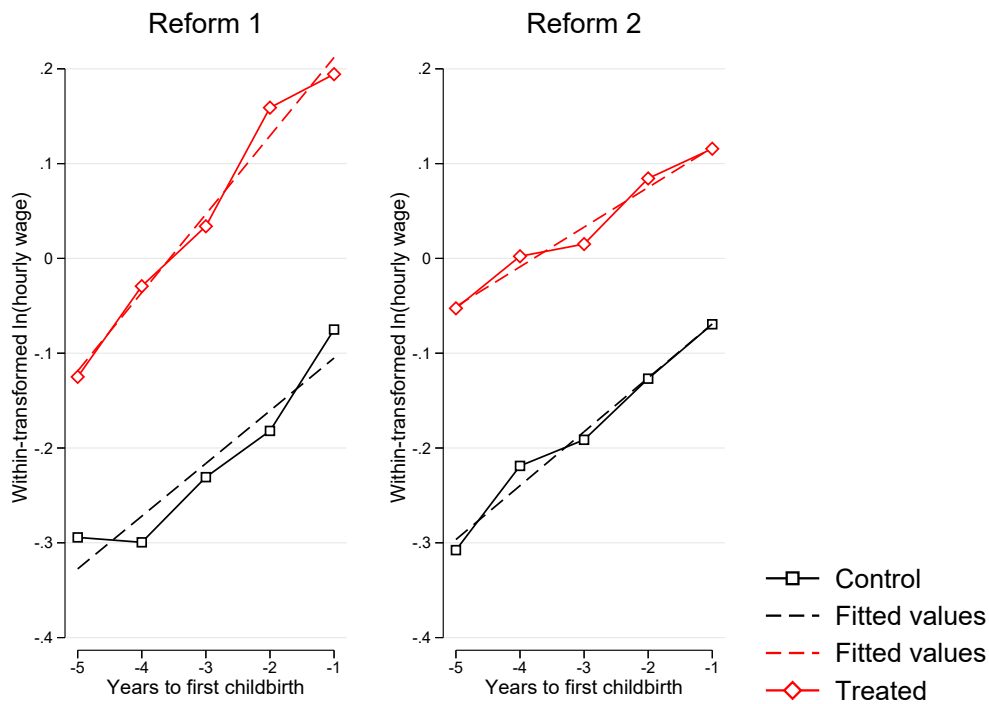


Figure 3A: Within-transformed log of real hourly wages plotted over years to first childbirth, for treated and control units across reform periods. Dashed lines are obtained fitting a linear trend for each treatment-reform group using Stata's `lfit` (SOEP 1985-2014).

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 β (SE)	Treated β (SE)	Difference 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)	(2)	(3)
	Controls β (SE)	Treated β (SE)	Difference 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 β (SE)	Treated β (SE)	Difference 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) Controls β (SE)	(2) Treated β (SE)	(3) 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 β (SE)	Treated β (SE)	Difference 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)	(2)	(3)
	Controls β (SE)	Treated β (SE)	Difference 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 β (SE)	(2) Treated β (SE)	(3) Difference 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) Controls β (SE)	(2) Treated β (SE)	(3) Difference 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).

Chapter 3

Policy, Compensating Differentials, and Gender Career Gaps: Evidence from a 'Right-to-Request' Reform

Policy, Compensating Differentials, and Gender Career Gaps: Evidence from a ‘Right-to-Request’ Reform*

Abstract

A commonly held account of the family gap in labour markets is that mothers favour flexible working-time arrangements over career attainment, yet little is known on whether this compensating differential is shaped directly by public policy. Relying on panel data and a difference-in-difference design, I examine the introduction of a ‘right to request’ flexible schedules for parents of small children in Britain. Fitting the theory of compensating differentials, mothers experience wage cuts combined with a reduction in working hours and accrued satisfaction with working-time arrangements. This is especially the case for mothers of children aged 0-2 at the time of the reform. Evidence also suggests that the negative economic impact of the reform might have deepened gender gaps in the British labour market. For mothers of older children though, I can only detect wage losses and not reductions in working hours or increases in satisfaction with working time.

*Data from the BHPS were made available through the UK Data Archive (University of Essex, Institute for Social and Economic Research, 2018). Neither the original collectors of the data nor the archive bear any responsibility for the analyses or interpretations presented here. I wish to thank Melanie Jones as well as participants in an invited seminar at the University of Bath (February 2018) and in the Gender Economics and the Workplace workshop (Nuremberg, November 2018) for their comments on previous versions of this paper. The paper is currently being prepared for submission to an international peer-reviewed journal.

A prominent explanation for persisting gender gaps in labour markets holds that women may favour flexible working-time arrangements even when less lucrative (Goldin, 2014). If flexible schedules combine with and offset lower career attainment (especially, lower wages), compensating differentials are said to arise. Employed mothers in particular may be willing to forgo their position on the career track and move toward flexible jobs, as a strategy to juggle childcare duties and paid work. Other than accounting for career gaps between women and men, compensating differentials may thus underlie motherhood penalties, that is, how women's career prospects worsen after as compared to before the birth of a child (Felfe, 2012a,b; Hotz et al., 2017).

Little is known however on if and how public policy may alleviate or rather reinforce the career costs attached to working flexibly (cf. Smith, 1979). Studies on Scandinavian countries often suggest that high-quality part-time may generate compensating differentials to the detriment of women's career attainment (Albrecht et al., 2003; Hardoy et al., 2017). In countries where part-time is of much poorer quality, its compensatory value is questionable. The UK presents a puzzle in this respect, with female part-timers clearly disadvantaged as compared to their full-time counterparts – on pay, training, promotion prospects – yet reporting higher job satisfaction (e.g. Arulampalam and Booth, 1998; Manning and Petrongolo, 2008; Booth and Van Ours, 2008; Matteazzi et al., 2014). In 2003, Britain introduced a 'right to request' changes to working-time arrangements, making it easier for parents of young children to work flexibly with their current employer. I ask therefore, first, if such a policy helped install compensating differentials, fostering moves to and satisfaction with flexible work arrangements at the expense of career attainment, particularly for mothers of small children. Second, I investigate whether these newly-minted compensating differentials deepened career gaps between women and men.

Taken together, I exploit the introduction of this new right to assess how working-time flexibility affects career gaps among women with and without children, as well as between women and men. The right-to-request policy under examination, implemented also in other European countries (e.g. Hegewisch, 2005; Fernández-Kranz and Rodríguez-Planas, 2013; Begall

and Grunow, 2015), ascribes in fact to a broader class of ‘women-’ and ‘family-friendly’ policies aimed at sustaining women’s participation in paid employment. Research has consistently suggested that such policies may achieve their purported goal while simultaneously hindering career attainment for employed women and mothers in particular (e.g. Ruhm, 1998; Albrecht et al., 2003; Arulampalam et al., 2007; Gupta et al., 2008; Blau and Kahn, 2013). I contribute here to this latter strand of literature, as few of these studies are longitudinal in design or attempt to evaluate the causal impact of a given policy.

1. Background

1.1. Flexible schedules and compensating differentials

As originally suggested by Smith (Smith, 1776), compensating differentials arise as jobs bundle together monetary and non-monetary features such as that losses on one end are counterbalanced by gains on the other (Rosen, 1986). Job amenities like flexible schedules might then be associated with lower wages and still attract workers that value flexibility over pay. Women and mothers in particular may discount career attainment in favour of family-friendly schedules, more so than men and fathers, opening up gender and family gaps in labour markets (Filer, 1985; Felfe, 2012b; Cha and Weeden, 2014; Wiswall and Zafar, 2017).

While some have argued for the ‘pervasive absence’ of compensating differentials in labour markets by analyzing men and women together (Bonhomme and Jolivet, 2009), other studies point to a trade-off between career attainment and flexibility for women and mothers specifically. As highlighted by Felfe for Germany (2012b), employed mothers display a ‘willingness to pay’ for flexible schedules (such as working during the evening or in rotating shifts) at the cost of lower wages, when also maintaining the right to return to their pre-pregnancy employer after taking parental leave. Coherent evidence also comes from Scandinavian countries, showing that motherhood exerts a (small) negative effect on wages only in the family-friendly public sector in Denmark (Simonsen and Skipper, 2006) or that job sorting accounts for most of the motherhood wage penalty in Norway (Petersen et al., 2010). Considering career pro-

gression, including that to supervisory and managerial roles, Kunze (Kunze, 2015) finds part-time work to explain part of the motherhood penalty she highlights, also for Norway (see also Hardoy et al., 2017).

Similar to their counterparts across Europe, British women work part-time in high proportions, reducing their working hours specially to reconcile work and family (Paull, 2008). In the British labour market though, part-time work is associated with particularly dismal working conditions compared to the aforementioned contexts of Germany or Scandinavia. Part-time jobs pay a substantial wage penalty to full-time equivalents (cf. Manning and Petrongolo, 2008; Hardoy and Schøne, 2006), are less likely to involve training (Arulam-palam and Booth, 1998), and are both vertically and horizontally segregated (Connolly and Gregory, 2008; Matteazzi et al., 2014). At the same time, British women holding part-time jobs report relatively high levels of job satisfaction, both as compared to full-timers (Booth and Van Ours, 2008) and to part-timers in other European countries (Gallie et al., 2016). Further, most female part-timers are ‘unconstrained’ in their job posts, meaning that they would not change their working hours if given the chance (Böheim and Taylor, 2004).

Combining career costs with high job satisfaction, part-time work may thus fit in a compensating differential story for British women and mothers in particular (see also Gangl and Ziefle, 2009). Such equilibrium may be sustained by a mix of institutional and cultural factors (e.g. Pollmann-Schult, 2016), from the lack of childcare coverage for children under 3 (Thévenon, 2011) to widespread social norms negatively sanctioning mothers of young children working full-time (O’Reilly et al., 2014). Whether regulations directly concerning part-time and flexible work also shape compensating differentials – and, as a consequence, gender gaps in the British labour market – is an empirical question.

1.2. The introduction of the ‘right to request’ flexible schedules

Under the provisions of the Employment Act, and effective since April 2003, parents of children under six years of age¹ have been granted the statutory right to request flexible working-time arrangements in Britain (Hegewisch, 2005). Eligibility criteria included being a dependent employee, continuously employed for at least 26 weeks, and not on a temporary-

work agency contract. Employers, who opposed more binding early drafts of this legislation (Lewis and Campbell, 2007), were charged with the duty to consider such requests and refuse them only on the basis of business reasons defined by the law. Following its implementation, monitoring suggested three stylized facts about the reform (Hegewisch, 2005; Hooker et al., 2011). First, awareness of the new right gradually built up and particularly among women, who requested flexible solutions primarily to deal with childcare. Second, the majority of requests under the new regime were fully granted by employers, with acceptance rates ranging from over 60% to 80% depending on the year and survey. Finally, a reduction in working hours and moves to part-time work were the most frequent changes requested by employees. The reform thus created a ‘before’ and ‘after’ for mothers of children under six. Adding new legal protections, it may have made part-time work more appealing and facilitated achieving hour flexibility with one’s current employer. This could have been of particular value in the UK, as hour *inflexibility* has been shown to be prominent within employers, meaning that workers who wish to adjust their working hours are better off switching employers altogether (Böheim and Taylor, 2004; Blundell et al., 2008). Women switch to part-time jobs across employers in particular after family-related work interruptions (Manning and Robinson, 2004; Connolly and Gregory, 2008), a move made more necessary given the aforementioned lack of coverage of childcare services for children aged 0-2.

If the reform thus fostered compensating differentials, I expect eligible mothers of children under six to become more likely to move to part-time jobs and experience accrued satisfaction with their working-time arrangements, after the reform and relative to women not eligible for the new right. These advantages should then mix up with career costs, such as a reduction in pay. As mothers more than fathers work flexibly, both generally and as a result of requests under the new law (Paull, 2008; Hooker et al., 2011), gender career gaps may have worsened after the reform as an unintended result of compensating differentials arising for women but not for men. Both within-gender and between-gender dynamics should be particularly pronounced when considering parents of children under 3.

2. Empirical approach

2.1. Data and samples

I rely on the British Household Panel Survey (BHPS), a multipurpose household panel focusing on the living conditions and life histories of UK residents running from 1991 to 2009 (Taylor et al., 2010; University of Essex, Institute for Social and Economic Research, 2018). The analyses are here restricted to waves 10 to 15 and thus cover the period 2000-2006. This time window excludes further normative changes, as the first expansion of the right was granted by the Work and Families Act of 2006 and came into effect in April 2007. The analytical sample comprises individuals aged 20 to 55, working as dependent employees, and with at least one valid person-year observation. Employees holding temporary-work agency (TWA) contracts were outside the scope of the right-to-request legislation, and person-year records in which individuals hold a TWA are therefore excluded from the analyses. Differently, and limiting the precision of my estimates, BHPS data does not allow me to precisely identify employees with at least 26 weeks of tenure, a second eligibility criterion. After retrieving individual fertility histories from the BHPS Consolidated Marital, Cohabitation, and Fertility Histories (1991-2009) file (Pronzato, 2011), I define women treated by the 2003 reform as those with at least one child under 6 years of age born prior to April 2003 and not giving birth in its aftermath. I thus focus on women with at least one of their youngest children born between 1998 and 2002. Such a restriction addresses two concerns. One is the potential for endogenous fertility behaviour among the treated, although, to the best of my knowledge, there is no evidence available on the fertility effects of this particular reform. Secondly, concomitant (and subsequent) reforms to maternity leave rights (Lewis and Campbell, 2007) may invalidate the evaluation strategy, yet they would necessarily apply only to women who give birth also from 2003 onwards, and these women are therefore not part of the analytical sample. The control group is drawn from the pool of ineligible women, thus comprising childless women and mothers of children aged 6 or older in 2003. Mirroring the selection made for the treated, mothers in the control group have given birth between 1993 and 1997 and not afterwards.

To assess whether the 2003 reform affected gender gaps in the British labour market, I follow the same sample restrictions to define treated and controls among men. Among the treated, I retain fathers of children under 6 years of age born prior to 2003. The control group comprises fathers of older children and childless men. Both within-gender and between-gender analyses are then repeated looking at more narrowly defined treated groups, one consisting only of parents of children born in 2001-2 and thus under 3 at the moment of the 2003 reform, and one consisting of parents of children born in 1998 to 2000 and thus aged 3 to 5 at the moment of the 2003 reform. All contrasts between treated and controls are summed up in Table 1, also reporting sample sizes in each cell.

2.2. Difference-in-differences and triple-difference strategies

Following the literature on compensating differentials (Rosen, 1986; Bonhomme and Jolivet, 2009; Felfe, 2012a), I estimate a number of empirical models, sharing the same specification, for outcomes capturing career attainment (e.g. hourly wages) on the one hand, and non-monetary job amenities (e.g. flexible schedules) on the other. Particularly for those women treated by the 2003 reform, I expect losses on the first class of outcomes balanced by gains on the second class, highlighting a trade-off between them.

To gauge whether, how, and for whom right-to-request policies gave rise to compensating differentials, I adopt a difference-in-differences (DD) framework. All models are linear or linear probability models for dichotomous outcomes, with the following specification for the 2003 reform

$$y_{it} = \alpha_1 POST03_t + \alpha_2 (POST03_t \times D_i) + \mathbf{X}'_{it} \gamma + \theta_i + \phi_t + \epsilon_{it} \quad (1)$$

where y_{it} is one of six outcomes. Monetary job features are classically involved in the trade-off underlying compensating differentials (Filer, 1985; Rosen, 1986) and are here modelled as the log of real² hourly wages. Hourly wages are obtained by dividing gross monthly pay by the sum of regular and overtime weekly working hours multiplied by 4.35 (e.g. Bryan and

Sevilla-Sanz, 2011). Wages below 1 or above 100 are trimmed.

Given that the expected trade-off centres around working hours and time flexibility, I also analyse career attainment in terms of access to managerial jobs, whose long hours are likely to prove incompatible with part-time work (Manning and Petrongolo, 2008; ONS, 2011). A second outcome in the analysis is therefore a dummy variable that equals 1 if individuals hold a managerial post and 0 otherwise, derived from a direct question concerning respondents' managerial responsibilities in their current job.

Moving to the other end of the trade-off, I first look at changes in working hours (excluding overtime) and transitions to part-time work (a dummy that equals 1 if employees work under 30 hours a week, e.g. Matteazzi et al., 2014). Further, I investigate subjective evaluations of one's current employment situation. Self-reported job satisfaction concerning working hours is my fifth outcome. In line with previous studies in the field (Rosen, 1986; Bonhomme and Jolivet, 2009), I take the original variable measured on a scale from 1 "Not satisfied at all" to 7 "Completely satisfied" and derive a dummy that distinguishes two levels of satisfaction, coded 1 if respondents' score is 5 or higher and 0 otherwise. Last, as an alternative measure of satisfaction with working-time arrangements, I look at the likelihood of feeling 'unconstrained' in the current job (Böheim and Taylor, 2004). Workers are considered to be unconstrained when reporting that they would not alter their working hours, upwards or downwards, if given the chance to do so keeping their current pay rate.

To fully exploit the panel nature of the data, I include individual fixed effects (FE, θ_i). Time-invariant individual characteristics are thereby netted out. Examples of such unobserved characteristics include the time-invariant components of individual productivity (e.g. ability) or of preferences regarding job (dis)amenities (Hotz et al., 2017). Notably, high-ability individuals might have both a higher earnings potential and be better poised to "purchase" job amenities in the labour market – being better at negotiation with their current employer or at job search. This can create a spurious positive correlation between earnings and job amenities (Heywood et al., 2007). Individual fixed effects take care of the fact that high-ability workers may sort in jobs that combine desirable monetary and non-monetary

features.

Adding individual fixed effects also allows to account for compositional changes over time among treated and controls, including changes due to attrition in unbalanced panel data, as long as these are due to time-invariant individual features (e.g. Francesconi and Van der Klaauw, 2007). Although the main effect of treatment status cannot be identified, since treatment status is time-invariant as well, one can still retrieve the DD estimator as it involves the interaction of treatment status ($D_i = 1$ if parent of a child under 6) and a time-varying dummy ($POST03_t$). The latter distinguishes the pre-reform and post-reform period, respectively before and after April 2003 for the 2003 reform. The FE-DD estimator α_2 thus expresses *within-individual* variation among the treated vis-à-vis the controls, after as compared to before the 2003 reform (Lechner et al., 2016). Since assignment to treatment and control group is here defined in terms of eligibility to the new right, all estimates are intention-to-treat ones.

The main empirical model for the assessment of the 2003 reform is completed by a set of control variables \mathbf{X}'_{it} , namely a quadratic for age, dummies for region of residence, and the log of the regional unemployment rate, as well as interview year fixed effects ϕ_t . Standard errors are clustered at the individual level to account for serial correlation within units over time.

In a second step of the analysis, I re-ran the model as specified in Equation 1 on the sample comprising both men and women to evaluate whether the 2003 reform affected gender gaps in the British labour market, installing compensating differentials for women and not (or more so than) for men. This amounts to let the reform-relevant parameters α_1 and α_2 vary by gender:

$$y_{it} = (1 + FEMALE_i) \times [(\alpha_1 POST03_t + \alpha_2 (POST03_t \times D_i))] + \mathbf{X}'_{it} \gamma + \theta_i + \phi_t + \epsilon_{it} \quad (2)$$

I thus retrieve a triple-difference estimator (DDD), namely the (within-individual) difference

for women vis-à-vis men in the difference between treated and controls, comparing the post- and pre-reform period. To gauge whether compensating differentials hold in particular for parents of small children, both DD and DDD analyses are replicated considering, on the one hand, only parents of children 0-2 and, on the other, only parents of children aged 3 to 5 at the time of the 2003 reform as alternative groups of treated units.

2.3. Parallel trends

The empirical strategy chosen here shields estimates from time-invariant sources of confounding via the inclusion of individual fixed effects. In DD settings, one then relies on the parallel trend assumption to exclude time-varying confounding. Prior to a given intervention, in my case the 2003 reform, outcomes in the treated and control group should evolve in parallel over time. Had there not been the intervention, one assumes outcomes for the two groups would have continued on their parallel paths. The DD strategy then retrieves an unbiased estimate of the causal impact of the intervention, expressing how much a given outcome deviates from the parallel path for the treated vis-à-vis the controls in the aftermath of the treatment (assignment) period.

I assess outcome trends prior to the intervention of interest and conditional on the covariates, time- and individual fixed effects comprised in the main statistical model. In practice, this entails running a series of models of the following form

$$y_{it} = \mathbf{X}'_{it}\gamma + \theta_i + \phi_t + \epsilon_{it} \quad (3)$$

for each of the six outcomes y_{it} and limiting the analysis to the pre-reform period (2000-2002). Similar to the main models, I include the vector \mathbf{X}'_{it} (a quadratic for age, dummies for region of residence, and the log of the regional unemployment rate), as well as individual and interview-year fixed effects (θ_i and ϕ_t , respectively).

Predicted values are then averaged across interview years and treatment groups. A graphical inspection of the obtained conditional outcome trends is portrayed in Figure 1, focusing

on the main contrast between mothers of children 0-5 (Treated) v. childless women and mothers of children older than 5 (Controls). Looking at the left panel and for most of the outcomes, outcome trends do not diverge substantially across groups even without any data pre-processing (“Unweighted”). Hourly wages are the only exception: the trend for women in the treated group is monotonically increasing, while control-group women experience somewhat of a decline in hourly wages during the pre-reform years. Since I expect a wage drop for the treated vis-à-vis the controls in the aftermath of the reform, I ran the risk of underestimating such reduction given that hourly wages for control-group women (but not the treated) were already declining in years prior to the reform.

Nevertheless, I also replicate my analyses addressing the potential violation of parallel trends. I augment both DD and DDD estimates via entropy balancing to increase the comparability of treatment groups with respect to time-varying confounders (e.g. Freier et al., 2015). Entropy balancing is a non-parametric algorithm (Hainmueller, 2012) that estimates weights such that, once re-weighted, the distribution of a set of variables in the control group resembles that of the treated, matching directly on sample moments (mean, variance, skewness). I balance each of the contrasts displayed in Table 1 on the means of a set of covariates measured in waves 11 and 12 (carried out in the pre-reform period 2000-2002). Overall, pre-treatment outcome trends in Figure 1 display a good balance, and I thus refrain from balancing samples on pre-treatment outcomes (see also Chabé-Ferret et al., 2017). Yet mothers of small children treated by the reform may still differ substantially from control-group women, including both mothers of older children and childless women, particularly with respect to labour supply patterns. It thus could be, for example, that one finds mothers of younger children to be more likely to reduce their working hours vis-à-vis the controls not because of the policy reform itself, but because of how childcare obligations and, thereby, labour supply simply differ between the two groups.

The set of covariates deployed for balancing is thus meant to capture factors related to women’s labour supply decisions. Covariates include a counter for the number of children in the household, a dummy for employment status (1 if employed), a dummy for marital status

(1 if married), a counter for the number of weeks spent in inactivity in the year prior to the interview, and household income (excluding the individual's own labour income).

Figure 2 depicts standardised differences between means, across treatment groups, of all variables used for entropy balancing. For illustrative purposes, I focus on the main treated group, that of mothers of children younger than 6, and their control group comprising mothers of older children and childless women. Positive differences indicate an imbalance “in favour” of the treated group, negative differences one “in favour” of controls. Prior to balancing, treated and control groups display evident imbalances in terms of number of children and proportion married, imbalances in the direction that could be expected given that the treated consist of mothers of young(er) children. Specifically, treated women have a higher number of children, are more often married, have spent longer periods inactive, and are less often employed than control-group women. After balancing, standardised differences for these and all the other variables reduce to 0³.

Parallel trends for the main sample, re-weighted by means of entropy balancing, are displayed in the right panel of Figure 1 (“Weighted”). Applying the balancing algorithm on a limited set of covariates, outcome trends for hourly wages now appear to follow parallel trends across treatment groups in the pre-reform period. Balancing is similarly applied to all other sample contrasts and seemingly ameliorates the observed component of parallel trends across all such contrasts (see Figures 1A, 2A, 3A in the Appendix).

3. Findings

The first panel of Table 2 displays how the introduction of the right to request flexible schedules in 2003 changed the outcomes of mothers of children 0-5 relative to a control group comprising mothers of older children and childless women. Predictions following the theory of compensating differentials would lead to expect, first, career penalties for the treated. The hourly wages of mothers in the treated group indeed fell by around 7% ($p < .001$) relative to the controls in the post-reform period. Differently, I cannot ascertain a relative penalty, despite the small coefficient being in the expected direction, in terms of access to managerial posts ($\alpha_2 = - .02$, $p = .405$).

Compensating monetary losses, I find some evidence for the hypothesised increased job amenity value of working-time flexibility. Mothers in the treated group experience a reduction in working hours, of around .6 hours on average ($p = .068$), although this does not seem to have been accomplished through increased access to part-time ($\alpha_2 = .012$, $p = .546$). Treatment effects for job satisfaction and unconstrained status are all positive, yet small in size (and $p = .403$, $p = .204$, respectively). Overall, eligible mothers in the treated group seem to have experienced a trade-off between wage attainment and working hours, but little evidence suggest this was accompanied by higher satisfaction with the newly reduced working schedule.

The second and third panel of Table 2 investigate whether the 2003 reform had heterogeneous effects depending on the age of the youngest child in the household at the time of the reform. Mothers of children aged 0-2 should confront the steepest search costs and hurdles in combining paid and unpaid work. Similar to the main estimates, mothers in this narrower treatment group experienced a pay cut of around 6% in the post-reform period ($p = .006$). Their chances of attaining managerial posts decline of around 5 percentage points vis-à-vis the controls (albeit $p = .114$). Results in columns 3, 5, and 6 indicate that the 2003 reform lead to strong changes in working hours (-3 weekly hours, $p < .001$) accompanied by accrued chances of feeling satisfied with working-time arrangements and feeling unconstrained with respect to working hours (8 and 10 percentage points, respectively). Chances of working part-time increase by around 10 percentage points ($p = .003$) for the group of mothers of children 0-2. Differently, moving to the bottom of Table 2, I can detect a relative wage penalty still standing at 7.5% ($p < .001$) for mothers of children 3-5. Yet, women in this treated group *increase* their working hours and their chances of working part-time decrease in the aftermath of the reform. This may signal the use of different sources of working-time flexibility other than reduced hours and part-time, still bearing a wage penalty and, in addition, not leading to heightened satisfaction (as per columns 5 and 6). On balance, results highlight how introducing the right to request in 2003 may have benefited employed mothers, especially mothers of toddlers at the time of the reform, in terms of working time

and the amenity value they associate with it. Career costs, in the form of wage drops, also emerge, albeit coherently with a compensating differential story only for mothers of children 0-2.

If compensating differentials arise for women, specially to ease (early) childcare duties, it might be that the 2003 reform had the unintended effect of deepening gender career gaps in the British labour market. Triple-difference estimates help assessing whether the effects of the 2003 reform were felt more among women than among men, thereby widening gender gaps. Finding negative DDD estimates for hourly wages might suggest, for example, that mothers vis-à-vis their controls have experienced harsher wage losses than fathers vis-à-vis their controls per effect of the 2003 reform. Column 1 of Table 3 points in this direction. Per effect of the reform, the pay cut for eligible mothers of children aged 0-5 vis-à-vis control-group women is around 5% larger than the wage change detected for similarly eligible fathers with respect to control-group men ($p = .022$). DDD wage estimates are similarly around 5.7% ($p = .061$) and 4% ($p = .142$) when considering only parents of children 0-2 or of children 3-5, respectively. Differently, I find small but noisier estimates associated with access to managerial posts suggesting nonetheless a stronger penalty for treated women as compared to female controls rather than treated men compared to male controls in all three sample contrasts. DDD estimates for career attainment mirror by and large the findings for women only, suggesting that the wage penalty for eligible mothers may contribute to gender career gaps more broadly.

As for working time and its amenity value, estimates in columns 3 to 6 are somewhat coherent with those of Table 2. The 2003 reform widened the gap in working hours between parents of small children and their controls, especially for women and especially when considering parents of toddlers. In the latter treated group, women also appear more likely than men to switch to part-time (around 10 percentage points, $p = .006$), as compared to their respective control groups and per effect of the 2003 reform. The likelihood of feeling unconstrained and satisfied with working time appears to have increased differentially for women and men, by and large in line with what found when analysing women only. Notably, estimates for

working hours and part-time in the DDD analysis for the subgroup of parents of children 3-5 suggest once again an increase in the working hours of eligible mothers. This seems not to be accompanied by substantial changes in job satisfaction or in the chances of feeling unconstrained, reflecting the findings of the DD analysis. This suggests that the wage cut for mothers of children aged 3 to 5 does not square well with a compensating differential mechanism.

4. Robustness

In Figures 3 and 4, I compare estimates for my main specifications in Tables 2 and 3 (“Un-weighted”) with those obtained after re-weighting via entropy balancing (“Weighting”). Overall, estimates prove robust to the weighting procedure here deployed to improve on parallel outcome trends across treatment groups prior to the reform.

For the DD analysis in particular (Figure 3), weighted estimates reinforce the main finding of a trade-off between hourly wages losses on the one hand, and reductions in working hours, access to part-time positions, and an accrued feeling of being ‘unconstrained’ with respect to working hours, on the other. This is primarily experienced by mothers of children under 2 and cannot be ascertained for mothers of older children still eligible for the new right. Gender gaps, highlighted by the DDD analysis in Figure 4, evolve in accordance to the patterns assessed for women only.

Discussion and conclusion

Introducing a right to request changes to working-time arrangements for parents of young children may have perverse effects on the career attainment of mothers. Fitting the theory of compensating differentials, eligible mothers after the reform decrease their working hours and become more satisfied with their working-time arrangements vis-à-vis the pool of ineligible women. Yet, this comes at an economic price in terms of wage losses. This is particularly true for mothers of children aged 0 to 2 at the time of the reform, whereas I cannot ascertain whether eligible mothers of older children are also compensated for wage losses. I also find some support for the idea that the reform, while widening the gap between women who were

eligible for the new right and those who were not, also widened gender (pay) gaps more at large.

Mothers of young children seem to have exploited the reform for reductions in working hours, as well as for moves to part-time. Previous research for Germany also shows shortened work-schedules, but not necessarily different working-time arrangements altogether, are a likely choice for new mothers (Felfe, 2012a). I have not assessed here whether and how the reform affected moves to a broader spectrum of flexible schedules also encompassed in the new right to request. This would include flexitime (flexible daily start and finishing times), working compressed or annualised hours, shift work, job sharing, or working from home. The career costs attached to such options are under-researched though (but see Heywood et al., 2007), whereas those attached to part-time and reduced hours are well established (e.g. Connolly and Gregory, 2008; Manning and Petrongolo, 2008; Matteazzi et al., 2014). Investigating compensating differentials, it is thus reasonable to start from the latter, also considering they have been the main options requested by women in the framework of the right-to-request policy (Hegewisch, 2005; Hooker et al., 2011).

Compensating differentials hold particularly for mothers of toddlers. These mothers arguably face the highest obstacles in combining childcare and paid work (e.g. Felfe, 2012b). Evidence on the increased job satisfaction of this group after the reform suggests that policies affecting working time help these mothers obtain effective flexible solutions. On the flip side though, mothers incur in career costs. Future research could shed a light on whether other policies, especially those centred around the availability and affordability of childcare for toddlers, may help mitigate this trade-off.

Working-time policies more broadly fall in the category of women-friendly policies previous studies have scrutinised for their boomerang effects on women's careers. Evidence in this paper suggests that detrimental effects may be particularly felt by mothers of young children via the instalment of compensating differentials, even in a context such as that of Britain where the main flexible option, part-time employment, is of particularly low quality. In the UK, expansions of the 'right to request' were carried out in 2006 and 2014, first encompassing

parents of older children and then all employees regardless of parental status. Building on evidence presented in this study, further research could assess whether policies targeting all women (and men), rather than mothers (parents) in particular, also have unintended effects on between- and within-gender inequality in labour markets.

Notes

¹Parents of disabled children under 18 were also entitled under the provisions of 2002-3, yet I cannot identify this group in my analyses due to data constraints.

²Throughout, wages are deflated at 2006 prices using the annual CHAW-RPI index provided by the Office for National Statistics.

³The same exact balancing on the means of past outcomes and selected covariates is achieved for all sample contrasts used for the evaluation of the 2003 reform (figures available upon request)

Tables & graphs

Table 1: Number of unique individuals (person-year records in parentheses) for all the contrasts between treated and controls.

	<i>Treated</i>	<i>Controls</i>
Mothers _{child₀₋₅} v. Childless women + Mothers _{child_{≥6}}	476 (1,759)	1,251 (5,270)
Fathers _{child₀₋₅} v. Childless men + Fathers _{child_{≥6}}	446 (1,822)	1,433 (5,785)
Mothers _{child₀₋₂} v. Childless women + Mothers _{child_{≥6}}	177 (628)	1,251 (5,007)
Fathers _{child₀₋₂} v. Childless men + Fathers _{child_{≥6}}	189 (823)	1,433 (5,785)
Mothers _{child₃₋₅} v. Childless women + Mothers _{child_{≥6}}	299 (1,004)	1,251 (5,007)
Fathers _{child₃₋₅} v. Childless men + Fathers _{child_{≥6}}	257 (999)	1,433 (5,785)

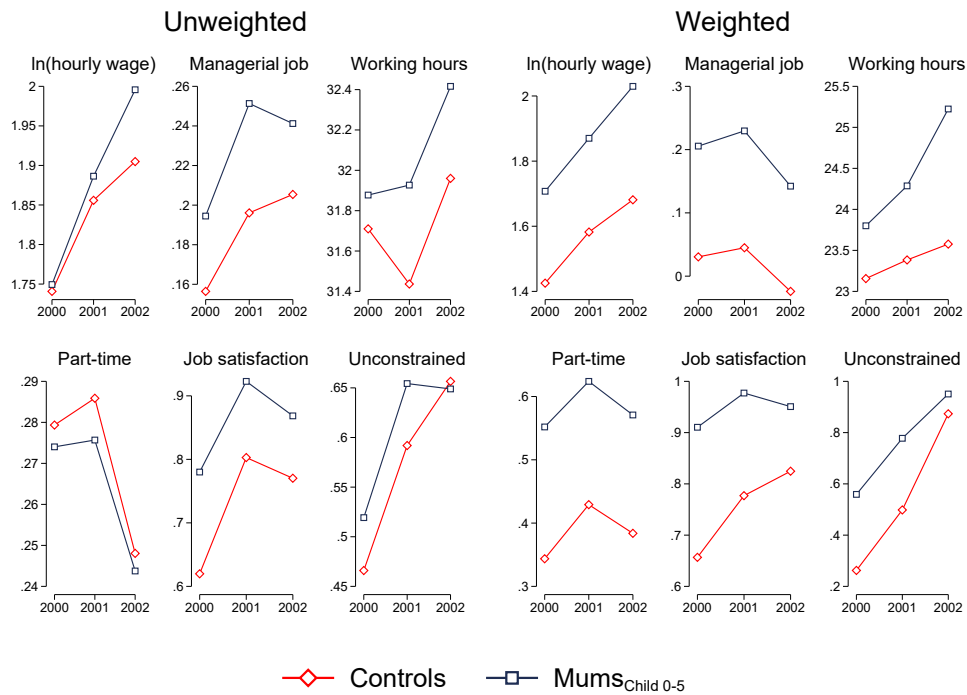


Figure 1: Predicted outcomes averaged over treatment groups and interview years, prior to the 2003 right-to-request reform (Equation 3). The left panel refers to estimates before weighting (“Unweighted”), the right panel after weighting by means of entropy balancing (“Weighted”). The group of “Controls” comprises childless women and mothers of children aged 6 or older (BHPS, 2000-2002).

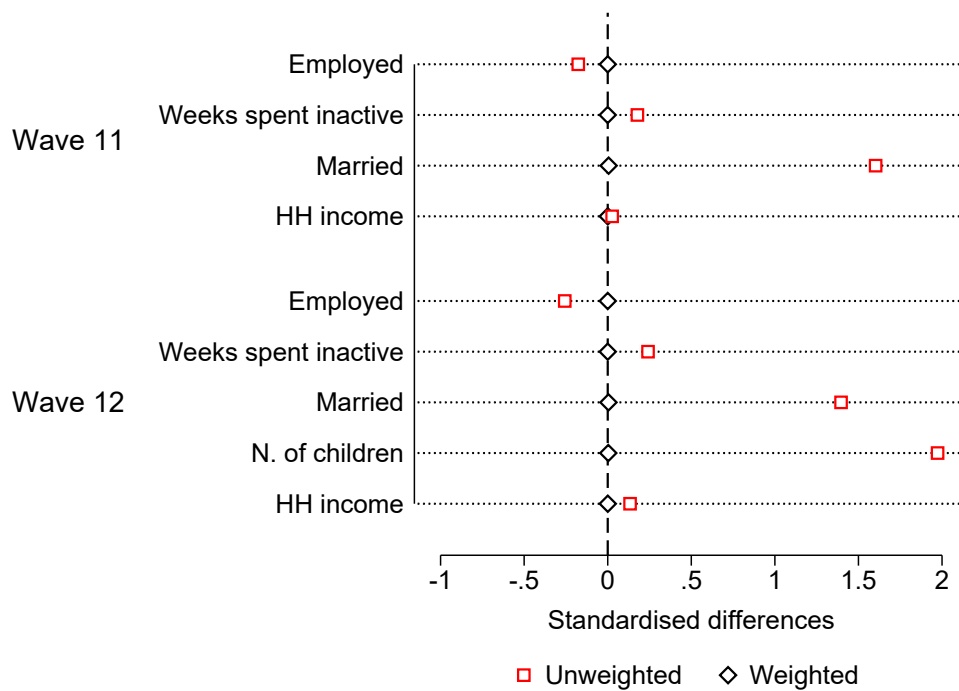


Figure 2: Standardised differences between the means of selected variables for treated and control groups. Selected variables are measured in pre-reform waves 11 and 12. Balanced differences are computed after weighting by means of entropy balancing. The treated group comprises mothers of children aged 0-5, controls comprise childless women and mothers of older children (BHPS waves 11 and 12).

Table 2: Fixed-effects (FE) DD estimates for the effects of the 2003 reform on mothers.

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(wage)	Managerial job	Working hours	Part-time	Job Satisfaction	Unconstrained
<i>POST03</i> × <i>Mum</i> _{child₀₋₅}	-0.072*** (0.016)	-0.017 (0.021)	-0.679* (0.371)	0.012 (0.020)	0.020 (0.024)	0.036 (0.028)
Number of individuals	1,727	1,727	1,727	1,727	1,727	1,727
Number of person-years	6,639	6,639	6,639	6,639	6,639	6,639
<i>POST03</i> × <i>Mum</i> _{child₀₋₂}	-0.066*** (0.024)	-0.049 (0.031)	-3.080*** (0.552)	0.103*** (0.035)	0.076* (0.040)	0.105** (0.044)
Number of individuals	1,428	1,428	1,428	1,428	1,428	1,428
Number of person-years	5,635	5,635	5,635	5,635	5,635	5,635
<i>POST03</i> × <i>Mum</i> _{child₃₋₅}	-0.075*** (0.020)	0.003 (0.025)	0.917** (0.418)	-0.049** (0.021)	-0.017 (0.027)	-0.009 (0.033)
Number of individuals	1,550	1,550	1,550	1,550	1,550	1,550
Number of person-years	6,011	6,011	6,011	6,011	6,011	6,011

Notes: All models include a post-reform dummy, a quadratic for age, dummies for region of residence, the log of regional unemployment rate, and individual and interview-year fixed effects. Estimates are weighted by means of entropy balancing as detailed in the text. Robust standard errors in parentheses.

* $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

Source: BHPS waves 10-15

Table 3: Fixed-effects (FE) DDD estimates for the effects of the 2003 reform on gender gaps.

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(wage)	Managerial job	Working hours	Part-time	Job Satisfaction	Unconstrained
<i>POST03</i> × <i>Parent</i> _{child₀₋₅ × <i>Female</i>}	-0.050** (0.022)	-0.026 (0.027)	-0.677 (0.470)	0.015 (0.021)	0.023 (0.034)	0.029 (0.038)
Number of individuals	3,606	3,606	3,606	3,606	3,606	3,606
Number of person-years	14,246	14,246	14,246	14,246	14,246	14,246
<i>POST03</i> × <i>Parent</i> _{child₀₋₂ × <i>Female</i>}	-0.057* (0.030)	-0.052 (0.039)	-3.138*** (0.670)	0.103*** (0.037)	0.086* (0.050)	0.047 (0.055)
Number of individuals	3,050	3,050	3,050	3,050	3,050	3,050
Number of person-years	12,243	12,243	12,243	12,243	12,243	12,243
<i>POST03</i> × <i>Parent</i> _{child₃₋₅ × <i>Female</i>}	-0.042 (0.028)	-0.014 (0.035)	0.923* (0.559)	-0.042* (0.022)	-0.019 (0.042)	0.033 (0.047)
Number of individuals	3,240	3,240	3,240	3,240	3,240	3,240
Number of person-years	12,795	12,795	12,795	12,795	12,795	12,795

Notes: All models include a post-reform dummy, a quadratic for age, dummies for region of residence, the log of regional unemployment rate, and individual and interview-year fixed effects. Estimates are weighted by means of entropy balancing as detailed in the text. Robust standard errors in parentheses.

* $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

Source: BHPS waves 10-15

DD analysis

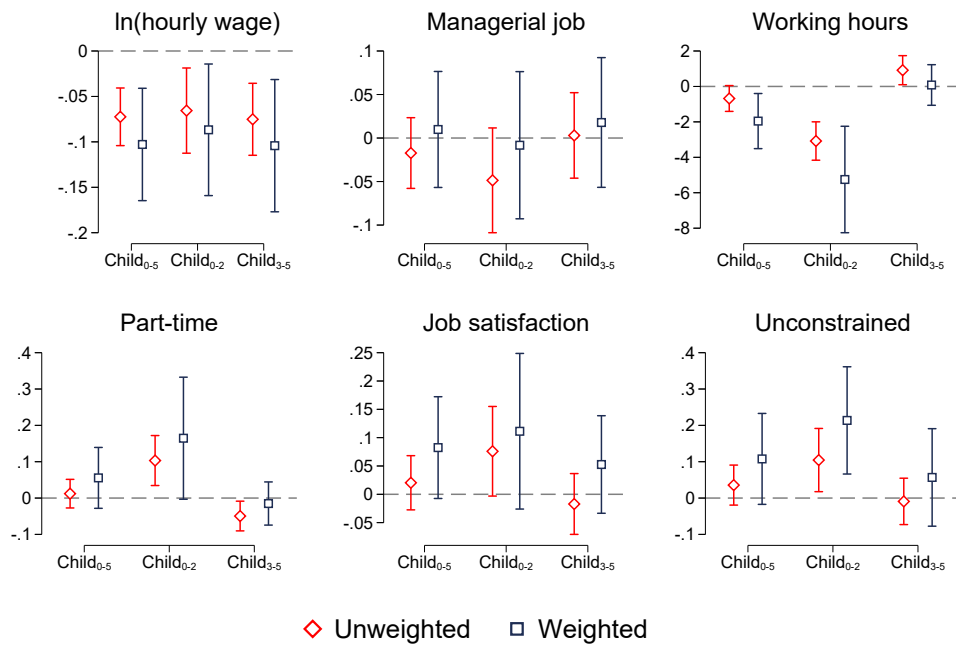


Figure 3: Fixed-effects (FE) DD estimates for the effects of the 2003 reform on mothers, without (“Unweighted”, see Table 2) and with entropy balancing (“Weighted”).

DDD analysis



Figure 4: Fixed-effects (FE) DDD estimates for the effects of the 2003 reform on gender gaps, without (“Unweighted”, see Table 3) and with entropy balancing (“Weighted”).

Appendix



Figure 1A: Predicted outcomes averaged over treatment groups and interview years, prior to the 2003 right-to-request reform (Equation 3). The left panel refers to estimates before weighting (“Unweighted”), the right panel after weighting by means of entropy balancing (“Weighted”). The group of “Controls” comprises childless women and mothers of children aged 6 or older (BHPS, 2000-2002).



Figure 2A: Predicted outcomes averaged over treatment groups and interview years, prior to the 2003 right-to-request reform (Equation 3). The left panel refers to estimates before weighting (“Unweighted”), the right panel after weighting by means of entropy balancing (“Weighted”). The group of “Controls” comprises childless women and mothers of children aged 6 or older (BHPS, 2000-2002).



Figure 3A: Predicted outcomes averaged over treatment groups and interview years, prior to the 2003 right-to-request reform (Equation 3). The left panel refers to estimates before weighting (“Unweighted”), the right panel after weighting by means of entropy balancing (“Weighted”). The group of “Controls” comprises childless men and fathers of children aged 6 or older (BHPS, 2000-2002).

Chapter 4

Gender, Parenthood, and Hiring Decisions in Sex-Typical Jobs: Insights from Two Survey Experiments

Gender, Parenthood, and Hiring Decisions in Sex-Typical Jobs: Insights from Two Survey Experiments*

Abstract

We ran two survey experiments with Dutch employers to investigate hiring discrimination in sex-typical jobs. We ask if women are especially discriminated against when they have children, whether discrimination applies similarly in different occupations, and whether statistical discrimination or status-characteristic theories best account for discriminatory practices (if any). Employers rate fictitious candidates for either a female-typical job (primary-school teacher) or a male-typical job (software engineer). Employers are found to display a slight preference for female candidates when filling a teacher post, although such bias is less strong for female applicants with children. No such ranking is found for a software engineer vacancy, nor do we find different salary offers across candidates and across vacancies. Employers do not appear to favour men over women for positions likely to be on the career track, as predicted by statistical discrimination theories, nor do they expect women to be less capable than men, as posited by status-characteristic theory. Female candidates with children, however, are expected to be less committed to their job and work fewer hours, especially in the teacher experiment. Such expectations seem to have small consequences for the hiring decisions and salary offers Dutch employers make in our study.

*This paper is co-authored with Ruud Luijkx. In this paper we make use of data from the LISS (Longitudinal Internet Studies for the Social sciences) panel administered by CentERdata (Tilburg University, The Netherlands). The LISS panel data were collected by CentERdata (Tilburg University, The Netherlands) through its MESS project funded by the Netherlands Organization for Scientific Research. Funding for this research project was provided by Tilburg University, School of Social and Behavioral Sciences. The authors wish to thank Marije Oudejans for managing the data collection process. We also wish to thank participants at the 2017 ESRA Conference in Lisbon, as well as in seminars at Tilburg University and at the University of Trento for their comments and suggestions. A slightly different version of this paper is currently under review at an international peer-reviewed journal.

Employers may shape gender disparities in the workplace when they treat differently otherwise equally-productive men and women. Favouring male candidates over comparable female ones at the hiring stage is one instance of such discriminatory practices. Hiring decisions are increasingly understood as key drivers of labour market inequalities (e.g. Bills et al., 2017; Protsch and Solga, 2015; Di Stasio and van de Werfhorst, 2016) and gender discrimination in hiring is well-documented (Azmat and Petrongolo, 2014; Neumark, 2016). Questions still loom, however, on whether women are especially discriminated against when they have children, whether discrimination applies similarly in different occupations, and which mechanisms drive discrimination. We address here this threefold gap by means of two survey experiments on the hiring decisions of Dutch employers.

First, previous research has yielded mixed support for gender-by-parental-status discrimination resulting in a motherhood penalty at the hiring stage. A seminal study combining a lab and a field experiment in the US found mothers to be the least-preferred candidate for hire in marketing and business jobs (Correll et al., 2007). Field experiments in the financial sector in France (Petit, 2007) and across a wide array of jobs in Sweden (Bygren et al., 2017) have not replicated this finding. Differently, a comparably large field experiment in Spain found lower callback probabilities for women especially if with children (González et al., 2019) and a vignette study in Switzerland found evidence of a motherhood penalty for women applying for a HR assistant position (Oesch et al., 2017). We add to this debate by examining how employers evaluate job candidates at the intersection of gender and parental status in the Netherlands. The Dutch context features a “residual” gender wage gap of around 8% that studies attribute in part to discrimination (Fransen et al., 2012), as well as persistent gender segregation by field of study and, consequently, occupation (van de Werfhorst, 2017; OECD, 2017b: p. 144). At the same time, the evidence for a motherhood penalty is mixed in the Dutch labour market (cf. Davies and Pierre, 2005; De Hoon et al., 2017) making the Netherlands a compelling *litmus test* for the analysis of gender-by-parental-status discrimination. Our second contribution consists in running one experiment for a male-typical job (software engineer) and one for a female-typical job (primary-school teacher). Research on the moth-

erhood penalty in hiring has focused on mixed occupations (for an exception, Bygren et al., 2017). Yet, experimental studies on gender discrimination at large have found differential treatment *in favour* of women in female-typical jobs and, albeit to a lesser extent, against women in male-typical jobs (Riach and Rich, 2002; Rich, 2014; Neumark, 2016). We complement these findings by assessing if and how parental status modifies gender discrimination in sex-typical occupations.

Finally, we develop and test the predictions of two theories, statistical discrimination theory and status-characteristic theory (e.g. Correll and Benard, 2006). While the former framework would predict particularly pronounced gender differentials in male-typical jobs and holding irrespective of parental status, the latter would suggest that mothers may be better off in female-typical jobs. Further, employers' propensity to discriminate statistically may vary depending on the length of the prospective employment relationship, with permanent contracts being conducive to higher risk-aversion on the part of employers and thereby lower hiring chances for the groups employers deem less reliable or productive. Alternatively, we measure the expectations employers form regarding different candidates' ability and commitment and assess the role of these constructs in accounting for hiring patterns, as predicted by status-characteristic theory.

In short, we ask if there is a motherhood penalty in hiring chances, whether it varies along occupational and type-of-contract lines, and if such variation fits with the predictions of either economic or socio-psychological models of discrimination. Rather than relying on undergraduate students as in most experimental simulations of the hiring process (Koch et al., 2015), participants in our study are real-world employers selected among the respondents of LISS, a web panel carried out since 2007 on a true probability sample of the Dutch population (Scherpenzeel, 2011). While field experiments involve real-world employers too, what is typically measured is only the outcome of employers' decision over a candidate, that is, whether or not a given job applicant gets a callback. We are able here not only to survey employers' decision over a candidate, but also to inquire about employers' motives.

1. Background

1.1. Statistical discrimination: risk aversion in hiring decisions

The design of our study is informed by economic and socio-psychological models of discrimination, namely statistical discrimination theory and status-characteristic theory, both commonly featured in the literature on motherhood penalties in labour markets (e.g. Correll et al., 2007; Gangl and Ziefle, 2009; Bygren et al., 2017; Oesch et al., 2017).

Under statistical discrimination, employers are rational actors aiming at maximizing expected profits in labour markets characterized by imperfect information (for a review, Fang and Moro, 2011). Specifically, employers may find it difficult or too expensive to access precise information on the individual productivity of job applicants. Group markers such as a candidate's sex are instead easily accessible – on CVs, at job interviews, etc. – and employers may thereby determine individual productivity by combining the expected productivity of a given female (male) candidate with the group-level productivity they estimate for women (men). Even if employers believe women and men to be equally productive on average, the productivity signals of a female applicant might be deemed more noisy (Aigner and Cain, 1977; Charles and Guryan, 2011). The underlying assumption employers make is that, despite equal educational credentials, accumulated work experience etc., female employees might be more likely than men to take career breaks, reduce their working hours, or leave their job altogether for family-related reasons.

To compensate for this “risk”, employers become more reluctant to hire women and pay them less than men all else equal. Economic discrimination arises then because of overshooting, as female employees deciding against motherhood or whose productivity is not affected by the presence of children will be discriminated against. In this framework, employers are thus forward-looking and their concern lies with women *potentially* becoming parents and altering their labour supply in ways harmful to productivity. Women of childbearing age, irrespective of current parental status, could be discriminated against with respect to men (Gupta and Smith, 2002; Petit, 2007; Yip and Wong, 2014; Baert, 2014; Biewen and Seifert, 2016). We name this our *family-risk hypothesis*.

What can be done to modify employers' propensity to discriminate statistically? Two variables can be manipulated, namely the information available for each job applicant and the risk employers associate with hiring him or her. Supplying precise and cheap information on individual productivity could help employers correct the signal-to-noise ratio in their productivity estimates (e.g. Altonji and Pierret, 2001; Pinkston, 2006). Although information on job candidates has indeed been manipulated to detect statistical discrimination in experimental settings, doing so is problematic for two reasons (Neumark, 2016). First, researchers may not know which information employers need to forgo using group markers as proxies for individual productivity. Besides, even if such information is known to the experimenters, withholding it in a "low-information" experimental condition may prove to be unrealistic with respect to real-world job applications¹.

Instead of information then, we turn to the manipulation of risk as a function of *a*) the length of the work contract being offered and *b*) the type of occupation. For the statistically-discriminating employer, investing in female job applicants may be more risky in permanent employment relationships as compared to temporary ones. While a temporary contract may be used to screen employees and get rid of bad matches at a low or no cost when the contract itself expires, permanent contracts are typically harder to terminate. This especially holds in contexts such as the Dutch one, where recent labour market reforms have relaxed the level of employment protection attached to temporary contracts but left largely untouched that of permanent contracts (e.g. Mooi-Reci and Dekker, 2015) – and the latter remains above the European average in the Netherlands (OECD, 2014). Other than with a costly dismissal, permanent contracts are also more likely to be associated with employer-funded training than are temporary contracts (Fouarge et al., 2012; Akgündüz and van Huizen, 2015). Therefore, given that hiring an employee on a permanent contract entails higher economic costs, risk-averse Dutch employers will be more prone to discriminate statistically against female applicants when filling a permanent rather than a temporary job vacancy (*contract-risk hypothesis*).

In the same spirit, we might expect both our *family-risk* hypothesis and our *contract-risk*

hypothesis to find more support in male-typical rather than in female-typical lines of work (*occupation-risk hypothesis*). Occupations that are held predominantly by men often have steeper career ladders than female-dominated ones. Foreseeing further investments in terms of training and promotions, employers in male-typical jobs may thus be reluctant to hire female applicants, especially if permanently. The evidence for gender discrimination in hiring in male-dominated jobs is mixed though, with some field experiments finding lower callback rates for women vis-à-vis men and others finding such rates to be substantially and statistically indistinguishable (e.g. Riach and Rich, 2006; Carlsson, 2011; Bygren et al., 2017; for reviews, Riach and Rich, 2002; Rich, 2014; Neumark, 2016). To the best of our knowledge, only one previous study incorporated the type of work contract as an additional experimental condition, finding hiring discrimination against young childless women for permanent vacancies in the financial sector in France (Petit, 2007). More to this point, Baert and colleagues (2016) find substantially lower callback rates for young women for jobs that imply promotion prospects in a large field experiment in Belgium. Keeping this in mind, we argue that the long-term investment of hiring an applicant for a ‘career’ rather than simply for a ‘job’ makes filling permanent posts in male-typical occupations more prone to statistical discrimination against women of childbearing age.

1.2. Status-characteristic theory: the role of status beliefs penalizing mothers

According to status-characteristic theory, no matter information asymmetries and risk aversion, status beliefs may bias employers’ evaluations against women and, particularly, mothers (Ridgeway and Correll, 2004; Correll and Benard, 2006; Correll et al., 2007). Status beliefs are conceived as a particular class of stereotypes (e.g. Fiske et al., 2002) by virtue of which individuals categorize members of social groups on the basis of perceived competence. Any nominal characteristic that groups together individuals in a social setting – sex, ethnicity, sexual orientation, and so forth – may become a status characteristic if actors share beliefs regarding that group’s competence, further conceptualized as the sum of *ability* and *commitment* (Berger et al., 1977; Ridgeway and Correll, 2006; Mark et al., 2009; Ridgeway et al., 2009). Seeking coordination with one another, individuals may activate performance

expectations regarding how capable and committed others will be with respect to the task at hand, depending on the salient social memberships. Performance expectations may then drive the distribution of rewards such that, in the hiring setting, low-status actors end up penalized in terms of hiring chances or salary offers (Correll et al., 2007; Pedulla, 2016, 2018; for a review, Ridgeway, 2011).

Motherhood is such a status characteristic insofar as it amplifies status beliefs morphed along gender lines (Ridgeway and Correll, 2004; Correll et al., 2007), following a pattern of “amplified congruence” between status characteristics (Pedulla, 2018). Specifically, women, especially if mothers, are perceived to be less capable than men in workplace settings (Cuddy et al., 2004; Correll et al., 2007; Thébaud, 2015). Superior ability may be granted to mothers only for tasks that involve nurturance and care (Ridgeway and Correll, 2004), in line with stereotypes broadly associating women with communion, i.e. being selfless and concerned with others, rather than with agency, i.e. being assertive and motivated to master a task (for a review, Ellemers, 2018). As for commitment, women and mothers in particular are not expected to prioritize work over family obligations nor to make sacrifices to build a career as much as men would (Correll et al., 2007; Rivera and Tilcsik, 2016; Ridgeway and Correll, 2004). We can therefore expect mothers to be the least-preferred candidate for hire and to receive the lowest salary offers (*motherhood penalty* hypothesis). Employers will attribute the lowest levels of expected competence, i.e. of ability and commitment, to female job applicants who are also mothers (*status belief* hypothesis).

Yet, in female-typical jobs where nurturing and caring skills are crucial (teaching, nursing, social work etc.), proponents of status-characteristic theory suggest women and mothers might actually be highly regarded in terms of competence (Ridgeway and Correll, 2004). This proposition may help explain the large body of evidence on gender discrimination in favour of women in female-typical lines of work (e.g. Booth and Leigh, 2010; Carlsson, 2011; Rich, 2014; Carlsson and Eriksson, 2017). In reviewing such findings, Neumark notes that they pose “a bit of a puzzle for labor economists, since our models of discrimination do not naturally predict this pattern. That said, perspectives on discrimination from other fields,

such as those emphasizing norms regarding who does which job, might fit these facts better” (Neumark, 2016: p. 77). Our hypothesis is that women, and mothers in particular, may top employers’ ratings of competence in female-dominated jobs and be therefore more likely to be hired and receive higher salary offers than men (see also Koch et al., 2015). We name this our *status reversal* hypothesis.

2. Empirical approach

2.1. Data, sample, and setup

Participants in our study are sample members of the Longitudinal Internet Study for the Social Sciences (LISS), a web household panel carried out since October 2007 on a true probability sample of the Dutch population (e.g. Scherpenzeel, 2011). The core modules of the LISS questionnaire allow to track the lifecourse changes and living conditions of panel members on a yearly basis, much like in traditional household panel surveys. *Ad hoc* modules can be submitted for consideration to CentERdata, a survey research institute affiliated with Tilburg University, and appended to one of the monthly rounds of interviews. From December 2016 to August 2017, we were thus able to ran our survey experiments thanks to the LISS infrastructure.

To select our “employers”, we administered four filter questions targeting respondents aged 25 to 65 and in paid work at the time of the most recent interview of LISS Work & Schooling module (Wave 8, April-May 2015). Of the 2,985 individuals on target, 2,252 (75.4%) were re-interviewed in December 2016 and 2,207 (73.9%) answered to all of our filter questions that month.

Four yes/no filter questions covered respondents’ HR responsibilities in their current job. Specifically, we asked whether respondents 1) had the power to hire/fire employees, 2) took part in any phase of the recruitment process (screening of CVs, job interviews, etc.), 3) could set or influence the rate of pay received by other employees, and 4) could have influence on or decide over the promotion of other employees. 749 respondents, who answered “Yes” to at least one of these four questions, form the pool of potential participants in our study. We

refer to them as employers in the study jargon.

Employers were later involved in a simulation of the hiring process (e.g. Correll et al., 2007; Di Stasio and Gërkhani, 2015; Di Stasio and van de Werfhorst, 2016). Employers were solely instructed that they would be taking part in such a simulation, with no explanation regarding the goals of the study. We ran two separate survey experiments, identical in design and questionnaire, but different in terms of the type of job vacancy employers were asked to rate candidates for (software engineer or primary-school teacher). After being randomly allocated to one of the two job vacancies, employers were presented with a CV for a first job applicant and asked to fill in a survey to evaluate him or her for the position. Employers were then presented with a second CV and repeated the questionnaire for a second candidate for the same position.

In March 2017, 150 employers were contacted for the pre-test phase and 132 (88%) completed the task, 66 for the software engineer vacancy and 66 for the teacher vacancy. In June and August 2017, a total of 480 respondents were targeted for the test phase, so that each survey experiment would have a sample size of 240 employers². Restricting our analysis to complete cases, we work with 239 participants – and 478 observations, since each employer evaluates a pair of CVs – for the teacher experiment and 237 (474 observations) for the software engineer experiment. As reported in Table 1, employers in both experiments are more likely to be men, highly-educated, and to have a partner (either married or cohabiting). Looking at their real-world HR responsibilities, the overwhelming majority of employers is involved in the screening of job applicants, matching well with our focus on employers' ratings of job candidates based on their CVs.

2.2. Choice of jobs and job ads

We picked *software engineer* and *primary-school teacher* for our male-typical and female-typical job, respectively. Both are examples of gender-segregated occupations in top-income countries. In the Netherlands, women made up to 86% of primary-school teachers as recently as 2015 (OECD, 2017a: p. 400), while only around 12% of Dutch women figured among computing professionals in 2007 (broadly defined, ISCO-88 code 213; see Bettio and

Verashchagina, 2009)³.

Beyond being on opposite sides of the spectrum in terms of gender segregation, our two jobs require similar educational qualifications (namely HBOs, the degrees obtained in the vocational track of the Dutch university system) yet widely differ in task content (e.g. Gathmann and Schönberg, 2010). Teaching primarily entails interactive tasks, with nurturance and care of small children being key at the primary-school level. Status beliefs on women's and mothers' higher competence might thus be activated in these realms, as predicted by status-characteristic theory. Further, even if in the public sector, primary-school teachers are hired directly by schools in the Netherlands, a feature that makes the hiring process comparable to that of private-sector jobs⁴. Software engineering jobs, on the other hand, not only primarily involve analytical tasks, but are also training-intensive and typically installed on well-structured career ladders, making them ideal to test our statistical discrimination argument.

As for previous studies on gender discrimination in these lines of work, only Bygren and colleagues (2017) have investigated the intersection of sex of the applicant and parental status in similar occupations. As part of a larger correspondence study in Sweden, they have found similar odds of a callback by sex and parental status for “computer specialists” and for “elementary school teachers”. More generally, evidence is often nil (Petersen et al., 2000; Carlsson, 2011; Di Stasio and van de Werfhorst, 2016; Fernandez and Campero, 2017; but see Riach and Rich, 2006) on gender discrimination in hiring in tech jobs. Our choice of jobs therefore provides a stringent test for our theoretical predictions.

Job ads in our study consisted of paragraphs of around seven lines of text (see Supplementary material). The wording and information provided was derived from real ads accessible through online databases⁵ and from fake ads used in previous studies (e.g. Di Stasio and Gërkhani, 2015). Ads included a range of possible salary offers for applicants with different levels of experience, as well as a range of possible weekly working hours (32 to 40). Employers were referred to such ranges when asked about their salary offer to a given candidate and about the weekly working hours they would have expected the candidate to put in if hired.

Ads also comprised our first manipulation, posting either a *permanent* or a *temporary* job vacancy to manipulate the risk associated with the prospective employment relationship.

2.3. CVs, job candidates, and manipulations

The CVs we presented to employers are stylized syntheses of real CVs accessed through free online databases (e.g. `indeed.nl`). CVs included the following information regarding the candidate: name, date of birth, marital status, previous work experience, maximum level and field of education attained, language skills, training courses for teachers/programming languages for engineers, and hobbies. Following previous studies (Correll et al., 2007; Petit, 2007), we sought to create a pair of similar but not identical CVs (CV A and CV B) for each job type, to minimize participants' suspicion about manipulations and about the purpose of the experiment. Specifically, we tweaked the precise date of birth, the dates, locations, and content of previous work experiences, where and when candidates got their degree, the content and dates of training courses for teachers, which programming languages engineers knew and with which proficiency, and the hobbies listed.

In the pre-test phase, we made sure that such heterogeneity across the CV pair did not result in systematic differences in the evaluation of candidate A vis-à-vis candidate B. Employers evaluated each candidate separately on a set of items for ability, commitment, and expressed the likelihood with which they would hire each candidate on a scale from 0 to 100 (see Section 2.4 for details on our dependent variables). Similar to later in the test, we looked for differences between candidates A and B in terms of each outcome by means of multilevel linear models. None of these differences were statistically significant at conventional levels nor sizeable (for all outcomes, Cohen's $d < .2$). Nevertheless, to further ensure that CV type (A or B) did not influence our results, treatments were assigned orthogonally to CV type in the test phase.

In line with common practices in the literature (for a review, Neumark, 2016), we manipulated our second binary treatment variable, sex of the job applicant, by varying the names of job candidates. Four pairs of Dutch names and surnames were randomly extracted from a list compiled in a previous study similarly manipulating names, albeit to gauge the presence

of ethnic discrimination in the Dutch labour market (Blommaert et al., 2014). The four resulting pairs were Roos Kosters and Eline Vos for women, Thijs Blom and Bram Ouwehand for men. We stuck to Dutch-native names and surnames, as the analysis of the intersection of sex and ethnic discrimination is beyond the scope of our experiment. Considering that the average age at first childbirth for Dutch women is around 29 years old (OECD, 2016), we set the age of candidates at around 30 years old. This choice makes it credible for a job candidate to have one child or, if not, to be of childbearing age anyway. Indeed, under our model of statistical discrimination, we expect employers to display risk aversion against women of childbearing age, regardless of actual parental status. As for the exact manipulation of parental status, our third binary treatment, the candidate is either referred to as “married, with one child” for parents or simply “married” otherwise (e.g. Petit, 2007; Bygren et al., 2017).

To recap, our two survey experiments share a 2 (sex) by 2 (parental status) by 2 (type of contract) factorial design. Sex of the applicant is a between-subjects factor, to avoid excessive suspicion on the purpose of the survey on the part of participants. The type of contract is also a between-subjects factor to ease participants’ task so that they would have to read only one job ad. Finally, parental status is a within-subjects factor coherently with previous studies (Correll et al., 2007). This means that, for instance, a given participant could start from a job ad for a primary-school teacher, with a temporary contract being offered, and evaluate first a CV (type B) for the candidate Roos Kosters without kids and secondly evaluate a CV (type A) for the candidate Eline Vos with kids⁶. The order in which CV types A and B, names of the candidates, and mentions of parental status appeared was counterbalanced across participants in such a way to avoid any systematic order effect.

2.4. Outcomes and models: candidate evaluations and hiring decisions

Once presented with a CV, employers evaluated the candidate first in terms of expected *competence*, further decomposed into the two constructs of ability and commitment (Ridgeway and Correll, 2004; Correll et al., 2007; Ridgeway, 2011). To capture the ability component we adopted seven items from previous studies (e.g. Fiske et al., 2002; Correll et al., 2007) and

asked how “competent”, “confident”, “independent”, “competitive”, “intelligent”, “skilled”, and “well-trained” employers found a given candidate. For commitment (e.g. Correll et al., 2007; Heilman and Okimoto, 2008), we asked to what extent employers expected the candidate “to be very committed to the company (school)”, “to make sacrifices for the job”, and “to make work a top priority” if hired. Each of these items was measured on a 7-point scale, with responses ranging from “not at all” (1) to “extremely” (7). Both in the pre-test and test phases we achieved high reliability for the resulting composite measures of ability and commitment, with Cronbach’s α s ranging from a minimum of 0.77 to a maximum of 0.93. We also included a more factual measure of work commitment by asking employers how many hours they would expect a given candidate to work on a weekly basis. Similar to the online job ads we surveyed, our fictitious job ads featured a 32- to 40-hour working week. Given that part-time work (under 35 h) is notoriously widespread in the Netherlands, especially among women and not limited to those among them with small children (e.g. Bosch et al., 2010), we are cautious nonetheless on the extent to which this variable may tap expected commitment.

Finally, employers were asked to estimate with what likelihood they would recommend a candidate for hire on a scale from 0 to 100 (e.g. Di Stasio and Gërkhani, 2015) and what monthly salary offer they would make to the candidate (e.g. Correll et al., 2007). Both outcomes are transformed, dividing by 100 the likelihood of hire and taking the logarithm of the salary offer, to enable the interpretation of regression coefficients in terms of percentage-point changes⁷. Hiring decisions (likelihood of hire, salary) and candidate ratings (ability, commitment, and expected hours) are our dependent variables in multilevel linear regression models⁸, with job candidates (CVs) nested within employers. For each of our dependent variables, our specification of choice includes the main effects of job candidates’ sex and parental status. The interaction between the two is our main focus as it captures the motherhood penalty, if any (e.g. Correll et al., 2007). For both job types, we ran our models first on the full sample and secondly separate by type of work contract. Full estimation results, including a simple bivariate model comprising only the main effect for sex of the job candidate, are

available in the Appendix.

3. Findings

Evidence from our primary-school teacher experiment is summarized in Figure 1 (see column 2 of Tables A1-A5 in the Appendix). For the likelihood of hire, we find a small positive effect of candidate's sex (≈ 4 percentage points, $p = .030$) indicating a preference for women, conditional though on the fact that they are childless. Looking at the interaction between candidate's sex and parental status, parental status depresses the likelihood of hire more for women than for men (-2.5 percentage points, $p = .062$). The ranking of candidates for hire in the teacher vacancy sees childless women come first, followed by mothers, and with men ranked at the bottom. No such pattern is discernible in the second panel of Figure 1, when looking at salary offers. Hence, while we do find some support for a pro-women bias in female-typical jobs, candidates giving evidence of being a mother share smaller advantages over men than childless women do. This stands in partial contrast with our *status reversal* hypothesis, since a motherhood penalty is found for our primary-school teacher vacancy, even if only with respect to childless women.

To make sense of hiring decisions we then look at ability and commitment ratings, as well as at expected working hours in the third to fifth panel of Figure 1. While we reach no conclusion on ability rankings, we find similar patterns to the ones highlighted for the likelihood of hire for commitment ratings. These are slightly skewed in favour of women, but again parental status moderates this advantage to the detriment of mothers. As for expected working hours, candidates' parental status reduces employers' estimate and particularly so for women as compared to men, as indicated by the interaction coefficient (≈ -2 hours, $p < .001$). All in all, little support is found for a *status reversal* in the primary-school teacher job. Mothers are perceived as less committed and this has (small) negative consequences in terms of their hiring chances, with respect to childless women but not men.

Turning to the software engineer job vacancy in Figure 2 (see column 2 in Tables A6 to A10), we surprisingly find the same ranking of candidates in terms of likelihood of hire, with childless women topping employers' preferences, followed by mothers and men. Differences

between candidates are smaller in size, as compared to the teacher experiment, and none of them reaches statistical significance at conventional levels. Also similar to our female-typical vacancy, we cannot detect any gradient by sex of the candidate and parental status in terms of salary offers or ability.

For our measures of commitment, evidence is mixed. The presence of children reduces employers' expectation regarding a candidate's commitment more for women than for men (gender-by-parental status interaction = $-.199$, $p = .053$). Differently, employers rank men at the top of expected hours, followed by childless women, and then by male and female parents. As shown in the fifth panel of Figure 2, both being female and being a parent decrease a candidate's expected working hours, but parental status seems not to hurt women more than men (gender-by-parental status interaction = $-.692$, $p = .186$).

Taken together, findings from both experiments fit with neither economic nor sociopsychological models of discrimination. Contrary to our *family-risk* hypothesis, core to the statistical discrimination scenario, women are not disadvantaged irrespective of parental status. Importantly, this lack of pro-male bias is more evident in male-typical jobs where our *occupation-risk* hypothesis would have predicted harsher statistical discrimination against women. As for status-characteristic theory, we do find some support for a *motherhood penalty*, but surprisingly only with respect to childless women and for primary-school teacher job vacancies. Also, if *status beliefs* are at work, they do so similarly across our male- and female-typical jobs. We cannot discern a status ranking in terms of ability, but only in terms of expected commitment, disadvantaging mothers albeit with respect to childless women only.

In line with our last, *contract-risk* hypothesis, we investigate heterogeneity by work contract. This may unveil patterns coherent with statistical discrimination, which we expected to be more prominent when permanent contracts are offered in male-typical jobs. We examine this prediction in Figures 3 and 4 (corresponding to columns 3 and 4 of Tables A1-A10). Comparing coefficients across models separate by type of contract, we find little evidence of heterogeneous effects. For our male-typical job vacancy in particular (Figure 4), female candidates are no less (nor more) likely to be hired or paid lower salaries than men when a

permanent contract is offered.

4. Discussion and conclusions

Confronting fictitious and identical CVs, Dutch employers in our study would hire women slightly more than men, and childless women more than mothers, to fill a primary-school teacher vacancy. Such preference ranking is not detected in a second experiment, this time for a software engineer job vacancy, nor do we detect discrimination in salary offers in both experiments. Female candidates that give evidence of being a mother are expected to be less committed to their job, but again their disadvantage is small in size. Clear-cut and sizeable effects are only detected for expected working hours, with mothers (both parents) expected to work the lowest amount of weekly hours when applying for a primary-school teacher job (software engineer job).

Our conclusions are, first of all, sample-dependent. We were able to draw on a sample of Dutch individuals, participating in a pre-existing online panel survey and with HR responsibilities in their current job. The latter feature in particular enhances the external validity of our study, improving on previous research typically relying on undergraduates in controlled experimental settings (Correll et al., 2007; Koch et al., 2015). Since our employers came from all sectors of the economy, however, they may not have had specialised knowledge of the two occupation-specific job markets of primary school teacher and software engineer. Yet, previous studies could not detect a preference for male over female candidates for the position of software engineer, even when involving Dutch employers in the IT sector (Di Stasio and van de Werfhorst, 2016). For the teacher vacancy, our findings are in line with those of two field experiments ran in Sweden for similar job titles (Carlsson, 2011; Bygren et al., 2017). Effect *sizes* are also in line with those found investigating gender-by-parental-status interactions in previous studies for other European countries (e.g. Bygren et al., 2017; Oesch et al., 2017). The small magnitude of these effects stands out, particularly in comparison with that highlighted for the US context (Correll et al., 2007).

Findings from our two survey experiments are also bounded to the successful manipulation of our treatments. Once employers finished rating CVs, we asked them *a*) whether the can-

didates they just evaluated were both female or both male, *b*) whether one of the candidates had a kid and the other did not, and *c*) whether the vacancy they were required to fill was permanent or temporary. While more than 70% of employers correctly recalled features *a*) and *b*) of job candidates in both experiments, only 46-47% correctly recalled the type of work contract being offered in the job ad. While many factors can bias recall, our estimates should nonetheless be regarded as intention-to-treat.

Overall we find little evidence of a motherhood penalty in hiring in two sex-typical jobs in the Netherlands, in line with what field experiments and vignette studies have shown in other European countries (Petit, 2007; Bygren et al., 2017; Oesch et al., 2017). In our experiments, this holds even if Dutch employers expect mothers to be less committed to their job and work fewer hours a week than men and women without kids. This may speak to the relevance of context for hiring decisions (e.g. Bills et al., 2017). In the Netherlands, women in part-time jobs have similar chances to receive firm-sponsored training (Picchio and van Ours, 2016) and experience comparatively small wage penalties (Fouarge and Muffels, 2009) vis-à-vis their full-time counterparts. The spread of part-time work arrangements (Bosch et al., 2010) and the weakening of the ‘work obligation’ norm (Wielers and Raven, 2013), one giving work the priority over all other spheres of life, may also explain why Dutch employers’ expectations regarding a candidate’s commitment and working hours have small or nil spillover effects on hiring decisions. Future research could shed light on whether employers hold similar expectations, but to much greater detriment for the hiring of female job candidates, in contexts where women (mothers) are expected to work par-time yet part-time jobs are of much poorer quality career-wise.

On balance, results in this study fail to meet our predictions under both statistical and status-based models of discrimination. Our study is, to the best of our knowledge, the first to explicitly pit these theories against each other while examining gender-by-parental-status hiring discrimination (cfr. Correll et al., 2007). We particularly focused on the screening of CVs: most of our employers carried out real-world screenings in their jobs and such stage of recruitment arguably presents a favourable opportunity structure for discrimination (e.g.

Petersen and Saporta, 2004). Yet, we cannot assess here whether discrimination, either of the statistical or of the status-based kind, may occur later in the hiring process or whether and how it may affect career progression. The current debate could thus benefit from accumulating tests of these two theories not only in other contexts, or for different job titles, but also considering more fully the stages in which employers reward or penalise job candidates and current employees.

Although not the main focus in our survey experiments, investigating gender discrimination in hiring in sex-typical jobs may also highlight whether or not employers contribute to occupational segregation in labour markets. In this respect, Dutch employers' preference for female candidates *all else equal* for a primary-school teacher vacancy is consistent with previous experimental research on hiring discrimination in female-typical jobs (Booth and Leigh, 2010; Carlsson, 2011; Rich, 2014). Yet, the presence of, if anything, very small effects may rather strengthen the stance of those pointing to persistent gender segregation by field of study (Barone, 2011; van de Werfhorst, 2017) as the main source of occupational segregation.

Notes

¹For example, Dutch employers particularly value a candidate's field of study and occupation-specific degrees (Di Stasio and van de Werfhorst, 2016). We felt that omitting such a commonplace information in order to simulate a low-information setting would have further compromised the external validity of our study.

²Both for the pre-test and test phase sample size was determined with reference to previous studies with a similar design (Correll et al., 2007) and drawing from the total pool of 749 potential participants previously identified in LISS through our filter questions. Prior to data collection, we defined clear stopping rules for both the pre-test and the test keeping into account that the response rate in LISS usually exceeds 80%. Employers involved in the pre-test phase were not re-contacted for the test.

³These numbers largely reflect persistent gender segregation in the correspondent fields of study (Barone, 2011; van de Werfhorst, 2017), yet we ask here if employers practice hiring discrimination even if confronted with candidates with equal educational qualifications (work experience) in the relevant field (occupation).

⁴For Dutch teachers, salaries are set with reference to a national standard nonetheless (*Salarisschaal*)

⁵We made inquiries into online jobs postings of websites such as `indeed.nl`, `nationalevacaturebank.nl`, and `meesterbaan.nl`.

⁶Throughout the test phase, employers give individual ratings to job applicants, meaning that there are no instructions that would explicitly encourage any comparison or choice between the two candidates forming a given pair. Previous research on gender discrimination in hiring has found no evidence of heterogeneous effects depending on the type of rating (individual v. comparative) participants have been asked to give (Koch et al., 2015).

⁷Such transformations do not alter the substantive or statistical significance of our results. Model estimates deploying untransformed outcomes are available upon request.

⁸Multilevel models are chosen to account for the nested nature of our data (CVs nested within employers/subjects) and to allow for the estimation of both between- and within-subjects factors. Alternatively, one can opt to include employer fixed effects in simple linear regression and thereby retain the possibility of estimating coefficients for within-subjects factors and for the interaction(s) of within- and between-subjects factors. Such fixed-effects estimates are available upon request. They are indistinguishable from their counterparts in our multilevel models of choice.

Tables & graphs

Table 1: Summary of sample features for each survey experiment.

	Primary-school teacher	Software engineer
Female	0.44	0.37
Age	46.57	47.08
Degree (HBO or WO)	0.60	0.62
Have a partner	0.81	0.78
Have children	0.49	0.49
Have the power to hire/fire	0.47	0.43
Screen candidates (CVs, interviews)	0.90	0.87
Set or influence pay	0.44	0.35
Influence or decide on promotions	0.48	0.45
Number of employers	239	237

Source: LISS 2016.

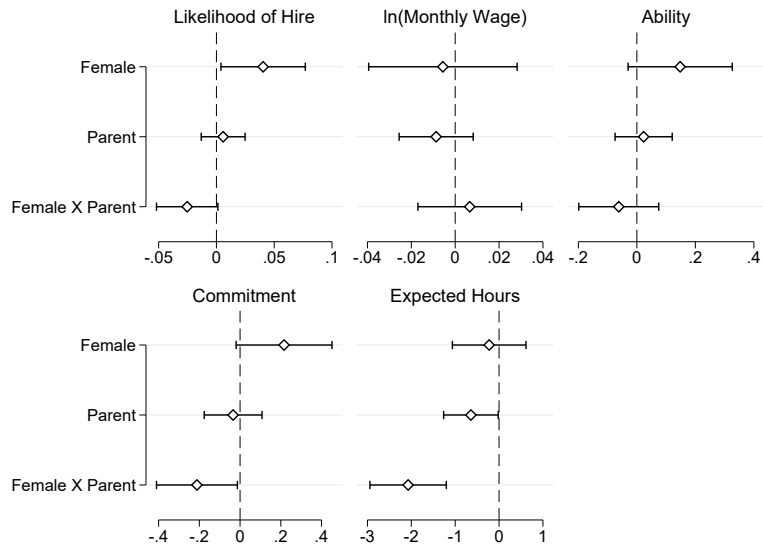


Figure 1: Point estimates and 95% confidence intervals from multilevel linear models for hiring decisions and ratings in the primary-school teacher experiment.

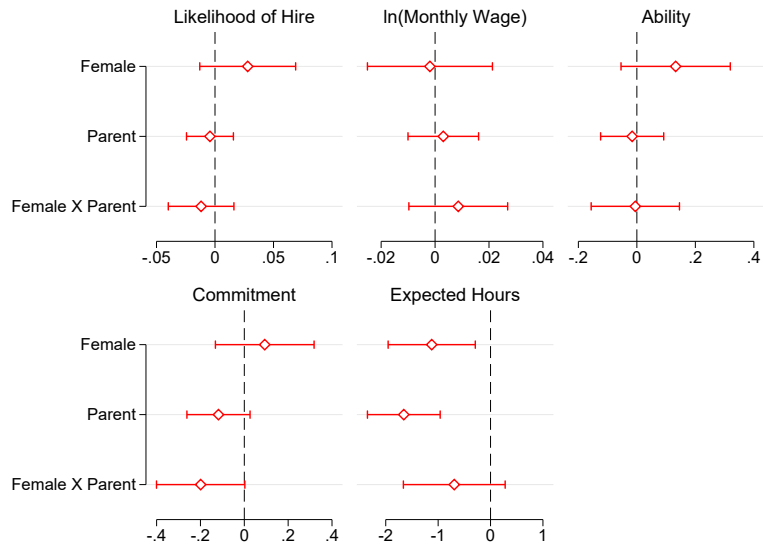


Figure 2: Point estimates and 95% confidence intervals from multilevel linear models for hiring decisions and ratings in the software engineer experiment.

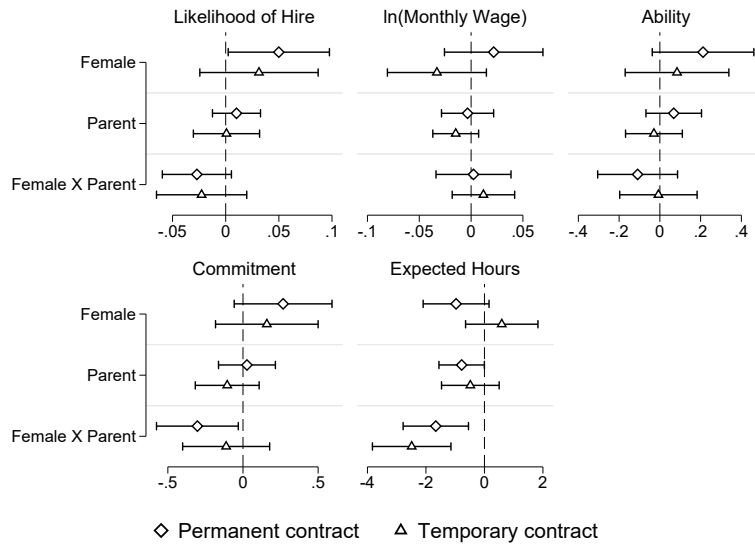


Figure 3: Point estimates and 95% confidence intervals from multilevel linear models for hiring decisions and ratings in the primary-school teacher experiment, separate by work contract.

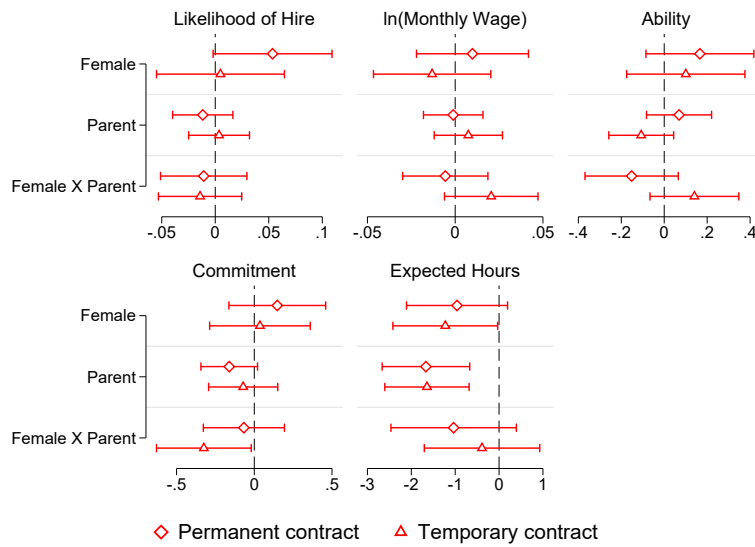


Figure 4: Point estimates and 95% confidence intervals from multilevel linear models for hiring decisions and ratings in the software engineer experiment, separate by work contract.

Appendix

Table 1A: Multilevel linear models for the likelihood of hire. Primary-school teacher experiment (LISS 2017).

	(1)	(2)	(3)	(4)
	Whole sample	Whole sample	Temporary contract	Permanent contract
	β (SE)	β (SE)	β (SE)	β (SE)
Female	0.028 (0.017)	0.040** (0.019)	0.031 (0.028)	0.050** (0.024)
Parent		0.006 (0.010)	0.001 (0.016)	0.010 (0.012)
Female \times Parent		-0.025* (0.014)	-0.023 (0.022)	-0.027 (0.017)
LR test (1,2)		4.55 ($p = .103$)		
ICC	0.731	0.735	0.709	0.767
Number of employers	239	239	117	122
Number of observations	478	478	234	244

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

Table 2A: Multilevel linear models for the monthly wage offer (logged). Primary-school teacher experiment (LISS 2017).

	(1)	(2)	(3)	(4)
	Whole sample	Whole sample	Temporary contract	Permanent contract
	β (SE)	β (SE)	β (SE)	β (SE)
Female	-0.002 (0.016)	-0.006 (0.017)	-0.033 (0.024)	0.022 (0.024)
Parent		-0.009 (0.009)	-0.015 (0.011)	-0.003 (0.013)
Female \times Parent		0.007 (0.012)	0.012 (0.015)	0.002 (0.018)
LR test (1,2)		1.07 ($p = 0.586$)		
ICC	0.755	0.756	0.801	0.710
Number of employers	239	239	117	122
Number of observations	478	478	234	244

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

Table 3A: Multilevel linear models for ability ratings. Primary-school teacher experiment (LISS 2017).

	(1)	(2)	(3)	(4)
	Whole sample	Whole sample	Temporary contract	Permanent contract
	β (SE)	β (SE)	β (SE)	β (SE)
Female	0.117 (0.084)	0.148 (0.091)	0.084 (0.130)	0.212* (0.127)
Parent		0.023 (0.050)	-0.029 (0.071)	0.068 (0.070)
Female \times Parent		-0.062 (0.070)	-0.007 (0.097)	-0.109 (0.100)
LR test (1,2)		0.84 ($p = .656$)		
ICC	0.705	0.705	0.721	0.691
Number of employers	239	239	117	122
Number of observations	478	478	234	244

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

Table 4A: Multilevel linear models for commitment ratings. Primary-school teacher experiment (LISS 2017).

	(1)	(2)	(3)	(4)
	Whole sample	Whole sample	Temporary contract	Permanent contract
	β (SE)	β (SE)	β (SE)	β (SE)
Female	0.110 (0.109)	0.216* (0.120)	0.159 (0.174)	0.267 (0.166)
Parent		-0.034 (0.072)	-0.105 (0.108)	0.026 (0.097)
Female \times Parent		-0.212** (0.101)	-0.112 (0.148)	-0.303** (0.139)
LR test (1,2)		11.96 ($p = .003$)		
ICC	0.629	0.644	0.640	0.650
Number of employers	239	239	117	122
Number of observations	478	478	234	244

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

Table 5A: Multilevel linear models for expected working hours. Primary-school teacher experiment (LISS 2017).

	(1)	(2)	(3)	(4)
	Whole sample	Whole sample	Temporary contract	Permanent contract
	β (SE)	β (SE)	β (SE)	β (SE)
Female	-1.261*** (0.366)	-0.225 (0.428)	0.593 (0.631)	-0.969* (0.575)
Parent		-0.641** (0.317)	-0.481 (0.503)	-0.778** (0.396)
Female \times Parent		-2.072*** (0.444)	-2.487*** (0.685)	-1.663*** (0.570)
LR test (1,2)		69.28 ($p < .001$)		
ICC	0.341	0.462	0.412	0.508
Number of employers	239	239	117	122
Number of observations	478	478	234	244

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

Table 6A: Multilevel linear models for the likelihood of hire. Software engineer experiment (LISS 2017).

	(1)	(2)	(3)	(4)
	Whole sample	Whole sample	Temporary contract	Permanent contract
	β (SE)	β (SE)	β (SE)	β (SE)
Female	0.022 (0.020)	0.028 (0.021)	0.005 (0.030)	0.054* (0.028)
Parent		-0.004 (0.010)	0.004 (0.015)	-0.012 (0.014)
Female \times Parent		-0.012 (0.014)	-0.014 (0.020)	-0.011 (0.021)
LR test (1,2)		2.78 ($p = 0.249$)		
ICC	0.763	0.766	0.788	0.736
Number of employers	237	237	120	117
Number of observations	474	474	240	234

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

Table 7A: Multilevel linear models for the monthly wage offer (logged). Software engineer experiment (LISS 2017).

	(1)	(2)	(3)	(4)
	Whole sample	Whole sample	Temporary contract	Permanent contract
	β (SE)	β (SE)	β (SE)	β (SE)
Female	0.002 (0.011)	-0.002 (0.012)	-0.013 (0.017)	0.010 (0.016)
Parent		0.003 (0.007)	0.007 (0.010)	-0.001 (0.009)
Female \times Parent		0.009 (0.009)	0.021 (0.014)	-0.006 (0.012)
LR test (1,2)		3.36 ($p = 0.186$)		
ICC	0.684	0.688	0.683	0.710
Number of employers	237	237	120	117
Number of observations	474	474	240	234

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

Table 8A: Multilevel linear models for ability ratings. Software engineer experiment (LISS 2017).

	(1)	(2)	(3)	(4)
	Whole sample	Whole sample	Temporary contract	Permanent contract
	β (SE)	β (SE)	β (SE)	β (SE)
Female	0.130 (0.087)	0.133 (0.095)	0.100 (0.141)	0.166 (0.128)
Parent		-0.016 (0.055)	-0.107 (0.077)	0.069 (0.077)
Female \times Parent		-0.005 (0.077)	0.141 (0.106)	-0.152 (0.111)
LR test (1,2)		0.24 ($p = .887$)		
ICC	0.673	0.673	0.718	0.626
Number of employers	237	237	120	117
Number of observations	474	474	240	234

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

Table 9A: Multilevel linear models for commitment ratings. Software engineer experiment (LISS 2017).

	(1)	(2)	(3)	(4)
	Whole sample	Whole sample	Temporary contract	Permanent contract
	β (SE)	β (SE)	β (SE)	β (SE)
Female	-0.006 (0.103)	0.093 (0.115)	0.036 (0.165)	0.148 (0.159)
Parent		-0.118 (0.073)	-0.071 (0.114)	-0.161* (0.093)
Female \times Parent		-0.199* (0.103)	-0.324** (0.155)	-0.067 (0.133)
LR test (1,2)		21.03 ($p < .001$)		
ICC	0.570	0.599	0.557	0.649
Number of employers	237	237	120	117
Number of observations	474	474	240	234

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

Table 10A: Multilevel linear models for expected working hours. Software engineer experiment (LISS 2017).

	(1)	(2)	(3)	(4)
	Whole sample	Whole sample	Temporary contract	Permanent contract
	β (SE)	β (SE)	β (SE)	β (SE)
Female	-1.468*** (0.344)	-1.122*** (0.424)	-1.225** (0.609)	-0.958 (0.588)
Parent		-1.655*** (0.354)	-1.643*** (0.490)	-1.667*** (0.509)
Female \times Parent		-0.692 (0.496)	-0.388 (0.671)	-1.035 (0.729)
LR test (1,2)		59.53 ($p < .001$)		
ICC	0.201	0.318	0.393	0.229
Number of employers	237	237	120	117
Number of observations	474	474	240	234

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

References

- Abendroth, A. K., M. L. Huffman, and J. Treas (2014). The Parity Penalty in Life Course Perspective: Motherhood and Occupational Status in 13 European Countries. *American Sociological Review* 79(5), 993–1014.
- Adda, J., C. Dustmann, and K. Stevens (2017). The career costs of children. *Journal of Political Economy* 125(2), 293–337.
- Adler, M. A. and K. Lenz (2017). *Father involvement in the early years: An international comparison of policy and practice*. Bristol University Press.
- Ahn, N. and P. Mira (2002). A note on the changing relationship between fertility and female employment rates in developed countries. *Journal of Population Economics* 15(4), 667–682.
- Aigner, D. J. and G. G. Cain (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review* 30(2), 175–187.
- Aisenbrey, S., M. Evertsson, and D. Grunow (2009). Is there a career penalty for mothers' time out? A comparison of Germany, Sweden and the United States. *Social Forces* 88(2), 573–605.
- 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.
- Akgündüz, Y. E. and T. van Huizen (2015). Training in two-tier labor markets: The role of job match quality. *Social Science Research* 52, 508–521.
- 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.
- Altintas, E. and O. Sullivan (2016). 50 years of change updated: cross-national gender convergence in housework. *Demographic Research* 35(16).
- Altonji, J. G. and C. R. Pierret (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics* 116(1), 313–350.
- 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 (2002). The motherhood wage penalty: Which mothers pay it and why? *American Economic Review* 92(2), 354–358.
- 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.
- Angelov, N., P. Johansson, and E. Lindahl (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics* 34(3), 545–579.
- Anxo, D., C. Fagan, I. Cebrian, and G. Moreno (2007). Patterns of labour market integration in Europe — a life course perspective on time policies. *Socio-Economic Review* 5(2), 233–260.
- 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.
- Aronow, P. M. and C. Samii (2016). Does regression produce representative estimates of causal effects? *American Journal of Political Science* 60(1), 250–267.
- Arulampalam, W. and A. L. Booth (1998). Training and labour market flexibility: is there a trade-off? *British Journal of Industrial Relations* 36(4), 521–536.
- 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.
- Athey, S. and G. W. Imbens (2017). The Econometrics of Randomized Experiments. In E. Duflo and A. Banerjee (Eds.), *Handbook of Economic Field Experiments*, Volume 1, pp. 73–140. Elsevier.
- Atkinson, A. B. (2015). *Inequality. What can be done?* Harvard University Press.
- Avdic, D. and A. Karimi (2018). Modern family? Paternity leave and marital stability. *American Economic Journal: Applied Economics* 10(4), 283–307.
- Azmat, G. and B. Petrongolo (2014). Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics* 30, 32–40.
- Baert, S. (2014). Career lesbians. Getting hired for not having kids? *Industrial Relations Journal* 45(6), 543–561.
- Baert, S., A.-S. De Pauw, and N. Deschacht (2016). Do employer preferences contribute to sticky floors? *ILR Review* 69(3), 714–736.
- Baird, M. and M. O’Brien (2015). Dynamics of parental leave in Anglophone countries: the paradox of state expansion in liberal welfare regimes. *Community, Work & Family* 18(2), 198–217.

- Baker, M. and K. Milligan (2008). How does job-protected maternity leave affect mothers' employment? *Journal of Labor Economics* 26(4), 655–691.
- Balbo, N., F. C. Billari, and M. Mills (2013). Fertility in advanced societies: A review of research. *European Journal of Population/Revue européenne de Démographie* 29(1), 1–38.
- Bar-Haim, E., L. Chauvel, J. C. Gornick, and A. Hartung (2018). The persistence of the gender earnings gap: Cohort trends and the role of education in twelve countries. *LIS Working Paper Series No. 737*.
- Bardasi, E. and J. C. Gornick (2008). Working for less? Women's part-time wage penalties across countries. *Feminist Economics* 14(1), 37–72.
- Barone, C. (2011). Some things never change: Gender segregation in higher education across eight nations and three decades. *Sociology of Education* 84(2), 157–176.
- 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.
- Bayard, K., J. Hellerstein, D. Neumark, and K. Troske (2003). New evidence on sex segregation and sex differences in wages from matched employee-employer data. *Journal of Labor Economics* 21(4), 887–922.
- Beblo, M., S. Bender, and E. Wolf (2009). Establishment-level wage effects of entering motherhood. *Oxford Economic Papers* 61, i11–i34.
- Becker, G. S. (1964). Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. *New York: Columbia University Press*.
- Becker, G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal* 75(299), 493–517.
- Becker, G. S. (1981). *A Treatise on the Family*. Cambridge, MA: Harvard University Press.
- Becker, G. S. (1985). Human capital, effort, and the sexual division of labor. *Journal of Labor Economics* 3(1), S33–S58.
- Becker, P. E. and P. Moen (1999). Scaling back: Dual-earner couples' work-family strategies. *Journal of Marriage and the Family* 61(4), 995–1007.
- Begall, K. and D. Grunow (2015). Labour Force Transitions around First Childbirth in the Netherlands. *European Sociological Review* 31(6), 697–712.
- Bergemann, A. and R. T. Riphahn (2017). Maternal Employment Effects of Paid Parental Leave. *SOEPpapers on Multidisciplinary Panel Data Research 900–2017*.
- Berger, J., J. Berger, M. H. Fisek, and R. Z. Norman (1977). *Status characteristics and social interaction: An expectation-states approach*. New York: Elsevier.

- 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.
- Bertrand, M., C. Goldin, and L. F. Katz (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics* 2(3), 228–255.
- Bettendorf, L. J., K. Folmer, and E. L. Jongen (2014). The dog that did not bark: The EITC for single mothers in the Netherlands. *Journal of Public Economics* 119, 49–60.
- Bettio, F. and A. Verashchagina (2009). *Gender segregation in the labour market: Root causes, implications and policy responses in the EU*. Luxembourg: Publications Office of the European Union.
- Bhalotra, S. and D. Clarke (2018). Twin Birth and Maternal Condition. *IZA Discussion Paper No. 11742*.
- Bianchi, S. M. and M. A. Milkie (2010). Work and family research in the first decade of the 21st century. *Journal of Marriage and Family* 72(3), 705–725.
- Bick, A. and N. Fuchs-Schündeln (2017). Taxation and Labour Supply of Married Couples across Countries: A Macroeconomic Analysis. *The Review of Economic Studies* 85(3), 1543–1576.
- Biegert, T. (2014). On the outside looking in? Transitions out of non-employment in the United Kingdom and Germany. *Journal of European Social Policy* 24(1), 3–18.
- 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.
- Bielby, D. D. V. and W. T. Bielby (1984). Work commitment, sex-role attitudes, and women’s employment. *American Sociological Review* 49(2), 234–247.
- Biewen, M. and S. Seifert (2016). Potential Parenthood and Career Progression of Men and Women: A Simultaneous Hazards Approach. *IZA Discussion Paper No. 10050*.
- Bills, D. B., V. Di Stasio, and K. Gërkhani (2017). The demand side of hiring: Employers in the labor market. *Annual Review of Sociology* 43, 291–310.
- Blackburn, M. L., D. E. Bloom, and D. Neumark (1993). Fertility timing, wages, and human capital. *Journal of Population Economics* 6(1), 1–30.
- Blau, F. D., M. C. Brinton, and D. B. Grusky (2006). The Declining Significance of Gender? In F. D. Blau, M. C. Brinton, and D. B. Grusky (Eds.), *The Declining Significance of Gender?* New York, NY: Russell Sage Foundation.

- Blau, F. D. and L. M. Kahn (2000). Gender Differences in Pay. *The Journal of Economic Perspectives* 14(4), 75–99.
- Blau, F. D. and L. M. Kahn (2006). The US gender pay gap in the 1990s: Slowing convergence. *ILR Review* 60(1), 45–66.
- Blau, F. D. and L. M. Kahn (2013). Female labor supply: Why is the United States falling behind? *The American Economic Review* 103(3), 251–256.
- Blau, F. D. and L. M. Kahn (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature* 55(3), 789–865.
- Blommaert, L., M. Coenders, and F. Van Tubergen (2014). Discrimination of Arabic-named applicants in the Netherlands: an internet-based field experiment examining different phases in online recruitment procedures. *Social Forces* 92(3), 957–982.
- Blum, S., A. Koslowski, A. Macht, and P. e. Moss (2018). *International Review of Leave Policies and Research 2018*. Available at: http://www.leavenetwork.org/lp_and_r_reports/.
- Blundell, R., M. Brewer, and M. Francesconi (2008). Job changes and hours changes: understanding the path of labor supply adjustment. *Journal of Labor Economics* 26(3), 421–453.
- Blundell, R., A. Gosling, H. Ichimura, and C. Meghir (2007). Changes in the distribution of male and female wages accounting for employment composition using bounds. *Econometrica* 75(2), 323–363.
- Böheim, R. and M. P. Taylor (2004). Actual and preferred working hours. *British Journal of Industrial Relations* 42(1), 149–166.
- Bonhomme, S. and G. Jolivet (2009). The pervasive absence of compensating differentials. *Journal of Applied Econometrics* 24(5), 763–795.
- Booth, A. and A. Leigh (2010). Do employers discriminate by gender? A field experiment in female-dominated occupations. *Economics Letters* 107(2), 236–238.
- Booth, A. L. and J. C. Van Ours (2008). Job satisfaction and family happiness: the part-time work puzzle. *The Economic Journal* 118(526), F77–F99.
- Booth, A. L. and J. C. van Ours (2013). Part-time jobs: What women want? *Journal of Population Economics* 26(1), 263–283.
- Borusyak, K. and X. Jaravel (2016). Revisiting Event Study Designs. *Working Paper*, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2826228.

- Bosch, N., A. Deelen, and R. Euwals (2010). Is Part-time Employment Here to Stay? Working Hours of Dutch Women over Successive Generations. *Labour* 24(1), 35–54.
- Bowles, S., H. Gintis, and M. Osborne (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature* 39(4), 1137–1176.
- Breen, R. and J. Ermisch (2017). Educational reproduction in Great Britain: a prospective approach. *European Sociological Review* 33(4), 590–603.
- Brinton, M. C., Y.-J. Lee, and W. L. Parish (1995). Married women’s employment in rapidly industrializing societies: Examples from East Asia. *American Journal of Sociology* 100(5), 1099–1130.
- Brüderl, J. and V. Ludwig (2015). Fixed-effects panel regression. In H. Best and C. Wolf (Eds.), *The Sage Handbook of Regression Analysis and Causal Inference*, pp. 327–358. London: Sage.
- Brülle, J., M. Gangl, A. Levanon, and E. Saburov (2018). Changing labour market risks in the service economy: Low wages, part-time employment and the trend in working poverty risks in Germany. *Journal of European Social Policy*, <https://doi.org/10.1177/0958928718779482>.
- Bryan, M. L. (2007). Workers, Workplaces and Working Hours. *British Journal of Industrial Relations* 45(4), 735–759.
- Bryan, M. L. and A. Sevilla (2017). Flexible working in the UK and its impact on couples’ time coordination. *Review of Economics of the Household* 15(4), 1415–1437.
- Bryan, M. L. and A. Sevilla-Sanz (2011). Does housework lower wages? Evidence for Britain. *Oxford Economic Papers* 63(1), 187–210.
- Buck, N. and S. McFall (2011). Understanding Society: design overview. *Longitudinal and Life Course Studies* 3(1), 5–17.
- Budig, M. J. and P. England (2001). The wage penalty for motherhood. *American Sociological Review* 66(2), 204–225.
- Budig, M. J., J. Misra, and I. Boeckmann (2016). Work–family policy trade-offs for mothers? Unpacking the cross-national variation in motherhood earnings penalties. *Work and Occupations* 43(2), 119–177.
- 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.
- Bukodi, E. and S. Dex (2010). Bad start: Is there a way up? Gender differences in the effect of initial occupation on early career mobility in Britain. *European Sociological Review* 4(26), 431–446.

- Bukodi, E., J. Goldthorpe, H. Joshi, and L. Waller (2017). Why have relative rates of class mobility become more equal among women in Britain? *The British Journal of Sociology* 68, 512–532.
- Buligescu, B., D. De Crombrughe, G. Menteşoğlu, 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.
- Bünning, M. and M. Pollmann-Schult (2016). Family policies and fathers’ working hours: cross-national differences in the paternal labour supply. *Work, Employment and Society* 30(2), 256–274.
- Burgess, S., P. Gregg, C. Propper, and E. Washbrook (2008). Maternity rights and mothers’ return to work. *Labour Economics* 15(2), 168–201.
- Butikofer, A., S. Jensen, and K. G. Salvanes (2018). The Role of Parenthood on the Gender Gap Among Top Earners. *European Economic Review* 109.
- Bygren, M., A. Erlandsson, and M. Gähler (2017). Do Employers Prefer Fathers? Evidence from a Field Experiment Testing the Gender by Parenthood Interaction Effect on Callbacks to Job Applications. *European Sociological Review* 33(3), 337–348.
- Bygren, M. and M. Gähler (2012). Family formation and men’s and women’s attainment of workplace authority. *Social Forces* 90(3), 795–816.
- Card, D., A. R. Cardoso, and P. Kline (2015). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics* 131(2), 633–686.
- Card, D., A. R. Cardoso, and P. Kline (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics* 131(2), 633–686.
- Carlsson, M. (2011). Does hiring discrimination cause gender segregation in the Swedish labor market? *Feminist Economics* 17(3), 71–102.
- Carlsson, M. and S. Eriksson (2017). The effect of age and gender on labor demand – evidence from a field experiment. *IFAU Working Paper* (2017:8).
- Cha, Y. and K. A. Weeden (2014). Overwork and the slow convergence in the gender gap in wages. *American Sociological Review* 79(3), 457–484.
- Chabé-Ferret, S. et al. (2017). Should we combine difference in differences with conditioning on pre-treatment outcomes? *TSE Working Paper* 17-824.

- Charles, K. K. and J. Guryan (2011). Studying discrimination: Fundamental challenges and recent progress. *Annual Review of Economics* 3(1), 479–511.
- Charles, K. K., J. Guryan, and J. Pan (2018). The Effects of Sexism on American Women: The Role of Norms vs. Discrimination. *NBER Working Paper No. w24904*.
- Charles, M. (2011). A world of difference: international trends in women’s economic status. *Annual Review of Sociology* 37, 355–371.
- Charles, M. and D. B. Grusky (2004). *Occupational Ghettos: The Worldwide Segregation of Women and Men*. Stanford, CA: Stanford University Press.
- Cheng, S. (2016). The Accumulation of (Dis)advantage: The Intersection of Gender and Race in the Long-Term Wage Effect of Marriage. *American Sociological Review* 81(1), 29–56.
- Christofides, L. N., A. Polycarpou, and K. Vrachimis (2013). Gender wage gaps, ‘sticky floors’ and ‘glass ceilings’ in europe. *Labour Economics* 21, 86–102.
- Chung, Y., B. Downs, D. H. Sandler, and R. Sienkiewicz (2017). The Parental Gender Earnings Gap in the United States. *Center for Economic Studies WP 17-68*.
- Cohen, P. N. and M. L. Huffman (2007). Working for the woman? Female managers and the gender wage gap. *American Sociological Review* 72(5), 681–704.
- Connolly, S., M. Aldrich, M. O’Brien, S. Speight, and E. Poole (2016). Britain’s slow movement to a gender egalitarian equilibrium: Parents and employment in the UK 2001–13. *Work, Employment & Society* 30(5), 838–857.
- Connolly, S. and M. Gregory (2008). Moving Down: Women’s Part-Time Work and Occupational Change in Britain 1991–2001. *The Economic Journal* 118(526), F52–F76.
- Cooke, L. P. (2011). *Gender-Class Equality in Political Economies*. London: Routledge.
- Cooke, L. P. (2014). Gendered parenthood penalties and premiums across the earnings distribution in Australia, the United Kingdom, and the United States. *European Sociological Review* 30(3), 360–372.
- Cooke, L. P. and J. Baxter (2010). “Families” in International Context: Comparing Institutional Effects Across Western Societies. *Journal of Marriage and Family* 72(3), 516–536.
- Cooke, L. P. and S. Fuller (2018). Class differences in establishment pathways to fatherhood wage premiums. *Journal of Marriage and Family* 80(3), 737–751.
- Cooke, T. J., P. Boyle, K. Couch, and P. Feijten (2009). A longitudinal analysis of family migration and the gender gap in earnings in the United States and Great Britain. *Demography* 46(1), 147–167.

- Cools, S., S. Markussen, and M. Strøm (2017). Children and Careers: How Family Size Affects Parents' Labor Market Outcomes in the Long Run. *Demography* 54(5), 1773–1793.
- Correll, S. J. (2001). Gender and the Career Choice Process: The Role of Biased Self-Assessments. *American Journal of Sociology* 106(6), 1691–1730.
- Correll, S. J. (2004). Constraints into preferences: Gender, status, and emerging career aspirations. *American Sociological Review* 69(1), 93–113.
- Correll, S. J. and S. Benard (2006). Biased estimators? Comparing status and statistical theories of gender discrimination. *Social Psychology of the Workplace: Advances in Group Processes* 23, 89–116.
- Correll, S. J., S. Benard, and I. Paik (2007). Getting a Job: Is There a Motherhood Penalty? *American Journal of Sociology* 112(5), 1297–1339.
- Creighton, C. (1999). The rise and decline of the “male breadwinner family” in Britain. *Cambridge Journal of Economics* 23(5), 519–541.
- Crompton, R. (1999). *Restructuring Gender Relations and Employment: The Decline of the Male Breadwinner*. Oxford: Oxford University Press.
- Crompton, R. (2006). *Employment and the family: The reconfiguration of work and family life in contemporary societies*. Cambridge, UK: Cambridge University Press.
- Cuddy, A. J., S. T. Fiske, and P. Glick (2004). When professionals become mothers, warmth doesn't cut the ice. *Journal of Social Issues* 60(4), 701–718.
- Cygan-Rehm, K. (2016). Parental leave benefit and differential fertility responses: Evidence from a German reform. *Journal of Population Economics* 29(1), 73–103.
- Cygan-Rehm, K. and M. Maeder (2013). The effect of education on fertility: Evidence from a compulsory schooling reform. *Labour Economics* 25, 35–48.
- Dariotis, J. K., J. H. Pleck, N. M. Astone, and F. L. Sonenstein (2011). Pathways of early fatherhood, marriage, and employment: a latent class growth analysis. *Demography* 48(2), 593.
- Davies, R. and G. Pierre (2005). The family gap in pay in Europe: a cross-country study. *Labour Economics* 12(4), 469–486.
- De Hoon, S., R. Keizer, and P. Dykstra (2017). The influence of motherhood on income: do partner characteristics and parity matter? *Community, Work & Family* 20(2), 211–225.
- Deaton, A. and N. Cartwright (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine* 210, 2–21.

- Dechant, A. and A. Rinklake (2016). Anticipating motherhood and fatherhood – German couples’ plans for childcare and paid work. In D. Grunow and M. Evertsson (Eds.), *Couples’ Transitions to Parenthood: Analysing Gender and Work in Europe*. Edward Elgar Publishing.
- Del Boca, D., S. Pasqua, and C. Pronzato (2009). Motherhood and market work decisions in institutional context: a European perspective. *Oxford Economic Papers* 61, i147–i171.
- Del Bono, E. and D. Vuri (2011). Job mobility and the gender wage gap in Italy. *Labour Economics* 18(1), 130–142.
- Di Nallo, A. (2018). Gender Gap in Repartnering: The Role of Parental Status and Custodial Arrangements. *Journal of Marriage and Family*.
- Di Stasio, V. and K. Gërkhani (2015). Employers’ social contacts and their hiring behavior in a factorial survey. *Social Science Research* 51, 93–107.
- Di Stasio, V. and H. G. van de Werfhorst (2016). Why Does Education Matter to Employers in Different Institutional Contexts? A Vignette Study in England and the Netherlands. *Social Forces* 95(1), 77–106.
- 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.
- Dieckhoff, M., V. Gash, and N. Steiber (2015). Measuring the effect of institutional change on gender inequality in the labour market. *Research in Social Stratification and Mobility* 39, 59–75.
- DiPrete, T. A. (2002). Life Course Risks, Mobility Regimes, and Mobility Consequences: A Comparison of Sweden, Germany, and the United States. *American Journal of Sociology* 108(2), 267–309.
- Dotti Sani, G. M. (2015). Within-couple inequality in earnings and the relative motherhood penalty. A cross-national study of European countries. *European Sociological Review* 31(6), 667–682.
- Dougherty, C. (2006). The marriage earnings premium as a distributed fixed effect. *Journal of Human Resources* 41(2), 433–443.
- Drasch, K. (2012). Educational attainment and family-related employment interruptions in Germany: do changing institutional settings matter? *European Sociological Review* 29(5), 981–995.
- Eggebeen, D. J. and C. Knoester (2001). Does fatherhood matter for men? *Journal of Marriage and Family* 63(2), 381–393.

- Ejrnaes, M. and A. Kunze (2013). Work and wage dynamics around childbirth. *The Scandinavian Journal of Economics* 115(3), 856–877.
- Ellemers, N. (2018). Gender Stereotypes. *Annual Review of Psychology* 69, 275–298.
- Elwert, F. and C. Winship (2014). Endogenous selection bias: The problem of conditioning on a collider variable. *Annual Review of Sociology* 40, 31–53.
- Elzinga, C. H. and A. C. Liefbroer (2007). De-standardization of family-life trajectories of young adults: A cross-national comparison using sequence analysis. *European Journal of Population/Revue européenne de Démographie* 23(3-4), 225–250.
- Ermisch, J. F. and R. E. Wright (1993). Wage offers and full-time and part-time employment by British women. *Journal of Human Resources* 28(1), 111–133.
- Esping-Andersen, G. (1990). *The Three Worlds of Welfare Capitalism*. John Wiley & Sons.
- Esping-Andersen, G. (1999). *Social Foundations of Postindustrial Economies*. Oxford: Oxford University Press.
- Esping-Andersen, G. (2009). *The incomplete revolution: Adapting welfare states to women's new roles*. Cambridge: Polity Press.
- Estevez-Abe, M. (2005). Gender bias in skills and social policies: The varieties of capitalism perspective on sex segregation. *Social Politics: International Studies in Gender, State & Society* 12(2), 180–215.
- Euwals, R. (2001). Female labour supply, flexibility of working hours, and job mobility. *The Economic Journal* 111(471), 120–134.
- Euwals, R., M. Knoef, and D. Van Vuuren (2011). The trend in female labour force participation: what can be expected for the future? *Empirical Economics* 40(3), 729–753.
- 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.
- Evertsson, M. and A.-Z. Duvander (2011). Parental Leave—Possibility or Trap? Does Family Leave Length Effect Swedish Women's Labour Market Opportunities? *European Sociological Review* 27(4), 435–450.
- Evertsson, M., D. Grunow, and S. Aisenbrey (2016). Work interruptions and young women's career prospects in Germany, Sweden and the US. *Work, Employment and Society* 30(2), 291–308.

- Eydal, G. B. and T. Rostgaard (Eds.) (2016). *Fatherhood in the Nordic welfare states: Comparing care policies and practice*. Bristol, UK: Policy Press.
- Fagan, C. and H. Norman (2012). Trends and social divisions in maternal employment patterns following maternity leave in the UK. *International Journal of Sociology and Social Policy* 32(9/10), 544–560.
- Fang, H. and A. Moro (2011). Theories of Statistical Discrimination and Affirmative Action: A Survey. In J. Benhabib, M. O. Jackson, and A. Bisin (Eds.), *Handbook of Social Economics. Volume 1A*. North Holland: Elsevier.
- Farbmacher, H., R. Guber, and J. Vikström (2018). Increasing the credibility of the twin birth instrument. *Journal of Applied Econometrics* 33(3), 457–472.
- Felfe, C. (2012a). The motherhood wage gap: What about job amenities? *Labour Economics* 19(1), 59–67.
- Felfe, C. (2012b). The willingness to pay for job amenities: Evidence from mothers’ return to work. *Industrial & Labor Relations Review* 65(2), 427–454.
- Fernandez, R. M. and S. Campero (2017). Gender sorting and the glass ceiling in high-tech firms. *ILR Review* 70(1), 73–104.
- 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.
- Fernández-Kranz, D. and N. Rodríguez-Planas (2013). Can Parents’ Right to Work Part-Time Hurt Childbearing-Aged Women? A Natural Experiment with Administrative Data. *IZA Discussion Paper No. 7509*.
- Filer, R. K. (1985). Male-female wage differences: The importance of compensating differentials. *Industrial & Labor Relations Review* 38(3), 426–437.
- Fiske, S. T., A. J. Cuddy, P. Glick, and J. Xu (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology* 82(6), 878.
- Fitzenberger, B., K. Sommerfeld, and S. Steffes (2013). Causal effects on employment after first birth — A dynamic treatment approach. *Labour Economics* 25, 49–62.
- Flabbi, L. and A. Moro (2012). The effect of job flexibility on female labor market outcomes: Estimates from a search and bargaining model. *Journal of Econometrics* 168(1), 81–95.

- Fortin, N. M., B. Bell, and M. Böhm (2017). Top earnings inequality and the gender pay gap: Canada, Sweden, and the United Kingdom. *Labour Economics* 47, 107–123.
- Fouarge, D., A. Manzonni, R. Muffels, and R. Luijkx (2010). Childbirth and cohort effects on mothers' labour supply: a comparative study using life history data for Germany, the Netherlands and Great Britain. *Work, Employment & Society* 24(3), 487–507.
- Fouarge, D. and R. Muffels (2009). Working part-time in the British, German and Dutch labour market: Scarring for the wage career? *Schmollers Jahrbuch* 129(2), 217–226.
- Francesconi, M. and M. Parey (2018). Early gender gaps among university graduates. *European Economic Review* 109, 63–82.
- Francesconi, M., H. Rainer, and W. Van Der Klaauw (2009). The Effects of In-Work Benefit Reform in Britain on Couples: Theory and Evidence. *The Economic Journal* 119(535), F66–F100.
- 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.
- Fransen, E., J. Plantenga, and J. D. Vlasblom (2012). Why do women still earn less than men? Decomposing the Dutch gender pay gap, 1996–2006. *Applied Economics* 44(33), 4343–4354.
- Fraser, N. (1994). After the family wage: Gender equity and the welfare state. *Political Theory* 22(4), 591–618.
- Freier, R., M. Schumann, and T. Siedler (2015). The earnings returns to graduating with honors — evidence from law graduates. *Labour Economics* 34, 39–50.
- Fuegen, K., M. Biernat, E. Haines, and K. Deaux (2004). Mothers and fathers in the workplace: how gender and parental status influence judgments of job-related competence. *Journal of Social Issues* 60(4), 737–754.
- Fuller, S. (2017). Segregation across Workplaces and the Motherhood Wage Gap: Why Do Mothers Work in Low-Wage Establishments? *Social Forces* 96(4), 1443–1476.
- Fuller, S. and C. E. Hirsh (2018). “Family-Friendly” Jobs and Motherhood Pay Penalties: The Impact of Flexible Work Arrangements Across the Educational Spectrum. *Work and Occupations (OnlineFirst)*.
- Gallie, D., M. Gebel, J. Giesecke, K. Halldén, P. Van der Meer, and R. Wielers (2016). Quality of work and job satisfaction: Comparing female part-time work in four European countries. *International Review of Sociology* 26(3), 457–481.

- 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.
- García-Manglano, J. (2015). Opting Out and Leaning In: The Life Course Employment Profiles of Early Baby Boom Women in the United States. *Demography* 52(6), 1961–1993.
- Gathmann, C. and U. Schönberg (2010). How general is human capital? A task-based approach. *Journal of Labor Economics* 28(1), 1–49.
- Gayle, G.-L. and L. Golan (2012). Estimating a dynamic adverse-selection model: Labour-force experience and the changing gender earnings gap 1968–1997. *The Review of Economic Studies* 79(1), 227–267.
- 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.
- Gibbons, R. and M. Waldman (1999). Careers in organizations: Theory and evidence. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, pp. 2373–2437. Amsterdam: Elsevier.
- Glauber, R. (2012). Women's Work and Working Conditions: Are Mothers Compensated for Lost Wages? *Work and Occupations* 39(2), 115–138.
- 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>.
- Goldin, C. (1990). *Understanding the Gender Gap*. New York: Oxford University Press.
- Goldin, C. (2006). The Quiet Revolution That Transformed Women's Employment, Education, and Family. *American Economic Review* 96(2), 1–21.
- Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review* 104(4), 1091–1119.

- Goldin, C., S. P. Kerr, C. Olivetti, and E. Barth (2017). The expanding gender earnings gap: evidence from the LEHD-2000 Census. *American Economic Review* 107(5), 110–14.
- Goldin, C. and J. Mitchell (2017). The new life cycle of women’s employment: Disappearing humps, sagging middles, expanding tops. *The Journal of Economic Perspectives* 31(1), 161–182.
- González, M. J., C. Cortina, and J. Rodríguez (2019). The role of gender stereotypes in hiring: A field experiment. *European Sociological Review* 35(2), 187–204.
- Görlich, D. and A. De Grip (2009). Human capital depreciation during hometime. *Oxford Economic Papers* 61, i98–i121.
- Gornick, J. C. and M. K. Meyers (2003). *Families that work: Policies for reconciling parenthood and employment*. New York, NY: Russell Sage Foundation.
- Gough, M. and M. Noonan (2013). A review of the motherhood wage penalty in the United States. *Sociology Compass* 7(4), 328–342.
- Greenland, S., J. Pearl, and J. M. Robins (1999). Causal diagrams for epidemiologic research. *Epidemiology* 10(1), 37–48.
- Gregg, P., M. Gutiérrez-Domènech, and J. Waldfogel (2007). The employment of married mothers in Great Britain, 1974–2000. *Economica* 74(296), 842–864.
- Grinza, E., F. Devicienti, M. Rossi, and D. Vannoni (2017). How Entry into Parenthood Shapes Gender Role Attitudes: New Evidence from Longitudinal UK Data. *IZA DP No. 11088*.
- Gronau, R. (1974). Wage comparisons—A selectivity bias. *Journal of Political Economy* 82(6), 1119–1143.
- Grunow, D., F. Schulz, and H.-P. Blossfeld (2012). What determines change in the division of housework over the course of marriage? *International Sociology* 27(3), 289–307.
- Gupta, N. D. and N. Smith (2002). Children and career interruptions: the family gap in Denmark. *Economica* 69(276), 609–629.
- Gupta, N. D., N. Smith, and M. Verner (2008). The impact of Nordic countries’ family friendly policies on employment, wages, and children. *Review of Economics of the Household* 6(1), 65–89.
- Gustafsson, S. (2001). Optimal age at motherhood. Theoretical and empirical considerations on postponement of maternity in Europe. *Journal of Population Economics* 14(2), 225–247.

- 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.
- Hakim, C. (2000). *Work-Lifestyle Choices in the 21st Century: Preference Theory*. Oxford: Oxford University Press.
- Hakim, C. (2002). Lifestyle preferences as determinants of women’s differentiated labor market careers. *Work and Occupations* 29(4), 428–459.
- Hardoy, I. and P. Schøne (2006). The Part-Time Wage Gap in Norway: How Large is It Really? *British Journal of Industrial Relations* 44(2), 263–282.
- Hardoy, I., P. Schøne, and K. M. Østbakken (2017). Children and the gender gap in management. *Labour Economics* 47, 124–137.
- 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.
- Hart, R. K. (2015). Earnings and first birth probability among Norwegian men and women 1995-2010. *Demographic Research* 33, 1067–1104.
- Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica* 47(1), 153–161.
- Heckman, J. J., R. J. LaLonde, and J. A. Smith (1999). The economics and econometrics of active labor market programs. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics*, pp. 1865–2097. Amsterdam, North Holland: Elsevier Science.
- Hegewisch, A. (2005). Individual Working Time Rights in Germany and the UK: How a Little Law Can Go a Long Way. In A. Hegewisch (Ed.), *Working Time for Working Families: Europe and the United States*. Washington DC: Friedrich-Ebert-Stiftung.
- Heilman, M. E. and T. G. Okimoto (2008). Motherhood: a potential source of bias in employment decisions. *Journal of Applied Psychology* 93(1), 189.
- Hellerstein, J. K., D. Neumark, and K. R. Troske (1999). Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations. *Journal of Labor Economics* 17(3), 409–446.
- Hernán, M. A. (2018). The C-word: Scientific euphemisms do not improve causal inference from observational data. *American Journal of Public Health* 108(5), 616–619.

- Hernán, M. A., S. Hernández-Díaz, and J. M. Robins (2004). A structural approach to selection bias. *Epidemiology* 15(5), 615–625.
- Hernan, M. A. and J. M. Robins (forthcoming). *Causal Inference*. Boca Raton: Chapman & Hall/CRC.
- Heywood, J. S., W. S. Siebert, and X. Wei (2007). The implicit wage costs of family friendly work practices. *Oxford Economic Papers* 59(2), 275–300.
- Hirsch, B., T. Schank, and C. Schnabel (2010). Differences in labor supply to monopsonistic firms and the gender pay gap: An empirical analysis using linked employer-employee data from Germany. *Journal of Labor Economics* 28(2), 291–330.
- Hobson, B. (Ed.) (2002). *Making Men into Fathers: Men, Masculinities, and the Social Politics of Fatherhood*. Cambridge, NY: Cambridge University Press.
- Hodges, M. J. and M. J. Budig (2010). Who gets the daddy bonus? Organizational hegemonic masculinity and the impact of fatherhood on earnings. *Gender & Society* 24(6), 717–745.
- Holland, J. A. (2017). The timing of marriage vis-à-vis coresidence and childbearing in Europe and the United States. *Demographic Research* 36, 609–626.
- Hondralis, I. (2017). Does Maternity Leave Pay Off? Evidence from a Recent Reform in Australia. *Social Politics: International Studies in Gender, State & Society* 24(1), 29–54.
- Hooker, H., F. Neathey, J. Casebourne, and M. Munro (2011). *The third work-life balance employee survey: Main findings (revised edition)*. Brighton: Institute for Employment Studies.
- Hotz, V. J., P. Johansson, and A. Karimi (2017). Parenthood, family friendly firms, and the gender gaps in early work careers. *NBER Working Paper 24173*.
- Hotz, V. J., C. H. Mullin, and S. G. Sanders (1997). Bounding causal effects using data from a contaminated natural experiment: Analysing the effects of teenage childbearing. *The Review of Economic Studies* 64(4), 575–603.
- 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*.
- Huerta, M. C., W. Adema, J. Baxter, W.-J. Han, M. Lausten, R. Lee, and J. Waldfogel (2014). Fathers' leave and fathers' involvement: evidence from four OECD countries. *European Journal of Social Security* 16(4), 308–346.
- 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.
- Ioannidis, J. P., T. D. Stanley, and H. Doucouliagos (2017). The power of bias in economics research. *The Economic Journal* 127(605), F236–F265.
- Iversen, T. and F. M. Rosenbluth (2010). *Women, Work, and Politics: The Political Economy of Gender Inequality*. New Haven & London: Yale University Press.
- Jalovaara, M., G. Neyer, G. Andersson, J. Dahlberg, L. Dommermuth, P. Fallesen, and T. Lappegård (2018). Education, gender, and cohort fertility in the nordic countries. *European Journal of Population*, 1–24.
- Jee, E., J. Misra, and M. Murray-Close (2018). Motherhood penalties in the us, 1986–2014. *Journal of Marriage and Family*, DOI:10.1111/jomf.12543.
- 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.
- Joshi, H. (2002). Production, reproduction, and education: Women, children, and work in a British perspective. *Population and Development Review* 28(3), 445–474.
- Kahn, J. R., J. García-Manglano, and S. M. Bianchi (2014). The Motherhood Penalty at Midlife: Long-Term Effects of Children on Women’s Careers. *Journal of Marriage and Family* 76(1), 56–72.
- Kalmijn, M., A. Loeve, and D. Manting (2007). Income dynamics in couples and the dissolution of marriage and cohabitation. *Demography* 44(1), 159–179.
- Kan, M. Y. (2007). Work Orientation and Wives’ Employment Careers An Evaluation of Hakim’s Preference Theory. *Work and Occupations* 34(4), 430–462.
- Karu, M. and D.-G. Tremblay (2018). Fathers on parental leave: an analysis of rights and take-up in 29 countries. *Community, Work & Family* 21(3), 344–362.
- Killewald, A. (2013). A reconsideration of the fatherhood premium: marriage, coresidence, biology, and fathers’ wages. *American Sociological Review* 78(1), 96–116.
- Killewald, A. and M. Gough (2013). Does specialization explain marriage penalties and premiums? *American Sociological Review* 78(3), 407–502.
- Killewald, A. and I. Lundberg (2017). New Evidence against a Causal Marriage Wage Premium. *Demography* 54(3), 1007–1028.
- King, G., R. O. Keohane, and S. Verba (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton University Press.

- Kleven, H. J., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2019). Child Penalties Across Countries: Evidence and Explanations. *Unpublished manuscript* (https://www.henrikkleven.com/uploads/3/7/3/1/37310663/klevenetal_aea-pp_2019.pdf).
- Kleven, H. J., C. Landais, and J. E. Sogaard (2017). Children and gender inequality: Evidence from Denmark. *Working Paper*, available at <http://econ.lse.ac.uk/staff/clandais/cgi--bin/Articles/Gender.pdf>.
- Kleven, H. J., C. Landais, and J. E. Sogaard (2018). Children and gender inequality: Evidence from Denmark. *NBER Working Paper No. 24219*.
- Kluge, 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.
- Kluge, 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.
- Knight, C. R. and M. C. Brinton (2017). One egalitarianism or several? Two decades of gender-role attitude change in Europe. *American Journal of Sociology* 122(5), 1485–1532.
- Koch, A. J., S. D. D'Mello, and P. R. Sackett (2015). A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *Journal of Applied Psychology* 100(1), 128.
- Kolinsky, E. (1989). *Women in West Germany: life, work, and politics*. Oxford, UK: Berg.
- Korenman, S. and D. Neumark (1991). Does marriage really make men more productive? *Journal of Human Resources* 26(2), 282–307.
- Koslowski, A., S. Blum, and M. P. (2016). International Review of Leave Policies and Research 2016. Available at: http://www.leavenetwork.org/lp_and_r_reports/.
- 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.
- Kravdal, Ø. and R. R. Rindfuss (2008). Changing relationships between education and fertility: A study of women and men born 1940 to 1964. *American Sociological Review* 73(5), 854–873.
- Kreyenfeld, M. and D. Konietzka (2017). *Childlessness in Europe: Contexts, causes, and consequences*. Springer.
- Krueger, A. B. (2017). Where have all the workers gone?: An inquiry into the decline of the us labor force participation rate. *Brookings Papers on Economic Activity* 2017(2), 1–87.

- Kühhirt, M. (2012). Childbirth and the long-term division of labour within couples: how do substitution, bargaining power, and norms affect parents' time allocation in West Germany? *European Sociological Review* 28(5), 565–582.
- 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.
- Kunze, A. (2005). The evolution of the gender wage gap. *Labour Economics* 12(1), 73–97.
- Kunze, A. (2014). Are All of the Good Men Fathers? The Effect of Having Children on Earnings. *IZA Discussion Paper No. 8113*.
- Kunze, A. (2015). The Family Gap in Career Progression. *Research in Labor Economics* 41, 115–142.
- Kunze, A. (2018). The Gender Wage Gap in Developed Countries. In A. L. M. H. S. D. Averett, S. L. (Ed.), *The Oxford Handbook of Women and the Economy*. Oxford: Oxford University Press.
- Kunze, A. and K. R. Troske (2012). Life-cycle patterns in male/female differences in job search. *Labour Economics* 19(2), 176–185.
- 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*.
- Lagakos, D., B. Moll, T. Porzio, N. Qian, and T. Schoellman (2018). Life cycle wage growth across countries. *Journal of Political Economy* 126(2), 797–849.
- Lalive, R., A. Schlosser, A. Steinhauer, and J. Zweimüller (2013). Parental leave and mothers' careers: The relative importance of job protection and cash benefits. *Review of Economic Studies* 81(1), 219–265.
- Lalive, R. and J. Zweimüller (2009). How does parental leave affect fertility and return to work? Evidence from two natural experiments. *The Quarterly Journal of Economics* 124(3), 1363–1402.
- Lechner, M., N. Rodriguez-Planas, and D. Fernández Kranz (2016). Difference-in-difference estimation by FE and OLS when there is panel non-response. *Journal of Applied Statistics* 43(11), 2044–2052.
- Leopold, T., J. Skopek, and F. Schulz (2018). Gender Convergence in Housework Time: A Life Course and Cohort Perspective. *Sociological Science* 5, 281–303.
- Lepinteur, A., S. Fleche, and N. Powdthavee (2016). My Baby Takes the Morning Train: Gender Identity, Fairness, and Relative Labor Supply Within Households. *IZA Discussion Paper No. 10382*.

- Lersch, P. M., M. Jacob, and K. Hank (2017). Parenthood, Gender, and Personal Wealth. *European Sociological Review* 33(3), 410.
- Lesner, R. V. (2018). Testing for Statistical Discrimination based on Gender. *Labour* 32(2), 141–181.
- Lewis, J. (1992). Gender and the development of welfare regimes. *Journal of European Social Policy* 2(3), 159–173.
- Lewis, J. (2002). The problem of fathers: policy and behavior in Britain. In B. Hobson (Ed.), *Making Men into Fathers: Men, Masculinities and the Social Politics of Fatherhood*. Cambridge: Cambridge University Press.
- Lewis, J. and M. Campbell (2007). UK work/family balance policies and gender equality, 1997–2005. *Social Politics* 14(1), 4–30.
- Livermore, T., J. Rodgers, and P. Siminski (2011). The effect of motherhood on wages and wage growth: evidence for australia. *Economic Record* 87, 80–91.
- Looze, J. (2014). Young Women’s Job Mobility: The Influence of Motherhood Status and Education. *Journal of Marriage and Family* 76(4), 693–709.
- 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.
- Lucifora, C., D. Meurs, and E. Villar (2017). Children, earnings and careers in an internal labor market. *Working Paper*, available at <http://www.aiel.it/cms/cms--files/submission/all20170831111248.pdf>.
- Ludwig, V. (2019). The linear Fixed-Effects model with Individual-Specific Slopes (FEIS). *Stata Journal*, forthcoming.
- Ludwig, V. and J. Brüderl (2018). Is There a Male Marital Wage Premium? New Evidence from the United States. *American Sociological Review OnlineFirst*, <https://doi.org/10.1177/0003122418784909>.
- Lundberg, S. (2005). Men and islands: dealing with the family in empirical labor economics. *Labour Economics* 12(4), 591–612.
- Lundberg, S. and E. Rose (2000). Parenthood and the earnings of married men and women. *Labour Economics* 7(6), 689–710.
- Lundborg, P., E. Plug, and A. W. Rasmussen (2017). Can women have children and a career? IV evidence from IVF treatments. *American Economic Review* 107(6), 1611–37.

- Lyngstad, T. H. and M. Jalovaara (2010). A review of the antecedents of union dissolution. *Demographic Research* 23, 257–292.
- Machado, C. (2017). Unobserved selection heterogeneity and the gender wage gap. *Journal of Applied Econometrics* 32(7), 1348–1366.
- 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.
- Manning, A. and B. Petrongolo (2008). The Part-Time Pay Penalty for Women in Britain. *The Economic Journal* 118(526), F28–F51.
- Manning, A. and H. Robinson (2004). Something in the way she moves: a fresh look at an old gap. *Oxford Economic Papers* 56(2), 169–188.
- Manning, A. and J. Swaffield (2008). The gender gap in early-career wage growth. *The Economic Journal* 118(530), 983–1024.
- Mark, N. P., L. Smith-Lovin, and C. L. Ridgeway (2009). Why do nominal characteristics acquire status value? A minimal explanation for status construction. *American Journal of Sociology* 115(3), 832–862.
- Martino, E. M. (2017). The Labor Cost of Motherhood and the Length of Career Break around Childbirth. *Working Paper*, available at <http://www.aiel.it/cms/cms--files/submission/all20170901180802.pdf>.
- Mas, A. and A. Pallais (2017). Valuing alternative work arrangements. *American Economic Review* 107(12), 3722–59.
- Matteazzi, E., A. Pailhé, and A. Solaz (2014). Part-time wage penalties for women in prime age: A matter of selection or segregation? Evidence from four European countries. *Industrial & Labor Relations Review* 67(3), 955–985.
- Maume, D. J. (1999). Glass ceilings and glass escalators occupational segregation and race and sex differences in managerial promotions. *Work and Occupations* 26(4), 483–509.
- McCall, L. and C. Percheski (2010). Income inequality: New trends and research directions. *Annual Review of Sociology* 36, 329–347.

- McLaughlin, H., C. Uggen, and A. Blackstone (2012). Sexual harassment, workplace authority, and the paradox of power. *American Sociological Review* 77(4), 625–647.
- McMunn, A., R. Lacey, D. Worts, P. McDonough, M. Stafford, C. Booker, M. Kumari, and A. Sacker (2015). De-standardization and gender convergence in work-family life courses in Great Britain: A multi-channel sequence analysis. *Advances in Life Course Research* 26, 60–75.
- Miller, A. R. (2011). The effects of motherhood timing on career path. *Journal of Population Economics* 24(3), 1071–1100.
- Mills, M., R. R. Rindfuss, P. McDonald, E. Te Velde, et al. (2011). Why do people postpone parenthood? Reasons and social policy incentives. *Human Reproduction Update* 17(6), 848–860.
- Mincer, J. and S. Polachek (1974). Family investments in human capital: Earnings of women. *The Journal of Political Economy* 82(2), S76–S108.
- Mincy, R., J. Hill, and M. Sinkewicz (2009). Marriage: Cause or mere indicator of future earnings growth? *Journal of Policy Analysis and Management* 28(3), 417–439.
- Moen, P. (2011). From ‘work–family’ to the ‘gendered life course’ and ‘fit’: Five challenges to the field. *Community, Work & Family* 14(1), 81–96.
- Mooi-Reci, I. and R. Dekker (2015). Fixed-Term Contracts: Short-Term Blessings or Long-Term Scars? Empirical Findings from the Netherlands 1980–2000. *British Journal of Industrial Relations* 53(1), 112–135.
- Morgan, S. L. and C. Winship (2007). *Counterfactuals and causal analysis: Methods and principles for social research*. Cambridge: Harvard University Press.
- Mulligan, C. B. and Y. Rubinstein (2008). Selection, investment, and women’s relative wages over time. *The Quarterly Journal of Economics* 123(3), 1061–1110.
- 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.
- Murphy, E. and D. Oesch (2015). The feminization of occupations and change in wages: A panel analysis of Britain, Germany, and Switzerland. *Social Forces* 94(3), 1221–1255.
- Napari, S. (2009). Gender differences in early-career wage growth. *Labour Economics* 16(2), 140–148.
- Neal, D. (2004). The measured black-white wage gap among women is too small. *Journal of Political Economy* 112(S1), S1–S28.

- Neuburger, J., D. Kuh, and H. Joshi (2011). Cross-cohort changes in gender pay differences in Britain from 1972 to 2004: accounting for selection into employment using wage imputation. *Longitudinal and Life Course Studies* 2(3), 260–285.
- Neumark, D. (2016). Experimental research on labor market discrimination. *NBER Working Paper 22022*.
- Nicoletti, C. and M. L. Tanturri (2008). Differences in delaying motherhood across European countries: Empirical evidence from the ECHP. *European Journal of Population/Revue européenne de Démographie* 24(2), 157–183.
- Nieuwenhuis, R. and L. C. Maldonado (2018). *The triple bind of single-parent families: resources, employment and policies to improve wellbeing*. Policy Press.
- Nosek, B. A., G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, et al. (2015). Promoting an open research culture. *Science* 348(6242), 1422–1425.
- OECD (2013a). *Closing the Gender Gap: Act Now*. Paris: Oecd.
- OECD (2013b). *OECD Employment Outlook*. Paris: Oecd.
- OECD (2014). *OECD Employment Outlook*. Paris: Oecd.
- OECD (2017a). *Education at a Glance 2017: OECD Social Indicators*. Paris: OECD.
- OECD (2017b). *The Pursuit of Gender Equality: An Uphill Battle*. Paris: OECD Publishing.
- Oesch, D., O. Lipps, and P. McDonald (2017). The wage penalty for motherhood: Evidence on discrimination from panel data and a survey experiment for Switzerland. *Demographic Research* 37, 1793–1824.
- 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.
- Olivetti, C. and B. Petrongolo (2017). The economic consequences of family policies: lessons from a century of legislation in high-income countries. *The Journal of Economic Perspectives* 31(1), 205–230.
- O’Neill, J. and S. Polachek (1993). Why the gender gap in wages narrowed in the 1980s. *Journal of Labor Economics* 11(1, Part 1), 205–228.
- ONS (2011). *Hours Worked in the Labour Market, 2011*. London: Office for National Statistics (ONS).
- O’Reilly, J., T. Nazio, and J. M. Roche (2014). Compromising conventions: attitudes of dissonance and indifference towards full-time maternal employment in Denmark, Spain, Poland and the UK. *Work, employment and society* 28(2), 168–188.

- Palier, B. and K. Thelen (2010). Institutionalizing dualism: Complementarities and change in France and Germany. *Politics & Society* 38(1), 119–148.
- Paul, M. (2016). Is There a Causal Effect of Working Part-Time on Current and Future Wages? *The Scandinavian Journal of Economics* 118(3), 494–523.
- Paull, G. (2008). Children and women’s hours of work. *The Economic Journal* 118(526), F8–F27.
- Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika* 82(4), 669–688.
- Pearl, J., M. Glymour, and N. P. Jewell (2016). *Causal inference in statistics: a primer*. John Wiley & Sons.
- Pedulla, D. S. (2016). Penalized or Protected? Gender and the Consequences of Nonstandard and Mismatched Employment Histories. *American Sociological Review* 81(2), 262–289.
- Pedulla, D. S. and S. Thébaud (2015). Can We Finish the Revolution? Gender, Work-Family Ideals, and Institutional Constraint. *American Sociological Review* 80(1), 116–139.
- Percheski, C. and C. Wildeman (2008). Becoming a Dad: Employment Trajectories of Married, Cohabiting, and Nonresident Fathers. *Social Science Quarterly* 89(2), 482–501.
- Perelli-Harris, B., M. Kreyenfeld, W. Sigle-Rushton, R. Keizer, T. Lappegård, A. Jasilioniene, C. Berghammer, and P. Di Giulio (2012). Changes in union status during the transition to parenthood in eleven European countries, 1970s to early 2000s. *Population Studies* 66(2), 167–182.
- Perelli-Harris, B., W. Sigle-Rushton, M. Kreyenfeld, T. Lappegård, R. Keizer, and C. Berghammer (2010). The educational gradient of childbearing within cohabitation in Europe. *Population and Development Review* 36(4), 775–801.
- Petersen, T., A. M. Penner, and G. Høgsnes (2010). The Within-Job Motherhood Wage Penalty in Norway, 1979–1996. *Journal of Marriage and Family* 72(5), 1274–1288.
- Petersen, T., A. M. Penner, and G. Høgsnes (2011). The male marital wage premium: Sorting vs. differential pay. *Industrial & Labor Relations Review* 64(2), 283–304.
- 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.
- Petersen, T. and I. Saporta (2004). The Opportunity Structure for Discrimination. *American Journal of Sociology* 109(4), 852–901.
- Petersen, T., I. Saporta, and M.-D. L. Seidel (2000). Offering a job: Meritocracy and social networks. *American Journal of Sociology* 106(3), 763–816.

- Petit, P. (2007). The effects of age and family constraints on gender hiring discrimination: A field experiment in the French financial sector. *Labour Economics* 14(3), 371–391.
- Pettit, B. and J. L. Hook (2009). *Gendered Tradeoffs: Women, Family, and Workplace Inequality in Twenty-One Countries*. Russell Sage Foundation.
- Phelps, E. S. (1972). The Statistical Theory of Racism and Sexism. *American Economic Review* 62(4), 659–661.
- Picchio, M. and J. C. van Ours (2016). Gender and the effect of working hours on firm-sponsored training. *Journal of Economic Behavior & Organization* 125, 192–211.
- Pinkston, J. C. (2006). A test of screening discrimination with employer learning. *ILR Review* 59(2), 267–284.
- Polachek, S. W. (1981). Occupational self-selection: A human capital approach to sex differences in occupational structure. *The Review of Economics and Statistics* 63(1), 60–69.
- Polavieja, J. G. (2012). Socially Embedded Investments: Explaining Gender Differences in Job-Specific Skills. *American Journal of Sociology* 118(3), 592–634.
- Polavieja, J. G. and L. Platt (2014). Nurse or mechanic? The role of parental socialization and children’s personality in the formation of sex-typed occupational aspirations. *Social Forces* 93(1), 31–61.
- Pollmann-Schult, M. (2011). Marriage and earnings: why do married men earn more than single men? *European Sociological Review* 27(2), 147–163.
- Pollmann-Schult, M. (2016). What mothers want: The impact of structural and cultural factors on mothers’ preferred working hours in Western Europe. *Advances in Life Course Research* 29, 16–25.
- Pollmann-Schult, M. and J. Reynolds (2017). The Work and Wishes of Fathers: Actual and Preferred Work Hours among German Fathers. *European Sociological Review* 33(6), 823–838.
- 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.
- Pronzato, C. (2011). *British Household Panel Survey Consolidated Marital, Cohabitation and Fertility Histories, 1991-2009. 3rd Edition*. University of Essex: Institute for Social and Economic Research. UK Data Service. SN: 5629, <http://doi.org/10.5255/UKDA-SN-5629-1>.

- Protsch, P. and H. Solga (2015). How employers use signals of cognitive and noncognitive skills at labour market entry. insights from field experiments. *European Sociological Review* 31(5), 521–532.
- Puhani, P. (2000). The Heckman correction for sample selection and its critique. *Journal of Economic Surveys* 14(1), 53–68.
- 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.
- Reed, W. R. and K. Harford (1989). The marriage premium and compensating wage differentials. *Journal of Population Economics* 2(4), 237–265.
- Rege, M. and I. F. Solli (2013). The impact of paternity leave on fathers' future earnings. *Demography* 50(6), 2255–2277.
- Riach, P. A. and J. Rich (2002). Field experiments of discrimination in the market place. *The Economic Journal* 112(483), F480–F518.
- Riach, P. A. and J. Rich (2006). An experimental investigation of sexual discrimination in hiring in the English labor market. *Advances in Economic Analysis & Policy* 5(2).
- Rich, J. (2014). What do field experiments of discrimination in markets tell us? A meta analysis of studies conducted since 2000. *IZA Discussion paper No. 8584*.
- Ridgeway, C. L. (2011). *Framed by Gender: How Gender Inequality Persists in the Modern World*. Oxford: Oxford University Press.
- Ridgeway, C. L., K. Backor, Y. E. Li, J. E. Tinkler, and K. G. Erickson (2009). How easily does a social difference become a status distinction? Gender matters. *American Sociological Review* 74(1), 44–62.
- Ridgeway, C. L. and S. J. Correll (2004). Motherhood as a status characteristic. *Journal of Social Issues* 60(4), 683–700.
- Ridgeway, C. L. and S. J. Correll (2006). Consensus and the creation of status beliefs. *Social Forces* 85(1), 431–453.
- Rivera, L. A. and A. Tilcsik (2016). Class advantage, commitment penalty: The gendered effect of social class signals in an elite labor market. *American Sociological Review* 81(6), 1097–1131.
- Rønsen, M. and M. Sundström (2002). Family policy and after-birth employment among new mothers—A comparison of Finland, Norway and Sweden. *European Journal of Population/Revue européenne de démographie* 18(2), 121–152.

- Rosen, S. (1986). The Theory of Equalizing Differences. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics*. Amsterdam, North Holland: Elsevier Science.
- Rossin-Slater, M. (2017). Maternity and Family Leave Policy. *NBER Working Paper 23069*.
- Rubery, J. (2011). Towards a gendering of the labour market regulation debate. *Cambridge Journal of Economics* 35(6), 1103–1126.
- Ruhm, C. J. (1998). The economic consequences of parental leave mandates: Lessons from Europe. *The Quarterly Journal of Economics* 113(1), 285–317.
- Scherpenzeel, A. (2011). Data collection in a probability-based internet panel: how the LISS panel was built and how it can be used. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique* 109(1), 56–61.
- Schmelzer, P. and A. V. Ramos (2015). Varieties of wage mobility in early career in Europe. *European Sociological Review* 32(2), 175–188.
- Schober, P. S. (2013). The parenthood effect on gender inequality: Explaining the change in paid and domestic work when British couples become parents. *European Sociological Review* 29(1), 74–85.
- 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.
- Semykina, A. and J. M. Wooldridge (2010). Estimating panel data models in the presence of endogeneity and selection. *Journal of Econometrics* 157(2), 375–380.
- Shadish, W. R., T. D. Cook, and D. T. Campbell (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.
- Sigle-Rushton, W. and J. Waldfogel (2007). Motherhood and women's earnings in Anglo-American, Continental European, and Nordic countries. *Feminist Economics* 13(2), 55–91.
- Simonsen, M. and L. Skipper (2006). The costs of motherhood: an analysis using matching estimators. *Journal of Applied Econometrics* 21(7), 919–934.
- Simonsen, M. and L. Skipper (2012). The family gap in wages: What wombmates reveal. *Labour Economics* 19(1), 102–112.
- Sipila, J., K. Repo, and T. Rissanen (2010). *Cash for childcare: The consequences for caring mothers*. Edward Elgar Publishing.
- Smith, A. (1776). *An Inquiry Into the Nature and Causes of the Wealth of Nations* (1976 ed.). Chicago: The University of Chicago Press.
- Smith, H. L. (1986). *War and social change: British society in the Second World War*. Manchester, UK: Manchester University Press.

- Smith, N., S. Dex, J. D. Vlasblom, and T. Callan (2003). The effects of taxation on married women's labour supply across four countries. *Oxford Economic Papers* 55(3), 417–439.
- Smith, R. S. (1979). Compensating wage differentials and public policy: a review. *ILR Review* 32(3), 339–352.
- Smith Koslowski, A. (2011). Working Fathers in Europe: Earning and Caring. *European Sociological Review* 27(2), 230–245.
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics* 87(3), 355–374.
- 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>).
- Struffolino, E., M. Studer, and A. E. Fasang (2016). Gender, education, and family life courses in East and West Germany: Insights from new sequence analysis techniques. *Advances in Life Course Research* 29, 66–79.
- 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*.
- Tanaka, S. and J. Waldfogel (2007). Effects of parental leave and work hours on fathers' involvement with their babies: Evidence from the millennium cohort study. *Community, Work & Family* 10(4), 409–426.
- Taylor, M. F., J. Brice, N. Buck, and E. Prentice-Lane (2010). *British Household Panel Survey: User Manual Volume A*. Colchester: University of Essex.
- Thébaud, S. (2015). Status Beliefs and the Spirit of Capitalism: Accounting for Gender Biases in Entrepreneurship and Innovation. *Social Forces* 94(1), 61–86.
- Thévenon, O. (2011). Family policies in OECD countries: A comparative analysis. *Population and Development Review* 37(1), 57–87.
- Thévenon, O. and A. Solaz (2012). Labour Market Effects of Parental Leave Policies in OECD Countries. *OECD Social, Employment, and Migration Working Papers No. 141* (<http://dx.doi.org/10.1787/5k8xb6hw1wjf-en>).
- Townsend, N. W. (2002). *The package deal: Marriage, Work and Fatherhood in Men's Lives*. Philadelphia: Temple.

- 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.
- Trimarchi, A. and J. Van Bavel (2017). Education and the transition to fatherhood: The role of selection into union. *Demography* 54(1), 119–144.
- Triventi, M. (2013). The gender wage gap and its institutional context: a comparative analysis of European graduates. *Work, Employment & Society* 27(4), 563–580.
- University of Essex, Institute for Social and Economic Research (2018). Understanding Society: Waves 1-7, 2009-2016 and Harmonised BHPS: Waves 1-18, 1991-2009. 10th Edition. *UK Data Service SN: 6614*, <http://doi.org/10.5255/UKDA--SN--6614--11>.
- Van Bavel, J., C. Schwartz, and A. Esteve (2018). The Reversal of the Gender Gap in Education and its Consequences for Family Formation. *Annual Review of Sociology* 44, 341–360.
- van de Werfhorst, H. G. (2017). Gender Segregation across Fields of Study in Post-Secondary Education: Trends and Social Differentials. *European Sociological Review* 33(3), 449–464.
- Vargha, L., R. I. Gál, and M. O. Crosby-Nagy (2017). Household production and consumption over the life cycle: National Time Transfer Accounts in 14 European countries. *Demographic Research* 36, 905–944.
- Visser, J. (2002). The first part-time economy in the world: a model to be followed? *Journal of European Social Policy* 12(1), 23–42.
- Waldfogel, J. (1995). The price of motherhood: family status and women’s pay in a young British cohort. *Oxford Economic Papers* 47(4), 584–610.
- Waldfogel, J. (1997). The Effect of Children on Women’s Wages. *American Sociological Review* 62(2), 209–217.
- Waldfogel, J. (1998a). 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.
- Waldfogel, J. (1998b). Understanding the “family gap” in pay for women with children. *The Journal of Economic Perspectives* 16(3), 137–156.
- Wang, R. and A. Weiss (1998). Probation, layoffs, and wage-tenure profiles: A sorting explanation. *Labour Economics* 5(3), 359–383.
- Warren, T. (2015). Work–life balance/imbalance: the dominance of the middle class and the neglect of the working class. *The British journal of sociology* 66(4), 691–717.

- Westreich, D., J. K. Edwards, C. R. Lesko, S. R. Cole, and E. A. Stuart (2018). Target Validity and the Hierarchy of Study Designs. *American Journal of Epidemiology*, **10.1093/aje/kwy228**.
- Wielers, R. and D. Raven (2013). Part-time work and work norms in the Netherlands. *European Sociological Review* *29*(1), 105–113.
- Wilde, E. T., L. Batchelder, and D. T. Ellwood (2010). The mommy track divides: The impact of childbearing on wages of women of differing skill levels. *NBER Working Paper No. 16582*.
- Wilner, L. (2016). Worker-firm matching and the parenthood pay gap: Evidence from linked employer-employee data. *Journal of Population Economics* *29*(4), 991–1023.
- Wiswall, M. and B. Zafar (2017). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics* *133*(1), 457–507.
- Wooldridge, J. M. (1995). Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics* *68*(1), 115–132.
- Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT press.
- Wright, E. O., J. Baxter, and G. E. Birkelund (1995). The gender gap in workplace authority: A cross-national study. *American Sociological Review* *60*(3), 407–435.
- Yerkes, M. A. and J. Javornik (2018). Creating capabilities: Childcare policies in comparative perspective. *Journal of European Social Policy*, <https://doi.org/10.1177/0958928718808421>.
- Yip, C. M. and R. S.-K. Wong (2014). Gender-oriented statistical discrimination theory: Empirical evidence from the Hong Kong labor market. *Research in Social Stratification and Mobility* *37*, 43–59.
- Young, R. and D. R. Johnson (2015). Handling missing values in longitudinal panel data with multiple imputation. *Journal of Marriage and Family* *77*(1), 277–294.
- Yu, J. and Y. Xie (2018). Motherhood penalties and living arrangements in China. *Journal of Marriage and Family* *80*(5), 1067–1086.
- Yu, W.-h. and J. C.-L. Kuo (2017). The Motherhood Wage Penalty by Work Conditions: How Do Occupational Characteristics Hinder or Empower Mothers? *American Sociological Review* *82*(4), 744–769.
- 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 Childcare Availability and Mothers' Return-to-Work Behaviour in East and West Germany. *European Sociological Review* 33(5), 693–707.

Summary

That ‘parenthood will change your life’ is common sense. Less self-evident is how women and men grow unequal in labour markets as a consequence of parenthood. Once confined to ‘home-making’, women have made strides in labour markets, increasingly so across cohorts and continuously along their lifecycle. Gender gaps in pay have shrank and women and men distribute more evenly across jobs, but only up to a point. Parenthood, in particular, divides today the careers of women and men. A *family gap* has emerged in labour markets: Women pay economic and career prices for motherhood, while the career progression of men marches on come fatherhood.

Gender inequality in paid work persists despite institutional change aimed at mitigating it or curbing it altogether. Labour market and welfare institutions have variously departed from the *family wage* model once supporting male breadwinning through secure, well-paid employment, surrounded by social protections. In particular, the United Kingdom, Germany, and the Netherlands moved away from this family wage model in recent decades. This move has been marked by two main transformations, namely the expansion of family leave rights and the flexibilisation of employment relationships. Yet, beyond commonalities, policy trajectories have diverged in the three countries and so have their consequences for the family gap and gender inequality more broadly.

Hence, I ask here how the family gap has shaped in the midst of akin and yet distinct changes in the labour market and welfare institutions formerly devoted to the family wage principle in the UK, Germany, and the Netherlands. By highlighting progress and stall in the ways these three countries came to modify their male breadwinner order, my main tenet is that policies aimed at women and families are not by default women- or family-friendly. The family gap, I argue, is often the unintended or perverse by-product of gradual and selective institutional change. Throughout, this overarching question is addressed drawing on panel data analysis, quasi-experimental designs, and experimental data.

In Chapter 1, I address whether fatherhood causally affects the wages of men in the modified male breadwinner societies of Germany and the UK, relying on long-running household panel data. Not unlike previous research for Nordic countries but differently

from US-based studies, I cannot detect (substantial) wage responses to the transition to fatherhood in both countries, on average as well as across cohorts. Positive selection into fatherhood on the basis of prior wage levels and prior wage growth seemingly accounts for much of the apparent premium. Overall, these results question whether fatherhood “bonuses” contribute to family gaps and suggest to (re)cast the focus of policies to motherhood “penalties” instead.

In Chapters 2 and 3, I turn to the effects of specific policies combining panel data and a difference-in-differences design. Chapter 2 assesses how changes to parental leave mandates have shaped the motherhood wage penalty in Germany over recent decades. Overall stability in the motherhood wage penalty is found comparing before and after the early parental leave expansion, culminated in 1992. The new benefit installed in 2007, combining shortened and earnings-related benefit receipt and a “daddy bonus”, seems to have ameliorated instead the average wage penalty for German mothers. Such patterns are examined in the light of human capital, effort, and signalling mechanisms triggered by each reform.

Chapter 3 turns from family leave rights to the flexibilisation of employment relationships, with a focus on working-time flexibility and the labour market outcomes of parents in the UK. I analyse a 2003 reform granting parents of young children the right to request changes to their working-time arrangements with their current employer. Coherent with a compensating differential story, eligible mothers of young children are the only group that ends up trading off wage losses for shorter work schedules, while also feeling more satisfied with working-time arrangements, in the aftermath of the reform.

A survey experiment involving Dutch employers is then the subject of Chapter 4. Hiring decisions in a female-typical and a male-typical job are investigated. The experimental design allows to single out differential treatment of female and male job applicants depending on their parental status, testing the implications of an economic and a socio-psychological model of employer discrimination. Data in this study support neither, as little differences in likelihood of hire or salary offers are found across candidates and job types. Strikingly though, and in contrast with previous US-based research, such results hold even if Dutch employers still expect mothers to be less committed to their job and

to ask for the shortest working week available.

All in all, I attempt here to bridge some of the gaps in the extant literature. I highlight links between institutional change and the well-established family gap in labour markets. Motherhood penalties are key to persistent gender gaps in labour markets, but the effects that new ‘family-friendly’ policies have on such gaps may vary over time and context. Based on the findings of this dissertation, no “welfare model” or policy recipe (or laundry list even) is identified and proposed as a panacea to counteract the family gap. Rather, I more modestly provide an appraisal of how policy effects are subject to the logic of *trial and error*, much like best practices in the science that aims at studying such policy effects. In doing so, I lay down a – hopefully transparent – causal approach aimed at shedding light on the assumptions needed to uncover whether and how parenthood and policies contribute, to this day, to gendered economic inequality.

Acknowledgements

To thank someone requires quite the leap of faith. By default we are drawn to counterfactual statements such as ‘If not for you, this wouldn’t have been possible!’. But, as with all counterfactuals, we simply do not know that: Without *you*, we might have fared worse, yes, but perhaps also better, or just differently. And to state that things could not have gone in any other way than the one we just experienced, well, that is just self-evident.

Perhaps then acknowledging who made this dissertation possible is the trickiest attempt at causal inference the unfortunate reader will find in these pages. Or perhaps where logic cannot suffice, respect, empathy, and love can. And so, if I have ever deserved even a tiny bit of your respect, your empathy, or your love during these years, thank *you*.

Across three countries, from Trento to Tilburg, Bath, and now Rotterdam, thanks for the times in which you have mentored me without preconceptions, persuaded me based on evidence and rigour, and lead by example with your kindness. Thanks to all those who have asked questions to me or cared to answer mine, to the students I learned from more than taught to, and to those who kept me passionate.

Thanks, most of all, to those who have listened to me or wanted me, of all people, to listen to them. Especially if in front of a (good) beer. Terrified by the prospect of leaving someone out – or, just as much, of including each and every one of you – I will just say thank you once more. Chances are I have already thanked you, or will do shortly, in person, the way I prefer it to be – although I might not be good at it.

A final word. As established once and for all by the film trilogy *Back to the Future*, there are some people without whom this truly would not have been possible. Thanks to my mum and dad. And, finally, to my brother, my inspiration.

About the author

Gabriele Mari (born in Mirandola, Italy, 1990) holds a Master's degree in Sociology and Social Research from the University of Trento and a MSc degree in Sociology from Tilburg University. In 2014 he started his PhD at the University of Trento, where he also taught Statistics at the graduate and undergraduate level. His PhD was part of a *co-tutelle* programme with Tilburg University. He has taken part in several international conferences and in summer schools in statistics at the University of Essex and at Universitat Pompeu Fabra (UPF, Barcelona). He was also visiting researcher at the ERC project NEWFAMSTRAT, University of Bath. His research interests centre around gender disparities in labour markets, social policies, parenting, and child development. And everything causal inference. He is now a postdoctoral researcher at Erasmus University Rotterdam.