

Jessen, Jonas; Jessen, Robin; Kluve, Jochen

Article — Accepted Manuscript (Postprint)

Punishing potential mothers? Evidence for statistical employer discrimination from a natural experiment

Labour Economics

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Jessen, Jonas; Jessen, Robin; Kluve, Jochen (2019) : Punishing potential mothers? Evidence for statistical employer discrimination from a natural experiment, Labour Economics, ISSN 0927-5371, Elsevier, Amsterdam, Vol. 59, pp. 164-172, <https://doi.org/10.1016/j.labeco.2019.04.002>

This Version is available at:

<https://hdl.handle.net/10419/233119>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Punishing Potential Mothers? Evidence for Statistical Employer Discrimination From a Natural Experiment*

Jonas Jessen[†]
Robin Jessen[‡]
Jochen Kluge[§]

ACCEPTED MANUSCRIPT

2 April 2019

Abstract

Before 2006, large firms in Germany were obliged to pay for the generous maternity protection of female employees, such that firms' expected costs depended on employees' gender and age. From 2006 onward, all firms paid for maternity protection by contributing to the statutory health insurance system, where the contribution depends only on the number of employees and their wages and is thus independent of gender and age. This had been the regulation for small firms already before the reform. Using data from linked employer-employee administrative records, we provide evidence that the reform was followed by an increase in female relative wages within large firms. This reform effect provides evidence for statistical employer discrimination in the pre-2006 setup.

Keywords Maternity protection · Statistical discrimination · Natural experiment

JEL Classification J31 · J71

*Thanks to David Card, Mathias Huebener, seminar participants at UC Berkeley and DIW Berlin and conference participants at EALE 2018 Lyon and COMPIE 2018 Berlin for helpful feedback. We also thank two anonymous referees and the editor, Éric Maurin, for valuable comments and suggestions.

[†]DIW Berlin and Free University Berlin, jjessen@diw.de

[‡]RWI, robin.jessen@rwi-essen.de

[§]RWI and Humboldt University Berlin, jochen.kluve@hu-berlin.de

1 Introduction

Theory predicts that employers may discriminate statistically and pay female employees of child-bearing age lower wages than their male counterparts, *ceteris paribus*, if motherhood imposes costs on employers. This discriminatory behaviour will be more pronounced if firms face direct costs of motherhood as per law. We test this prediction using a natural experiment created by a reform of German maternity protection. Before the reform, large firms had to pay mothers' wage continuation around childbirth; this made their expected costs arising from maternity protection wage payments a direct function of their employees' probability to become mothers, i.e. effectively depending on gender and age of their workforce. The reform regulated that all wage continuation to mothers comes centrally from the social security system, and firms' contributions are merely a function of the number of employees and their wages, regardless of gender and age.

A driving mechanism behind the gender wage gap is the fact that, on average, mothers take substantially more time off work after childbirth than fathers. These career interruptions result in lower labour market experience and lower wages (Gangl and Ziefle 2009; Goldin 2014; Adda et al. 2017). However, a substantial pay gap exists even when controlling for observables such as labour market experience or tenure and an—admittedly substantially smaller—gap also exists pre-birth. One reason for this might be that maternity is not only punished *ex-post*, but also *ex-ante* via statistical discrimination of women of childbearing age (i.e. potential mothers). Employers in general bear some of the costs of motherhood: they need to find a replacement during leave, the accumulation of firm-specific human capital is disrupted, and existing skills of their employees deteriorate during leave. Also, mothers may not return to their job (full-time) afterwards. This can make profit-maximising companies reluctant to promote them or hire them in the first place. Under standard assumptions, disincentives to hiring potential mothers lead to adjustments along two margins: fewer potential mothers are hired, and those that are hired receive lower wages than men, *ceteris paribus*. In theory this effect is stronger if companies have to pay wage continuation directly. Using a natural experiment, we estimate to what extent this effect can be counter-acted if the state pays wage continuation centrally through the social security system.

Before 1 January 2006, large firms in Germany (more than 30 employees) were obliged to pay for the generous maternity protection—14 weeks of 100 per cent wage continuation around the date of delivery—of their female employees. From 2006 onward, each firm contributed to maternity protection through a contribution to the countrywide health insurance system, which is paid for every worker in the firm, irrespective of whether the worker is male or female (i.e. firms with an all-male workforce contribute the same as firms with an all-female workforce). This had been the regulation for small firms (≤ 30 workers) already before the reform. In its ruling declaring the previous regime as unconstitutional, the German Federal Constitutional Court stated it was unable to say with certainty "how large the probability is that due to this burden employers who have to pay maternity protection wage continuation decided not to employ women" (Bundesverfassungsgericht 2003a, Section 120). However, not only hiring decisions but promotions and, more generally, wage setting may be affected. We use comprehensive data—from annual linked employer-employee administrative records—covering over 1 million workers in more than 10,000 firms to estimate the wage effects of the reform.

Gender wage and employment differences have received extensive coverage in the literature. Despite signs of a narrowing over time these differences have persisted in all Western economies, see, e.g. Blau and Kahn (2003), Weichselbaumer and Winter-Ebmer (2005), Olivetti and Petrongolo (2008) and OECD (2019). Card et al. (2015) look at firm-specific pay premiums as a source of (gender) wage inequality and point out that if firms have some control over the wages offered,

relative wages of women will be influenced by both a potential sorting of women into higher or lower paying firms and on their relative bargaining power. In Germany, the share of females is higher in small firms, which on average pay lower wages. [Heinze and Wolf \(2010\)](#) confirm some selection of women into lower paying firms more generally, i.e. for firms of all sizes.

To our knowledge, no previous study links the gender wage gap to *statistical* discrimination of *potential* mothers. So far, the literature has focused on *actual* mothers.¹ However, not only *being* a mother but already the *possibility of becoming* a mother can have adverse effects in the labour market. This links the topic to the literature on statistical discrimination, which predicts wage differences between groups even in the absence of taste-based discrimination as in [Becker \(1957\)](#). The basic idea of statistical discrimination—originally developed by [Phelps \(1972\)](#)—is that profit-maximising employers have imperfect information about the productivity of potential employees and rationally use group statistics as proxies for unobserved exogenous characteristics (see [Fang and Moro 2010](#)).² In the case of statistical discrimination against potential mothers, firms do not know whether a specific woman will become a mother, but use the information that a share of women of childbearing age will have children (about 80% in Germany), which imposes additional expected costs on firms.³ Therefore the firm is expected to be willing to pay women of childbearing age a lower wage than men of same age. This is a case of statistical discrimination, because even women who will not have a child are paid a lower wage than men due to the use of group statistics for gender and age to determine expected costs due to childbearing.

While in Germany anti-discrimination laws regulate that employers are not allowed to ask (potential) employees about a current or planned pregnancy, it is an evident possibility that a woman of childbearing age will become a mother: the annual average probability to give birth for working women in the age bracket 24-35 is 5.5%, peaking at an average of 7% p.a. for women aged 30-32. Adverse labour market prospects for women can then arise without any taste-based discrimination, and, as [Phelps \(1972, p. 661\)](#) notes, "[d]iscrimination is no less damaging to its victims for being statistical."

In the empirical analysis, we implement a *difference-in-differences (DID) model* to capture the post-reform effect, as well as a *trend-break (TB) model* that distinguishes between general convergence (or divergence) in female and male wages at large firms over time and the divergence from this trend post-reform. The treatment effect would imply, and measure, the corresponding degree of statistical employer discrimination pre-2006.

Our results confirm the theoretical prediction: the DID estimates show a statistically significant increase in female wages relative to male wages of 1.1 per cent in the post-reform period. The magnitude of the effect is in line with a back-of-the-envelope calculation, which puts the expected effect at around 0.8 per cent. The trend-break model carves out this pattern in more detail and shows that the general time trend in the convergence of the gender wage gap in large firms is

¹A recent study by [Biewen and Seifert \(2018\)](#) is an exception: while not focusing on the gender wage gap, it quantifies the association of the probability of parenthood on career transitions for men and women in Germany, and finds a negative relationship between the contemporaneous probability to have a child and horizontal career transitions for women. Horizontal career transitions are defined as job changes in which the number of subordinates does not change by more than two. These transitions might thus still be associated with substantial wage increases.

²[Arrow \(1973\)](#) develops a similar model in which, in contrast, average group differences in characteristics are endogenous.

³Specifically, before the reform the costs per hour worked of a mother were higher for large firms, because the amount of hours worked in the year a woman had a child were reduced due to maternity protection, but firms' expenses (wage and wage continuation) were not reduced proportionally. In contrast, small firms did not have to pay wage continuation during maternity protection and therefore the expenses per hour worked are the same regardless of whether a woman has a child. See also Section 3.

not statistically significantly different from zero and the post-2006 trend estimates a total reform effect of 2.5 per cent over five years. Hence, the estimated effect confirms the hypothesis that the statistically discriminatory behaviour of large firms pre-reform reflects the costs for the firm arising from the regulation. Overall, the findings imply that policy makers should try to identify factors that could result in statistical discrimination against potential mothers. If the aim is to prevent negative labour market effects *ex-ante* and *ex-post* for (potential) mothers, it is worthwhile for the public to compensate firms for costs that occur asymmetrically to mothers only.

2 Institutional background

Mothers in Germany are entitled to paid maternity protection, which lasts from six weeks before the predicted date of childbirth until eight weeks after. During those 14 weeks expecting mothers are entitled to 100 per cent wage replacement. This paid maternity protection regulation is the focus of our analysis. In addition, several other regulations are in place that protect mother and child and the mother's employment relationship: these include, for instance, dismissal protection starting with the onset of the pregnancy, and maternity leave for up to three years post-childbirth with the option to return to the same job.

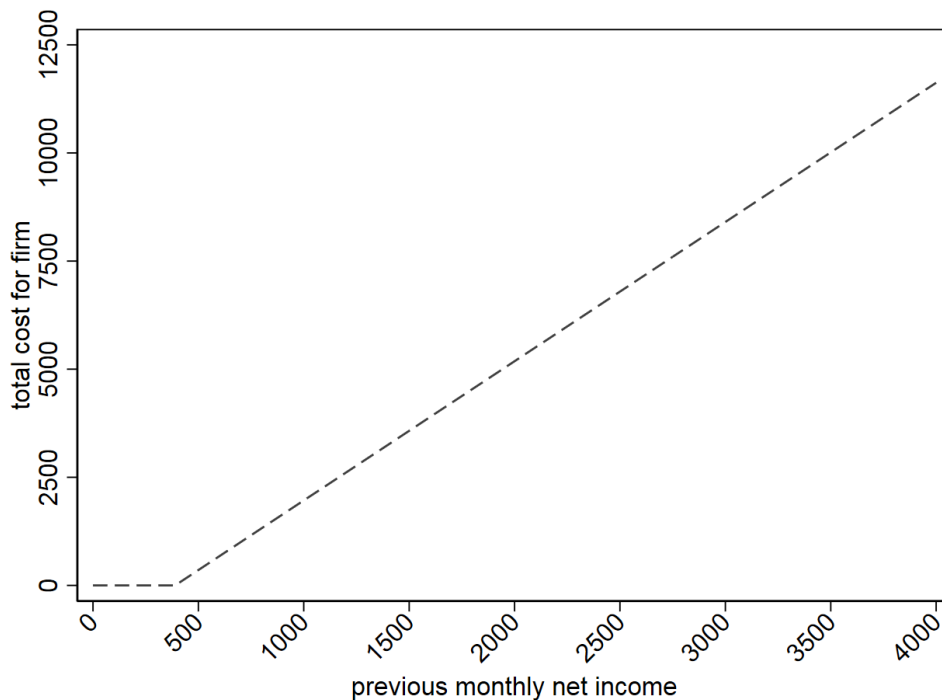
Before 2006, employers in Germany with more than 30 full-time-equivalent employees had to pay a substantial share of the wage replacements of mothers during maternity protection themselves:⁴ women who were insured by a statutory health insurance company received a fixed amount of 13 Euro per calendar day (roughly 400 Euro monthly) and additionally, if they were employed previously, the difference to their previous net earnings from their employer. In 2001, firms paid a total of about 1.48 billion Euro for maternity protection wage continuation ([Bundesverfassungsgericht 2003a](#)).⁵

Figure 1 displays the costs firms faced for the 14 weeks of maternity protection as a function of previous net earnings. From a monthly net income of about 400 Euro the costs are a linearly increasing function of the previous net wage. Note that the figure displays firms' direct, effective total costs for one female worker in maternity protection: the fixed share of 13 Euro per calendar day (a total of about 1,400 Euro for the 14 weeks) covered by the statutory health insurance is deducted, and the cost curve begins its upward sloping part at monthly earnings of about 400 Euro. As this 13 Euro daily contribution paid by insurers had remained unchanged since 1968, whereas wages had risen substantially, the share paid by large firms had increased strongly over time.

⁴The regulation originally set the threshold at 20 and granted the statutory health insurers the flexibility to increase that threshold from 20 to up to 30. We contacted the different regional entities of the largest statutory health insurer, AOK, which were responsible for executing the maternity protection payments. The vast majority of them set the limit to 30, therefore we use this effective cut-off in our analysis to ensure that all firms are correctly assigned. The relevant measure and threshold of full-time equivalent employees (FTE) is determined by the number of employees weighted by hours worked. A person working less than 10 hours counted 0.25 FTE, a person working 10 to 20 hours 0.5 FTE, 20-30 hours 0.75 FTE and a person working more than 30 hours counted as 1 FTE. In our data (see Section 4) a part-time variable indicates whether a person has worked more or less than 18 hours per week, and the FTE categories are approximated. We use the German Socio-Economic Panel (SOEP, [Goebel et al. 2018](#)) to impute the respective shares of workers falling into the working hour groups. We use information from 2003, in line with our specification of small and large firms (see below), and implement the imputation separately by gender.

⁵Note that individuals with private health insurance were not covered by this regulation and also not affected by the change in law. However, since only about 10 per cent of Germans are privately insured, and since our data do not contain civil servants (see Section 4), of which many have a private insurance, this is not a major issue. Concerning our estimates, we still identify lower bounds in absolute terms for individuals covered by public health insurance, as we cannot determine the type of insurance an individual had.

Figure 1: Firms' costs per maternity protection period



Notes: Costs of maternity protection imposed on firms for different levels of employee's pre-birth monthly net income in Euro.

For small firms, in contrast, statutory health insurance companies paid the entire wage continuation. These firms had to pay a social security contribution per employee to compensate for the costs, a pay-as-you-go system called *Umlage U2—Mutterschaft* (Contribution U2—Motherhood). Small firms had to contribute to *U2* per worker regardless of gender and age, a set-up explicitly designed to prevent adverse employment effects for women of childbearing age. In 2003, around 90 per cent of all firms were covered by *U2*, but this included only one-third of female and one-quarter of male employees.

Following a constitutional complaint of a large firm, on 18 November 2003 the German Federal Constitutional Court ruled the legislation as unconstitutional. The judges argued that, since it created a disincentive for large firms to hire women, it violated the constitutional principles of equal opportunities for men and women and the right to choose the workplace freely ([Bundesverfassungsgericht 2003b](#)). The court demanded that a new regulation be conceived in line with the principles of the constitution. Theoretically, anticipation effects might have played a role from the moment of the ruling onward, although the court gave the government flexibility concerning the specific design of the legislation. Eventually on 1 January 2006 a new law came into effect regulating that firms of all sizes have to take part in the pay-as-you-go system *U2*. That is, large firms are treated in the same way as small firms; they have to pay the social security contribution for maternity protection payments—a pure function of the number of employees and their wages, irrespective of gender and age composition—and the statutory health insurance companies reimburse firms for the wage replacements.

3 Mechanisms and expected effects

This section first explains how the pre-reform regime influenced large firms' expected costs per week of work of potential mothers. Second, it describes the mechanism through which the 2006 reform is expected to impact on wages of female and male workers at large firms.

When calculating the costs of employing a potential mother, employers take into account the firm's expected total costs of motherhood. These include costs that were not affected by the 2006 reform, including e.g. finding a replacement for the mother while she is on leave, or reductions in output during this period. The reform did affect whether companies had to contribute to the 14 weeks of wage continuation during maternity protection. When firms contribute, this increases their costs per week worked of mothers. If male and female workers are perfect substitutes, profit maximising firms are expected to pay potential mothers such that the expected costs per week worked are the same as those for men of the same age, *ceteris paribus*. The magnitude of total wage continuation $Cont$ paid by large firms before the reform is given by Equation (1):

$$Cont = \begin{cases} 0 & \text{if } X(W) \leq 13 \times 365 \\ X(W) \times 14/52 - 13 \times 14 \times 7 & \text{if } X(W) > 13 \times 365, \end{cases} \quad (1)$$

where W denotes the annual gross wage and $X()$ is a function that converts gross wages to net wages. 13 Euro per day of the 14 weeks of maternity protection were paid by insurers.

Denote the weekly gross wage by w . The *expected* effective costs per week of work including employer social security contributions of a potential mother are given by

$$c_{exp} = p \left(\frac{Cont}{weeks} + w + SSC \right) + (1 - p)(w + SSC), \quad (2)$$

where SSC denotes employer social security contributions per week, about 20 per cent of the gross wage, p is the probability that a potential mother will give birth during her current contract, 5.5% per year in our sample,⁶ and $weeks$ are the weeks effectively worked under the current contract. For instance, if the mother has a one year contract and returns to work immediately after maternity protection, this figure is 38.⁷ Note that the effective costs per week of work for a mother under the old regime increase substantially if she takes a long parental leave and thus works fewer weeks under her contract. In the post-reform regime, where $Cont = 0$, the effective costs per week are independent from the length of parental leave. The first term of Equation (2) gives the effective costs per week if the woman gives birth multiplied by the probability of this event. The second term gives the probability of not giving birth multiplied by the weekly wage including social security contributions.

As a benchmark for the order of magnitude of the treatment effect, consider a woman whose monthly gross earnings equal the average in our sample, 2,440 Euro: given the 2006 tax regime and single filing her labour income net of taxes and social security contributions is 1,547 Euro.⁸ Using Equation (2) and assuming a probability of giving birth of 5.5% and that she would return

⁶We implement a method developed by Müller and Strauch (2017) to identify births in German social security data. The annual births probabilities are in line with those calculated by Raute (2019), but slightly higher as we restrict the upper age limit to 35 in our calculations.

⁷Adjusting 52 weeks per year for vacations would increase the terms $Cont/weeks$, w , and SSC by the same proportion. Thus it would not change the ratio of the expected effective costs for potential mothers pre-reform and post-reform.

⁸This figure can be obtained using the tax calculator provided by the Federal Ministry of Finance and applying social security contributions of 0.2 of gross earnings.

to work directly after maternity protection, the expected costs per week of work are 681.08 Euro. In the post-reform regime, where *Cont* is zero and the equation simplifies to $w + SSC$, the costs per week of work would be 675.69 Euro, about 0.8 per cent lower. In practice, women on average take considerable time off after motherhood, therefore this figure is a lower bound for the change in expected costs. Thus, if the gains of the reform are entirely reaped by female workers through higher wages, one would expect wages of potential mothers to increase by at least 0.8 per cent as a result of the reform.

In practice, the adjustment processes may affect both wages and the share of employed women. The reform decreased the expected costs of employing female workers. At constant wages, this leads to firms hiring more women. If the elasticity of substitution of male and female workers equals unity, i.e. they are neither substitutes nor complements, and abstracting from other inputs in the production function, the ratio of expenditure on female and male workers is unchanged and thus the number of employed males remains constant. If instead men and women are substitutes, the number of employed males decreases. If they are complements, the number of both males and females will increase and the sign of the impact on the ratio of male and female workers is ambiguous. The effects on wages depend on the bargaining power of women. If female workers have at least some bargaining power, their wages will increase if their labour costs decrease. Then if female and male workers are substitutes, female wages will increase relative to male wages.

4 Data and empirical approach

The empirical analysis uses linked employer-employee data taken from social security records from the German Institute for Employment Research (IAB). The data set is constructed by first selecting a sample of establishments that took part in the IAB's annual establishment survey and then selecting individuals who have worked at any of these establishments for at least one day between 2002 and 2012. For those individuals we observe the full employment biography from 1993 to 2014 (Heining et al. 2016). As the employment data are based on social security records, civil servants, self-employed and participants in higher education are not covered.

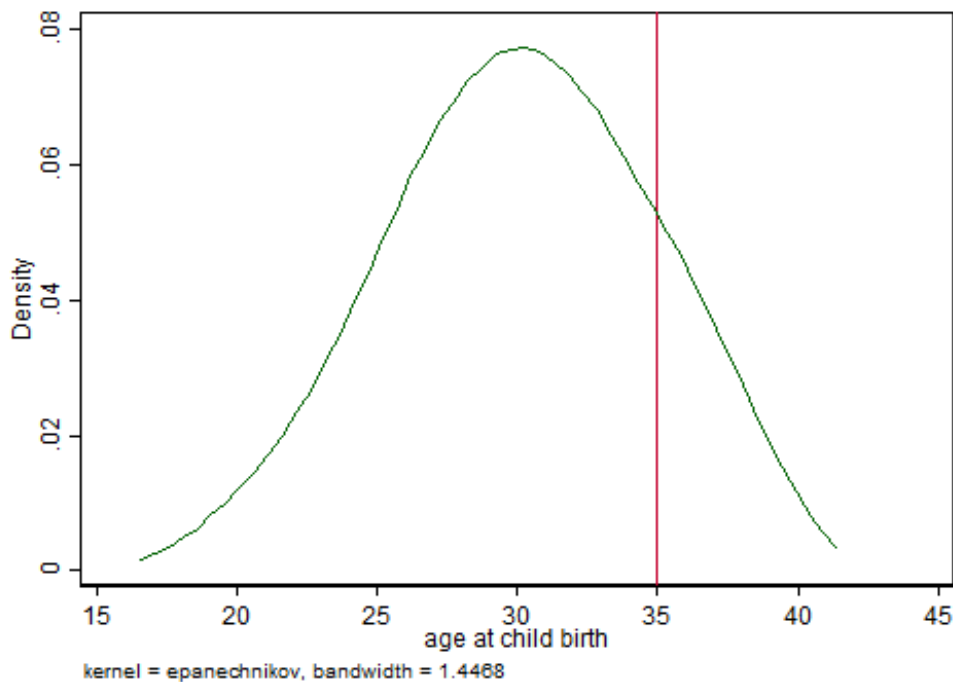
The individual spell data include detailed information on workers' employment history and gross daily wages or benefits and contain a limited set of socio-demographic characteristics such as gender, age and other variables with less complete coverage. By law employers are obliged to report the beginning and end of each employment relationship along with an annual report at the end of each year. The reported daily wage comprises all gross earnings including premiums and allowances. As is common for social security data, the wage is censored from above at the upper earnings limit in the statutory pension fund. The threshold is adjusted in most years and differs between East and West Germany, e.g. in 2010 the threshold was 66,000 Euro in West Germany. The data also contain information on full-time, part-time and marginal employment, but do not contain more detailed information on hours of work. Hence one cannot conclude unambiguously whether a change in daily wages is due to a change in the hourly wage or in hours worked.

At the establishment level, we observe the sector, the geographical location as well as the founding and—if applicable—closing year. In total we observe more than 10,000 establishments in each year, which have been linked to individuals through a unique establishment ID, corresponding to more than 1 million individual observations per year.

We exclude person-level observations with daily wages below 1 Euro, which we assume are either dormant employment relationships or spells attributable to measurement error. We keep both regularly and marginally employed individuals, because the definition of marginal employment changed in 2003, transforming some regular employment relationships into marginal ones.

Excluding marginally employed individuals would lead to a drop in the observed average wage in 2003. We include only individual spells that cover June 30 of a given year, as this is the point in time when the surveys at the establishments are conducted. Finally we exclude firms with nine employees or fewer when briefly contrasting small and large firms, since these very small firms arguably are too different from large ones to constitute an adequate comparison. All analyses restrict the sample to individuals up to 35 years of age. This age threshold is chosen based on the child-births identified in our sample: the age distribution of mothers at birth is plotted in Figure 2 and indicates that after the age of 35 only a small share of women is likely to give birth in the future.

Figure 2: Density plot of mothers' age at childbirth



Notes: Authors' calculations based on the administrative data described in Section 4; births are identified using a method developed by Müller and Strauch (2017). Only childbirths for mothers who were employed subject to social security contributions (pre-birth) are covered. Pooled over the years 2001 to 2010.

We conduct the analysis of (full-time) wages at the individual level.⁹ We use the cut-off of 30 full-time equivalent employees defined by the maternity protection legislation before 2006 to ensure that firms were affected by the change in law. Assignment to the group of large firms is based on firm size in the year 2003, prior to the ruling of the constitutional court. Note that this implies that individuals can move between large and small firms. Therefore, effects on wages could reflect a composition effect to some extent.

In principle, one alternative would have been to assign workers to whether they are in a small or large firm (i.e. not part or part of the analysis sample) depending on the size of their firm in 2003. However, this leads to a substantial problem; one could then only include individuals who were employed in 2003, which would lead to an ageing sample by construction and therefore one could not analyze the effect on women of childbearing age in later years. While the research design for

⁹Note that this is equivalent to conducting an analysis of the gender wage gap at the firm level and weighting annual firm observations by the number of employees used to calculate average wages at the firm level.

testing the hypothesis on statistical discrimination is thus based on the assignment to the analysis sample conditional on the firm size, we make use of the advantages of merged worker-establishment data by combining individual and firm information (and fixed effects) to analyse individual wages conditional on the size of the establishment.

Table 1: Summary statistics: firms by size and sector

	All firms		Manufacturing		Service		Public sector		Other sectors	
	10-30	> 30	10-30	> 30	10-30	> 30	10-30	> 30	10-30	> 30
Number of employees										
Monthly full-time wage, women	1859	2773	1714	2947	2224	2790	2030	2397	1470	2496
	(708)	(675)	(640)	(670)	(799)	(722)	(644)	(474)	(602)	(627)
Monthly full-time wage, men	2105	2963	2045	3110	2537	3085	2226	2631	1755	2539
	(731)	(684)	(592)	(630)	(1026)	(804)	(774)	(626)	(607)	(615)
Share of female employees	0.40	0.35	0.27	0.22	0.50	0.48	0.66	0.66	0.31	0.23
	(0.35)	(0.25)	(0.30)	(0.15)	(0.31)	(0.19)	(0.33)	(0.19)	(0.31)	(0.19)
Share of female new hires	0.41	0.39	0.28	0.27	0.49	0.51	0.67	0.66	0.37	0.26
	(0.41)	(0.26)	(0.37)	(0.19)	(0.38)	(0.23)	(0.37)	(0.22)	(0.38)	(0.24)
Share of women working full-time	0.69	0.82	0.73	0.87	0.75	0.79	0.57	0.67	0.71	0.86
	(0.37)	(0.20)	(0.37)	(0.14)	(0.35)	(0.22)	(0.37)	(0.23)	(0.37)	(0.20)
Share of men working full-time	0.87	0.93	0.92	0.98	0.83	0.89	0.73	0.82	0.87	0.94
	(0.27)	(0.16)	(0.19)	(0.08)	(0.31)	(0.20)	(0.38)	(0.23)	(0.26)	(0.14)
Share women of FT employees	0.35	0.28	0.24	0.17	0.43	0.37	0.57	0.56	0.29	0.18
	(0.29)	(0.23)	(0.24)	(0.13)	(0.24)	(0.17)	(0.29)	(0.20)	(0.23)	(0.16)
Number of new hires	3.5	214.5	2.2	293.5	9.6	122.4	2.7	94.4	3.8	46.6
	(17.4)	(480.9)	(3.7)	(588.6)	(45.4)	(169.9)	(4.4)	(209.0)	(5.7)	(48.5)
At least one new hire (=1)	0.70	0.98	0.69	0.98	0.72	0.98	0.71	0.98	0.74	0.96
	(0.46)	(0.14)	(0.46)	(0.13)	(0.45)	(0.14)	(0.46)	(0.15)	(0.44)	(0.20)
At least one woman hired (=1)	0.41	0.91	0.30	0.89	0.52	0.95	0.59	0.96	0.43	0.79
	(0.49)	(0.29)	(0.46)	(0.31)	(0.50)	(0.22)	(0.49)	(0.2)	(0.50)	(0.41)
Individual-year observations	47,466	1,139,324	26,213	700,156	6,965	147,259	8,506	234,900	5,782	57,009
Firm-year observations	10,658	17,822	5,428	8,679	1,458	2,209	2,539	5,648	1,233	1,286

Notes: Table entries are means and standard deviations (in parentheses) of the variables. All variables are calculated using only employees below 35 years of age. Bold variables indicate dependent variables in the analysis. New hires refer to employees who joined the firm in the 12 months up to 30 June of a given year. Unit of observation for summary statistics is the firm, pooled over the years 2001-2010. Individual-year observations refer to full-time employees which are used in the analysis. Number of employees denotes full-time equivalent employees. Sectors are based on the 2003 edition of the German Classification of Economic Activities (WZ 2003) at the firm level. Manufacturing consists of codes D-G, service of K and L, public sector of N-Q, other sectors capture the residual sectors (agriculture, forestry, mining, hospitality industry and transportation).

Table 1 displays summary statistics of small and large firms. We contrast firms of different sizes to emphasise the relevance of looking at large firms. As these pay higher wages, improving employment prospects for women at larger firms can be an effective way to reduce the overall gender wage gap in the economy. We focus on the years 2001 to 2010 in the analysis in order to take into account a sufficiently long enough period prior to the 2006 reform (and the ruling of the constitutional court); at the same time, given the mechanisms laid out in the previous section it seems plausible that the treatment effect will have materialised and the adjustment process completed within five years post-reform. The table also presents summary statistics by main sectors, in particular classifying workers and firms into manufacturing, services, the public sector and a residual of other sectors using a three-digit sector classification. The firm level summary statistics are based on our analysis sample (individuals below 35 years).

Looking at the mean monthly gross full-time wage, several pronounced patterns are apparent: first, male average monthly wages are consistently higher than female average wages, with an absolute difference of around 200 to 300 Euro, around 10 per cent in relative terms. Second, for both men and women, average wages are substantially higher in large firms than in small firms: in the full sample, average wages at large firms are about 915 Euro higher than at small firms (approx. 50 per cent) for women and about 860 Euro higher (approx. 40 per cent) for men. These differences,

in turn, vary significantly by sector. In the manufacturing sector, the large-small firm differential is more than 1,000 Euro in absolute terms (for both men and women), in service occupations it is around 550 Euro, and in the public sector it is around 400 Euro.

The average share of female workers is lower at large firms (35 per cent) than at small firms (40 per cent). The main pattern of the share of female workers by firm size is maintained across the main three sectors shown in Table 1, albeit at different levels: females make up about one fourth of the workforce in manufacturing firms, about half of the workforce in service firms, and about two thirds in the public sector. Table 1 also shows that the average share of female workers among new hires is about 40 per cent in both large and small firms. Again, differences are visible by main sector: in manufacturing, the share of new hires who are female is just over one fourth, in service occupations it is about half, and in the public sector it is two thirds.

The share of female workers working full-time, unsurprisingly, is consistently lower than the share of male workers working full time. *Within* gender, i.e. looking at the share of women in the firm who work full-time, this share is substantially *higher* in large firms than in small firms: the difference amounts to 13 percentage points in the full sample (82 per cent vs. 69 per cent); it is of similar size in the manufacturing (87 vs. 73 per cent) and public sectors (67 vs. 57 per cent), but smaller in services occupations (79 vs. 75 per cent). *Across* gender, however, i.e. looking at the share of women among all full-time employees, this share is typically *lower* in large firms than in small firms.

We estimate the causal effect of discontinuing the regime of letting large firms pay for maternity protection wage continuation on wages of women up to 35 years of age at large firms using two specifications. As shown in Table 1, wages of men and women differ fundamentally. A DID specification allows for differences among groups, but crucially relies on the common trend assumption, i.e. in the absence of the reform the groups would have followed the same trend and the difference between them would have remained constant. By definition this assumption cannot be tested. The following specification gives the DID estimator for the outcome variable log monthly wages $\ln(w)$ for individual i in firm j at time t :

$$\ln(w_{ijt}) = \gamma_1 \text{female}_i + \gamma_2 \text{female}_i \text{post}_t + \text{year}_t + \omega_j + \epsilon_{ijt}. \quad (3)$$

year_t and ω_j are year and firm fixed effects, and the indicator female_i equals one for women. post_t is a binary indicator that takes on the value one from 2006 onward. The coefficient γ_2 is the DID estimator, and ϵ_{ijt} is the error term. Potential control variables at the firm level are limited in the administrative data and candidates (e.g. founding year, state, or sector) are absorbed by firm fixed effects.

For the comparison of male and female wages the common trend assumption of the DID estimator might be problematic: instead of a common trend there could be some convergence over time, even in the absence of the reform. We thus propose an alternative estimator, which nests DID. Specifically, we assume that the rate of convergence (or divergence) in the outcome between treatment and control group would be constant in the absence of the reform. The reform effect manifests itself in a break of this trend. The *trend-break model* is specified as

$$\ln(w_{ijt}) = \gamma_1 \text{female}_i + \gamma_2 \text{female}_i \text{post}_t + \delta_1 \text{female}_i \text{trend}_t + \delta_2 \text{female}_i \text{posttrend}_t + \text{year}_t + \omega_j + \nu_{ijt}. \quad (4)$$

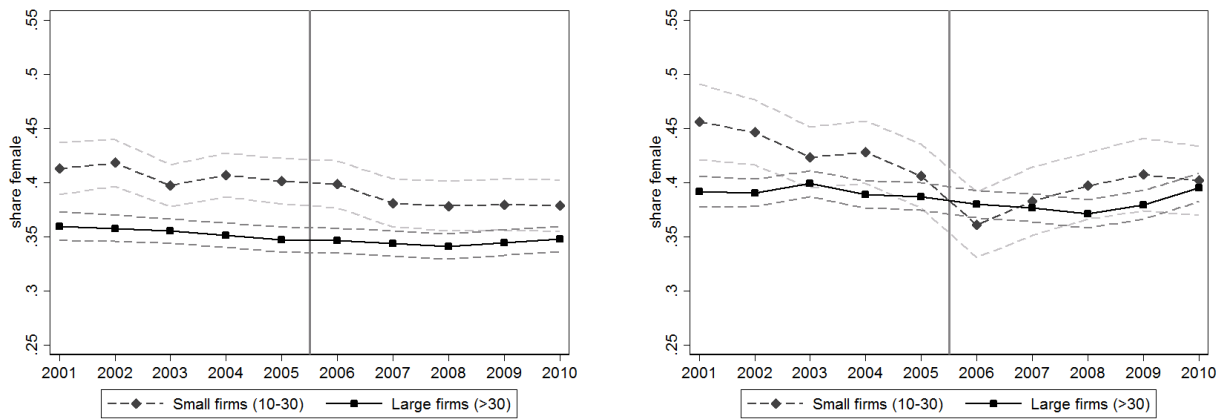
Here $\text{posttrend}_t = \text{year}_t - 2005$ if $t > 2005$ and zero otherwise and $\text{trend}_t = \text{year}_t - 2000$. The coefficient δ_1 thus gives the annual convergence or divergence between the control and treatment groups, and the coefficient δ_2 gives the diversion from this longer-term trend for the post-reform

years. In case of a common trend in the years before the reform, δ_1 would equal zero. The DID estimator γ_2 indicates a one-time shift in the outcome variable due to the reform. δ_2 indicates the additional average *annual* effect of the reform on the treated. For instance, the effect in the first post-reform year is $\gamma_2 + \delta_2$, in the second year it is $\gamma_2 + 2 \times \delta_2$, etc. The TB specification has two advantages. First, it does not rely on the common trend assumption, and second, it may be more in line with the expectation that wages adjust gradually to new regimes.

5 Results

This section presents the empirical results. Figure 3 shows the share of female employees in large and small firms for all employees and for newly hired employees. Two things are worth pointing out; first, the share of female employees in large and small firms does not appear to follow a common trend pre-reform.¹⁰ Second, both the female employment share and the female share in hiring are virtually constant in large firms before and after the reform. An interesting finding in itself, the latter gives the wage analysis within large firms more legitimacy, as concerns about a compositional bias in the wage estimations are alleviated. The clearest adjustment mechanism—and, hence, evidence for statistical employer discrimination—can be observed by contrasting female and male wages *within* large firms. Moreover, this analysis is not contaminated by potential spill-overs from large to small firms.

Figure 3: Share of female employees in small and large firms, 2001-2010



(a) All employees

(b) New hires

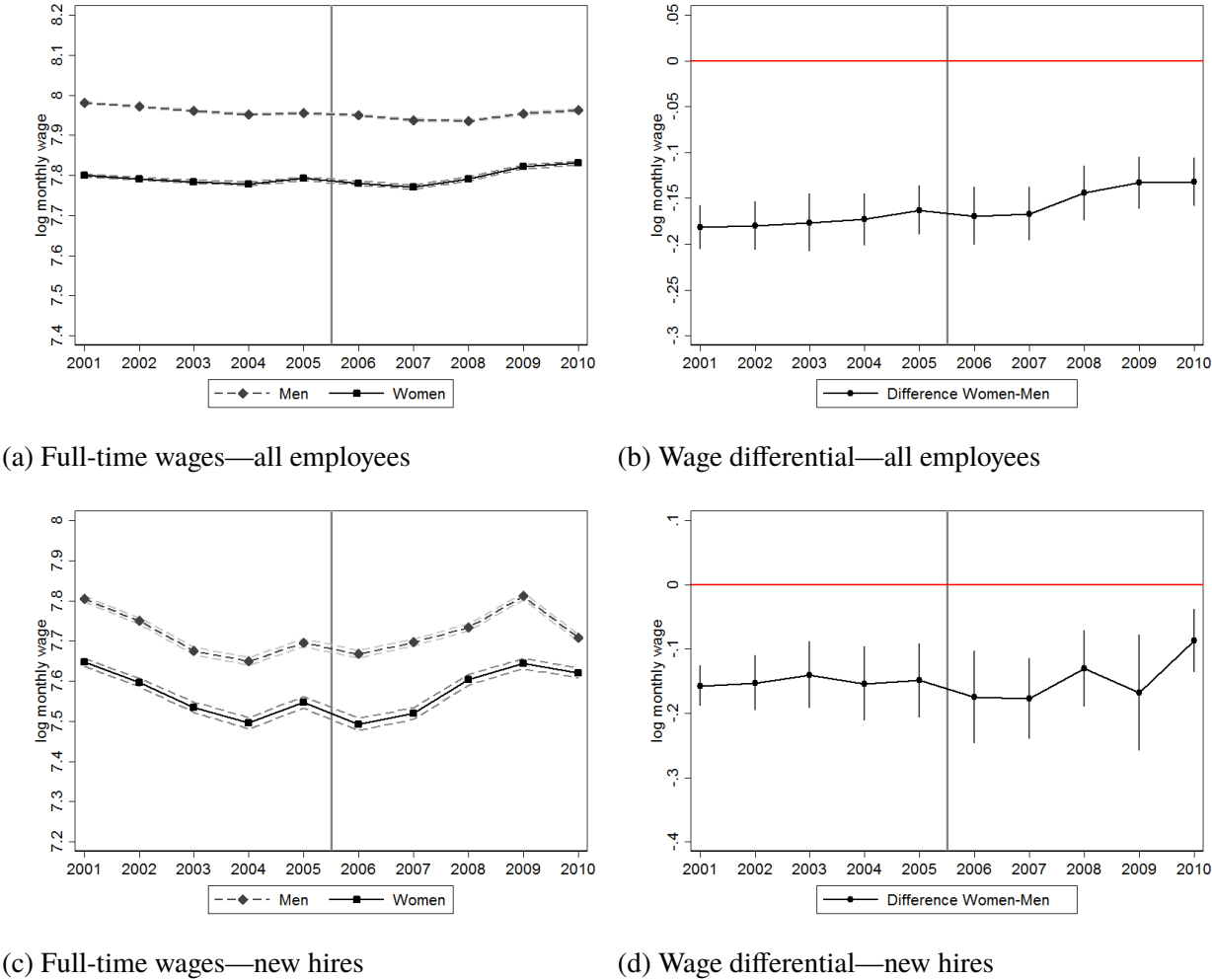
Notes: Figures display the share of female employees in small (10-30 employees) and large firms (>30). Panel (a) plots shares for all employees up to 35 years of age, Panel (b) for newly hired employees specifically. The vertical line indicates the reform. Firms are weighted by their number of full-time equivalent employees. Light grey dashed lines indicate 95 per cent confidence intervals.

Figure 4 begins with a descriptive investigation, plotting log full-time monthly wages for men and women and the respective female-male differential in large firms over the observation period. Panel (a) of the figure shows a parallel—i.e. flat—development of the respective male and female wages for all employees during the pre-reform period, and an increase in female wages vs. a

¹⁰We explored an analysis of the reform effect using the share of female employees in large vs. small firms as a dependent variable, but due to the lack of a common trend this analysis was inconclusive.

continued flat curve for male wages during the post-reform years. Panel (b) illustrates this pattern for the gender wage differential, and again shows its narrowing during the post-reform years. The figure therefore gives some indicative evidence for the expected adjustment mechanism and statistical discrimination by employers.

Figure 4: Average full-time wages of men and women in large firms (> 30 employees), 2001-2010



Notes: Figures on the left display annual averages of log monthly wages for female and male full-time employees up to 35 years of age in large firms (>30 employees). Light grey dashed lines indicate 95 per cent confidence intervals. Figures on the right show the coefficient for the difference between the groups and 95% confidence intervals. Panels (a) and (b) plot wages for all employees, Panels (c) and (d) for newly hired employees. The vertical line indicates the reform.

When looking specifically at wages for new hires in the bottom panels, given the resulting smaller sample sizes, the patterns shown are somewhat noisier than the top panel. Panel (c) indicates a slightly decreasing trend in wages for both male and female new hires—a trend that would possibly be in line with a German labour market that was at the time (early to mid 2000s) characterised by the highest unemployment rate since the 1950s and by wage moderation (the same trend is also visible for female wages in small firms, which are not reported here). Perhaps more importantly for this study, however, the bottom panel of Figure 4 shows that during the pre-reform

Table 2: Estimates of the reform effect on log full-time wages within large firms

	Full sample		Above median wage		Below median wage	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All employees						
Post-period \times female (γ_2)	0.0113*** (0.0040)	-0.0132*** (0.0043)	0.0098* (0.0054)	-0.0142** (0.0055)	0.0143** (0.0063)	-0.0126* (0.0066)
Time trend \times female (δ_1)		0.0004 (0.0016)		-0.0013 (0.0019)		0.0039* (0.0020)
Post-2006 trend \times female (δ_2)		0.0076*** (0.0024)		0.0102*** (0.0029)		0.0026 (0.0030)
Clusters	2,007	2,007	528	528	1,477	1,477
Observations	1,139,324	1,139,324	622,968	622,968	516,301	516,301
Panel B: Newly hired employees						
Post-period \times female (γ_2)	0.0151* (0.0091)	-0.0074 (0.0152)	0.0207 (0.0156)	0.0060 (0.0281)	0.0118 (0.0099)	-0.0155 (0.0128)
Time trend \times female (δ_1)		0.0011 (0.0036)		-0.0070 (0.0067)		0.0053 (0.0040)
Post-2006 trend \times female (δ_2)		0.0056 (0.0046)		0.0167** (0.0079)		0.0002 (0.0061)
Clusters	1,975	1,975	521	521	1,452	1,452
Observations	217,202	217,202	99,952	99,952	117,235	117,235
Year FEs	Y	Y	Y	Y	Y	Y
Firm FEs	Y	Y	Y	Y	Y	Y

Notes: Table entries are the interaction terms of Equations (3) and (4). Dependent variable is log monthly wage. Sample is restricted to full-time employees up to 35 years of age. Observations denote individual by year observations. Columns (1) and (2) use the full sample, columns (3)-(6) stratify according to the median wage among all large firms in 2003. Standard errors (in parentheses) are clustered at the firm level. */**/** denote statistical significance at the 10/5/1% level, respectively.

period also for new hires male and female wages in large firms display a parallel development and that the corresponding gender wage differential is narrowing slightly during the post-reform years.¹¹ The confidence intervals in Panel (d) indicate that this is less precisely estimated than for all workers, but the main pattern is still visible.

The estimation results in Table 2 indicate that the adjustment process in large firms anticipated in theory has indeed taken place in practice: first, the DID estimate in column (1) shows an increase in female wages relative to male wages of 1.1 per cent, statistically significant at the 1 per cent level. The TB model in column (2) carves out the pattern more finely: the general time trend in the convergence (or divergence) of the gender wage gap in large firms is not statistically different from zero. The estimated reform effect in the TB model is composed of the coefficients γ_2 and δ_2 . That is, the reform effect is estimated to be -0.56 per cent in 2006 ($1 \times 0.76-1.32$), 0.2 per cent in 2007 ($2 \times 0.76-1.32$), 0.96 per cent in 2008, 1.72 per cent in 2009, and 2.48 per cent in 2010. The latter value is the total reform effect after five years. An F-test yields that the estimated reform effect is

¹¹One might wonder whether the overall macroeconomic trend—decreasing employment rates before 2006, increasing employment rates from 2006 onward—might affect wage growth differentially for men and women. Burda and Seele (2017) show that, while the overall downward trend in real wages between 2003 and 2011 is the same for men and women, it is slightly more pronounced for women. If anything, this would make our analysis underestimate the reform effect on wages.

statistically significant at the 5 per cent level from 2008 onward and insignificant before. This result indicates that indeed under the pre-2006 regulation employers showed statistically discriminating behaviour against women, depressing their wages, which was then reversed with the reform.

The bottom panel of Table 2 presents estimation results for the wages of new hires. The point estimates for the DID (column (1)) and the TB models (column (2)) are similar in size to those for all employees, but less precisely estimated. The DID point estimate indicates a positive reform effect of 1.5 per cent (significant at the 10 per cent level) on wages of newly hired female workers. Columns (3) through (6) of Table 2 investigate further these reform effects, by distinguishing large firms that pay high wages from large firms that pay lower wages: specifically, we classify firms as above and below median wage if their average wage was above or below the sample median in the year 2003 (such that this classification, again, is independent of the reform).

Table 3: Estimates of the reform effect on log full-time wages in large firms—by sector

	Manufacturing		Service		Public sector		Other sectors	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All employees								
Post-period \times female (γ_2)	0.0209*** (0.0048)	-0.0003 (0.0061)	0.0174** (0.0075)	0.0041 (0.0110)	0.0176** (0.0079)	-0.0034 (0.0095)	0.0113 (0.0155)	0.0116 (0.0199)
Time trend \times female (δ_1)		0.0017 (0.0024)		-0.0015 (0.0029)		0.0009 (0.0022)		0.0029 (0.0062)
Post-2006 trend \times female (δ_2)		0.0044 (0.0031)		0.0070* (0.0041)		0.0053 (0.0036)		-0.0050 (0.0062)
Clusters	1,004	1,004	259	259	599	599	145	145
Observations	700,156	700,156	147,259	147,259	234,900	234,900	57,009	57,009
Panel B: Newly hired employees								
Post-period \times female (γ_2)	0.0207 (0.0130)	0.0202 (0.0187)	0.0343* (0.0180)	0.0067 (0.0360)	0.0245 (0.0158)	-0.0103 (0.0229)	-0.0200 (0.0202)	-0.0163 (0.0462)
Time trend \times female (δ_1)		-0.0135* (0.0077)		0.0045 (0.0069)		0.0127** (0.0059)		-0.0062 (0.0121)
Post-2006 trend \times female (δ_2)		0.0233** (0.0106)		0.0017 (0.0068)		-0.0099 (0.0083)		0.0090 (0.0114)
Clusters	994	994	254	254	586	586	141	141
Observations	109,276	109,276	35,626	35,626	58,695	58,695	13,605	13,605
Year FEs	Y	Y	Y	Y	Y	Y	Y	Y
Firm FEs	Y	Y	Y	Y	Y	Y	Y	Y

Notes: Table entries are the interaction terms of Equations (3) and (4). Dependent variable is log monthly wage. Sample is restricted to full-time employees up to 35 years of age. Observations denote individual by year observations. Sectors based on the 2003 edition of the German Classification of Economic Activities (WZ 2003). Manufacturing consists of codes D-G, service of K and L, public sector of N-Q, other sectors capture the residual sectors (agriculture, forestry, mining, hospitality industry and transportation). Standard errors (in parentheses) are clustered at the firm level. */**/** denote statistical significance at the 10/5/1% level, respectively.

For both firm types, the respective DID specifications—columns (3) and (5)—show a positive and statistically significant impact on female wages after the reform (the respective point estimates for new hires in the bottom panel are insignificant). Results from the TB model indicate that the overall pattern found in column (2) is determined to a larger extent by firms that pay above median wages: column (4) estimates that the annual convergence (divergence) between treatment and control group is zero, but that due to the reform there was a positive total average wage effect of 3.68 per cent for women over the 5-year post-reform period ($5 \times 1.02-1.42$), statistically significant at the 5 per cent level. For firms paying below median wages—column (6)—the reform effect is less precisely estimated and indicates an overall effect of zero. Again, these empirical findings

echo *ex ante* expectations: Section 2 has shown that the costs of maternity protection for the firm are higher for high-wage earners, and this feature of the regulation appears to have led firms to statistically discriminate against this group accordingly.

Additional results on wage effects by sectors are presented in Table 3, again distinguishing between all employees and new hires. Several patterns are worth noting: first, estimation results from the DID model across sectors again point towards a relatively large—1.74 to 2.09 per cent—and statistically significant treatment effect on female wages, as indicated in columns (1), (3), and (5). Thus, the main finding of a positive reform effect wages of young women is not driven by a particular sector. Second, estimation results from the TB model indicate that the time trend of convergence (divergence) between male and female wages is estimated to be zero (columns (2), (4), (6), and (8)). The reform effect in the TB model by sector is relatively imprecisely estimated, and indicates a notable pattern only for the services sector, with an overall reform effect of 3.9 per cent over the 5-year post reform, significant at the 10 per cent level.

As sensitivity analysis we additionally included an interaction of the female binary variable and an indicator that equals one for the crisis years 2008-09. While the coefficient of this additional variable showed that men were hit harder than women by the crisis, the main result did not change. Moreover, the main result also proved robust to controlling for a level shift in female relative to male wages due to a reform of the parental leave system in 2007.

6 Conclusion

Despite substantial improvements in women's labour market prospects in the past decades, women are still paid, on average, lower wages. Motherhood has been identified as one driving mechanism. While a large literature examines the *ex-post* career cost of motherhood, theory predicts that maternity may also be punished *ex-ante* through statistical discrimination by employers: specifically, negative effects on all women of childbearing age may arise if firms face direct costs of employing potential mothers. To test the hypothesis of statistical employer discrimination, we use the natural experiment of a maternity protection reform in Germany.

The empirical results contrast female and male wages within large firms and confirm the theoretical prediction: first, DID estimates show a statistically significant increase in female wages relative to male wages of 1.1 per cent in the post-reform period. This treatment effect estimate is in line with a back-of-the-envelope calculation, which puts the expected effect at around 0.8 per cent. Second, the trend-break model indicates that the general time trend in the convergence of female and male wages in large firms is not statistically different from zero, and the post-reform trend estimates a total reform effect of 2.5 per cent over five years. This result indicates that indeed under the pre-reform regulation employers showed statistically discriminating behaviour against potential mothers due to maternity protection wage continuation.

This evidence for statistical discrimination implies several policy conclusions. Since we find that labour market prospects improved significantly due to the reform, policy makers should try to identify factors that could result in statistical discrimination against potential mothers. A prominent example is the German parental leave legislation, where the current incentive structure leads women, who on average earn less than their partners, to take longer leave periods. Moreover, our findings support that it is worthwhile for the public to finance, through taxes, costs that occur asymmetrically to mothers only, if the aim is to prevent negative labour market effects *ex-ante* and *ex-post* for (potential) mothers, and to reduce gender-specific earnings differences.

References

- Adda, J., Dustmann, C., and Stevens, K. (2017). The career costs of children. *Journal of Political Economy*, 125(2):293–337.
- Arrow, K. J. (1973). *The Theory of Discrimination*. Princeton University Press.
- Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago press.
- Biewen, M. and Seifert, S. (2018). Potential parenthood and career progression of men and women—a simultaneous hazards approach. *The BE Journal of Economic Analysis & Policy*, 18(2).
- Blau, F. D. and Kahn, L. M. (2003). Understanding international differences in the gender pay gap. *Journal of Labor Economics*, 21(1):106–144.
- Bundesverfassungsgericht (2003a). 1 bvr 302/96.
- Bundesverfassungsgericht (2003b). Press release 98 / 2003-12-09.
- Burda, M. C. and Seele, S. (2017). Das deutsche Arbeitsmarktwunder: Eine Bilanz. *Perspektiven der Wirtschaftspolitik*, 18(3):179–204.
- Card, D., Cardoso, A. R., and Kline, P. (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.
- Fang, H. and Moro, A. (2010). Theories of statistical discrimination and affirmative action: A survey. In Benhabib, J., Bisin, A., and Jackson, M., editors, *Chapter 5 of Handbook of Social Economics*, volume IA of *Handbook of Labor Economics*, chapter 27, pages 133–200. North-Holland.
- Gangl, M. and Ziefle, A. (2009). Motherhood, labor force behavior, and women’s careers: An empirical assessment of the wage penalty for motherhood in Britain, Germany, and the United States. *Demography*, 46(2):341–369.
- Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C., and Schupp, J. (2018). The German socio-economic panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik*.
- Goldin, C. (2014). A grand gender convergence: Its last chapter. *The American Economic Review*, 104(4):1091–1119.
- Heining, J., Klosterhuber, W., Lehnert, P., Seth, S., et al. (2016). Linked-Employer-Employee-Daten des IAB: LIAB-Längsschnittmodell 1993-2014 (LIAB LM 9314). Technical report, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany].
- Heinze, A. and Wolf, E. (2010). The intra-firm gender wage gap: A new view on wage differentials based on linked employer–employee data. *Journal of Population Economics*, 23(3):851–879.

- Müller, D. and Strauch, K. (2017). Identifying mothers in administrative data. Technical report, Institut für Arbeitsmarkt-und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany].
- OECD (2019). Gender wage gap (indicator). doi: 10.1787/7cee77aa-en, accessed on 20 March 2019.
- Olivetti, C. and Petrongolo, B. (2008). Unequal pay or unequal employment? A cross-country analysis of gender gaps. *Journal of Labor Economics*, 26(4):621–654.
- Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review*, 62(4):659–661.
- Raute, A. (2019). Can financial incentives reduce the baby gap? Evidence from a reform in maternity leave benefits. *Journal of Public Economics*, 169:203–222.
- Weichselbaumer, D. and Winter-Ebmer, R. (2005). A meta-analysis of the international gender wage gap. *Journal of Economic Surveys*, 19(3):479–511.