

Supplemental Appendix: Forward-Looking Labor Supply Responses to Changes in Pension Wealth: Evidence from Germany

Elisabeth Artmann* Nicola Fuchs-Schündeln[†] Giulia Giupponi[‡]

March 2026

Contents

Appendix A Additional Figures and Tables	2
Appendix B Details on Institutional Setting	13
B.1 The 2019 Mütterrente Reform	13
B.2 Maternity Leave Legislation in Germany	13
Appendix C Details on Data Construction	15
C.1 Simulation of Pension Wealth Effect in VSKT Data	15
C.2 Imputation of Pension Points in IEB Data	16
C.3 Matching of Individuals to Their Spouses	18
Appendix D Comparison of Extensive-Margin Employment Effects with Estimates in Becker et al. (2022)	21
Appendix E Heterogeneity Analysis	24
E.1 Returns to Tenure and Costs of Working	24
E.2 Ability to Self-Insure Against Longevity Risk	25
E.3 Liquidity Constraints	25
E.4 Age at the Time of the Reform	26
Appendix F Intra-household Spillover Effects	31

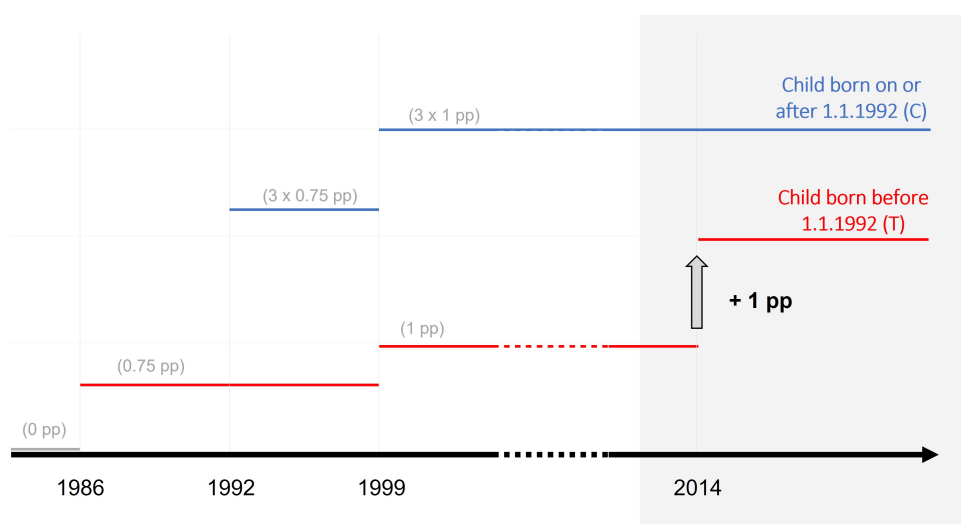
*Institute for Employment Research (IAB). Email: elisabeth.artmann2@iab.de

[†]WZB Berlin Social Science Center, Goethe University Frankfurt and CEPR. Email: nicola.fuchs@wzb.eu

[‡]Bocconi University and CEPR. Email: giulia.giupponi@unibocconi.it

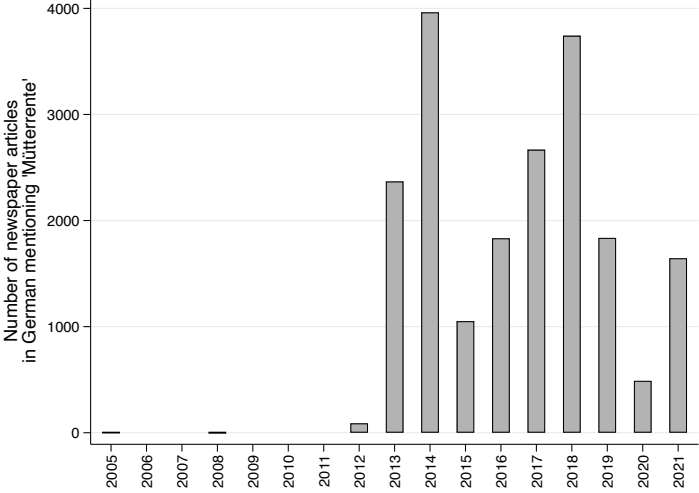
Appendix A. Additional Figures and Tables

Figure A1. HISTORY OF MÜTTERRENTE UP TO 2018



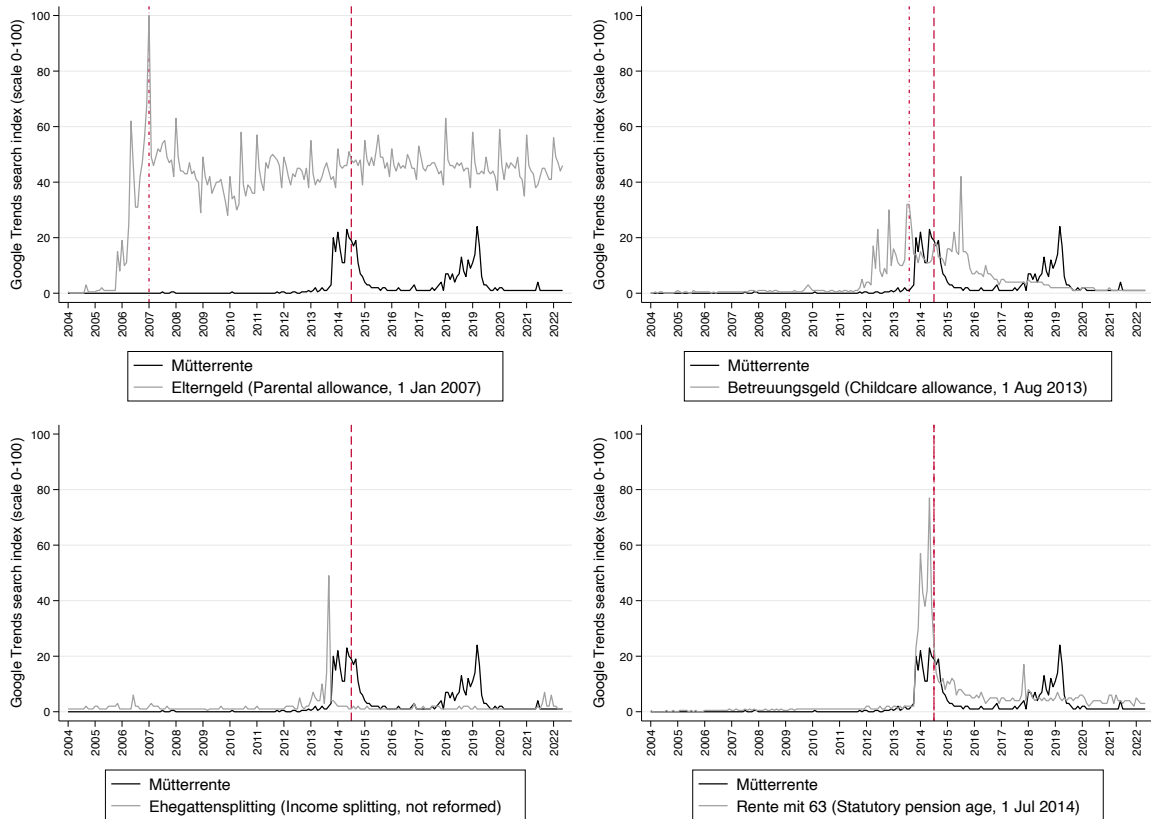
Notes: The chart illustrates the evolution of Mütterrente from 1986 to 2018. From 1986 to 1991, mothers were credited a maximum of 0.75 pension points per child for the first year after childbirth; this also applied to childbirths before 1986. The full 0.75 pension point benefit was granted to mothers who were not employed in the first year after childbirth. For employed mothers, childcare pension points were fully withdrawn against pension points from regular employment. From 1992, mothers of children born on or after January 1, 1992 were entitled to a maximum of 2.25 pension points (three years \times 0.75 pension points). Mothers of children born before January 1, 1992 continued to receive a maximum of 0.75 pension points per child. Childcare pension points were still withdrawn against employment contributions at a 100% rate. A reform in 1999 increased the generosity of the childcare pension benefit from 0.75 to one pension point per year and repealed the earnings test. The 1999 pension reform was retroactive. Starting July 2014, mothers whose children were born before January 1, 1992 got one extra pension point per child and, as a result, reached a total of two pension points per child born before 1992. The effect of this reform is analyzed in this paper.

Figure A2. NUMBER OF NEWSPAPER ARTICLES INCLUDING THE WORD 'MÜTTERRENTE'



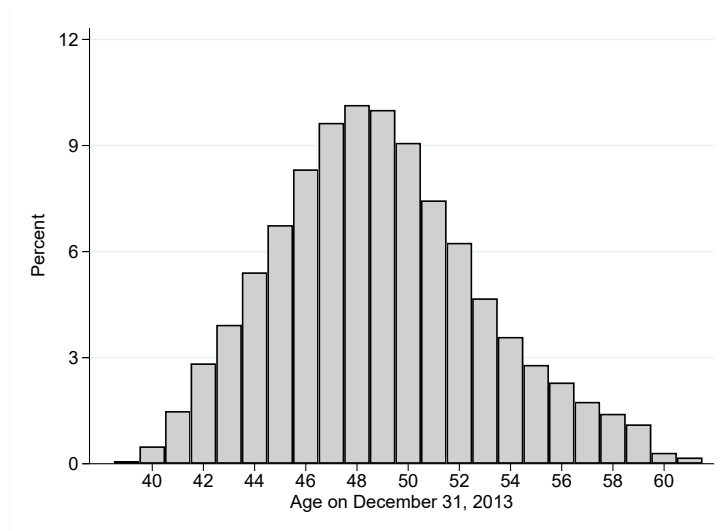
Notes: The graph reports the number of newspaper articles in German including the word 'Mütterrente'. Data are sourced from Factiva.

Figure A3. GOOGLE SEARCHES FOR ‘MÜTTERRENTE’ AND OTHER GENDER- AND PENSION-RELATED POLICIES



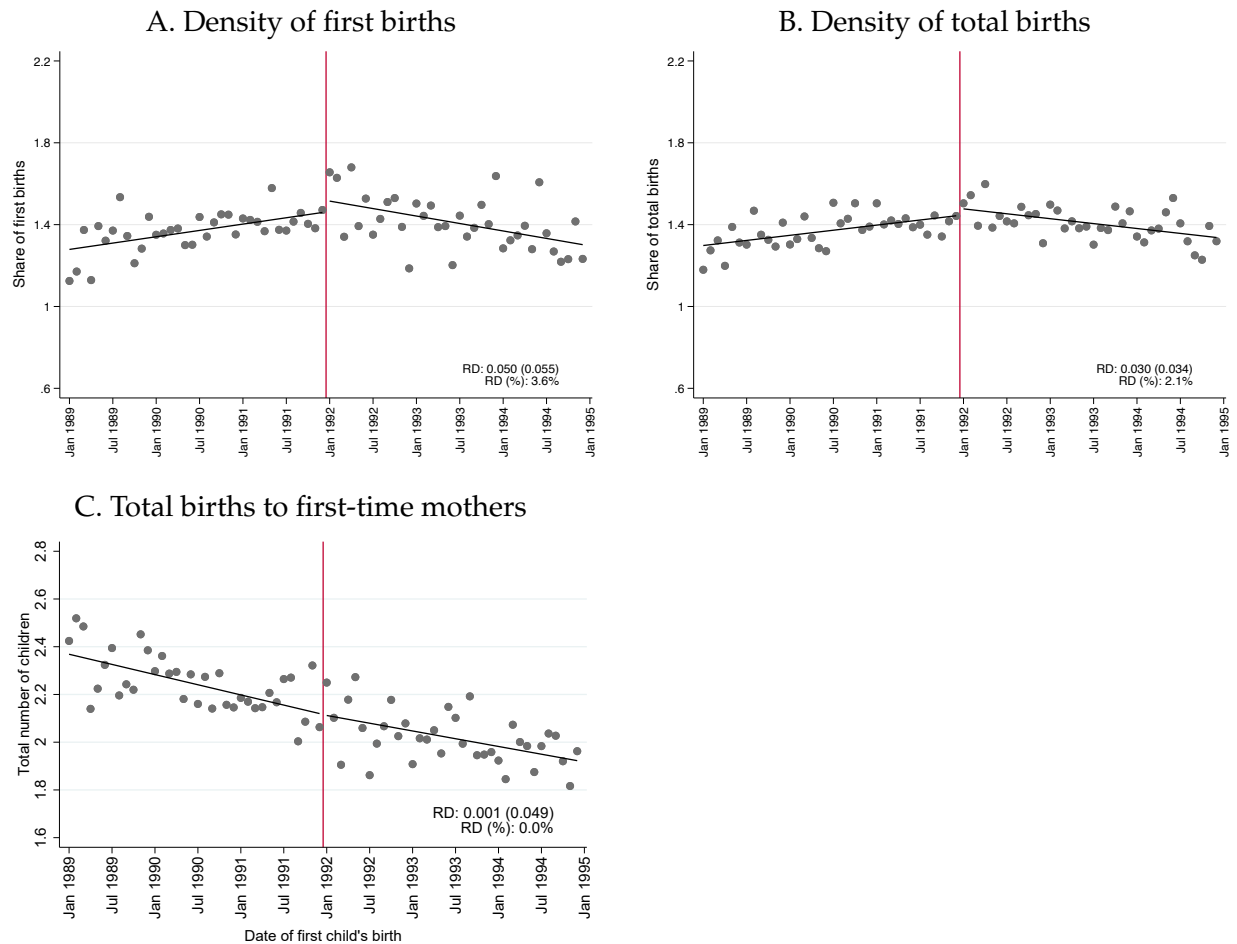
Notes: The figure reports the evolution of Google Trends search indices for the term ‘Mütterrente’ and for terms referring to other gender- and pension-related policies implemented in Germany. The Google Trends search index is normalized to take values between 0 and 100, where 100 corresponds to the point in time in which the largest number of searches took place within the location specified in the query. Our query covers Google searches across Germany between 2004 and the beginning of 2022. The index is comparable across searches. Each panel reports a comparison of the Google Trends search index for the term ‘Mütterrente’ (black line) and other gender-related policies (grey line). Panel A reports searches for ‘Elterngeld’, a parental allowance introduced on January 1, 2007; Panel B for ‘Betreuungsgeld’, a childcare allowance introduced on August 1, 2013; Panel C for ‘Ehegattensplitting’, the German system of joint taxation of married couples, which was not reformed during this period but is often discussed; Panel D for ‘Rente mit 63’, a reduction of the statutory seniority-pension age to 63 down from 65 which took place on July 1, 2014. The dashed vertical line corresponding to July 2014 indicates the time of implementation of Mütterrente; the dash-dotted vertical line indicates the time of implementation of the other policy, when relevant.

Figure A4. AGE DISTRIBUTION OF MOTHERS IN THE MAIN SAMPLE IN DECEMBER 2013



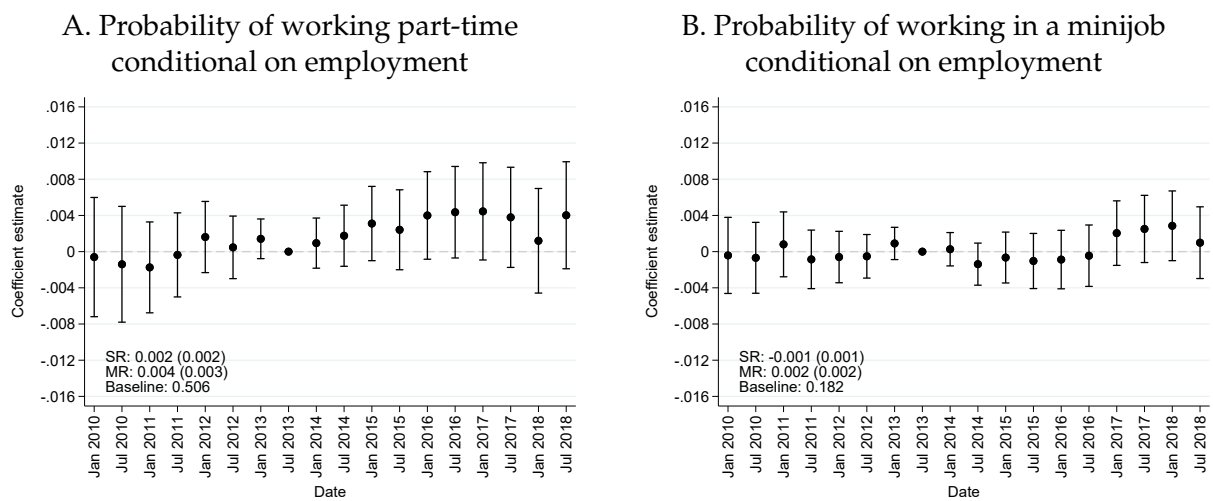
Notes: The graph shows the age distribution for women in the main sample as of December 31, 2013.

Figure A5. EFFECT OF 1992 MATERNITY-LEAVE AND MÜTTERRENTE REFORMS ON FERTILITY OUTCOMES



Notes: The graphs report parametric regression-discontinuity-design (RDD) estimates of the effect of the 1992 maternity-leave and Mütterrente reforms on fertility outcomes. We use data from VSKT for all charts. Panels A and B report the density of first births and total births, respectively. The gray dots plot the outcome by month, residualized against month of birth fixed effects. The solid line is fitted based on a parametric RDD with linear fit, conditional on month of birth fixed effects. The graph also reports the point estimate of the discontinuity at the January 1, 1992 cutoff and associated robust standard error. We also report the discontinuity as a percent of the mean outcome on the left of the cutoff. Panel C reports the total number of children born to a mother as of December 2019. The gray dots plot the average number of children born to mothers by month of birth of their first child, residualized against mother's cohort and child's month of birth fixed effects. The solid line is fitted based on a parametric RDD with linear fit, also conditional on cohort and month of birth fixed effects. The graph also reports the point estimate of the negative of the discontinuity at the January 1, 1992 cutoff and associated robust standard error, clustered at month of birth level. We report the negative of the discontinuity so that it reflects the impact on the total fertility of the treatment group (relative to the control group). We also report the discontinuity as a percent of the mean outcome on the right of the cutoff.

Figure A6. EFFECT OF 2014 MÜTTERRENTE REFORM ON THE PROBABILITY OF WORKING PART-TIME OR IN A MINIJOB CONDITIONAL ON EMPLOYMENT



Notes: The graphs report estimates of the coefficients γ_s from equation (2). The dots represent the point estimates and the capped vertical bars 95% confidence intervals based on robust standard errors clustered at the individual level. Each graph also reports estimates of the short-run effect over the 1-2.5 years after the 2014 reform (SR) and of the medium-run effect over the 3-5 years after the reform (MR), with robust standard errors clustered at the individual level in parentheses. The baseline statistic is the average value of the outcome in the treatment group in July-December 2013.

Table A1. DESCRIPTIVE STATISTICS: IEB SAMPLE

	Mothers with first child born		Difference	p-value of diff.
	before 1.1.1992 (3 months)	after 1.1.1992 (3 months)		
	Mean (1)	Mean (2)	(3)	(4)
<i>Individual characteristics</i>				
Age at first birth	27.30	27.46	-0.164	0.000
Age on December 31, 2013	49.44	49.32	0.125	0.000
High education	0.22	0.23	-0.004	0.109
<i>Labor market status in December 2013</i>				
Any employment	0.72	0.72	-0.007	0.010
Full-time employment	0.21	0.20	0.002	0.475
Part-time employment	0.38	0.39	-0.011	0.000
Marginal employment	0.13	0.13	0.002	0.344
UI benefits and/or jobseeker	0.05	0.05	0.003	0.059
Not in IEB data	0.23	0.22	0.005	0.069
<i>Cumulative duration (in months) as of December 31, 2013</i>				
Any employment	240.44	241.21	-0.775	0.194
<i>Unconditional earnings July to December 2013</i>				
Any employment	7760.14	7870.71	-110.56	0.028
Observations	50,428	48,676		
<i>Earnings July to December 2013 conditional on employment of the respective type</i>				
Any employment	10,452.71	10,511.84	-59.13	0.283
Observations	37,438	36,446		
Regular employment	12,155.55	12,149.61	5.94	0.917
Observations	30,426	29,830		
Regular FT employment	15,931.47	15,889.06	42.41	0.695
Observations	10,839	10,384		
Regular PT employment	9,909.16	10,001.16	-92.01	0.099
Observations	19,793	19,652		
Marginal employment	1,905.09	1,910.04	-4.95	0.697
Observations	7,536	7,152		

Notes: The table reports summary statistics from the IEB sample, separately for mothers of children born before (treatment group) and after (control group) January 1, 1992 (columns 1 and 2 respectively), omitting all births in the two weeks around January 1, 1992. The table also reports the difference in means between the treatment and control groups (column 3), and the p-value of the difference (column 4). High education is defined as having either a university (of applied science) degree or A-levels plus vocational training.

Table A2. DESCRIPTIVE STATISTICS: VSKT SAMPLE

	Mothers with first child born		Difference (3)	p-value of diff. (4)
	before 1.1.1992 (12 months) Mean (1)	after 1.1.1992 (12 months) Mean (2)		
<i>Demographics and pension points as of December 2013</i>				
Age	49.10	48.36	0.745	0.000
Number of children	1.96	1.94	0.027	0.474
Cumulated pension points	23.32	24.60	-1.283	0.002
<i>Labor market status in December 2013</i>				
Employed	0.73	0.74	-0.016	0.443
Regular employment	0.60	0.62	-0.018	0.440
Marginal employment	0.13	0.13	0.002	0.897
Unemployed	0.05	0.05	0.003	0.767
Observations	1,300	1,487		

Notes: The table reports summary statistics from the VSKT sample, separately for mothers of children born before (treatment group) and after (control group) January 1, 1992 (columns 1 and 2 respectively). The table also reports the difference in means between the treatment and control groups (column 3), and the p-value of the difference (column 4).

Table A3. REPRESENTATIVENESS OF MAIN SAMPLE

	Sample 1 (1)	Sample 2 (2)	Sample 3 (3)
<i>Sample restrictions</i>			
Birth year	1952-1974	1952-1974	1952-1974
First child born in	1991 or 1992	Unrestricted	Before 1992
<i>Individual characteristics</i>			
Year of birth	1964	1963	1959
Age at first birth	27.4	27.4	24.4
Total number of children	2.3	2.1	2.3
Years of education	12.3	12.4	11.8
Total labor market experience	14.6	15.7	17.9
Experience before age 30	6.9	5.9	5.7
Observations	238	3,412	1,436

Notes: The table reports summary statistics for different samples of mothers using data from the German Socio-Economic Panel (SOEP) wave 38. For comparability with our main IEB sample, all samples exclude mothers who ever lived in East Germany and who ever worked in the mining industry. Sample 1 covers mothers born 1952-1974 who gave birth to their first child in 1991 or 1992. This is analogous to our main IEB sample. Sample 2 includes mothers born 1952-1974 leaving unrestricted the birth date of their first child. This sample can be viewed as covering "mothers in general". Sample 3 covers mothers born 1952-1974 who gave birth to their first child before 1992 and can be viewed as covering "mothers affected by the reform", conditional on the same cohorts as our main sample. Variables are measured as of 2013, except for the total number of children and years of education that are measured as of 2019. Sample sizes are maximum sample sizes in each column. Observations for labor market experience may be lower.

Table A4. MAIN ESTIMATES BY PRE-REFORM EMPLOYMENT STATUS

	Unconditional earnings (1)	Any employment (2)	Conditional earnings (3)	Cond. fulltime employment (4)
<i>Panel A. Sample of mothers employed in July-December 2013</i>				
T × Post (yr 1-2.5)	-49.70 (23.50)	-0.003 (0.002)	-39.47 (19.35)	-0.003 (0.002)
T × Post (yr 3-4.5)	-103.56 (34.54)	-0.003 (0.002)	-82.39 (27.72)	-0.005 (0.002)
Pre-reform mean of dep. var.	10,452.71	0.966	10,452.71	0.277
N individuals	73,884	73,884	73,845	73,845
<i>Panel B. Sample of mothers not employed in July-December 2013</i>				
T × Post (yr 1-2.5)	-13.59 (28.35)	-0.001 (0.003)		
T × Post (yr 3-4.5)	-39.99 (41.57)	-0.002 (0.004)		
Pre-reform mean of dep. var.	0.00	0.000		
N individuals	25,220	25,220		

Notes: The table reports estimates of the coefficients γ_s from equation (2) for post-reform years. The coefficients γ_s are estimated pooling post-reform years together, distinguishing between a short-run effect over the 1-2.5 years after the 2014 reform and a medium-run effect over the 3-5 years after the reform. Robust standard errors clustered at the individual level are reported in parentheses. The estimates in Panel A are based on the sample of individuals who had positive labor earnings (i.e. were employed) in the second half of 2013. Those in Panel B are based instead on the sample of those with zero labor earnings (i.e. not employed) in the second half of 2013. In Panel B, we do not report results for earnings conditional on employment and full-time status conditional on employment, since those are not defined in the base period for the sample analyzed. The table reports the average value of the outcome in the treatment group in July-December 2013.

Table A5. PLACEBO ESTIMATES

	Unconditional earnings (1)	Any employment (2)	Conditional earnings (3)	Cond. fulltime employment (4)
<i>Panel A. Placebo threshold: January 1, 1994</i>				
T × Post (yr 1-2.5)	-23.48 (21.71)	0.001 (0.002)	-21.73 (21.66)	0.002 (0.002)
T × Post (yr 3-5)	-24.43 (31.15)	0.003 (0.002)	-30.34 (30.44)	-0.002 (0.002)
Pre-reform mean of dep. var.	7,938.53	0.737	10,410.72	0.254
N individuals	84,676	84,676	72,538	72,538
<i>Panel B. Placebo threshold: January 1, 1995</i>				
T × Post (yr 1-2.5)	-7.77 (23.08)	0.001 (0.002)	-11.75 (22.96)	0.000 (0.002)
T × Post (yr 3-5)	-17.53 (33.32)	0.002 (0.002)	-20.27 (32.45)	-0.000 (0.003)
Pre-reform mean of dep. var.	8,037.86	0.745	10,435.18	0.243
N individuals	77,636	77,636	66,954	66,954

Notes: Panels A and B report estimates of the coefficients γ_s from versions of equation (2), where we use January 1, 1994 and January 1, 1995 as placebo thresholds to define (placebo) treatment and control groups, respectively. The coefficients γ_s are estimated pooling post-reform years together. Row 1 of each panel reports estimates of the short-run effect over the 1-2.5 years after December 2013 (SR) and row 3 of the medium-run effect over the 3-5 years after the reform (MR), with robust standard errors clustered at the individual level in parentheses. Row 5 reports the average value of the outcome in the (placebo) treatment group in July-December 2013.

Table A6. ROBUSTNESS TO DIFFERENT DOUGHNUT-HOLE WIDTHS

	Unconditional earnings (1)	Any employment (2)	Conditional earnings (3)	Cond. fulltime employment (4)
<i>Panel A. Zero weeks</i>				
T × Post (yr 1-2.5)	-51.76 (17.82)	-0.002 (0.001)	-40.42 (17.88)	-0.002 (0.002)
T × Post (yr 3-4.5)	-86.30 (25.81)	-0.001 (0.002)	-84.69 (25.24)	-0.004 (0.002)
Pre-reform mean of dep. var. N individuals	7,742.31 117,554	0.717 117,554	10,431.27 98,220	0.276 98,220
<i>Panel B. Four weeks</i>				
T × Post (yr 1-2.5)	-55.54 (21.57)	-0.003 (0.002)	-50.48 (21.56)	-0.005 (0.002)
T × Post (yr 3-4.5)	-115.73 (31.22)	-0.003 (0.002)	-114.03 (30.39)	-0.008 (0.002)
Pre-reform mean of dep. var. N individuals	7,783.90 81,054	0.717 81,054	10,487.23 67,769	0.277 67,769
<i>Panel C. Six weeks</i>				
T × Post (yr 1-2.5)	-39.84 (24.13)	-0.002 (0.002)	-39.07 (24.13)	-0.005 (0.002)
T × Post (yr 3-4.5)	-106.55 (34.98)	-0.002 (0.002)	-108.15 (34.15)	-0.009 (0.003)
Pre-reform mean of dep. var. N individuals	7,790.85 64,253	0.718 64,253	10,489.34 53,716	0.279 53,716

Notes: The table reports estimates of the coefficients γ_s from equation (2) for post-reform years for samples with different doughnut-hole widths around the January 1, 1992 cutoff. Our baseline sample excludes observations with imputed childbirths in a two-week window around January 1, 1992 (see Table 1). Panel A reports estimates for a sample with no doughnut hole; Panel B for a sample with a four-week doughnut whole and Panel C for a sample with a six-week doughnut hole. The coefficients γ_s are estimated separately for the 1-2 years and the 3-4.5 years after the July 2014 reform. Robust standard errors clustered at the individual level are reported in parentheses.

Appendix B. Details on Institutional Setting

B.1. The 2019 Mütterrente Reform

Following the 2014 Mütterrente reform, the number of pension points credited for children born before 1992 was equal to two, while that for children born in or after 1992 was equal to three. In order to progressively equalize the treatment of mothers of children born before and after January 1, 1992, additional 0.5 pension points were credited to mothers of children born before January 1, 1992 starting from January 1, 2019, under the Mütterrente scheme. Appendix Figures A2 and A3 show substantial coverage and salience of this additional reform, too. Due to the smaller wealth shock generated by the 2019 reform and the short post-reform time window, we do not analyze this reform here and restrict our analysis to the years up to 2018 only.

We would argue that it is highly unlikely that individuals had anticipated the 2019 reform and expected a full equalization of childcare pension points already in 2014, or at any point before 2018. First, the 2014 reform received quite some criticism regarding its financing and the burden it would impose on the pension system. Second, the 2019 reform was not included in the CDU's official election manifesto, in contrast to the 2014 reform. Moreover, it became part of a CDU-SPD coalition agreement only in March 2018, when the coalition was, however, planning to grant one additional pension point per child only to women with 3 or more children born before 1992, while the actual reform granted half a pension point to all women who gave birth before 1992. This implies that the average woman in our sample (who has two children) would not have expected to benefit from the reform as stated in the coalition agreement. The law crediting 0.5 points to all mothers was then passed in November 2018. More generally, the mid-2010s were not characterized by intense pension policy activity nor policy uncertainty around retirement issues.¹ The last pension reform had in fact been passed in 2007.

B.2. Maternity Leave Legislation in Germany

To provide a complete picture of the institutional context, we briefly illustrate the features of German legislation in the matter of maternity leave. In Germany, mothers have been entitled to fully paid maternity leave for the 6 weeks before and the 8 weeks after childbirth since the mid-1950s. During maternity leave, the firm cannot dismiss the mother, and the mother has the right to return to a job comparable to the one she held before childbirth in the same firm.

In 1979, Germany introduced additional parental leave, which in practice has been used by mothers in the vast majority of cases. This parental leave extended the job protection period to 6 months after birth, and introduced parental pay (at a fixed amount) for the same time period.² From the late 1980s, Germany enacted a series of reforms to this scheme. In July 1989 and July

¹By contrast, the 1990s were a decade of high policy uncertainty regarding pension reforms (Giavazzi and McMahon, 2008). See Chapter 2 in Giupponi and Seibold (2024) for a review of pension reforms in Germany since the early 1990s.

²The first eight weeks after childbirth are still fully paid.

1990, the job protection and maximum benefit periods were extended to 15 and 18 months, respectively. In January 1992, the duration of job-protected leave was increased from 18 to 36 months, while that of leave payments was extended to 24 months in January 1993. [Schönberg and Ludsteck \(2014\)](#) analyze the effect of these reforms on mothers' labor supply decisions after childbirth. They find that the 1992 reform — whose timing coincides with one of the childcare pension benefit reforms — reduced the share of mothers who had returned to work 19 months after childbirth by 9.6 percentage points. It also slightly reduced the probability that a mother had returned to work 6 years after childbirth (by 0.8 percentage points).

Even though the child's birth date of January 1, 1992 is a relevant cutoff date with respect to both the Mütterrente and the parental leave reforms, this does not pose a problem for identifying the effect of the July 2014 pension reform, since we compare the behavior of mothers who had children before versus after the cut-off date of January 1, 1992, before versus after July 2014. We assume that any confounding effect of the 1992 maternity leave reforms is accounted for by the second, time difference.

Appendix C. Details on Data Construction

C.1. Simulation of Pension Wealth Effect in VSKT Data

We use the VSKT data to simulate the absolute and proportional increase in the present discounted value of pension wealth generated by the 2014 reform at the time of the coalition agreement in December 2013, which is the baseline period in our difference-in-differences analysis. We employ the sample restrictions described in Section B to make the VSKT sample as close as possible to our IEB one.

The absolute effect of the reform on the present discounted value of pension wealth is simulated based on the following formula:

$$(C1) \quad \Delta W_{\text{Dec 2013}} = \beta^J \cdot \Delta P_{\text{Dec 2013}} \cdot (1 + \gamma)^J \cdot R_{\text{Dec 2013}} \cdot \sum_{k=0}^{240-1} \beta^k (1 + \gamma)^k$$

where $\Delta W_{\text{Dec 2013}}$ indicates the change in the present discounted value of pension wealth induced by the 2014 reform as of December 2013. $\Delta P_{\text{Dec 2013}}$ is the mechanical change in cumulated pension points due to the reform, which is of one pension point per child born prior to January 1, 1992. In our sample of analysis $\Delta P_{\text{Dec 2013}} = 1.03$, which is slightly larger than 1 due to multiple births.

In order to map the pension point increase into a monetary quantity, we multiply it by the value of a pension point R (Rentenwert), which was EUR 28 per month in West Germany in December 2013 ([Deutsche Rentenversicherung Bund, 2020](#); [Statistisches Bundesamt, 2022](#)). We account for the fact that the value of a pension point R is expected to change over time by multiplying its monthly value in 2013 by a growth factor γ over (i) the J months that separate a mother from her full retirement age and (ii) the following months over which she is expected to receive pension benefits. We estimate γ empirically by regressing the natural logarithm of the CPI-adjusted R in each month on a time trend and use the coefficient on the trend as our measure of γ .³ We estimate $\hat{\gamma} = -0.0001619$ over the years from 2002 (the introduction of the Euro) to 2019.⁴ According to official life expectancy tables from the German Federal Statistical Office ([Statistisches Bundesamt, 2025](#)), the average woman in our sample is expected to live – and hence receive pension payments – for 20 years (240 months) after her full retirement age (67). Last, we discount the change in cumulated pension payments due to the reform to December 2013. For discounting, we use $\beta = 0.9983$ as monthly discount factor.

This formula provides an increase in the present discounted value of pension wealth of EUR 3,830, corresponding to a 4.4% increase of pension wealth as of December 2013.

³In practice, the Rentenwert is updated each year by the government as a function of gross wage growth.

⁴If we consider the period between 2002 and 2013 – to reflect the time period over which individuals in the sample are likely to have formed their expectations of the Rentenwert’s growth – we obtain a more negative $\hat{\gamma}$, which would imply a slightly larger estimate of the MPE.

C.2. Imputation of Pension Points in IEB Data

In the German pension system, there are three types of pension-relevant periods: contribution periods (Beitragszeiten), non-contributory periods (beitragsfreie Zeiten) and consideration periods (Berücksichtigungszeiten) ([Deutsche Rentenversicherung Bund, 2021b](#)).⁵ Contribution periods are periods during which contributions to the pension insurance are made, most importantly periods of employment subject to social security contributions, of childcare as outlined in Section B, or of specific times of unemployment. Non-contributory periods include periods during which no contributions can be made, e.g. because of long-term sickness, rehabilitation, maternity leave, education, or certain periods of unemployment. As will be explained below, some non-contributory periods can also be credited with pension points. Consideration periods are important for determining how many pension points someone might get attributed for non-contributory periods. The most important ones are related to childcare until the child is 10 years old, as explained in more detail below.

Some of the events associated with pension-relevant periods, such as long-term sickness or rehabilitation, can either not be (reliably) identified in the IEB or have eligibility conditions that cannot be directly verified in the data. For our analysis, the quantitatively most important contributory events are employment and care for the first child, so that we discuss the imputation of pension points for these periods in more detail. For brevity, the imputation of pension points for other pensionable periods like unemployment is only discussed briefly.

In order to impute pension points from employment subject to social security contributions (SSC), we compute monthly earnings based on gross daily wage and monthly employment duration. While wages in the IEB are in principle capped at the social security contribution ceiling, it is possible for calculated monthly earnings to exceed the cap in some cases. This can be due to approximation error of the monthly wage, or, for instance, bonus payments that are added up to regular income. In all these cases, we cap earnings at the pension contribution ceiling. Monthly pension points from employment subject to SSC are then calculated as the ratio of monthly earnings to the average annual income in the population as determined by the government (see also Section A).

Earnings from marginal employment are subject to employer pension contributions, while employee contributions were voluntary until the end of 2012.⁶ Although pension contributions became the default for marginal employees in 2013, employees can still opt-out of making contributions. In the data, we cannot observe whether an employee actually contributes to pension insurance for marginal employment. In March 2020, more than 80% of marginal employees did not pay pension contributions ([Minijob-Zentrale, 2020](#)). Therefore, only employer contributions

⁵Non-contributory periods comprise so-called Anrechnungszeiten and Ersatzzeiten. In the following, we omit the latter as these are periods during which individuals cannot make contributions due to, for example, captivity of war or persecution.

⁶Under certain conditions (e.g. earnings above the income threshold), marginal employment is subject to regular social security contributions. In these cases, income from marginal employment is added to income from employment subject to SSC above.

are taken into account when imputing pension points. The employer contribution rate was 12% until June 30, 2006, and subsequently increased to 15%. Marginal employment in private households is subject to an employer contribution rate of 5%. Pension points from marginal employment per spell are calculated as $\frac{\text{Income from marginal employment}}{\text{Average annual income}} \times \frac{\text{Sector employer contribution rate}}{\text{Total pension contribution rate}}$. The total pension contribution rate is the sum of the employer and employee pension contribution rates for regular employment subject to SSC and amounted to 18.6% in 2018. The monthly pension points from marginal employment are the sum of pension points for all spells in a month.

The regulations on how periods of receipt of different types of unemployment benefits (unemployment insurance, unemployment assistance, maintenance benefits, unemployment benefits II) are credited with pension points have changed repeatedly over time (Steffen, 2011; Deutsche Rentenversicherung Bund, 2022). In the imputation procedure, we take these changes into account as far as the data permits to impute the pension entitlements derived from unemployment as accurately as possible.

To impute pension points for contribution periods of childcare before the Mütterrente reform, we assign 0.0833 (i.e. 1/12) pension points per month for the first 12 (36) months to the treatment (control) group starting from the month after the birth of the child. If the annual cap on pension points is exceeded by the sum of pension points from other contribution periods and childcare, the pension points from childcare are reduced accordingly (Deutsche Rentenversicherung Bund, 2021a). In the analysis, we only use childcare pension points as of December 31, 2013, and only assign them for the first child, since we cannot identify the majority of higher order births.

There are two types of non-contributory periods, those that are credited with pension points (bewertete Anrechnungszeiten) and those that are not credited with pension credits (unbewertete Anrechnungszeiten). For instance, non-contributory periods like school-based education are not credited with pension points, while specific periods of unemployment are credited with pension points.^{7,8} How many pension points are credited for these periods is determined in the overall assessment procedure explained further below. Both of these non-contributory periods have in common that they are subtracted from the pensionable period and therefore indirectly increase an individuals' average monthly pension contribution over the insured life calculated in the overall assessment procedure. While this procedure is normally performed upon retirement, we use it to calculate individuals' total pension entitlements from contribution, consideration and non-contributory periods as of December 31, 2013. We do this for both partners, though childcare-related pension points are only imputed for women.

Consideration periods for childcare play a role in the overall assessment procedure and are intended to close gaps in the insurance records or compensate for reduced contributions that

⁷We cannot observe school-based education in the data, but assume that individuals with A-levels have two years of education after their 17th birthday and individuals with a university (of applied science) degree have seven years.

⁸Additionally, months of vocational training are considered reduced-value contribution periods (beitragsgeminderte Zeiten), so that the pension entitlements derived from these periods are topped up to 75% of the average monthly pension value, but subject to a cap of 0.0625 pension points obtainable per month.

are caused by childcare obligations. The consideration period for childcare begins on the child's day of birth and ends when the child turns 10. This period counts towards the qualifying period necessary to be eligible for certain types of pensions, e.g. the seniority pension that requires 45 years of contributions. The consideration period does not directly increase pension entitlements, but is 'fictitiously' credited with 0.0833 pension points per month subject to the annual cap on pension points (Beye, 2009).

Upon retirement, the so-called overall assessment procedure (Gesamtleistungsbewertung) is performed to determine the pension entitlements for non-contributory periods that are credited with pension points. The overall assessment procedure first calculates an individual's sum of actual pension points from contribution periods and fictitious pension points from consideration periods. These points are then divided by the total pensionable period (belegungsfähiger Zeitraum) to calculate individuals' average monthly pension contribution over their insured life (Gesamtleistungswert) (Beye, 2009). Months of creditable non-contributory periods are then assigned pension points to the value of the average monthly pension contribution, thus increasing total pension entitlements. Therefore, a shorter pensionable period is beneficial in the overall assessment procedure. The pensionable period starts the day after the individual turns 17, but includes pension-relevant periods before age 17 (if any), and typically ends when entering retirement (Beye, 2009). Non-contributory periods are subtracted from the pensionable period, while other gaps in the insurance record (e.g. periods out of the labor force for non-pension-relevant reasons) are not.

As shown in Appendix Table C1, our procedure accurately imputes pension points for employment and unemployment spells. Due to the difficulty in identifying multiple and second or higher order births, we underestimate total childcare pension points, but we precisely estimate the difference in total childcare pension points between mothers whose first child was born right before and right after January 1, 1992. We note that the table reports pension points for the contributory events that we can impute in the IEB data.

C.3. Matching of Individuals to Their Spouses

In order to match spouses, Bächmann et al. (2021) create cross-sections from the Integrated Employment Biographies as of June 30 of each year from 2001 to 2014 (Institut für Arbeitsmarkt- und Berufsforschung, 2021b). They restrict the sample to individuals who are at least 16 years old, the legal age to marry in Germany. Using information on surnames and geo-coded residential addresses, two individuals of opposite sex are matched as spouses if they meet all three of the following conditions: (i) having the same surname, (ii) having the exact same residential location, and (iii) being uniquely identified as a couple.

The first condition requires that couples have the same surname recorded in the data, including double surnames if they are hyphenated (e.g. Müller-Schmidt). The procedure thus cannot identify married couples who do not share a surname nor cohabiting couples. Second, the

matching procedure requires that married couples share the exact same residential location consisting of municipality, postal code, street name and street number. Spouses who do not cohabit, report different addresses or report the same address in different ways can therefore not be linked. Third, the uniqueness condition requires that only two individuals aged 16 or older with the same surname can report the exact same residential location. Finally, matched individuals with an absolute age difference of more than 15 years are excluded from the sample. Appendix Table C2 reports, for each year, the number of couples that can be identified for the last time in that year, using the matching procedure described above. More than half of the couples can be identified at least once in the five years preceding 2014, which correspond to the pre-reform years in our analysis. Overall, out of the 99,104 women in our sample, 41,646 (42%) can be matched to a spouse. As for women, we exclude partners with employment spells in East Germany and potential times in the Knappschaft, due to incomplete working histories and different pension rules, which make it difficult to precisely impute pension points. We exclude partners who died before July 2014, since they are no longer alive after the reform. Finally, even though we do not (and cannot) restrict the sample to partners who are also the father of women's children, we do exclude male partners born after 1979, since those would be 12 or younger at the time of the female partner's first childbirth. These restrictions reduce the final sample size of couples to 37,427 (38%).

Table C1. VALIDATION OF PENSION POINT IMPUTATION

	VSKT		IEB	
	Mothers with first child born before 1.1.1992 (1)	after 1.1.1992 (2)	Mothers with first child born before 1.1.1992 (3)	after 1.1.1992 (4)
<i>Cumulative pension points for contribution periods on December 31, 2013</i>				
Income subject to SSC	11.60	11.48	11.47	11.56
Unemployment	0.321	0.313	0.240	0.246
Childcare	4.00	5.90	0.995	2.971
Observations	1,300	1,487	50,428	48,676

Notes: The table reports the number of cumulated pension points for different types of contributory events, separately for mothers whose first child was born before and after January 1, 1992, and separately for the VSKT and IEB samples. Pension points are directly available in VSKT; they are, instead, imputed in IEB. In the IEB data, childcare pension points only include those for the first child.

Table C2. MATCHING OF SPOUSES

Last year of observation as couple (June 30)	Frequency	Percent	Cumulative percent
2001	1,581	3.80	3.80
2002	1,273	3.06	6.85
2003	1,259	3.02	9.88
2004	1,347	3.23	13.11
2005	1,383	3.32	16.43
2006	1,486	3.57	20.00
2007	3,357	8.06	28.06
2008	4,356	10.46	38.52
2009	3,750	9.00	47.52
2010	3,282	7.88	55.41
2011	3,384	8.13	63.53
2012	2,520	6.05	69.58
2013	2,912	6.99	76.57
2014	9,756	23.43	100.00
N couples	41,646	100.00	
N couples final sample	37,427	89.87	

Notes: The table reports, for each year, the number of couples that can be identified for the last time in that year, using the matching procedure and sample restrictions described in Appendix C.3. The final sample excludes couples in which the male partner fulfils any of the following conditions: (i) at least one employment spell in East Germany, (ii) potential times at the Knappschaft, (iii) died before July 2014, or (iv) born after 1979.

Appendix D. Comparison of Extensive-Margin Employment Effects with Estimates in [Becker et al. \(2022\)](#)

In contemporaneous work, [Becker et al. \(2022\)](#) use VSKT data to analyze the effect of the 2014 Mütterrente reform on the probability of employment. In their preferred specification, they compare mothers who had *any* child in 1990 or 1991, to mothers who had their *first* child in 1992 or 1993. This specification leads to a reduced form effect of -0.011 with a standard error of 0.004, i.e. points to a statistically significant reduction of 1.1 percentage points in the likelihood of employment. Our estimate of the effect of the reform on the probability of employment is of -0.002 with a standard error of 0.002 (not significant at conventional levels). Two elements may be responsible for the differing headline estimates in the two papers: (i) the definition of treatment (any versus first child before 1992) and (ii) the childbirth bandwidth around the January 1, 1992 cutoff. In addition, (iii) the IEB data having a larger sample size (100% of the population versus the 2% of VSKT data) is likely responsible for the improved precision of our estimates. To understand what the main drivers of the differing findings are, we compile a set of estimates of the reduced form effect of the policy on the probability of employment, starting from [Becker et al. \(2022\)](#)'s preferred specification and altering one specification or sample parameter at a time, until we get to our main specification.

Appendix Table D1 reports the results of this exercise. Column 1 provides a brief description of the specification; column 2 the sample used, indicating the data source (VSKT or IEB data) and the sample coverage (2% or 100%); column 3 indicates the treatment definition, where the main difference across specifications is whether the treatment group is defined as having *any* child before 1992 or the *first* child before 1992; column 4 indicates the time bandwidth over which treatment and control childbirths are defined; column 5 reports the reduced form estimate with the associated standard error in parentheses. Standard errors are always clustered at the individual level.

We start in row 1 with [Becker et al. \(2022\)](#)'s preferred specification, based on the VSKT 2% sample and a comparison of mothers who had *any* child in 1990 or 1991, and mothers who had their *first* child in 1992 or 1993. This specification leads to a reduced form effect of -0.011 with a standard error of 0.004, i.e. points to a reduction in the likelihood of employment of 1.1 percentage points. Row 2 reports results from a similar specification – taken from [Becker et al. \(2022\)](#) – in which the treatment group is now defined as mothers having their *first* child in 1990 or 1991. The control group is defined as in row 1. The point estimate is almost halved in size compared to row 1 and becomes statistically insignificant. Row 3 restricts the sample to mothers who had their first child in 1991 or 1992 (12-month bandwidth), leading to a further halving of the point estimate, which remains statistically insignificant.

Row 4 acts as a “bridge” from [Becker et al. \(2022\)](#)'s to our sample. We implement the same specification reported in row 3 using a 2% random extract of our IEB data. We estimate an

insignificant effect of 0.001, with a standard error that is identical to the one in row 3.⁹ Row 5 restricts the childbirth bandwidth to three months on either side of the cutoff, as we do in our main specification. Finally row 6 implements our main specification, using the full 100% IEB sample. The point estimates in rows 5 and 6 are both equal to -0.002, with the 100% sample in row 6 giving us a nearly seven-fold increase in precision.

These results indicate that (i) the definition of treatment (any vs first child before 1992) and (ii) the childbirth bandwidth are both significant drivers of the difference between [Becker et al. \(2022\)](#)'s main results and our main results. In fact, even when looking at [Becker et al. \(2022\)](#)'s results only, we can see that moving from the specification in row 1 to the specification in row 3 leads to a dramatic reduction in the coefficient estimate (in absolute value), with an associated loss of significance. The comparison of rows 5 and 6 highlights that having a 100% sample allows us to achieve a great degree of precision in our estimate, with the standard error dropping from 0.013 to 0.002.

⁹We note that, when running this specification, we need to correct for differential pre-trends between the treatment and control group.

Table D1. COMPARISON OF SPECIFICATIONS

Specification	Sample	Treatment	Bandwidth	Estimate
1. Becker main	VSKT 2%	Post x Any child before	1990-1993	-0.011 (0.004)
2. Becker 1 st child	VSKT 2%	Post x First child before	1990-1993	-0.006 (0.005)
3. Becker 1 st child 12m	VSKT 2%	Post x First child before	1991-1992 (12 months)	-0.003 (0.006)
4. IAB 2% 1 st child 12m	IAB 2 %	Post x First child before	1991-1992 (12 months)	0.001 (0.006)
5. IAB 2% 1 st child 3m	IAB 2%	Post x First child before	1991-1992 (3 months)	-0.002 (0.013)
6. IAB 100% 1 st child 3m	IAB 100%	Post x First child before	1991-1992 (3 months)	-0.002 (0.002)

Notes: The table reports a series of estimates of the reduced form effect of the policy on the probability of employment, starting from [Becker et al. \(2022\)](#)'s preferred specification and altering one sample or specification parameter at a time, until we get to our main specification. Column 1 provides a brief description of the specification; column 2 the sample used, indicating the data source (VSKT or IEB data) and the sample coverage (2% or 100%); column 3 indicates the treatment definition, where the main difference across specifications is whether the treatment group is defined as having *any* child before 1992 or the *first* child before 1992; column 4 indicates the time bandwidth over which treatment and control childbirths are defined; column 5 reports the reduced form estimate with the associated standard error in parentheses. Standard errors are always clustered at the individual level. The estimation in row 4 is based on a specification that corrects for differential pre-trends between the treatment and control group.

Appendix E. Heterogeneity Analysis

E.1. Returns to Tenure and Costs of Working

We start by considering whether individuals with different labor market returns to tenure respond differentially to the wealth shock. We measure returns to tenure as the late-career firm-level tenure gradient of earnings in the occupation in which an individual is employed. We first estimate occupation-specific wage tenure gradients in a 2% sample of the male population employed in the private sector in Germany ([Institut für Arbeitsmarkt- und Berufsforschung, 2021a](#)).¹⁰ Returns to tenure are estimated using a Mincerian regression in which log daily earnings are a function of individual demographics (education, age and age squared), year fixed effects, tenure and tenure squared (in months). To account for the intensive margin of work, months of tenure at a given establishment are weighted by working time, with weights taking value 1 for full-time, 0.5 for part-time, and 0.2 for marginal employment. This regression is estimated for each occupation separately using data for the years 2010 to 2018. Having estimated occupation-specific late-career returns to tenure, we assign women in our IEB sample to the occupation they were employed at in the second half of 2013. We rank individuals based on the predicted return to continuing working full-time in that occupation until the end of the sample period. We then median-split them into two groups, such that half of the women are in the above-median returns group and half in the below-median returns group. Panel A of Appendix Table E1 reports estimates of an augmented version of equation (2), fully interacted with a dummy for having above-median predicted returns to tenure. Therein, we report the coefficients on the term $D_i \cdot \mathbb{I}[t > \text{Jul-Dec 2013}]$ and its interaction with the heterogeneity indicator. Consistent with forward-looking behavior, the results show that individuals with higher predicted returns to tenure respond less strongly at the intensive margin, in that they are less likely to leave full-time employment for a part-time job or a minijob.

On the cost side, we study whether individuals who are employed in more physically demanding occupations, for whom presumably the cost of working is thus larger, respond more strongly to the wealth shock. We classify occupations based on an index of physical strain constructed using information from O*NET ([National Center for O*NET Development, 2022a](#)).¹¹ The latter provides occupation-specific ratings of (i) the amount of time spent in tiring body positions (e.g. kneeling, bending, standing or making repetitive motions), (ii) the frequency of exposure to extreme environmental conditions (e.g. cramped work spaces, bright or inadequate lighting, hot or cold temperatures, body vibrations or discomforting sounds) and (iii) the likelihood of

¹⁰We estimate returns to tenure on the sample of men for two reasons. First, men tend to have uninterrupted careers until retirement, which allows to better estimate tenure gradients. Second, men are not directly affected by the reform, implying that our estimates of the tenure gradients are not endogenous to the reform itself. We restrict the male sample to individuals in the same age range as women in our IEB sample (i.e. ages 39 to 61 in December 2013) and weight the male sample to be representative of the female one in terms of the age distribution.

¹¹For this analysis, we use the fine-grained 5-digit classification of occupations available in the IEB (not the 127 aggregated categories used above) as only this can be merged with the data from O*NET ([National Center for O*NET Development, 2022b](#); [Institute for Structural Research, 2022](#)).

job hazards (e.g. exposure to disease or infections, to hazardous conditions or equipment, or to radiation). We construct an index of physical strain as an average of the O*NET rankings over these three dimensions, and study heterogeneous responses across mothers who were employed in occupations with above versus below median values of the physical strain index. The results reported in Panel B of Table Appendix E1 do not reveal any substantial difference in labor supply responses by physical strain of the occupation, indicating that the cost of work effort does not seem to be mediating the observed response to the wealth shock.¹²

E.2. Ability to Self-Insure Against Longevity Risk

The labor supply response to a change in pension wealth may depend on the amount of total resources that individuals expect to be able to rely upon once retired. In other words, we may expect individual labor supply responses to vary depending on the ability to self-insure against the 'risk of longevity' via one's own or the partner's savings and pension wealth. We explore this question by estimating heterogeneous responses to the 2014 reform by measures of individual and household wealth. Since in the data we cannot observe wealth from sources other than public pension, we employ as proxies for wealth the number of non-childcare pension points cumulated by the end of 2013.

We start by analyzing heterogeneous responses by the number of non-childcare pension points cumulated by the mother as of December 2013. Since cumulated pension points are likely correlated with age, we split the sample according to cumulated pension points by cohort. As a result, we define two groups of mothers, one with cumulated non-childcare pension points above the median in their cohort, and one below the median. Panel C of Appendix Table E1 shows that women with above-median pension points experience larger reductions in labor supply at the intensive margin. In Panel D of Appendix Table E1, we focus on couples that we can identify in the data. We consider the sum of both partners' pre-reform cumulated non-childcare pension points as a proxy for household wealth and investigate heterogeneous responses along this dimension. Here the above versus below median split is within partners' cohort. Consistent with our results on heterogeneous responses by own pension wealth, we find qualitatively larger responses at the intensive margin for individuals with higher household pension wealth. The point estimates are however insignificant. Overall, these results point to larger reductions in labor supply by individuals who have larger wealth to start with.

E.3. Liquidity Constraints

Labor supply responses to the pension wealth change may also differ across individuals who face different liquidity constraints. Since for the majority of workers, labor earnings are the

¹²This result helps qualify the evidence on heterogeneity by returns to tenure described in the previous paragraph. To the extent that higher predicted returns to tenure could be driven by lower effort costs, our results in Panel A of Table Appendix E1 might partly reflect variation in costs. The evidence in Panel B of the same table would seem to exclude that this is the case.

main source of income for consumption, we analyze heterogeneity in labor supply responses by the level of partners' labor earnings. We exclude from the sample couples in which the male spouse cannot be observed in the IEB data during the second half of 2013.¹³ We then split the sample of mothers by partners' median earnings within partners' birth cohorts. The results reported in Panel E of Appendix Table E1 show that intensive margin responses are only visible among mothers with above-median partners' labor earnings, suggesting that liquidity constraints limit the ability to reduce labor supply among mothers married to less well-earning husbands.

E.4. Age at the Time of the Reform

We then consider whether different ages at the time of the reform lead to differential responses to the pension wealth shock. The labor supply response to a pension wealth shock might be larger for individuals at older ages, for example due to the cost of working being higher and the time to reach the full retirement age being shorter. In Panel F of Appendix Table E1, we report estimates of heterogeneous responses by age as of December 2013. The results do not reveal differential responses by age. We should note, however, that age at the time of the reform is collinear with age at the time of the first childbirth and with the number of years to the full retirement age. This has two implications. First, it implies that we cannot separate out the impact of age from that of time to retirement. Second, heterogeneity by age in December 2013 might also be capturing differences in the degree of labor market attachment. Mothers who are older in 2013 had their first child relatively late, possibly indicating a stronger attachment to the labor market. This could counteract a potential, larger labor supply response by individuals who are older and closer to the full retirement age.

In Appendix Figure E1, we test the parallel trends assumption for all estimates reported in Appendix Table E1. Each panel in the figure corresponds to the same-letter panel in the table. For each outcome Y_{it} , we regress Y_{it} against a time trend, its interaction with a treatment indicator, and the interaction of both those terms with an indicator taking value one if the observation belongs to the above-median group of the relevant dimension of heterogeneity. Individual fixed effects are also included. The estimation sample is restricted to the pre-reform period. In Appendix Figure E1, estimates in black (dots) refer to the coefficient on the interaction between the trend and the treatment indicator, while estimates in gray (crosses) to the interaction between the trend, the treatment indicator and the above-median indicator. As can be appreciated from the graphs, pre-reform differential trend estimates are all statistically insignificant, the only exception being the estimates for the probability of being employed in Panel A, which are marginally significant.

¹³This sample restriction is motivated by the fact that we cannot distinguish non-employment from self-employment or civil service, and it reduces the sample by 10%. Results are qualitatively similar, but less precisely estimated, if we retain all couples and impute zero earnings to partners who are not in the IEB data in December 2013.

Table E1. HETEROGENEITY ANALYSIS

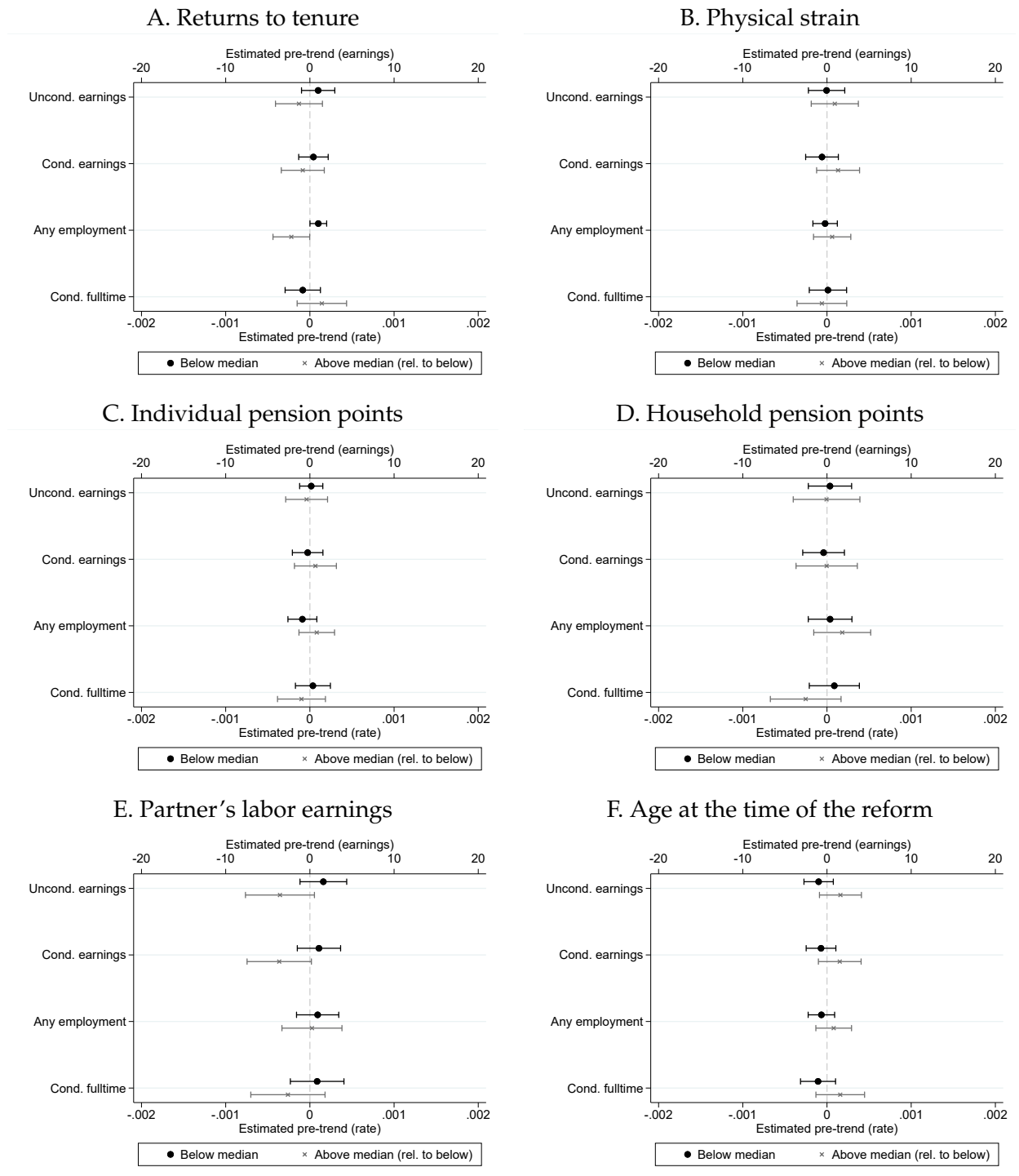
	Unconditional earnings (1)	Any employment (2)	Conditional earnings (3)	Cond. fulltime employment (4)
<i>Panel A. Returns to tenure in pre-reform occupation</i>				
T × Post	-106.60 (41.46)	-0.004 (0.002)	-60.75 (31.66)	-0.008 (0.003)
T × Post × above median	56.87 (54.76)	0.001 (0.003)	3.16 (44.16)	0.007 (0.004)
Pre-reform mean of dep. var. (below median)	13,997.40	0.985	13,997.40	0.396
Pre-reform mean of dep. var. (above median)	7,112.69	0.949	7,112.69	0.165
N individuals	73,178	73,178	73,139	73,139
<i>Panel B. Physical strain of pre-reform occupation</i>				
T × Post	-44.57 (43.10)	-0.000 (0.002)	-49.79 (34.14)	-0.005 (0.003)
T × Post × above median	-45.55 (54.90)	-0.004 (0.003)	-4.74 (44.09)	0.001 (0.004)
Pre-reform mean of dep. var. (below median)	12,596.27	0.975	12,596.27	0.328
Pre-reform mean of dep. var. (above median)	8,575.82	0.959	8,575.82	0.234
N individuals	72,826	72,826	72,787	72,787
<i>Panel C. Individual pre-reform non-childcare pension points</i>				
T × Post	-53.99 (25.62)	-0.005 (0.002)	-64.11 (32.38)	-0.000 (0.003)
T × Post × above median	-32.60 (44.70)	0.004 (0.003)	0.70 (43.79)	-0.007 (0.004)
Pre-reform mean of dep. var. (below median)	3,074.85	0.561	5,211.72	0.119
Pre-reform mean of dep. var. (above median)	12,579.63	0.878	13,990.00	0.383
N individuals	99,104	99,104	82,876	82,876
<i>Panel D. Household pre-reform non-childcare pension points</i>				
T × Post	-68.77 (46.72)	-0.004 (0.004)	-86.25 (44.14)	-0.001 (0.004)
T × Post × above median	-46.15 (72.08)	0.005 (0.005)	-8.43 (64.70)	-0.007 (0.005)
Pre-reform mean of dep. var. (below median)	6,315.00	0.794	7,679.81	0.196
Pre-reform mean of dep. var. (above median)	10,766.48	0.876	12,018.78	0.269
N individuals	37,427	37,427	34,962	34,962

Table E1 (cont.): HETEROGENEITY ANALYSIS

	Unconditional earnings (1)	Any employment (2)	Conditional earnings (3)	Cond. fulltime employment (4)
<i>Panel E. Partner's pre-reform labor earnings</i>				
T × Post	-31.87 (50.94)	-0.005 (0.004)	-22.30 (45.85)	0.001 (0.004)
T × Post × above median	-90.92 (74.57)	0.007 (0.005)	-119.19 (67.88)	-0.010 (0.006)
Pre-reform mean of dep. var. (below median)	8,280.42	0.837	9,593.91	0.250
Pre-reform mean of dep. var. (above median)	8,718.33	0.845	10,032.05	0.201
N individuals	33,605	33,605	31,563	31,563
<i>Panel F. Age at the time of the reform</i>				
T × Post	-78.09 (30.92)	-0.002 (0.002)	-71.49 (30.64)	-0.005 (0.003)
T × Post × above median	14.32 (44.67)	0.001 (0.003)	16.47 (44.06)	0.002 (0.004)
Pre-reform mean of dep. var. (below median)	7,362.87	0.729	9,720.26	0.281
Pre-reform mean of dep. var. (above median)	8,149.95	0.705	11,200.90	0.272
N individuals	99,104	99,104	82,876	82,876

Notes: The table reports estimates from versions of equation (2), where we interact the term $D_i \cdot \mathbb{I}[t > \text{Jul-Dec 2013}]$ with indicators capturing heterogeneous treatment effects. The coefficient γ_s is estimated pooling all post-reform years together. Robust standard errors clustered at the individual level are reported in parentheses. In Panel A, we interact the term $D_i \cdot \mathbb{I}[t > \text{Jul-Dec 2013}]$ with a dummy for being employed in an occupation with above-median predicted returns to tenure in the second half of 2013; in Panel B, with a dummy for being employed in an occupation with above-median index of physical strain in the second half of 2013; in Panel C, with a dummy for having above median non-childcare pension points within birth cohort in December 2013; in Panel D, with a dummy for having above median household-level non-childcare pension points within partners' birth cohort in December 2013; in Panel E, with a dummy for having above-median partners' labor earnings within partners' birth cohort in the second half of 2013; and in Panel F, a dummy for having above-median age as of December 2013.

Figure E1. TEST OF PARALLEL TRENDS FOR HETEROGENEITY ANALYSIS



Notes: The graph reports estimates of pre-reform trends for our main outcomes of interest across the heterogeneity dimensions investigated in Appendix Table E1. For each outcome Y_{it} , we regress Y_{it} against a time trend, its interaction with a treatment indicator, and the interaction of both those terms with an indicator taking value one if the observation belongs to the above-median group of the relevant dimension of heterogeneity. Individual fixed effects are also included. The estimation sample is restricted to the pre-reform period. The dots/crosses represent the point estimates and the capped vertical bars the associated 95% confidence intervals based on robust standard errors clustered at the individual level. Estimates in black (dots) refer to the coefficient on the interaction between the time trend and the treatment indicator, while estimates in gray (crosses) to the interaction between the trend, the

treatment indicator and the above-median indicator. Estimates for unconditional and conditional earnings are plotted along the top x-axis, while those for the employment rate and the conditional full-time rate on the bottom x-axis. Each panel corresponds to the same-letter panel in Appendix Table [E1](#).

Appendix F. Intra-household Spillover Effects

Our ability to match the working histories of spouses allows us to explore behavioral spillovers in labor supply choices within the household.

We first analyze whether mothers' labor supply responses differ by their partner's age at the time of the reform. In December 2013, partners in the sample are on average 51.5 years old and are thus approximately 15 years away from the full retirement age.¹⁴ As reported in Panel A of Appendix Table F1, we find that the intensive-margin labor supply responses of women in our sample are driven by mothers whose partner is older, consistent with within-household interactions in labor supply choices. Estimates are large in magnitude but not statistically significant. To the extent that age is correlated across spouses, heterogeneity by partner's age could also reflect heterogeneity by own age. Since, in our setting, one's own age does not seem to matter for labor supply responses (Panel F of Appendix Table E1), we believe that the heterogeneity reported in Panel A of Appendix Table F1 is likely to reflect genuine variation by partner's age. The estimates reported in Panel A of Appendix Figure F1 provide supporting evidence for the parallel trends assumption.

We can also investigate whether the pension wealth shock generated by the 2014 reform has spillover effects on partners' labor supply choices. To this end, we estimate equation (2) on the sample of male spouses whom we can match to our original sample of women. The treatment indicator D_i takes value one for the male spouses of women in our treatment group, and zero for the male spouses of women in our control group. Panel B of Appendix Table F1 reports the estimated coefficients. All point estimates are negative, suggesting that male spouses also reduce their labor supply, with effects that become more pronounced in the longer run. For total unconditional earnings, we estimate an effect size that is approximately half of the one found among women in absolute terms (see Table 1).¹⁵ Also in the case of male spouses, the effect appears to be driven by reductions in employment at the intensive margin. However, due to the small sample size, most coefficients are imprecisely estimated, with the exception of the negative coefficient on the likelihood of working full time conditional on employment in the medium run, which is statistically significant. We test the common trends assumption for partner's outcomes in Panel B of Appendix Figure F1. Here, we regress each outcome Y_{it} against a time trend, its interaction with a treatment indicator and individual fixed effects, restricting the sample to the pre-reform period, and plot the estimated coefficient on the interaction. We detect a marginally significant differential pre-trend for unconditional earnings only (driven by effects in 2010 and early 2011), with all remaining outcomes passing the pre-trends test.

Even though we acknowledge that spousal earnings responses are imprecisely estimated, we provide here calculations of the household-level MPE implied by our estimates of mothers' and

¹⁴Less than 1% of partners had reached the full retirement age in December 2013.

¹⁵We note here that, even though we can match only around 40% of women in our main sample to a partner, the effect of the reform on total unconditional earnings is comparable in the full sample of mothers and the sample of married mothers for whom we can identify a partner. The latter respond more strongly in the short run (EUR -85 versus EUR -51), but have a similar medium-run response (EUR -100 versus EUR -98); see Appendix Table F2.

partners' responses. Taking those estimates at face value, discounted earnings drop by EUR 612 for mothers and EUR 301 for partners over the 5 years after reform. Rescaling their sum by $5/17$ of the pension wealth gain, we obtain a household-level MPE of 0.81.

Table F1. HOUSEHOLD-LEVEL SPILLOVERS

	Unconditional earnings (1)	Any employment (2)	Conditional earnings (3)	Cond. fulltime employment (4)
<i>Panel A. Heterogeneity by partner's age at the time of the reform</i>				
T × Post	-40.35 (50.19)	0.000 (0.003)	-62.59 (46.12)	-0.000 (0.004)
T × Post × above median	-97.45 (72.06)	-0.004 (0.005)	-55.04 (65.09)	-0.008 (0.005)
Pre-reform mean of dep. var. (below median)	7,894.84	0.829	9,220.95	0.236
Pre-reform mean of dep. var. (above median)	9,145.94	0.840	10,614.93	0.233
N individuals	37,427	37,427	34,962	34,962
<i>Panel B. Partner's labor supply response</i>				
T × Post (yr 1-2.5)	-19.49 (48.46)	0.001 (0.002)	-49.84 (32.05)	-0.001 (0.001)
T × Post (yr 3-5)	-44.49 (74.15)	0.000 (0.003)	-77.30 (47.75)	-0.004 (0.002)
Pre-reform mean of dep. var.	19,982.79	0.844	23,177.00	0.916
N individuals	37,427	37,427	34,604	34,604

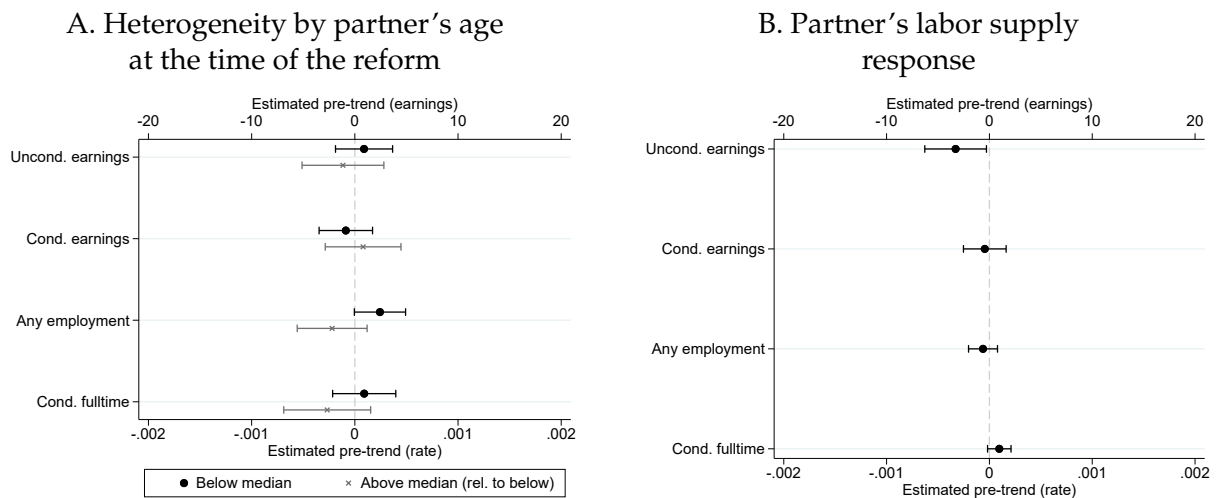
Notes: Panel A reports estimates from versions of equation (2), where we interact the term $D_i \cdot \mathbb{I}[t > \text{Jul-Dec 2013}]$ with a dummy taking value one for individuals whose partner has above-median age as of December 2013. The coefficient γ_s is estimated pooling all post-reform years together. Panel B reports estimates of the coefficients γ_s from equation (2) for post-reform years using the sample of male partners. The coefficients γ_s are estimated separately for the 1-2.5 years and the 3-5 years after the reform. Robust standard errors clustered at the individual level are reported in parentheses.

Table F2. MAIN ESTIMATES FOR SAMPLE OF MOTHERS MATCHED TO A PARTNER

	Unconditional earnings (1)	Any employment (2)	Conditional earnings (3)	Cond. fulltime employment (4)
T × Post (yr 1-2.5)	-84.52 (31.49)	-0.003 (0.002)	-77.14 (28.69)	-0.004 (0.003)
T × Post (yr 3-4.5)	-100.14 (45.18)	-0.001 (0.003)	-107.45 (40.60)	-0.005 (0.003)
Pre-reform mean of dep. var.	8,535.93	0.835	9,937.45	0.234
N individuals	37,427	37,427	34,962	34,962

Notes: The table reports estimates of the coefficients γ_s from equation (2) for post-reform years using the sample of mothers who can be matched to a partner. The coefficients γ_s are estimated separately for the 1-2 years and the 3-4.5 years after the July 2014 reform. Robust standard errors clustered at the individual level are reported in parentheses.

Figure F1. TEST OF PARALLEL TRENDS FOR HOUSEHOLD SPILLOVERS



Notes: Panel A reports estimates of pre-reform trends for our main outcomes of interests across the heterogeneity dimension investigated in Panel A of Appendix Table F1. For each outcome Y_{it} , we regress Y_{it} against a time trend and its interaction with an indicator taking value one if the observation belongs to the group with above-median partner's age as of December 2013. Individual fixed effects are also included. The estimation sample is restricted to the pre-reform period. The dots/crosses represent the point estimates and the capped vertical bars 95% confidence intervals based on robust standard errors clustered at the individual level. Estimates in black (dots) refer to the coefficient on the interaction between the time trend and the treatment indicator, while estimates in gray (crosses) to the interaction between the trend, the treatment indicator and the above-median indicator. Panel B reports estimates of pre-reform trends for the outcomes analyzed in Panel B of Appendix Table F1. For each outcome Y_{it} , we regress Y_{it} against a time trend, its interaction with a treatment indicator and individual fixed effects, restricting the sample to the pre-reform period. The dots represent the point estimates on the interaction and the capped vertical bars the associated 95% confidence intervals based on robust standard errors clustered at the individual level. Estimates for unconditional and conditional earnings are plotted along the top x-axis. Estimates for the employment rate and the conditional full-time rate are plotted on the bottom x-axis.

References

- Becker, Sebastian, Hermann Buslei, Johannes Geyer, and Peter Haan.** 2022. "The Effect of Pension Wealth on Employment." Institute of Labor Economics (IZA) IZA Discussion Papers 15836.
- Beye, Claus-Jürgen.** 2009. *Gesetzliche Rentenversicherung als Teil der Altersvorsorge*. C.F. Müller Verlag, Heidelberg.
- Bächmann, Ann-Christin, Corinna Frodermann, Benjamin Lochner, Michael Oberfichtner, and Simon Trenkle.** 2021. "Identifying Couples in Administrative Data for the Years 2001–2014." FDZ FDZ-Methodenreport 03/2021 (en).
- Deutsche Rentenversicherung Bund.** 2020. "Rentenversicherung in Zeitreihen." Deutsche Rentenversicherung Bund, Berlin.
- Deutsche Rentenversicherung Bund.** 2021a. "Rente: So wird sie berechnet - alte Bundesländer." Deutsche Rentenversicherung Bund, Berlin.
- Deutsche Rentenversicherung Bund.** 2021b. "Rentenrechtliche Zeiten." Deutsche Rentenversicherung Bund, Berlin.
- Deutsche Rentenversicherung Bund.** 2022. "Arbeitslos - was Sie beachten sollten." Deutsche Rentenversicherung Bund, Berlin.
- Giavazzi, Francesco, and Michael McMahon.** 2008. "Policy Uncertainty and Precautionary Savings." National Bureau of Economic Research Working Paper 13911.
- Giupponi, Giulia, and Arthur Seibold.** 2024. *Rethinking Pension Reform*. CEPR Press.
- Institute for Structural Research.** 2022. "Occupation Classifications Crosswalks – From O*NET-SOC to ISCO." <https://ibs.org.pl/en/resources/occupation-classifications-crosswalks-from-onet-soc-to-isco/>.
- Institut für Arbeitsmarkt- und Berufsforschung.** 2021a. "2%-Stichprobe der Integrierten Erwerbsbiografien (IEB) stp2pct_V15.00_00.201912."
- Institut für Arbeitsmarkt- und Berufsforschung.** 2021b. "Paardatensatz."
- Minijob-Zentrale.** 2020. "Aktuelle Entwicklungen im Bereich der Minijobs. 1. Quartalsbericht 2020." Deutsche Rentenversicherung Knappschaft-Bahn-See, Essen.
- National Center for O*NET Development.** 2022a. "Browse by Work Context. O*NET OnLine." <https://www.onetonline.org/find/descriptor/browse/4.C>.
- National Center for O*NET Development.** 2022b. "Crosswalk 2010 to 2019. O*NET Resource Center." <https://www.onetcenter.org/taxonomy/2019/walk.html>.
- Schönberg, Uta, and Johannes Ludsteck.** 2014. "Expansions in Maternity Leave Coverage and Mothers' Labor Market Outcomes after Childbirth." *Journal of Labor Economics*, 32(3): 469–505.
- Statistisches Bundesamt.** 2022. "GENESIS-Online Database: Table 12621-0003." <https://www-genesis.destatis.de/datenbank/online/url/aab71b9f>, Accessed: 2022-03-25.
- Statistisches Bundesamt.** 2025. "GENESIS-Online Database: Table 12621-0003." <https://www-genesis.destatis.de/datenbank/online/statistic/12621/table/12621-0003/table-toolbar>, Accessed: 2025-12-19.
- Steffen, Johannes.** 2011. "Arbeitslosigkeit und Rente." Arbeitnehmerkammer Bremen.