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

Elisabeth Artmann† Nicola Fuchs-Schündeln‡ Giulia Giupponi§

November 20, 2023

Abstract

We provide new evidence of forward-looking labor supply responses to changes in pension wealth. We exploit a 2014 German reform that increased pension wealth for mothers by an average of 4.4% per child born before January 1, 1992. Using administrative data on the universe of working histories, we implement a difference-in-differences design comparing women who had their first child before versus after January 1, 1992. We document significant reductions in labor earnings, driven by intensive margin responses. Our estimates imply that, on average, an extra euro of pension wealth in a given period reduces unconditional labor earnings by 54 cents.
Keywords: labor supply, social security, pension wealth
JEL Codes: H55, J22, J26

*We thank Camille Landais, Nathaniel Hendren and Attila Lindner for helpful comments. We also thank participants in seminars at Bocconi University and Columbia University, conference participants at ESPE 2023, EEA 2023, EALE 2023 and the 2023 meeting of the Verein für Socialpolitik, and workshop participants at the BI Norwegian Business School, the University of Cambridge, Central European University, the CEPR Public Economics Annual Symposium, CESifo, Erasmus University Rotterdam, the European Labor Symposium for Early Career Economists, the FDZ-RV, the Forschungsnetzwerk Alterssicherung, the IAB, the NBER Summer Institute, the University of Cologne, the University of Oslo and the Social Policy Workshop of the Ausschuss für Sozialpolitik for helpful discussions and suggestions. We are grateful to the staff at the Research Data Center of the German Pension Insurance (DRV) for helpful assistance with the VSKT data. Celina Proffen provided excellent research assistance. This paper uses administrative data sampled from the Integrated Employment Biographies (IEB) data of the Institute for Employment Research (IAB). The data are social data with administrative origin which are processed and kept by IAB according to the German Social Code, Book III. Fuchs-Schündeln and Giupponi gratefully acknowledge financial support from the Forschungsnetzwerk Alterssicherung. Fuchs-Schündeln additionally gratefully acknowledges financial support from the German Research Foundation through a Gottfried Wilhelm Leibniz-Prize, and from the ERC through Consolidator Grant No. 815378. Giupponi additionally gratefully acknowledges financial support from Bocconi University, Baffi CAREFIN and the Steven H. Sandell Grant Program sponsored by the Center for Retirement Research at Boston College and funded by the U.S. Social Security Administration.
†Institute for Employment Research (IAB). Email: elisabeth.artmann2@iab.de
‡Goethe University Frankfurt and CEPR. Email: fuchs@wiwi.uni-frankfurt.de
§Bocconi University and CEPR. Email: giulia.giupponi@unibocconi.it

1. Introduction

Demographic trends have increased the potential labor supply of the elderly, but are simultaneously exerting rampant pressure on the fiscal sustainability of pay-as-you-go public pension systems. To counteract these trends, governments have implemented – or are planning to implement – different types of pension reforms.[1] A large literature investigates how the design of pension systems affects the labor supply choices of individuals close to retirement. However, if individuals are forward-looking, pension reforms can affect behavior also at younger ages, with potential aggregate labor supply implications. For example, more generous systems might induce workers to reduce their lifetime labor supply by working fewer hours. In light of the demographic and economic trends common to most developed economies, understanding the size and nature of these forward-looking effects has become increasingly important.

In this paper, we study how changes in the generosity of public pension systems affect labor supply behavior far from retirement. In particular, we provide new estimates of forward-looking labor supply responses to changes in pension wealth. Knowledge of income (or wealth) effects is key for the evaluation of pension policies and other old-age income support schemes. In addition, income effects are important for welfare analyses and normative assessments of the distributional effects of retirement policies, since they are directly related to the insurance value of social security (Chetty, 2004; 2008; Kolsrud et al., 2021).

The identification of forward-looking labor supply responses to changes in pension wealth is challenging for various reasons. First, separately identifying income and substitution incentives is notoriously difficult, since most policy changes conflate the two.[2] Secondly, the gradual phase-in of reforms implies that they typically generate small discontinuities, which are unlikely to trigger meaningful responses. Third, individuals far from retirement are usually all identically treated, leaving little room for exogenous variation in incentives. Finally, changes in benefit generosity are generally achieved through changes in pension rules or formulae, and it is hard to isolate behavioral responses that are not attenuated by cognitive or information frictions.

This paper contributes to our understanding of how changes in pension wealth affect labor supply in middle-age years by overcoming these limitations. We leverage unique variation stemming from a German pension reform to the pension scheme popularly known as 'Mütterrente' (mothers' pension). The reform exogenously changed the amount of pension contributions credited to mothers for the time spent raising their children. From July 2014, mothers of children born before January 1, 1992 had their pension contributions increased by an amount corresponding to 4.4% of average pension wealth per child; the pension contributions of mothers of children

[1]Pension reforms typically entail a combination of a tightening of the contribution-benefit link, a reduction in the overall generosity of the system, and an increase of statutory retirement ages.

[2]Most social security reforms entail both changes in benefit levels and in the net marginal returns to additional earnings, making it difficult to separate income (or wealth) and substitution effects. For example, the well-known Social Security Notch, which led to large cuts in Old Age and Survivors Insurance (OASI) benefits for individuals born after 1917 in the United States, decreased both the level of pension benefits (income effect) and the returns to additional earnings (substitution effect); see Gelber, Isen and Song (2016).

born on or after January 1, 1992 were left unchanged. The timing and nature of the Mütterrente reform generate a pure pension wealth effect on affected mothers, which is permanent, is not confounded by changes in substitution incentives, and is sizable when compared to pension reforms in other countries. Importantly, the Mütterrente reform did not alter the pension benefit formula. If leisure is a normal good, the increase in pension wealth is expected to reduce the lifetime labor supply of recipients.

Using administrative data from the Integrated Employment Biographies (IEB) of the German Institute for Employment Research (IAB) covering the universe of employment histories in the private sector, we select women who gave birth to their first child in a small window around January 1, 1992. We analyze their labor supply responses to the change in pension wealth induced by the Mütterrente reform of July 2014 over the five years after the policy was enacted, when the affected women were on average aged 50 to 55. To identify the effect of interest, we implement a difference-in-differences design and compare the employment dynamics of women who had their first child just before (*treatment group*) and just after (*control group*) January 1, 1992, from before to after the 2014 reform.

We document significant reductions in total labor earnings unconditional on employment among treated mothers following the 2014 reform. In the 2.5 years following the policy change, unconditional labor earnings drop on average by EUR 101 per year or 0.7% of pre-policy levels. Consistent with the notion that optimization frictions may limit the ability of individuals to adjust their labor supply in the short run, we find larger medium-run responses – of EUR 196 per year or 1.3% of the pre-policy mean – in the 3-5 years after the reform. To interpret the magnitude of these reduced-form effects, we rescale the estimated medium-run earnings response by the pension wealth shock due to the 2014 reform. For an average increase in the present discounted value of pension wealth of EUR 3,830 or 4.4%, the implied pension wealth effect is EUR -5.1 per year per EUR 100 of wealth, and the implied pension wealth elasticity is -0.3.[3] The size of the pension wealth effect that we describe above can be hard to assess, since it does not account for the time horizon over which individuals plausibly allocate the increase in pension wealth. In order to better gauge the size of the wealth effect, we adopt the annuitization method described in Golosov et al. (2021) and compute a marginal propensity to earn (MPE) out of pension wealth far from retirement. Our estimates imply a moderately large MPE of -0.54, which is in line with recent estimates of MPEs out of lottery wealth by Golosov et al. (2021) (-0.5) and out of social security income by Gelber, Isen and Song (2016; 2017) (-0.6 for men and -0.9 for women). These papers, however, focus on current rather than future wealth changes.

We probe the anatomy of the labor supply response by decomposing the change in unconditional labor earnings into its extensive and intensive margin components. We find that the response is entirely concentrated along the intensive margin: in the medium run, labor earnings conditional on employment drop by 0.9%, and the probability of working full-time (as opposed to working

[3] As we explain in more detail in Section 3, we use a second administrative data source, the Insurance Account Sample Versicherungskontenstichprobe (VSKT) from the German Pension Insurance (DRV), to calculate the change in pension wealth generated by the Mütterrente reform.

part-time or in marginal employment) by 1.8%.[4] By contrast, the probability of being employed is not affected by the reform.

The richness of the administrative data allows us to investigate various dimensions of heterogeneity. First, we consider whether individuals with different labor market returns to tenure and costs of working respond differently to the wealth shock. Consistent with forward-looking behavior, individuals with higher predicted returns to tenure are less likely to leave full-time employment. On the other hand, work disutility – measured by how physically demanding an individual's occupation is – does not seem to matter substantially for the size and nature of the response. Second, we show that the intensive margin response is stronger among those who, within their birth cohort, have higher pre-reform pension wealth, which can be viewed as a proxy for the income an individual expects to rely upon once retired. Third, we show that labor supply responses are muted among those who are likely to face stronger liquidity constraints, as proxied by their partner's pre-reform labor earnings.

For approximately 40% of our sample, we can match mothers to their spouses. Thanks to the possibility of jointly observing their working histories, we can explore the presence of behavioral spillovers across spouses. We find some evidence that male spouses also reduce their labor supply along the intensive margin, with effects that become quantitatively more pronounced in the medium run, are approximately half in magnitude compared to those of women, but are not always precisely estimated due to the small sample size. We also show that the labor supply responses of women are amplified when their partner is closer to the full retirement age. Both these pieces of evidence are consistent with some degree of within-household interaction in labor supply choices (Goux, Maurin and Petrongolo, 2014; Carta and De Philippis, 2021).

A large literature has investigated the effects of social security and other government old-age support programs on labor supply close to retirement (for reviews, see Gruber and Wise, 1999; Krueger and Meyer, 2002; Blundell, French and Tetlow, 2016). Instead, only a small number of papers have analyzed labor supply responses to changes in pension benefits far from retirement. Among these, existing studies examine the role of changes in the statutory retirement age (Hairault, Sopraseuth and Langot, 2010; Carta and De Philippis, 2021), the contribution-benefit link (French et al., 2022), and pension rules more generally (Bovini, 2019). Our paper contributes to this emerging strand of the literature by providing new evidence on 'forward-looking' labor supply responses to permanent changes in pension wealth.[5] In contemporaneous work, Becker et al. (2022) use data from the 2% Insurance Account Sample Versicherungskontenstichprobe (VSKT) from the German Pension Insurance (DRV) to analyze the effect of the 2014 Mütterrente reform on the probability of employment. Comparing women who gave birth to a child in the two-year window before or after the January 1, 1992 cut-off, they document a 2% drop in participation, corresponding to a pension wealth elasticity of participation of -0.22. The effect

[4]Marginal employment – also called minijob – is a low-wage contract with monthly earnings below given thresholds. Employees in marginal employment are exempt from employee social security contributions.

[5]Other work has looked at savings responses to pension wealth changes. See, for instance, Attanasio and Brugiavini (2003); Attanasio and Rohwedder (2003); Bottazzi, Jappelli and Padula (2006); Lachowska and Myck (2018).

in their paper is driven by mothers with two or more eligible births in the treatment period. In contrast to their paper, we focus on a much more homogeneous group by analyzing only women who gave birth to their first child, and by restricting the births to a bandwidth of 3 months rather than 2 years around the cut-off date of January 1, 1992. We do not find any employment effects, but significant intensive margin effects.

This paper is also related to work that attempts to separately identify income/wealth effects and substitution incentives in the context of retirement benefits (e.g., Costa, 1997; French, 2005; Liebman, Luttmer and Seif, 2009; Danzer, 2013; Gelber, Isen and Song, 2016; Fetter and Lockwood, 2018; Gelber, Jones and Sacks, 2020) and other welfare transfers (Deshpande, 2016; Giupponi, 2019; Jones and Marinescu, 2022), using quasi-experimental or structural approaches. Our paper complements this literature in three ways. First, our setting allows us to compellingly identify the wealth effect of retirement income through quasi-experimental policy variation, rather than through the estimation or simulation of structural models that typically require strong functional-form assumptions. Second, thanks to the nature of the policy variation that we exploit, we can provide a point estimate of the wealth effect, thus complementing previous work that set bounds on it.[6] In this respect, our parameter estimates are similar in nature to those that can be obtained studying lottery winnings (Imbens, Rubin and Sacerdote, 2001; Cesarini et al., 2017; Picchio, Suetens and van Ours, 2018; Golosov et al., 2021). Finally, in contrast to the papers cited above, we identify a forward-looking response, as opposed to a response to contemporaneous (or a mix of contemporaneous and future) wealth changes.

Finally, this paper complements a literature that has predominantly examined men's labor supply, by placing its contextual focus on female labor supply in the early twenty-first century. As recent demographic trends have given greater significance to the aggregate potential of female labor force at middle ages, it is ever more important to understand how the design of pension systems affects the labor supply decisions of this group.

The paper proceeds as follows. Section 2 outlines the institutional details of the German pension system and the 2014 Mütterrente reform. Section 3 describes the data used in the analysis. Section 4 discusses the expected effects of the reform on individual labor supply and illustrates the empirical strategy. Estimates of the pension wealth effect on labor supply are presented in Section 5. Section 6 concludes.

[6]This is the case, for example, in Deshpande (2016), Giupponi (2019), and Gelber, Isen and Song (2016; 2017).

2. Institutional Context

2.1 The German Pension System

The German statutory public pension system is an earnings-related Pay-As-You-Go (PAYG) system.[7] Pension payments are determined by a formula based on cumulated pension contributions as measured by pension points, the monetary value assigned to each pension point, and an adjustment factor based on age at the time of retirement.[8] One pension point represents the annual pension contributions made by a reference contributor earning the average income in a given year. The amount of the monthly pension is calculated as the total number of cumulated pension points multiplied by a monthly 'pension-point value', i.e. the monthly monetary equivalent of a pension point. The value of a pension point is set annually by the government in relation to gross wage growth and is thus independent of individual conditions (e.g. individual earnings or contributions).

Eligibility for retirement pension is conditional on a minimum qualifying period of 5 years of contributions. Childcare periods count as qualifying periods. For the majority of individuals in our sample, the full retirement age is 67.[9] Individuals can retire at most four years earlier under the early retirement regulations, which require 35 years of contributions and imply a 3.6% reduction in pension entitlements for each year of retirement before the full retirement age.[10]

In Germany, individuals insured with the public pension system are regularly informed about their pension entitlements. Since 2002, the German Pension Insurance automatically sends information about pension entitlements on an annual basis to all insured persons aged 27 and over who have acquired at least five years of contribution periods. This information includes the amount of cumulated pension contributions and the projected level of old-age pension one would obtain if the individual kept contributing as in the previous five years until the full retirement age.[11]

[7] Pensions from the statutory public pension system are the main form of retirement income in Germany. In 2019, about 91% of West German women aged 65 or older received payments from the statutory pension system. The average gross pension entitlement per recipient amounted to EUR 1,037. Only around 28% of women received retirement income from company-provided pension plans, averaging EUR 357 per recipient (Deutsche Rentenversicherung Bund, 2022*b*).

[8] In contrast to old-age pensions, disability pensions are subject to a second adjustment factor.

[9] More precisely, the full retirement age was 65 years and 6 months for women born in 1952, the oldest cohort in our sample. It progressively increased for subsequent cohorts, reaching 67 for women born in or after 1964. The average woman in our sample was born in 1964.

[10] An additional pathway into retirement is the seniority pension. Eligibility for seniority pension requires 45 years of contributions and 63 years of age for individuals born before 1953. The seniority pension age threshold increases step-wise to 65 for subsequent cohorts. Only a small fraction of individuals in our sample are likely to satisfy the seniority pension requirements (Börsch-Supan, Coppola and Rausch, 2015).

[11] See Figure A7 in Seibold (2021) for an illustration of the information letter (*Renteninformation)*.

2.2 The Childcare Contribution Benefit 'Mütterrente'

In Germany, parents are credited pension contributions for the time spent raising their children under the Mütterrente childcare pension benefit scheme. This childcare pension benefit was first introduced in 1986. Since then, parents have been credited pension points for certain periods after childbirth, the generosity and length of the benefits having changed over time. Whilst, in principle, either fathers or mothers are entitled, childcare pension points are by default attributed to mothers and only assigned to fathers on the mother's request. Around 98% of the pension benefits go to mothers rather than fathers (Deutsche Rentenversicherung Bund, 2014).

The childcare benefit scheme has been subject to various reforms, which are visually summarized in Appendix Figure A1.[12] 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 this first year after childbirth. For employed mothers, childcare pension points were fully withdrawn against pension points from regular employment. Hence, if mothers earned half of a pension point through work during the first year after childbirth, they would be credited only 0.25 pension points as childcare benefit.

The 1992 Pension Reform Act, which was passed in December 1989, introduced an important change to the Mütterrente. Benefit duration increased from one to three years, but only for mothers of children born on or after January 1, 1992. These mothers were thus entitled to a maximum of 2.25 pension points (three years $\times$ 0.75 pension points) instead of a maximum of 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 introduced two main changes: (i) it increased the generosity of the childcare pension benefit from 0.75 to one pension point per year; (ii) it repealed the earnings test. Consequently, pension contributions from childcare were no longer withdrawn against those from employment.[13] The 1999 pension reform was retroactive, hence both current retirees and individuals with legal rights to future pension payments benefited from it.

We explore in this paper the reform implemented in 2014. Starting July 2014, mothers whose children were born before January 1, 1992, got one extra pension point per child, totaling two pension points per child born before 1992. In 2014, the accreditation of one additional pension point was equivalent to a EUR 336/312 higher pension per year in the West/East, respectively. As we describe in Section 5.2, for the average mother in our sample, who was 27 at the time

[12] See also Thiemann (2015).

[13] This is true as long as the sum of childcare and employment contributions does not exceed the 'contribution ceiling'. The contribution ceiling varies over time. It is adjusted each year by the German government, mostly based on the gross salary growth rate. The contribution ceiling is roughly equivalent to twice the average income in the country and is defined separately for East and West Germany. In 2019, for example, the contribution ceiling was EUR 80,400 in West Germany and EUR 73,850 in East Germany. In our estimation sample, only 0.7% of women ever hit the contributory ceiling in the three years after childbirth.

of her first childbirth and close to 50 at the time of the reform, one additional pension point implied an increase in real pension wealth of approximately EUR 3,830 at the time of the reform. This increase corresponds to 4.4% of pension wealth, a sizable effect compared to other changes analyzed in the literature. By comparison, the US Social Security 'Notch' – the largest discontinuous change in Old Age and Survivors Insurance benefits in the US – generated a reduction in average lifetime discounted benefits of 2% for women and of 5% for men (Gelber, Isen and Song, 2016; 2017).[14] At the time of reform, individualized letters were sent to affected mothers informing them of the accreditation of additional childcare periods for each of their eligible children (the letter was sent separately from the regular yearly information letter we mentioned before). Moreover, the German Pension Insurance (DRV) provided plenty of easily accessible information on the reform and its implications on its website.

The 2014 reform – popularly known as 'Mütterrente' reform – was an item of the Christian Democratic Union (CDU) throughout the electoral campaign for the 2013 elections. The main narrative of the CDU was to attenuate the unequal treatment of mothers with children born before or after January 1, 1992. On September 22, 2013, the CDU won the election by a 16 percentage point margin and subsequently formed a coalition with the Social Democratic Party (SPD). The Mütterrente reform became a central part of the coalition agreement signed in December 2013; this agreement already stated the implementation of the reform as of July 1, 2014. It was then legislated on May 23, 2014 and implemented only six weeks later on July 1, 2014. The reform was highly debated during the electoral campaign and received ample news coverage around the time of its announcement and implementation. Appendix Figure A2 reports the number of online and newspaper articles in German including the word 'Mütterrente' over time. There was an abrupt increase in newspaper coverage of Mütterrente in 2013, the year of the political campaign and the election, followed by even greater coverage in 2014, the year of legislation and adoption. The salience of the policy is also confirmed by the substantial increase in the intensity of Google searches for the term 'Mütterrente' in the German territory over the months preceding the reform implementation, as shown in Appendix Figure A3. The chart compares the evolution of the index of Google search intensity for Mütterrente with those for other salient gender- and pension-related policies. In light of this evidence, we consider December 2013, the signature date of the coalition agreement, as the time in which the enactment of the Mütterrente reform became almost certain, and use it as the relevant baseline period in our empirical strategy (see Section 4 for more details). From now on, we will refer to the period until December 2013 as pre-reform, and the period from January 2014 onward as post-reform.

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

[14] Ye (2022) studies the effects of a German pension subsidy program that increased the pension benefits by 15% on average. Such a large effect is due to the subsidy targeting the most disadvantaged retirees.

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 kids 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 kids) 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.[15] The last pension reform had in fact been passed in 2007.

To provide a complete picture of the institutional context, Appendix B describes reforms of maternity leave legislation in Germany that took place during the same time period as the reforms to the Mütterrente scheme. Specifically, a 1992 maternity leave reform took place concurrently with the Mütterrente reform, and increased the duration of job-protected leave from 18 to 36 months. As we argue in Section 4, this concurrent maternity leave reform does however not pose a problem for identifying the effect of the July 2014 pension reform on labor supply behavior in our empirical design.

3. Data and Sample Definition

Our empirical analysis is based on two administrative data sources. The first and main source are the Integrated Employment Biographies (IEB) of the Institute for Employment Research (IAB). The second source is the Insurance Account Sample Versicherungskontenstichprobe (VSKT) from the German Pension Insurance (DRV), which we use exclusively to calculate the change in pension wealth caused by the 2014 reform, and to check for differential fertility around the January 1, 1992 cut-off date. Pension points and information on all births are included in the latter data set, but not in the former. Yet, the sample size of the latter data set is much smaller.

3.1 Integrated Employment Biographies (IEB)

The IEB provide administrative data on *working histories* for the *entire* population of individuals who have ever worked in private sector employment subject to social security contributions or in

[15]By contrast, the 1990s were a decade of high policy uncertainty regarding pension reforms (Giavazzi and McMahon, 2008).

marginal employment.[16] It thus covers approximately 80% of the German working population starting from 1975 (1992) for West (East) Germany, excluding individuals who never worked, or always worked as either self-employed or civil servants.[17] The data include earnings and spell data on regular employment, marginal employment, receipt of unemployment insurance (unemployment benefit I) or of unemployment assistance (unemployment benefit II), registered job search, and participation in active labor market policies. Earnings are given as gross daily wages. Some spells indicate bonus payments which we include in our measure of total earnings. With this information, it is possible to construct the entire working and earnings history of individuals at daily frequency.

For our analysis, we determine individuals' main labor market status in a month selecting the status with the longest monthly duration.[18] Indicators for the different labor market statuses are coded as 1 if it is an individual's monthly main status and 0 otherwise. The monthly labor market indicators are collapsed at the semester level to obtain the average participation rate per status and semester. Monthly earnings in the main status are determined by multiplying the gross daily wage and the monthly spell duration. Earnings per semester are obtained by adding up all earnings in individuals' main status in a semester.

Rich information on occupation is also included in the data. Occupations are classified according to the 5-digit German classification of occupations (Klassifikation der Berufe 2010). The first four digits indicate an increasingly fine-grained categorization of occupations, while the fifth digit indicates the qualification level of the occupation (helper, professional, specialist, expert). In our analysis, an occupation is defined as the interaction between (i) the first two digits and (ii) the fifth digit, for a total of 127 categories.[19] As commonly the case in administrative data, hours worked are not recorded in the IEB data. We can, however, differentiate between part-time and full-time employment, and we will use that as our main indicator of intensive margin responses. The decision of marking a spell as part- vs. full-time is up to the reporting firm.

The IEB data do not provide direct information on childbirths, which we impute following the methodology developed by Müller and Strauch (2017). A woman aged 38 or younger is identified as giving birth if she features an employment interruption of at least 14 weeks (which is the mandatory maternity leave period) or an unemployment insurance de-registration, both due to entitlements from the statutory health insurance.[20] Since maternity leave is mandated to

[16]Marginal employment is covered by the IEB since April 1999.

[17]We cannot distinguish civil service, self-employment, and non-employment in the data.

[18]We break ties according to the following ranking: unemployment benefit I receipt > regular employment > unemployment benefit II receipt > marginal employment > registered job search > ALMP participation > not in IEB data. In a second step, if regular employment was determined to be the monthly main status in the first step, but spells of equal length in full-time and part-time employment exist, full-time employment is prioritized over part-time employment.

[19]A few very small categories have been combined with similar larger ones to obtain 127 groupings.

[20]Employment interruptions due to compensation by the statutory health insurance provider could be due to maternity leave or long-term sickness, which we cannot distinguish in the administrative data. The imputation procedure rests on the assumption that women of childbearing age are unlikely to experience employment interruptions due to long-term sickness. For women who move out of unemployment insurance receipt, we can distinguish de-registrations due to maternity from those due to sickness.

start six weeks before the estimated date of childbirth, the expected date of delivery is identified by adding 42 days to the last date of employment or the date of unemployment de-registration.[21] With this procedure, we can identify (expected) childbirths only for women with a record in the administrative data source immediately preceding childbirth, since we do not observe maternity leave spells for the self-employed, civil servants, or women out of the labor market.[22] This makes it likely that we miss second or higher order births, since many mothers do not return to the labor market between subsequent childbirths. We therefore focus in our analysis on first births only. Using a 1% sample of the IEB data that supplements the social security records with direct information on childbirth from the German Pension Insurance, Schönberg (2009) shows that approximately 54% of women in the IEB sample who gave birth in 1991 or 1992 took maternity leave and can thus be identified as mothers with the imputation procedure. Considering that Schönberg (2009) estimates that women are 37 percentage points less likely to take maternity leave for their second than for their first (recorded) child, 46 for their third and 49 for their fourth or more child, and taking the birth order percentages from 1991 and 1992 into account, we can conclude that the share of first-time mothers who take maternity leave and can therefore be identified with our imputation strategy is around 75% of all first-time mothers.[23] Given that 20% of West-German women born in the same cohorts as those in our sample are childless, our sample is likely representative of 60% of women in West Germany.[24]

In the IEB data, individuals can be matched to their spouses using a statistical matching procedure based on surnames and exact geo-referenced addresses (Bächmann et al., 2021). We describe the matching procedure in detail in Appendix C.3.

3.2 Insurance Account Sample (VSKT)

The second data source used in the analysis is the Insurance Account Sample Versicherungskontenstichprobe (VSKT) from the German Pension Insurance (DRV). VSKT is a stratified random sample of 2% of individuals living in Germany and registered with the statutory pension system. The data set provides the entire *contributory histories* of individuals aged between 30 and 67 in the reference year of the data. We use the 2020 wave of VSKT, which includes individuals born between 1952 and 1989.

The data set contains the full contributory histories of the sampled individuals up to the reference year, starting from age 14 and ending at age 67. For each individual, a monthly history

[21]We are unable to observe if a woman goes on maternity leave earlier than six weeks before the estimated due date, e.g. due to pregnancy issues. In those cases, we would underestimate the date of childbirth, potentially misclassifying as treated individuals in the control group.

[22]In addition, we cannot distinguish between single and multiple births, nor between live and still births.

[23]The birth order percentages are only available for births to married couples, which make up 88% of all births (Statistisches Bundesamt, 1991; 1992). Schönberg (2009) also shows that approximately 56% of women were working at some point in the nine months before childbirth, the criterion we use to be included in our VSKT sample. Again, this is likely a lower bound for first births.

[24]Data on the proportion of childless women by cohort are sourced from the Federal Statistical Office (Statistisches Bundesamt).

of employment, unemployment, sickness, childcare and other contributions to the pension system is recorded. With this information, we can compute the exact number of pension points (both relating to the specific time period or cumulated) at each point in time, both in total and from different contributory events. We use these data to calculate the pension wealth effect of the 2014 Mütterrente reform.[25] The VSKT data also include detailed information on the number of children and the exact month of birth of each child, irrespective of the previous labor market status of the mother. When computing statistics based on the VSKT, we always use sample weights, because the VSKT oversamples certain groups.

3.3 Sample Selection

IEB Sample We restrict the main sample of analysis in the IEB data to women born between 1952 and 1974, who gave birth to their first child in the last quarter of 1991 (treatment group) or the first quarter of 1992 (control group). We exclude mothers who gave birth in the two weeks right before and after January 1, 1992, to deal with mismeasurement in childbirth imputation, which could lead us to misassign treatment status. Since East Germans were only systematically included in the IEB data starting in 1992, we cannot identify births in the treatment group for East Germans, and therefore exclude them from the analysis. To do that, we exclude all women who had their first spell recorded after 1989, at least one spell recorded in East Germany, or missing district of employment. We also exclude women who ever worked in the mining sector, which is subject to different pension contribution rules. Last, we exclude women who died before July 2014. Our final sample comprises 99,104 women.

Appendix Table A1 reports summary statistics for our analysis sample as of December 2013 (with earnings relating to the semester July to December 2013), distinguishing individuals in the treatment and control groups. The average woman in our sample was 27 at the time of her first childbirth and close to 49.5 at the end of 2013. In December 2013, 72% of women are employed, 21% full-time and 38% part-time, with the remainder in marginal employment. About 23% of women have no employment or unemployment spell in the data in December 2013, meaning they are either out of the labor force, in civil service, or self-employed. Earnings when in work average EUR 10,500 for the second semester of 2013. Differences in means between the treatment and control groups are very small in magnitude. Out of those that are statistically significant at 5% or less, most differences are below 2% of the mean in the control group.

VSKT Sample We follow the same sample selection criteria in the VSKT data as in the IEB data, with a few exceptions. Due to the smaller sample size by two orders of magnitude, we restrict the sample in the VSKT to women who gave birth to their first child in the twelve

[25] For our heterogeneity analysis, we also impute pension points in the IEB data based on individual employment histories as described in more detail in Appendix C.2. Therein we also validate the quality of our imputation by comparing it with administrative records from the Insurance Account Sample of the German Pension Insurance.

months, rather than three months, before or after January 1, 1992.[26] To increase consistency with the IEB sample of mothers, we further restrict the sample to mothers who have at least one employment spell in the nine months prior to giving birth. Our final VSKT sample comprises 2,787 women.

The summary statistics reported in Appendix Table A2 show that the average woman in the VSKT sample was approximately 48.5 in December 2013 with total fertility of 2 children and a total of about 24 pension points.

4. Expected Effects of Mütterrente and Identification Strategy

Our objective is to identify the effect of pension wealth on individual labor supply behavior far from retirement. Before delving into the identification strategy and its empirical implementation, it is useful to consider the potential impact of the 2014 childcare pension benefit reform on mothers' labor supply decisions.

Consider two identical women giving birth to a child right before and right after January 1, 1992, and with earnings below the pension contribution limit in 1991 and 1992. As a consequence of the 1992 pension reform and the subsequent 1999 reform, the situation until July 2014 is the following: the mother who gave birth before January 1, 1992, receives one pension point for this birth, while the mother who gave birth on or after January 1, 1992, receives three points. The 2014 reform attributes one extra pension point to the first mother who gave birth before January 1, 1992, raising her points to two. In other words, the 2014 reform increased the number of cumulated pension points of mothers of children born before the cutoff date by one (per child), while leaving the number of cumulated pension points of mothers of children born after that date unchanged. We thus assign mothers of children born before January 1, 1992 to the treatment group, and mothers of children born January 1, 1992 or later to the control group.

Since the increase in pension points mechanically increases pension wealth, the reform generates a positive and permanent wealth effect for mothers of children born before January 1, 1992. This is a pure wealth effect, not confounded by changes in substitution incentives. The reason why the reform creates a pure wealth effect is that the accreditation of pension points is independent of current and future labor supply choices. It is also fully exogenous to fertility choices, since eligibility is uniquely based on the birth date of existing children, which obviously cannot be manipulated ex post. In addition, the accreditation of the additional pension point does not affect the qualifying period for pension eligibility under either the early retirement or the seniority pension scheme.[27] Importantly, the pension wealth effect stems from a change in

[26] We do not exclude mothers who gave birth in the two weeks around the January 1, 1992 cutoff, since we observe the exact month of birth of children in the VSKT, thus not running the risk of mismeasurement.

[27] As we explain in Appendix C.2, mothers are credited *consideration* periods for child care until the child turns 10. Such consideration periods do not increase pension entitlements (i.e. pension wealth), but count towards the qualifying period for eligibility for early retirement or seniority pension. On the other hand, *contribution* periods raise pension entitlements and count towards pension eligibility. The Mütterrente reform transformed the second

pension points for a constant pension benefit formula. The stability of pension rules makes it likely that information or cognitive frictions are minimal. As a result, the reform provides useful variation to neatly identify the effect of pension wealth on labor supply decisions.

If leisure is a normal good, the increase in pension wealth is expected to reduce the lifetime labor supply of recipients, at either the intensive or the extensive margin and including anticipating retirement, relative to a counterfactual scenario in which the 2014 reform did not take place. Clearly, we do not expect any labor supply response from individuals who had already retired by 2014, though the fraction of retired individuals among those who gave birth in 1992 is negligible.[28] Even though the increase in pension wealth could in principle induce individuals to anticipate retirement benefit claiming, this is a margin that we are unlikely to uncover since less than 5% of our sample is old enough to qualify for retirement benefits over the time period of analysis.[29]

Our objective is to identify the effect of pension wealth on labor supply behavior prior to retirement. The regression equation that describes the causal relationship of interest is:

$$Y_{it} = \alpha_0 + \alpha_1 W_{it} + X_{it}' \alpha_2 + \varepsilon_{it} \tag{1}$$

where Y_{it} is the labor supply outcome of interest Y for individual i at time t, W_{it} is the present discounted value of pension wealth cumulated by i at time t, X_{it} represents a vector of individual controls, and ε_{it} is an error term. For $Y = z$, where z is labor earnings, the parameter α_1 identifies the wealth effect of retirement benefits. Given the potential endogeneity of W, we exploit exogenous variation in cumulated pension wealth due to the 2014 pension reform, and use the timing and eligibility requirements of said reform as an instrumental variable (IV) for pension wealth. In practice, we estimate the coefficient α_1 by indirect least squares – the ratio of the reduced form to the first stage – where, due to data availability, the first stage is estimated in the VSKT data and the reduced form in the IEB data.

To identify the reduced-form effect of the reform on labor supply, we implement a difference-in-differences design. We compare labor supply outcomes between mothers who had their first child before versus after January 1, 1992, and trace out the dynamics of this difference from before to after the 2014 reform.[30] More formally, for each semester t running from the first semester of 2010 to the second semester of 2018, our reduced form is estimated using the

year after childbirth for mothers of children born before 1992 from a consideration period to a contribution period, thus increasing pension wealth without affecting the overall qualifying period for retirement eligibility.

[28]In our VSKT sample, we see that around 2.7% of women who gave birth to their first child around January 1, 1992 had retired by December 31, 2013, primarily receiving pensions for (partial) disability.

[29]Appendix Figure A4 reports the age distribution of mothers in our main sample as of December 2013. Only mothers aged 57 and over in 2013 are potentially eligible for retirement benefits by the end of 2018 (subject to meeting other eligibility requirements).

[30]Given that our sample excludes mothers who gave birth in the two weeks around January 1, 1992, our empirical strategy is in essence a *doughnut* difference-in-differences.

following specification:

$$Y_{it} = \sum_{s \neq Jul-Dec\ 2013} \gamma_s \cdot D_i \cdot \mathbb{I}[t = s] + \delta_t + \delta_i + \nu_{it} \tag{2}$$

where Y_{it} is defined as above, D_i is an indicator taking value 1 if individual i's first child was born before January 1, 1992, and zero otherwise, δ_t and δ_i are calendar time and individual fixed effects, $\mathbb{I}[\cdot]$ is an indicator function and ν_{it} an error term. Standard errors are clustered at the individual level. The set of coefficients γ_s identifies the difference in outcomes between the treatment ($D_i = 1$) and control ($D_i = 0$) group in each semester relative to the July-December 2013 one, which we take as baseline as explained in Section 2.2. The magnitude of the γ_s coefficients quantifies the reduced-form effect of the reform on outcome Y. When presenting the empirical estimates, we plot the set of estimated γ_s coefficients from January-June 2010 to July-December 2018, which allows to visually assess pre-policy trends and the dynamics of post-policy responses. We also report the estimated coefficients of a version of equation (2) in which we replace the first term with a set of interactions between D_i and indicators for 1-2.5 years after the baseline period and 3-5 years after the baseline period. We label the former as short-run effect and the latter as medium-run effect. Our indirect least squares estimates will be based on the medium-run reduced-form effect.

Our first stage effect is the change in the present discounted value of pension wealth generated by the 2014 reform as of December 2013. As we will explain in more detail below, we simulate the absolute and proportional increase in pension wealth using detailed data on pension contributions from the VSKT.

The key assumptions for the γ_s coefficients in equation (2) to be identified are a standard common-trends assumption and the lack of contamination effects. The former requires that the evolution of labor supply outcomes of women with children born after January 1, 1992 (control group) offers a good counterfactual for that of mothers of children born before that date (treatment group). This assumption can be corroborated by the lack of pre-policy differential trends between the treatment and control groups. To strengthen the case for a compelling counterfactual, we focus on women giving birth to their first child in a three-month bandwidth around the January 1, 1992 cutoff. Concerning contamination effects, the imputation procedure to impute childbirth is prone to measurement error around the January 1, 1992 cutoff, implying that the treatment/control group classification may be incorrect in a narrow window around the threshold. As described above, to overcome this issue, we implement a doughnut difference-in-differences estimation in which we exclude women with imputed childbirths in a two-week window around the cutoff. One may also worry about general equilibrium effects: if treated individuals reduce their labor supply in response to the reform, this may affect equilibrium wages and tax rates. However, it seems reasonable to believe that these effects will be small and, in any case, identical for individuals in both treatment and control groups, who should be perfect substitutes in the labor market.

For our indirect least squares estimate to have a causal interpretation, we also require the instrument to be relevant and to satisfy the exclusion restriction. We document below that the reform generated a meaningful increase in the present discounted value of pension wealth of EUR 3,830 or 4.4% of total pension wealth as of December 2013. As for the exclusion restriction, we are not aware of other reforms taking place around the same time and with the same eligibility criteria. Moreover, we believe it is unlikely that the timing and eligibility requirements of the reform affect employment outcomes through channels other than pension wealth. To the extent that it may do so, for example through general equilibrium effects, we expect those to be identical across the treatment and control groups.

The main advantage of this empirical strategy is that it neatly identifies a pure wealth effect, which is typically hard to do using policy variation. Moreover, our difference-in-differences design allows us to account for the potential confounding effect on labor supply of (i) the 1992 Mütterrente reform, which gave more generous pension contributions to mothers of children born after January 1, 1992, and (ii) a 1992 maternity leave reform, which increased the duration of job-protected leave for those same mothers. The 1992 pension reform was legislated in 1989 and, as such, was anticipated by mothers giving birth in a window around the January 1, 1992 cutoff. Knowledge of future benefits might have induced mothers to time the birth date of their children in order to take advantage of the more generous regime. In fact, evidence from Schönberg and Ludsteck (2014) seems to rule out this strategic behavior. The 1992 maternity leave reform is, instead, likely not anticipated, since the draft bill for the reform was proposed after children born within three months of the policy reform were conceived (Schönberg and Ludsteck, 2014). Nonetheless, by affecting the likelihood and timing of mothers returning to work after childbirth, both reforms are likely to have had an impact on the working history and, hence, on cumulated contributions for women with children born after January 1, 1992. The time difference of our design accounts for these potential effects.

5. Effect of Pension Wealth on Individual and Spousal Labor Supply

5.1 Reduced-Form Effects

Figure 1 reports estimates of equation (2) for our main outcome of interest: total unconditional earnings. These are defined as the sum of labor income from dependent employment (full-time, part-time and marginal employment), and take value zero when individuals are not recorded as employed (e.g. because they are registered as job-seeker or because they are out of the labor force and hence not covered by the data) or receive unemployment benefits.[31] As such, they provide an all-encompassing measure of employment at the extensive and intensive margin. Total unconditional earnings drop significantly for the treatment group after the reform. Our estimate of the short-run effect over the 2.5 years following the policy change is of a reduction of

[31] As a result, we work here with a balanced panel.

EUR 51 per semester, or 0.7% of the pre-policy mean in the treatment group (see the first row of Panel A of Table 1). Consistent with the notion that optimization frictions may limit the ability of individuals to adjust their labor supply in the short run, we estimate a larger medium-run effect in the 3-5 years after the reform, when the magnitude of the treatment effect appears to stabilize at around EUR 98 per semester, or 1.3% of the pre-policy mean.[32] The treatment leads are statistically insignificant, indicating that there are no differential pre-trends. Thus, we conclude that the positive pension wealth effect of the childcare pension benefit reform significantly reduces total unconditional earnings of the affected mothers.[33]

Appendix Table A3 reports estimates of the effect of the reform on total unconditional earnings for the sample of individuals employed in the baseline period in Panel A and the sample of the non-employed in Panel B. The table shows that the effect we detect on unconditional earnings in Figure 1 is entirely driven by individuals who are working before the reform and reduce their labor supply afterwards, rather than by fewer entries into the labor market following the reform.

Panels A to C of Figure 2 and Panels A and B in Table 1 uncover the anatomy of the labor supply response. As illustrated in Panel A of Figure 2, the reform has essentially no effect on the extensive margin, as measured by the probability of being employed. There is also no change in the probability of receiving unemployment benefits, nor of not appearing in the IEB records (because of non-employment, self-employment or employment in the civil service), as reported in Panel A of Table 1. Rather, individuals seem to respond at the intensive margin: Panel B of Figure 2 shows that labor earnings conditional on employment drop by 0.4% in the short run and 0.9% in the medium run. Whilst the data do not allow us to measure hours worked – and thus disentangle whether the drop in conditional earnings is due to a change in hours or wages – the estimates in Panel C document a drop in the likelihood of working full-time (as opposed to working part-time or in marginal employment) conditional on employment in the aftermath of the reform. Although less precisely estimated, the coefficients reported in Panel C indicate a drop in full-time employment of 1.1% in the short run and of 1.8% in the medium run. In Panel B of Table 1, we can see positive, but insignificant increases in the probability of working part-time and, to a smaller extent, of being in marginal employment in the medium run – effects that are also visible in Panels A and B of Appendix Figure A5. These results suggest that affected mothers reduce their hours of work in response to the positive pension wealth shock.

Finally, we investigate whether individuals adjust their labor supply by switching employer or type of job. As reported in Panel C of Table 1, we do not find evidence of an increase in the probability of switching establishment, nor in the likelihood of being employed in a

[32]According to VSKT data – where we observe all children born to a woman – the youngest child of the average woman in our sample is 18.3 years old in December 2013. This suggests that childcare constraints should not play a role in labor supply adjustment.

[33]In principle, the reduced-form effect that we estimate for 2018 could encompass both the effect of the 2014 reform and some anticipated responses to the 2019 one. In practice, though, the stability of the estimated effects in 2017 and 2018, coupled with the lack of evidence of anticipated effects of the 2014 reform, leads us to conclude that such anticipated responses are likely minimal.

part-time intensive occupation, that is an occupation with above-median share of employees working part-time in the pre-reform period.[34] In practice, though, the probability of working in a part-time intensive occupation is close to 90% at baseline, indicating that the vast majority of mothers already work in occupations that are suitable for part-time arrangements.

5.2 Implied Marginal Propensity to Earn out of Pension Wealth

To interpret the magnitude of the reduced-form estimates illustrated in the previous subsection, we rescale the estimated earnings response by the simulated pension wealth shock, in the spirit of indirect least squares. To simulate the pension wealth shock, we recur to the VSKT data. We describe our simulation in detail in Appendix C.1. Based on an assumed monthly discount factor of 0.9983, assumed retirement at age 67 and an assumed retirement period of 22.5 years (270 months), as well as an estimated monthly growth rate of the real value of a pension point of -0.0001619, we calculate that the present discounted value of pension wealth for mothers in the treatment group increases by EUR 3,830 as of December 2013, corresponding to a 4.4% increase in the present discounted value of their pension wealth.

To obtain measures of the wealth effect and the wealth elasticity that are plausibly not attenuated by optimization frictions, we consider the medium-run estimates of the unconditional earnings response reported in the first row of Panel A of Table 1. Total unconditional earnings dropped by EUR 196 per year or 1.3% of their pre-policy level. For an increase in the present discounted value of pension wealth of EUR 3,830, or 4.4%, the implied wealth effect is of EUR -5.1 per year per EUR 100 of wealth, and the implied wealth elasticity is of -0.3.

The size of the pension wealth effect that we estimate above can be hard to assess, since it does not account for the time horizon over which individuals plausibly allocate the increase in pension wealth. In order to obtain an effective estimate of the marginal propensity to earn (MPE) out of pension wealth in a given period, we first estimate the allocation of the pension wealth increase over time and then compute the MPE in a given year. We follow the annuitization method described in Golosov et al. (2021) to allocate the wealth windfall over time. We assume that individuals perfectly smooth the pension wealth increase over the years that separate them from the full retirement age, which correspond to 17 years on average in our sample. Under this assumption, individuals allocate 29% (5/17) of the windfall to the 5 years covered by our analysis. Our estimates of the unconditional earnings response from the first row of Panel A of Table 1 imply that, over the 5 years after the reform, earned income drops cumulatively by EUR 612 on average, discounting the per-period changes to December 2013.[35] Based on these computations, we estimate an MPE of -0.54: an extra euro of pension wealth leads to a 54

[34] To determine the pre-reform share of part-time employees in an occupation, we use a representative 2% sample of the IEB and only consider employment spells that include December 31, 2013. In case a worker has more than one relevant spell in an occupation, we keep the one with the highest working time. For each of our 127 occupations, we then calculate the fraction of part-time workers as the share of part-time and marginal employees among all employees.

[35] For discounting, we use a monthly discount factor of 0.9983. This is the same discount factor used to simulate the pension wealth change in Appendix C.1.

cent reduction in labor earnings in middle-age years.[36] The size and salience of the economic incentives generated by the pension reform could both be important for the magnitude of the response that we document (Seibold, 2021). Our estimate of the MPE is close to the one found by Golosov et al. (2021) for US lottery winners (-0.5). It is also close to estimates of the wealth effect of Old Age and Survivors Insurance benefits close to retirement, estimated out of the US Social Security 'Notch' by Gelber, Isen and Song (2016) and Gelber, Isen and Song (2017). They find, respectively, MPEs out of social security wealth of -0.6 for men and -0.9 for women.[37]

The labor supply responses that we document are concentrated along the intensive margin. Studies that examine pension wealth changes closer to the retirement age (and disentangle margins of adjustment) tend to find, instead, responses along the extensive margin (Gelber, Isen and Song, 2016; Becker et al., 2022).[38] Our evidence of intensive margin responses is in line with the findings in Deshpande (2016), who documents large intensive margin responses by parents to the income effect induced by the loss of Supplemental Security Income payments for their children, when the latter are below the age of 13. Taken together, this evidence points towards age patterns in the use of different margins of adjustment, which could be explained by varying (utility) costs of adjusting labor supply with time to retirement. To the extent that extensive margin responses signal larger costs of adjustment, they reveal higher implicit valuations of social insurance transfers for a given level of the wealth effect (Chetty, 2004; 2008).

By reducing labor supply in response to the reform, mothers are implicitly reducing the amount of contributions paid into the system. We quantify the pension point loss associated with the unconditional earnings response, assuming that the unconditional earnings reduction remains stable at EUR 196 per year for the 12 years that separate the average individual from the full retirement age after our period of analysis. The pension point loss amounts to a total of 0.074 points over the entire post-reform period, which corresponds to approximately 1/13 of the childcare pension point increase due to the Mütterrente reform.

5.3 Identification Tests and Robustness Checks

Our difference-in-differences strategy rests on the 'common trends' assumption that the labor supply outcomes of treated and control individuals would have evolved in parallel absent the 2014 Mütterrente reform. As we already noted in Section 5.1, the fact that we do not see differential earnings and employment changes between the treatment and control groups in the pre-policy period is reassuring in this respect.

[36] If we are willing to assume that the medium-run response that we estimate also holds in the long run (i.e. for the years until the full retirement age that we do not observe in the data), our estimate of the MPE becomes -0.69.

[37] It is worth noting that the estimate in Gelber, Isen and Song (2017) is based on a selected sample of women born in 1916-1917 who had particularly high lifetime earnings relative to their husbands.

[38] Gelber, Isen and Song (2016) find that the probability of being employed decreases by -0.67 (standard error 0.18) percentage points per 10,000 dollars of additional pension wealth. For an equivalent pension wealth shock, our estimates imply an employment effect of -0.005 (standard error 0.005), and thus allow us to rule out effects lower than -0.015.

As a further test for the validity of our assumption, we report a battery of placebo tests in Panels A and B of Appendix Table A4, where we use – respectively – January 1, 1994, and January 1, 1995, as placebo thresholds to define (placebo) treatment and control groups.[39] The tables report estimates of equation (2) for our main outcomes of interest. We do not find any significant differential effects of the placebo reform across our placebo treatment and control groups, which is reassuring of the fact that our main estimates are picking up a genuine causal effect of the reform.

As discussed in Section 2, the January 1, 1992, threshold is not only relevant for the 2014 Mütterrente reform, but also for (i) the 1992 Mütterrente reform, which gave additional pension points per child born on or after, as opposed to before, January 1, 1992, and (ii) a maternity leave reform, which doubled the duration of job-protected leave for mothers of children born on or after January 1, 1992.[40] One may worry that differential fertility responses to those 1992 reforms might make our treatment and control groups different with respect to their childcare constraints as of 2014 and, more generally, their life-cycle labor supply trajectories. In Appendix Figure A6, we use data from the VSKT to document that the combination of the 1992 reforms did not affect the total fertility of women in our sample. The gray dots plot the average number of children born to mothers who had their first child in a given month around the January 1, 1992 cutoff, conditional on mother's cohort fixed effects and child's month of birth fixed effects. The number of children is computed as of December 2019, that is approximately 28 years after the first childbirth, and can therefore be considered a close proxy for total fertility. Whilst total fertility is on a declining trend, we do not observe any meaningful difference in the total number of children of women in a three-year bandwidth around the cutoff. This is further corroborated by the very small and statistically insignificant regression-discontinuity estimate reported in the graph.

To account for measurement error in childbirth imputation, our main doughnut difference-in-differences specification excludes women with imputed childbirths in a two-week window around the cutoff. We assess the robustness of our estimates to different doughnut-hole widths (zero, four and six weeks) in Appendix Table A5. Point estimates are stable across specifications.

As a final robustness check, we consider the soundness of conditioning some of our main outcomes on employment. In Section 5.1, we probed the anatomy of the unconditional earnings response by decomposing the latter into a participation response (extensive margin) and a conditional earnings response (intensive margin). In general, conditioning on an outcome – employment in this case – may raise concerns of endogenous selection due to compositional changes. In our specific case, the lack of employment effects of the reform suggests that endogenous selection should not be a concern. To further probe this point, Panel A of Appendix Table A3 replicates our main analysis on the sample of mothers who were already in employment

[39] We omit January 1, 1993, since a maternity leave reform was implemented on that date.

[40] More precisely, the 1992 Mütterrente reform gave 2.25 pension points per child born on or after January 1, 1992, leaving the amount of pension points per child born before that date at 0.75. The 1999 Mütterrente reform increased retroactively the number of pension points per child to 3 for those born on or after January 1, 1992, and to 1 for those born prior to that date.

in the second semester of 2013. Albeit imperfectly, conditioning on pre-reform employment implicitly allows to isolate the effect of the reform on earnings (and the likelihood of working full-time) from that of potential compositional changes due to extensive-margin responses to the reform itself. Results are very similar to the ones for the baseline sample in the paper, suggesting that endogenous sample selection is not confounding our estimates.

5.4 Heterogeneity Analysis

The observed response to a change in pension wealth may vary across individuals depending on a range of factors, such as their idiosyncratic returns to tenure and costs of working, their liquidity constraints, or their ability to self-insure against the longevity risk through their savings. In this section, we study heterogeneous responses to the pension wealth effect along these and other dimensions.

5.4.1 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.[41] 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 semester 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 Table 2 reports estimates of an augmented version of equation (2), fully interacted with a dummy for having above-median predicted returns to tenure. In Table 2, 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

[41] We 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.

the intensive margin, in that they are less likely to leave full-time employment for a part-time job or a minijob.[42]

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.[43] 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 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 2 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.[44]

5.4.2 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 Table 2 shows that women with above-median pension points experience larger reductions in labor supply at the

[42]Note that, being weighted by working time, our measure of the returns to tenure is a proxy for returns at both the extensive and intensive margins of employment.

[43]For 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.

[44]This 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 2 might partly reflect variation in costs. The evidence in Panel B of the same table would seem to exclude that this is the case.

intensive margin. In Panel D of Table 2, 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. Overall, these results point to larger reductions in labor supply by individuals who have larger wealth to start with.

5.4.3 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 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 semester of 2013.[45] We then split the sample of mothers by partners' median earnings within partners' birth cohorts. The results reported in Panel E of Table 2 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.

5.4.4 Time to Retirement Age

We then consider whether different working horizons lead to differential responses to the pension wealth shock. The labor supply response to a pension wealth shock might be larger if the time to statutory retirement is shorter. In Panel F of Table 2, we report estimates of heterogeneous responses by the time to the full retirement age, measured as of December 2013. We take into account that individuals born in different years face different full retirement ages.[46] The results do not reveal differential responses by time to statutory retirement. We should note, however, that the number of years to the full retirement age in December 2013 is collinear with age at the time of the first childbirth and age at the time of the reform. 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 time to statutory retirement might also be capturing differences in the degree of labor market attachment. Mothers with a shorter time to the full retirement age in 2013 had their first child relatively late, possibly indicating a stronger attachment to the labor

[45] This 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.

[46] Note that the inclusion of individual fixed effects in the regression implicitly accounts for the fact that shorter working horizons may correspond to higher baseline levels of pension wealth.

market. This could counteract a potential larger labor supply response of individuals closer to the full retirement age.

In Appendix Figure A7, we test the parallel-trends assumption for all estimates reported in Table 2. 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 A7, 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.

5.5 Evidence on Effects across Spouses

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 distance to the full retirement age. 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.[47] As reported in Panel A of Table 3, we find that the intensive-margin labor supply responses of women in our sample are driven by mothers whose partner is closer to the full retirement age (and thus older), consistent with within-household interactions in labor supply choices.[48] The estimates reported in Panel A of Appendix Figure A8 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 Table 3 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

[47] Slightly less than 1% of partners had reached the full retirement age in December 2013.

[48] To the extent that age is correlated across spouses, time to full retirement age could also be correlated across spouses. As such, heterogeneity by partner's time to full retirement age could also reflect heterogeneity by own time to full retirement age. Since, in our setting, one's own time to full retirement age does not seem to matter for labor supply responses (Panel E of Table 2), we believe that the heterogeneity reported in Panel A of Table 3 is likely to reflect genuine variation in partner's time to full retirement age.

found among women in absolute terms (see Table 1).[49] 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, which is statistically significant.[50] We test the common trends assumption for partner's outcomes in Panel B of Appendix Figure A8. 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.

6. Conclusion

In this paper, we study how changes in the generosity of public pension systems affect labor supply behavior far from retirement. We exploit the 2014 reform of the Mütterrente, which increased the pension wealth of mothers of children born before January 1, 1992 by 4.4% per child on average. We document significant reductions in labor earnings when affected women are on average 50-55 years old, driven by shifts out of full-time employment. Our estimates imply a marginal propensity to earn out of pension wealth of -0.54, which can be considered a moderately large effect. We also provide evidence of within-household interactions in labor supply choices. Our findings contribute to a growing strand of the literature that analyzes how the design of pension benefits affects labor supply far from retirement. In particular, our setting allows us to overcome several empirical challenges in the identification of pension wealth effects far from retirement.

Our results have important policy implications. We show that (i) individuals are forward-looking and (ii) spillovers within the household can amplify responses to individual incentives. This implies that pension reforms can have aggregate labor supply effects well beyond the direct impact on individuals on the verge of retirement. Another important policy implication of our results is that the true government cost of the Mütterrente reform is higher than its mechanical cost. If we restrict our focus on mothers' labor supply responses and are willing to assume that our estimates of the medium-run effect of the policy on earnings hold for all years until full

[49] We 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 A6.

[50] In Germany, married individuals and civil partners can choose whether to be taxed on their individual or joint income. Under joint taxation, mothers' earnings reductions in response to the Mütterrente reform could, in principle, lower the marginal tax rate faced by their partners. Our estimates of partners' labor supply responses to the pension wealth change could therefore be attenuated (in absolute value) due to the concurrent change in substitution incentives. We use a tax simulator to quantify the change in the marginal tax rate due to the mothers' earnings response (Bick et al., 2019). Our simulation indicates that the marginal tax rate changes by less than 1 percentage point across all deciles of the household income distribution. Thus, in practice, substitution incentives are negligible in our context.

retirement age, our estimates indicate that increasing pension wealth by EUR 3,830 decreases labor earnings by EUR 2,660 in present discounted value. Assuming a 45% marginal income tax and social security contribution rate, the net impact of the policy on government expenditures is EUR 5,027 (= EUR 3,830 + EUR 1,197). Equivalently, the policy generates a fiscal externality (FE) of 31 cents per euro of transfer (= EUR 1,197/EUR 3,830).[51]

We can use our estimate of the fiscal externality to assess the welfare implications of the reform, in the spirit of Hendren (2016)'s policy elasticity approach. This approach quantifies the welfare impact of a policy by the marginal value of public funds (MVPF), which is the ratio of the beneficiaries' willingness to pay for the policy to its net government cost. For a willingness to pay of EUR 1 for a EUR 1 increase in the net present value of pension wealth, the MVPF of the Mütterrente reform is $\frac{1}{1+\text{FE}} = \frac{1}{1+0.31} = 0.76$. This value is in line with the MVPFs for retirement benefits (0.8) and disability insurance for adults (0.87 on average) reported in Hendren and Sprung-Keyser (2020) and Policy Impacts (n.d.). These MVPFs are below 1 due to the earnings crowd-out that they entail. Welfare weights on pension recipients would have to be substantial to justify the Mütterrente benefit increase, compared to alternative policies with larger MVPFs.

[51] Assuming that labor is the only source of taxable income, tax simulations indicate that the average mother in our sample faces a 42% marginal income tax and social security contribution rate when filing taxes individually and of 47.5% when filing taxes jointly with her spouse. For our calculation, we consider an average marginal tax rate of 45%. The fiscal externality is 0.29 and 0.33 for a marginal tax rate of 42% and 47.5% respectively.

References

Attanasio, Orazio P., and Agar Brugiavini. 2003. "Social Security and Households' Saving." *The Quarterly Journal of Economics*, 118(3): 1075–1119.

Attanasio, Orazio P., and Susann Rohwedder. 2003. "Pension Wealth and Household Saving: Evidence from Pension Reforms in the United Kingdom." *American Economic Review*, 93(5): 1499–1521.

Bächmann, Ann-Christin, Corinna Frodermann, Benjamin Lochner, Michael Oberfichtner, and Simon Trenkle. 2021. "'Identifying Couples in Administrative Data' for the years 2001-2014." FDZ-Methodenreport 03/2021 (en).

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.

Bick, Alexander, Bettina Brüggemann, Nicola Fuchs-Schündeln, and Hannah Paule-Paludkiewicz. 2019. "Long-term Changes in Married Couples' Labor Supply and Taxes: Evidence from the U.S. and Europe since the 1980s." *Journal of International Economics*, 118: 44–62.

Blundell, R., E. French, and G. Tetlow. 2016. "Retirement Incentives and Labor Supply." In *Handbook of the Economics of Population Aging*. Vol. 1 of *Handbook of the Economics of Population Aging*, , ed. John Piggott and Alan Woodland, Chapter 0, 457–566. Elsevier.

Bottazzi, Renata, Tullio Jappelli, and Mario Padula. 2006. "Retirement expectations, pension reforms, and their impact on private wealth accumulation." *Journal of Public Economics*, 90(12): 2187–2212.

Bovini, Giulia. 2019. "Life-Cycle Labor Supply and Social Security Wealth: Evidence from the 1995 Italian Pension Reform." London School of Economics and Political Science Working Paper.

Börsch-Supan, Axel, Michela Coppola, and Johannes Rausch. 2015. "Die „Rente mit 63“: Wer sind die Begünstigten?" *Perspektiven der Wirtschaftspolitik*, 16(3): 264–288.

Carta, Francesca, and Marta De Philippis. 2021. "Working horizon and labour supply: the effect of raising the full retirement age on middle-aged individuals." Bank of Italy, Economic Research and International Relations Area Temi di discussione (Economic working papers) 1314.

Cesarini, David, Erik Lindqvist, Matthew J. Notowidigdo, and Robert Östling. 2017. "The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries." *American Economic Review*, 107(12): 3917–3946.

Chetty, Raj. 2004. "Optimal Unemployment Insurance When Income Effects are Large." National Bureau of Economic Research Working Paper 10500.

Chetty, Raj. 2008. "Moral Hazard versus Liquidity and Optimal Unemployment Insurance." *Journal of Political Economy*, 116(2): 173–234.

Costa, Dora L. 1997. "Displacing the Family: Union Army Pensions and Elderly Living Arrangements." *Journal of Political Economy*, 105(6): 1269–1292.

Danzer, Alexander M. 2013. "Benefit Generosity and the Income Effect on Labour Supply: Quasi-Experimental Evidence." *The Economic Journal*, 123(571): 1059–1084.

Deshpande, Manasi. 2016. "The Effect of Disability Payments on Household Earnings and Income: Evidence from the SSI Children's Program." *The Review of Economics and Statistics*, 98(4): 638–654.

Deutsche Rentenversicherung Bund. 2014. "Rentenbestand am 31.12.2014." Deutsche Rentenversicherung Bund, Berlin.

Deutsche Rentenversicherung Bund. 2021*a*. "Rente: So wird sie berechnet - alte Bundesländer." Deutsche Rentenversicherung Bund, Berlin.

Deutsche Rentenversicherung Bund. 2021*b*. "Rentenrechtliche Zeiten." Deutsche Rentenversicherung Bund, Berlin.

Deutsche Rentenversicherung Bund. 2022*a*. "Arbeitslos - was Sie beachten sollten." Deutsche Rentenversicherung Bund, Berlin.

Deutsche Rentenversicherung Bund. 2022*b*. "Rentenversicherung in Zahlen 2022." Deutsche Rentenversicherung Bund, Berlin.

Fetter, Daniel K., and Lee M. Lockwood. 2018. "Government Old-Age Support and Labor Supply: Evidence from the Old Age Assistance Program." *American Economic Review*, 108(8): 2174–2211.

French, Eric. 2005. "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour." *Review of Economic Studies*, 72(2): 395–427.

French, Eric, Attila S Lindner, Cormac O'Dea, and Tom A Zawisza. 2022. "Labor Supply and the Pension-Contribution Link." National Bureau of Economic Research Working Paper 30184.

Gelber, Alexander, Adam Isen, and Jae Song. 2016. "The Effect of Pension Income on Elderly Earnings: Evidence from Social Security and Full Population Data." Goldman School of Public Policy Working Paper.

Gelber, Alexander, Adam Isen, and Jae Song. 2017. "The Role of Social Security Benefits in the Initial Increase of Older Women's Employment: Evidence from the Social Security Notch." *Women Working Longer: Increased Employment at Older Ages*, 239–268. University of Chicago Press.

Gelber, Alexander M., Damon Jones, and Daniel W. Sacks. 2020. "Estimating Adjustment Frictions Using Nonlinear Budget Sets: Method and Evidence from the Earnings Test." *American Economic Journal: Applied Economics*, 12(1): 1–31.

Giavazzi, Francesco, and Michael McMahon. 2008. "Policy Uncertainty and Precautionary Savings." National Bureau of Economic Research Working Paper 13911.

Giupponi, Giulia. 2019. "When Income Effects are Large: Labor Supply Responses and the Value of Welfare Transfers." Centre for Economic Performance, London School of Economics CEP Discussion Papers dp1651.

Golosov, Mikhail, Michael Graber, Magne Mogstad, and David Novgorodsky. 2021. "How

Americans Respond to Idiosyncratic and Exogenous Changes in Household Wealth and Unearned Income." National Bureau of Economic Research Working Paper 29000.

Goux, Dominique, Eric Maurin, and Barbara Petrongolo. 2014. "Worktime Regulations and Spousal Labor Supply." *American Economic Review*, 104(1): 252–76.

Gruber, Jonathan, and David Wise. 1999. *Social Security and Retirement around the World.* National Bureau of Economic Research, Inc.

Hairault, Jean-Olivier, Thepthida Sopraseuth, and François Langot. 2010. "Distance to Retirement and Older Workers' Employment: The Case for Delaying the Retirement Age." *Journal of the European Economic Association*, 8(5): 1034–1076.

Hendren, Nathaniel. 2016. "The Policy Elasticity." *Tax Policy and the Economy*, 30(1): 51–89.

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. "A Unified Welfare Analysis of Government Policies." *Quarterly Journal of Economics*, 135(3): 1209–1318.

Imbens, Guido W., Donald B. Rubin, and Bruce I. Sacerdote. 2001. "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players." *American Economic Review*, 91(4): 778–794.

Jones, Damon, and Ioana Marinescu. 2022. "The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund." *American Economic Journal: Economic Policy*, 14(2): 315–40.

Kolsrud, Jonas, Camille Landais, Daniel Reck, and Johannes Spinnewijn. 2021. "Retirement Consumption and Pension Design." CEPR Discussion Paper 16420.

Krueger, Alan B., and Bruce D. Meyer. 2002. "Chapter 33 Labor supply effects of social insurance." In . Vol. 4 of *Handbook of Public Economics*, 2327–2392. Elsevier.

Lachowska, Marta, and Michał Myck. 2018. "The Effect of Public Pension Wealth on Saving and Expenditure." *American Economic Journal: Economic Policy*, 10(3): 284–308.

Liebman, Jeffrey B., Erzo F.P. Luttmer, and David G. Seif. 2009. "Labor supply responses to marginal Social Security benefits: Evidence from discontinuities." *Journal of Public Economics*, 93(11): 1208–1223.

Minijob-Zentrale. 2020. "Aktuelle Entwicklungen im Bereich der Minijobs. 1. Quartalsbericht 2020." Deutsche Rentenversicherung Knappschaft-Bahn-See, Essen.

Müller, Dana, and Katharina Strauch. 2017. "Identifying Mothers in Administrative Data." FDZ-Methodenreport 13/2017 (en).

Picchio, Matteo, Sigrid Suetens, and Jan C. van Ours. 2018. "Labour Supply Effects of Winning a Lottery." *The Economic Journal*, 128(611): 1700–1729.

Policy Impacts. n.d.. "The Policy Impacts Library." https://policyimpacts.org/, Accessed April 6 2023.

Schönberg, Uta. 2009. "Does the IAB Employment Sample Reliably Identify Maternity Leave Taking? A Data Report." *Journal for Labour Market Research*, 42: 49–70.

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.

Seibold, Arthur. 2021. "Reference Points for Retirement Behavior: Evidence from German Pension Discontinuities." *American Economic Review*, 111(4): 1126–65.

Statistisches Bundesamt. 1991. “Statistisches Jahrbuch 1991.” Statistisches Bundesamt, Wiesbaden.

Statistisches Bundesamt. 1992. “Statistisches Jahrbuch 1992.” Statistisches Bundesamt, Wiesbaden.

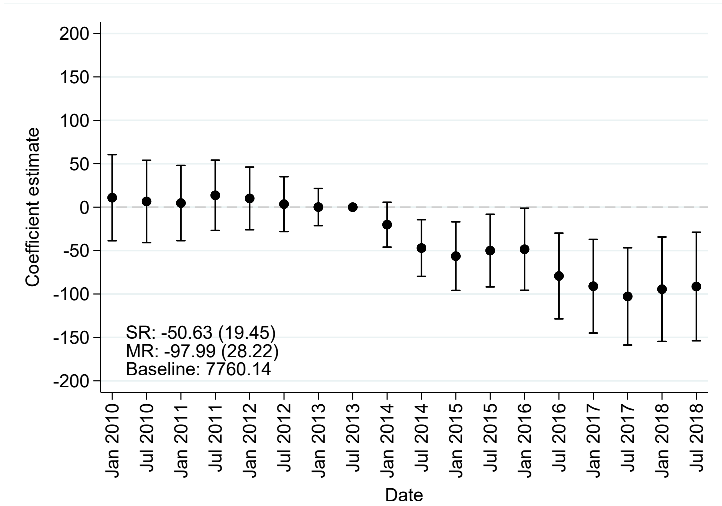
Steffen, Johannes. 2011. “Arbeitslosigkeit und Rente.” Arbeitnehmerkammer Bremen.

Thiemann, Andreas. 2015. “Pension Wealth and Maternal Employment: Evidence from a Reform of the German Child Care Pension Benefit.” DIW Berlin, German Institute for Economic Research Discussion Papers of DIW Berlin 1499.

Ye, Han. 2022. “The Effect of Pension Subsidies on the Retirement Timing of Older Women.” *Journal of the European Economic Association*, 20(3): 1048–1094.

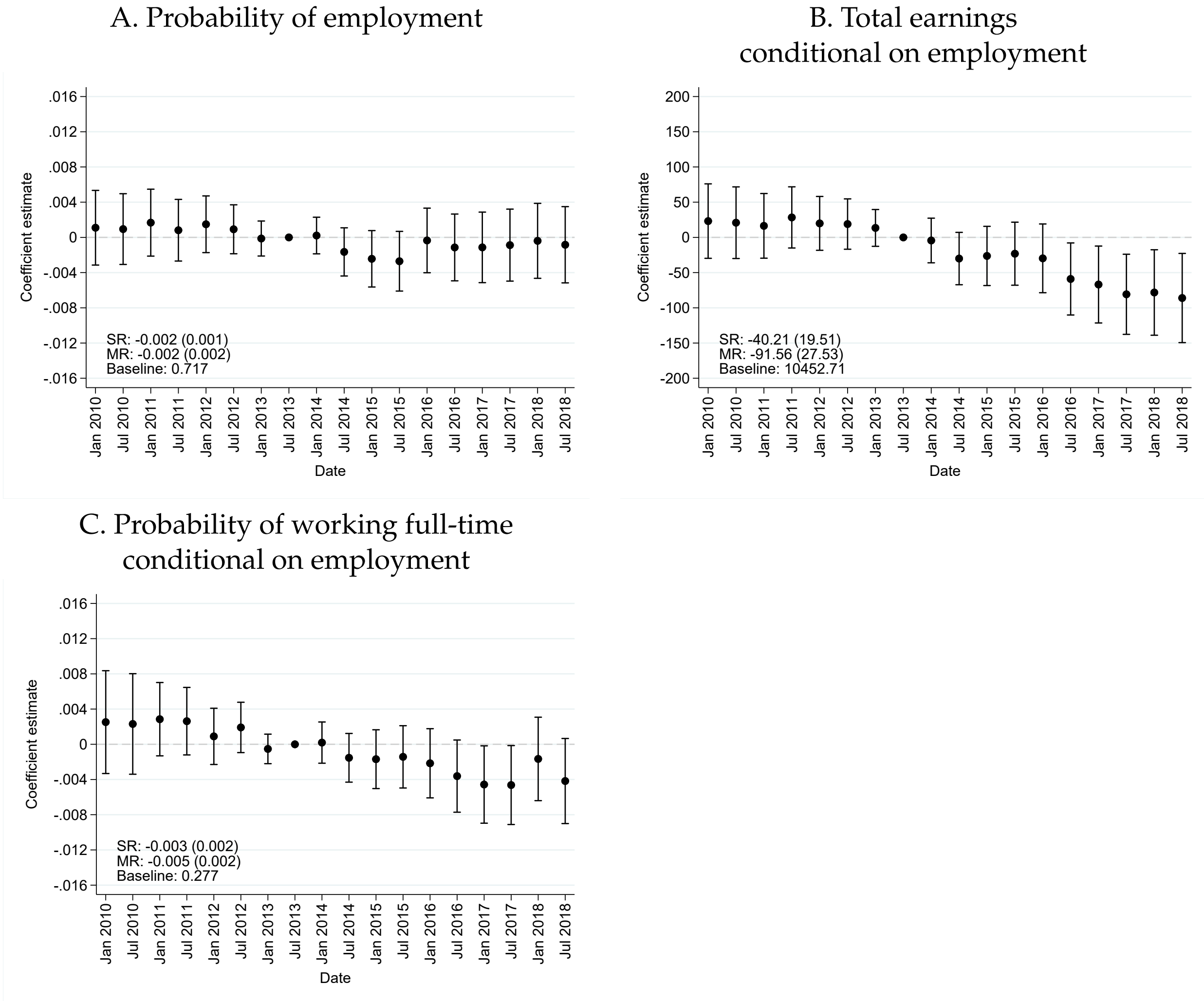
Figures

Figure 1. EFFECT OF 2014 MÜTTERRENTE REFORM ON TOTAL EARNINGS UNCONDITIONAL ON EMPLOYMENT



Notes: The graph reports 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. The 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 parenthesis. The baseline statistic is the average value of the outcome in the treatment group in July-December 2013.

Figure 2. EFFECT OF 2014 MÜTTERRENTE REFORM ON LABOR SUPPLY OUTCOMES



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 parenthesis. The baseline statistic is the average value of the outcome in the treatment group in July-December 2013.

Tables

Table 1. MAIN RESULTS

	T × Post (yr 1-2.5) (1)	T × Post (yr 3-5) (2)	Mean of dep. var. (3)	N of individuals (4)
A. Unconditional earnings and labor market status				
Unconditional earnings	-50.63***	-97.99***	7,760.14	99,104
	(19.45)	(28.22)		
Prob. of employment	-0.002	-0.002	0.717	99,104
	(0.001)	(0.002)		
Prob. of UI receipt/job-seeker status	-0.000	-0.001	0.053	99,104
	(0.001)	(0.001)		
Prob. of not being in IEB data	0.002*	0.002	0.230	99,104
	(0.001)	(0.002)		
B. Earnings and working time conditional on employment				
Conditional earnings	-40.21**	-91.56***	10,452.71	82,876
	(19.51)	(27.53)		
Cond. prob. of working full-time	-0.003*	-0.005**	0.277	82,876
	(0.002)	(0.002)		
Cond. prob. of working part-time	0.002	0.004	0.506	82,876
	(0.002)	(0.003)		
Cond. prob. of marginal employment	-0.001	0.002	0.182	82,876
	(0.001)	(0.002)		
C. Other outcomes				
Prob. of switching establishment	0.000	0.001	0.052	82,250
	(0.001)	(0.001)		
Cond. prob. of being employed in part-time intensive occupation	-0.000	0.001	0.898	82,986
	(0.001)	(0.001)		

Notes: The table reports estimates of the coefficients γ_s from equation (2) for post-reform years and for a set of different outcomes. The coefficients γ_s are estimated pooling post-reform years together. Column 1 reports estimates of the short-run effect over the 1-2.5 years after the 2014 reform (SR) and column 2 of the medium-run effect over the 3-5 years after the reform (MR), with robust standard errors clustered at the individual level in parenthesis. Column 3 reports the average value of the outcome in the treatment group in July-December 2013. Part-time intensive occupations are defined as occupations with above-median share of part-timers in the German economy. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 2. Heterogeneity analysis

	Unconditional earnings	Any employment	Conditional earnings	Cond. fulltime employment
	A. Returns to tenure in pre-reform occupation			
T × Post	-106.60**	-0.004**	-60.75*	-0.008***
	(41.46)	(0.002)	(31.66)	(0.003)
T × Post × above median	56.87	0.001	3.16	0.007*
	(54.76)	(0.003)	(44.16)	(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
	B. Physical strain of pre-reform occupation			
T × Post	-44.57	-0.000	-49.79	-0.005*
	(43.10)	(0.002)	(34.14)	(0.003)
T × Post × above median	-45.55	-0.004	-4.74	0.001
	(54.90)	(0.003)	(44.09)	(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
	C. Individual pre-reform non-childcare pension points			
T × Post	-53.99**	-0.005*	-64.11**	-0.000
	(25.62)	(0.002)	(32.38)	(0.003)
T × Post × above median	-32.60	0.004	0.70	-0.007*
	(44.70)	(0.003)	(43.79)	(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
	D. Household pre-reform non-childcare pension points			
T × Post	-68.77	-0.004	-86.25*	-0.001
	(46.72)	(0.004)	(44.14)	(0.004)
T × Post × above median	-46.15	0.005	-8.43	-0.007
	(72.08)	(0.005)	(64.70)	(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 2 (cont.): HETEROGENEITY ANALYSIS

	Unconditional earnings	Any employment	Conditional earnings	Cond. fulltime employment
	E. Partner's pre-reform labor earnings			
T × Post	-31.87	-0.005	-22.30	0.001
	(50.94)	(0.004)	(45.85)	(0.004)
T × Post × above median	-90.92	0.007	-119.19*	-0.010*
	(74.57)	(0.005)	(67.88)	(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
	F. Time to full retirement age			
T × Post	-60.30*	-0.002	-50.82	-0.003
	(32.50)	(0.002)	(31.92)	(0.003)
T × Post × above median	-20.91	-0.000	-24.05	-0.002
	(44.71)	(0.003)	(44.09)	(0.004)
Pre-reform mean of dep. var. (below median)	8,159.65	0.705	11,221.98	0.272
Pre-reform mean of dep. var. (above median)	7,364.86	0.730	9,722.14	0.281
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 parenthesis. 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 semester of 2013; in Panel B, with a dummy for being employed in an occupation with above-median index of physical strain in the second semester 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 semester of 2013; and in Panel F, a dummy for being at above-median distance from full retirement age, as of December 2013. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 3. HOUSEHOLD-LEVEL SPILLOVERS

	Unconditional earnings	Any employment	Conditional earnings	Cond. fulltime employment
	A. Heterogeneity by partner's time to full retirement age			
T × Post	-143.49***	-0.004	-120.63***	-0.009**
	(52.12)	(0.003)	(46.27)	(0.004)
T × Post × above median	107.07	0.005	59.96	0.009*
	(72.09)	(0.005)	(65.09)	(0.005)
Pre-reform mean of dep. var. (below median)	9,157.64	0.840	10,627.70	0.234
Pre-reform mean of dep. var. (above median)	7,901.56	0.829	9,228.65	0.235
N individuals	37,427	37,427	34,962	34,962
	B. Partner's labor supply response			
T × Post (yr 1-2.5)	-19.49	0.001	-49.84	-0.001
	(48.46)	(0.002)	(32.05)	(0.001)
T × Post (yr 3-5)	-44.49	0.000	-77.30	-0.004**
	(74.15)	(0.003)	(47.75)	(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 is at above-median distance from the full retirement age. 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 parenthesis. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.