CERGE - EI
Center for Economic Research and Graduate Education –
Economics Institute

Essays in Personnel Economics: Aging Workforce and Worker Substitutability

Sona Badalyan

Dissertation

Prague 2026

Dissertation Committee:
Štěpán Jurajda, Ph.D. (CERGE-EI, Chair)
Wolfgang Dauth, Ph.D. (Otto Friedrich University in Bamberg, IAB)
Randall K. Filer, Ph.D. (City University of New York)
Paolo Zacchia, Ph.D. (Ca' Foscari University of Venice)

Referees:
Simon Jäger (Princeton University)
Johannes Geyer (DIW Berlin)

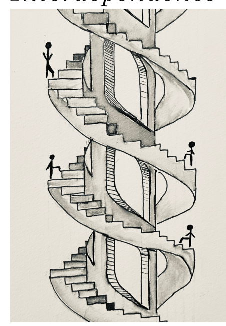
To the workforce—past, present, and future

Contents

Acknowledgements

This dissertation is about workforce interdependence. So was the path to writing it.

Interdependence



I am grateful to my dissertation chair, Štěpán Jurajda, for his feedback and optimism, which were truly motivating in achieving the completion of this dissertation, and to my committee members—Wolfgang Dauth, Randall Filer, and Paolo Zacchia. I thank my external referees, Johannes Geyer and Simon Jäger, for their insightful and detailed reviews. I also thank Deborah Nováková and Grayson Krueger from the Academic Skills Center for editing my texts.

At the Institute for Employment Research, I am grateful to Wolfgang Dauth and Michael Moritz for recognizing the potential in the early drafts of my papers and for their belief in me at a time when it made all the difference. I thank Dan Black for hosting me at the University of Chicago, where parts of this dissertation were written, and for our many discussions. I thank Nikolas Mittag for facilitating the contact with Dan, for his time discussing this dissertation with me, and for his feedback on my early drafts.

My intellectual foundations in the field were shaped during my master's studies at Central European University in Budapest, whose academic environment inspired me to pursue doctoral research and continues to influence my work.

I am grateful to Oscar Tatosian and Renita Esayian for becoming my American Armenian family, and to Randall Filer and Barbara Forbes for their encouragement and support throughout my PhD. Ich danke meiner Deutschlehrerin Ruzanna für unsere Kurse. I thank Dorota for her support.

I thank my mother, Sofya, for her unwavering support. Her quiet determination shaped my path, and I am grateful to have someone I can call even with the smallest mathematics questions. The memory of my late father, Varuzhan, gave me strength throughout this journey. I thank my brother, Ayk, for our conversations during trips.

This dissertation marks the end of a journey that changed me beyond academic growth. It took me across countries, languages, and disciplines—and into the landscapes of perseverance, self-reliance, and self-trust. This work is dedicated to all those who continue their path despite difficulties.

To the workforce—past, present, and future.

Prague, Czech Republic
2026

Sona Badalyan

This dissertation was supported by Charles University, GAUK project No. 333221, the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 870245, the Armenian General Benevolent Union, contributions from the Armenian diaspora, the CERGE-EI Foundation, and the mobility fund.

Abstract

This dissertation studies how retirement age reforms reshape labor markets beyond individual retirement decisions, focusing on firms, coworkers, and workplace interactions. Using quasi-experimental evidence from German pension reforms that raised women's early retirement age, I examine how delayed retirements affect internal labor markets, worker substitutability, and peer behavior at older ages. The first chapter shows that retaining older workers generates substantial intra-firm spillovers. Delayed retirements reduce internal promotions and external hiring of younger coworkers, with the largest losses among middle-aged workers closest to older workers on career ladders. These effects are highly structured: promotion crowd-out arises in dense internal labor markets with intense competition and operates primarily within jobcells, while coworkers in other jobcells can benefit when retained older workers possess firm-specific human capital. Hiring declines, by contrast, are concentrated in thin external labor markets where replacement is difficult, highlighting that without accounting for hiring frictions, crowd-out effects cannot be interpreted as pure measures of worker substitutability. The second chapter demonstrates that retirement age reforms help firms retain less substitutable workers with job-specific skills and managerial responsibilities, preserving valuable human capital that would otherwise be lost through voluntary exits. The third chapter shows that reform effects are further amplified through peer behavior: women are more likely to remain employed at older ages when their coworkers do so, explained by conformity, information, and work complementarity mechanisms. Together, these chapters show the interdependence of coworker careers through slot constraints, human capital, and social interactions.

Introduction

Population aging has prompted many OECD countries to raise statutory and early retirement ages in an effort to sustain public pension systems and expand the labor supply at older ages. While a large body of literature studies how such reforms affect individual retirement behavior, much less is known about how delayed retirements reshape labor markets within firms and workplaces. However, retirement age reforms are not only labor supply shocks at the individual level: by altering who remains employed, these reforms change internal career ladders, turnover dynamics, peer interactions, and the allocation of human capital inside firms. Understanding these broader spillover effects is essential for evaluating the true economic consequences of policies that extend working lives.

This dissertation studies how retirement age reforms affect firms and workers beyond the directly treated individuals. Using a series of quasi-experimental designs based on a 1999 German pension reform, the three chapters examine complementary but distinct mechanisms through which delayed retirement propagates through internal labor markets and workplace networks. Together, these chapters show that raising the retirement age reshapes employment outcomes through slot constraints, human capital preservation, worker substitutability, and peer effects, with implications that extend well beyond the retirement decisions of older workers themselves.

Chapter 1 ("Crowded Career Ladders? Intra-firm Spillovers of Raised Retirement Age") studies how retaining older workers affects coworkers' careers and firms' hiring decisions. Leveraging a 1999 German reform that raised women's early retirement age by at least three years, the chapter uses firm-level variation in exposure to the reform to identify spillovers on internal promotions and external hiring. The results show that delayed retirements substantially increase the retention of older women and reduce both internal promotions and external hiring of younger coworkers. These effects are highly structured. Promotion crowd-outs are strongest in thick internal labor markets with dense career ladders and intense competition for advancement, while hiring declines are concentrated in thin external labor markets where replacement is costly. Spillovers arise primarily within narrowly defined jobcells—where workers compete for the same positions—while coworkers in other jobcells can benefit when retained older workers possess firm-specific human capital.

This chapter provides new evidence that retirement age reforms operate through slot constraints inside firms, but that these constraints interact with turnover frictions and human capital specificity. The findings challenge the view that retaining older workers is uniformly harmful for younger employees and show that internal labor market structure critically shapes who loses and who gains.

Chapter 2 ("Retirement Age Reforms and Worker Substitutability: Implications for Employment of Older Workers") focuses on why firms retain certain older workers after a retirement age increase and how worker substitutability shapes firms' responses. Using a regression discontinuity design based on the same 1999 German reform, the chapter shows that firms disproportionately retain workers who are harder to replace—those with job-specific skills, managerial responsibilities, or limited internal and external substitutes. The chapter demonstrates that delayed retirement helps firms preserve valuable human capital that would otherwise be lost through voluntary exits. At the same time, the loss of early pension eligibility weakens workers' outside options, allowing firms to adjust wages, often via partial retirement arrangements.

These findings provide a micro-foundation for the mechanisms identified in Chapter 1: the workers whose retention generates promotion bottlenecks within jobcells are often the same workers whose firm-specific knowledge can benefit coworkers across jobcells. Taken together, Chapters 1 and 2 show that retirement age reforms reshape firms not only by blocking career ladders, but also by altering the composition of retained human capital. Slot constraints alone cannot explain the observed patterns without accounting for worker substitutability and firm-specific skills.

Chapter 3 ("Peer Effects in Old-Age Employment Among Women") studies a different but complementary channel through which retirement age reforms propagate: peer effects within workplaces. Exploiting the gender-neutralization of retirement ages in Germany as a norm-shifting shock, the chapter examines how women's employment at older ages responds to the retirement behavior of their coworkers. Using universe data on workplace networks, the chapter finds strong peer effects: women are more likely to remain employed at older ages when their peers do so, with particularly large effects in regions of the former West Germany, where traditional gender norms historically discouraged female employment at older ages. These peer effects amplify the direct impact of the reform and contribute to regional convergence in late-career employment. I find evidence of conformity, information, and work complementarity channels explaining the presence of peer effects in old-age employment patterns.

This chapter highlights that retirement age reforms do not operate solely through firms' personnel decisions or individual incentives. Instead, they also reshape social norms and expectations within workplaces, generating spillovers that extend the reach of policy changes beyond directly treated individuals.

Together, the three chapters provide a unified view of how retirement age reforms affect labor markets inside firms and workplaces. Chapter 1 documents how delayed retirements crowd career ladders and reshape hiring, Chapter 2 explains which workers are retained and why preserving their human capital matters, and Chapter 3 shows how reforms are amplified through peer interactions and social norms. By combining evidence on internal labor markets, worker substitutability, human capital specificity, and peer effects, this dissertation moves beyond a narrow focus on individual retirement behavior. It shows that the consequences of population aging and pension reforms depend critically on workplace structure and interactions—and that policies extending working lives can generate both costs and benefits that are unevenly distributed across workers.

1 Crowded Career Ladders? Intra-Firm Spillovers of Raised Retirement Age

Single-authored job market paper.[1]

1.1 Introduction

Population aging has prompted many OECD countries to raise the official retirement ages, often through transitions to gender-neutral rules that sharply increase women's statutory retirement ages. These reforms change not only the age at which individuals exit work, but also offer rare opportunities to observe how internal and external labor markets operate, and how substitution and complementarity between older and younger workers unfold. While existing studies imply that older workers may crowd out the young, they provide limited insight into the structure of these spillovers. In this paper, I leverage a German reform that increased women's early retirement age by at least three years and a well-defined firm-level identification strategy to study how retaining older women affects both internal promotions and external hiring. By examining spillovers within occupational career ladders and across occupations, I document who is most exposed to crowding-out and who may benefit from

[1]This chapter revises earlier versions previously circulated as a conference proceeding (Badalyan, 2024), as a draft on personal and job market webpages (Badalyan, 2024), and appeared as a working paper (Badalyan, 2025). Earlier drafts appeared under the titles "Firm Responses to Raising Women's Retirement Age", "Employer Responses to Raising the Retirement Age: Spillovers on Coworkers and External Hiring", "Spillover Effects of Labor Force Retainment", and the work received first prize at the Young Economists Seminar (Croatian National Bank, 2024) and second prize for Young Economist of the Year 2024 in the Czech Republic. This study uses the factually anonymous data sampled from the Integrated Employment Biographies (IEB) database of the Institute for Employment Research (IAB). Due to their administrative origin, these data are highly confidential and can only be processed on-site at the IAB by IAB staff and guest researchers. IAB has established a process to grant access to the data in case of reasonable doubt against the validity of published results. This study was supported by Charles University, GAUK project No. 333221. I thank Dan Black, Wolfgang Dauth, Manasi Deshpande, Steven Durlauf, Randall Filer, Johannes Geyer, Michael Gibbs, Joshua Gottlieb, James Heckman, Jörg Heining, Anders Humlum, Jonas Jessen, Štěpán Jurajda, Simon Jäger, Dmitri Koustas, Nikolas Mittag, Bruce Meyer, Magne Mogstad, Michael Moritz, Matthew Notowidigdo, Shogher Ohannessian, Ben Ost, Birgitta Rabe, Evan Rose, Alexander Torgovitsky, Simon Trenkle, and Paolo Zacchia for feedback; Dan Black for hosting me at the University of Chicago; Michael Moritz and Wolfgang Dauth at the IAB "Regional Labour Markets" Department for support. I thank Philipp vom Berge, Wolfgang Dauth, Jonas Jessen, Hannah Illing, Michael Moritz, Duncan Roth, Niklas Vetterer, and Katja Wolf from IAB for their help with the data; Deborah Nováková from the Academic Skills Center for language editing. This paper benefited from presentations at the WZB Workshop on the Economics of Ageing and Pensions, the RFBerlin Workshop on the Economics of Aging, the IAB NU-DE Workshop on Applied Microeconomics, the EEA Congress, the 21st International Conference on Pension, Insurance and Savings, IZA/Leiden/OECD Workshop on Recent Advances in Labor Economics Using Linked Employer-Employee Data, HUN REN, and the ifo Institute (all 2025); EALE (2024, 2023); EWMES, IAB-DiskAB, ESPE, the IZA Summer School, the Dutch National Bank, and the Young Economists Seminar by Dubrovnik Economic Conference (all 2024); SITES and the U Chicago Student Workshop (all 2023); AIEL (2023, 2022); CERGE-EI Brown Bag (2023); BSE Summer School, the Czech Economic Society, the Armenian Economic Association, the CERGE-EI Reading Group, DW (all 2022), and DPW (2020).

worker retention, opening the "black box" behind these mechanisms and challenging the prevailing view that retention of older workers is uniformly detrimental to the careers of their younger coworkers.

My analysis starts by estimating retention and overall spillover effects of the reform based on the quasi-random age structure of job cells within firms near the retirement reform age cutoff. A unique reform setting with a large rise in retirement age allows me to quantify its impacts on firms and workers. I study the 1999 reform that abolished the female pathway to early retirement in Germany, raising the early retirement age (hereafter ERA- the minimum age that a worker can start claiming a pension) by at least three years, starting from the cohort of women born in 1952. This marks the largest rise in retirement age for consecutive cohorts in Germany. Studying the direct impact of this reform, Geyer and Welteke (2021) find that abolishing the female pathway to early retirement in Germany led to a 13.5 percentage point increase in employment of directly affected women at age 60-62 (an approximately 30% increase relative to the pre-reform cohort mean). Badalyan (2025) finds slightly larger effects of 17.3 p.p., conditional on employment immediately before reaching the age of 60, that is, for women who were more attached to the labor market. Although the reform was announced when the first affected cohort of women was only 47, Geyer and Welteke (2021) do not find differential labor market behavior of older women around the cutoff until they reached 60. The pre-announcement provides the opportunity to examine not only the "*main reform*" effects, that is, in 2012-2017 as the reform was enacted, but also the "*upstream*" (that is, before affected workers reach 60 years of age) adjustments in 1998-2011, following the terminology of labor supply literature (Rabaté, Jongen, and Atav, 2024).

Causally identifying the effects of an aging workforce on firm-level demand for incumbent workers and external hires is challenging because older workers do not randomly choose to be employed at older ages, and firms that employ older workers may differ significantly from those that do not; hence, the amount of time an older worker spends working at a firm at older age may not correspond to random exposure at the firm level. To identify spillover effects, I leverage the fact that the rise in retirement age affected only cohorts born in and after 1952. For firms employing the same number of women who were born around the reform cutoff, there is a random variation in the number of women born just before or after the cutoff. This gives rise to quasi-random treatment intensities in older worker retention, thanks to the reform. I focus on firms that employed at least one woman born near the reform cutoff, that is, in 1950-1953 (I call these *focal* workers). I compare firms with similar observable characteristics (including the total number of focal workers and total workforce) but that plausibly differ exogenously in terms of the numbers of their workers who were born just after the reform cutoff (1952–1953 cohorts, i.e., *treated focal workers*). I utilize a generalized difference-in-difference approach in the retirement setting (Hut, 2024) to analyze the effect of an *additional* treated focal worker, who was subject to the rise in the ERA, on the demand for workers within the workplace and from external sources.

Germany offers large, high-quality social security data on establishments—single locations of multisite firms.[2] The data allow me to observe the universe of workers of affected

[2]Throughout this paper, I use the terms establishments and firms interchangeably.

establishments (that is, those with at least one focal worker) and the full employment histories of all workers employed at these establishments. I use two samples because I am interested in upstream adjustment strategies and in the main reform effects during the years when focal workers reached the ages of 60-65. The data for upstream effects comprise private-sector firms observed in 1998 (the pre-reform year). The data include approximately 140,000 establishments, about 250,000 focal workers, and roughly ten million coworkers employed in those establishments. The data for the main reform effects are sampled similarly, except that they are based on firms that employed at least one focal worker in 2008, two years before all the focal worker cohorts were younger than 60. It consists of approximately 160,000 establishments of all sizes, and over 400,000 focal workers and eight million workforce employed in 2008.

I establish the following set of results. First, I start with the upstream period and study the direct effects of this reform on retention at older ages of focal workers. An additional treated focal worker employed in 1998 leads to increased focal worker retention, but only after workers turn 55 years old. Second, I test whether such increased retentions and competition for internal promotions lead to downward pressure on the promotions of younger workers in the upstream period, thus hindering the career progression of coworkers and influencing external hiring practices. I do not find significant spillover effects on coworker promotions or on external hiring in the upstream period.

Because firms show virtually no adjustment in the upstream period, I re-sample the analysis in 2008—before any treated workers reached pensionable ages—which provides a fresh firm cohort and improves precision. Turning to the main reform period, I find that having an additional treated worker exposed to the rise in the ERA in 2008 leads to approximately 0.163 more focal worker retentions, generating 0.075 fewer coworker promotions and 0.103 fewer external hires in 2012–2017. Scaled by the increase in focal worker retention, these estimates imply that each additional older worker retained reduces external hiring by about 0.63 workers and about 0.46 fewer coworker promotions. The magnitude of these spillovers aligns with evidence from shocks that remove rather than retain workers.[3]

Next, I examine which demographic segments (by gender and age) are most affected by the reform. Lazear and Oyer (2004) find that internal hiring (promotions) constitutes 44-88% of job postings, with the highest percentage occurring among the highest level of occupational hierarchy. If older workers occupy high-ranked jobs, their retention could slow the career progression of middle-aged workers who are closer to them on the career ladder. The results reveal that middle-aged workers—especially women—experience the strongest crowd-out in promotions, consistent with their being the closest substitutes for older women who remain employed.[4] The crowd-outs on external hiring do not display such gendered patterns.

[3]Employers replace workers who die suddenly by increasing hiring by about 0.4 workers (Jäger and Heining, 2022), and hire 0.35 additional workers when women go on maternity leave in Denmark (Brenøe et al., 2024), and about 0.3 in Germany (Huebener et al., 2024).

[4]This pattern is consistent with Carta et al. (2024) in parental-leave settings, who show that firms tend to replace women on leave with female workers, and with Ginja, Karimi, and Xiao (2023), who document gender-specific spillovers following paternity leave.

Understanding how delayed retirements affect firms requires distinguishing between two competing theories of internal labor markets. One view, rooted in incentive-contract models (Lazear, 1979), holds that long-tenure workers are costly to retain because their wages exceed their marginal product of labor; when retirement is postponed, firms face higher labor costs. These effects arise even in the absence of deferred compensation, because many firms operate under *slot constraints*—a fixed number of positions within each occupational ladder—so that when an older worker stays longer, the slots available for promotions or external hires shrink mechanically, reducing opportunities for internal promotions and external hiring (Boeri and van Ours, 2021). This mechanism predicts negative spillovers on younger workers, especially in hierarchical labor markets such as Germany's, where a large share of lifetime wage growth reflects career ladders (Bayer and Kuhn, 2018).

A second view emphasizes the productivity advantages of older long-tenure workers arising from accumulated firm- and job-specific human capital (Bartel et al., 2014; Friedrich and Hackmann, 2021; Jäger and Heining, 2022; Jaravel, Petkova, and Bell, 2018). When skills are highly specific, internal or external hires are imperfect substitutes (Chan, 1996; Herrmann and Rockoff, 2012; Waldman, 2003), making retention of older workers valuable to employers and potentially beneficial for coworkers through complementarities. Replacement hiring is costly—around two months' wages for high-skilled workers (Muehlemann and Pfeifer, 2016), even understating true turnover costs (Bertheau et al., 2022)—therefore, firms often move workers internally rather than recruiting externally (Becker, 1962; Bertheau, 2021).

Recent evidence supports the relevance of these frictions. Badalyan (2025) shows that the 1999 reform disproportionately increased retention among older workers who are costly to replace—managers, workers in highly specific occupations, and those in firms with few internal or external substitutes—underscoring that turnover frictions and firm-specific skills shape which older workers remain employed. In this paper, I show that these same forces also determine how delayed retirements spill over onto promotions of coworkers and onto external hiring, helping to differentiate between the incentive-contract view and the human-capital/complementarity view.

To shed light on these theories, I estimate the structure of worker substitution across occupations and age groups. I define "true" substitution elasticity as the elasticity that applies in an ideal case when a suitable younger worker is easily available (at low search costs) on the internal or external labor market. To uncover this "true" elasticity, I decompose the estimated crowd-out (spillover) effects of older worker retention into the element corresponding to labor market frictions (availability of suitable replacements) and the remaining part, which corresponds to the "true" elasticity of substitution between older focal workers and external hires. I do this separately for promotions and for external hiring, because I have ideal proxies for labor market thickness across both dimensions of worker replacement. My elasticities differ from those in the retirement literature because I can control the frictions firms face in external labor markets and the competition on career ladders. I outline a conceptual framework that characterizes the firm's decision problem, explicitly incorporating these frictions.

Recruiting new staff depends on the availability of suitable external candidates, which varies sharply across space because workers are not perfectly mobile across sectors or regions (Yi, Müller, and Stegmaier, 2024). In a frictionless market, spillovers from delayed retirements would directly reveal the elasticity of substitution between older and younger labor. But when external hiring is costly or constrained, firms cannot easily replace older workers retained, and turnover frictions inflate the observed crowd-out. Thin external labor markets—commuting zones and industries in which suitable replacement workers are scarce—should therefore exhibit larger hiring declines, allowing me to separate true substitutability from constraints imposed by local hiring frictions. In line with this mechanism, I find that the negative hiring effects are concentrated in thin external labor markets. In thick markets, where turnover frictions are weaker, hiring responses are much smaller. These patterns show that external labor market thickness is a key mediator of spillovers from delayed retirements and that interpreting baseline crowd-out estimates as structural elasticities of substitution would be misleading unless one accounts for local hiring frictions.

Internal labor-market thickness also plays a central role. Several predictions that are consistent with those outlined in the model follow. First, crowding-out should be larger for promotions than for hiring, because incumbent workers are more substitutable than external hires. Second, crowding-out should be strongest within the occupations of older workers—especially in jobcells[5] with many coworkers competing for the same rungs on the ladder—while workers in other occupations may benefit from cross-occupation complementarities (Jäger and Heining, 2022). In establishments in which many coworkers share the older worker's occupation, promotion losses are substantially larger, reflecting more intense competition for limited slots. Because the reform increased promotions of focal workers by about 0.02 per additional focal worker, it pushed more workers into top positions and intensified congestion along the career ladder. This scarcity of advancement slots contributed to the fourfold crowd-out of coworker promotions. These patterns also appear in wage-bill adjustments: while focal workers' wages rise, coworker wage bills decline, especially in thick internal labor markets. Employers thus adjust the composition, rather than the scale, of their workforce, implying compressed hierarchies and persistent career costs for coworkers.

Imperfect substitutability among workers within a firm (Chan, 1996; Herrmann and Rockoff, 2012; Waldman, 2003) implies that retaining older workers with accumulated firm-specific knowledge can yield benefits, in line with human-capital theories. To test whether such complementarities operate across rather than within occupations, I decompose spillovers in multi-jobcell firms. Within jobcells—where tasks and ladders overlap—the promotion crowd-out is substantial. Across jobcells, however, I find no systematic declines and, when older workers possess substantial firm- or task-specific human capital, I find small positive wage spillovers. This highlights that complementarities among workers on distinct internal ladders can offset crowd-out pressures, consistent with classic and modern theories of firm-specific human capital (Becker, 1962; Lazear, 2009).

Taken together, the results show that spillover magnitudes and signs depend jointly on limited external substitutability, cautioning against interpreting crowd-out estimates on external hiring as pure elasticities of substitution, and on the concentration of skills within internal

[5] I define jobcells as 3-digit occupation groups within establishments.

labor markets, highlighting the value of older workers.

This paper contributes to the retirement literature by shifting focus from individual labor supply responses[6] to the smaller but growing body of work on intra-firm spillovers of rising retirement ages, thereby linking the retirement and internal labor market literature (Doeringer and Piore, 1971; Lazear and Oyer, 2004).[7] Existing evidence—mainly from Italy (Bianchi et al., 2023; Boeri, Garibaldi, and Moen, 2022; Carta, D'Amuri, and Von Wachter, 2024), and the Netherlands (Hut, 2024; Ferrari, Kabátek, and Morris, 2023)—offers mixed findings on whether older worker retention crowds out or complements younger colleagues.[8] One explanation for these discrepancies lies in firm heterogeneity: positive effects in Carta, D'Amuri, and Von Wachter (2024) may reflect differences in firm characteristics, such as firm size and underlying hiring practices correlated with such characteristics, as their heterogeneity analysis highlights that positive impacts are concentrated in larger firms. This paper provides new evidence from Germany, a large and institutionally distinct labor market, and offers a reconciliation of prior puzzles through incorporating internal and external labor markets into spillovers and novel mechanisms. The evidence presented here also contributes to earlier related literature that examines aggregate trade-offs between retirement and youth unemployment; policies in the 1980s and 1990s promoted early exits with this objective (Gruber and Wise, 2010), but the earlier studies find little systematic support and provide limited insight into underlying mechanisms. The nuanced answer to this question appears to be that the effect depends on the state of the internal labor markets and on the human capital of the older workers.

Unlike settings in which financial frictions shape workforce adjustments (Hut, 2024), the long pre-announcement horizon of the German reform makes liquidity constraints unlikely.[9] My findings are broadly consistent with slot-constraint logic (Bianchi et al., 2023): additional retention of older workers is nearly fully offset by reduced promotions and hiring among younger workers. Nevertheless, the average pattern conceals important heterogeneity. When suitable external replacements are scarce, turnover frictions amplify hiring crowd-out—echoing

[6]See, among others, Geyer and Welteke (2021), Lalive, Magesan, and Staubli (2023), Manoli and Weber (2016), Mastrobuoni (2009), and Ye (2020).

[7]Related work studies firm-level spillovers and worker substitutability in settings of labor force exit, such as sudden worker death (Becker and Hvide, 2022; Bennedsen, Pérez-González, and Wolfenzon, 2020; Bertheau et al., 2022; Jäger and Heining, 2022; Isen, 2013; Illing, Schwank, and Tô, 2024; Poege et al., 2025; Sauvagnat and Schivardi, 2024), workers quitting (Kuhn and Yu, 2021), emigration (Dicarlo, 2022), childbirth and parental-leave absences (Bonney, Pistaferri, and Voena, 2025; Brenøe et al., 2024; Carta et al., 2024; Schmutte and Skira, 2023; Gallen, 2019; Ginja, Karimi, and Xiao, 2023; Corekcioglu, Francesconi, and Kunze, 2025; Friedrich and Hackmann, 2021; Huebener et al., 2024), and the introduction of paternity leave (Johnsen, Ku, and Salvanes, 2023).

[8]Further related work studies spillovers of retirement or labor force exit at different levels of aggregation. Evidence on intra-firm and local labor market spillovers includes studies from Portugal (Martins, Novo, and Portugal, 2009), Norway (Hernæs et al., 2023), and the 1992 German reform (Berg et al., 2025), as well as commuting-zone analyses in the US (Mohnen, 2025) and Italy (Bertoni and Brunello, 2021), and cross-country evidence (Kalwij, Kapteyn, and De Vos, 2010). Differences in institutional context, data coverage, and identification strategies make these studies informative but not fully comparable to the setting examined here.

[9]While prior work examines upstream labor supply responses to pension reforms (Carta and De Philippis, 2024; Mastrobuoni, 2009; Rabaté, Jongen, and Atav, 2024; Staubli and Zweimüller, 2013), I provide the first evidence on upstream spillovers within firms.

evidence on high replacement costs and thin labor markets (Ginja, Karimi, and Xiao, 2023; Jäger and Heining, 2022; Huebener et al., 2024; Schmutte and Skira, 2023). Inside a firm, promotion spillovers depend on internal structure and human-capital specificity: using establishment–occupation identifiers, I show that within jobcells—where tasks and ladders overlap—retained older workers intensify competition and reduce promotions of coworkers, whereas across jobcells their firm- and task-specific expertise can complement those of younger coworkers. Together, these patterns suggest that neither liquidity constraints nor pure slot-constraint mechanisms alone explain firms' responses. Instead, delayed retirements interact with thin external markets and concentrated internal skill hierarchies—consistent with human-capital and internal labor-market theories (Baker, Gibbs, and Holmstrom, 1994; Huitfeldt et al., 2023). This unified view helps reconcile divergent findings across countries and firm types as being possibly related to variations in hiring frictions and competition hidden in internal labor market structure, and cautions against treating crowd-out estimates as direct elasticities of substitution.

The rest of the paper is organized as follows. Subsection 1.2 describes the institutional setting. Subsection 1.3 describes the data source, the sample construction, and the corresponding identification strategy. Subsection 1.4 presents the baseline results, followed by subsection 1.5, which outlines a simple model of firm decisions and quantifies the elasticity of substitution. Subsection 1.6 shows intra- and inter-jobcell spillovers. Finally, I conclude in subsection 1.7.

1.2 Institutional Setting

Features of the labor market in Germany.

Wage setting and rigidity. Germany features relatively decentralized wage setting, allowing firms to deviate from collective agreements (Dustmann et al., 2014; Jäger and Heining, 2022). Despite this flexibility, wages remain downward-rigid due to unions and incentive contracts for lower-skilled workers, and to firm-specific human capital for higher-skilled workers (Franz and Pfeiffer, 2006).

Employment protection. Older workers are strongly protected under the Equal Treatment Act (AGG),[10], and severance pay rises steeply with tenure (Hut, 2024). Germany has relatively stable jobs and relatively high severance pay, compared to countries such as the US. Together, wage rigidity and dismissal protections imply that deferred compensation (Lazear, 1979) makes separating from older workers costly, so firms are more likely to adjust to workforce aging by reducing hiring rather than by dismissing older workers.

Industry segregation by gender. Germany exhibits pronounced gender segregation across industries. Table 1.D.1 shows that women are overrepresented in service-oriented sectors, while men dominate goods-producing and infrastructure-related industries. A comparison between Panels A and B indicates that this pattern remains largely stable over time.

[10]General Act on Equal Treatment of 14 August 2006 (Federal Law Gazette I, p. 1897) as last amended by Article 4 of the Act of 19 December 2022 (Federal Law Gazette I, p. 2510).

Key features of the German public pension system. The German pension system has three main pillars: public, occupational pensions, and private provisions. Public pension insurance is the most popular choice among the working population, covering approximately 90% of the German workforce (Geyer and Welteke, 2021; Zwick et al., 2022). The public pension system consists of a pay-as-you-go scheme, where the retirement pensions are financed by social security contributions of the insured workers and taxes.

Germany has two statutory retirement ages: the early retirement age (ERA), the earliest age at which a pension can be claimed, and the normal retirement age (NRA), the earliest age to claim a full pension without actuarial deductions. On the regular pathway to retirement, which requires only five years of social-security contributions, the ERA equals the NRA. Reduced ERAs exist only for special pathways targeted at specific groups— including women, long-term insured workers, and the unemployed—upon them meeting the relevant eligibility criteria. Retiring between the ERA and NRA entails actuarial deductions of 0.3% per month.[11] Workers respond strongly to these statutory ages, which function as reference points and generate pronounced bunching (Seibold, 2021). Consequently, reforms that shift the ERA or NRA induce significant adjustments in labor supply (Riphahn and Schrader, 2021; Geyer and Welteke, 2021). Importantly, the 0.3% monthly deduction is low by international standards and not actuarially neutral (Queisser and Whitehouse, 2006), making ERA claims particularly attractive. Cross-country evidence shows that ERA are more successful in raising the effective age of retirement than changes in NRA (Boeri and van Ours, 2021). Although Germany does not mandate retirement, continued employment beyond the NRA requires contract renewal, which employers can decline. These institutional features make early-retirement reforms especially well-suited for studying intra-firm spillovers of workforce aging.

The 1999 reform: Abolishment of the women's early-retirement pathway. Before 1999, women could claim the *"Old-Age Pension for Women"* at age 60 if they had accumulated at least 15 years of social security contributions, including ten after age 40; roughly 60% of the 1951 cohort qualified (Geyer and Welteke, 2021). The 1999 reform abolished this pathway for women born on or after January 1, 1952, creating a sharp discontinuity in retirement eligibility. Another common route into early retirement was the *"Long-Insurance Pathway"*, available to workers with sufficiently long contribution histories. For most affected women—about 90% of those previously eligible, who also met the 35-year contribution requirement for the long-insurance pathway—the early retirement age (ERA) rose from 60 to 63. Women eligible only for the *"Regular Pathway"* experienced an even larger increase of up to 5.5 years.[12] Further institutional details, such as the description of pathways, are provided in section 1.7.

Discontinuity in birth cohorts and labor supply responses. This discontinuous three-year rise in ERA is the largest statutory shift affecting adjacent cohorts in recent years in Germany (Panel A of Figure 1.1). It generated substantial labor-supply responses: Geyer and Welteke (2021) document a 13.3 p.p. increase in employment at ages 60–62 with no offsetting rise in disability, unemployment, or inactivity. Conditional on employment at ages 58–59, the

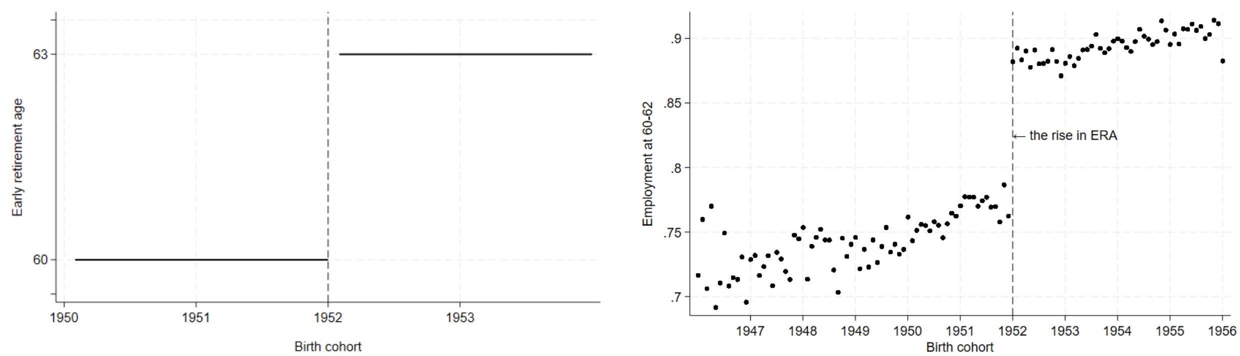
[11] For a woman in the 1951 cohort, retiring at age 60 implies a 10.8% permanent reduction.

[12] The regular pathway retirement age also increased due to the 2007 reform, which raised the NRA in small increments; thus, the NRA rose from 65 for the 1951 cohort to 65.5 for the 1952 cohort.

employment discontinuity at ages 60–63 reaches 17.3 p.p. (Badalyan, 2025) (Panel B in Figure 1.1). Consistent with these patterns, Figure 1.C.1 shows pronounced bunching at the relevant ERA (60 and 63) and NRA thresholds (65 and 65.5) for the 1951 and 1952 cohorts.

Figure 1.1: Early retirement age and employment at ages 60–62: Discontinuities across birth cohorts

Panel A: Assignment rule **Panel B:** Scatter plot (Badalyan, 2025)



Notes: **Panel A** shows the policy rule for the earliest age a woman could claim a pension by birth cohort. **Panel B** plots the fraction of women employed at ages 60–62 across birth cohorts 1947–1956. The dashed line presents the birth cohort cutoff, January 1952, starting from which the ERA rose by at least three years.

Reform timing: upstream, main reform, and downstream periods. The 1999 reform was approved on January 1, 1999 (Gohl, 2023), when the first affected cohort was only 47 (see Figure 1.A.1). Because the first treated cohort reached the pre-reform early retirement age (ERA) of 60 in 2012, the policy generates three analytically distinct periods. The upstream period spans 1999–2012, when firms could anticipate the reform and adjust their workforces before workers reached ERA. The main reform period covers the years when the 1952 cohort transitioned through the pre- and post-reform ERAs (ages 60–62), during which the delayed-retirement shock materializes. The downstream period begins when affected cohorts pass the post-reform ERA and regain the option to exit at age 63. While the main reform period is the core focus, the upstream and downstream periods are informative about anticipatory adjustments and post-shock recovery.

The reform offers a useful case to study firm responses to workforce aging. Its sharp, cohort-based three-year increase in early retirement age provides clean identification and primarily affected women, limiting general-equilibrium concerns, such as shifts in industry composition by gender. The pre-announcement enables analysis of anticipatory behavior. Because firms cannot foresee exits at the early retirement age—contracts end only at the NRA—the reform generated unexpected retention, plausibly affecting younger workers who would normally be replacing retirees.

1.3 Data and Empirical Framework

Identifying the effects of increased older-worker employment on coworkers and external hires is difficult due to nonrandom selection of workers into late-career employment based on their unobserved characteristics and of firms that employ more older women. I leverage the 1999 reform, which exogenously raised women's early retirement age. I first describe the data source and sample construction—private-sector establishments with 5 to 500 employees that employed at least one focal worker in 1998 (upstream analysis) and 2008 (main reform period). I then outline the outcome variables, descriptive patterns, and the identification strategy used to estimate the reform's impact on coworker promotions and on external hiring.

1.3.1 Data Source and Sample Construction

I proceed in two steps. First, I describe the data source, and next, the sample construction necessary to identify the intra-firm spillovers.

Integrated Employment Biographies Database. The source of the dataset used in this paper is the Integrated Employment Biographies (IEB) database, provided by the Data- and IT-Management (DIM) at the Institute for Employment Research (IAB).[13] It is based on the integrated notification procedure for health, pension, and unemployment insurance. Data are collected from employers on all of their employees subject to Social Security. Hence, such data exclude workers with self-employment spells and civil servants. Such exclusion does not matter for this study, because participation in the public pension system in Germany is mandatory for everyone except self-employed and civil servants. Employers must file notifications of their workforce at least once each year, by June 30$^{\text{th}}$, or whenever there is a change in employment spells, such as the start of employment, exit, or change of a contract. These data include all the workers subject to social security in Germany until 2021, with a starting date in 1975 for West Germany and 1992 for East Germany.[14]

The data include precise day-to-day information on the start and end dates of employment spells and wages, which include overtime pay and bonus payments.[15] The data are rich in demographic, occupational, and establishment-level variables. The demographic variables include birth month and year, gender, education level, nationality, and district-level place of residence. The workplace data include detailed 3-digit occupational codes based on the 1988

[13] I use data from the full universe of German employment records *(IEB, version 17_00_00_202212)* of the IAB. Due to its administrative origin, these data are confidential and can only be accessed on-site at IAB. Access for guest researchers requires clearance from the German Federal Ministry of Labour and Social Affairs.

[14] Throughout the paper, I refer to East Germany to define New Länder (including Berlin), and to West Germany as the current regions of the former Federal Republic territory.

[15] Wages are top-coded. Because the reform I study affected primarily women, and given that due to occupational and industry segregation by gender, women are less likely to pass the threshold, potential wage correction would impact only a small fraction of workers in my sample (Drechsler, Ludsteck, and Moczall, 2023); hence, I refrain from performing such imputations.

classification of occupations (5-digit occupational variables starting from 2011), contract type by the number of working hours (part-time, full-time), and employment type (for example, regular employment opposed to traineeships), etc. The establishment variables include 3-digit industries and the district-level location of the establishments. The detailed occupation and industry codes allow me to observe teams within establishments and to count the number of available internal and external substitutes for retiring workers, as described in subsection 1.5.

The data consist of a universe of single locations of multi-site firms, that is, establishments.[16] This data feature is a significant advantage in my research, because I can identify the spillovers on local coworkers. Nevertheless, I follow the existing literature and use the terms establishment and firm interchangeably throughout this paper (Card, Heining, and Kline, 2013; Dustmann, Ludsteck, and Schönberg, 2009; Jäger and Heining, 2022), because single-establishment firms constitute the majority of the data (Jäger and Heining, 2022). More information about the data can be found in Jacobebbinghaus and Seth (2007).

Sample Construction for the Upstream Period

First, I construct a worker–year panel following Dauth and Eppelsheimer (2020), aggregating annual records as of June 30$^{\text{th}}$—the date when employers submit annual workforce notifications and the reference point used in the IAB Establishment Panel.[17] I define *focal workers* as female employees born within two years of the 1952 cutoff—the first cohort affected by the abolition of the female early-retirement pathway—excluding miners and sailors, whose special retirement rules are not identifiable in the data (Lorenz et al., 2018).[18] I retain all establishments that employed at least one focal worker in 1998. Sampling firms prior to the reform avoids endogeneity arising from workforce adjustments after the policy change. The identification strategy compares firms with similar baseline characteristics and the same total number of focal workers, but quasi-random variation in the number born just after the cutoff (the treated focal workers; see subsection 1.3).

I restrict the sample to private-sector establishments[19] with 5–500 employees in 1998. I exclude the public sector because employment dynamics differ markedly (Oberfichtner and Schnabel, 2019) and substitutability responses may be muted by politically fixed budgets (Ginja, Karimi, and Xiao, 2023). Firms with fewer than five workers are removed to avoid cases in which focal workers are firm owners. Establishments with more than 500 employees are excluded due to administrative data limits; this threshold corresponds to the 98$^{\text{th}}$ percentile of the size distribution and is standard in spillover studies using similar identification strategies (Ginja, Karimi, and Xiao, 2023). After defining the firm sample, I keep all workers employed in these establishments from 1995 to 2019—four pre-announcement years for baseline trends and extending until all focal workers have reached ERA and NRA. I restrict the sample to workers in standard employment relationships (employees liable to social security and those

[16]The assignment of establishment identifiers is based on ownership, location at the municipality level, and industry. Firms may be assigned multiple establishments if they belong to different sectors and/or locations.

[17]June 30$^{\text{th}}$ aligns reporting across datasets and captures the administrative employment stock.

[18]For details, see Deutsche Rentenversicherung Knappschaft-Bahn-See.

[19]Public-sector industries are defined as 5-digit classifications beginning with 84–85 and 99, based on 2008 classification of industries.

in vocational training). As a result, individuals in partial retirement are not classified as employed in this analysis.[20] These restrictions (sampling in 1998, at least one focal worker, private sector, and firm size 5–500) and their impacts on sample size are summarized in Panel A of Table 1.D.2.

Sample Construction for the Main Period

To estimate the effects of the reform between the pre-reform ERA and post-reform NRA, I employ the same sampling steps as above, except that now, instead of sampling the firms in the pre-reform year 1998, I sample them in 2008- a year in which all focal workers were under the age of 60. Resampling is necessary because many firms appear only after 1998 and employ focal workers. Panel B in Table 1.D.2 records the sample size after each sample restriction. The sample for the main reform period consists of 160,667 establishments that employed 1,234,969 workers in 2008, 414,209 of which were focal workers. Over 94% of these focal workers would be eligible for the women's pathway to retirement if it were not abolished.[21] Focal workers constitute approximately 10% of the total workforce on average (Panel A in Figure 1.C.7).

To assess how the selection criterion affects my sample, Table 1.D.3 and Table 1.D.4 show comparisons of establishments in the analysis sample with the random sample of all the establishments in Germany in 1998 (Panel A), the pre-reform year, and 2008 (Panel B). The sampled establishments have, on average, more women and part-time workers, which is expected, given that I keep only firms that have at least one focal worker in my sample in 1998 or 2008. The sampled establishments and random samples have roughly equal likelihood of being located in the West or East, and overall, the industry composition of the establishments is also similar, except the construction sector is underrepresented, while the health sector is overrepresented in my analysis sample. This is likely because there is some gender segregation in industry employment in Germany (see the discussion in subsection 1.2). As noted above, public sector establishments are absent in the data by construction.

1.3.2 Main Outcome Variables

Below, I describe the main outcome variables used in this paper. There are two main outcome variable groups- turnover and profitability outcomes, and they are defined similarly for the upstream and main reform period.

To measure the turnover variables, I count the number of workers in given year t who are *hired* and *separated* from the establishment. To count establishment-level hiring, promotions, etc, I need to define these outcomes at the individual level and then only aggregate them to

[20] Partial retirement corresponds to separate employment status categories in the data (e.g., codes 103 and 119) and is therefore excluded by construction. While partial retirement was an important pathway during this period, it reflects a transition out of standard employment relationships. The focus of this paper is on how retention of older workers in standard employment affects internal labor markets within firms.

[21] Own calculations based on full employment biographies of focal workers.

years and firms.

To capture heterogeneity in career-stage exposure to the reform, these turnover variables are further decomposed by age group, distinguishing young workers (ages 15–34) from middle-aged workers (ages 35–54), and by gender, as the reform primarily extended the working lives of women.

External hires and separations. I define *external hires* as workers employed in year t but not $t-1$. These workers can have employment biographies. For example, they may come from different establishments, from nonemployment, or be new graduates who have just entered the labor force. Similarly, *separation* is recorded if employment is recorded in a current year but not in the following year.

Promotions. I define *promotion* based on Ginja, Karimi, and Xiao (2023) and Bronson and Thoursie (2019), which rely on the relative real wage growth of individuals within a firm.[22] First, I deflate wages by consumer price indices in 2015, and, following the recommendation of Drechsler, Ludsteck, and Moczall (2023), I allocate lump-sum payments to regular employment spells weighted by spell length. Own-wage growth is computed as a logarithmic difference in real wages relative to the previous year. As a next step, I compute the mean wage growth of coworkers in the establishment. The difference between one's own and coworkers' wage growth shows the relative real wage growth. I define promotions as a dummy if an employee's own real wage growth is at least ten log points higher than their mean coworker wage growth. This measure of promotions captures wage growth through establishment-level wage hierarchies and excludes wage increases that are due to reasons unrelated to promotions, such as collective bargaining or firm performance. This measure captures upward movements within the firm's internal wage hierarchy rather than formal job title changes. In settings with structured wage ladders, such relative wage increases are closely aligned with promotions, as they reflect movement to higher-paying positions within the establishment. While this proxy does not directly observe occupational transitions, it is widely used in administrative data where job titles are either unavailable or are measured with error (Ginja, Karimi, and Xiao, 2023; Bronson and Thoursie, 2019).[23]

I construct promotion measures separately for focal workers and coworkers (nonfocal workers). I define the coworker promotion measure excluding focal workers, so that it captures spillover effects on other workers without mixing in direct effects of the reform.

[22]Although German social security data include 5-digit occupation codes, they are not a reliable measure of promotions. Establishments often fail to update codes when workers advance, many occupations lack a hierarchical 5-digit structure that maps onto internal ladders, and the detailed codes are only available from 2011. As a result, administrative occupation changes would misclassify both upward and lateral mobility. I thank Katja Wolf and Wolfgang Dauth for these insights.

[23]In the papers on intra-firm spllovers, Bianchi et al. (2023) and Ferrari, Kabátek, and Morris (2023) apply a 10% increase of wages relative to the previous year. I prefer relative wage growth over individual wage growth because the measure is defined relative to coworkers within the same establishment, so that it nets out common wage shocks, including collective bargaining and firm-wide pay adjustments. While individual bonuses or renegotiations may still contribute to observed wage growth, such changes are unlikely to systematically generate large relative deviations across workers within the same firm.

Changes in coworker wages could reflect both working hours and promotions. It is difficult to test for the intensive margin adjustments in German social security data because working hours are available only for certain years, and the only variable available throughout the years is the part-time versus full-time indicator.[24] I define promotion conditional on a constant working-time status (part-time or full-time). I exclude workers who switch between part-time and full-time employment from the promotion measure to avoid mechanically classifying such transitions as promotions. Negative wage growth may still occur due to changes in hours within these categories, but such adjustments are unlikely to generate large relative wage increases that meet the promotion threshold. Hut (2024) finds no effect on hours worked or hourly wages in a similar reform in the Netherlands, and argues that these outcomes are set in collective labor agreements and are difficult to change (Cahuc, Carcillo, and Le Barbanchon, 2019); therefore, most of the wage effects captured likely reflect the promotions.

After defining the main treatment and outcome variables, such as individual-level hiring and promotion, among others, I aggregated them at the establishment level to observe the employment and wage dynamics of the entire workforce employed from 1995.

Wage bills. Given that I do not observe profit outcomes for the establishments, I rely on the existing literature (Dustmann et al., 2022; Huebener et al., 2024) and use (1) wage bills, and (2) firm closure, defined as a transition to zero employment with no subsequent re-employment (Huebener et al., 2024).

1.3.3 Identification: Generalized Difference-in-Differences

In an ideal experiment, firms employing workers near retirement age would be randomly assigned to a regime in which the pensionable age is increased or left unchanged. In Germany, however, the pension reform applied uniformly to all women born after 1951, implying that exposure to the reform varies across firms only through differences in workforce composition. Firms employing substantially younger or older workers may differ systematically, creating a key identification challenge.

To address this concern, I restrict attention to establishments that employed at least one woman born within a narrow two-year window around the 1952 cutoff. This restriction ensures that firms are comparable in workforce composition while generating quasi-random variation in exposure to the reform through the number of women born just after the cutoff. Firms that happened to employ more women born in 1952–1953 in 1998 (for upstream analyses) or 2008 (for the main reform period) were mechanically more exposed to the increase in the early retirement age.

I exploit this variation to estimate the effect of employing an additional treated focal worker—i.e., a woman whose retirement age was raised—on firm-level outcomes, including older-worker retention, coworker promotions, and external hiring. The strategy is implemented

[24] Fitzenberger and Seidlitz (2020) argue that more than half of women have been employed part-time in recent years in Germany.

separately for the upstream period and the main reform period, when no treated worker had yet reached retirement eligibility.

Upstream effects. To identify the firm responses to their labor inputs (coworkers and hired workers), I follow an identification similar to that employed by Hut (2024) and compare establishments with a similar workforce composition (total number of focal workers in the reform year) but with a variation in the number of treated focal workers, that is, workers who experienced a rise in the ERA (see Figure 1.A.2 for the graphical illustration of identification strategy). The resulting estimation strategy is generalized difference-in-differences.

I estimate a generalized Difference-in-Differences (DiD):[25]

$$y_{jt} = \alpha_j + \lambda_t + \sum_{\substack{t=1995 \\ t\neq 1998}}^{2019} \mathbb{1}\{year = t\} \left(\beta_t \cdot N\_TreatedFocal_j + \gamma_t \cdot N\_Focal_j + \zeta_t \cdot N_j\right) + \epsilon_{jt} \tag{1}$$

where y_{jt} are the outcomes of interest (number of hires, promotions of both focal workers and their coworkers, number of separations, etc.), $N\_TreatedFocal_j$- number of workers in 1998 that belong to 1952-1953 (treated focal) cohorts, N_Focal_j- number of workers in 1998 that belong to 1950-1953 (treated and control focal) birth cohorts, N_j- total number of workers in 1998. λ_t- year fixed effects controlling for time-varying shocks common to all establishments and α_j- firm fixed effects.

The coefficient of interest, β_t, shows the difference in the evolution of the outcome variable across firms with a similar workforce composition (including, the number of focal workers, number of old workers (men who will reach ERA and NRA around the main reform period), and the total number of workers) but different exposure to the reform (number of workers who experienced the rise in ERA among the focal workers) before and after the intervention. In other words, it estimates the effect of having an additional worker who experienced a rise in the ERA (born to the right of the cutoff, that is, 1952–1953) employed in the pre-reform year, 1998, on the outcomes of interest. The reference period is 1998 (treatment construction year).

I interact the treatment variables with time variables: either the flexible time dummies $\mathbb{1}\{year = t\}$ to observe how the treatment effects evolve over the years (where $\forall t = 1995, ..., 2019$). The interactions with time dummies help me to visually analyze the detailed effects of the reform over the years. To aid interpretation, I pool together the years when the workers in the 1950-1953 cohorts turn 60-65. In the simplified DiD model, I aggregate the time dummies to *Post*, which stands for the years from 2012, when the first treated cohort (born in 1952) turns 60, until 2017, when she turns 65.

[25] The difference from the identification strategy employed by Hut (2024) is that my 1999 reform affected primarily women, while the Dutch reform in Hut (2024) affected both genders. In addition, the German reform was uniform by industries and other characteristics, while in the Dutch reform the raise of retirement age and its amount depended on the industry and other characteristics.

Main reform and downstream effects. For the main reform period, I resample the firms with at least 1 focal worker in 2008. The choice of 2008 reflects a trade-off between measuring workforce composition sufficiently close to retirement ages and avoiding excessive noise in treatment assignment. Defining the sample too early would rely on workforce composition at ages far from retirement, which provides a weak signal of actual retention behavior at older ages. By 2008, all focal cohorts (1950–1953) remain below retirement eligibility, while being sufficiently close to retirement margins to yield a more informative measure of exposure to the reform (see subsubsection 1.3.1 for details). The challenge is that this design falls short of the ideal experiment for two reasons. First, firms with many affected workers may differ systematically from others in unobservable ways—such as size, technology, or HR policies—that are correlated with employment dynamics. Second, even conditional on observables, the composition of the workforce evolves endogenously over time. To restore the experimental ideal, I include firm fixed effects that absorb all time-invariant firm characteristics (ν_j), and year fixed effects that capture macroeconomic shocks common to all firms (μ_t). I additionally control for the evolution of establishment age categories, East German location, broad industry groups, and baseline hiring levels in 2008 (X_j). These controls account for systematic differences in firm characteristics correlated with workforce composition and labor demand, which is particularly important in the main reform sample, where firms are observed several years after the reform announcement. Including these controls improves comparability across establishments and supports the identifying assumptions.[26] The coefficient on the interaction between treatment intensity and post-reform enaction years is therefore identified from within-firm changes in employment outcomes relative to pre-reform trends, across firms with differing exposure intensities.

The corresponding identification resembles that of upstream effects:

$$y_{jt} = \nu_j + \mu_t + \sum_{\substack{t=2005 \\ t \neq 2008}}^{2020} \mathbb{1}\{year = t\} \left(\delta_t \cdot N\_TreatedFocal_j + \rho_t \cdot N\_Focal_j + \omega_t \cdot N_j + \xi_t \cdot X_j\right) + u_{jt} \tag{2}$$

This specification mimics the ideal randomized experiment: conditional on firm and year fixed effects and controls, the remaining variation stems from plausibly exogenous differences in exposure to the reform determined by historical workforce composition, not by contemporaneous decisions. In some parts of the paper, I run the same regressions at the firm-occupation level, adding the subscript c-occupations.

Identification assumption. The key identifying assumption is a parallel-trends condition: absent the reform, outcomes such as hiring and promotions would have evolved similarly across firms with different treatment intensities (i.e., different numbers of treated focal workers). While counterfactual post-treatment trends are unobservable, I assess this assumption by showing that event-study coefficients in pre-treatment years (1995–1997 for the upstream

[26] Specifications without controls are less comparable due to observable differences across establishments, and I therefore do not report them as baseline results.

period and 2005–2007 for the main reform period) are statistically indistinguishable from zero.

I further support the identification strategy with falsification tests by estimating the results on placebo cohorts and placebo gender in subsection 1.4. In addition, I exploit the structure of internal labor markets by estimating both intra- and inter-jobcell spillovers. If the estimated effects reflected unobserved confounders or aggregate shocks, they would plausibly appear across occupations within firms. Instead, I find crowd-out effects only within jobcells, while spillovers across jobcells are absent (subsection 1.6), lending credibility to the causal interpretation.

Finally, Panel B of Figure 1.C.7 shows that, within the estimation sample, the share of workers born after the cutoff is close to 50 percent and varies smoothly across firms with different treatment intensities, consistent with quasi-random exposure to the reform.

Standard errors. I cluster standard errors at the level of treatment variation. In the baseline specifications (Equation 1 and Equation 2), standard errors are clustered at the establishment level. In specifications at the establishment-by-occupation level (subsection 1.6), standard errors are clustered at the establishment-by-occupation level to account for within-cell correlation.

Effect heterogeneity. The average effects that the Equation 2 estimates could hide substantial heterogeneity across different firms, industry, and/or labor market characteristics. I estimate the same regressions on the corresponding subsamples to study the effect of heterogeneity.

1.4 Intra-Firm Personnel Decisions

1.4.1 Upstream Effects

As discussed in subsection 1.2, there are three main effects of interest- the upstream effects (before the pre-reform ERA at the age of 60), the main reform effects (between the pre-reform ERA and NRA, that is, 60-65 years old), and the downstream effects after reaching the NRA, corresponding to the three periods in Figure 1.A.1. First, I present the upstream effects on focal worker retention, followed by spillovers on promotions, external hiring, and the wage bill.

Although Carta and De Philippis (2024) find significant upstream effects of employment, Rabaté, Jongen, and Atav (2024) and Mastrobuoni (2009) find no sizeable effects in the Netherlands and the US. Neither of these studies analyzed spillover effects. I find upstream effects on employment and retention of focal workers starting from around 55 years old, and almost no significant upstream spillover effects.

Focal worker retentions. Panel A in Figure 1.C.3 shows that establishments with more treated focal workers (1952–1953 birth cohorts) in 1998 retained more focal workers (1950–1953 birth cohorts) from around 2007. This pattern indicates that treated workers increasingly stayed with their employer beyond the age of 60, the former early retirement threshold. An additional treated focal worker employed in 1998 leads to around 0.140 more focal worker retentions. Although some of the retained workers were on part-time contracts in 2012-2017 (31%), most were on full-time contracts (69%). While Figure 1.C.3 suggests that retention responses begin prior to 2008, this earlier variation is relatively small and noisy. Measuring treatment closer to retirement age for the main reform period improves the signal-to-noise ratio in retention and resulting spillovers.[27]

Null upstream spillover effects on coworker promotions and hiring. A natural follow-up question is whether the increase in the number of focal workers spills over to coworker promotions and external hiring. Having an additional treated focal worker in 1998 does not lead to any changes in coworker promotions (Panel B of Figure 1.C.3) or in the number of external hires (Panel C in Figure 1.C.3). There is a small decrease in the number of hired workers up to three years after the reform, but overall, there are no significant effects afterward, even in the years after the focal workers reach 60. For the analysis of spillover effects, I focus on the 2008 sample, which provides a more representative cross-section of firms at the time when retention becomes binding. While the first-stage retention effects are similar across the 1998 and 2008 samples, the 2008 sample better captures contemporaneous firm adjustment along promotion and hiring margins.[28]

Firm closure. Panel D of Figure 1.C.3 shows the effects of an additional treated worker employed in 1998 on the probability of firm closure. There are no effects up to 2010, followed by a small negative effect around 2010-2014, around the main reform period. These results show that, if anything, having an additional focal worker who faced an increase in retirement age (due to gender-neutral retirement ages) does not lead to firm closure, but, on the contrary, has a positive effect on firm closure around the reform years.

1.4.2 Main Reform Effects

In the previous section, I show that upstream firm responses occur only after focal workers reach older ages (around 55+), with almost no significant spillover effects on coworker promotions or external hiring before age 60. I now turn to the "main reform" period, when the cohorts directly exposed to the increase in the early retirement age (ERA) reach the ages of 60–65, and examine how firms adjust internal personnel policies and hiring in response to

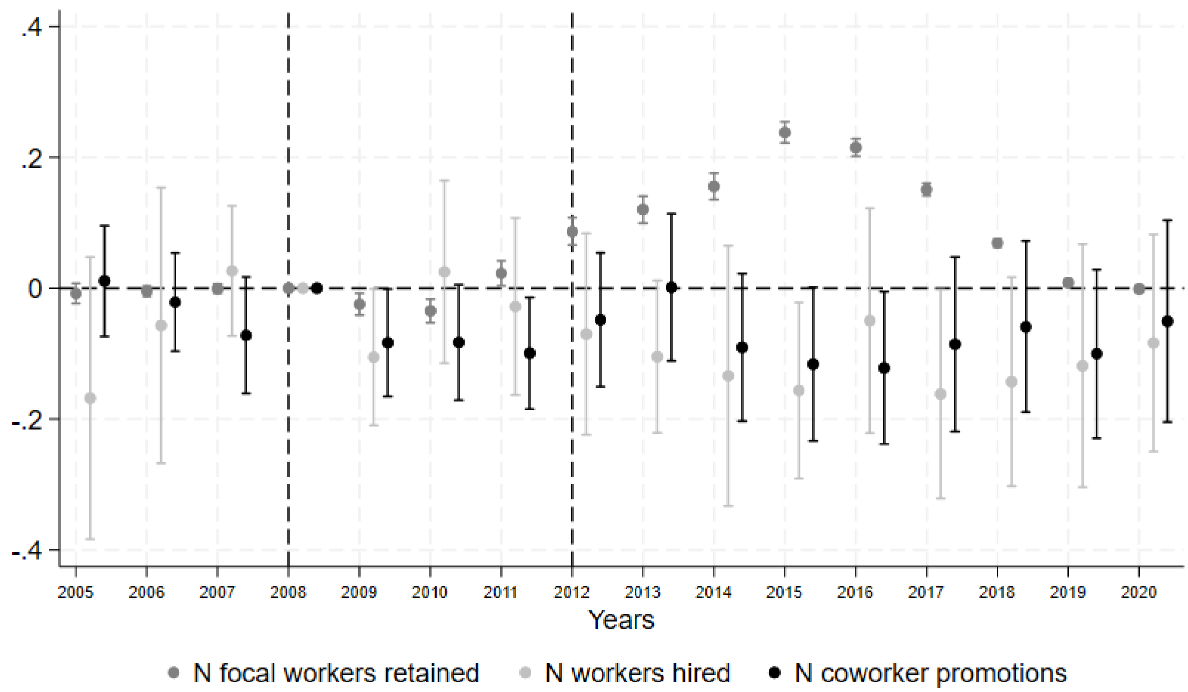
[27] Because the sample is constructed after the reform announcement, I additionally control for observable firm characteristics to improve comparability across establishments, see details in subsection 1.3.

[28] The 1998 sample consists of firms observed long before the reform becomes binding and therefore disproportionately includes surviving, more established firms by the time outcomes are measured in 2012–2017. Bianchi et al. (2023) argue that spillover effects are more likely to be concentrated in non-growing firms. This sample composition may attenuate observed spillover effects on promotions and hiring, even though first-stage retention effects are similar across samples. For this reason, I focus on the 2008 sample for second-stage estimates, as it includes a large chunk of the labor market that is missing in the 1998 sample.

delayed retirements.

Positive direct effects on retention and promotion of focal workers. Figure 1.2 shows that having an additional treated focal worker (a woman born in 1952–1953) employed in 2008 leads to a sizeable increase in the number of focal workers (1950–1953 birth cohorts) who remain employed at the same establishment once they reach age 60. The event-study coefficients lie between 0.08 and 0.25 additional retentions per year during 2012–2017. When I aggregate over these main reform years, one additional treated focal worker in 2008 increases focal retention by $\Delta R = 0.163$ workers.[29]

Figure 1.2: The effect of an additional treated focal worker employed in 2008 on focal worker retentions, external hiring, and coworker promotions



Notes: This figure represents the effect of having one additional treated worker (1952-1953 birth cohorts) in 2008 on the number of focal worker retentions, external hiring, and coworker promotions in each year. The points represent the estimated coefficients δ_t in Equation 2 and the vertical bars represent 95% confidence intervals. The dashed vertical line represents the year before policy enactment, when all focal workers (1950-1953 birth cohorts) were under the age of 60. Standard errors are clustered at the establishment level.

Negative spillover effects on coworker promotions and hiring. Next, I examine whether the retention of older workers comes at the expense of younger and middle-aged employees. Figure 1.2 summarizes the main margins: focal worker retentions, coworker promotions, and numbers of external hires. The aggregated treatment effects over 2012–2017

[29] All treatment effects in this section refer to the impact of one additional treated focal worker employed in 2008, with coefficients averaged over the years 2012–2017.

show clear evidence of negative spillovers: one focal worker employed in 2008 generates -0.075 fewer coworker promotions and -0.103 fewer external hires.

The timing of promotion effects reflects cohort-specific retirement eligibility rather than an immediate response to the reform. In particular, untreated focal workers (e.g., born in 1950–1951) reach the pre-reform early retirement age earlier and begin to exit before 2012, while treated cohorts (1952–1953) remain employed. This already generates differences in workforce composition and promotion opportunities prior to 2012. The effects become fully pronounced from 2012 onward, when treated cohorts reach their pre-reform designated retirement ages but are no longer able to exit. For this reason, I focus on the period 2012–2017 when the reform is fully binding. These dynamics are therefore driven by differential retirement eligibility across cohorts rather than by the choice of baseline year.

Substitution rates: how many hires and promotions are forgone per retained older worker? To quantify how firms substitute between older workers and other labor inputs, I relate the reform-induced increase in focal workers' retentions to the corresponding decline in hires and promotions. For any outcome Y, I compute the ratio $-\frac{\Delta Y}{\Delta R}$, which measures how many forgone units of Y correspond to one additional older worker retained due to the reform.[30] For total external hiring, the substitution rate is 0.63. In other words, retaining one additional older worker reduces total hiring by about 0.63 workers. For coworker promotions, the corresponding rate is 0.46, so roughly 0.46 coworker promotions are lost per extra retention of an older worker.

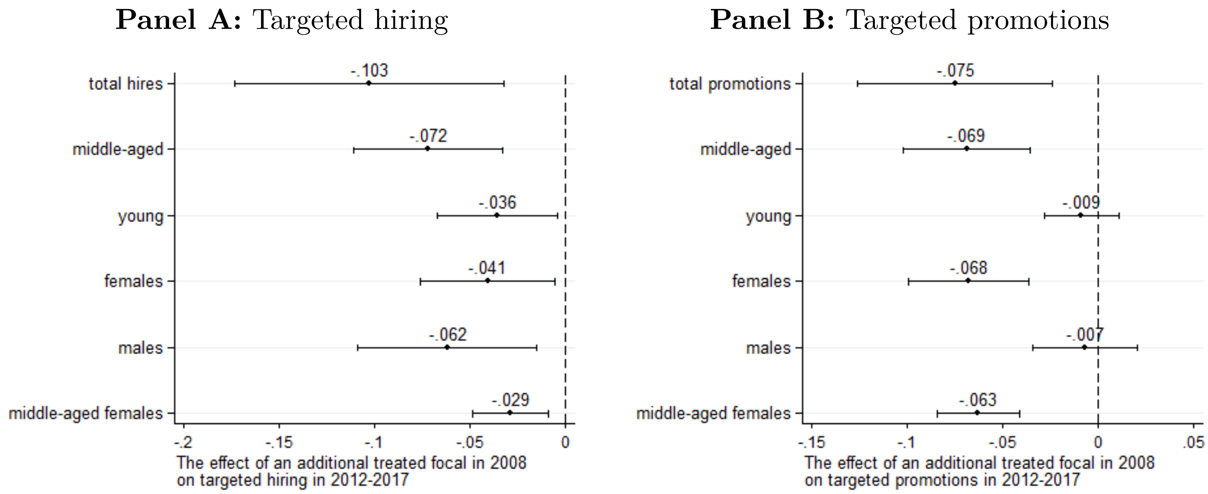
Allocation of promotion "slots". An alternative normalization uses the increase in focal promotions, 0.017, as the denominator and captures how promotion "slots" are reallocated within a firm. The implied ratio of 4.4 suggests that, for each additional promotion of an older focal worker, about 4.4 coworker promotions are forgone. This large magnitude reflects that focal promotions are a relatively rare margin—most retained older workers remain in their positions—so the denominator is small. It also reflects that promotions and operate along hierarchical ladders, where a single blocked senior position can prevent multiple downstream promotions. Together, these patterns highlight that the reform does not simply keep older workers in place but reshapes internal promotion ladders.

Who is crowded out? Heterogeneity by age and gender. To understand which groups of workers are most affected, I estimate treatment effects on hiring by age and gender. The aggregated coefficients for 2012–2017 are displayed in Figure 1.3.

To quantify how firms substitute between older workers and other labor inputs, I relate the reform-induced increase in the retention of focal workers to the corresponding decline in hiring across different groups. For each group g, I compute the ratio $-\Delta H^g/\Delta R$, which measures how many forgone hires of type g correspond to one additional older worker retained due to

[30] Formally, this ratio is a reduced-form *substitution rate in headcounts* rather than a structural Hicksian elasticity of substitution, because both ΔY and ΔR are estimated in levels (numbers of workers). Under the approximation that the changes are small relative to baseline employment, this ratio can be interpreted as an elasticity-like object, but I avoid a structural interpretation and use the ratio as a transparent measure of crowd-out per additional retained older worker.

Figure 1.3: The effect of an additional treated focal worker employed in 2008 on targeted external hiring and coworker promotions



Notes: Coefficient plots. Each row corresponds to the effect of having one additional treated worker (1952-1953 birth cohorts) in 2008 on the number of external hiring (**Panel A**) and coworker promotions (**Panel B**) in 2012-2017, between the new ERA and NRA, decomposed by age- and gender-based demographic groups (displayed in rows). The points represent the mean estimated coefficients δ_t in Equation 2 over 2012-2017, and the bars represent 95% confidence intervals. Standard errors are clustered at the establishment level.

the reform. The resulting substitution responses are economically meaningful. Firms reduce overall hiring by 0.63 workers for each additional focal worker retained. The largest component is the decline in hiring of middle-aged workers (0.44), followed by a smaller substitution away from young workers (0.22). Such a result could be driven by better substitutability of middle-aged workers with older focal workers, suggesting that the relevant margin of adjustment is movement within the core workforce rather than at the point of entry into the firm. Gender-specific estimates indicate that substitution occurs against both women (0.25) and men (0.38), with slightly stronger crowd-out among men. See Figure 1.C.6 for the dynamic effects of hiring by gender and age groups.

Hence, the crowd-out is strongest for middle-aged hires overall, and somewhat more pronounced for men than for women. These patterns suggest that the relevant margin of adjustment is among workers who are the closest substitutes for the retained older women—middle-aged workers—rather than those at the bottom of the age distribution.

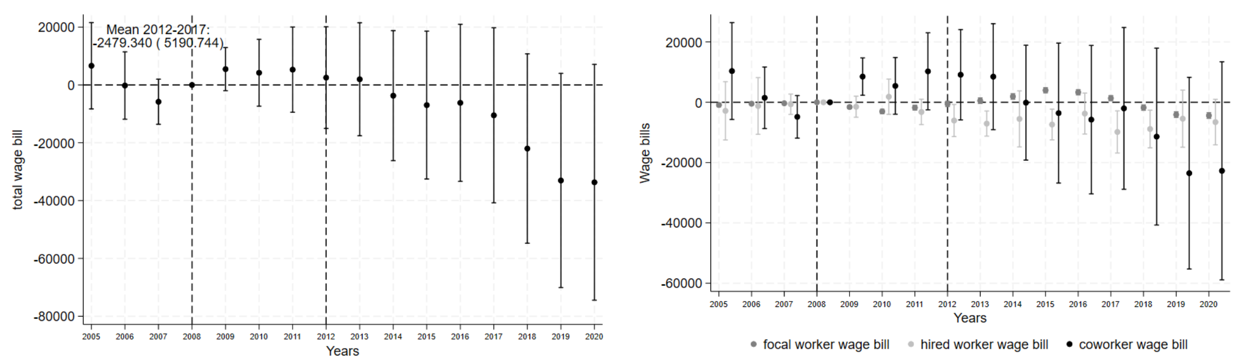
Panel B in Figure 1.3 reveals that crowd-out in promotions is concentrated among middle-aged and female coworkers, with no responses for young and male coworkers. Combined with the hiring results, these findings point to compressed career ladders in the middle of the age distribution rather than at entry.

Null effects on wage bills and probability of establishment closure. As a next step, I assess whether the aging workforce enlarged by the reform raises overall labor costs. Panel

A of Figure 1.4 plots the effect of one additional treated focal worker on total wage bills at the establishment level. The aggregated effect over 2012–2017 is slightly negative: -2,479, indicating that firms do not experience higher wage bills despite retaining more older workers. To investigate how firms maintain low wage bills, Panel B of Figure 1.4 decomposes wage bills by worker type. Wage bills for focal workers increase by about 1,841, while wage bills of external hires decline sufficiently to more than offset this increase. Overall, the rise in focal wage bills is absorbed by lower wage bills of coworkers and external hires. This pattern is consistent with firms re-optimizing the composition of their workforce rather than expanding total employment. Overall, firms adjust primarily along the composition margin: they retain and pay more older focal workers, but economize on other labor inputs so that total wage costs do not rise.

Figure 1.4: The effect of an additional treated focal worker employed in 2008 on wage bills: total and by worker type

Panel A: Total wage bills **Panel B:** Wage bills by worker types



Notes: This figure represents the effect of having one additional treated worker (1952-1953 birth cohorts) in 2008 on the total establishment wage bills in each year. **Panel A** displays the total wage bills, **Panel B** decomposes the total wage bill into the focal worker, coworker, and hired worker wage bills. The points represent the estimated coefficients δ_t in Equation 2 and the vertical bars represent 95% confidence intervals. The dashed vertical line represents the year before policy enactment, when all focal workers (1950-1953 birth cohorts) were under the age of 60. Standard errors are clustered at the establishment level.

Turning to firm closure, I estimate the impact of the reform on the probability of an establishment closure on Figure 1.C.4. The corresponding event-study coefficients are statistically insignificant. This suggests that, if anything, having more older workers whose retirements are delayed does not increase the likelihood that firms exit the market.

1.4.3 Downstream Effects

Finally, I examine whether treated and less-treated firms converge when the affected cohorts pass the new ERA. The event-study coefficients become small and statistically insignificant after 2017 for most outcomes in Figure 1.2, suggesting that firms gradually return to similar trajectories when the main reform window closes. However, downstream effects should be interpreted with caution, because the abolishment of the women's early retirement pathway also raised the normal retirement age from 65 to 65.5, so part of the adjustment beyond 2017 reflects changes at the normal retirement margin.

In sum, the German reform that raised the early retirement age for women leads to (i) higher retention and slightly higher promotion rates for older focal workers, (ii) unchanged or slightly lower total wage bills due to offsetting declines in coworkers' and external hires' wage bills, and (iii) sizeable crowd-out of coworker promotions and external hiring, particularly among middle-aged workers. One additional treated focal worker crowds out roughly 0.63 external hires and 0.46 coworker promotions per establishment during the main reform period. These magnitudes are close to, although somewhat smaller than, those found in the Netherlands by Hut (2024), consistent with the idea that the more strongly pre-announced German reform allowed firms more time to plan their adjustment.

A natural follow-up question is whether the crowding-out promotion effects on middle-aged women cascade to younger workers. Though this is difficult to test empirically, the downward trend in the coworker wage bills in Panel B in Figure 1.4 and the limited separation pattern from the firm in Panel B of Figure 1.C.5 imply pronounced and persistent negative effects that increase with retirement delay.

1.4.4 Robustness and Falsification Checks

Below, I conduct a set of robustness and sensitivity analyses that assess the stability of the baseline results to alternative modeling choices. Details on the corresponding sample construction and resulting sample sizes are provided in section 1.7.

Falsification tests: placebo birth cutoff and gender. A concern is that contemporaneous macroeconomic shocks, such as the 2008–2009 Great Recession, could differentially affect firms with more treated workers. This is unlikely, as such shocks should not correlate with the share of women born around the 1952 cutoff, and industry $\times$ year fixed effects absorb sector-specific business-cycle dynamics.

Nevertheless, to assess whether the estimated spillovers reflect the reform rather than coincident shocks, I conduct placebo tests that redefine the treatment along dimensions unaffected by the reform. Panel A of Figure 1.C.8 uses placebo birth cohorts (1952–1955), assigning a false cutoff at 1954; Panel B redefines focal workers as men born between 1950 and 1953. In both cases, I find no evidence of intra-firm spillovers. These null results support the interpretation that the baseline effects are driven by the retirement reform rather than by macroeconomic conditions or unrelated trends.

Robustness check: altering the estimation bandwidth. The baseline specification uses a two-year bandwidth around the reform cutoff. In Figure 1.C.9, I re-estimate the main results using a narrower one-year bandwidth. Focusing on cohorts born in 1951–1952 increases the share of firms with a single focal worker, effectively reducing the design to a difference-in-discontinuities framework with a binary treatment. The estimates remain qualitatively similar, though less precise, indicating that the results are robust to the choice of bandwidth and treatment definition.

Pre-reform behavior of focal workers. One of the identification concerns is that firms employing focal workers in 2008 may have non-random treatment intensity, potentially biasing the baseline estimates in the main reform period. To assess this, I analyze focal workers directly—independent of firm sampling—and test whether treated workers (born after January 1952) exhibit differential employment, hiring, or separation behavior prior to retirement ages.

Regressions of employment outcomes on treatment status show no differences until the mid-50s (Panel A of Figure 1.C.2). The number of days worked (conditional on employment) is unaffected (Panel B of Figure 1.C.2), indicating that responses operate on the extensive margin. Panels C and D of Figure 1.C.2 show no differential hiring or separation by treatment status before the pre-reform ERA. Overall, there is no evidence of anticipatory worker or firm adjustments before age 55, consistent with prior findings (Badalyan, 2025; Geyer and Welteke, 2021). This supports the identifying assumption that sampling firms in 2008—before any treated worker reaches pensionable age—is plausible.

1.5 Elasticity of Substitution

The baseline results show that extending the employment of older women induces sizable within-firm reallocations: establishments retain more older focal workers and reduce both promotions and external hiring of younger employees. These average effects, however, mask substantial heterogeneity arising from differences in firms' internal structures and their access to an external labor supply. Understanding these heterogeneities is essential for interpreting whether the observed crowd-out reflects technological substitutability, turnover frictions, or constraints imposed by internal career ladders.

To organize this analysis, I distinguish between two complementary dimensions of adjustment capacity: the thickness of the internal labor market (ILMT) and the thickness of the external labor market (ELMT). ILMT captures the extent to which coworkers within a firm represent viable substitutes for older incumbents on internal ladders, whereas ELMT captures the availability of suitable replacement candidates in the broader labor market. Variation along these dimensions provides a lens through which crowding-out of promotions and hiring can be mapped into economic parameters of interest—most notably, the elasticity of substitution between older and younger workers.

1.5.1 Definitions: Internal and External Labor Market Structures

Internal labor market thickness (ILMT). Internal labor market thickness is defined as the establishment-level concentration of employment in its largest occupation, following Ginja, Karimi, and Xiao (2023) and Cortes and Salvatori (2019). For establishment j in year t,

$$s_{jt} = \frac{N_{jt}^{\text{largest occupation}}}{N_{jt}}, \tag{3}$$

where N_{jt} denotes the establishment's total employment. An establishment is classified as having a "thick" internal labor market when this share exceeds the sample median (slightly above 0.5), meaning that a large share of workers occupy the same occupational ladder. Such settings are expected to exhibit strong internal competition and greater sensitivity of promotions to delayed retirements.

External labor market thickness (ELMT). To capture hiring frictions arising from local labor supply, I compute an index of external labor market thickness for each commuting zone. Using the full population of social security records, I define 141 commuting zones based on mobility patterns, following Kosfeld and Werner (2012).[31] For industry k in zone c at time t, external market thickness is measured as:

$$\theta_{kct} = \frac{N_{kct}/N_{ct}}{N_{kt}/N_t}, \tag{4}$$

where N_{kct} is local employment in industry k, N_{ct} is total local employment, and N_{kt} and N_t are the corresponding national values. Values above one indicate that an industry is locally overrepresented relative to the national distribution, implying a thicker external market and lower turnover frictions. Thin external markets, by contrast, constrain a firm's ability to replace or expand its workforce and are therefore expected to amplify hiring crowd-out. Figure Figure 1.C.10 illustrates these patterns for motor vehicles and hospital activities.[32]

1.5.2 A Slot-Constraint Model with Human Capital and Turnover Frictions

Setup. A firm fills a fixed number of job slots in each period. For any given slot, the firm may (i) retain an incumbent older worker, (ii) promote an internal junior coworker, or (iii) hire externally. The indicator variables for these choices are denoted by $r, p, h \in [0, 1]$ with

[31] I choose the finer 141-zone classification, rather than broader regional aggregates as in Jäger and Heining (2022), because women are typically less mobile across regions (Meekes and Hassink, 2022), making finer spatial units more relevant in this context.

[32] I thank Niklas Vetterer for help getting started creating these maps.

$$r + p + h = 1 \tag{5}$$

The firm maximizes current net output from filling the slot:

$$\Pi(r, p, h) = rf_d + pf_p(\sigma_p) + hf_h(\sigma_h) - C(p, h; \tau) \tag{6}$$

where f_d is the output from retaining an older incumbent, $f_p(\sigma_p)$ the output from promoting a younger coworker (with σ_p measuring how close a promoted worker is to the productivity of an older incumbent), and $f_h(\sigma_h)$ the output from hiring externally (with σ_h measuring substitutability of a new hire).

Turnover frictions enter through the adjustment cost function $C(p, h; \tau)$, with

$$\frac{\partial C}{\partial p} > 0, \qquad \frac{\partial C}{\partial h} > 0, \qquad \frac{\partial C}{\partial \tau} < 0 \tag{7}$$

so that thicker external labor markets (larger τ, corresponding to thicker ELMT) reduce the cost of replacing a worker.

Solution. Because (5) implies $h = 1 - r - p$, the problem can be rewritten in (r, p) only. The first-order conditions for an interior solution are

$$f_d - f_h(\sigma_h) - C_r + C_h = 0 \tag{8}$$

$$f_p(\sigma_p) - f_h(\sigma_h) - C_p + C_h = 0 \tag{9}$$

These conditions determine the firm's allocation (r, p, h) as functions of the primitives (f_d, f_p, f_h) and of turnover frictions τ.

Comparative statics: effects of a higher retention incentive. The reform increases the value of retaining older workers, f_d. dr denotes the induced marginal change in retention. Differentiating (8)–(9) gives the responses of promotion and hiring:

$$\frac{dp}{dr} = \psi(\sigma_p, \sigma_h, \tau), \qquad \frac{dh}{dr} = -\left(1 + \frac{dp}{dr}\right) \tag{10}$$

where $\psi(\cdot)$ is decreasing in external labor market thickness τ (i.e., increasing in external frictions).:

$$\frac{\partial}{\partial \tau}\left(\frac{dp}{dr}\right) < 0, \qquad \frac{\partial}{\partial \tau}\left(\frac{dh}{dr}\right) > 0 \tag{11}$$

Thus, in thin external labor markets (low τ), external adjustment is costly (i.e., hiring is a less attractive margin), and the firm relies more on internal adjustment, while reducing hiring more strongly because external hiring is a costly margin.

Observed crowd-out and its decomposition. Empirically, the observed effect of one additional retained older worker on younger-worker outcomes is

$$\Delta Y_{\text{obs}} = \sigma_p \frac{dp}{dr} + \sigma_h \frac{dh}{dr} \tag{12}$$

Substituting (10) yields

$$\Delta Y_{\text{obs}} = -\sigma_h + (\sigma_p - \sigma_h)\frac{dp}{dr} \tag{13}$$

Equation (13) highlights two determinants of observed spillovers:

(i) *Relative substitutability*: if promoted workers are closer substitutes than external hires ($\sigma_p > \sigma_h$), the promotion channel amplifies crowd-out; the opposite holds when $\sigma_p < \sigma_h$.

(ii) *Turnover frictions*: in thin markets (low τ), hiring is a costly adjustment margin, so firms reduce hiring more strongly when retention increases (i.e., dh/dr is more negative), while in thick markets hiring adjusts more smoothly.

Implications for interpreting elasticities of substitution. A naïve elasticity-of-substitution estimate maps ΔY_{obs} directly into a structural substitution parameter. Equation (13) shows that such an interpretation is biased whenever firms face frictions (τ finite) or when promotion and hiring substitute imperfectly for each other.

Underestimation. If external hires are the closer substitutes (σ_h large), but hiring is difficult (low τ), firms reduce hiring strongly (i.e., dh/dr is more negative). However, because adjustment is distorted by frictions, the observed response reflects both substitutability and constraints, which may lead to an underestimation of the underlying technological substitutability.

Overestimation. If promoted juniors are the closer substitutes (large σ_p) and the firm strongly reallocates internally when external hiring is constrained, the promotion response may exaggerate crowd-out relative to the frictionless benchmark, overstating substitutability.

In summary, observed spillovers equal a combination of technological substitutability (σ_p, σ_h) and turnover frictions (τ). The empirical heterogeneity by ILMT and ELMT maps directly onto these mechanisms: thick ILMT affects σ_p (internal substitution), while thin ELMT raises adjustment costs $C(\cdot;\tau)$ and shifts the firm toward internal responses.

Mapping to empirical estimates. The empirical specification estimates the effect of an additional retained older worker on promotions and hiring at the establishment level. In the model, these correspond to $\frac{dp}{dr}$ and $\frac{dh}{dr}$, respectively. The heterogeneity analysis by internal and external labor market thickness maps to variations in (σ_p, σ_h) and τ: thicker internal labor markets increase the scope for internal substitution (σ_p), while thinner external labor markets increase turnover frictions (low τ) and shift adjustment toward promotions. The empirical estimates can therefore be interpreted as reduced-form counterparts of these structural relationships, capturing how adjustment is distributed across margins.

Link to empirical specification. The model is written at the level of a single job slot, while the empirical analysis is conducted at the establishment level. To connect the two, I interpret establishments as consisting of multiple slots subject to similar decision problems. Under the assumption that slot-level decisions are additively separable and that the reform affects the marginal value of retention f_d uniformly across slots, aggregate firm-level responses can be approximated as the sum of slot-level adjustments. In this case, the empirical coefficients can be interpreted as averages of $\frac{dp}{dr}$ and $\frac{dh}{dr}$ across slots within the establishment. This interpretation highlights that the model is intended as a conceptual framework to guide the empirical analysis rather than as a fully structural estimation.

1.5.3 Results

Below, I show the heterogeneous effects of the reform by internal and external labor market structures. The heterogeneity results should be interpreted with caution. While several patterns are consistent with the proposed mechanisms, differences across subgroups are not always statistically distinguishable, and the evidence should therefore be viewed as suggestive rather than definitive.

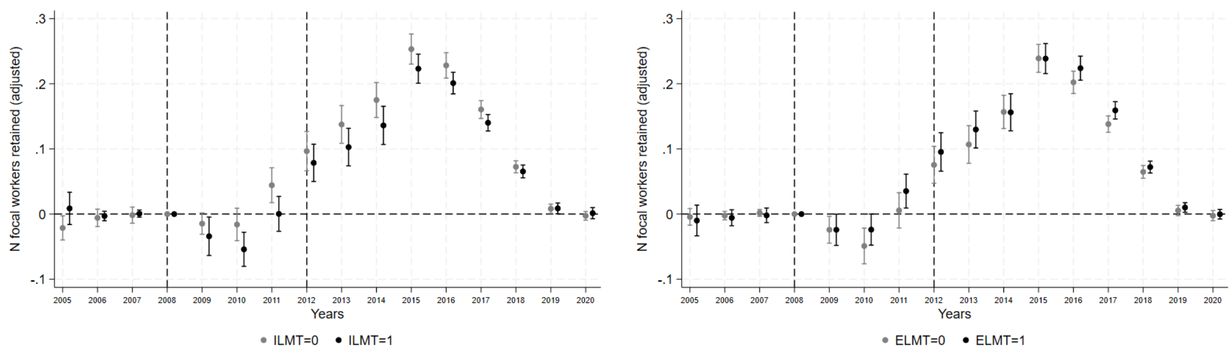
Focal worker retentions. Before analyzing the mechanisms behind crowd-out effects, it is important to understand how internal and external labor market thickness shapes the retention decisions of focal workers. Retention appears to vary systematically with internal labor structure. Panel A in Figure 1.5 shows that older workers tend to be retained more often when the ILMT is thin (0.175), where few internal substitutes exist, and the loss of experienced workers would be particularly costly. Estimated retention is somewhat lower when the ILMT is thick (0.147), meaning that internal candidates are more abundant. By contrast, the thickness of the external labor market does not appear to strongly affect retention: the effects are similar in magnitude in thin (0.153) and thick (0.167) ELMTs (Panel B).

These results are closely aligned with Badalyan (2025), which, using individual-level regressions, finds that internal substitutability appears to strongly shape labor supply responses to this reform, while industry-based ELMT generates relatively little heterogeneity. That study also documents heterogeneity by occupation-based ELMT, showing that older workers in occupations with scarce external substitutes are more likely to remain employed.

Together with the firm-level evidence here, the pattern is consistent with the interpretation that retention decisions are driven primarily by firm-specific human capital and internal knowledge complementarities, rather than by external hiring frictions for given industries. Such specific skills and low substitutability may help explain the higher retention and lower hiring rates for older workers documented in the literature (Hutchens, 1986).

Figure 1.5: The effect of an additional treated focal worker employed in 2008 on retentions by internal and external labor market thicknesses

Panel A: Retentions by ILMT **Panel B:** Retentions by ELMT



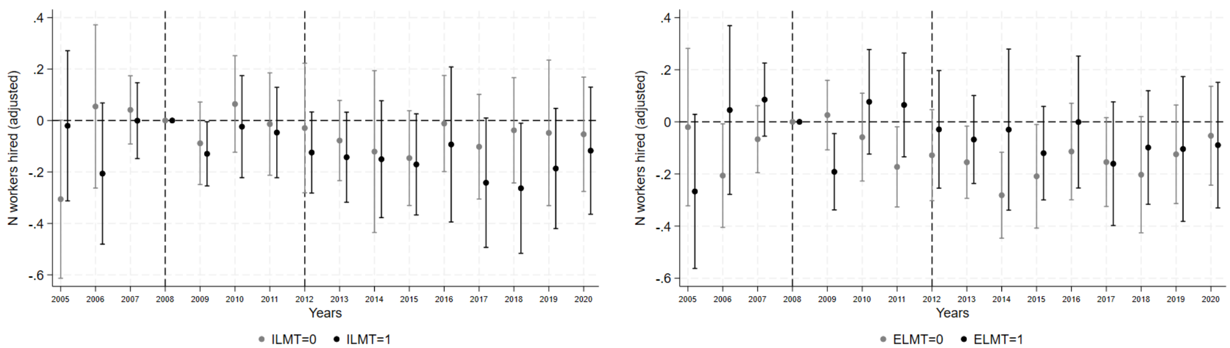
Notes: This figure represents the effect of having one additional treated worker (1952-1953 birth cohorts) in 2008 on the number of retentions of focal workers. **Panel A** represents subsample analysis by internal labor market thickness- the share of the largest employment occupation in the establishment in the total workforce. **Panel B** represents subsample analysis by external labor market thickness (ELMT). The ELMT is categorized into two groups based on the commuting zone being at least as concentrated as the country-level concentration ($ELMT > 1$). The points represent the estimated coefficients δ_t in Equation 2 and the bars represent 95% confidence intervals. Standard errors are clustered at the establishment level.

Hiring. The crowd-out of external hiring appears to vary with labor market thickness. In thin external markets, where firms face severe turnover frictions, hiring declines more strongly (–0.174). In thick external markets, the hiring response is smaller in magnitude (–0.068). Because retention rates are similar across ELMT groups, these differences suggest that failing to account for external frictions may lead to overstating the substitutability between older workers and external hires. What may be interpreted as "strong crowd-out" could, in thin markets, partly reflect limited hiring possibilities rather than purely technological substitution.

Internal labor market structure also appears to shape hiring responses, although in a different direction: hiring declines more strongly when ILMTs are thick (–0.923) and less so when ILMT is thin (–0.081). When many workers share the same occupation, firms may adjust internally by reallocating career progression rather than by expanding the workforce. This is consistent with larger hiring reductions relative to retention effects in thick ILMTs, suggesting an apparent substitution that may partly reflect internal hierarchy constraints rather than purely technological substitution.

Figure 1.6: The effect of an additional treated focal worker employed in 2008 on hiring by internal and external labor market thicknesses

Panel A: Hiring by ILMT **Panel B:** Hiring by ELMT



Notes: This figure represents the effect of having one additional treated worker (1952-1953 birth cohorts) in 2008 on the number of hired workers. The points represent the estimated coefficients δ_t in Equation 2 and the bars represent 95% confidence intervals. **Panel A** represents subsample analysis by internal labor market thickness- the share of the largest employment occupation in the establishment in the total workforce, and **Panel B** represents subsample analysis by external labor market thickness (ELMT). The ELMT is categorized into two groups based on the commuting zone being at least as concentrated as the country-level concentration ($ELMT > 1$). Standard errors are clustered at the establishment level.

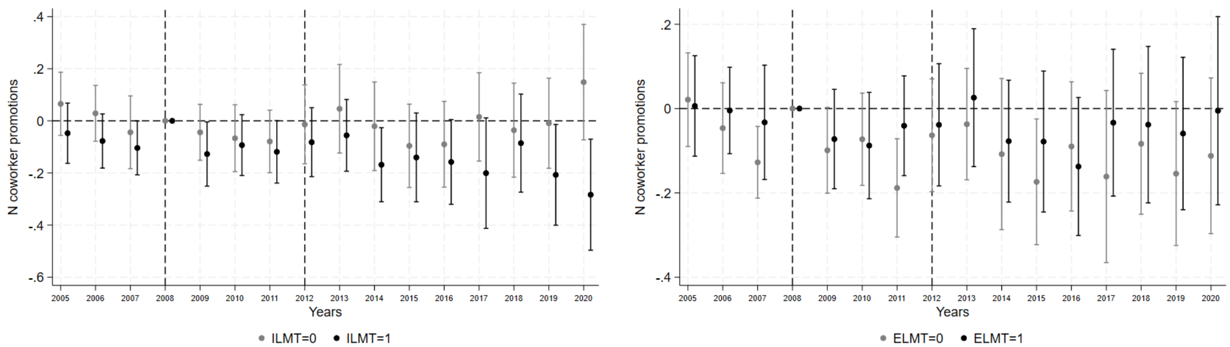
Promotions. Promotion responses provide additional evidence that internal bottlenecks may mediate the impact of delayed retirement. In high-ILMT establishments, coworker promotions decline more strongly (–0.134)—substantially more than in thin ILMTs (–0.026). Because focal-worker retention is similar across ILMT categories, this gap is consistent with tighter congestion along internal career ladders in thick markets. By contrast, promotion effects are more similar across ELMTs (–0.106 in thin vs. –0.057 in thick), suggesting that promotion crowd-out is more closely related to internal rank constraints than to external hiring frictions.

These patterns are consistent with Bertheau (2021), who shows that a large share of firms fill vacancies internally, reflecting imperfect substitutability between internal and external candidates. Such imperfect substitution implies that coworker promotions may be more sensitive to delayed retirements than external hiring. Taken together, these results suggest that internal coworkers may be closer substitutes for older workers and therefore may experience larger crowd-out effects than external hires, although these differences should be interpreted with caution.

Industry tradability. To provide additional suggestive evidence on the role of turnover frictions, I analyze whether firms in more tradable industries exhibit different retention patterns for focal workers. Because production can be relocated across borders, tradable industries may offer greater scope for worker substitution through outsourcing than non-tradable sectors

Figure 1.7: The effect of an additional treated focal worker employed in 2008 on coworker promotions by internal and external labor market thicknesses

Panel A: Promotions by ILMT **Panel B:** Promotions by ELMT



Notes: This figure represents the effect of having one additional treated worker (1952-1953 birth cohorts) in 2008 on the number of coworker promotions. The points represent the estimated coefficients δ_t in Equation 2 and the bars represent 95% confidence intervals. **Panel A** represents subsample analysis by internal labor market thickness- the share of the largest employment occupation in the establishment in the total workforce, and **Panel B** represents subsample analysis by external labor market thickness (ELMT). The ELMT is categorized into two groups based on the commuting zone being at least as concentrated as the country-level concentration ($ELMT > 1$). Standard errors are clustered at the establishment level.

(Drenik et al., 2023). I classify industries by tradability following Gregory, Salomons, and Zierahn (2022).[33]

Figure 1.C.11 shows that establishments in non-tradable industries tend to exhibit larger increases in the retention of older workers, which is consistent with their more limited exposure to external competitive pressures. By contrast, the crowd-out of younger workers' promotions and hiring appears more pronounced in tradable industries. These patterns are in line with earlier evidence suggesting that external market conditions shape the adjustment margin: tradable industries may rely more on external hiring and therefore exhibit stronger displacement when retention rises.

These differences across industries should be interpreted with caution, as confidence intervals overlap and the estimates are imprecise. Nevertheless, the observed patterns suggest that firm-level averages may mask meaningful within-firm differences in adjustment. This motivates the subsequent analysis at the jobcell level, which more directly captures internal bottlenecks and the role of occupation- and task-based constraints in shaping spillovers.

Wage bills. If retention of focal workers is concentrated among workers who are less

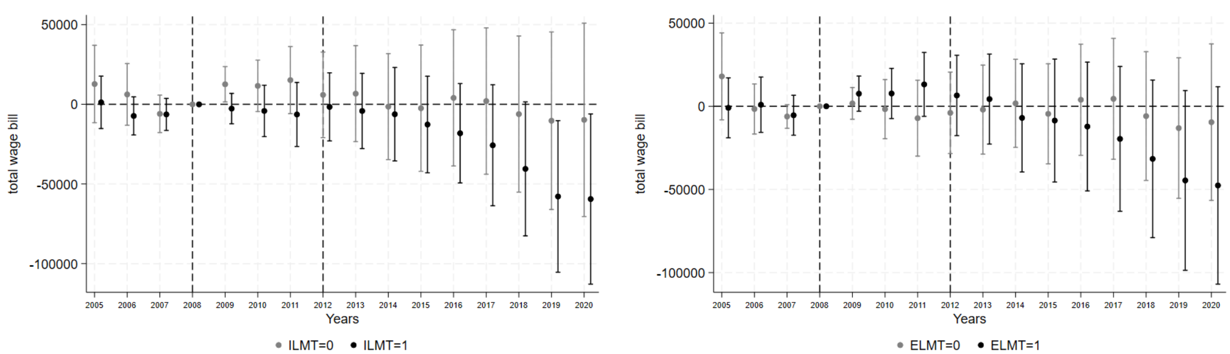
[33] *Tradable industries* are: Mining (WZ08: B); Manufacturing (WZ08: C); Electricity, water supply (WZ08: D, E); Transport, storage (WZ08: H); Financial services (WZ08: K); Real estate (WZ08: L); Agriculture (WZ08: A); Information and communication (WZ08: J); Scientific and technical services (WZ08: M). I thank Duncan Roth for the help with the data.

substitutable internally, this may affect firm-level wage bills, particularly in settings with thin internal labor markets. The average effects of an aging workforce may mask heterogeneity in how firms adjust total labor costs.

Panel A in Figure 1.8 shows that, in thick internal labor markets, the estimated wage bill response is negative (though not statistically significant): an additional treated focal worker is associated with a reduction in total wage bills of roughly EUR 11,415. This decline may reflect reductions in coworker wage bills documented in the previous section, consistent with compressed promotion opportunities and slower wage growth in more competitive settings.[34].

Figure 1.8: The effect of an additional treated focal worker employed in 2008 on wage bills by internal and external labor market thicknesses

Panel A: Wage bills by ILMT **Panel B:** Wage bills by ELMT



Notes: This figure represents the effect of having one additional treated worker (1952-1953 birth cohorts) in 2008 on wage bills. The points represent the estimated coefficients δ_t in Equation 2 (2012-2017 pooled together) and the bars represent 95% confidence intervals. **Panel A** represents subsample analysis by internal labor market thickness- the share of the largest employment occupation in the establishment in the total workforce. **Panel B** represents subsample analysis by external labor market thickness (ELMT). The ELMT is categorized into two groups based on the commuting zone being at least as concentrated as the country-level concentration ($ELMT > 1$). Standard errors are clustered at the establishment level.

Implications. Taken together, these results suggest that observed spillovers depend on market structure—both inside and outside the firm. This pattern can be related to the conceptual framework introduced earlier, where adjustment occurs through internal promotion, external hiring, or continued retention, and where these margins are shaped by internal substitutability and external turnover frictions. Two implications emerge.

First, not accounting for external labor market thickness may lead to overstating substitutability between older workers and young hires. In thin ELMTs, firms reduce hiring more

[34] By contrast, according to Panel B in Figure 1.8, in thin internal labor markets, the wage bill effect is small and positive (EUR 2,478), which may reflect the limited pool of internal substitutes and the higher relative value of retaining experienced workers. External labor market thickness appears to play a more limited role: wage bill responses are near zero in thin external markets (EUR –35), but reach a negative value (EUR –6,062) in thick external markets, where hiring adjustments may be easier to implement, although both estimates are statistically insignificant

strongly than in thick ELMTs despite similar increases in retention. This stronger hiring response is consistent with the interpretation that it reflects the limited availability of external candidates rather than a purely technological ability to substitute older workers for younger ones. Estimates that treat all hiring reductions as technological substitution may therefore overstate the implied elasticity of substitution. In my data, this corresponds to an increase in the implied substitutability of roughly a factor of 2.8 when ELMT is not accounted for. This pattern is consistent with the model case in which turnover frictions suppress the hiring margin, shifting adjustment toward other channels.

Second, not accounting for internal labor market thickness may understate the crowd-out of promotions. In thick ILMTs, delayed retirements are associated with substantially lower promotion rates (–0.134), whereas the same reform is associated with much smaller effects in thin ILMTs (–0.026). Averaging across firms may therefore mask environments in which internal competition is more intense and promotion ladders are more congested. Analyses that do not condition on ILMT may therefore understate the career costs borne by workers in occupations with dense hierarchies. In terms of the earlier conceptual framework, thick ILMTs can be interpreted as settings with high internal substitutability: when older workers stay longer, promotions appear to become the primary margin of adjustment.

Overall, the evidence suggests that spillovers from raising retirement ages reflect the interaction of: (i) internal substitutability and firm-specific human capital, (ii) turnover frictions in external labor markets, and (iii) hierarchical congestion within occupations. Adjustment across margins appears interrelated: when external hiring is constrained (thin ELMT), promotions tend to absorb more of the adjustment; when internal ladders are congested (thick ILMT), external hiring tends to decline, even when the labor supply is relatively abundant. These interactions suggest that reduced-form crowd-out estimates combine technological substitution with adjustment constraints, and should be interpreted in light of both internal and external market structure.

1.6 Intra- and Inter-Jobcell Personnel Decisions

The evidence that spillover effects vary systematically with the thickness of internal labor markets (ILMT) suggests that firms' adjustment mechanisms operate at a finer level than the firm-level used in the sections above. If internal promotions and hiring decisions depend on the pool of available coworkers within specific job ladders, then analyzing only firm-level outcomes may conceal important heterogeneity in how delayed retirements affect coworkers. This motivates a closer examination of intra- and inter-jobcell spillovers. By zooming into establishment–occupation cells (jobcell), I can disentangle whether crowd-out effects arise primarily within job ladders (intra-jobcell) or through broader reallocation across occupations or establishments (inter-jobcell).[35]

[35] The occupations are based on 3-digit classifications, as the ILMT was constructed above.

1.6.1 Baseline Inter- and Intra- Jobcell Effects

Previous literature finds conflicting results on intra-firm spillovers of an aging workforce, with most papers finding negative impacts (Bianchi et al., 2023; Ferrari, Kabátek, and Morris, 2023), while others, using slightly larger firms, find positive impacts (Carta, D'Amuri, and Von Wachter, 2024). The positive impacts on larger firms could be driven by a lack of availability of more granular data, such as occupations within establishments; therefore, such analyses could hide negative spillovers. On the other hand, larger firms may find it easier to spread work to incumbent workers and to grant promotions, due to their capacity to make more flexible internal organizational adjustments (Hensvik and Rosenqvist, 2019; Jäger and Heining, 2022), for example, due to human resource management systems (Holzer (1987) as cited in Schmutte and Skira (2023)). If internal adjustments are more muted in larger firms despite zooming into occupations, then this is evidence for their flexibility in making internal adjustments.

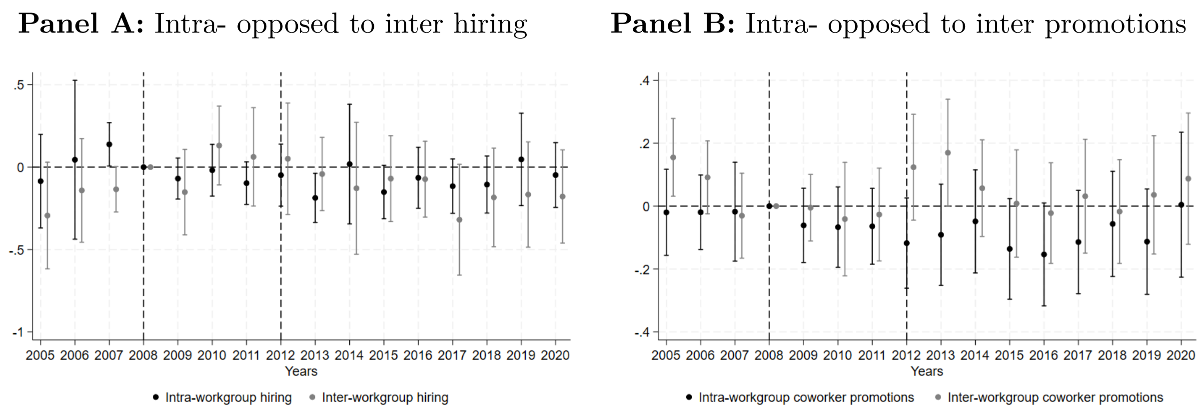
First, I show that the hiring responses are relatively larger in small establishments than in larger ones. Then, I zoom into the jobcells, proxied by occupations that interact within establishments. Zooming in on larger firms enables me to test whether there are negative effects hidden within workplaces that are not apparent in firm-level estimates. Such analyses help to highlight the differences in the sign of spillover effects on hired workers, even in studies that use the same reforms (Carta, D'Amuri, and Von Wachter, 2024; Bianchi et al., 2023).

Firm size heterogeneity. A firm's size is a first-order determinant of how it adjusts to delayed retirements. Larger establishments typically operate with deeper internal hierarchies, more diversified jobcells, and more flexible redeployment possibilities. These features allow them to absorb shocks through internal reallocation—via promotions or transfers across jobcells—rather than through hiring or separations. Smaller firms, by contrast, generally have thinner internal structures: occupational ladders are short, jobcells are narrow, and external hiring often plays a relatively larger role. As a result, the same increase in retention may generate different patterns of spillovers in small versus large firms. Examining heterogeneity by firm size, therefore, provides an important bridge between the baseline establishment-level results and the more granular intra-, as opposed to inter-jobcell analysis that follows. It helps to clarify whether the observed adjustment margins are driven by differences in organizational depth or by within-occupation constraints operating inside jobcells.

Figure 1.C.12 shows that establishment size also moderates firms adjustments to delayed retirements. Smaller firms exhibit larger increases in the retention of older incumbents, consistent with their more limited internal substitution possibilities. By contrast, crowd-out coefficients are noisier but tend to be larger in establishments with more than 30 workers; however, the estimates for small firms lie well within the confidence intervals of large firms, suggesting no statistically significant difference in spillovers across size groups. These patterns motivate the next step of the analysis, which decomposes spillovers into intra- and inter-jobcell responses to examine more finely how firms reallocate tasks and mobility opportunities when internal substitution options vary.

Intra- opposed to inter-jobcell spillovers. I next decompose the firm-level spillovers into within- and across-jobcell adjustments. Jobcells are defined as 3-digit occupations (*Klassifikation der Berufe (KldB) 1988*) within establishments. I keep establishments that had at least two jobcells in 2008 to enable intra- opposed to inter-jobcell analyses.

Figure 1.9: Intra- and inter-jobcell effects



Notes: This figure represents the effect of having one additional treated worker (1952-1953 birth cohorts) in 2008 on hiring (**Panel A**) and promotions (**Panel B**) in each year, decomposed by intra-jobcell (black) and inter-jobcell (gray) spillovers. The points represent the estimated coefficients δ_t in Equation 2 and the vertical bars represent 95% confidence intervals. The dashed vertical line represents the year before policy enactment, when all focal workers (1950-1953 birth cohorts) were under the age of 60. Standard errors are clustered at the establishment level.

The results reveal that crowd-out operates primarily within the focal worker's own jobcell. Panel A in Figure 1.9 shows that an additional treated older worker retained in 2008 reduces *intra-jobcell* hiring by approximately 0.086 workers, nearly twice the magnitude of the reduction in *inter-jobcell* hiring (0.052). Internal promotions show an even sharper segmentation (Panel B in Figure 1.9): promotions decline meaningfully within the focal jobcell (by 0.109 per retained older worker), whereas promotions in other jobcells are essentially unaffected, with estimates close to zero.

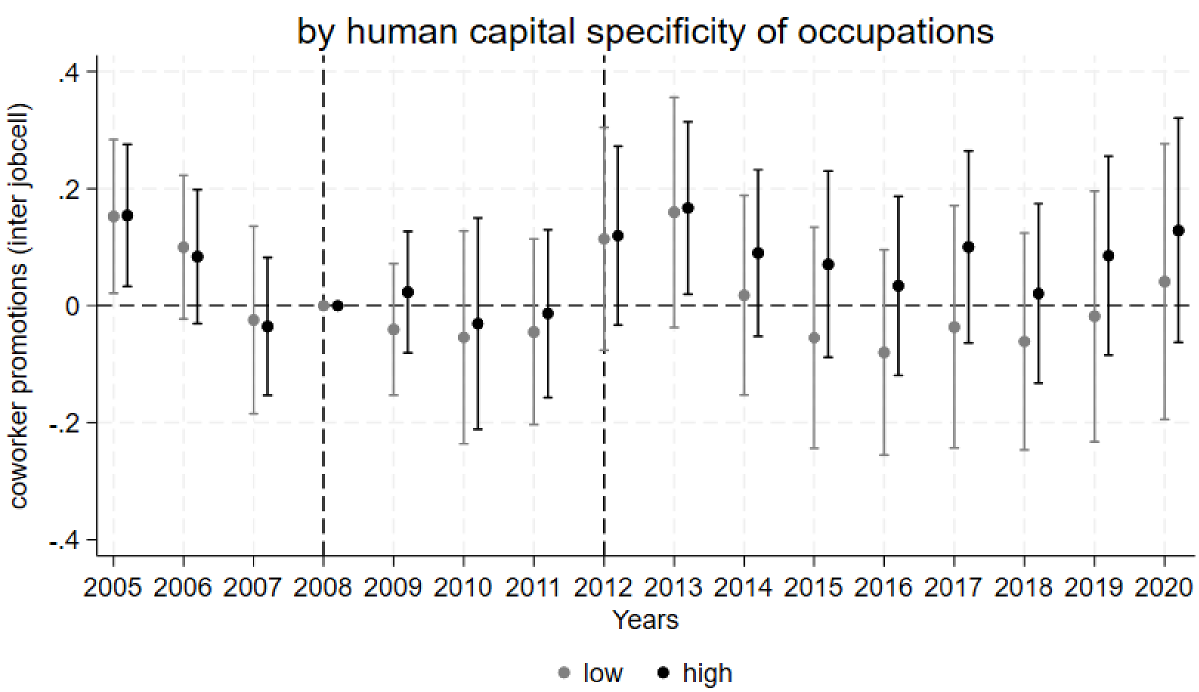
The lack of spillover impacts on hiring across occupations is in line with the previous literature on worker substitutability (Brenøe et al., 2024; Huebener et al., 2024; Jäger and Heining, 2022; Schmutte and Skira, 2023). Moreover, the absence of spurious effects across occupations confirms that the intra-firm adjustments are the result of the aging workforce as opposed to establishment-specific impacts of the crisis or other reforms.

These patterns indicate that adjustment frictions are highly localized. Firms substitute most strongly among workers who share tasks, supervisors, or career ladders, foreshadowing the heterogeneities to come: the importance of internal labor market structure, firm-specific human capital, and bottleneck occupations is concentrated precisely in those segments where intra-jobcell substitution is feasible.

1.6.2 Mechanisms: Human Capital and Value of Old Workers

Existing evidence suggests that delayed retirements disproportionately retain older workers who are costly to replace—such as managers and workers in occupations with strong job-specific skill requirements—highlighting the importance of firm-specific human capital and turnover frictions (Badalyan, 2025). If firms optimally retain these workers because they embody valuable knowledge, relationships, or organizational capital, their continued presence may have implications beyond mechanical crowd-out effects. In particular, while slot constraints may intensify competition for promotions among workers on the same career ladder, retained older workers may simultaneously generate positive spillovers for coworkers in complementary roles by preserving firm-specific human capital that is difficult to replicate. This insight motivates an explicit analysis of inter-jobcell spillovers: distinguishing between competitive effects within narrowly defined career ladders and complementarities across occupations that rely on shared expertise, coordination, or managerial oversight.

Figure 1.10: Inter-jobcell effects on coworker promotions by human capital specificity



Notes: This figure represents the effect of having one additional treated worker (1952-1953 birth cohorts) in 2008 on inter-jobcell promotions in each year by low (in gray) and high (in black) human capital specificity of occupations. The points represent the estimated coefficients δ_t in Equation 2 and the vertical bars represent 95% confidence intervals. The dashed vertical line represents the year before policy enactment, when all focal workers (1950-1953 birth cohorts) were under the age of 60. Standard errors are clustered at the establishment level.

Human capital specificity of occupations. Occupations differ by how much of a worker's productivity stems from job-specific skills. When human capital is highly occupation-specific, workers are harder to replace externally and may generate complementarities for coworkers whose tasks rely on their accumulated know-how. To measure this specificity, I follow the approach in Jäger and Heining (2022) and Bleakley and Lin (2012) and estimate occupation-level Mincer regressions for each 3-digit occupation.[36] I use the occupation-specific return to experience as a measure of how strongly wages depend on on-the-job learning. Occupations with returns above the median are classified as having high human-capital specificity.

Figure 1.10 displays the results. When older workers are employed in occupations with high firm-specific human capital, there are positive impacts on coworkers' promotions in intra-cell occupations. Overall, these results highlight that older workers can be particularly valuable to firms and can generate higher productivity for some coworkers.[37]

1.7 Conclusion

This paper exploits a large retirement reform—abolition of the female pathway to early retirement in Germany—to examine how firms adjust when a large group of older workers remains employed longer than prior cohorts. Using rich administrative data with detailed occupational information, I document how establishments respond to the extended retention of older female workers and how these adjustments cascade through internal promotion structures and external hiring pipelines.

The reform substantially raised older-worker retention when the affected cohorts reached pension-eligible ages, but firms made little systematic adjustment beforehand, suggesting that the long pre-announcement horizon muted any liquidity-driven responses. When older workers ultimately remained in their jobs longer, firms shifted their personnel decisions: internal promotions slowed, and external hiring declined, consistent with an internal substitution mechanism in which the delayed exit of older workers constrains advancement opportunities and reduces openings for new recruits. These effects were unevenly distributed. Promotion losses were concentrated in thick internal labor markets, where many workers compete for the same rungs on tightly structured ladders, whereas hiring declines were greater in thin external markets, where firms face limited replacement options.

A broader implication emerging from these findings is that the conventional view of older workers simply crowding out younger ones provides only a partial picture. Prior studies on worker exits highlight the costs of turnover—losses of tacit firm-specific human capital, the difficulty of replicating high-quality matches, and the time required for external hires to become productive insiders. My results illuminate the converse mechanism: when older workers remain, these same forces can preserve valuable expertise and, across occupations, can benefit coworkers whose tasks complement the know-how accumulated by senior employees.

[36] I run these regressions on a random sample of all the workers to classify occupations, and merge these classifications with my analysis data.

[37] Subsample analyses by alternative measures of human capital specificity, such as managerial status and tenure, are available upon request.

This helps to reconcile seemingly contradictory findings—negative spillovers within jobcells but neutral or positive effects across jobcells—by showing that the sign of spillovers depends on whether workers compete on the same ladder or operate in complementary roles. Taken together, the evidence provides a richer account of how firms navigate workforce aging and underscores the central role of turnover frictions, match quality, and internal labor market structure in shaping the distributional effects of retirement reforms.

Future research could examine whether these within-firm adjustments propagate across firms and whether they generate general equilibrium responses in labor markets. Moreover, the limited ability to hire or promote younger workers may have consequences for productivity or service quality—for example, in health and care professions—where staffing shortages have been shown to harm patient outcomes (Friedrich and Hackmann, 2021). Finally, the mechanisms documented here—career-stage substitution, internal bottlenecks, and frictions in external hiring—are likely to be relevant in other settings where a component of the workforce is retained, and warrant further comparative study.

1.A Appendix: The Public Pension System, Reform, and Identification Details

Pathways to retirement. There are several pathways to retirement in Germany, including regular, disability, long-term insurance, women's, and unemployment pathways. While the rules of some of these pathways changed or the pathways abolished altogether, the workers eligible for regular pathways to retirement were subject to a single statutory retirement age, because ERA and NRA are equivalent for them. ERA exists on pathways for more vulnerable groups, including women, the unemployed, and the long-insured workers with over 35 contribution years. More details can be found in Lorenz et al. (2018).

Birth cohorts affected by the 1999 reform. Panel A in Figure 1.1 shows how the retirement age increased discontinuously starting from the 1952 birth cohort. Although the reform also abolished pensions for the unemployed and persons on a progressive retirement plan (Lorenz et al., 2018), I focus primarily on the abolishing of women's pathways to early retirement because the other two categories are not recorded in the data.

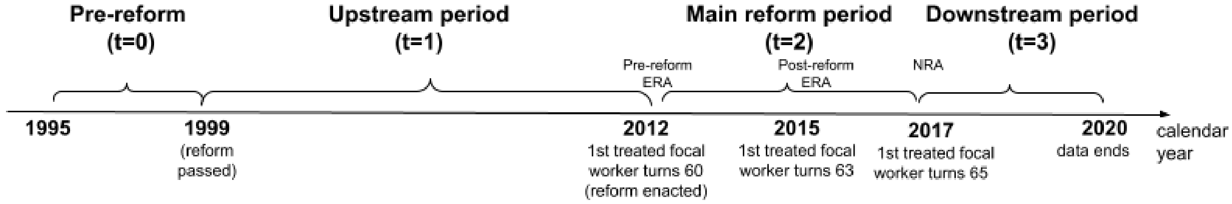
Timing of the 1999 reform. The 1999 Reform (*"Rentenreformgesetz 1999"*) abolished the women's pathway to early retirement. The reform was drafted in October 1997, and the affected cohorts were announced on December 17, 1997, through publication in the Federal Law Gazette; however, according to Etgeton, Fischer, and Ye (2023), there was uncertainty regarding implementation due to the federal elections in 1998, as the reform was drafted under the old government which might not have remained in power the following year.

The Social Democratic Party and the Green Party coalition promised to change the established reform during the election campaign, but even two months after the elections and their victory in September 1998, there was still uncertainty about which parts of the reform would be changed (Bulmahn, 1998; Etgeton, Fischer, and Ye, 2023). In the end, the new government did not revoke the abolition of the women's pathway to early retirement. Due to this uncertainty, the news of the reform is unlikely to have changed worker or firm behavior in advance (Etgeton, Fischer, and Ye, 2023). The reform became effective on January 1, 1999. The previous literature studying this reform, such as (Etgeton, Fischer, and Ye, 2023), uses 1998 as a pre-reform year. Given the uncertainty about implementation and use of the pre-reform period from 1998 in prior literature, I also use that year for the treatment construction to study the upstream period. Other papers (Geyer and Welteke (2021) and Badalyan (2025) among others) use the pre-reform enactment year, close to 60 years old, which I use for the main reform period.

The Figure 1.A.1 shows the reform timeline. Because the first affected cohort (1952) was only 47 years old when the reform became effective in 1999 and would turn 60 only in 2012, there is a large upstream period.

Illustration of empirical strategy. Below, I provide an illustration of the identification strategy in Equation 1 and Equation 2.

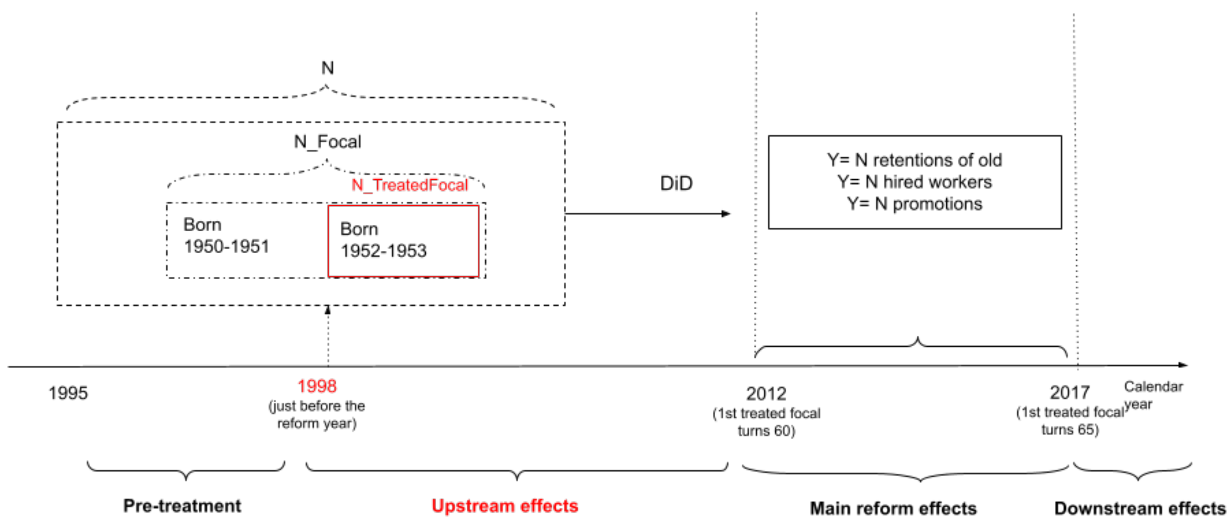
Figure 1.A.1: The reform timeline



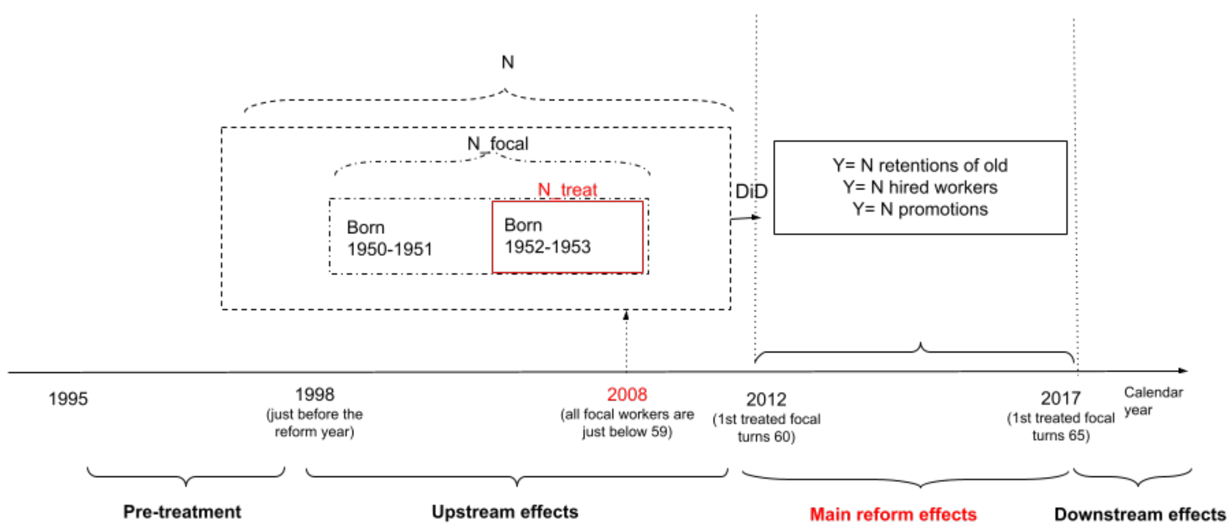
Notes: This figure represents the timeline of the reform. The upper section describes the periods, while the bottom part indicates the corresponding calendar years. For simplicity, the calendar years are written for the years of the ERA of the first affected cohort - the 1952 cohort (see Figure 1.1).

Figure 1.A.2: Illustration of identification strategy: difference-in-differences

Panel A: Illustration for identification of upstream effects



Panel B: Illustration for identification of main reform effects



Notes: This graph illustrates the identification strategy for upstream effects (**Panel A**, displayed in Equation 1) and main reform effects (**Panel B**, displayed in Equation 2).

1.B Appendix: Additional and/or Alternative Sample Definitions

As discussed in subsection 1.3, the universe of comparable establishments is reserved for the main reform analyses, where precise measurement of firm-level spillovers, such as within and across workgroups, and subgroup analyses are essential for mechanisms. For robustness and falsification tests, I use the Sample of Integrated Employer–Employee Data (SIEED), a 1.5% random sample of German establishments with complete employment biographies of workers ever employed in them, because it focuses only on intra-firm spillovers and provides the necessary longitudinal detail while ensuring data parsimony and compliance with data security constraints.

I create four additional samples. The Samples B-D below follow all the data creation steps as in the main reform period, and alter one attribute described. Sample E follows a similar sample construction to that in the upstream period, except that I observe full biographies of focal workers regardless of the establishment at which they are employed.

Sample B: placebo birth cutoff sample. These data are sampled analogously to the main period data, with the exception that I define focal workers as those women who were fully treated, that is, born in 1952-1955 cohorts.

Sample C: placebo gender sample. These data are also sampled analogously to the main period data, but I define focal workers as males rather than females.

Sample D: alternative window (1951-1952). In this section, I sample all the firms that had focal workers born within a 1-year bandwidth around the cutoff.

The sample sizes, including the number of coworkers, peers (by treatment), and workgroups, are recorded in Table 1.B.1.

Table 1.B.1: Sample sizes in baseline and alternative samples for the main reform period

	No. of establishments	No. of jobcells	No. of workers	No. of focal workers
Panel A: baseline sample				
Sample A	160,667	1,234,969	8,029,046	414,209
Panel B: alternative samples				
Sample B	2,621	19,565	127,381	7,276
Sample C	2,706	22,295	135,580	7,774
Sample D	1,637	13,655	98,722	3,290

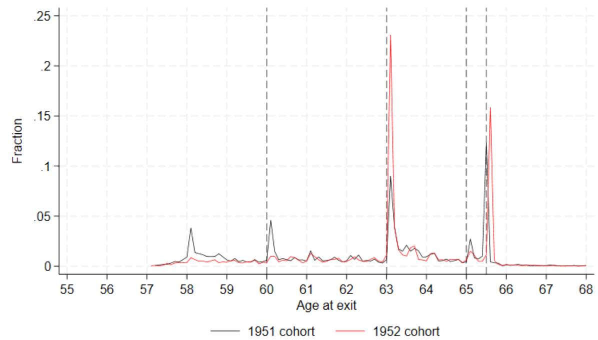
Notes: This table describes the number of establishments, jobcells, workers, and focal workers in the baseline (**Panel A**) and alternative samples (**Panel B**). For the details on baseline and alternative sample definitions, see subsection 1.3 and section 1.7.

Sample E: Individual-level focal worker biographies. I construct a complementary sample of focal workers born within a 2-year window of the January 1952 cutoff and follow their complete employment biographies, independent of the establishments in which they were employed in 1998 or 2008. This analysis is not feasible in the IEB data used for the baseline regressions, which restrict workers to their original establishments in order to identify intra-firm spillovers. Therefore, I rely on the SIEED 1975–2018, which provides complete employment histories for a random 1.5% sample of establishments. Unlike firm-level analysis, I do not condition on continued employment in a given establishment, since the objective here is to test whether treatment status affects workers' probabilities of employment, hiring, or separation. Treatment is defined by a dummy for being born after the cutoff. The resulting sample includes 14,707 focal workers observed across 23,264 establishments in which they were employed over their working lives.

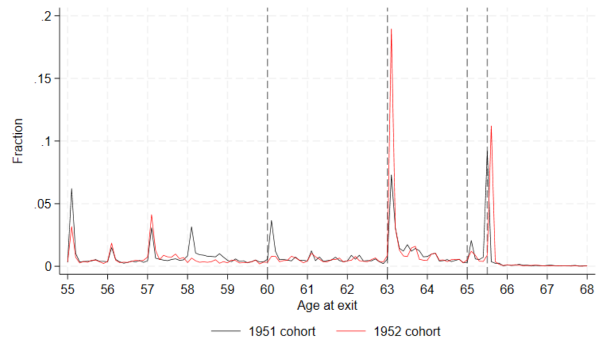
1.C Appendix: Figures

Figure 1.C.1: Retirement age distribution by birth cohorts

Panel A: Sample of workers employed at the age of 58



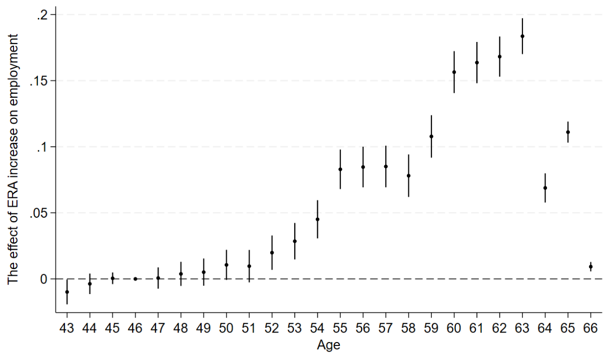
Panel B: Sample of workers employed in 1998



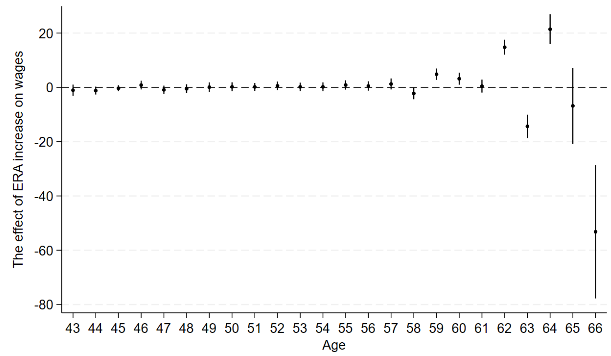
Notes: These graphs show the distribution of retirement ages. **Panel A** shows retirement ages based on focal workers employed at the age of 58. **Panel B** displays the distribution of retirement ages for workers employed in 1998. Both graphs are generated from the 2% random sample of IEB records.

Figure 1.C.2: Direct effects of the rise in ERA on employment, probability to become a new hire, or to separate from an employer

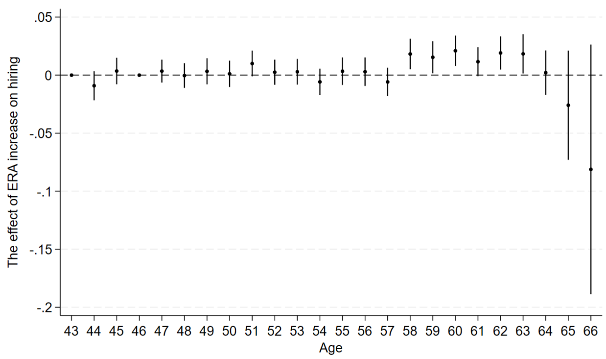
Panel A: Employment dummy



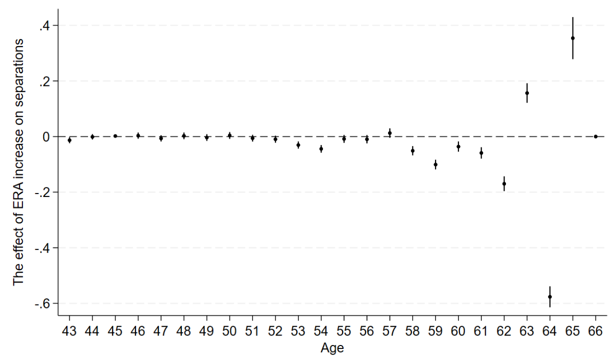
Panel B: No. of days employed at a given age



Panel C: Dummy for being a new hire



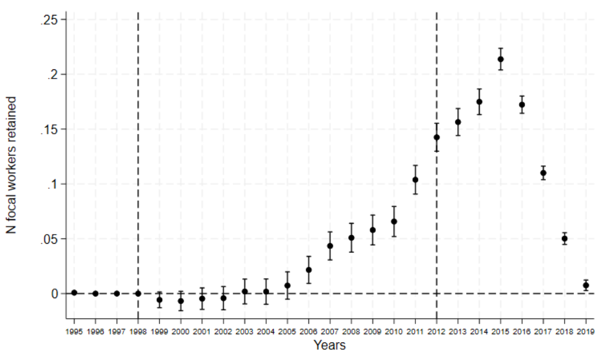
Panel D: Dummy for being separated



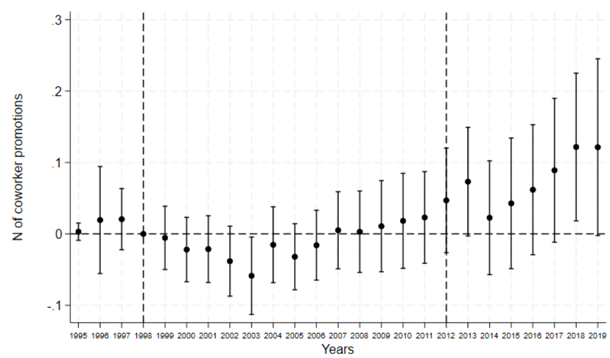
Notes: Coefficient plots. These graphs represent simple regressions of outcome variables on being born a year after the cutoff, separately for each age cohort. **Panel A**: employment dummy, **Panel B**: number of days employed at a given age, **Panel C**: dummy for being a new hire, **Panel D**: dummy for being separated. For sample construction details, see section 1.7.

Figure 1.C.3: The effect of an additional treated focal worker employed in 1998 on the probability of a firm closure, number of focal worker retentions, coworker promotions, and external hires

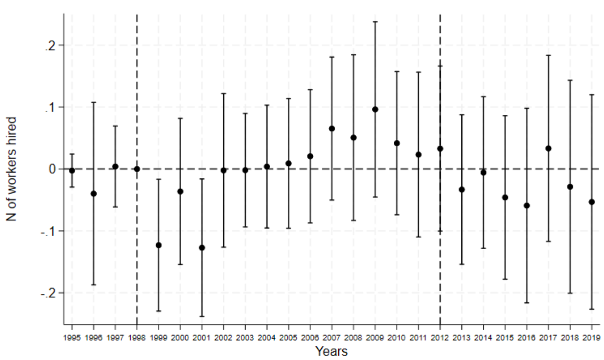
Panel A: No. of focal worker retentions



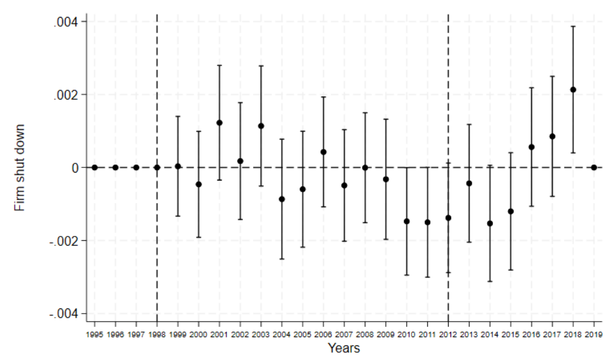
Panel B: No. of coworker promotions



Panel C: No. of hires

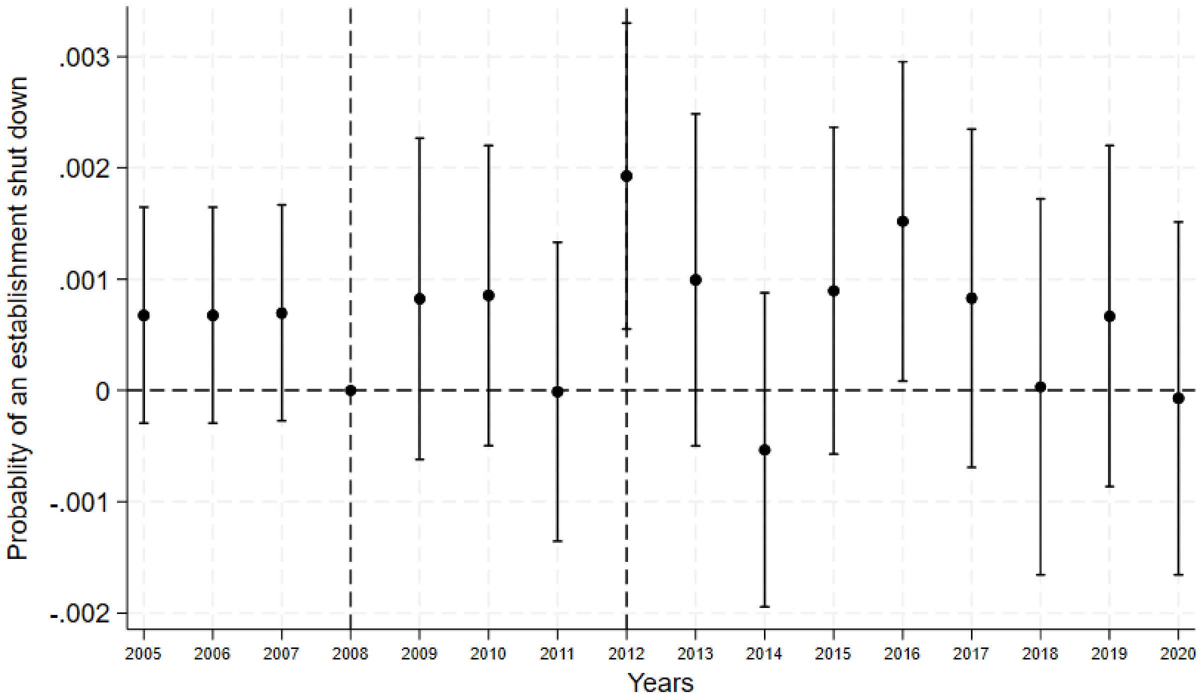


Panel D: Probability of a firm closure



Notes: This figure represents the effect of having one additional treated worker (1952–1953 birth cohorts) in 1998 on number of focal worker retentions (**Panel A**); coworker promotions (**Panel B**); external hires (**Panel C**); and firm closure (**Panel D**). The points represent the estimated coefficients β_t in Equation 1 and the vertical bars represent 95% confidence intervals. The dashed vertical line in 1998 represents the year before the reform passed, while the second dashed line represents when the first affected cohort (1952 birth cohort) reached the age of 60. Standard errors are clustered at the establishment level.

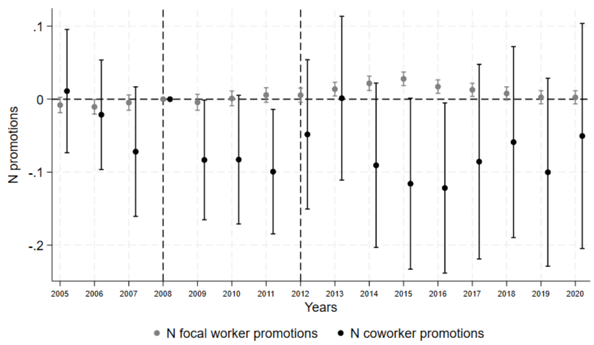
Figure 1.C.4: The effect of an additional treated focal worker employed in 2008 on the probability of an establishment closure



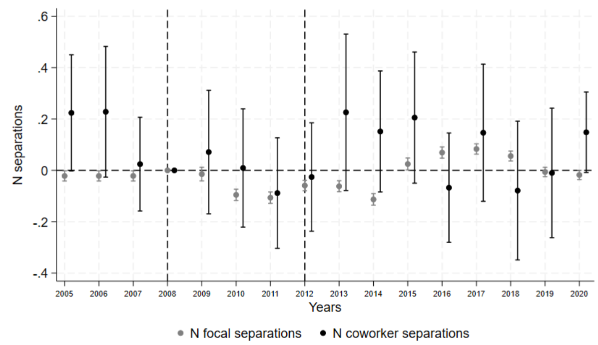
Notes: This figure represents the effect of having one additional treated worker (1952-1953 birth cohorts) in 2008 on the probability of a firm closure each year. The points represent the estimated coefficients δ_t in Equation 2 and the vertical bars represent 95% confidence intervals. The dashed vertical line represents the year before policy enactment, when all focal workers (1950-1953 birth cohorts) were under the age of 60. Standard errors are clustered at the establishment level.

Figure 1.C.5: The effect of an additional treated focal worker employed in 2008 on promotions and separations of focal workers and coworkers

Panel A: No. of promotions



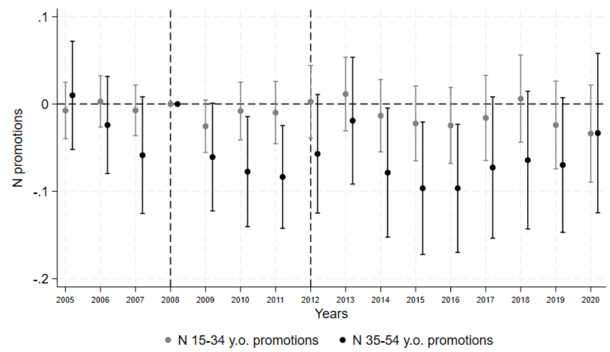
Panel B: No. of separations



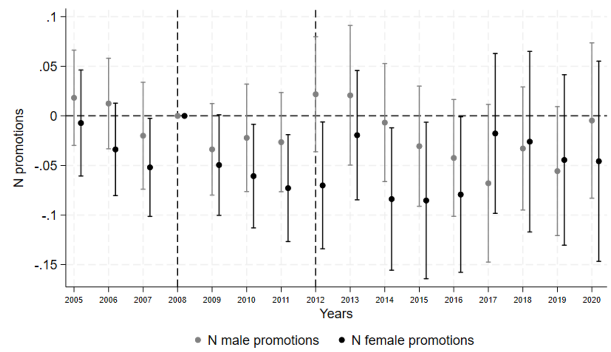
Notes: This figure represents the effect of having one additional treated worker (1952-1953 birth cohorts) in 2008 on the number of promotions (**Panel A**) and separations (**Panel B**) of focal workers (cohorts 1950-1953, in gray) and coworkers (in black) in each year. The points represent the estimated coefficients δ_t in Equation 2 and the vertical bars represent 95% confidence intervals. The dashed vertical line represents the year before policy enactment, when all focal workers (1950-1953 birth cohorts) were under the age of 60. Standard errors are clustered at the establishment level.

Figure 1.C.6: The effect of an additional treated focal worker employed in 2008 on coworker promotions, separations, and external hiring by age groups and gender

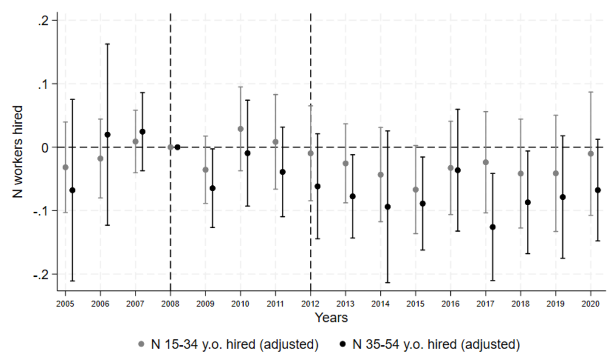
Panel A: No. of coworker promotions by age



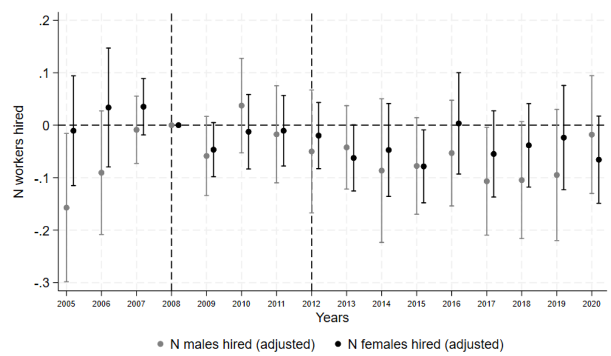
Panel B: No. of coworker promotions by gender



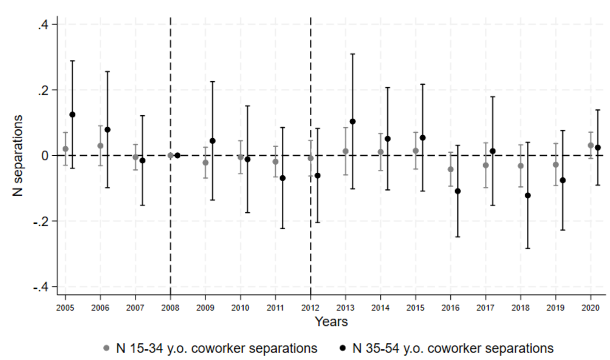
Panel C: No. of hires by age



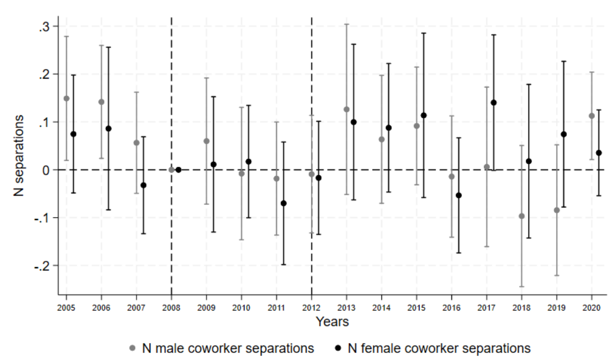
Panel D: No. of hires by gender



Panel E: No. of coworker separations by age



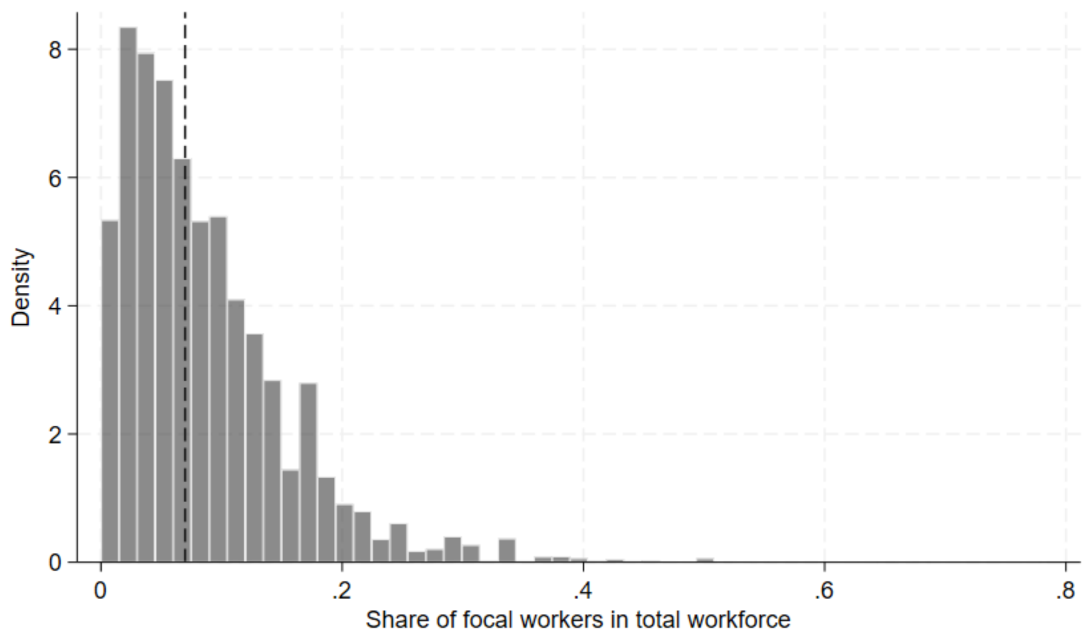
Panel F: No. of coworker separations by gender



Notes: This figure represents the effect of having one additional treated worker (1952-1953 cohorts) in 2008 on the number of coworker promotions by age (**Panel A**) and gender (**Panel B**); the number of hires by age (**Panel C**) and gender (**Panel D**); the number of coworker separations by age (**Panel E**) and gender (**Panel F**) in each year. The points represent the estimated coefficients δ_t in Equation 2 and the vertical bars represent 95% confidence intervals. The dashed vertical line represents the year before policy enactment, specifically when all the focal workers (1950-1953 cohorts) were under the age of 60. Standard errors are clustered at the establishment level.

Figure 1.C.7: Firm-level treatment variables

Panel A: Distribution of share of treated focal workers in total workforce



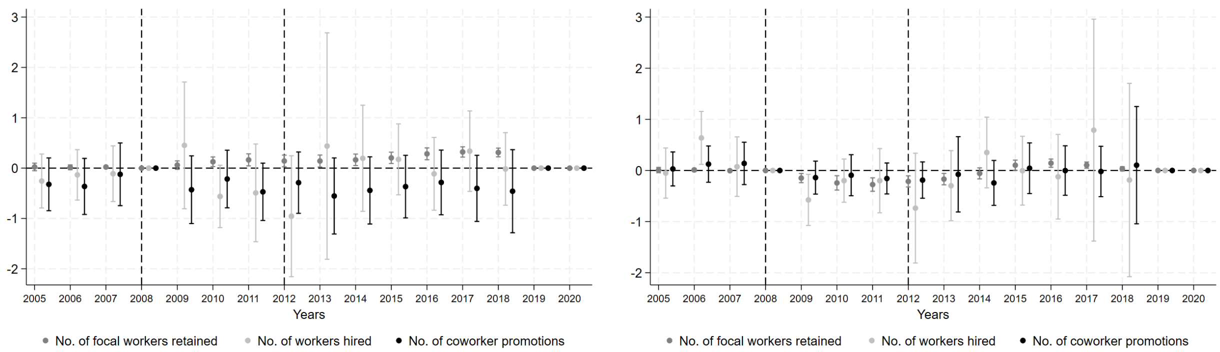
Panel B: Main treatment variables in 2008

Notes: **Panel A** displays the distribution of the share of focal workers in the total workforce in 2008. The vertical dashed line displays the median value in the distribution, which stands for almost 10% of the total workforce and is almost equivalent to the distribution mean. In both panels, the part-time workers are counted as 0.5 workers (the headcounts are adjusted for working hours). **Panel B** displays the share of 2008 workers (black circles show the share of treated focal workers (birth cohorts 1952-1953)) in focal workers (birth cohorts 1950-1953), medium-gray diamonds show the share of treated focal workers in the total workforce, and the light-gray triangles show the share of focal workers in the total workforce) over different establishment sizes. The horizontal dashed line at 0.5 indicates a point of no excess mass of share of treated focal workers among the total focal workers.

Figure 1.C.8: Falsification tests by birth cohorts and gender

Panel A: Placebo birth cohorts (1952-1955, all treated)

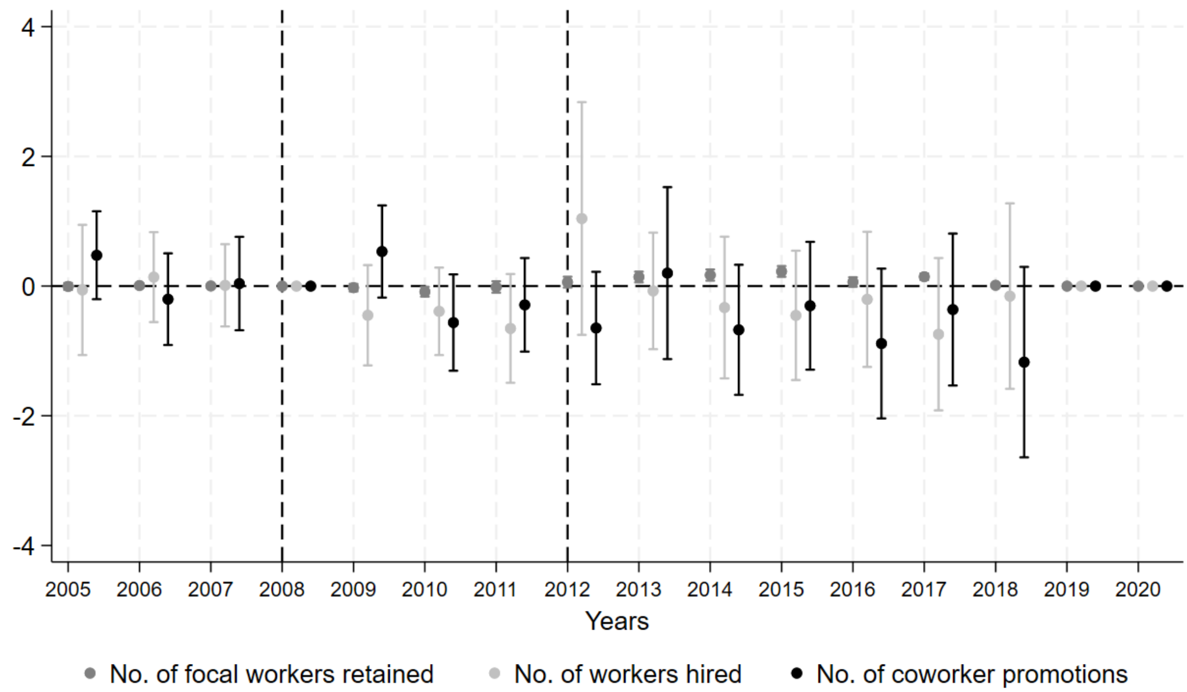
Panel B: Placebo gender (males)



Notes: This figure represents the effect of having one additional placebo-treated worker in 2008 on the number of focal worker retentions, number of external hires, and number of internal promotions. I perform falsification tests by redefining the focal workers as women born between 1952 and 1955, that is, all treated (**Panel A**) and as males born between 1950 and 1953 (**Panel B**). The points represent the estimated coefficients δ_t in Equation 2 and the vertical bars represent 95% confidence intervals. Standard errors are clustered at the establishment level. For sample construction details, see section 1.7.

Figure 1.C.9: Robustness check: altering the estimation bandwidth around the 1952 cutoff

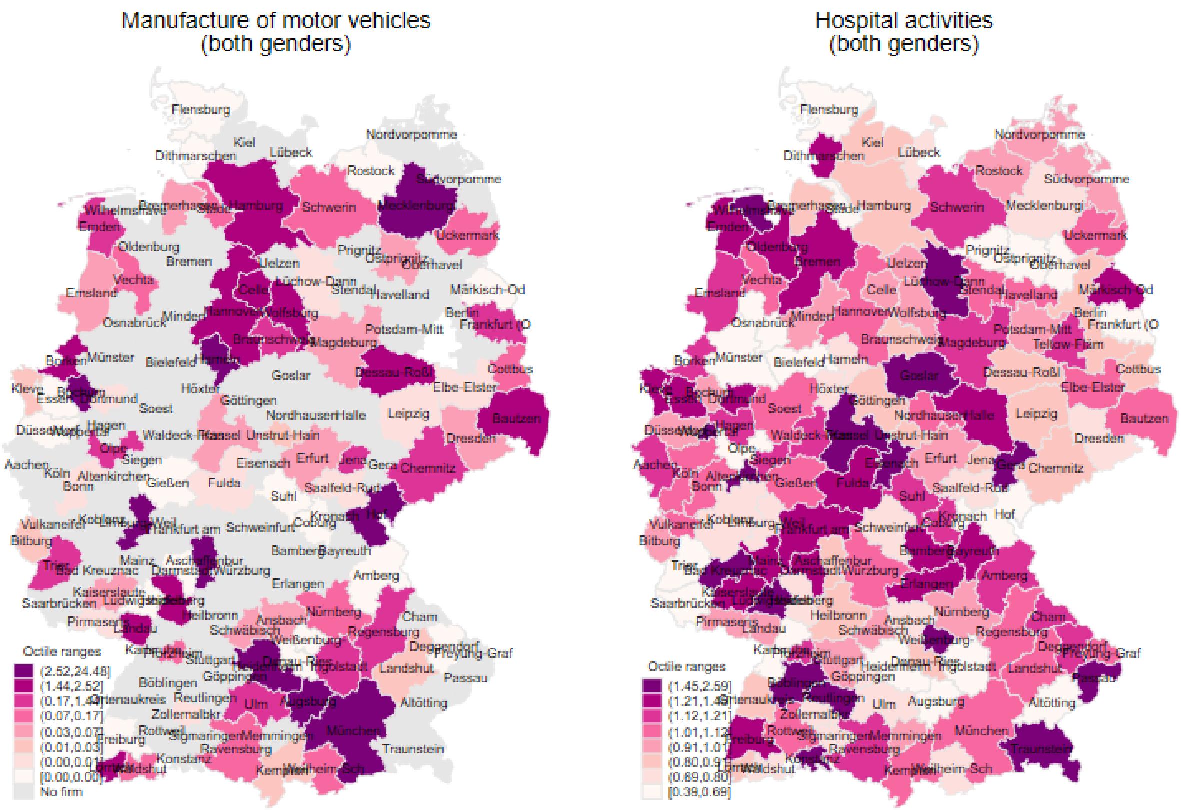
1-year bandwidth (1951-1952)



Notes: This figure represents the effect of having one additional treated worker (1952 cohorts) in 2008 on the number of focal worker retentions, number of external hires, and number of internal promotions. I perform a robustness test by redefining the window of focal workers as women born in 1951-1952, that is, a 1-year bandwidth. The points represent the estimated coefficients δ_t in Equation 2 and the vertical bars represent 95% confidence intervals. Standard errors are clustered at the establishment level. For sample construction details, see section 1.7.

Figure 1.C.10: Example of external labor market thickness in 2007

Panel A **Panel B**



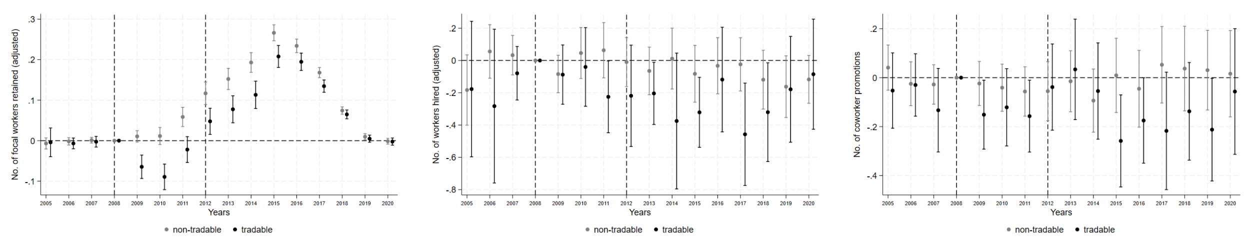
Notes: This map shows the computed external labor market thicknesses (ELMT) for each of the 141 German labor market regions based on the classification of Kosfeld and Werner (2012), based on high within-region commuting and low between-region commuting. I compute ELMT based on Equation 4 for the two large German industries: "manufacture of motor vehicles" (**Panel A**) and "hospital activities" (**Panel B**). I plot the ELMT indices on the map using the eight quantile ranges (octiles) shown in the left corner of each graph.

Figure 1.C.11: The effect of an additional treated focal worker employed in 2008 on focal worker retentions, coworker promotions, and external hiring by industry tradability

Panel A: No. of focal worker retentions

Panel B: No. of hires

Panel C: No. of promotions



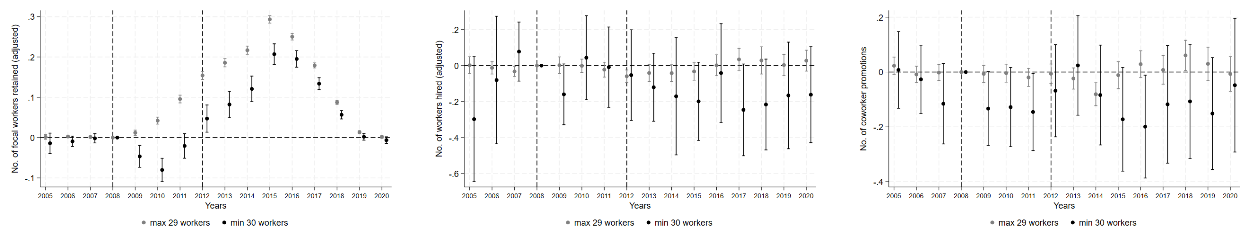
Notes: This figure shows the effects of an additional treated focal worker employed in 2008 on the number of retentions (**Panel A**), the number of external hires (**Panel B**), number of promotions (**Panel C**) by industry tradability. The points represent the estimated coefficients δ_t in Equation 2 and the vertical bars represent 95% confidence intervals. The dashed vertical line represents the year before policy enactment, when all the focal workers (1950-1953 cohorts) were under the age of 60. Standard errors are clustered at the establishment level.

Figure 1.C.12: The effect of an additional treated focal worker employed in 2008 on focal worker retentions, coworker promotions, and external hiring by establishment size

Panel A: No. of focal worker retentions

Panel B: No. of hires

Panel C: No. of promotions



Notes: This figure shows the effects of an additional treated focal worker employed in 2008 on the number of retentions (**Panel A**), the number of external hires (**Panel B**), number of promotions (**Panel C**) by establishment size categories. I split the establishments by those below 30 (gray color) and above 30 (black color) workers. The points represent the estimated coefficients δ_t in Equation 2 and the vertical bars represent 95% confidence intervals. The dashed vertical line represents the year before policy enactment, when all the focal workers (1950-1953 cohorts) were under the age of 60. Standard errors are clustered at the establishment level.

1.D Appendix: Tables

Table 1.D.1: Share of industry employment by gender in 1998 and 2007

Industry	Panel A: 1998 share women	Panel A: 1998 share men	Panel B: 2007 share women	Panel B: 2007 share men
Agriculture, forestry, and fishing	0.32	0.68	0.32	0.68
Mining and quarrying	0.09	0.91	0.09	0.91
Manufacturing	0.26	0.74	0.25	0.75
Electricity, gas, steam, air conditioning supply	0.21	0.79	0.23	0.77
Water supply; sewerage, waste management and remediation activities	0.18	0.82	0.18	0.82
Construction	0.12	0.88	0.12	0.88
Wholesale and retail trade; repair of motor vehicles and motorcycles	0.50	0.50	0.50	0.50
Transportation and storage	0.27	0.73	0.25	0.75
Accommodation and food service activities	0.58	0.42	0.57	0.43
Information and communication	0.38	0.62	0.36	0.64
Financial and insurance activities	0.54	0.46	0.55	0.45
Real estate activities	0.49	0.51	0.49	0.51
Professional, scientific, and technical activities	0.53	0.47	0.53	0.47
Administrative and support service activities	0.42	0.58	0.40	0.60
Public administration and defense; compulsory social security	0.59	0.41	0.61	0.40
Education	0.67	0.33	0.67	0.33
Human health and social work activities	0.80	0.20	0.80	0.20
Arts, entertainment, and recreation	0.48	0.52	0.50	0.50
Other service activities	0.66	0.34	0.66	0.34
Activities of household as employers; undifferentiated goods and services-producing activities of households for their own use	0.86	0.14	0.88	0.12
Activities of extraterritorial organizations and bodies	0.33	0.67	0.36	0.64

Notes: This table shows the female and male employment share in each of 22 industries. I aggregate 3-digit industries (based on the 2008 classification) into 22 groups following suggestions by Statistisches Bundesamt. The numbers are generated from the universe of full-time employed workers aged 18-64 employed in jobs subject to social security or vocational training as of June 30th, 1998 (**Panel A**) or 2007 (**Panel B**).

Table 1.D.2: Sample restrictions to obtain the original data extract for upstream and main reform periods

Restriction	No. of establishments	No. of workers
Panel A: upstream period		
Universe of establishment and workers in 1995-2019	8,611,676	69,208,790
+ observed in 1998	2,044,663	
+ employed at least 1 focal worker in 1998	413,995	
+ private sector	382,007	
+ at least 5 employees in 1998	221,853	
+ at most 500 employees in 1998	218,588	32,506,683
+ Workforce with positive wages and employment subject to social security in 1998, and redefined focal group (cohorts 1950-1953)	140,222	21,774,237
+ 5-199 workers	131,592	15,625,535
Panel B: main reform period		
Universe of establishments and workers in 1995-2020	8,241,529	69,296,143
+ observed in 2008	1,958,754	23,798,218
+ employed at least 1 focal worker in 2008	352,836	15,364,408
+ private sector	317,912	13,388,108
+ at least 5 employees in 2008	193,612	13,059,745
+ at most 500 employees in 2008	190,228	9,117,917
Universe of affected establishments, and their employed workers in 1995-2020	190,228	26,593,003

Notes: This table shows the number of establishments and workers after each restriction in the data extract requested, separately for the upstream period (**Panel A**) and main reform period (**Panel B**).

Table 1.D.3: Comparison of characteristics in analysis sample with a random sample of German establishments

	Panel A: upstream period		**Panel B: main reform period**	
	Random sample	Sampled establishments	Random sample	Sampled establishments
located in East Germany	0.234 (0.424)	0.264 (0.441)	0.209 (0.407)	0.221 (0.415)
No. of non-German workers	11.784 (84.822)	2.434 (6.693)	1.154 (19.372)	3.216 (10.174)
No. of female workers	5.322 (35.329)	18.755 (23.43)	6.677 (33.933)	24.335 (37.797)
No. of workers with university degree	1.519 (21.804)	3.835 (9.01)	2.516 (40.415)	6.643 (18.181)
No. of workers 15-34 y.o.	4.486 (36.634)	12.631 (15.084)	4.064 (31.398)	11.971 (20.64)
No. of workers 35-54 y.o.	6.722 (52.281)	21.375 (22.962)	9.583 (78.63)	30.381 (46.063)
No. of workers 55+ y.o.	1.475 (10.704)	4.864 (6.865)	1.991 (11.456)	7.622 (10.742)
No. of full-time workers	10.81 (87.239)	33.435 (36.106)	12.765 (107.217)	39.396 (62.526)
No. of part-time workers	1.842 (17.585)	5.085 (12.21)	2.867 (19.896)	10.562 (24.56)
Observations	30,296	131,592	21,581	160,667

Notes: This table shows the characteristics of a random sample of establishments (1.5% random sample based on SIEED7518 data and the sampled establishments (universe of affected firms sampled from IEB). IEB data are described in subsection 1.3, while the sampling is described in subsection 1.3. The comparison is performed separately for the upstream period (**Panel A**, variables measured in sampling year 1998), and main reform period (**Panel B**, variables measured in sampling year 2008).

Table 1.D.4: Comparison of the industry composition in the analysis sample with a random sample of German establishments.

	Panel A: upstream period		Panel B: main reform period	
	Random sample	Sampled establishments	Random sample	Sampled establishments
Agriculture, forestry, and fishing	0.017 (0.128)	0.016 (0.125)	0.016 (0.127)	0
Mining and quarrying	0.002 (0.043)	0.002 (0.048)	0.002 (0.049)	0
Manufacturing	0.111 (0.314)	0.232 (0.422)	0.109 (0.311)	0.199 (0.399)
Electricity, gas, steam and air conditioning supply	0.002 (0.044)	0.004 (0.066)	0.003 (0.051)	0.005 (0.067)
Water supply; sewerage, waste management, and remediation activities	0.004 (0.064)	0.007 (0.085)	0.004 (0.067)	0.008 (0.087)
Construction	0.121 (0.326)	0.078 (0.269)	0.107 (0.31)	0.049 (0.215)
Wholesale and retail trade; repair of motor vehicles and motorcycles	0.229 (0.42)	0.231 (0.421)	0.211 (0.408)	0.247 (0.431)
Transportation and storage	0.043 (0.204)	0.036 (0.185)	0.041 (0.199)	0.039 (0.195)
Accommodation and food service activities	0.07 (0.255)	0.042 (0.201)	0.056 (0.23)	0.043 (0.203)
Information and communication	0.015 (0.123)	0.022 (0.145)	0.021 (0.144)	0.02 (0.14)
Financial and insurance activities	0.024 (0.154)	0.038 (0.191)	0.027 (0.162)	0.034 (0.182)
Real estate activities	0.024 (0.152)	0.013 (0.113)	0.021 (0.142)	0.012 (0.107)
Professional, scientific, and technical activities	0.082 (0.275)	0.231 (0.421)	0.089 (0.285)	0.075 (0.263)
Administrative and support service activities	0.036 (0.186)	0.036 (0.185)	0.042 (0.201)	0.047 (0.211)
Public administration and defense; compulsory social security	0.018 (0.131)	0.042 (0.201)	0.019 (0.137)	0
Education	0.024 (0.154)	0.013 (0.115)	0.031 (0.173)	0
Human health and social work activities	0.093 (0.29)	0.082 (0.275)	0.118 (0.323)	0.15 (0.357)
Arts, entertainment, and recreation	0.013 (0.113)	0.008 (0.088)	0.012 (0.108)	0.01 (0.101)
Other service activities	0.056 (0.23)	0.043 (0.203)	0.059 (0.236)	0.049 (0.216)
Activities of the household as employers; undifferentiated goods and services-producing activities of households for their own use	0.014 (0.119)	0 (0.021)	0.009 (0.097)	0 (0.011)
Activities of extraterritorial organizations and bodies	0.002 (0.047)	0 (0.011)	0.001 (0.031)	0.014 (0.118)
Observations	30296	131592	21,581	160,667

Notes: This table shows the characteristics of a random sample of all the establishments in Germany (1.5% random sample based on SIEED7518 data and the sampled establishments (universe of affected firms sampled from IEB). IEB data are described in subsection 1.3, and the sampling is described in subsection 1.3. I aggregate 3-digit industries (based on the 2008 classification) into 22 groups following suggestions by Statistisches Bundesamt. The comparison is performed separately for the upstream period (**Panel A**, variables measured in sampling year 1998), and the main reform period (**Panel B**, variables measured in sampling year 2008).

2 Retirement Age Reforms and Worker Substitutability: Implications for Employment of Older Workers

Single-authored paper.[38]

2.1 Introduction

The dynamics of labor markets are profoundly influenced by the interplay between worker substitutability and firm-specific human capital. The ease with which workers can be replaced affects various labor supply decisions, including absences due to temporary illness (Hensvik and Rosenqvist, 2019), the duration of actual parental leave in reaction to extension of parental leave duration (Ginja, Karimi, and Xiao, 2023) and increase of paid parental leave eligibility coverage (Huebener et al., 2024), and labor supply following a coworker's death (Jäger and Heining, 2022). Worker substitutability has also been associated with wage losses after job displacement (Jacobson, LaLonde, and Sullivan, 1993), as workers with more specific skills, such as those tied to a particular industry or occupation, face greater difficulty finding comparable jobs in the external labor market. However, the role of worker substitutability in the context of retirement, a significant driver of workforce turnover, remains underexplored.

In this paper, I analyze a German retirement age reform that raised women's pensionable age by at least three years, to examine how worker substitutability shapes employment and wage outcomes. While the analysis focuses on employment, I interpret these outcomes through firms' incentives to retain older workers (i.e., to continue employing them) when turnover costs are high, linking individual labor supply decisions to firm-side labor demand and the interests of firms facing replacement costs. In this sense, employment at older ages reflects firms' willingness to retain workers given turnover costs, making employment outcomes informative about organizational decisions and stakeholder interests.

While a substantial literature examines how statutory retirement age reforms impact labor

[38]An earlier version of this paper circulated as a working paper (Badalyan, 2025). I thank Achim Ahrens, Dan Black, Wolfgang Dauth, Randall Filer, Johannes Geyer, Štěpán Jurajda, Andreas Mense, Nikolas Mittag, and Paolo Zacchia for their feedback; Dan Black for inviting me to the University of Chicago, where parts of this paper were written; Michael Moritz, Wolfgang Dauth, and the "Regional Labour Markets" Department at the Institute for Employment Research (IAB) for their belief in this project; Philipp vom Berge and Katja Wolf for help with understanding the data; Thomas Zwick for sharing a programming file that computes pension eligibility; Deborah Nováková from the CERGE-EI Academic Skills Center for language editing. This paper benefited from presentations at CERGE-EI Applied Brown Bag; ILRR workshop on Organizational Perspectives on an Aging Workforce (all 2026); HUN REN 2025; ifo Institute 2025; EWMES 2024; EALE 2024 & 2023; IAB Brown Bag and Regio Flash Talks 2024; ESPE 2024; IZA Summer School 2024; Dutch National Bank 2024; Young Economists Seminar (Croatian National Bank) 2024; SITES 2023; AIEL 2023 & 2022; Student Workshop at Harris School of Public Policy at UChicago 2023; BSE Summer School 2022; Czech Economic Society Biennial Conference 2022; Armenian Economic Association Annual Meeting 2022; CERGE-EI Brown Bag 2023, Applied Microeconometrics Reading Group 2022, DW 2022 & DPW 2020 Seminars.

supply (Atalay and Barrett, 2015; Brinch, Vestad, and Zweimüller, 2015; Geyer and Welteke, 2021; Hanel and Riphahn, 2012; Hernæs et al., 2016; Lalive and Staubli, 2015; Lalive, Magesan, and Staubli, 2023; Manoli and Weber, 2016; Mastrobuoni, 2009; Staubli and Zweimüller, 2013; Vestad, 2013), there is limited understanding of how labor demand mechanisms, such as job-specific skills and worker substitutability, shape employment responses to such reforms because existing analyses often assume that labor demand is perfectly elastic at the relevant margins. In contrast, my paper argues that labor demand is not uniformly elastic and highlights the role of worker substitutability in shaping firms' retention decisions. This paper aims to bridge this gap in the retirement literature by integrating insights from studies on worker substitutability with research on employment reactions to retirement reforms. Understanding this mechanism is crucial, as it offers deeper insights into how worker substitutability influences labor supply adjustments to retirement reforms and the coping strategies adopted by workers and firms. This challenges the standard assumption of uniformly elastic labor demand and offers new insights into the incidence and efficiency of retirement reforms.

The seminal study by Becker (1962) posits that firm-specific human capital renders incumbent workers less substitutable by external hires. Under such imperfect labor market conditions, the old-age employment decisions are driven by "push factors", mainly explained by labor demand (Boeri and van Ours, 2021). In the context of reforms that raise the retirement age, this theory suggests that employment responses by older workers may exhibit substantial heterogeneity based on their substitutability and the specificity of the human capital required for their roles. A pertinent question arises: When early retirement options are curtailed, do firms respond uniformly across worker types, or do employment gains disproportionately accrue to those with more specific skills and lower substitutability? Such differences may reflect how firms and workers coordinate—depending on their turnover costs—in response to extended employment horizons. The demand for workers rises due to firm- or job-specific human capital, or challenges in finding suitable replacements internally or externally. However, in the presence of outside options in the form of pensions, firms may have difficulties retaining such workers. Reforms raising the retirement age could help firms to retain such workers.[39]

Employment decisions at older ages are affected by many factors, including health, ability, income, and flexibility of contracts and firms; hence, in the absence of exogenous drivers, such decisions are likely endogenous at the individual level. Moreover, given an option to retire and receive a pension, workers may opt to exit the workplace and instead prioritize personal benefits (such as health, leisure time with family, etc) over firm factors (such as their substitutability and costs of replacement) in deciding to retire. A reform that raises the retirement age shifts the employment dynamics of those affected. I overcome this endogeneity challenge by studying the effects of a reform in Germany that abolished the women's pathway to early retirement by making the statutory retirement ages gender neutral. This reform

[39]Stole and Zwiebel (1996) and Stole and Zwiebel (1996) provided theoretical discussions of intra-firm bargaining and its relation to firm-specific human capital, while Lazear (2009) and Cahuc, Marque, and Wasmer (2008) extended the discussion by arguing that, similar to firm-specific human capital, the ease with which a firm can find a suitable replacement could affect the wages of workers. However, having lower bargaining power after the removal of the option to receive pensions, firms may be in a stronger position than workers.

resulted in a sharp rise of at least three years (from 60 to 63) in the Early Retirement Age (ERA), the earliest age women could begin to claim a pension. This discontinuous policy change, which impacted women born from 1952 onward, provides a natural experiment for causally identifying the effect of raising the retirement age on employment and wages using a Regression Discontinuity Design (RDD), and exploring the relationship of worker substitutability with a large labor supply increase.

The German labor market, characterized by substantial variation in worker substitutability[40] and strong dismissal protections, offers a suitable setting for investigating whether workers delay retirement based on their skills and substitutability. The availability of comprehensive German establishment data, which encompasses entire workforces and employment histories, together with job cell data (3-digit occupation groups within the establishments), enables analysis of internal markets, measurement of the availability of internal substitutes (workers sharing the same 3-digit occupation), and a study of personnel practices employed by the establishments.

To examine how employment responses to the rise in retirement age interact with worker substitutability, I start by sketching a simple model of the interplay between the reform that raises the age of the option to receive pensions, turnover costs, and employment decisions at 60-62. I also outline a Nash bargaining model with implications for the effects of the reform and of substitutability on wages conditional on employment at 60-62.

To test these implications empirically, I first construct several proxies for worker substitutability (and therefore turnover costs). First, I examine whether workers with specific skills are more likely to be retained at older ages. Specific human capital and managerial roles are key determinants of worker substitutability, as external replacements for these skills are often scarce (Baker, Gibbs, and Holmstrom, 1994)[41]. Consistent with theories of firm- and job-specific human capital (Becker, 1962), Bertheau (2021) shows that jobs requiring teamwork and training with senior workers are more often filled internally. In the context of retirement reform, this suggests that establishments where older workers' positions rely on job- or firm-specific human capital may benefit most from the extended retention of older workers. Next, I explore internal (coworkers in the same occupation) and external (potential hires in a commuting zone for a given occupation or industry) labor market thickness. According to Topel and Ward (1992), both internal and external labor markets affect workers' life-cycle labor market outcomes. In thin labor markets, finding suitable replacements is more challenging, making worker turnover costly for firms (Lazear, 1979). Automation can substitute for some types of labor, leading to reduced employment and wages, particularly in economies with aging populations like Germany (Acemoglu and Restrepo, 2022). Hence, I test whether the substitutability matters beyond the worker level, by dividing occupations by routineness, a proxy for substitution by automation. Finally, I consider the tradability of industries as another dimension of worker substitutability. Firms in tradable industries can replace workers

[40]Previous literature for Germany has shown that frictions in replacing workers are important (Jäger and Heining, 2022; Huebener et al., 2024).

[41]See also Bartel et al. (2014), Friedrich and Hackmann (2021), Jäger and Heining (2022), Jaravel, Petkova, and Bell (2018)

not only locally but also by outsourcing tasks globally, increasing substitutability (Drenik et al., 2023). While characteristics such as managerial status or skill specificity may reflect both firm-side costs and worker-side preferences, I interpret heterogeneity in the reform's effects primarily through the lens of firms' retention incentives — that is, the labor demand channel.

My findings confirm the implications of the model and indicate that the reform increased employment among women aged 60–62 by 17.3 percentage points (a 22% increase relative to the control mean of workers who were eligible to retire at 60). These results are robust to variations in model specification. To gauge the potential scale of the reform's impact, I conduct a back-of-the-envelope calculation. This treatment effect would translate into roughly 540,000 additional women remaining employed at ages 60–62 due to the reform.[42] Conditional on employment, the workers whose retirement age rose by the reform are less likely to bargain for higher wages at ages 60–62, compared to those previously eligible for pension benefits. The reform removed access to early retirement, weakening outside options and shifting bargaining power toward employers. This effect is likely amplified for older workers with specific skills and low substitutability, who already face limited mobility in the labor market. The observed decline in monthly wages partly reflects a compositional shift toward part-time or partial-retirement contracts, but also suggests a change in the wage-setting environment. These patterns are consistent with monopsony models, where firms exploit weak outside options to offer lower wages or fewer hours. Recent evidence for Germany supports this interpretation, showing that firms in more monopsonistic labor markets reduce wage costs when workers' alternatives are constrained (Plöger, 2024).

My findings reveal that raising the retirement age does not have a uniform effect across workers; instead, its impact depends on how easily firms can replace those approaching retirement. The largest employment gains are observed among women whose leave would be associated with high turnover costs for the employers, i.e., women with specific skills and those who are employed in occupations that are more difficult to replace both internally and externally. The findings suggest that reforms raising the retirement age are most effective in extending careers for workers who are less easily replaced, shedding light on the interplay between firm- and occupation-specific human capital, labor market frictions, and retirement decisions. It is noteworthy that substitutability by automation does not display post-reform differences. Moreover, external substitutability of industries does not show differences as widely as do the external substitutabilities of occupations, nor does the tradability of industries. These findings suggest that substitutability by humans is more likely to explain retirement decisions than substitutability by automation, and that skills and occupations are more linked to substitutability than industries. Importantly, heterogeneous RDD effects do not imply that firms actively dismiss substitutable workers; employment protection laws make such terminations unlikely. Rather, the reform removes the early retirement option, shifting the retention margin: firms now retain more older workers overall, but the largest increase occurs

[42] This is a rough calculation based on local treatment effects for women born in 1951–1952, who were employed continuously at 58-59 years old. The estimate assumes that the sample is nationally representative and that the effect generalizes across cohorts affected by the reform. It does not adjust for compositional differences or cohort trends and should be interpreted as illustrative.

among non-substitutable workers—those who previously left voluntarily when early retirement was available. Other explanations could be that some workers who are not retained can, for example, be encouraged by employers to use up to two years of unemployment as a bridge to retirement (Gudgeon et al., 2023), or choose self-employment or inactivity. Due to data limitations, I do not test for these alternative explanations, and only focus on employment. Because Geyer and Welteke (2021) do not find strong evidence for active substitution to unemployment for this specific reform, I conclude that the main explanation of these results are pre-reform voluntary exits of substitutable workers.

Before the reform, i.e., when these women were eligible to retire at 60, although firms faced greater constraints when substitutes were scarce, they generally could not prevent women from retiring at ages 60 to 62. Therefore, before the reform, retirement decisions were primarily driven by workers rather than employers. However, after the reform, firms became more likely to retain women who were less substitutable—even as they were offering them lower wages compared to their peers who were eligible to retire at 60. These findings imply that raising the retirement age shifts the dynamics of retirement decisions from the individual level to the firm level, conditional on turnover costs measured by low substitutability. In this context, reforms that raise the retirement age may help firms operating in imperfect labor markets to better manage workforce turnover and skill retention. This is a significant relief for firms, especially as, according to Muehlemann and Pfeifer (2016), firms in Germany bear sizable hiring costs for high-skilled labor, amounting to almost two months' wages.

The effects of the retirement age rise on wages vary across worker types. While overall wage bargaining power declines due to the reform, wage increases are observed among managers and workers in occupations with thin external labor markets, consistent with firms raising wages to retain more difficult-to-replace employees. However, this pattern does not hold for all subsamples that proxy for high turnover costs. With the early retirement path closed, older workers are effectively locked into the labor market, and for some, into their current firm, especially those in thin labor markets or with high job-specific skills. This weakens their outside options and enhances employers' monopsony power, allowing firms to suppress wages even for valuable workers. The reform effectively increases firms' monopsony power over older workers by removing early retirement as a credible outside option. This shift particularly affects less substitutable workers, who, before the reform, could leverage their high retention costs and scarce replacements to negotiate better conditions or to exit. With retirement no longer available until age 63, these workers become more reliant on their current employer, which limits their bargaining position despite their value to the firm. This explains why employment rises without proportional wage gains and highlights how policy-induced changes to outside options can amplify monopsony effects in segmented labor markets.

This study contributes to several strands of the literature. First, it adds to the empirical research on worker substitutability and workplace characteristics affecting labor market decisions (Jäger and Heining, 2022; Ginja, Karimi, and Xiao, 2023; Huebener et al., 2024). My results align with recent evidence on firm responses to retention shocks. Jäger and Heining (2022) find that firms facing the removal of their least substitutable workers experienced substantial wage growth and hiring strain. Similarly, Huebener et al. (2024) document how extended paid parental leave weakened the link between internal substitutability and return-to-work behavior, showing how such policy-induced frictions can distort employer–employee coordination. Opposed to a reform that increases worker absence due to parental leave, I study a reform that decreases worker absences due to a rise in retirement age. My findings suggest that the reform delaying retirement strengthens the employer-employee coordination: firm responses depend on skill-specific turnover costs, with large effects concentrated among workers with few internal or external substitutes. My paper introduces novel evidence on how substitutability mediates firm responses to a retirement age increase. A key distinction is the nature of the shock: whereas parental leave is temporary and expected (as the mothers usually return to their prior employers), raising the early retirement age (ERA) from 60 to 63 binds older workers more tightly to their jobs up until their new pensionable age. Anticipated parental leave absences allow firms to plan. By contrast, the removal of early retirement compresses exit options, increasing reliance on specific workers while weakening their leverage in wage negotiations. I find that firms are more likely to retain workers who are more difficult to replace, as these workers have weaker bargaining positions due to reduced outside options. The reform thus reshapes employment, especially for less substitutable workers, and average wage dynamics.

Second, this paper contributes to the literature on employment decisions at older ages by examining the novel labor demand mechanisms that shape them. Most existing studies focus on mechanisms related to individual and household characteristics, while less attention has been given to the role of firms and labor demand. For example, previous research shows that retirement decisions are often coordinated within households, particularly among couples (Atalay, Barrett, and Siminski, 2019; Bloemen, Hochguertel, and Zweerink, 2019; García-Miralles and Leganza, 2024; Johnsen, Vaage, and Willén, 2022; Lalive and Parrotta, 2017; Selin, 2017; Zweimüller, Winter-Ebmer, and Falkinger, 1996). I extend this literature by demonstrating that older women also effectively coordinate their retirement timing with their employers, depending on potential turnover costs. Additionally, I extend the seminal paper by Geyer and Welteke (2021) that analyzed the 1999 reform[43] by (1) studying workplace labor demand mechanisms, in particular those highlighting turnover costs, worker job-specific skills and substitutability, which have not been analyzed for retirement reforms; (2) analyzing employment responses beyond 62 years of age, including bunching at the Normal Retirement Age; (3) analyzing whether the option to receive a pension before the reform helps workers to bargain for higher wages, i.e., the link between wages and employment, which has not been analyzed previously for this reform. My findings also complement Boockmann, Kroczek, and Laub (2023), who find that the effects of this 1999 reform are smaller for workers in physically and psycho-socially demanding occupations and manual non-routine tasks.

Third, while several studies have examined the role of firms in shaping retirement behavior, they primarily focus on institutional constraints. Deshpande, Fadlon, and Gray (2024) show that firms contribute to the rigidity of retirement decisions, with many workers continuing to retire at the pre-reform statutory age despite policy changes in the U.S. Similarly, Rabaté, Jongen, and Atav (2024) find that automatic job termination policies in the Netherlands

[43] Geyer and Welteke (2021) use data on pension insurance, which is not linked to workplaces.

drive much of the observed bunching at the statutory retirement age. The only paper that provides evidence on replacement costs using a different reform in Germany is by Geyer et al. (2022). They show that employers with a high share of older worker inflow compared with their younger worker inflow, employers in sectors with few investments in research and development, and employers in sectors with a high share of collective bargaining agreements allow their employees to remain employed longer after a reform that raised the normal retirement age. Frimmel et al. (2018) show that firms that have higher labor costs for old workers relative to their productivity are more likely to incentivize them to leave early. These studies imply that firms play an important role in old-age employment of workers. My research builds on these insights by introducing worker substitutability as a key labor demand factor influencing retirement responses that has been overlooked by the literature due to the scarcity of workplace data linked to retirement decisions. Detailed job-cell data from German social security records, combined with employment data of monthly frequency, allow for such analyses. Several papers analyze spillovers of raising retirement age on hiring using Italian (Bianchi et al., 2023; Boeri, Garibaldi, and Moen, 2022; Carta, D'Amuri, and Von Wachter, 2025) and Dutch data (Hut, 2024; Ferrari, Kabátek, and Morris, 2023); however, due to limited data on occupations and job cells in these administrative records, these studies do not analyze the direct effects of a reform on older workers' employment by availability of internal and external substitutes, or by human capital specificity of occupations, which I focus on in this paper. This study also bolsters understanding of the findings of papers that argue that institutional constraints and firms explain retirement behavior (Deshpande, Fadlon, and Gray, 2024; Rabaté, Jongen, and Atav, 2024). I focus on workers aged 60–62, who fall between the pre- and post-reform early retirement ages. At these ages, employment is not determined by formal contract changes or layoffs by employers, but by more implicit coordination between firms and workers. Building on this insight, I examine how these dynamics interact with voluntary early retirement decisions. I show that the reform, by restricting early retirement eligibility, enables firms to selectively retain workers who are more difficult to replace. My study extends theirs by focusing on retirement choices, policy-induced separation risk, and voluntary exits.

The structure of this paper is as follows. In subsection 2.2, I describe the institutional setting, including details about the 1999 reform that raised women's retirement age and the conceptual framework with implications of employment and wage dynamics. Subsection 2.3 presents the data and sample construction. Subsection 2.4 specifies the identification strategy I employ to study the effect of the reform on labor supply and the mechanisms associated with employment, while subsection 2.5 shows the corresponding estimation results for employment at 60-62 and wages, and robustness checks. Subsection 2.6 studies the mechanisms associated with labor demand, skills, and worker substitutability, and is followed by the conclusion.

2.2 Institutional setting and conceptual framework

This section presents the German labor market institutions and the 1999 reform, followed by the conceptual framework and implications I aim to test empirically.

2.2.1 Institutional setting

The German labor market is characterized as a labor market with relatively decentralized wage setting (Jäger and Heining, 2022; Dustmann et al., 2014). This labor market feature makes it easier for wages to deviate from the levels set by bargaining agreements. Overall, the labor market structure during the years under study makes it unlikely for firms to easily fire older workers.[44] Such regulation implies that the older workers are more likely to either leave voluntarily or in a subtle agreement with their employers through offering differentiated contracts, working hours, or wages. The downward rigidity of wages implies that wages usually decrease through offering different contracts, for example, through lower working hour agreements.

There are three pillars of the German pension system: public pensions, occupational pensions, and private provisions. Public pension insurance is the most popular choice among the working population, covering about 90% of the German workforce, according to Zwick et al. (2022). Given that in this paper I analyze a reform that changed some attributes in the public pension system, it had an impact on many people in the country because participation in the public pension is mandatory for all workers except for civil servants and the self-employed.[45].

The early retirement age (ERA) serves as a key behavioral anchor and coordination point in the German retirement system. It marks the first age at which workers can begin claiming a pension, albeit with actuarial deductions, while the normal retirement age (NRA) determines when a full, undeducted pension can be drawn. Workers respond strongly to both thresholds: Seibold (2021) documents bunching at these statutory ages, and Riphahn and Schrader (2021) and Geyer and Welteke (2021) find large labor supply shifts when either is reformed. In my setting, a reform raised the ERA from 60 to 63 for certain cohorts, effectively eliminating a prominent and widely used exit option between ages 60 and 62. This creates a new period where workers must either continue working or negotiate alternative exit paths. Importantly, this window still lies below the NRA, so continued employment is legally possible and common, but less predictable. Understanding the institutional role of ERA helps motivate the analysis: it clarifies why changes at 60 to 62 years generate observable effects and why firm decisions about retention and wages matter in the absence of this early retirement channel.

There are several pathways to retirement, including regular, disability, long-term insurance, women's, and unemployment. While rules surrounding some pathways changed and some were abolished altogether, workers eligible for the regular pathways to retirement had a single statutory retirement age. ERA followed by NRA applied to women, the unemployed, and the long-insured workers. Some of these pathways were modified or abolished, including the women's pathway that I analyze in this paper.

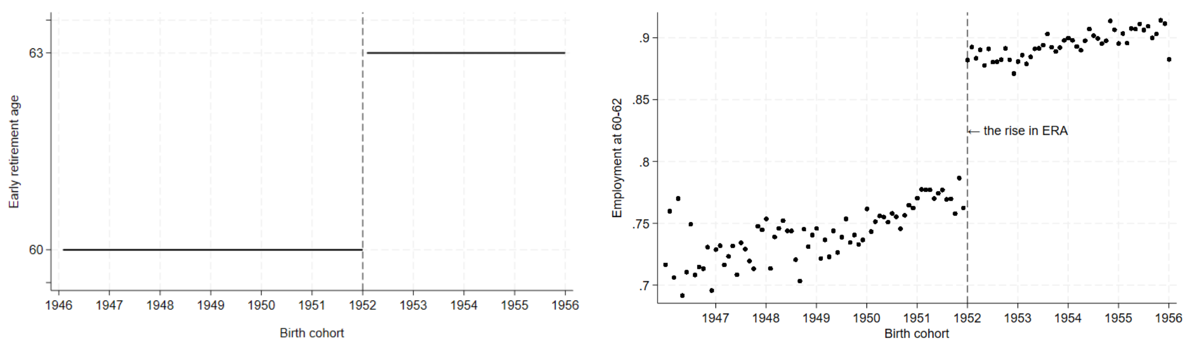
[44] The Equal Treatment Act protects older workers from unjustified dismissal (*Allgemeines Gleichbehandlungsgesetz – AGG*, General Act on Equal Treatment of 14 August 2006 (Federal Law Gazette I, p. 1897), as last amended by Article 4 of the Act of 19 December 2022, Federal Law Gazette I, p. 2510).

[45] The public pension system consists of a pay-as-you-go scheme. Pay-as-you-go scheme means that pensions of the old are paid by social security contributions of insured working people and taxes.

The 1999 reform abolished the women's pathway to early retirement at 60 years old by making the statutory retirement ages gender-neutral.[46] Before the reform, women could start claiming pensions earlier than men. Therefore, the gender-neutral statutory retirement ages introduced by the reform increased the early retirement age for women. The 1999 reform officially came into force on January 1, 1999, (Gohl et al., 2020), and affected women born from January 1, 1952. Hence, the change was discontinuous in terms of birth cohorts. For those who had accumulated enough contributions to be eligible for the long insurance pathway, the ERA rose by three years, while for workers on a regular pathway to retirement, the ERA rose by 5.5 years. Overall, the reform increased ERA for women by at least three years (Panel A of Figure 2.1).[47]

Figure 2.1: Discontinuity in birth cohorts

Panel A: ERA assignment rule **Panel B:** Fraction employed at 60-62



Notes: The **Panel A** shows the policy rule for the earliest age a person could claim pensions by birth cohort. The **Panel B** shows the scatter plot of the fraction of women employed at the ages 60-62 over the birth cohorts 1947-1956. The dashed line presents the birth cohort cutoff, January 1952, starting from which the ERA rose by at least three years.

2.2.2 Conceptual framework and implications

Firms operating in imperfect labor markets face frictions in replacing experienced workers, particularly those with occupation- or firm-specific skills. As these workers approach retirement age, firms risk productivity losses and incur hiring costs due to turnover. Early retirement eligibility grants workers considerable autonomy in deciding when to exit the labor

[46]While the reform also abolished early pathways to retirement for the unemployed and for persons under a progressive retirement plan (Lorenz et al., 2018), I focus primarily on the abolishment of women's pathway to early retirement because the other two categories are not recorded in the data.

[47]Before the 1999 reform, the NRA of women's pathway to retirement was fixed at 65. After the abolishment of women's pathway to early retirement, women were also affected by another reform that affected the regular pathway to retirement. In particular, due to the 2007 reform, workers on the regular pathway experienced a retirement age increase starting from 1946 by one month per birth year (Figure 2.A.1), and their retirement age is expected to reach 67 for the 1963 birth cohort by 2029. This 2007 reform affected the women under my study because the NRA of those born in 1951 was 65, while that of those born in 1952 became 65.5.

force. This paper studies how a policy reform that raised the early retirement age (ERA) from 60 to at least 63 alters the interaction between worker substitutability and retirement behavior.

The model builds on the idea that retirement is not only a worker's choice, but also reflects the relative bargaining power of workers and firms. When early retirement is an option, workers with valuable skills may leverage this as a bargaining chip in wage negotiations. When the option is removed, firms can retain even valuable workers without raising wages. To understand how firms respond to a rise in early retirement age, I first develop a static model in which the firm's decision to retain a worker depends explicitly on the worker's substitutability and the policy environment. I then extend the framework with a Nash bargaining model, allowing wages to be endogenously determined. This yields testable implications conditional on employment.[48]

2.2.2.1 Firm's problem of employment decisions

Setup. Consider a firm employing worker i, who is approaching retirement age. Continued employment at ages 60–62 depends on whether the match between the worker and the firm remains viable. I model this using a latent retention condition, in which both the firm and the worker must benefit from continued employment. While the firm's willingness to accommodate employment reflects the economic value of the match, the worker's outside option plays a key role in the joint decision. I do not model active dismissal, consistent with strong employment protections. Instead, I interpret "retention" as the match continuing when both parties find it preferable to separation. The reform removes a key voluntary exit channel (early retirement), extending employment among older workers, especially those with high specificity. I specify the firm maximization problem as follows:

$$\max_{d_i \in \{0,1\}} \left\{ \pi_i = d_i \cdot \underbrace{(y - o(R_i, s_i))}_{\substack{\text{surplus if} \\ \text{match continues}}} + (1 - d_i) \cdot \underbrace{(-c(s_i))}_{\substack{\text{cost if} \\ \text{match ends}}} \right\} \tag{14}$$

where:

y is the worker's output. I abstract from heterogeneity in output across workers, as substitutable workers may be either more or less productive depending on job fit and skill specificity.[49]

[48]The theoretical framework presented in this section builds on Nash bargaining models of labor market frictions (e.g., Pissarides (2000)), adapting them to retirement contexts by incorporating outside options shaped by policy. It also draws on Acemoglu and Pischke (1999) in the implications of firm-specific skills for wage setting and turnover, and from Gruber and Wise (2008) the responsiveness of retirement to institutional incentives. Lastly, the interaction between substitutability and tax incidence in determining the incidence of adjustment costs is conceptually linked to Gruber (1997), and is applied here to changes in retirement age.

[49]I thank Wolfgang Dauth for this discussion.

The model uses a single specificity parameter s_i to capture employer-side turnover costs. These costs increase when workers are more difficult to replace (due to specialized knowledge or task-specific skills) and when they have limited outside options (due to thin external markets for their skills). In this sense, specificity s_i represents a reduced-form measure encompassing both skill specificity and substitutability.

$o(R_i, s_i)$ is the outside option, shaped by the policy reform R_i and worker specificity s_i. It is decreasing in R_i ($\frac{\partial o(R_i,s_i)}{\partial R_i} < 0$) because pension eligibility is delayed post-reform, and decreasing in s_i because workers with specific skills face thinner external labor markets, especially after age 60 ($\frac{\partial o(R_i,s_i)}{\partial s_i} < 0$). Such specificity could, for example, decrease the likelihood of leaving the social insurance for other pathways than retirement (move to another country, start self-employment, etc.).[50]

$c(s_i)$ denotes the replacement cost of an employee, increasing in specificity, and reflecting dismissal, severance, and other compensation payments, as well as hiring, training, and productivity ramp-up costs that rise with the degree of human capital specificity ($\frac{\partial c(s_i)}{\partial s_i} > 0$).

Solution. The match continues if the joint surplus from continuing exceeds the cost of separation:

$$y - o(R_i, s_i) + c(s_i) > 0 \tag{15}$$

Interpretation. While the equation is modeled as a firm-side optimization problem, the interpretation reflects a joint agreement between the firm and the worker: continued employment occurs when both benefit relative to separation. The reform shifts this condition by removing early retirement as a fallback, thus altering the outside option $o(R_i, s_i)$ and increasing the likelihood of match continuation, especially for workers with high specificity.

Implication 1: A higher ERA rises employment of older workers.

$$\frac{\partial}{\partial R_i}\left(y - o(R_i, s_i) + c(s_i)\right) = -\underbrace{\frac{do(R_i, s_i)}{dR_i}}_{<0} > 0 \tag{16}$$

Delaying pension eligibility reduces $o(R_i, s_i)$, making continued employment more attractive to the firm and less avoidable for the worker, thereby increasing employment.

[50]Such argument is also in line with the finding of literature on displaced workers: those unemployed who switch to another industry or occupation experience much larger declines in earnings (e.g. Neal (1995); Addison and Portugal (1989)).

Implication 2: Workers with higher specificity (s_i) are more likely to remain employed.

$$\frac{\partial}{\partial s_i}\left(y - o(R_i, s_i) + c(s_i)\right) = -\underbrace{\frac{do(R_i, s_i)}{ds_i}}_{<0} + \underbrace{\frac{dc(s_i)}{ds_i}}_{>0} > 0 \tag{17}$$

Less substitutable workers are more likely to remain in employment due to both weaker outside options and higher replacement costs. As a result, the joint surplus of continued employment is larger, sustaining the match.

2.2.2.2. Wage determination under Nash bargaining

Conditional on retention ($d_i = 1$), the firm and the worker bargain over the wage w_i based on the total surplus generated by employment:

$$S_i = y - o(R_i, s_i) \tag{18}$$

With worker bargaining power $\beta \in (0, 1)$, the Nash wage splits the surplus between the worker and the firm. The wage thus depends on both the worker's productivity y and their outside option $o(R_i, s_i)$. The general form of the Nash-bargained wage is:

$$w_i = \underbrace{\beta \cdot y}_{\text{productivity-based reward}} + \underbrace{(1-\beta) \cdot o(R_i, s_i)}_{\text{outside-option fallback}} \tag{19}$$

Intuition. The worker's wage is a weighted average of what they contribute to the firm's output (through y) and what they could earn elsewhere (via $o(R_i, s_i)$). Workers with high productivity naturally command higher wages, all else equal. However, their outside option—such as retirement income or alternative employment—also determines their bargaining position. If their fallback option weakens, the firm can offer a lower wage even if the worker is productive.

To understand how wages change due to the reform and differences in substitutability, I examine how w_i responds to changes in R_i and s_i.

Implication 3: The reform lowers wages via weaker outside options.

$$\frac{\partial w_i}{\partial R_i} = (1-\beta) \cdot \underbrace{\frac{do(R_i, s_i)}{dR_i}}_{<0} < 0 \tag{20}$$

When the policy raises the early retirement age (i.e., increases R_i), the outside option $o(R_i, s_i)$ declines. This reduces the worker's fallback position in wage negotiations, shifting surplus toward the firm. The wage falls even though the worker remains employed. This mechanism is stronger when the worker has low bargaining power β, and when the reduction in outside options is large.

Implication 4: The effect of specificity on wages is ambiguous.

$$\frac{\partial w_i}{\partial s_i} = \beta \cdot \frac{dy(s_i)}{ds_i} + (1 - \beta) \cdot \frac{do(R_i, s_i)}{ds_i} \tag{21}$$

Higher specificity s_i affects both productivity and outside options. If more specific workers are more productive ($\frac{dy(s_i)}{ds_i} > 0$), then wages may increase through the first term. However, specificity also reduces outside options ($\frac{do(R_i,s_i)}{ds_i} < 0$), which lowers wages through the second term. If outside options deteriorate faster than productivity improves, or if the worker has low bargaining power, the overall effect on wages may be negative.

Intuition. Even if the worker is valuable to the firm (due to difficult-to-replace skills), the firm may exploit their lack of external alternatives. The reform amplifies this asymmetry by removing early retirement as a viable fallback, especially for workers in thin external labor markets (e.g., managers, specialists). This is a form of *monopsony power*, where the employer's ability to set wages below marginal product is strengthened by the worker's limited exit options.

Summary. Implications three and four jointly imply that the wage response to the reform depends on the interaction between substitutability and bargaining frictions. For workers with low specificity (who are easy to replace), wages fall mostly due to the loss of retirement options. For workers with high specificity, the story is more nuanced: while their productivity makes them costly to replace (increasing employment), their weakened fallback position gives the firm the leverage to suppress wages. Thus, employment may rise while wages fall or stagnate, despite high skill specificity, due to increased employer monopsony power post-reform. This prediction is aligned with recent evidence of wage setting in imperfect German labor markets.

2.3 Data

This section consists of two parts. First, I describe the data I utilize, its sampling procedure, and its suitability to my research question. Second, I describe how I constructed my sample, the reasoning behind each restriction, and the resulting sample size.

2.3.1 The Sample of Integrated Employer-Employee Data

I use the Sample of Integrated Employer-Employee Data (SIEED7518), a random 1.5% sample of all establishments in Germany. The establishment identifiers are fixed by industry, ownership, and location at the municipality level; hence, an establishment is not equivalent to a firm in all cases. Nevertheless, I use the terms firms and establishments interchangeably. Employers are obliged to report data on all of their employees subject to social security contributions. Self-employed and civil servants are not covered by the data. At the end of each year, employers report the start and end date of employment, wages, and other occupational, educational, and demographic indicators of all of their workers. Typically, the data is a snapshot of the employment state as of June 30$^{\text{th}}$ of each year. Employers are also obliged to report changes in employment contracts.[51]

For each of these establishments, the entire employment biographies of all employees are included over the observation period 1975-2018 for West Germany and 1992-2018 for East Germany. Hence, the data also include the establishments that did not constitute the random 1.5% of the establishments originally sampled, in case the workers from the establishments originally sampled were ever employed elsewhere. Observing the entire workforce of the sampled establishments is critical for my analyses, because I study substitutability mechanisms behind employment reactions to the raise in retirement age, which requires observing all coworkers of a given establishment. Schmidtlein, Seth, and Vom Berge (2020) describe the data sampling in more detail.

2.3.2 Sample construction for analyses

To construct the final sample for my analysis, I keep only women born in 1951, the control group, i.e., women who were potentially eligible for the women's pathway to early retirement, if they accumulated enough years of social security contributions in later life; and 1952, the treatment group, i.e., women who experienced the rise in the women's ERA. I drop women who were ever employed as miners and sailors (for clarity) because their retirement rules differ from those in other occupations.[52]

To address the issue of parallel spells in the data, which is possible, for example, due to dual earners (employed at several establishments simultaneously), I keep the spells in the randomly

[51] One of the data limitations is the lack of working hours; hence, I am limited to the analyses of only the extensive margin of employment.

[52] The seminal work by Geyer and Welteke (2021) on labor supply responses to the 1999 reform makes a restriction of keeping only women who are eligible for the women's pathway to retirement at the age of 60. I make restrictions that proxy for eligibility, following Lorenz et al. (2018). I do not explicitly make sample restrictions that keep the women eligible for the women's pathway (e.g., 15 years of contributions in total and ten years after 40 years old, etc), because I do not observe the unemployment spells that also contribute to the contribution years. Because unemployment spells still count towards the contributions to social security, not making this restriction results in smaller treatment effects in my sample, compared to that of Geyer and Welteke (2021).

selected 1.5% establishments. If both spells come from randomly sampled establishments, I keep the spells where the worker accumulated more tenure. In cases where the employee works in two randomly selected establishments and has accumulated an equal amount of tenure in each of them, I keep the job with the highest wage. Dropping parallel spells allows me to construct Panel data and study the firm mechanisms for only the establishments to which the dual workers are more attached.

The final data consists of person-age entries (in age-month), where I observe women from the age of 42 (age-month 504) until 66 (age-month 792). The choice of this time frame is driven by the fact that the first affected cohort was 47 years old at the time of the reform announcement in 1999, and in some of my analyses I want to observe employment (1) before the reform announcement, (2) between the reform announcement and its inaction at 60, (3) and workers who continue working beyond both the ERA (60 or at least 63) and NRA (65 or 65.5). First, studying employment before the reform announcement shows whether the treatment and control groups had different labor supply frequencies before the reform announcement. Second, studying employment between 47-60 can show whether the rise in ERA leads to different employment choices during middle age, in expectation of a longer employment period. Finally, studying the effects beyond the new ERA shows how the effect of raising ERA also spills over to post-ERA employment, which could show indirect employment effects beyond the age targeted by the reform, further increasing its effectiveness in keeping workers in employment longer.

I keep workers who are continuously (in each age-month) employed at 58 and 59. To make such a restriction plausible, I have to assume that the employment at 58-59 is unaffected by the reform. Geyer and Welteke (2021) show that there are no employment effects before the age of 60; therefore, such a restriction is not likely to lead to a selection bias. Because most of the main heterogeneity variables are constructed at the establishment level, this restriction helps me to obtain a sample of workers with sufficient attachment to their establishments. The final data consists of 32,770 workers, and 9,036,582 worker-age months (Table 2.B.1 records the number of workers after each restriction). Out of these workers, 15,640 are in the control group (born in 1951), and 17,130 are in the treatment group (born in 1952).

2.4 Identification

First, I describe the identification strategy based on reform discontinuity in birth dates, and then I provide some descriptive results that confirm the presence of discontinuity in the data.

2.4.1 Regression discontinuity design

I follow Geyer and Welteke (2021) to locally identify the effect of the reform that raised the ERA on employment, τ_m, in an RDD framework[53]:

$$\begin{aligned} y_{im} &= \alpha_m + \tau_m \mathbb{1}\{b_i \geq b^*\} + \\ &+ \beta_{0m}\mathbb{1}\{b_i < b^*\}(b_i - b^*) + \beta_{1m}\mathbb{1}\{b_i \geq b^*\}(b_i - b^*) + X_i'\beta_m + \epsilon_{im} \end{aligned} \tag{22}$$

where y_{im}- is employment state, recorded for each woman i at every age-months m; b_i is the birth cohort of the individual i; $\mathbb{1}\{b_i \geq b^*\}$ is an indicator showing that i was born after the cutoff b^* (January 1952), i.e., experienced the rise in the ERA (treatment group); while $\mathbb{1}\{b_i < b^*\}$ includes the individuals who are below the cutoff (control group). I use a local linear regression, and by interacting the running variable $(b_i - b^*)$ with the treatment indicator, I allow for different slopes in treatment and control groups. Figure 2.1 shows that a linear trend in the running variable is a plausible assumption, and there is a clear discontinuity that is unlikely to be attributed to a wrong functional form of polynomials. To compute the RDD estimates, I use a triangular kernel function and the optimal bandwidth choice based on mean square error (Imbens and Kalyanaraman, 2012). As a result, I calculate the bias-corrected RDD estimates with a robust variance estimator.

I also control for calendar month, a dummy for Western German residence, wages at the age of 46, and two education categories (out of 3), because previous literature confirms that education is an important determinant of employment at an older age (Geyer et al., 2022). I cluster the standard errors at the birth month level to account for the potential correlation of standard errors ϵ_{im} for the women belonging to the same birth cohort.[54] In robustness and sensitivity checks, I re-run the regressions, altering all the specification parameters- the procedures for estimating the parameters and covariance matrices, polynomial order, kernel weights, bandwidth choice, included covariates, and clustering level.

The baseline regressions pool the 60-62 age (720-756 age months) together, because this is the age frame that was affected by the ERA reform. This identification results in a local average treatment effect of higher ERA on employment outcomes at ages 60-62 (coefficient τ_m in equation Equation 22).[55]

[53]There are several differences from the identification in Geyer and Welteke (2021). First, I do not control for the presence of children in my RDD regression as I do not observe such variables in the data. Second, because the most recent year observed in my data is 2018, the data allow me to pool all the age months corresponding to 60-62 years of age in the baseline regression and beyond 63 in the supplementary analyses, while Geyer and Welteke (2021) pooled only 60-62 due to their right-censored data. Finally, I use the mean square-based optimal bandwidth, while they use a 12-month ad-hoc bandwidth selection procedure.

[54]Clustering at the level of birth dates aligns with literature suggesting clustering the standard errors at the treatment level.

[55]Because I cannot claim that all the women included in my sample were eligible for women's pathway to early retirement, the coefficient could also capture the Intention-to-Treat (ITT) effect. However, Lorenz et al. (2018) show which sample restrictions are likely to lead to eligibility imputations, and because most of my restrictions match their proposed restrictions, my sample likely captures most of women eligible for the

Identification assumptions. This identification relies on two main assumptions.

(1) Smoothness in density. This assumption requires continuity of the running variable (birth cohort) around the cutoff, which eliminates the possibility of strategic bunching (manipulation of the treatment status) at the cutoff. This assumption holds by construction because it is impossible to change one's own birth date.[56] Nevertheless, in the sensitivity tests, I re-estimate the main regressions by omitting the observations close to the cutoff and confirm the robustness of the results.

(2) Smoothness in covariates. This assumption requires continuity of the distribution of the observed and unobserved variables around the threshold, showing that the assignment of the treatment around the cutoff is as good as random. Table 2.B.2 shows that there is no sizeable significant discontinuity in pre-determined variables. In particular, I choose a variable showing whether a woman has Western origin (proxied by the place of living according to the first biographical spell) and nationality, as these variables are fixed over time and hence are pre-determined.

Main outcome variables. In terms of outcome variables, at each age month, I create three mutually exclusive main labor market categories - employment, nonemployment, and retirement. I further disentangle the employment into three groups- employees liable to social security, marginal part-time employment, and partial retirement. Nonemployment stands for a gap in the employment age-month spells. I proxy retirement with the last labor market activity of a worker. Figure 2.A.2 displays the evolution of the three main employment states over age by treatment status, i.e., the gap in employment and retirement statuses at 60-62.

In addition to these employment state categories, I also define wages, because I am also interested in wages conditional on employment.[57] Wages are created at the detailed monthly level, and are non-zero only if the worker is employed.

Effect heterogeneity. To study the mechanisms behind these effects, I perform subsample analysis using several categories of variables, which show turnover costs associated with retirement in the next section. Because the research question relates to the labor demand factors influencing employment at ages 60-62, I define these variables at the age of 58, just before the pre-reform retirement age of 60.

women's pathway to early retirement.

[56]One could argue that the reform cohorts could be chosen by policy-makers in a way that violates the assumption, for example, by the cohort of baby-boomers, etc. However, because I compare cohorts born around the cutoff, and the cutoff does not appear in any other reforms, policies, or characteristics (both of these cohorts are typically classified in the baby-boomer generation) that would make the 1951 cohort different from the 1952 cohort, there is no reason to believe that the assumption is likely to be violated.

[57]Although wages are top-coded in the social security data, this data feature is unlikely to constitute an issue for the analyses as women are less likely to cross the threshold for wage censoring.

2.4.2 Descriptive evidence of the discontinuity

The abolishment of women's pathway to early retirement led to a large increase in employment rates at 60-62, as shown in Panel B of Figure 2.1. While overall there is an upward-sloping employment trend at 60-62 over the birth cohorts, there is also a clear discontinuity around the 1952 cohort. Only around 75% of women born in 1946-1951 were employed at 60-62[58]. However, the employment rate jumped to approximately 90 percent starting with the 1952 cohort, the reform cutoff.

Figure 2.A.3 extends the analyses to display employment rates by treatment status at all age months (corresponding to the ages between 42 and 66), and confirms the presence of a discontinuity in employment rates at 60-62 (due to the 1999 reform that I study) and to a smaller magnitude of discontinuity at the ages 65-65.5 (due to the 2007 reform). Estimating the treatment effects of the 2007 reform is beyond the scope of this paper; hence, in the next section, I causally quantify the largest employment discontinuity that happens due to the 1999 reform, i.e., at 60-62.

2.5 Results

In this section, I first focus on the effect of the 1999 reform on employment, confirming the results of prior studies on this reform (Geyer and Welteke, 2021). I show the effects of retirement on employment trajectories before studying the labor demand mechanisms of employment, because I want to provide a general picture of the labor supply behavior overall before zooming in on the total employment mechanisms.

Based on the theoretical framework, I expect that the rise in the ERA should extend employment among affected workers, particularly those whose exit would impose high turnover costs on firms. These costs are likely higher for workers with specific skills or those employed in occupations with limited internal or external substitutes. Therefore, I expect the employment effects of the reform to be stronger for such workers. On wages, the model predicts ambiguous effects depending on workers' outside options and replacement difficulty: lower bargaining power due to the loss of pension eligibility may lead to wage decreases, while high replacement costs for specific or non-substitutable workers could result in wage premiums to incentivize retention.

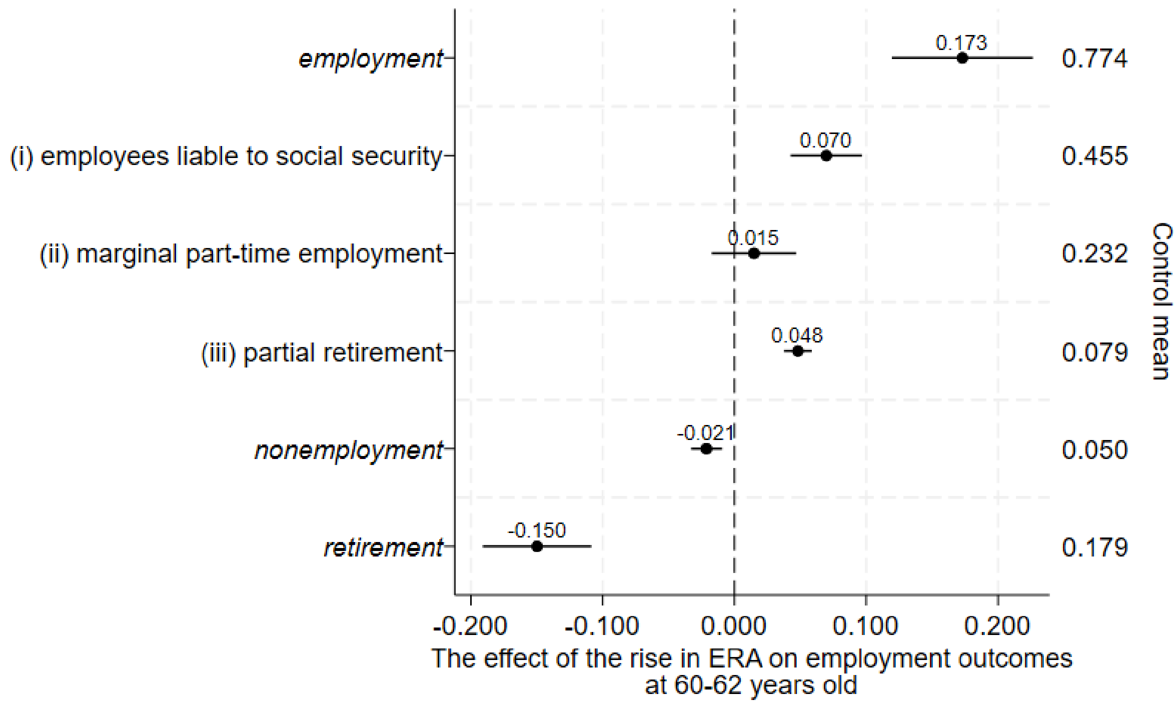
2.5.1 The effect of the rise in ERA on employment states

I start by analyzing how the employment states (employment, non-employment, and retirement) change at the ages targeted by the retirement reform, i.e., at 60-62; hence, I confirm the

[58]This control mean is higher than that in existing literature studying the labor supply response of this reform (Geyer and Welteke, 2021), likely because the sampling of SIEED and my sample restriction (employment at the ages 58-59) results in a sample of workers who are more attached to the labor force.

result of Geyer and Welteke (2021).[59] In this section, I analyze several employment states as outcome variables- (1) *employment* (which is further disentangled into *employment subject to social security*, *marginal part-time employment*, and *partial retirement*), (2) *nonemployment*, and (3) *retirement*[60].

Figure 2.2: The effect of the rise in ERA on the employment state (overall and from each category)



Notes: Coefficient plots. Each row corresponds to the RDD regression of the share of the employment state of the corresponding category (left axis) around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The points represent the estimated coefficients, and the bars represent the 95% confidence intervals. The control means (right column) are the means of the share of employment state in the corresponding category over the control group (born in 1951). A corresponding table with more details can be found in Table 2.B.3.

The right column of Figure 2.2 shows the causal effect of the rise in ERA on employment

[59] I pool the ages 60-62 together because when analyzing the employment effects separately by performing RDD for each age-month in Figure 2.3, there are 2 main periods of significant effects- at the ages of 60-62 and 65-65.5; the rest are either insignificant or very small. The widest gap in employment appears at 60-62, corresponding to the effects of the rise in ERA per the 1999 reform, while the rise at 65-65 years and 6 months corresponds to the 2007 reform's NRA response. Even though the 2007 reform resulted in an NRA rise for the same cohorts under study, the direction of effects is the same and is unlikely to cause any threat to the identification of the 1999 reform under study. Analyzing the effects of the 2007 reform is beyond the scope of this paper.

[60] See subsection 2.4 for more details about these variables

statuses at 60-62. I find that 77.4% of women in the control group (born in 1952) are employed at 60-62, while 5% are non-employed, and 17.9% are already retired. Higher ERA leads to an increased likelihood of being employed at 60-62 by 17.3 percentage points (pp) ($p < 0.01$; a 22.4% increase relative to the control mean).[61] Although most of the increase in employment is attributed to employment subject to social security, i.e., 7 p.p. ($p < 0.01$; a 15.4% increase relative to the control mean), there is also some evidence for an increase in partial retirement claims by 4.8 p.p. ($p < 0.01$; a 60.8% increase relative to the control mean). Hence, the employees respond at the extensive margin, but not necessarily the intensive margin of employment.

Figure 2.2 shows that the likelihood to retire at 60-62 falls by 15 p.p. ($p < 0.01$; an 83.8% decrease relative to the control mean), and there is a negative effect on nonemployment: 2.1 p.p. ($p < 0.01$; 42% decrease relative to the control mean). Overall, these results show that workers are likely to work longer in response to the reform. In the next subsection, I confirm that the results presented are robust to specification and have a credible specification.

Employment beyond 63. Figure 2.3 displays the RDD coefficients at each age month. The workers whose ERA rises do not only work until they reach pensionable age, but are also more likely to extend their employment beyond 63 and to bunch at their Normal Retirement Age of 65.5, before the effects fade away at 66.[62]

While the effects of the rise in ERA on employment beyond 62 are smaller than those at 60-62, they are significant. In most of the sections below, I concentrate on employment at 60-62, as this is the main retirement age shift impacted by the reform.

2.5.2 Robustness and sensitivity checks for the baseline RDD results

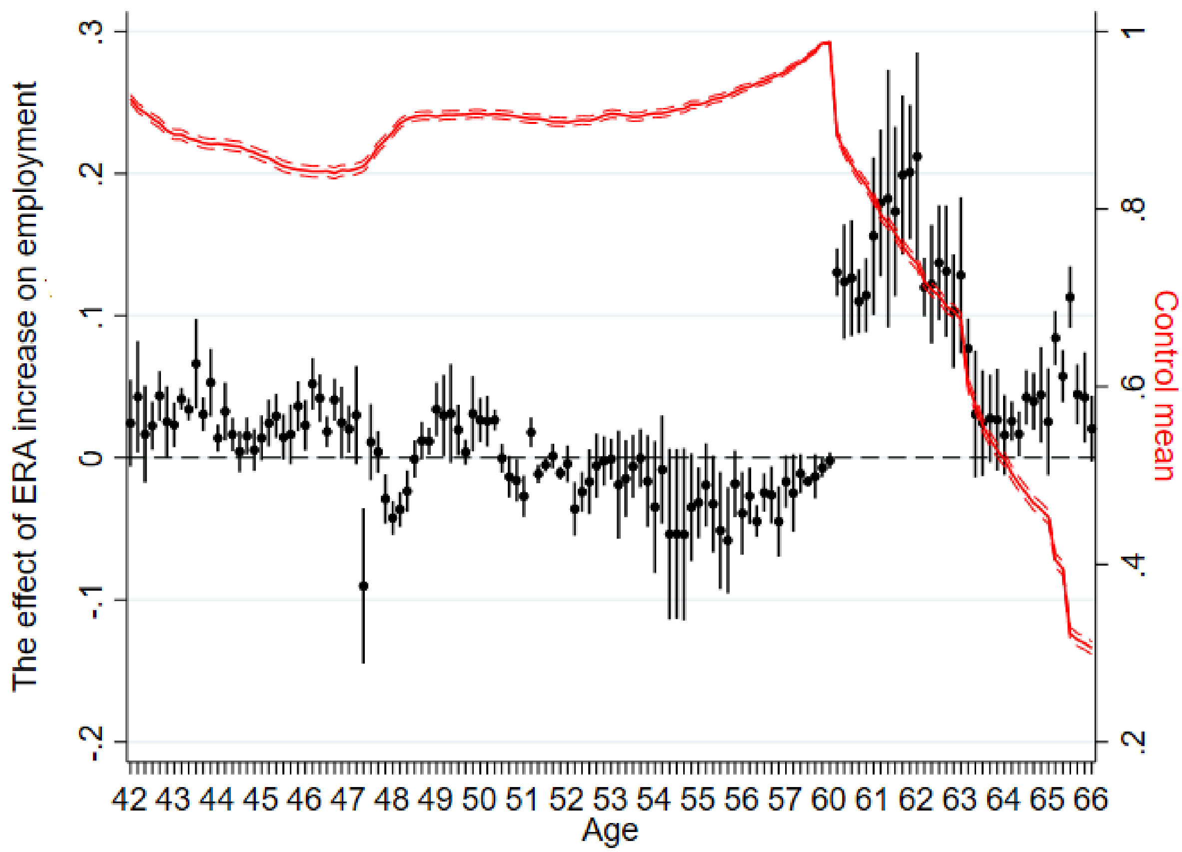
Below, I perform robustness and sensitivity tests that confirm the specified model findings by altering specific parameters of the model. In particular, I alter the estimation procedure for the coefficients or variance estimators, the polynomial order, and specified weights in RDD regressions, and cluster levels of standard errors. I also alter the number of covariates included in the baseline regression and remove the observations close to the cutoff to have a more robust estimation with respect to potential bunching. All the tests indicate that the coefficient estimates presented above are within the confidence intervals of all the alternative models below. Finally, I perform a falsification test by re-estimating the RDD regressions around placebo cutoffs and find no jumps, confirming the validity of the estimation strategy.

Sensitivity to the estimation procedure. Table 2.B.4 shows the sensitivity of estimates with respect to the three different coefficient and variance estimators procedures, and shows that the choice of bias-corrected or conventional coefficient estimates or robust vs. conventional

[61]Figure 2.A.4 zooms in on the *employment* outcome in a regression discontinuity plot, and confirms once more the presence of a discontinuous jump.

[62]The ineligibility of some women for long-insurance pathway could explain the bunching at the NRA, which is the age when workers on the regular pathway to retirement can begin to claim pensions.

Figure 2.3: RDD by age in months



Notes: Coefficient plots. Each vertical line corresponds to the RDD regression of the share of employed at a given age-month. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. The points represent the estimated robust coefficients, and the bars represent the 95% confidence intervals, clustered at the birth month level. The red solid line represents the control mean (with corresponding values displayed on the reversed y-axis), while the red dashed lines represent the confidence intervals for the control means.

variance estimators does not lead to significantly different results.

Sensitivity to the choice of polynomial order. Table 2.B.5 changes the linear regressions to the second-order polynomials. Even the 4th order polynomial choice shows discontinuity in the running variable (the second graph in Figure 2.A.4). Hence, the discontinuity in the running variable is not due to the wrong polynomial choice.

Sensitivity to the specified weights. Table 2.B.6 shows that the estimates with uniform and Epanechnikov kernel function specifications do not significantly differ from the baseline specification that uses triangular weights.

Ad-hoc bandwidths and "donut RD". One of the concerns related to RDD estimation is the potential bunching at the cutoff. To show that bunching would not alter the results, I repeat the estimation and inference without the data points in the area just around the treatment threshold, i.e., the December 1951 and January 1952 birth cohorts, and compare the results to the ad-hoc bandwidth of twelve months.[63] Table 2.B.7 confirms that excluding the observations close to the cutoff does not alter the results of regressions with 12-months bandwidth.

Sensitivity to the inclusion of covariates. Table 2.B.8 reports an RDD regression (1) controlling for Western German origin and nationality in addition to the covariates in baseline specification (calendar months, western residence and education dummies), and (2) the specification with no control variables at all. I do not have enough evidence to argue that the specification is sensitive to the included covariates, as the confidence intervals in all three specifications include the coefficient of the baseline specification.

Sensitivity to clustering level. In an alternative specification, I cluster the standard errors at the establishment level, which captures the main mechanisms discussed later, as opposed to the birth months in the baseline specification, which captures the treatment level. The significance of results does not change from the alternative clustering method, and the confidence intervals generated by this clustering method include the coefficients from the baseline regressions (Table 2.B.9).

Placebo cutoffs. Finally, I perform falsification tests by using placebo cutoffs. I test whether employment at 60-62 rises for women at the other birth cohort cutoffs who were not affected by the reform.[64] I use the cutoffs corresponding to January 1947, January 1948, January 1949, January 1950, and January 1951, as women were eligible for the women's pathway to retirement at these cutoffs. Table 2.B.10 shows that all the placebo cutoffs yield insignificant effects ($p > 0.05$).

2.5.3 The effect of the rise in ERA on wages

While the direct effect of the reform is to delay retirement and extend employment at older ages, its broader implications also depend on the quality of these additional years in the labor force. Wages provide a natural measure of labor market returns, productivity, potential employer valuation, and bargaining power. Studying wages allows me to assess whether not

[63]The "donut RD" does not work in combination with the optimal bandwidth selection procedure due to missing data around the cutoff; hence the necessity to perform such analyses in comparison with the ad-hoc bandwidth.

[64]Table 2.B.11 shows the RDD around the 1952 cutoff for male workers. Although male workers were affected by the 1999 reform to a lesser extent than women (due to the abolishment to early retirement programs for workers on other pathways), they do not constitute an ideal setting placebo group, because if they were on the regular pathway to retirement, their NRA could increase by one month around the cutoff, so at 60-62 they could extend their employment as a forward-looking approach towards retirement after 65 and five months vs. 65 and six months. Still, I report the results, and as expected, there is a discontinuity in employment at 60-62 years old, but it is very small in magnitude.

having an outside option for a pension at 60-62 results in lower bargaining power of workers, and hence, lower wages, as derived in a theoretical model in subsection 2.2. Below, I extend the analyses of the rise in ERA on wages, conditional on employment.

The estimated discontinuity in wages at 60-62 reflects the effect of the reform on those who remained employed at 60–62. The last column of Table 2.B.3 shows that, among those who remain employed, the rise in ERA is associated with 116.522 EUR lower wages ($p < 0.01$; 6.8% decrease relative to the control mean). However, because the reform extended employment, the composition of employed individuals may have changed, introducing selection into the observed wage sample. While the observed wage declines are consistent with reduced bargaining power due to fewer outside options, they may also partially reflect the increased incidence of part-time work or partial retirement among older workers. However, as shown in Figure 2.2, most of the employment responses are driven by employment that is subject to social security rather than other employment types.

2.6 Labor demand mechanisms: replacement costs

After analyzing the overall employment effects above, as a next step, I investigate the labor demand mechanisms of the employment response to the reform through subsample analyses. As outlined in the conceptual framework (subsection 2.2), workers associated with higher turnover costs are predicted to be significantly more likely to remain employed after age 60 following the reform. This suggests that firms retain older employees when replacement is costly, consistent with a labor demand mechanism in which internal firm frictions shape post-reform employment outcomes. Before the reform, older workers with low substitutability could leave the labor force due to generous early retirement options, despite their high firm-specific value. The reform thus can reduce outside options, making retirement less accessible and shifting the relative cost-benefit calculus in favor of retaining less substitutable workers whose departure would impose higher replacement costs on firms.

I create two groups of labor demand measures - worker job-specific skills and market-level worker substitutability, both of which proxy for turnover and replacement costs for the employers. While neither of these groups is preferred over the other, they show different dimensions of substitutability and complement each other for a fuller picture. Worker skills may be firm-specific; hence, with turnover, some information may be lost, making incumbent workers less substitutable by potential new hires. Meanwhile, the internal and external labor market thicknesses show the availability of potential hires. Searching for suitable replacements in the labor market or through internal hiring is costly due to hiring costs.

For the remainder of the paper, I focus only on *employment* as an outcome variable to study the labor demand mechanisms. I show that women who possess high skills and are more difficult to substitute internally (by coworkers) or externally (by external hires) are more likely to extend their employment years in response to the reform, confirming that the reform helps the firms to avoid replacement costs associated with worker turnover.

2.6.1 The role of job-specific skills

The first group of variables showing turnover costs and worker substitutability is worker skills. In the presence of firm- and occupation-specific human capital and knowledge that is difficult to substitute for, turnover can be costly for the establishments. Hence, workers possessing such skills may be more likely to remain employed at an older age. I create two measures: (1) return to occupation (which I interchangeably call human capital specificity of occupation); (2) managerial occupations.

Human capital specificity of occupation. Guvenen et al. (2020) show that wage growth is largely tied to firm and occupation-specific factors, supporting the idea that human capital specificity can shape workers' ability to remain employed after a retirement age increase. If a worker's human capital is very occupation-specific (skills tied closely to their current job/occupation), they are valuable in their current job, and their employer might want to keep them because their replacement would be costly. In terms of the worker's perspective, they may have higher returns to staying and face difficulties in switching occupations if needed. In contrast, if a worker's human capital is more general (easily transferable skills), they can more easily move to other jobs if needed, and employers might replace them more easily followed their retirement.

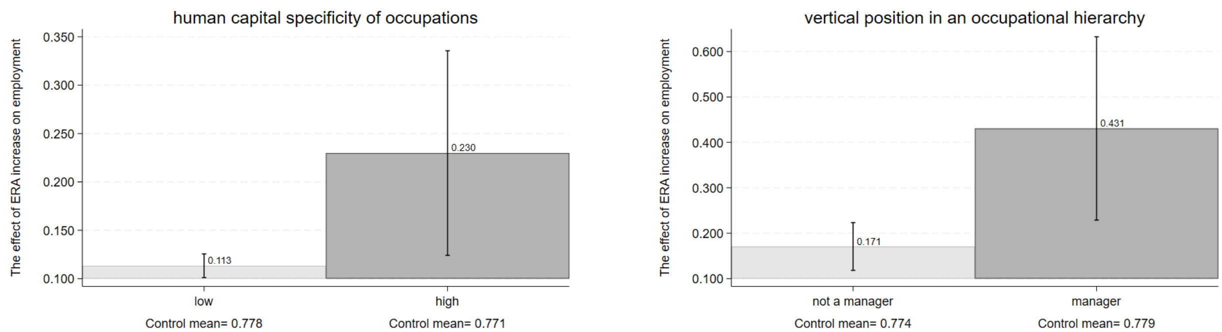
Human capital specificity, proxied by the return to experience, thus constitutes an important mechanism moderating the effects of retirement age reforms on employment outcomes at older ages. To obtain a measure of the human capital specificity of an occupation, I follow a strategy similar to those used by Jäger and Heining (2022) and Bleakley and Lin (2012) to estimate Mincer equations for each of 3-digit occupations.[65] I use the occupation-specific returns to experience, which essentially quantify the impact of on-the-job training and skill accumulation on an individual's wage, and classify the specialization as high if this return is greater than the median value (0.12, i.e., 12% increase in wages associated with an additional year of experience).

I examine treatment effects separately for occupations requiring high levels of specific human capital and those that require less specific human capital to investigate potential heterogeneity. Panel A of Figure 2.4 displays that the workers employed in occupations with above median value of returns to experience have significantly higher employment effects. Among the occupations requiring less specific human capital, the treatment increases employment by 11.3 percentage point (14.5% relative to the control mean of workers who did not experience the rise in ERA and were employed in occupations requiring low human capital specificity). For workers with high human capital specificity, the treatment effect is 23 p.p (29.8% increase relative to a control mean). The difference in point estimates (11.7 p.p.) suggests that the employment response to the treatment is substantially larger for the workers performing occupations requiring specific skills. Thus, human capital specificity (proxied by returns to experience) moderates the effect of retirement age reforms.

[65] Given my smaller sample size, I use only 3-digit occupations as opposed to Jäger and Heining (2022), who use 5-digit occupations.

Figure 2.4: The effect of the rise in ERA on employment at ages 60-62 by return to experience in a given occupation and occupational hierarchy level

Panel A: Human capital specificity **Panel B:** Hierarchical positions



Notes: Coefficient plots for RDD regressions around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and the highest education. The subsample analysis in **Panel A** is performed by the human capital specificity of occupation. **Panel B** stands for managerial status. The vertical lines indicate 95% confidence intervals based on robust standard errors clustered at the birth month level. The control means (on the x-axis) show the employment share at the ages of 60-62 in the corresponding subsample over the control group (born in 1951). A corresponding table with more details can be found in Table 2.B.12.

Managerial status. Managerial occupations often entail a higher degree of firm-specific and occupation-specific human capital, due to their reliance on accumulated institutional knowledge, leadership skills, and relationship-specific investments within the firm. Managers are typically more difficult to replace than are non-managers, particularly at older ages when experience and firm-specific knowledge peak. Therefore, distinguishing between managers and non-managers offers a meaningful way to capture heterogeneity in turnover costs and the value of worker retention following a rise in the early retirement age. Even if two workers have the same returns to experience, managerial roles may imply extra firm-specific value.

In a related study, Jäger and Heining (2022) find that the death of a manager or a worker in a specialized occupation results in more negative effects on the coworkers in other occupations. In my setting, if a worker is a manager, she likely has many coworkers under her hierarchy, communicates with them more, and has more information, making her less substitutable, and thus making the extension of her working life more valuable. I create a variable showing managerial or supervisory status based on the last two digits of the 5-digit occupations. I pool the supervisors and managers into the dummy variable *manager*.[66]

Panel B of Figure 2.4 shows that workers in managerial positions are significantly more likely to remain employed at older ages in reaction to the reform. The workers in managerial positions extend their employment by 43.1 p.p. (55.3% relative to the control mean- the

[66]Depending on occupation type, some occupational hierarchies have managers, while others have supervisors as the highest occupation level in a hierarchy.

managers whose retirement age was not altered by the reform), while the non-managers raise their retirement ages by 17.1 p.p. (22.1% increase relative to the control mean). The difference in point estimates (26 p.p.) suggests that the employment response to the treatment is substantially larger for workers performing managerial occupations.

Alternative measures of skills and specificity. To test whether the results presented above are sensitive to approximating worker skills, I explore alternative proxies for worker skill specificity. The baseline analysis relies on hierarchical job positions as indicators of skill-specific roles. As an alternative, I use an occupational classification by Blossfeld (1985). This classification groups occupations into ten categories and shows the occupational split by required skills- simple vs. professional[67] Figure 2.A.5 shows that across all occupational groups, workers in skilled (i.e., professional) categories exhibit greater employment gains after the reform than those in corresponding simple roles. Managers and professionals are particularly likely to remain employed longer, reinforcing the idea that skills and job specificity drive retention. Although one might suspect that this is driven by longer job tenure, Figure 2.A.6 shows that employment gains do not increase monotonically with tenure. This suggests that hierarchical position captures more than tenure alone.[68]

I further explore whether employment responses differ across occupational task types, offering another dimension of worker substitutability. Following Dengler, Matthes, and Paulus (2014), I categorize jobs along two dimensions: (i) skill content—analytical, interactive, cognitive, or manual tasks—and (ii) routineness—routine vs. non-routine.[69] High-skilled workers typically perform analytical or interactive tasks, while low-skilled workers perform manual non-routine tasks. Panel A of Figure 2.A.7 shows that workers in high-skill task occupations experience the largest post-reform employment extension. In contrast, workers in routine occupations—often more replaceable by automation—do not exhibit systematically different employment responses. These results are in line with the findings of Boockmann, Kroczek, and Laub (2023) for the same reform. They show that employment effects are smaller for workers performing manual non-routine tasks. I next perform a more granular analysis in Panel B of Figure 2.A.7, where I classify occupations by quartiles of task routineness and find higher employment responses across the upper three quartiles. Overall, these results suggest that, in the period studied, task routineness and the potential for substitution by automation play a more limited role in driving employment effects than overall skill specificity.

[67] I use the codes from material published by Schmieder, von Wachter, and Bender (2016) to implement this classification. Education level is not a suitable candidate for skill differentiation in this context, as it is directly controlled for in the baseline specification due to institutional reasons (subsection 2.2 and subsection 2.4). The results by education level are shown in Panel C of Table 2.B.22 and exhibit no meaningful differences.

[68] Tenure is an imperfect proxy for skills in this context because eligibility for retirement at age 63 depends on tenure. Thus, its use conflates eligibility rules with substitutability.

[69] This classification is matched to my main data using the 3-digit occupation identifier. Task types include analytical non-routine, interactive non-routine, cognitive routine, manual routine, and manual non-routine.

2.6.2 The roles of internal and external substitutability

The next group of variables showing the turnover costs and substitutability of workers is based on the markets- internal (by availability of coworkers in the same job cell as an older woman) and external (potential hires in the local labor market). The main motivation for studying internal labor market thickness is that the scarcer the job performed is, the more difficult it is for the employer to replace potential retirees with coworkers, thus leading to higher employment responses to the retirement reform. Internal substitutability is particularly important, as internal workers are imperfect substitutes for external workers (Jäger and Heining, 2022); hence, often the internal substitutes weigh more than the external substitutes. When fewer workers are working in the specific occupation of an older woman in the commuting zone, the less substitutable such a woman is. Similarly, when fewer workers are working in the specific industry of an establishment in local labor markets, the less substitutable the older women of such establishments are by external hires.

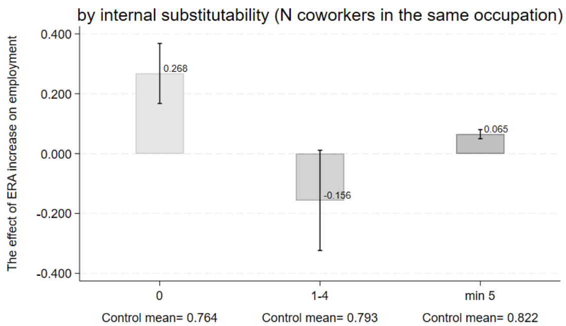
Availability of internal substitutes. To capture internal substitutability, I use the number of available coworkers in the same 3-digit occupation as women born around the reform cutoff. I count only workers in employment positions subject to social security. Following Huebener et al. (2024), I define three categories of such variables by the availability of coworkers in the same 3-digit occupation as the affected women: 0, 1-4, and five or more internal substitutes. I perform the analyses for establishments with fewer than 100 workers, as the levels of substitutability will be less dependent on establishment size (such a restriction also closely follows Huebener et al. (2024) definitions). Panel A of Figure 2.5 shows that when there are no coworkers who perform the same job as the older workers, the older workers are more likely to remain employed at 60-62 following a retirement reform. The group with more than five substitutes has significantly lower employment responses than those with 0 coworkers in the given job cell.

Workers who have no internal replacements respond to the reform by extending their employment by 26.8 p.p. (35% increase relative to the control mean of workers who were allowed to retire at 60 and were employed in non-substitutable establishments). While the effects are insignificant for the group of workers who have between one and four coworkers, the workers who have more than five coworkers in the same occupation extend their employment by 6.5 p.p. (7.9% increase relative to the control mean). The difference in point estimates (20.3 p.p.) suggests that the employment response to the reform that raised the ERA is substantially larger for workers who have no internal substitutes, relative to those who have at least five internal substitutes, in line with the prediction that firms retain workers who are more difficult to replace.[70]

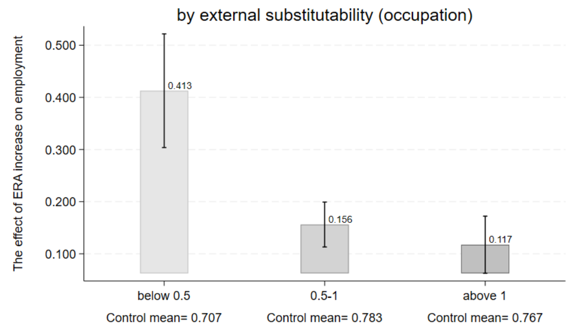
[70]Figure 2.A.9 repeats the analyses for all establishments, regardless of size. When at least five coworkers perform the same job as a woman, the effects are still large, despite being slightly smaller (but not significantly smaller) than those of women with no internal worker substitutes. This pattern could be driven by the variation in treatment effects by establishment size. Indeed, in larger establishments, women are more likely to work longer in reaction to the reform than those in smaller establishments (Figure 2.A.8); hence, when analyzing internal substitutability, it is important to account for the establishment size by restricting the sample to those with at most 100 workers. Even in large establishments, if there are no internal substitutes,

Figure 2.5: Subsample analyses for the effect of the rise in ERA on employment at ages 60-62 by number of internal and external substitutes for the given occupation

Panel A: Internal substitutability in the sample of small establishments



Panel B: External substitutability (occupations)



Notes: Coefficient plots from RDD regressions around the January 1952 cutoff. The estimates are obtained using local linear regressions with first-order polynomials, a triangular kernel, and mean square error–optimal bandwidth selection. Controls include calendar month of birth, Western German residence, wages at age 46, and education. Subsample analyses are conducted by internal substitutability in **Panel A** and external substitutability in **Panel B**. Internal substitutability in **Panel A** is measured by the number of coworkers in the same 3-digit occupation as the old worker, restricting the sample to establishments with fewer than 100 workers. **Panel B** shows *external labor market thickness (ELMT)*, based on the commuting zone at most half as concentrated in a given occupation relative to the country-level ($ELMT < 0.5$), or at least half as concentrated but less concentrated than the country-level ($0.5 < ELMT < 1$), and at least as concentrated as the country-level concentration ($ELMT > 1$). Vertical lines represent 95% confidence intervals based on robust standard errors clustered at the birth-month level. Control means (on the x-axis) refer to the average employment rate at ages 60–62 among the control group within each group's optimal bandwidth. The corresponding detailed tables are reported in Table 2.B.13 and Table 2.B.17.

To address concerns that internal labor market thickness (ILMT) may partly capture differences across occupations, such as comparing workers in occupations with inherently few coworkers (e.g., accountants) to those in larger occupational groups, I control for detailed occupation fixed effects and firm size in Panel B of Figure 2.A.9. The results remain statistically significant and similar in magnitude. This suggests that the estimated effects are not solely driven by cross-occupation differences in skill composition, but instead reflect within-occupation variation in internal substitutability. These findings reinforce the interpretation that internal replacement capacity, rather than solely occupation-specific characteristics per se, is also a key determinant of firms' retention responses.

External labor market thickness (ELMT). I define ELMT in two steps. First, I create 141 local labor markets based on high within-region and low between-region commuting for work, following Kosfeld and Werner (2012). Next, I create an index $ELMT_{kc}$, showing the

the effects are quite large, which highlights that, although in large firms workers stay in employment longer, those who have no substitutes still work longer regardless of the establishment size.

local labor market share of 3-digit occupation (or industry) employment (E_{kc}/E_c) over the national share of occupation (or industry) employment (E_k/E). I count only workers between 18 and 64 years old who are either in employment subject to social security contributions or trainee workers.

$$ELMT_{kc} = \frac{E_{kc}/E_c}{E_k/E} \tag{23}$$

where k is a 3-digit occupation (or industry), and c is a commuting zone, E_{kc} shows the number of workers employed in the occupation (or industry) k, and in the commuting zone c, E_c is the number of workers employed in the commuting zone c and all the occupations (or industries) together, E_k is the number of workers employed in the occupation (or industry) k in all the commuting zones together, while E is the number of workers employed in all the occupations (or industries) and all the commuting zones together (i.e., country).[71]

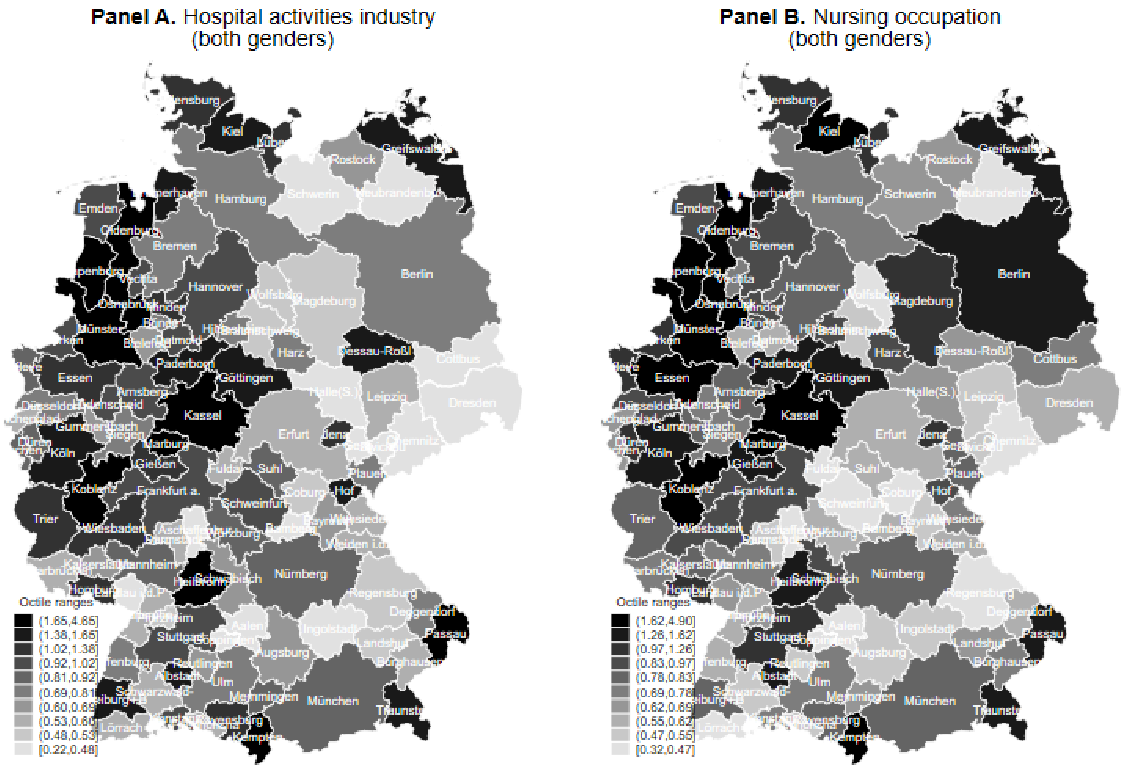
Figure 2.6 displays an example of this index construction for the nursing occupation and hospital activities industry. While Passau has many workers employed in these industries relative to the national level, Leipzig does not. This means that for an establishment located in Leipzig, an older worker in a given occupation and industry is more valuable (i.e., such a worker is associated with higher turnover costs) than for an establishment located in Passau. I call an external labor market thick if this index is over 1, i.e., if the thickness of an occupation (or industry) in a given commuting zone is denser than the thickness at the national level. Additionally, I define a group where the index $ELMT_{kc}$ is below 0.5 (i.e., the commuting zone is at most half as concentrated in a given occupation or industry as the country-level), between 0.5 and one (at least half as concentrated but less concentrated than the country-level).

Panel B of Figure 2.5 displays the RDD results split by external labor market thicknesses of occupations. If women are employed in a commuting zone at most half as concentrated in a given occupation as the country-level, they extend their employment by 41.3 p.p. (58.4% increase relative to the control group). I find that if a woman is employed in a commuting zone at least half as concentrated but less concentrated than the country-level, the employment increase is 15.6 p.p. (19.9% increase relative to the control mean). Finally, in the commuting zones in which a given occupation is more represented than at the national level, the reform leads to an 11.7 p.p. increase in employment at ages 60- 62 (15.3 % increase relative to the control mean). This increase in employment is 29.6 p.p. lower than in commuting zones at most half as concentrated in a given occupation as the country-level. This pattern indicates that employment responses are strongest in thin external labor markets, consistent with lower worker substitutability.

Next, I examine heterogeneity in employment effects along external labor market thickness

[71] All of these variables are defined based on my SIEED data, but because the sample is representative of all German establishments in the country (and the random sampling provides representativeness of workforce subject to social security at the commuting zone level), I expect these indices to proxy the country-level index well.

Figure 2.6: External labor market thickness by German industry and occupation in 2010



Notes: This map shows the computed external labor market thicknesses (ELMT) for each of the 141 local labor markets based on the Kosfeld and Werner (2012) classifications, which are constructed based on high within-region and low between-region commuting. I compute ELMT based on Equation 23 for the industry and occupation with the largest share of female employees: "Hospital activities industry" (**Panel A**) and "Nursing occupation" (**Panel B**). I plot the ELMT indexes (Equation 23) on the map based on the ten deciles presented in the left corner of each graph.

at the industry level (Figure 2.A.10). Unlike the baseline occupation-based results, which showed clear differences by substitutability, I find no significant heterogeneity in responses across industries with different levels of labor market thickness. One potential explanation is that industry-level measures are too broad to capture substitutability for specific skills or tasks. Additionally, larger firms, which are included in the full sample, may be less affected by external labor market conditions because they can rely more on internal replacement options. A likely explanation is that industry-level labor market thickness does not cleanly capture external replacement options, because it is influenced by the presence of large firms. In such settings (e.g., a large employer such as BMW that dominates a local industry, such as the motor vehicle sector within the Munich commuting zone), the measure reflects both external labor supply and firms' internal replacement capacity, attenuating the link between external substitutability and employment responses. This issue is less pronounced at the occupation level, where markets are more granular. As a result, industry-level thickness is a noisier proxy for substitutability, which may explain the weaker heterogeneity.

To account for this concern, I re-estimate the analysis for a subsample of establishments with fewer than 100 employees, where firms are less likely to rely on internal labor markets.[72] In this subsample, the effects of the reform do differ significantly by industry-level labor market thickness: I find that workers are more likely to remain employed in industries in which the external labor markets are thin (Figure 2.A.11). This finding is consistent with external substitutability becoming more salient when internal replacement options are limited. Overall, the occupation-based measure remains more informative, as industry-level thickness may reflect broader agglomeration patterns rather than job-specific substitutability.[73]

Does the external substitutability matter beyond the local level? Tradability of industries. The results above show that workers employed in less substitutable occupations in a given local labor market are more likely to extend their employment in response to the reform. I analyze the broad industry groups and discuss the results in terms of the conventional classification of industries by tradability to test whether the workers in tradable industries are more likely to respond to the raised retirement age. Such analyses allow me to test whether external substitutability matters beyond the local level. In tradable industries, firms can replace workers not only locally but also by outsourcing tasks globally, increasing substitutability (Drenik et al., 2023). I classify the industries by tradability following Gregory, Salomons, and Zierahn (2022).[74] Figure 2.A.12 shows no difference between tradable and non-tradable sectors. The result implies that substitutability does not matter beyond the local level when it comes to the effects of the reform on remaining in employment after 60.[75]

To conclude, I find that job-specific skills and low internal and external substitutability are associated with a stronger increase in employment at ages 60–62 following the reform. While the analysis captures equilibrium effects — that is, match-specific attributes shaped by both

[72] Studies examining worker substitutability in Germany—such as those focusing on coworker death (Jäger and Heining, 2022) and parental leave extensions (Huebener et al., 2024)—typically restrict their samples to small firms. While one approach would be to construct a leave-one-out measure of ELMT, I instead re-estimate the effects using a sample of small firms, thereby aligning the measurement more closely with the existing literature.

[73] In section 2.7, I extend the analysis to gender-specific substitution by constructing ELMT measures based on female workers only and by examining gender dominance of occupations and establishments. I find weaker differences across subgroups. Because the ELMT is constructed from a sample, restricting it to women introduces additional measurement noise. For this reason, I rely on ELMT measures based on the full workforce in the main analysis.

[74] *Tradable industries* are: Mining (WZ08: B); Manufacturing (WZ08: C), Electricity, water supply (WZ08: D, E); Transport, storage (WZ08: H); Financial services (WZ08: K); Real estate (WZ08: L); Agriculture (WZ08: A); Information and communication (WZ08: J); Scientific and technical services (WZ08: M). *Non-tradable* industries are Construction (WZ08: F), Wholesale and retail trade (WZ08: G), Hotels, restaurant (WZ08: I), Public administration (WZ08: O), Education (WZ08: P), Health and social services (WZ08: Q), Cultural, social and personal services (WZ08: R, S), Household-related services (WZ08: T), Other economic services (WZ08: N), Extraterritorial organizations (WZ08: U).

[75] In addition, the generalized categories of industries help me to test whether the external substitutability operates beyond the national level. I define industries by mapping based on the IAB establishment Panel, following the procedure described in Dauth and Eppelsheimer (2020). Figure 2.A.13 does not display significant differences by tradability.

worker and firm — the pronounced retention of managers and specific workers is consistent with higher replacement costs, pointing to an important role for labor demand frictions.

Extensions: bunching at the NRA. As an additional extension, I examine employment and wage responses at ages 64–65, when workers become eligible for full pensions. The results, reported in section 2.7, show a modest increase in employment driven primarily by part-time work, with limited evidence of heterogeneity across substitutability measures. This pattern is consistent with the theoretical framework: when pension eligibility is restored, workers' outside options improve, which weakens the role of substitutability in shaping employment and wage outcomes. However, these estimates should be interpreted with caution, as employment at ages 64–65 reflects a selected sample of workers who remained employed at earlier ages.

2.6.3 The effect of raised ERA on wages by replacement costs

The theoretical framework in subsection 2.2 predicts that raising the early retirement age weakens older workers' outside options, most directly by removing the fallback of early pension access. Such elimination of outside options in the form of pensions reduces wages for affected workers on average. In this section, I analyze whether the effects on wages display heterogeneity by substitutability and job-specific skills.

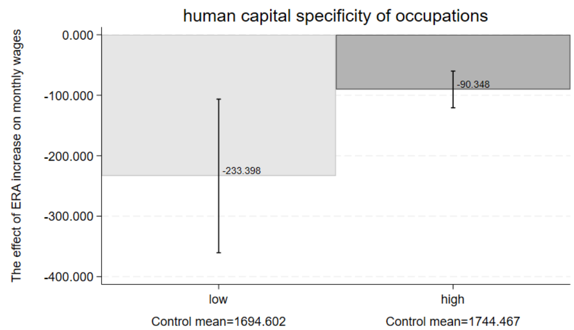
There are two main opposite forces that display heterogeneity. On the one hand, the negative effect might be more pronounced for workers with high specificity (e.g., job-specific skills or high-level managerial roles), because their outside options may be especially limited. On the other hand, if such workers are more productive, firms may have incentives to offer wage premia to retain them, potentially offsetting the negative effect on their wages (see the derivations in subsection 2.2). Hence, the effect of the rise in ERA on wages by substitutability and specificity of skills required to perform the given job may be both positive and negative.

I test this implication by estimating RDD regressions with monthly wages as the outcome, focusing on subsamples that differ in job-specificity and substitutability. Figure 2.7 presents the results. As expected, the overall wage effect is negative, consistent with reduced outside options weakening employee bargaining power, but effects vary across groups. Among the more replaceable workers, wages decline post-reform. In contrast, managers and those in occupations that are difficult to replace externally sometimes experience wage gains after the reform, likely reflecting firms' reluctance to lose strategically important employees.

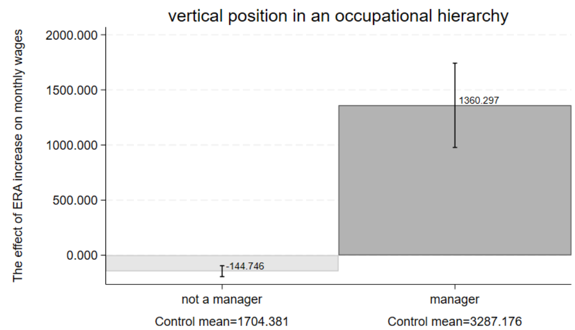
This result may reflect firm retention motives: when specific workers contribute more to firm profits, firms may offer wage premia despite weak outside options. However, selection into employment, whereby only the most productive or critical workers remain, may bias the upward wage effects observed in these groups. Overall, these findings highlight that the wage effects of the retirement reform are shaped by a complex interplay between retention needs and bargaining power, conditional on continued employment.

Figure 2.7: Subsample analyses for the effect of the rise in ERA on wages at ages 60-62 by substitutability measures

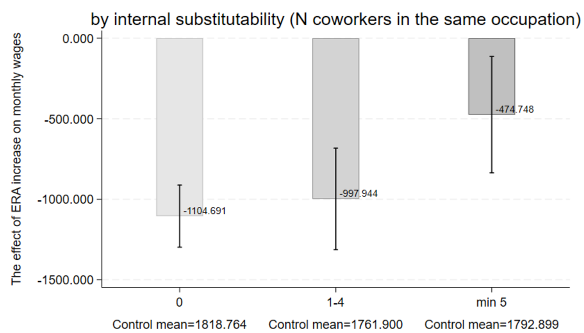
Panel A: Human capital specificity



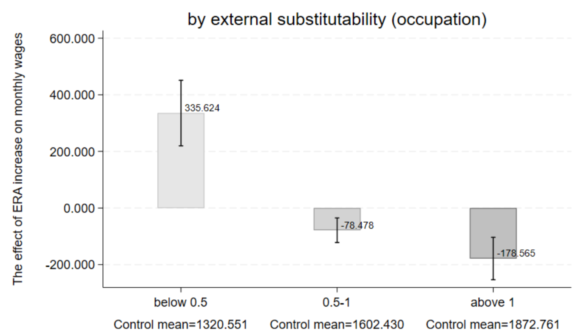
Panel B: Hierarchical positions



Panel C: Internal substitutability in the sample of small establishments



Panel D: External substitutability (occupations)



Notes: Coefficient plots from RDD regressions around the January 1952 cutoff. The estimates are obtained using local linear regressions with first-order polynomials, a triangular kernel, and mean square error–optimal bandwidth selection. Controls include calendar month of birth, Western residence, wages at age 46, and education. The vertical lines represent 95% confidence intervals based on robust standard errors clustered at the birth-month level. Control means (on the x-axis) refer to the average employment rate at ages 60–62 among the control group within each group's optimal bandwidth. The corresponding detailed tables are reported in Table 2.B.23, Table 2.B.24, and Table 2.B.25.

2.7 Conclusion

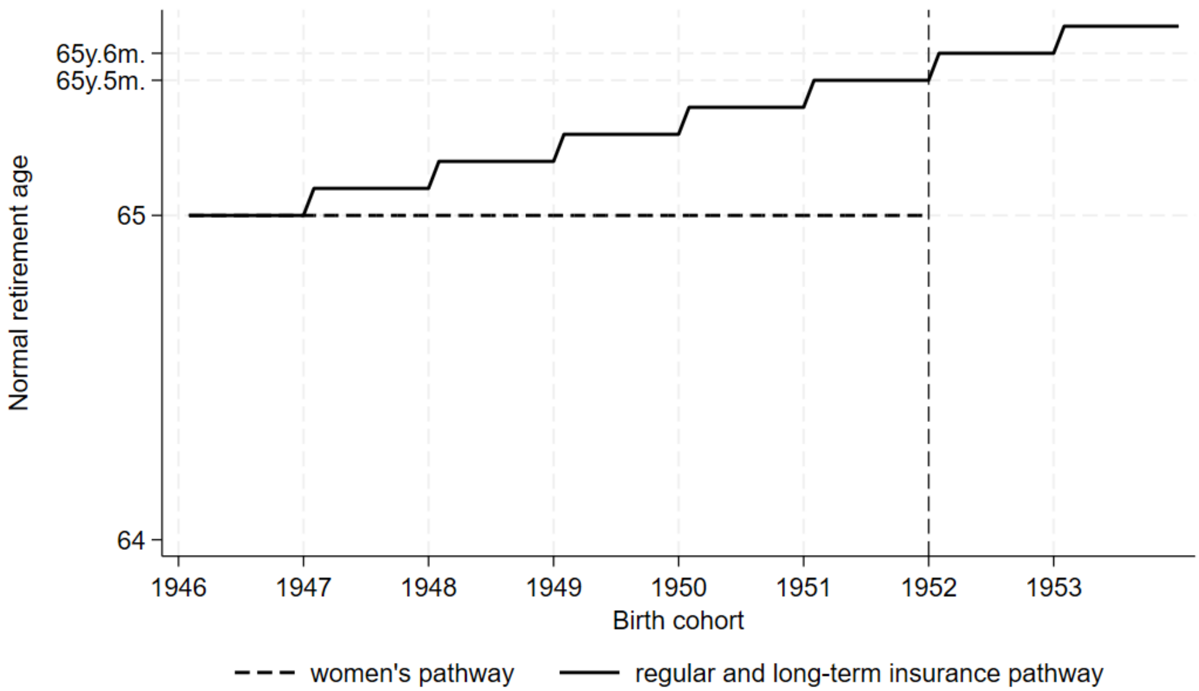
This paper highlights the often-overlooked role of worker substitutability in shaping firm responses to retirement age reforms. While raising the early retirement age extends working lives among older workers on average, this result masks substantial heterogeneity driven by differences in skill specificity to perform a given job and substitutability of a given occupation internally (coworkers in the same occupation) and externally (potential external hires in commuting zone for a given occupation or industry).

The results show that workers in occupations with high skill specificity required to perform a given job and limited substitutability are more likely to be retained post-reform. In these settings, raising the early retirement age reduces staffing frictions and extends working lives. However, these effects are not uniformly beneficial for the workers: less substitutable workers are retained more often, but may experience weaker wage growth due to reduced outside options, while more easily replaced workers face greater employment risk. These findings suggest that retirement age reforms can alleviate staffing constraints in rigid labor markets, but may also enhance firms' wage-setting power, especially when older workers have fewer fallback options. Evaluating such policies thus requires attention to both labor supply and firm-side frictions. These findings also underscore that retirement age reforms affect not only older workers but also the interests of organizational stakeholders—particularly employers and managers—as firms' employment decisions reflect trade-offs between workforce stability, turnover costs, and wage-setting power.

Future research could investigate the exit routes taken by workers who leave employment before the statutory retirement age, including transitions into self-employment, unemployment, or other non-employment states. This is particularly important because selective movement of more substitutable workers into unemployment—potentially encouraged by employers as a bridge to retirement—could represent an alternative mechanism behind the employment patterns observed and thus would pose a threat to the interpretation of the results. While my data do not allow me to observe unemployment spells directly, existing evidence for this reform does not point to substantial substitution into unemployment (Geyer and Welteke, 2021), suggesting that the main findings are unlikely to be driven by such behavior. Nevertheless, understanding these margins remains important for fully disentangling firm retention from potential displacement effects.

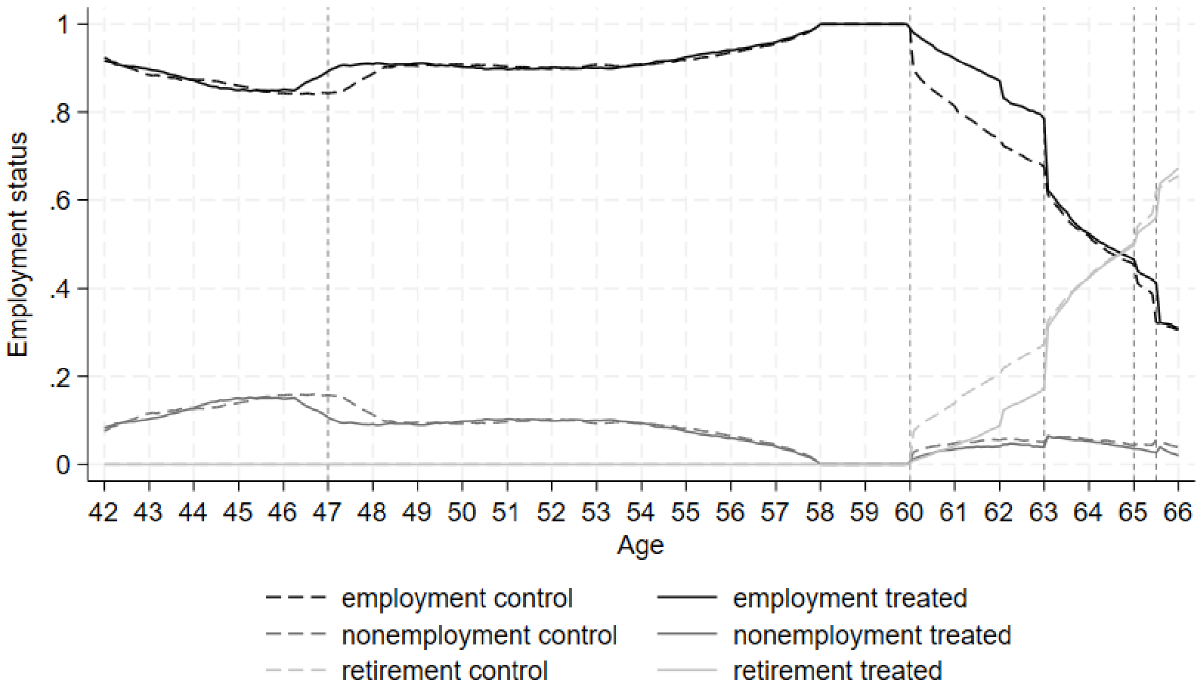
2.A Appendix: Figures

Figure 2.A.1: The assignment of normal retirement age by birth cohorts



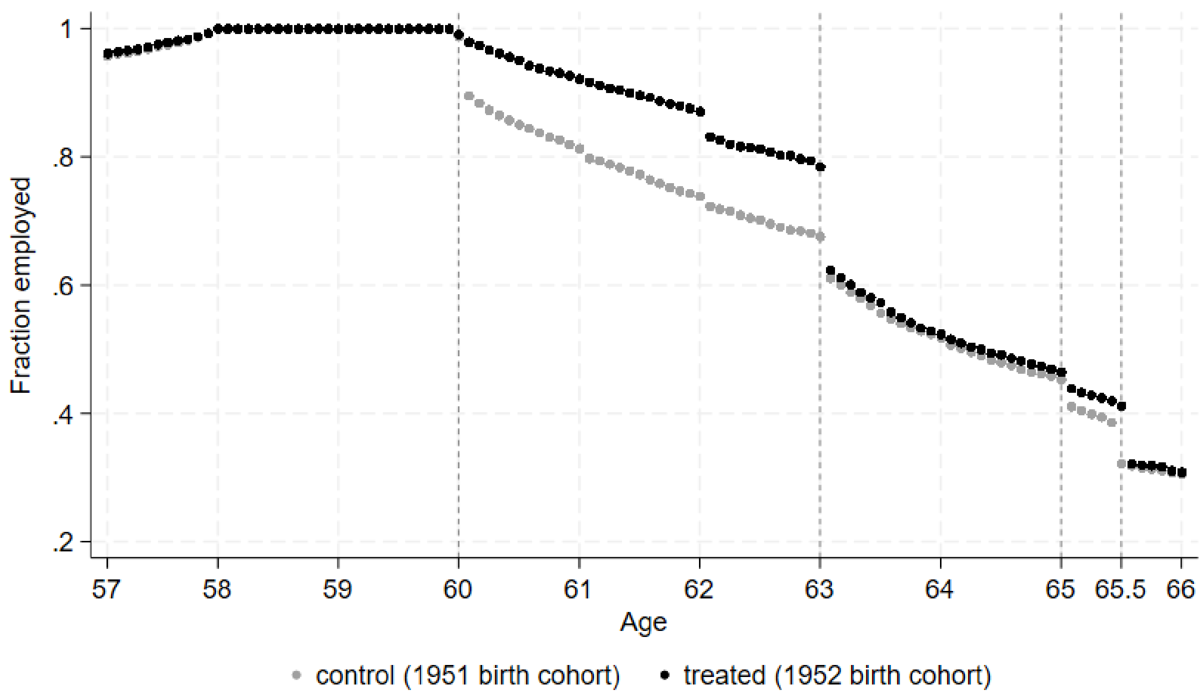
Notes: This figure depicts the assignment rule of normal retirement age by birth cohorts. Before the 1952 cohort, there was a women's pathway to retirement (dashed line). The vertical dashed line at the January 1952 cohort indicates the birth cutoff from which the women's pathway to early retirement was abolished. Starting from the 1952 cohort, the NRA for people eligible for the regular pathway to retirement is equal to the NRA for long-term insured, which used to be 65, but increased by monthly increments per birth year starting from the 1947 cohort (black line).

Figure 2.A.2: Fraction of women employed, nonemployed, and retired at each age-month by treatment and control group



Notes: This figure displays the evolution of three main employment states (*employment* in black, *nonemployment* in dark gray, and *retirement* in light gray- see subsection 2.4 for more details) over age by treatment status: (i) *treated* - women born in 1952 (solid lines), and (ii) *control*- women born in 1951 (dashed lines). The first short-dashed vertical line (at age 47) corresponds to the age of the 1st treated cohort in 1999. The next two short dashed vertical lines show the age frame between the old ERA scheme (at age 60) and the new one (at least age 63) per the 1999 reform, while the last two short-dashed vertical lines show the old NRA scheme (at age 65) and the new one (at age 65 years and six months) per the 2007 reform.

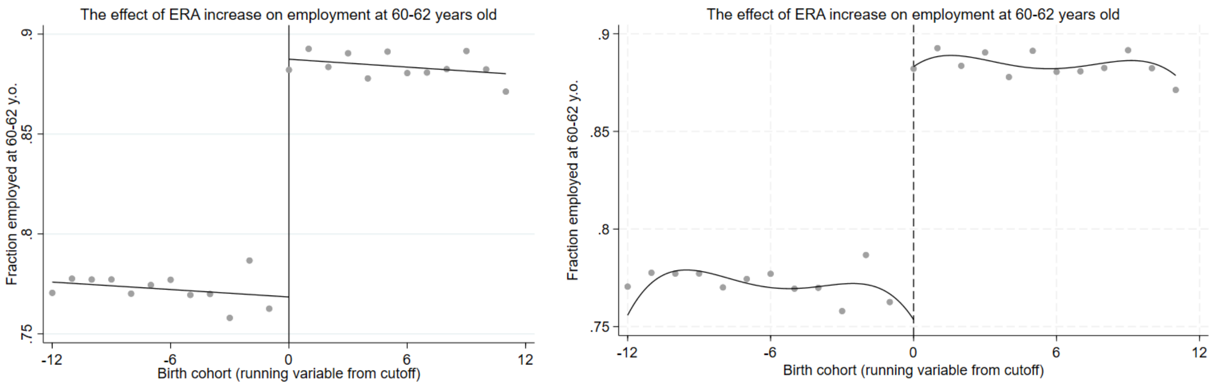
Figure 2.A.3: Fraction of women employed at each age-month by treatment and control group



Notes: This figure displays the fraction of women employed at each age month by two treatment statuses: treated (the 1952 birth cohort, in black) and control (the 1951 birth cohort, in gray). The period between the two dashed lines at 60 and 63 years old indicates the gaps between the two groups due to the 1999 reform under study.

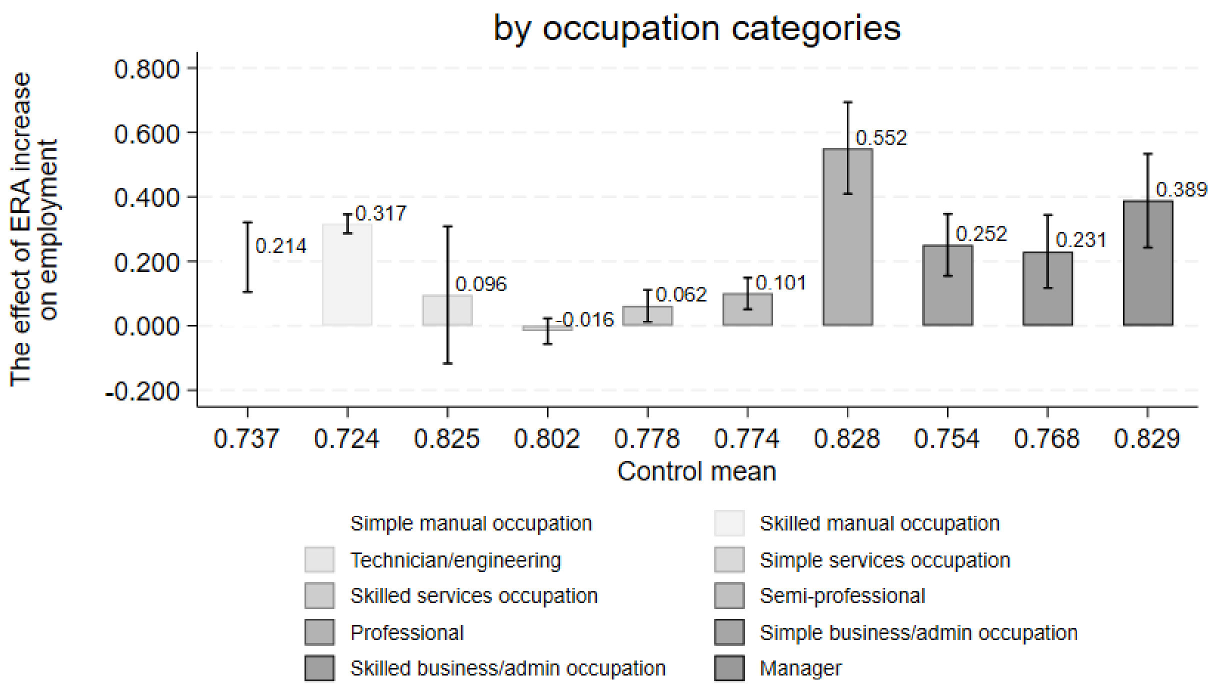
Figure 2.A.4: The effect of the rise in ERA: RDD plot

Panel A: first-order polynomials **Panel B:** fourth-order polynomials



Notes: RDD regression of the share of employed at ages 60-62 around the 1952 cutoff. For computing the RDD estimates, I use first-order polynomials (**Panel A**) or automatic 4^{th} order (**Panel B**), triangular kernel function, and mean square-based optimal bandwidth selection procedure. The vertical line marks the birth cohort threshold 1952 (e.g., 0 corresponds to January 1952, -6 corresponds to people born six months before, in June 1951).

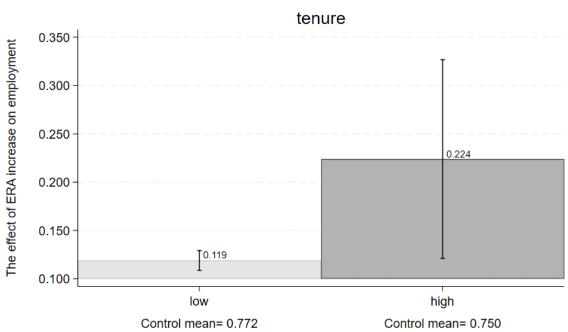
Figure 2.A.5: Subsample analyses for the effect of the rise in ERA on employment at ages 60-62 by aggregate occupations



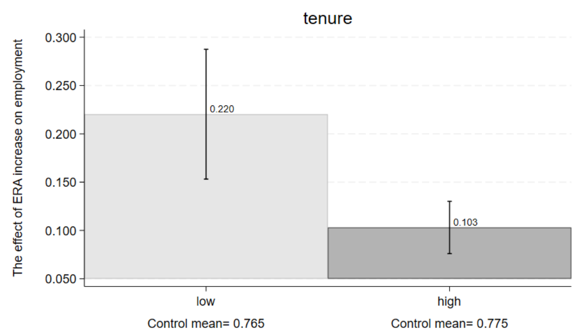
Notes: Coefficient plots for RDD regressions around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. I perform subsample analyses by ten categories of occupations based on occupational classification. The vertical lines indicate 95% confidence intervals based on robust standard errors clustered at the birth month level. The control means (on the x-axis) show the employment share at the ages of 60-62 in the corresponding subsample over the control group (born in 1951). A corresponding table with more details can be found in Table 2.B.15.

Figure 2.A.6: Subsample analyses for the effect of the rise in ERA on employment at ages 60-62 by tenure

Panel A: tenure measured at 46 years old (Me=4.5 years)

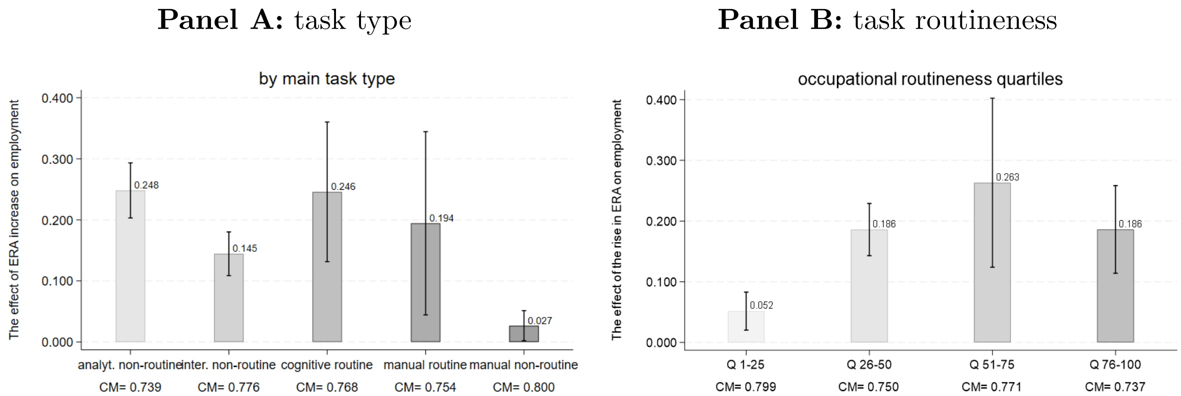


Panel B: tenure measured at 58 years old (Me=7.7 years)



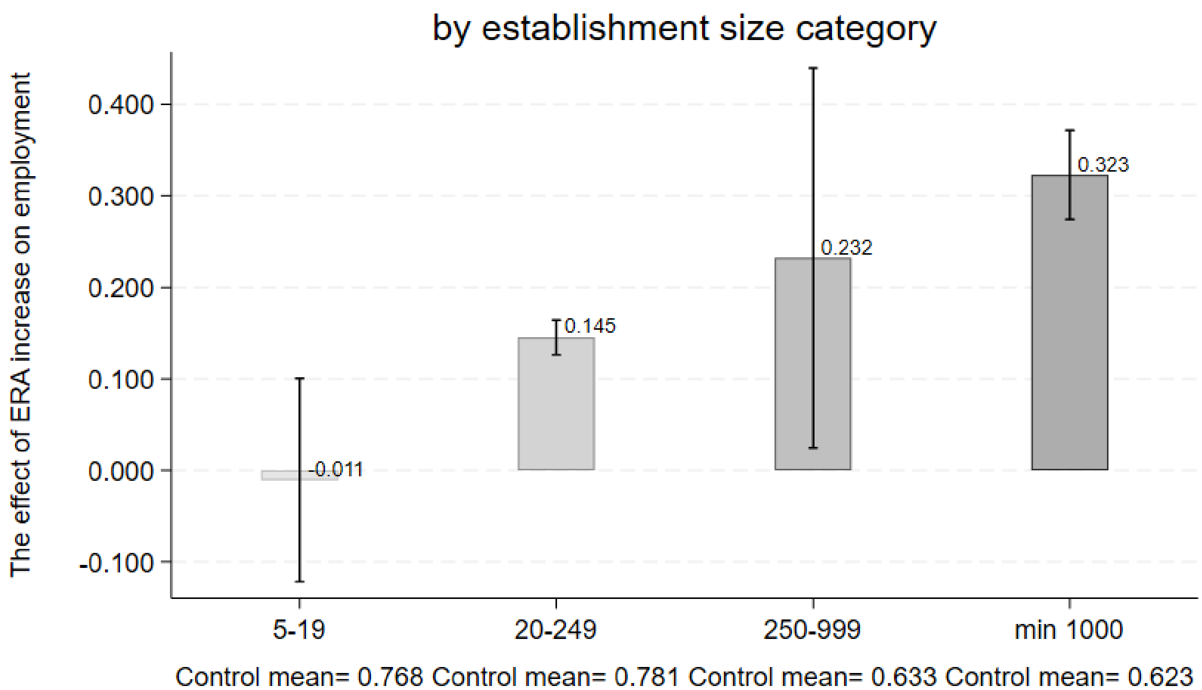
Notes: Coefficient plots for RDD regressions around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. I perform subsample analyses by median split of tenure recorded at 46 years old (**Panel A**), and 58 years old (**Panel B**)- 4.5 and 7.7 years, respectively. The vertical lines indicate 95% confidence intervals based on robust standard errors clustered at the birth month level. The control means (on the x-axis) show the employment share at the ages of 60-62 in the corresponding subsample over the control group (born in 1951). A corresponding table with more details can be found in Table 2.B.14.

Figure 2.A.7: Subsample analyses for the effect of the rise in ERA on employment at ages 60-62 by task type and quartiles of occupational task routineness



Notes: Coefficient plots for RDD regressions around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. I perform subsample analyses by five task-type categories (**Panel A**) and quartiles of task routiness (**Panel B**). The vertical lines indicate 95% confidence intervals based on robust standard errors clustered at the birth month level. The control means (on the x-axis, abbreviated as “CM”) show the employment share at the ages of 60-62 in the corresponding subsample over the control group (born in 1951). A corresponding table for **Panel A** with more details can be found in Table 2.B.18.

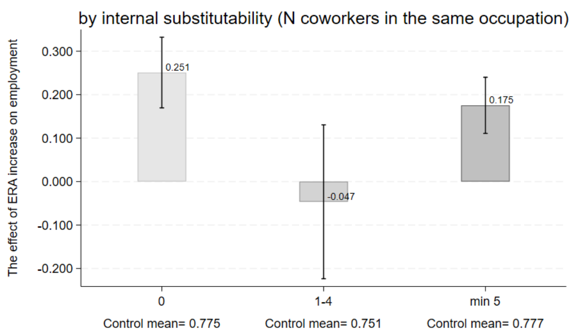
Figure 2.A.8: The effect of the rise in ERA on employment at ages 60-62 by establishment size



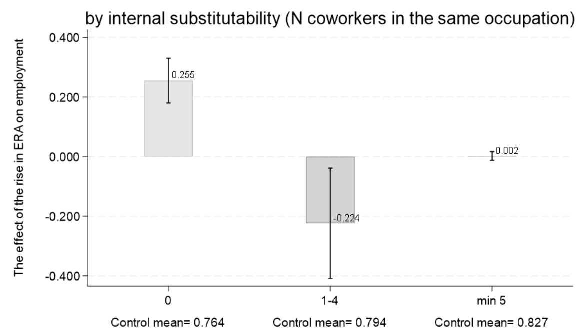
Notes: Coefficient plots for RDD regressions around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. I perform subsample analyses by *establishment size* categories. The three categories of establishment size are (1) up to 19, (2) 20-249, (3) 250-999, and (4) more than 1,000 workers employed at the establishment. The vertical lines indicate 95% confidence intervals based on robust standard errors clustered at the birth month level. The control means (on the x-axis) show the employment share at the ages of 60-62 in the corresponding subsample over the control group (born in 1951). A corresponding table with more details can be found in Table 2.B.16.

Figure 2.A.9: Subsample analyses for the effect of the rise in ERA on employment at ages 60-62 by number of internal substitutes

Panel A: ILMT in firms of all sizes



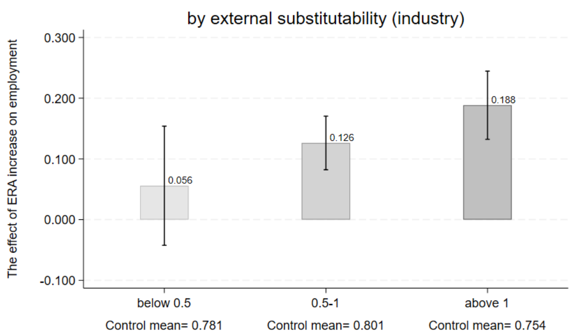
Panel B: ILMT in small firms, controlling for occupation and firm size dummies



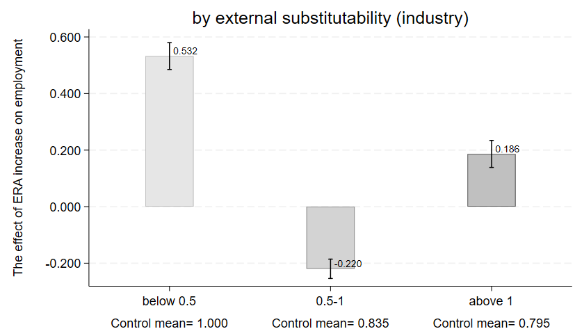
Notes: Coefficient plots for RDD regressions around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. I perform subsample analyses by the number of coworkers in the same 3-digit occupation. **Panel A** includes establishments of all sizes, while **Panel B** restricts the sample to establishments with fewer than 100 workers, and includes occupation and firm size fixed effects. The vertical lines indicate 95% confidence intervals based on robust standard errors clustered at the birth month level. The control means (on the x-axis) show the employment share at the ages of 60-62 in the corresponding subsample over the control group (born in 1951). A corresponding table with more details can be found in Table 2.B.13.

Figure 2.A.10: Subsample analyses for the effect of the rise in ERA on employment at ages 60-62 by external substitutability of a given industry

Panel A: industry-based ELMT, all establishment sizes



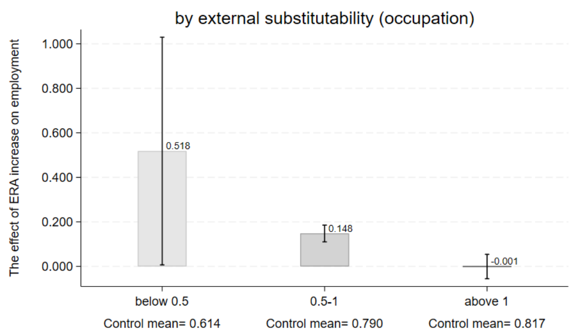
Panel B: industry-based ELMT, small establishments



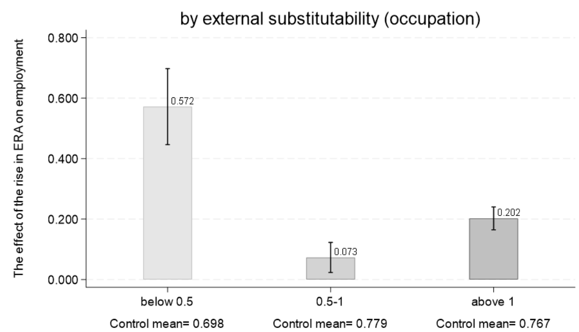
Notes: Coefficient plots from RDD regressions around the January 1952 cutoff. The estimates are obtained using local linear regressions with first-order polynomials, a triangular kernel, and mean square error–optimal bandwidth selection. Controls include calendar month of birth, Western residence, wages at age 46, and education. The *external labor market thickness (ELMT)* is categorized in three groups based on the commuting zone being at most half as concentrated in a given industry as the country-level ($ELMT < 0.5$), or at least half as concentrated but less concentrated than the country-level ($0.5 < ELMT < 1$), and at least as concentrated as the country-level concentration ($ELMT > 1$). **Panel A** displays the subsample analyses in the sample of establishments of all sizes, while **Panel B** is restricted to small establishments. Vertical lines represent 95% confidence intervals based on robust standard errors clustered at the birth-month level. Control means (on the x-axis) refer to the average employment rate at ages 60–62 among the control group within each group's optimal bandwidth. The corresponding detailed table is reported in Table 2.B.17.

Figure 2.A.11: Subsample analyses for the effect of the rise in ERA on employment at ages 60-62 by external substitutability of a given occupation: sample of small establishments vs. all establishments with firm size fixed effects

Panel A: occupation-based ELMT, small establishments



Panel B: occupation-based ELMT, all establishment sizes, with firm size fixed effects

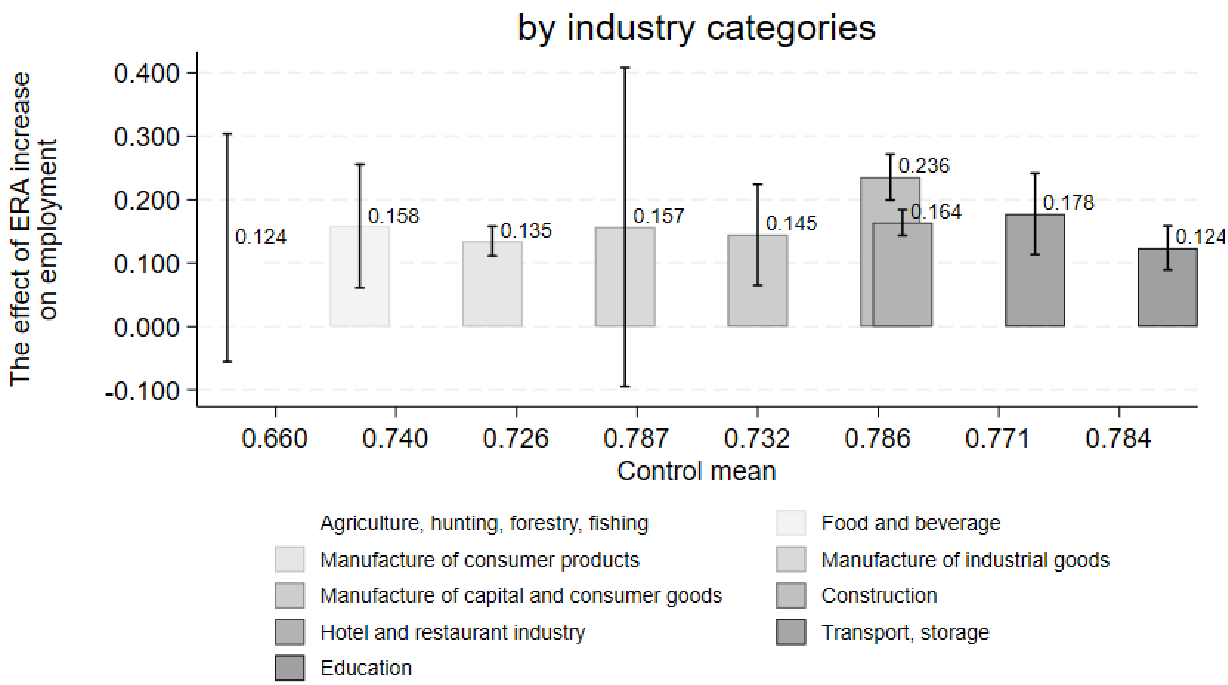


Notes: Coefficient plots for RDD regressions around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. Both Panels show the subsample analyses by *external labor market thickness (ELMT)*, based on the commuting zone being at most half as concentrated in a given occupation (**Panel A**) or industry (**Panel B**) as the country-level ($ELMT < 0.5$), or at least half as concentrated but less concentrated than the country-level ($0.5 < ELMT < 1$), and at least as concentrated as the country-level concentration ($ELMT > 1$). The vertical lines indicate 95% confidence intervals based on robust standard errors clustered at the birth month level. The control means (on the x-axis) show the employment share at the ages of 60-62 in the corresponding subsample over the control group (born in 1951). A corresponding table with more details can be found in Panels C and D of Table 2.B.17.

Figure 2.A.12: Subsample analyses for the effect of the rise in ERA on employment at ages 60-62 by tradability of industries

Notes: Coefficient plots for RDD regressions around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. I perform subsample analyses by tradability of industries. The vertical lines indicate 95% confidence intervals based on robust standard errors clustered at the birth month level. The control means (on the x-axis) show the employment share at the ages of 60-62 in the corresponding subsample over the control group (born in 1951). A corresponding table with more details can be found in Table 2.B.19.

Figure 2.A.13: Subsample analyses for the effect of the rise in ERA on employment at ages 60-62 by aggregate industry categories



Notes: Coefficient plots for RDD regressions around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. I perform subsample analyses by aggregated industry categories. The vertical lines indicate 95% confidence intervals based on robust standard errors clustered at the birth month level. The control means (on the x-axis) show the employment share at the ages of 60-62 in the corresponding subsample over the control group (born in 1951). A corresponding table with more details can be found in Table 2.B.20.

2.B Appendix: Tables

Table 2.B.1: Baseline sample size after each restriction in German social security data

	N women, (birth cohort 1951)	N women, (birth cohort 1952)	N total
unrestricted	**34570**	**36776**	**71346**
delete miners	34562	36771	71333
delete sailors	34560	36768	71328
delete parallel spells	-	-	-
delete age-months below 42 years old	32236	34166	66402
delete age-months above 66	31988	33936	65924
delete repeating age-months	-	-	-
delete if not employed at 58-59	**15640**	**17130**	**32770**

Notes: This table records the sample size after each of the restrictions in German social security data. The first column names the restrictions. The second and third columns list the sample size of treated and control groups, while the last column records the total sample size, i.e., the sum of the two preceding columns.

Table 2.B.2: Balance check. The effect of the rise in ERA on covariates

	(1) West origin	(2) non-German
The rise in ERA	-0.007	0.013***
	(0.009)	(0.005)
Bandwidth	2.8	3.4
Observations	1179720	1179720

Notes: This table shows the effect of the rise in ERA on *Western German origin* (column 1) and *non-German* nationality (column 2) (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA was raised by at least 3 years. I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday. I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. Robust standard errors in parentheses are clustered at the birth-month level.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.3: The effect of the rise in ERA on employment outcomes at 60-62 years old

	(1) employ-ment	(2) employees liable to social security	(3) marginal part-time employment	(4) partial retirement	(5) non-employ-ment	(6) retire-ment	(7) monthly wage
The rise in ERA	0.173***	0.070***	0.015	0.048***	-0.021***	-0.150***	-116.522***
	(0.027)	(0.014)	(0.016)	(0.005)	(0.006)	(0.021)	(23.368)
Bandwidth	2.9	3.9	3.9	4.5	3.0	3.0	3.4
Control mean	0.774	0.455	0.232	0.079	0.050	0.179	1719.644
Observations	1179720	1179720	1179720	1179720	1179720	1179720	980014
N workers	32770	32770	32770	32770	32770	32770	31346

Notes: These tables show the regression discontinuity design estimates around the cutoff of 1952, starting from which ERA rose by at least 3 years (Equation 22). I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). There are 3 mutually exclusive outcome variables: *employment* (column 1), *nonemployment* (column 5), and *retirement* (column 6). *Employment* can be further decomposed into columns 2-4. I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951 (the control group). Robust standard errors in parentheses are clustered at the birth-month level. The corresponding coefficient plot can be found in Figure 2.2.

* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.4: Robustness and sensitivity checks. The effect of the rise in ERA on employment outcomes at 60-62 years old by altering the estimation procedure

	(1) employ-ment	(2) employees liable to social security	(3) marginal part-time employment	(4) partial retirement	(5) non-employ-ment	(6) retire-ment	(7) monthly wage
Panel A: bias-corrected RD estimates with robust variance estimator (baseline)							
Robust	0.173***	0.070***	0.015	0.048***	-0.021***	-0.150***	-116.522***
	(0.027)	(0.014)	(0.016)	(0.005)	(0.006)	(0.021)	(23.368)
Panel B: conventional RD estimates with conventional variance estimator							
Conventional	0.166***	0.078***	0.003	0.051***	-0.020***	-0.144***	-64.181***
	(0.002)	(0.009)	(0.014)	(0.003)	(0.002)	(0.002)	(21.622)
Panel C: bias-corrected RD estimates with conventional variance estimator							
Bias-corrected	0.173***	0.070***	0.015	0.048***	-0.021***	-0.150***	-116.522***
	(0.002)	(0.009)	(0.014)	(0.003)	(0.002)	(0.002)	(21.622)
Bandwidth	2.9	3.9	3.9	4.5	3.0	3.0	3.4
Control mean	0.774	0.455	0.232	0.079	0.050	0.179	1719.644
Observations	1179720	1179720	1179720	1179720	1179720	1179720	980014
N workers	32770	32770	32770	32770	32770	32770	31346

Notes: These tables show the regression discontinuity design estimates around the cutoff of 1952, starting from which ERA rose by at least 3 years (Equation 22). I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). There are 3 mutually exclusive outcome variables: *employment* (column 1), *nonemployment* (column 5), and *retirement* (column 6). *Employment* can be further decomposed into columns 2-4. I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951 (the control group). **Panel A** shows the bias-corrected RD estimates with robust variance estimator, **Panel B** -conventional RD estimates with conventional variance estimator, **Panel C** -bias-corrected RD estimates with conventional bias estimator. Standard errors in parentheses are clustered at the birth-month level.

* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.5: Robustness and sensitivity checks. The effect of the rise in ERA on employment outcomes at 60-62 years old by specified polynomial order

	(1) employ-ment	(2) employees liable to social security	(3) marginal part-time employment	(4) partial retirement	(5) non-employ-ment	(6) retire-ment	(7) monthly wage
Panel A: polynomial function of order 1 (baseline)							
The rise in ERA	0.173***	0.070***	0.015	0.048***	-0.021***	-0.150***	-116.522***
	(0.027)	(0.014)	(0.016)	(0.005)	(0.006)	(0.021)	(23.368)
Bandwidth	2.9	3.9	3.9	4.5	3.0	3.0	3.4
Control mean	0.774	0.455	0.232	0.079	0.050	0.179	1719.644
Panel B: polynomial function of order 2							
The rise in ERA	0.254***	0.063***	0.131***	0.056***	-0.040***	-0.215***	-145.377***
	(0.047)	(0.022)	(0.023)	(0.010)	(0.014)	(0.032)	(33.774)
Bandwidth	3.3	4.6	3.2	4.6	3.4	3.3	4.9
Control mean	0.769	0.458	0.232	0.079	0.050	0.181	1724.441
Observations	1179720	1179720	1179720	1179720	1179720	1179720	980014
N workers	32770	32770	32770	32770	32770	32770	31346

Notes: These tables show the regression discontinuity design estimates around the cutoff of 1952, starting from which ERA rose by at least 3 years (Equation 22). I pool all observations from the month after a worker's 60th birthday to their 63rd birthday (age months corresponding to ages 60–62). There are 3 mutually exclusive outcome variables: *employment* (column 1), *nonemployment* (column 5), and *retirement* (column 6). *Employment* can be further decomposed into columns 2-4. I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I use a first-order polynomial function in **Panel A**, and a second-order polynomial in **Panel B**. I control for calendar month, a dummy for Western residence, wages at 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951 (the control group). Robust standard errors in parentheses are clustered at the birth-month level. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.6: Robustness and sensitivity checks. The effect of the rise in ERA on employment outcomes at 60-62 years old by the specified kernel function

	(1) employ-ment	(2) employees liable to social security	(3) marginal part-time employment	(4) partial retirement	(5) non-employ-ment	(6) retire-ment	(7) monthly wage
Panel A: triangular weights (baseline)							
The rise in ERA	0.173***	0.070***	0.015	0.048***	-0.021***	-0.150***	-116.522***
	(0.027)	(0.014)	(0.016)	(0.005)	(0.006)	(0.021)	(23.368)
Bandwidth	2.9	3.9	3.9	4.5	3.0	3.0	3.4
Control mean	0.774	0.455	0.232	0.079	0.050	0.179	1719.644
Panel B: Epanechnikov kernel							
The rise in ERA	0.171***	0.071***	0.008	0.048***	-0.020***	-0.148***	-99.628***
	(0.029)	(0.016)	(0.017)	(0.006)	(0.006)	(0.022)	(22.389)
Bandwidth	2.9	3.8	4.0	4.3	3.0	3.0	3.5
Control mean	0.774	0.455	0.231	0.079	0.050	0.179	1719.644
Panel C: uniform kernel							
The rise in ERA	0.168***	0.076***	0.002	0.047***	-0.023***	-0.146***	-133.632***
	(0.029)	(0.021)	(0.019)	(0.004)	(0.006)	(0.024)	(25.635)
Bandwidth	2.7	3.3	3.1	2.8	2.7	2.9	2.6
Control mean	0.774	0.455	0.232	0.080	0.047	0.179	1723.688
Observations	1179720	1179720	1179720	1179720	1179720	1179720	980014
N workers	32770	32770	32770	32770	32770	32770	31346

Notes: These tables show the regression discontinuity design estimates around the cutoff of 1952, starting from which ERA rose by at least 3 years (Equation 22). I pool all observations from the month after a worker's 60th birthday to their 63rd birthday (age months corresponding to ages 60–62). There are 3 mutually exclusive outcome variables: *employment* (column 1), *nonemployment* (column 5), and *retirement* (column 6). *Employment* can be further decomposed into columns 2-4. I use a triangular kernel function in **Panel A**, Epanechnikov kernel in **Panel B**, and uniform weights in **Panel C**. I use a mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951 (the control group). Robust standard errors in parentheses are clustered at the birth-month level.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.7: Robustness and sensitivity checks. The effect of the rise in ERA on employment outcomes at 60-62 years old by ad-hoc bandwidth choices

	(1) employ-ment	(2) employees liable to social security	(3) marginal part-time employment	(4) partial retirement	(5) non-employ-ment	(6) retire-ment	(7) monthly wage
Panel A: all the birth cohorts							
The rise in ERA	0.132***	0.091***	-0.010	0.051***	-0.013***	-0.119***	-4.547
	(0.014)	(0.014)	(0.016)	(0.005)	(0.004)	(0.011)	(38.270)
Bandwidth	12.0	12.0	12.0	12.0	12.0	12.0	12.0
Control mean	0.772	0.458	0.228	0.086	0.050	0.178	1744.540
Observations	1179720	1179720	1179720	1179720	1179720	1179720	980014
N workers	32770	32770	32770	32770	32770	32770	31346
Panel B: excluding December 1951 and January 1952 birth cohorts							
The rise in ERA	0.115***	0.123***	-0.055***	0.047***	-0.006	-0.109***	97.799***
	(0.023)	(0.018)	(0.015)	(0.007)	(0.007)	(0.017)	(23.570)
Bandwidth	12.0	12.0	12.0	12.0	12.0	12.0	12.0
Control mean	0.773	0.458	0.229	0.086	0.050	0.177	1744.764
Observations	1077408	1077408	1077408	1077408	1077408	1077408	895417
N workers	29928	29928	29928	29928	29928	29928	28662

Notes: These tables show the regression discontinuity design estimates around the cutoff of 1952, starting from which ERA rose by at least 3 years (Equation 22). I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). There are 3 mutually exclusive outcome variables: *employment* (column 1), *nonemployment* (column 5), and *retirement* (column 6). *Employment* can be further decomposed into columns 2-4. I use a triangular kernel function and a 12-month ad-hoc bandwidth choice. **Panel A** displays the regressions with all cohorts born 1 year before or after the January 1952 cutoff, while **Panel B** removes the observations of women born 1 month around the cutoff. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951 (the control group). Robust standard errors in parentheses are clustered at the birth-month level. The corresponding coefficient plot can be found in Figure 2.2. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.8: Robustness and sensitivity checks. The effect of the rise in ERA on employment outcomes at 60-62 years old by the choice of covariates included

	(1) employment
Panel A: baseline (month dummies, education, and western German residence)	
The rise in ERA	0.173***
	(0.027)
Panel B: additionally controlling for regional origin and foreigner (non-German) status	
The rise in ERA	0.173***
	(0.027)
Panel C: no controls	
The rise in ERA	0.152***
	(0.024)
Bandwidth	2.8
Control mean	0.772
Observations	1179720
N workers	32770

Notes: This table shows the effect of rise in the ERA on *employment* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. In **Panel A**, I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. In **Panel B**, I additionally control for western origin and foreigner status. I have no control variables in **Panel C**. The control means are the average values of the outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth-month level.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.9: Robustness and sensitivity checks. The effect of the rise in ERA on employment outcomes at 60-62 years old by the specified clustering method for standard errors

	(1) employ-ment	(2) employees liable to social security	(3) marginal part-time employment	(4) partial retirement	(5) non-employ-ment	(6) retire-ment	(7) monthly wage
Panel A: clustering at the birth date level (baseline)							
The rise in ERA	0.173***	0.070***	0.015	0.048***	-0.021***	-0.150***	-116.522***
	(0.027)	(0.014)	(0.016)	(0.005)	(0.006)	(0.021)	(23.368)
Bandwidth	2.9	3.9	3.9	4.5	3.0	3.0	3.4
Control mean	0.774	0.455	0.232	0.079	0.050	0.179	1719.644
Observations	1179720	1179720	1179720	1179720	1179720	1179720	1179720
N workers	32770	32770	32770	32770	32770	32770	32770
Panel B: clustering at the establishment level							
The rise in ERA	0.148***	0.070**	0.022	0.051**	-0.017	-0.131***	-136.181
	(0.027)	(0.035)	(0.022)	(0.021)	(0.011)	(0.026)	(100.397)
Bandwidth	3.3	3.2	3.3	3.6	3.3	3.4	2.9
Control mean	0.769	0.455	0.232	0.081	0.050	0.181	1723.688
Observations	1179720	1179720	1179720	1179720	1179720	1179720	980014
N workers	32770	32770	32770	32770	32770	32770	31346

Notes: These tables show the regression discontinuity design estimates around the cutoff of 1952, starting from which ERA rose by at least 3 years (Equation 22). I pool all observations from the month after a worker's 60th birthday to their 63rd birthday (age months corresponding to ages 60–62). There are 3 mutually exclusive outcome variables: *employment* (column 1), *nonemployment* (column 5), and *retirement* (column 6). *Employment* can be further decomposed into columns 2-4. I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western German residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951 (the control group). Robust standard errors in parentheses are clustered at the birth-month level in **Panel A** and establishment level in **Panel B**. The corresponding coefficient plot can be found in Figure 2.2.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.10: Falsification test: RDD on employment at 60-62 years old around placebo cutoffs

	(1) employment
Panel A: 1948 cohort females	
Robust RDD	-0.025
	(0.101)
Bandwidth	4.0
Observations	728892
N workers	20247
Panel B: 1949 cohort females	
Robust RDD	0.004
	(0.784)
Bandwidth	3.7
Observations	853812
N workers	23717
Panel C: 1950 cohort females	
Robust RDD	-0.004
	(0.438)
Bandwidth	3.0
Observations	985104
N workers	27364
Panel D: 1951 cohort females	
Robust RDD	0.021 *
	(0.062)
Bandwidth	3.2
Observations	1083420
N workers	30095

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22). **Panel A** performs RDD for the women born in 1947–1948, around the January 1948 cutoff; **Panel B** - born in 1948–1949, around the January 1949 cutoff; **Panel C** - born in 1949–1950, around the January 1950 cutoff; and **Panel D** - born in 1950–1951, around the January 1951 cutoff. I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. Robust standard errors in parentheses are clustered at the birth month level.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.11: Falsification test: RDD on employment at 60-62 years old around the reform cutoff for males

	(1) employment
Robust RDD	0.051***
	(0.016)
Bandwidth	3.2
Observations	1230624
N workers	34184

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22) for males. The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60$^{\text{th}}$ birthday to their 63$^{\text{rd}}$ birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to men born in 1951. Robust standard errors in parentheses are clustered at the birth month level.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.12: The effect of the rise in ERA on employment at 60-62 years old by measures of worker skills

	employment (1)	(2)
Panel A: human capital specificity of occupations		
	low	high
The rise in ERA	0.113***	0.230***
	(0.006)	(0.054)
Bandwidth	4.8	2.6
Control mean	0.778	0.771
Observations	547164	632340
N workers	15199	17565
Panel B: by hierarchical vertical position		
	not a manager	manager
The rise in ERA	0.171***	0.431***
	(0.027)	(0.103)
Bandwidth	2.9	3.1
Control mean	0.774	0.779
Observations	1165896	13824
N workers	32386	384

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60th birthday to their 63rd birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. **Panel A** is performed by "*HK specificity*"- which stands for human capital specificity of occupation. It is based on the returns to experience in Mincer equations performed separately for each of the 3-digit occupations. Then, I create a dummy variable based on a median split across all occupations. **Panel B** shows managerial status, which is created as a dummy from the last 2 digits of the 5-digit occupational variables. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth-month level. The corresponding coefficient plot can be found in Figure 2.4.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.13: The effect of the rise in ERA on employment at 60-62 years old by internal substitutability (number of coworkers in the same occupation)

	employment		
	(1)	(2)	(3)
	0	1-4	at least 5
Panel A: all the establishment categories			
The rise in ERA	0.251***	-0.047	0.175***
	(0.041)	(0.090)	(0.033)
Bandwidth	4.0	2.3	2.7
Control mean	0.775	0.751	0.777
Observations	53784	39888	1055808
N workers	1494	1108	29328
Panel B: establishments with fewer than 100 workers			
The rise in ERA	0.268***	-0.156*	0.065***
	(0.051)	(0.085)	(0.008)
Bandwidth	3.3	2.7	4.0
Control mean	0.764	0.793	0.822
Observations	22896	24156	56412
N workers	636	671	1567

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I perform subsample analyses by 3 categories of internal substitutes: 0, 1-4, and at least 5 workers. **Panel A** displays the results for all the sizes of establishments, while **Panel B** zooms in on smaller establishments with fewer than 100 workers. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth month level. The corresponding coefficient plot can be found in Figure 2.5.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.14: The effect of the rise in ERA on employment outcomes at 60-62 years old by tenure

	employment	
	(1) low tenure	(2) high tenure
Panel A: at 46 years old (Me=4.5 years)		
The rise in ERA	0.220***	0.103***
	(0.034)	(0.014)
Bandwidth	2.9	3.4
Control mean	0.765	0.775
Observations	600444	579276
N workers	16679	16091
Panel B: at 58 years old (Me=7.7 years)		
The rise in ERA	0.119***	0.224***
	(0.005)	(0.052)
Bandwidth	4.2	2.8
Control mean	0.776	0.745
Observations	511092	503352
N workers	14197	13982

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60th birthday to their 63rd birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I perform subsample analyses by tenure, which is created as a dummy based on a median split across all workers - 4.5 and 7.7 years for the measure created at 46 and 58 years old, respectively. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth month level. The corresponding coefficient plot can be found in Figure 2.A.6 in the Appendix.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.15: The effect of the rise in ERA on employment at 60-62 years old by occupation at 58 y.o.

	employment									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
	Simple manual	Skilled manual	Technician/ engineering	Simple services	Skilled services	Semi-profes-sional	Profes-sional	Simple business /adminis-tration	Skilled business /adminis-tration	M
The rise in ERA	0.214***	0.317***	0.096	-0.016	0.062**	0.101***	0.552***	0.252***	0.231***	0
	(0.055)	(0.015)	(0.108)	(0.020)	(0.025)	(0.025)	(0.073)	(0.049)	(0.058)	(
Bandwidth	4.2	2.8	2.8	3.2	3.3	4.0	3.2	3.0	2.7	
Control mean	0.737	0.724	0.825	0.802	0.778	0.774	0.828	0.754	0.768	
Observations	87228	45576	23796	262908	82476	145692	20592	201744	288792	2
N workers	2423	1266	661	7303	2291	4047	572	5604	8022	

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22) by Blossfield categories. The cutoff is January 1952, star which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages use a triangular kernel function and a mean square error-based optimal bandwidth choice. I perform subsample analyses by 10 categories of occupations based on E occupational classification. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average val outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth month level. he corresponding coefficient pl found in Figure 2.A.5.

* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.16: The effect of the rise in ERA on employment at 60-62 years old by establishment size category

	employment			
	(1)	(2)	(3)	(4)
	small	medium	large	mega large
	$N \in [5; 19]$	$N \in [20; 249]$	$N \in [250; 999]$	$N \in [999, -]$
The rise in ERA	-0.011	0.145***	0.232**	0.323***
	(0.057)	(0.010)	(0.106)	(0.025)
Bandwidth	2.8	4.1	2.9	5.7
Control mean	0.768	0.789	0.633	0.623
Observations	48204	108936	36360	28080
N workers	1339	3026	1010	780

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60th birthday to their 63rd birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I perform subsample analyses by 3 categories of establishment size: small (5–19 workers), medium (20–249 workers), large (250-999 workers), and mega large (above 1,000 workers). I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth month level. The corresponding coefficient plot can be found in Figure 2.A.8 in the Appendix.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.17: The effect of the rise in ERA on employment at 60-62 years old by external substitutability measures

	employment		
	(1)	(2)	(3)
	below 0.5	0.5-1	above 1
Panel A: external labor market thickness (industry)			
The rise in ERA	0.056	0.126***	0.188***
	(0.050)	(0.023)	(0.029)
Bandwidth	3.1	2.9	3.0
Control mean	0.781	0.801	0.754
Observations	64836	419904	687348
N workers	1801	11664	19093
Panel B: external labor market thickness (occupation)			
The rise in ERA	0.413***	0.156***	0.117***
	(0.056)	(0.022)	(0.028)
Bandwidth	3.1	2.8	3.8
Control mean	0.707	0.783	0.767
Observations	47808	513396	610632
N workers	1328	14261	16962
Panel C: external labor market thickness (industry) for small firms			
The rise in ERA	0.532***	-0.220***	0.186***
	(0.024)	(0.017)	(0.024)
Bandwidth	2.8	2.9	4.3
Control mean	1.000	0.835	0.795
Observations	3132	33444	73476
N workers	87	929	2041
Panel D: external labor market thickness (occupation) for small firms			
The rise in ERA	0.518**	0.148***	-0.001
	(0.261)	(0.019)	(0.028)
Bandwidth	3.7	3.7	3.4
Control mean	0.614	0.790	0.817
Observations	2232	50688	57132
N workers	62	1408	1587

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. **Panel A** shows subsample analyses by external labor market thickness (ELMT) for a given *occupation*, based on the index taking values below 0.5, 0.5-1, and above 1. **Panel B** shows subsample analyses by ELMT for a given *industry*. **Panel C** and **Panel D** display the same regressions for the establishments with fewer than 100 workers. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth month level. The corresponding coefficient plot can be found in Figure 2.5, Figure 2.A.10 and Figure 2.A.11.

* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.18: The effect of the rise in ERA on employment by task type

	employment				
	(1) analytic non-routine	(2) interactive non-routine	(3) cognitive routine	(4) manual routine	(5) manual non-routine
The rise in ERA	0.248***	0.145***	0.246***	0.194**	0.027**
	(0.023)	(0.018)	(0.058)	(0.077)	(0.013)
Bandwidth	4.2	3.1	2.9	2.9	3.7
Control mean	0.739	0.776	0.768	0.754	0.800
Observations	91152	218952	417384	88416	320724
N workers	2532	6082	11594	2456	8909

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I perform subsample analyses in five task-type categories. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth-month level. The corresponding coefficient plot can be found in Figure 2.A.7.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.19: The effect of the rise in ERA on employment at 60-62 years old by industry tradability

	employment	
	(1)	(2)
	non-tradable	tradable
The rise in ERA	0.127***	0.227***
	(0.015)	(0.057)
Bandwidth	3.1	3.0
Control mean	0.784	0.737
Observations	838044	334044
N workers	23279	9279

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I perform subsample analyses by tradability of sectors. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth-month level. The corresponding coefficient plot can be found in Figure 2.A.12.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.20: The effect of the rise in ERA on employment at 60-62 years old by industry categories

	employment								
	(1) Agriculture, hunting and forestry, fishing	(2) Food and beverage	(3) Manu-facture of consumer products	(4) Manu-facture of industrial goods	(5) Manufacture of capital and consu mer goods	(6) Cons-truc-tion	(7) Hotel and res-taurant	(8) Trans-port, storage	(9) Edu-catio
The rise in ERA	0.124	0.158***	0.135***	0.157	0.145***	0.236***	0.164***	0.178***	0.124**
	(0.092)	(0.050)	(0.012)	(0.128)	(0.041)	(0.018)	(0.010)	(0.033)	(0.018
Bandwidth	3.2	4.3	3.2	2.7	4.8	3.5	3.5	3.4	2.9
Control mean	0.660	0.740	0.726	0.787	0.732	0.786	0.786	0.771	0.784
Observations	17136	34668	30960	41400	44748	21960	279036	252252	42408
N workers	476	963	860	1150	1243	610	7751	7007	1178

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by least 3 years. I pool all observations from the month after a worker's 60$^{\text{th}}$ birthday to their 63$^{\text{rd}}$ birthday (age months corresponding to ages 60–62). I use a triangu kernel function and a mean square error-based optimal bandwidth choice. I perform subsample analyses by industry categories. I control for calendar month, a dummy Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 19 Robust standard errors in parentheses are clustered at the birth-month level. The corresponding coefficient plot can be found in Figure 2.A.13

* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.21: The effect of the rise in ERA on employment at 60-62 years old by gender domination

	employment		
Panel A: gender domination in occupation			
	gender-integrated	female-dominated	male-dominated
The rise in ERA	0.122***	0.288***	0.245***
	(0.018)	(0.027)	(0.029)
Bandwidth	2.8	3.7	4.5
Control mean	0.736	0.778	0.724
Observations	174600	76752	20376
N workers	4850	2132	566
Panel B: gender domination in establishment			
	gender-integrated	female-dominated	male-dominated
The rise in ERA	0.188***	0.207***	0.178**
	(0.015)	(0.025)	(0.072)
Bandwidth	4.4	4.0	4.1
Control mean	0.741	0.782	0.681
Observations	144000	95184	19656
N workers	4000	2644	546

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60$^{\text{th}}$ birthday to their 63$^{\text{rd}}$ birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. The subsample analyses are performed by *gender dominance* of occupations (**Panel A**) and establishments (**Panel B**). *Gender-integrated* occupations and establishments are defined as those in which the proportion of men and women ranges from 21% to 79%. *Gender-dominated* occupations/establishments are those in which the share of one of the genders exceeds 80%. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth-month level. The corresponding coefficient plot can be found in Figure 2.C.2 in the Appendix.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.22: The effect of the rise in ERA on employment at 60-62 years old by demographic characteristics of employees

	employment		
Panel A: residence			
	East	West	
The rise in ERA	0.164***	0.174***	
	(0.039)	(0.024)	
Bandwidth	2.8	3.0	
Observations	228168	949392	
N workers	6338	26372	
Panel B: residence of origin			
	East	West	
The rise in ERA	0.172***	0.171***	
	(0.038)	(0.025)	
Bandwidth	2.9	2.9	
Observations	232776	945756	
N workers	6466	26271	
Panel C: education			
	high school	vocational	university
The rise in ERA	0.165***	0.183***	0.094***
	(0.029)	(0.039)	(0.024)
Bandwidth	3.4	2.8	3.7
Observations	160740	897840	155340
N workers	4545	24940	4315

Notes: This table shows the effect of the rise in ERA on *employment* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. **Panel A** performs subsample analyses by the residence of the workers (dummy variable); **Panel B** divides the workers by Eastern and Western German origin, proxied by the place of residence of the first worker as observed in the employment biography; and **Panel C** divides the sample by educational categories. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. Robust standard errors in parentheses are clustered at the birth month level.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.23: The effect of the rise in ERA on monthly wages at 60-62 years old by measures of worker skills

	monthly wages	
	(1)	(2)
Panel A: human capital specificity of occupations		
	low	high
The rise in ERA	-233.398***	-90.348***
	(64.777)	(15.485)
Bandwidth	2.7	3.9
Control mean	1694.602	1744.467
Observations	458872	520927
N workers	14619	16721
Panel B: by hierarchical position		
	not a manager	manager
The rise in ERA	-144.746***	1360.297***
	(25.361)	(195.102)
Bandwidth	3.4	3.4
Control mean	1704.381	3287.176
Observations	968243	11771
N workers	30976	370

Notes: This table shows the effect of the rise in ERA on *monthly wages* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. **Panel A** is performed by "*HK specificity*"- which stands for human capital specificity of occupation. It is based on the return of experience in Mincer equations performed separately for each of the 3-digit occupations. Then, I create a dummy variable based on a median split across all the occupations. **Panel B** stands for managerial status, which is created as a dummy from the last 2 digits of the 5-digit occupational variables. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth-month level. The corresponding coefficient plot can be found in Panel A and Panel B of Figure 2.7.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.24: The effect of the rise in ERA on monthly wages at 60-62 years old by internal substitutability (number of coworkers in the same occupation)

	monthly wages		
	(1) 0	(2) 1-4	(3) at least 5
Panel A: all the establishment categories			
The rise in ERA	-217.352	-960.250***	-33.783***
	(327.504)	(156.779)	(6.537)
Bandwidth	3.5	3.4	3.6
Control mean	1566.372	1349.531	1754.040
Observations	44454	33085	877485
N workers	1427	1054	28054
Panel B: establishments with fewer than 100 workers			
The rise in ERA	-1104.690***	-997.944***	-474.748**
	(98.434)	(160.777)	(184.439)
Bandwidth	3.7	3.3	2.8
Control mean	1818.764	1761.9	1792.3
Observations	18601	20033	47272
N workers	610	641	1503

Notes: This table shows the effect of the rise in ERA on *monthly wages* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. I perform subsample analyses by 3 categories of internal substitutes: 0, 1-4, and at least 5 workers. The **Panel A** displays the results for all the sizes of establishments, while **Panel B** zooms in on the smaller establishments with fewer than 100 workers. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth month level. The corresponding coefficient plot can be found in Panel C and Panel D of Figure 2.7.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 2.B.25: The effect of the rise in ERA on monthly wages at 60-62 years old by external substitutability measures

	monthly wages		
	(1)	(2)	(3)
	below 0.5	0.5-1	above 1
Panel A: external labor market thickness (industry)			
The rise in ERA	-155.193	-81.246***	-90.073***
	(214.400)	(12.265)	(21.603)
Bandwidth	4.1	4.0	4.1
Control mean	1403.030	1518.215	1903.207
Observations	54534	352647	567063
N workers	1738	11245	18164
Panel B: external labor market thickness (occupation)			
The rise in ERA	335.624***	-78.478***	-178.565***
	(59.037)	(22.142)	(38.253)
Bandwidth	4.1	4.0	3.3
Control mean	1320.551	1602.430	1872.761
Observations	38515	430331	505147
N workers	1264	13728	16148

Notes: This table shows the effect of the rise in ERA on *monthly wages* (RDD regression in Equation 22). The cutoff is January 1952, starting from which ERA rose by at least 3 years. I pool all observations from the month after a worker's 60^{th} birthday to their 63^{rd} birthday (age months corresponding to ages 60–62). I use a triangular kernel function and a mean square error-based optimal bandwidth choice. **Panel A** shows subsample analyses by external labor market thickness (ELMT) for a given *occupation*, based on the index taking values below 0.5, 0.5-1, and above 1. **Panel B** shows subsample analyses by ELMT for a given *industry*. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The control means are the average values of the outcomes when I limit the sample to women born in 1951. Robust standard errors in parentheses are clustered at the birth month level. The corresponding coefficient plot can be found in Panel E and Panel F of Figure 2.7.

* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

2.C Appendix: Extensions: Gender-specific substitutability

Gender-specific substitutability. The main results rely on gender-neutral measures of external labor market thickness (ELMT), pooling employment densities of both men and women. However, Germany exhibits pronounced occupational and industry segregation by gender, and the reform exclusively affected women. If women face limited competition or hiring barriers in male-dominated fields, their effective substitutability may depend on the gender composition within occupations and establishments.[76]

To explore this, I construct a gender-specific version of the ELMT index using only female employment densities (a modification of Equation 23) and re-estimate the main analysis across the previously defined three ELMT categories. As expected, the variation in the female-specific ELMT is smaller, and the results become statistically insignificant (Figure 2.C.1). One possible explanation is that employers do not confine their replacement pool to women and may consider male hires instead. In such cases, a gender-neutral ELMT measure may better reflect the labor supply elasticity firms actually face.

However, this interpretation is not definitive. Relying solely on female data reduces statistical power, and the resulting ELMT measure may be noisier and less correlated with true substitutability. Therefore, I cannot fully assess gender-specific substitutability with precision in this setting.

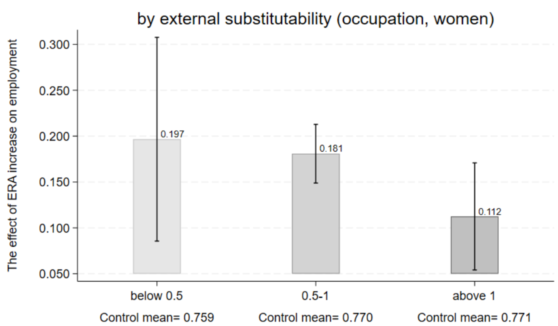
Nonetheless, to test whether gender segregation interacts with employment responses, I perform additional subsample analyses by the gender dominance of occupations and establishments in Figure 2.C.2.[77] I find no significant differences in employment responses between male- vs. female-dominated contexts. One possible interpretation is that, conditional on occupation and firm size, women and men are generally substitutable from the firm's perspective, and the substitutability measures used in the baseline are robust to gender composition.

[76] For example, Illing, Schwank, and Tô (2024) find gender gaps in wages at the hiring stage for vacancies created by worker deaths in Germany.

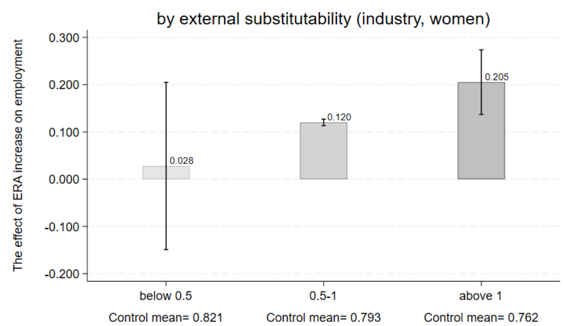
[77] I follow Tophoven et al. (2015) and define gender-integrated occupations or establishments as those in which the proportion of men and women ranges from 21% to 79%. Gender-dominated occupations or establishments are those in which the share of one gender exceeds 80%.

Figure 2.C.1: Subsample analyses for the effect of the rise in ERA on employment at ages 60-62 by external substitutability, constructed for female workers only

Panel A: occupation-based ELMT, women only



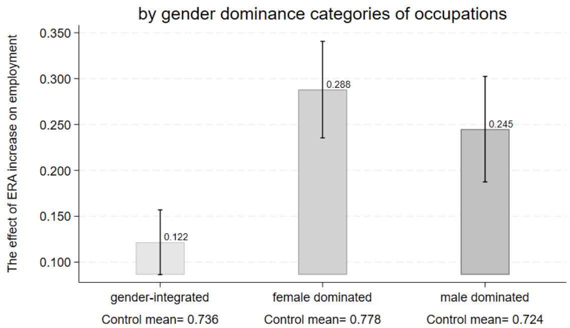
Panel B: industry-based ELMT, women only

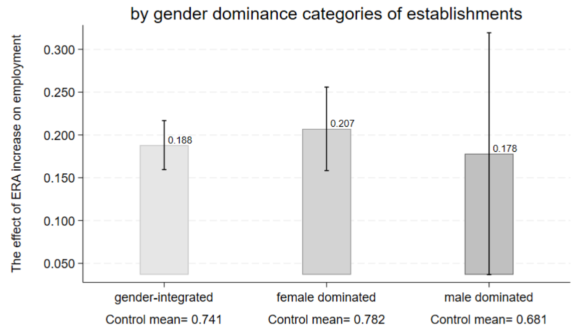


Notes: Coefficient plots for RDD regressions around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. Both Panels show the subsample analyses by *external labor market thickness (ELMT)*, based on the commuting zone being at most half as concentrated in a given occupation (**Panel A**) or industry (**Panel B**) relative to the country-level ($ELMT < 0.5$), or at least half as concentrated but less concentrated than the country-level ($0.5 < ELMT < 1$), and at least as concentrated as the country-level concentration ($ELMT > 1$). The difference from the baseline definitions in the paper is that I use data on women only to construct ELMT. The vertical lines indicate 95% confidence intervals based on robust standard errors clustered at the birth month level. The control means (on the x-axis) show the employment share at the ages of 60-62 in the corresponding subsample over the control group (born in 1951).

Figure 2.C.2: The effect of the rise in ERA on employment at ages 60-62 by gender-composition of occupations and establishments

Panel A: gender dominance of occupations **Panel B:** gender dominance of establishments





Notes: Coefficient plots for RDD regressions around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The subsample analyses are performed by *gender dominance* of occupations (**Panel A**) and establishments (**Panel B**). *Gender-integrated* occupations/establishments are defined as those in which the proportion of men and women ranges from 21% to 79%. *Gender-dominated* occupations/establishments are those in which the share of one of the genders exceeds 80%. The vertical lines indicate 95% confidence intervals based on robust standard errors clustered at the birth month level. The control means (on the x-axis) show the employment share at the ages of 60-62 in the corresponding subsample over the control group (born in 1951). A corresponding table with more details can be found in Table 2.B.21.

2.D Appendix: Extensions: Re-estimating the main results for bunching at the Normal Retirement Age

In this section, I extend the findings to bunching at the normal retirement age (NRA), when workers become eligible for full pensions and their outside options increase. Importantly, employment at ages 64–65 is partly conditional on prior employment at ages 60–62, implying that the estimates may reflect both the direct effects of the reform and selection into continued employment. The results should therefore be interpreted with this caveat in mind.

The effect of the rise in ERA on employment states at 64-65. I perform the same analysis for employment at ages 64–65 to quantify bunching at the NRA. The likelihood of being employed at 64–65 rises by 3.5 percentage points (a 7.3% increase relative to the control group). Unlike the earlier results, most of this employment increase is attributed to marginal part-time employment.

The effect of the rise in ERA on wages at 64-65. When estimating wage effects at ages 64–65, I find no significant results. This finding is consistent with the predictions: at these ages, workers are eligible for pensions—often full pensions, as they coincide with the normal retirement age, which strengthens their outside options and bargaining power. Moreover, since much of the employment at ages 64–65 is driven by marginal part-time work, this pattern is consistent with improved bargaining positions of workers, as they supply less labor while maintaining stronger outside options.

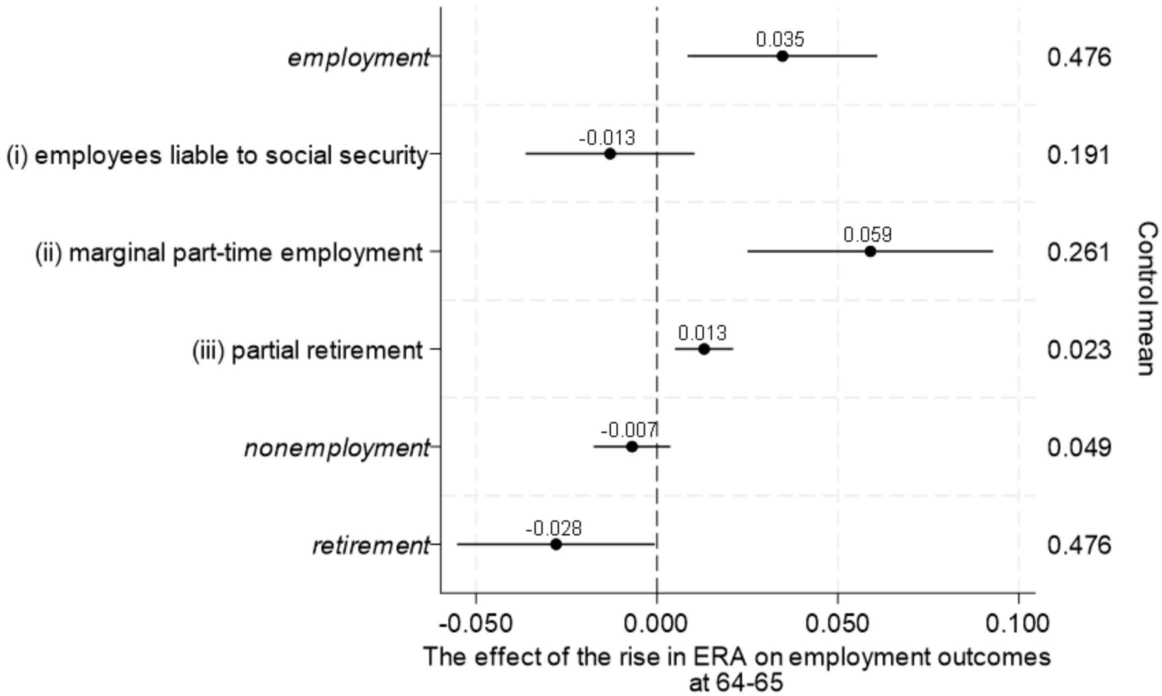
The role of labor demand mechanisms for employment at 64-65. Throughout the main body of the paper, I have focused on employment at ages 60–62. Comparing these patterns to employment responses at ages 64–65 in Figure 2.D.2 reveals an important shift when workers become eligible for (full) pensions. At ages 60–62, the reform operates primarily through a reduction in outside options, making continued employment more likely, especially for workers with high specificity s_i, consistent with Implication 2 in subsection 2.2.

At ages 64–65, however, pension eligibility is restored, increasing outside options, $o(R_i, s_i)$, and weakening this mechanism. In addition, employment at these ages reflects selection from earlier employment decisions, as workers observed at 64–65 are more likely to be those who remained employed at 60–62. This selection attenuates differences across groups and helps explain why heterogeneity by human capital specificity largely disappears at the NRA.

Internal substitutability remains relevant, as firms are less likely to retain workers who can be easily replaced when pension eligibility is restored. Overall, these patterns suggest that, while employment at ages 60–62 is primarily shaped by constrained outside options, employment at ages 64–65 reflects both restored outside options and prior selection into employment.

The effect of raised ERA on wages at 64-65 by replacement costs. Figure 2.D.3 shows the effects of the rise in ERA on wages at ages 64–65 by replacement costs. In contrast to the results at ages 60–62, I find little evidence of systematic heterogeneity across groups. This pattern is consistent with the theoretical framework: at the normal retirement age,

Figure 2.D.1: The effect of the rise in ERA on the employment state (overall and from each category) at ages 64-65



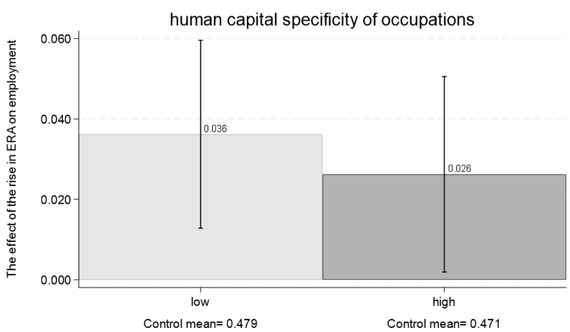
Notes: Coefficient plots. Each row corresponds to the RDD regression of the share of the employment state of the corresponding category (left axis of the graphs) at 64-65 around the 1952 cutoff. For computing the RDD estimates, I use local linear regressions, a triangular kernel function, and mean square error-based optimal bandwidth choice. I control for calendar month, a dummy for Western residence, wages at the age of 46, and education. The points represent the estimated coefficients, and the bars represent the 95% confidence intervals. The control means (right column of the graphs) are the means of the share of employment state in the corresponding category over the control group (born in 1951).

both treated and control workers are eligible for pensions, restoring their outside options and weakening the mechanism through which the reform affected wage bargaining.

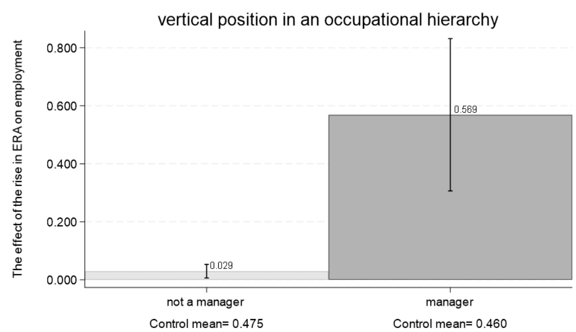
At the same time, wage estimates at ages 64–65 are conditional on continued employment and thus subject to selection, which further limits the scope for detecting heterogeneity across groups. As a result, differences in substitutability and job-specific skills play a less prominent role in shaping observed wage responses, and the heterogeneity observed at earlier ages largely disappears.

Figure 2.D.2: Subsample analyses for the effect of the rise in ERA on employment at age 64-65 by substitutability measures

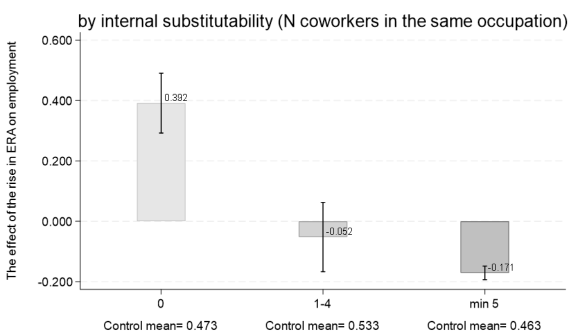
Panel A: Human capital specificity



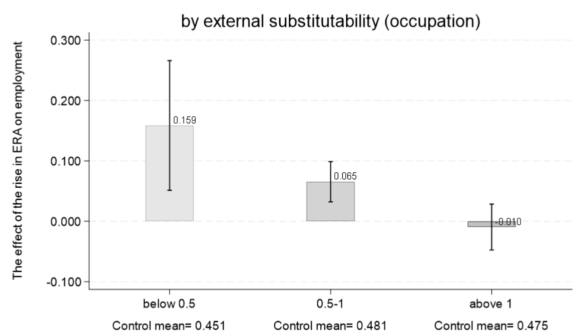
Panel B: Hierarchical positions



Panel C: Internal substitutability in the sample of small establishments



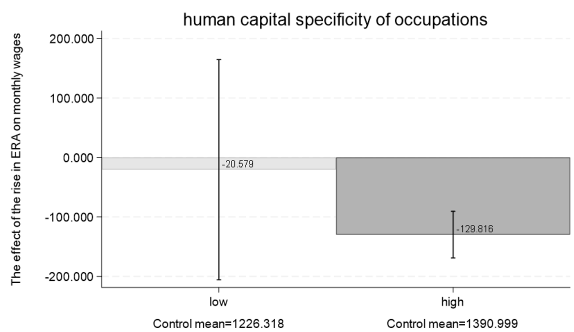
Panel D: External substitutability (occupations)



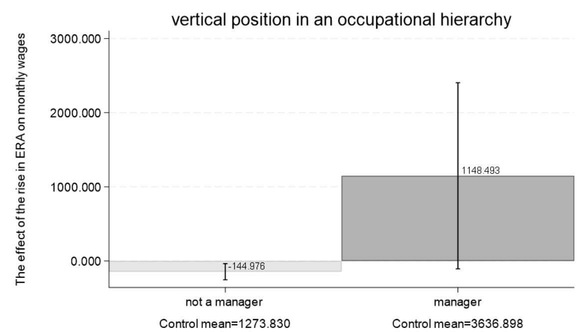
Notes: Coefficient plots from RDD regressions around the January 1952 cutoff. The estimates are obtained using local linear regressions with first-order polynomials, a triangular kernel, and mean square error–optimal bandwidth selection. Controls include calendar month of birth, Western residence, wages at age 46, and education. The vertical lines represent 95% confidence intervals based on robust standard errors clustered at the birth-month level. Control means (on the x-axis) refer to the average employment rate at ages 64-65 among the control group within each group's optimal bandwidth.

Figure 2.D.3: Subsample analyses for the effect of the rise in ERA on wages at age 64-65 by substitutability measures

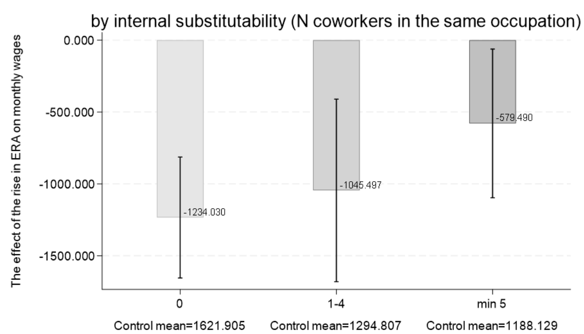
Panel A: Human capital specificity



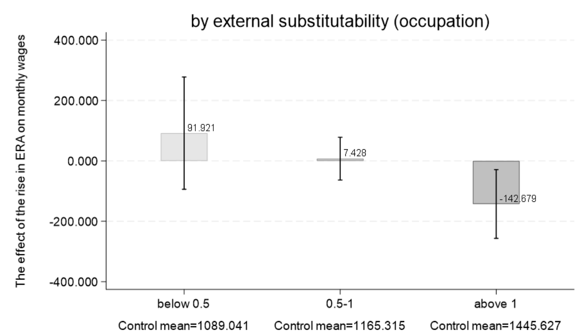
Panel B: Hierarchical positions



Panel C: Internal substitutability in the sample of small establishments



Panel D: External substitutability (occupations)



Notes: Coefficient plots from RDD regressions around the January 1952 cutoff. The estimates are obtained using local linear regressions with first-order polynomials, a triangular kernel, and mean square error–optimal bandwidth selection. Controls include calendar month of birth, Western residence, wages at age 46, and education. The vertical lines represent 95% confidence intervals based on robust standard errors clustered at the birth-month level. Control means (on the x-axis) refer to the average employment rate at ages 64-65 among the control group within each group's optimal bandwidth.

3 Peer Effects in Old-Age Employment Among Women

Single-authored paper.[78]

3.1 Introduction

Gender disparities in labor markets have long attracted policy attention. Firms and institutions have implemented measures such as equal pay legislation, family support policies, and corporate diversity initiatives to narrow these gaps (Blau and Kahn, 2017). Recently, demographic shifts, aging in particular, have led many OECD countries to dramatically raise retirement ages for women by adopting gender-neutral pensionable ages.[79] The workplace is likely to play an important role in shaping the change in gender-specific retirement norms that these reforms aim to achieve (Bramoullé, Djebbari, and Fortin, 2020).[80] However, we still know relatively little about how the interaction of workers within the workplace shapes old-age employment decisions. In this paper, I provide causal evidence of the presence and magnitude of coworker-peer effects in old-age employment among women due to a shift to gender-neutral retirement ages in Germany—a norm-shifting setting in which perceptions of women's employment at older ages were at a relatively formative stage. I examine how workplace peers affect women's employment decisions at older ages by leveraging the quasi-random age composition of worker peer groups before the reform's enactment.

Studying women in the context of peer effects in old-age employment is important for several reasons. First, rising old-age poverty has become a key policy concern (Börsch-Supan and Coile, 2018). Research consistently finds that older women face a persistently higher risk

[78]An earlier version of this paper circulated as a working paper (Badalyan, 2025). I thank Dan Black, Wolfgang Dauth, Johannes Geyer, Štěpán Jurajda, Simon Jäger, Nikolas Mittag, Kerstin Ostermann, and Paolo Zacchia for their feedback; Grayson Krueger from the CERGE-EI Academic Skills Center for language editing. This study uses factually anonymous administrative data from the Integrated Employment Biographies (IEB) of the Institute for Employment Research (IAB). Due to the confidential nature of these data, they can only be accessed and processed on-site at the IAB by authorized staff and guest researchers. The IAB has established procedures to grant data access for replication or verification purposes in cases of reasonable doubt regarding published results. I therefore kindly request an exemption from providing the data in a public online repository. I thank Wolfgang Dauth and his team for their support at the IAB, where this paper was written, and Dan Black for inviting me to the University of Chicago, where part of this research was written or developed, and I received helpful feedback from the U Chicago faculty. This paper benefited from presentations at the IAB Brown Bag and Regio Flash Talks 2025; the CERGE-EI Brown Bag 2025, Applied Student Research Lunch 2024, DW 2022, and DPW 2020. This study was supported by Charles University, GAUK project No. 333221.

[79]For example, Austria plans to raise the retirement age for women from 60 to 65 between 2024 and 2033, while Switzerland aims to achieve full equalization by 2028.

[80]Some survey evidence exists. Using household survey data, Lancee and Radl (2012) show that social connectedness influences the timing of retirement among German workers. In particular, informal social participation, such as gatherings with friends and relatives, is associated with earlier retirement, whereas formal participation in voluntary associations tends to delay retirement in Germany. Vermeer, van Rooij, and van Vuuren (2019) show a positive correlation between preferred retirement within social networks, such as those forced by coworkers.

of poverty than men, due to lower lifetime earnings and reduced pension entitlements from shorter careers and more frequent employment interruptions (Ginn and Arber, 1999; Jefferson, 2009). In particular, Germany is the fifth country in the EU with the highest pension gap among retirees above 65 as of 2018, amounting to 37%.[81] While my analysis does not study pension levels directly, understanding why some women remain employed longer—and how peers and social norms influence this decision—is important for informing policies aimed at reducing old-age economic vulnerability. Second, addressing gender disparities in old-age employment in the labor market remains crucial not only for equity, given women's greater financial vulnerability and higher poverty rates in old age, but also for efficiency, because gender gaps can reflect a misallocation of talent and under-utilization of human capital (Hsieh et al., 2019).

Employment rates at older ages differ markedly even when workers have the same institutions and retirement rules (Krueger and Pischke, 1992), in particular for women (Blau and Goodstein, 2010; Börsch-Supan and Coile, 2018; Gruber and Wise, 2008). Although institutional incentives, such as gender-neutral statutory retirement ages, shape retirement behavior, women's labor market participation at older ages remains limited, partly due to persistent gender norms surrounding work and retirement. These social norms are reinforced by several factors. Retirement is frequently coordinated within couples, and women—who tend to have older spouses—often retire early to synchronize retirement timing for joint leisure (Lalive, Magesan, and Staubli, 2023). Second, women are more likely to shoulder caregiving responsibilities for spouses, grandchildren, or elderly relatives, or engage in home production activities, [82] which can limit their continued attachment to the labor market. Therefore, some women could exit employment before reaching pensionable ages.

A key but often overlooked factor in old-age employment decisions is the role of coworkers in peer effects, which can shape individual retirement choices. Work occupies far more time than any other daily activity; therefore, work attitudes have a central influence on a person's life (Smith, 1965). Coworker peer groups are important to study in the context of employment-related decisions. For example, Meekes and Lent (2025) find that peer effects in working hours are larger in coworker networks than in neighbor and family networks. A possible explanation for understudied co-worker peer effects in the context of old-age employment is the difficulty in detecting peer effects in naturally occurring coworker groups. Difficulties arise because peer groups are not formed exogenously due to three common problems: simultaneity, correlated effect bias, and endogenous group formation (Blume et al., 2011; Manski, 1993). For example, selecting specific occupations and establishments or experiencing common shocks can be wrongly attributed to peer effects. It is also difficult to argue which peer influenced the other. Moreover, the limited data on workgroups, i.e., sufficiently granular occupational structures within establishments that could proxy close worker interactions, further limit the scope of research on coworker peer effects.

To circumvent these problems related to causal inference and data availability, I employ

[81] https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20200207-1

[82] Ciani (2016) shows that women are more likely to engage in home production upon retirement than married men in Italy.

a quasi-experiment in Germany, in which a reform in 1999 raised the early retirement age (hereafter, ERA, i.e., the age at which people can start claiming pensions) of women by at least three years (from 60 to 63) starting from the 1952 birth cohort, and thereby abolishing women's pathway to early retirement. There are several advantages to focusing on this specific quasi-experiment for estimating peer effects. First, the three-year increase in statutory retirement age is the largest increase for two consecutive cohorts that has occurred in recent German public policy and thus is the most suitable reform for detecting peer effects in retirement, as opposed to a step-wise increase in retirement ages. Second, the reform is unique because it abolished the gender-dependent retirement age system, allowing for a shifting norm related to promoting old-age employment among women. This is particularly relevant in Germany—a country with heterogeneous gender norms that is still more inclined towards the breadwinner model than the dual-earner model, in particular among the older generations from West Germany.[83] Finally, Vermeer, van Rooij, and van Vuuren (2019) show that willingness to postpone labor market exit given peer group preferences is higher if the planned retirement age is below the statutory retirement age, motivating the study of peer effects in the context of reforms targeting ERA.

The identifying variation stems from the exposure of coworkers to peer women who were born in a narrow window either before or after the reform cutoff of 1952. By exploiting exogenous variations in retirement eligibility rules and assigning the peer groups before reaching their retirement ages, I prevail over the three problems of peer effect estimation. The universe of detailed German social security data enables the assignment of workers to their workplace peers within job cells based on (1) establishments—single locations of multisite firms—and (2) detailed 4-digit occupational codes. I define *peers* as workers employed in a given establishment who are directly affected by the reform, and *coworkers* as their colleagues who are younger and thus will reach old age after observing their peers' employment decisions, and being influenced by them. I thereby overcome the reflection problem, where the observed coworker outcome may be both a cause and a consequence of peer retirement behavior.

The identification strategy employed in this paper rests on the assumption that establishment-level characteristics are conditionally exogenous to the reform exposure, ensuring that the reform can be used as an exogenous shifter of old-age employment. To further strengthen the identification, I control for a rich set of observable worker, establishment, sector, and regional characteristics, all measured when the peer was 57. I start by estimating the direct effect of the reform on the average employment rate at age 62 among the peers. I find that raising the ERA from 60 to at least 63 leads to 11.6 percentage point (p.p.) higher likelihood of being employed at age 62.[84] Having established the significant effect of the reform on employment at older ages, I turn to estimating the peer effects through two-stage least squares. I find significant peer effects in old-age employment among women. An individual's probability of retiring increases by 1.5 p.p. when their immediate coworkers are ineligible for claiming pensions at 60. Because the first-stage estimate on employment at age 62 is 11.6 p.p., such

[83] Throughout the paper, I refer to East Germany to define New Länder (and additionally Berlin in the data), and to West Germany as the current regions of the former Federal Republic territory.

[84] This direct effect of the reform is slightly smaller than that found in Geyer and Welteke (2021) (13.5 p.p.) and Badalyan (2025) (17.3 p.p.) because their sample of women has higher labor force attachment.

results translate into 13.3 p.p. of peer effects, i.e., a higher likelihood of staying employed at 62 if the peer decides to do so. The results remain robust across various specifications, including the definitions of coworker groups, peer groups, workgroups, and treatment. They also hold after including a comprehensive set of covariates. Placebo tests do not display peer effects, further supporting the credibility of my results.

There can be many reasons why peer effects exist. I analyze whether peer interactions shape individual choices through (1) conformity (Bernheim, 1994) and social norms (Stutzer and Lalive, 2004); (2) information diffusion about career concerns (Dahl, Løken, and Mogstad, 2014; Johnsen, Ku, and Salvanes, 2024; Krstic and Hideg, 2019; Welteke, 2015) or the reform (Nicoletti, Salvanes, and Tominey, 2018), and (3) work complementarities within teams. Importantly, it is difficult to fully disentangle these mechanisms from one another in social security data; hence, I provide suggestive evidence about them by performing subsample analyses.

First, because the reform effectively raised the labor force participation of older women, it could shift social norms regarding women's employment at older ages by changing norms about the appropriate retirement age. Conformity then acts as the behavioral channel through which such norms spread: individuals imitate their peers in order not to deviate from the group. In my setting, the two are therefore difficult to disentangle empirically. The East–West comparison is consistent with this interpretation: in West Germany, where female employment rates are lower and the breadwinner model more prevalent,[85] peer effects are stronger because conformity accelerates the diffusion of new, more egalitarian norms. In East Germany, where higher female labor force attachment is more common, the same conformity mechanism yields weaker incremental effects.

Second, retirement occurs within an institutional context where coworkers exchange knowledge about the reaction to and consequences of retiring at a certain age, such as pension schemes, employer reactions, health implications, wages, and financial preparedness. The behavior of coworkers thus provides information that can reduce uncertainty about the transition to retirement. If this mechanism prevails, the effects should be the largest under higher uncertainty, such as for less tenured coworkers, high turnover, and younger establishments, and if the peer is more informative, for example, a manager.

Finally, novel to the coworker peer effects literature, I test whether work complementarities, such as firm incentives related to team productivity, are crucial in coworker peer effects. Peer effects influence labor supply decisions, potentially creating spillover effects through strategic complementarities within the team across older workers and their younger counterparts, such as collaboration benefits, productivity, and shared workgroup-specific human capital and workload (Bartel et al., 2014; Jaravel, Petkova, and Bell, 2018; Jäger and Heining, 2022). I test this mechanism by proxying interactions by the main tasks performed in occupations,

[85] For example, at least for older cohorts, West Germany is known as a country with a breadwinner model, opposed to Scandinavian countries with dual-income households, and has been used for studying peer effects in other contexts before (Dustmann et al., 2016; Pink, Leopold, and Engelhardt, 2014; Welteke and Wrohlich, 2019).

and by measuring the potential harm to the workgroup in employment interruptions due to turnover costs. I also test whether there are significant effects on old-age employment in the same workgroup, because staying in the same workgroup retains the workgroup-specific human capital, further confirming the importance of this channel.

I find that the peer effect is particularly pronounced in settings with traditional gender norms (West Germany), suggesting that social norms regarding women's old-age employment serve as the main channels for the peer effects. I find only limited evidence for information transmission and work complementarities channels.

I proceed to compute the social multiplier: for every woman employed at an older age, an additional 0.13 coworker women remained employed at an older age due to peers. I also show how such a shift in social norms regarding old-age employment of women could lead to regional convergence in old-age employment among women in the West and the East of Germany. Nevertheless, disentangling the mechanisms or attributing certain subsample analyses to one specific mechanism is difficult; hence, these channels are only suggestive.

I contribute to three strands of the literature. First, I extend research on retirement reforms and older workers' labor supply (Carta and De Philippis, 2024; Deshpande, Fadlon, and Gray, 2024; Geyer and Welteke, 2021; Lalive, Magesan, and Staubli, 2023; Manoli and Weber, 2016; Mastrobuoni, 2009; Rabaté, Jongen, and Atav, 2024; Staubli and Zweimüller, 2013; Ye, 2020). Complementing evidence on the direct effects of the reform under this study (Badalyan, 2025; Geyer and Welteke, 2021), I show that peers amplify delayed retirement: women close in age are more likely to remain employed when their workplace peers do. The reform thus operated through a dual mechanism—firms retained workers due to substitutability incentives (Badalyan, 2025), while peers reinforced employment through conformity and social norms regarding old-age employment among women. More broadly, this paper demonstrates that retirement is not solely an individual decision but is shaped by group-level norms and behavioral responses to statutory retirement ages (Behaghel and Blau, 2012; Blundell, French, and Tetlow, 2016; Seibold, 2021). In other words, the retirement behavior of older workers can be perceived as a reference point for one's own behavior. Recognizing these peer dynamics is essential for policy design, as they can amplify the impact of social insurance changes (Dahl, Løken, and Mogstad, 2014), alter long-run reform effects as norms evolve, and complicate inference since aggregate outcomes combine direct and peer responses (Glaeser, Sacerdote, and Scheinkman, 2003; Grodner and Kniesner, 2008; Welteke, 2015).

Second, I contribute to the literature using quasi-experiments to estimate coworker peer effects in labor market decisions and social insurance programs.[86] Existing work on retirement-related peer effects has largely examined couples and within-family dynamics—often showing

[86]Examples include disability pension participation (Rege, Telle, and Votruba, 2012), job search (Dustmann et al., 2016; Glitz, 2017; Saygin, Weber, and Weynandt, 2021), productivity (Bandiera, Barankay, and Rasul, 2009; Cornelissen, Dustmann, and Schönberg, 2017; Herbst and Mas, 2015; Mas and Moretti, 2009; Messina, Sanz-de Galdeano, and Terskaya, 2023), parental leave and labor supply (Casarico et al., 2025; Cavapozzi, Francesconi, and Nicoletti, 2021; Dahl, Løken, and Mogstad, 2014; Welteke and Wrohlich, 2019; Nicoletti, Salvanes, and Tominey, 2018; Carlsson and Reshid, 2022), and welfare take-up (Bertrand, Luttmer, and Mullainathan, 2000).

that women time their exits around their husbands' retirement (Atalay, Barrett, and Siminski, 2019; Bloemen, Hochguertel, and Zweerink, 2019; García-Miralles and Leganza, 2024; Johnsen, Vaage, and Willén, 2022; Lalive and Parrotta, 2017; Oral, Rabaté, and Seibold, 2024; Selin, 2017; Zweimüller, Winter-Ebmer, and Falkinger, 1996). Given that women are usually younger than their spouses, such coordination results in earlier exit from the labor force. I show that norms related to old-age employment of women are shifting: women increasingly align their retirement with workplace peers. This result represents a broader transition from domestic coordination to workplace-based coordination, underscoring how workgroup norms shape late-career decisions. Unlike most coworker peer effect studies, which focus on narrow firm-level cases or U.S. field experiments (Duflo and Saez, 2002, 2003; Brown and Laschever, 2012), I exploit a universal German pension reform that reshaped retirement behavior across the labor market and analyze mechanisms that relate to coworkers, using detailed coworker networks defined by establishment and occupation codes.

This paper also complements the recent study by Oral, Rabaté, and Seibold (2024), which examines peer effects in retirement decisions in the Netherlands across a broad range of networks, including spouses, siblings, neighbors, and coworkers. My analysis differs along two main dimensions. First, I study a different institutional setting: a universal reform that equalized retirement ages across genders in Germany. In contrast, the Dutch reform affected both genders. Exploring peer effects in gender-neutral retirement-age reforms is a new area of research. My paper is the first to test for the peer effects of a reform that made retirement ages gender-neutral. Germany is one of the first countries to transfer from gender-dependent to gender-neutral retirement ages, which allows me to analyze this research question and the mechanisms that relate to women's employment at older ages. Second, I exploit rich administrative data containing detailed occupation and establishment identifiers, which allows me to construct peer groups within narrowly defined workgroups. The occupational details available in German social security data enable a more detailed analysis of workplace peer mechanisms than does the Dutch data, which lacks such occupation identifiers. In particular, I examine channels including social norms, conformity, information, and work complementarities. Work complementarities have received little attention in prior peer effects literature.

Finally, I contribute to the literature on reforms aimed at reducing gender gaps in the labor market. While most studies focus on women of childbearing age (Goldin, Kerr, and Olivetti, 2021; Kleven, Landais, and Søgaard, 2019), much less is known about gender dynamics at older ages. Existing research shows that raising women's retirement age increases their employment, but I demonstrate that these effects propagate through workplace peer networks, especially in regions with traditionally stronger gender norms, such as those in the former West Germany. Closest to my mechanism, Boelmann, Raute, and Schönberg (2025) document that women from the former East Germany migrating west after reunification raised local women's employment by carrying their higher labor force attachment with them. My findings complement theirs by providing micro-level evidence that pension reforms also diffuse through coworkers within establishments. Together, these results highlight migration and peer spillovers due to reforms extending the careers of women as distinct yet reinforcing channels through which policy can reshape social norms around older women's employment, thereby narrowing the gender gaps in late-career labor supply.

The remainder of this paper is structured as follows. Subsection 3.2 outlines the institutional setting in Germany. Subsection 3.3 details the sample construction, peer group assignment, and identification strategy for causal inference. Subsection 3.4 presents the main results, followed by an analysis of underlying mechanisms in subsection 3.5, discussion in subsection 3.6, and the conclusion in subsection 3.7.

3.2 Institutional setting

In this section, I provide details on the labor market and pension system in Germany, which helps contextualize the peer effects. I also describe the 1999 reform, which motivates the identification strategy in the next section.

Labor markets in Germany. Compared to most Western European countries, West Germany has had a low employment rate for women and a high gender wage gap; the gender wage gap is still large and is decreasing only slowly. The gender gap in employment rates was particularly low in 1999, and shrank over two decades until 2019 (see Figure 3.1). It is noteworthy that although gender gaps around childbirth have received high attention, the gender employment gaps at older ages are just as striking (see Panel A of Figure 3.1). West German women traditionally had lower employment rates compared to East German women, due to the historic divide between the Federal Republic of Germany and the German Democratic Republic (GDR), religiosity in the West, and communism in the East (Rosenfeld, Trappe, and Gornick, 2004). As shown in Panel B of Figure 3.1, women in East Germany have significantly higher employment-to-population ratios than women in West Germany (e.g., in 2009, the difference was 2.1 p.p.). In 2009, women in West Germany were 11.6 p.p. less likely to be employed than men, whereas in East Germany this gap was only 4.5 p.p., confirming the breadwinner household model in West Germany and the more egalitarian dual earner model in East Germany.

Since East and West Germany share largely similar retirement rules, differences in norms surrounding old-age employment create a useful setting for studying peer effects. I expect these effects to be stronger in West Germany, where baseline employment rates before the reform were lower. Larger peer effects could, over time, contribute to a regional convergence in employment patterns at older ages.

Public pension system in Germany. The public pension system in Germany covers over 90% of the workforce, and operates on a "pay-as-you-go" basis, where social security contributions of the insured workers and taxes pay for the pensions of the old. There are two statutory retirement ages: the early retirement age (ERA) and the normal retirement age (NRA).[87] The ERA is the age at which a worker can begin claiming pensions (Panel A in Figure 3.2), while the NRA is the age at which full pensions can be claimed without any

[87] There is also a "Full Retirement Age", which exists only for very specific cohorts of workers, such as those with very long employment records.

Figure 3.1: Employment over population by gender and territory over time

Panel A: Employment over population by gender and age **Panel B:** Employment over population over time by East and West Germany

Notes: **Panel A** displays employment shares by 5-year age groups and gender. Pink bars represent women and black bars represent men; narrow bars correspond to 1999, while wide bars correspond to 2019. **Panel B** shows the evolution of employment shares among the working-age population (ages 15–65) from 1991 to 2023, again disaggregated by gender. Pink lines represent women and black lines represent men; dashed lines refer to East Germany, while solid lines refer to West Germany. Employment share is defined as the number of employed individuals divided by the total population in each respective group. The underlying data are sourced from the Federal Statistical Office.

deductions (Panel B in Figure 3.2). Retiring between the ERA and NRA results in a 3.6% deduction in pension benefits for each year taken early.[88] Because early retirement deductions in Germany were smaller than actuarially fair levels, many workers retired as soon as they became eligible. In a peer effects setting, this means peers' retirement timing is strongly clustered at the ERA, making their behavior more salient and influential on coworkers' own decisions.

Old-age pension for women. There are several pathways to retirement, which individuals use depending on eligibility. Up until the 1952 birth cohort, conditional on having at least 15 years of contribution to social security, ten of which were contributed after the age of 40, women could claim pensions as early as 60 years old, thanks to the pathway to early retirement at 60 available to women but not men.[89] Geyer and Welteke (2021) show that in the total sample of women born in 1951, around 21% retired before 63.[90]

Employment exits before ERA. While most German workers transition directly from employment into retirement with pensions, several alternative exit routes remain available. Some studies highlight unemployment insurance (UI) as a bridge to retirement (Gudgeon

[88]For example, retiring three years before the NRA results in an 18% pension deduction.

[89]The women's pathway to early retirement was a popular pathway for women who wanted to exit the labor force early, because the other pathways either required more contribution years to the social security system or implied a later ERA.

[90]Almost 60% of all the women born in 1951 were eligible for the old-age pension for women, 35% of which retired before 63 by utilizing the old-age pension program for women.

et al., 2023). This option is attractive because UI benefits replace roughly 60% of prior wages, the period on UI counts toward pension contributions, and job search requirements are less strict for older workers (Geyer and Welteke, 2021). For those aged 57 and above, UI duration was generous for the cohorts under study—up to two years—allowing, for example, a woman born in 1951 to exit employment at 58 and receive UI until retiring at 60. However, Gudgeon et al. (2023) find that because of the already-generous retirement pathway for women, using unemployment as a bridge to retirement was not as common for this cohort of women as for men. Other routes include disability insurance (DI) and inactivity due to caregiving, illness, self-employment, or personal preference. Women with sufficient contribution years (5 for the regular pathway and 35 for the long-insurance pathway) may choose inactivity before their ERA to care for family members, grandchildren, or to coordinate retirement with spouses (Lalive and Parrotta, 2017) before claiming pensions. These alternatives are relevant for the peer effects setting, as norm-driven increases in old-age employment and labor market activity could decrease early exits, thereby improving compliance with and enforcement of the higher old-age employment targeted by the reform.

The 1999 reform that abolished women's pathway to early retirement. Starting from the 1952 cohort, the women's old-age pension pathway was abolished, and the earliest age at which women could claim pension benefits rose by at least three years (see Table 3.B.1).[91] Moreover, starting from the 1952 cohort, there were fewer opportunities to bridge the gap between employment and pension claiming with UI. Overall, this is the largest increase in retirement ages for two consecutive cohorts in Germany, as other reforms increase the retirement ages gradually, in incremental steps over a larger span of birth cohorts. This large and discontinuous increase in retirement rules helps me circumvent the problem of separating reform effects from time or cohort effects, facilitating the identification strategy in subsection 3.3. Importantly, unlike reforms in some other countries, such as the Netherlands, this reform was uniform across industries, occupations, and regions, consistently and homogeneously affecting the retirement ages, which supports the comparability of treated and control workers discussed in the identification strategy below.

Geyer and Welteke (2021) find that women born in 1952 extended employment at ages 60–62 by 13.5 p.p.—about a 30% increase relative to the 1951 cohort—without affecting employment before age 60 despite the reform's pre-announcement. They also show rises in unemployment and inactivity, driven mainly by extensions of existing statuses rather than active substitution. In related work (Badalyan, 2025), I re-estimate the effects for women by their employment characteristics at the ages 58-59 and find a 17.3 p.p. increase at 60–62 (a 22% increase relative to the control mean).[92] Overall, these findings point to a strong direct effect of the reform on employment at ages 60–62, which may also generate peer effects.

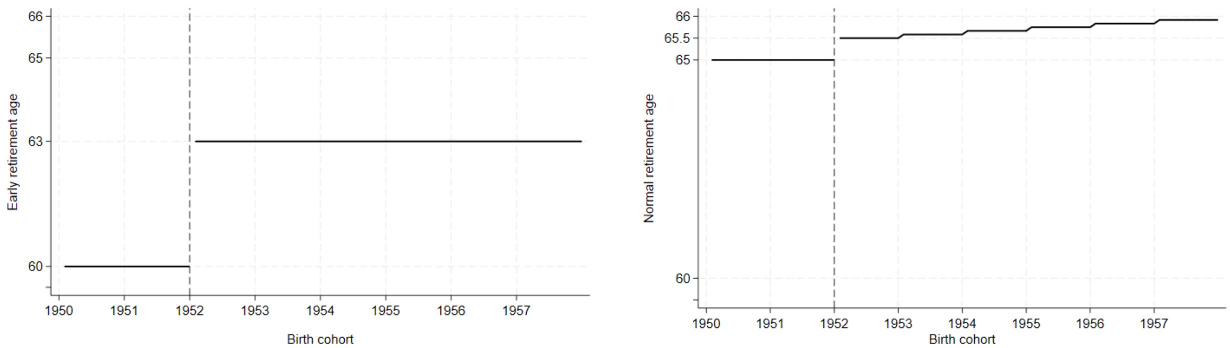
Figure 3.A.1 displays the distribution of retirement age (proxied by the age at the last labor

[91] For women with at least 35 years of social security contributions, the retirement age rose by three years; for those with just over five years of contributions, the retirement age rose by five years. Nevertheless, Geyer and Welteke (2021) show that 90% of women eligible for women's old-age pensions were also eligible for pensions for long-insured.

[92] The higher estimate reflects stronger labor force attachment in my sample, where over 70% of the control group remain employed at 60–62.

Figure 3.2: Early and normal retirement age rules by cohorts

Panel A: Early retirement age **Panel B:** Normal retirement age



Notes: **Panel A** shows the policy rule for ERA, the earliest age a person could claim pensions, by birth cohorts. The dashed line presents the birth cohort cutoff, January 1952, starting from which the ERA rose by at least three years. **Panel B** depicts the assignment rule of NRA, the age at which workers can claim full pensions, by birth cohorts. Before the 1952 cohort, there was a women's pathway to retirement, which had a fixed NRA at the age of 65. The vertical dashed line at the January 1952 cohort indicates the birth cutoff from which the women's pathway to early retirement was abolished. Starting from the 1952 cohort, the NRA for people eligible for the regular pathway to retirement is equal to the NRA for long-term insured, which used to be 65, but was raised by monthly increments per birth year starting from the 1947 cohort. Since women born before 1952 had their own NRA rules, their NRA was set at 65, and started increasing only from the 1952 cohort, with a discontinuous 6-month rise from the 1952 cohort. The detailed tables with ERA and NRA by cohorts can be found in Table 3.B.1.

market activity spell) for the women who were employed at age 58, and reveals that there are employment exit peaks before age 60 for only the 1950-1951 cohorts, i.e., those women born before the reform cutoff who were eligible for women's pathway allowing earlier retirement. These peaks shift to later ages for the 1952-1957 cohorts, and the earlier exits are more evenly distributed before the age of 62. The gradual spikes shown at different months at the age of 65 can be attributed to the increase in the normal retirement age (see the assignment rule in panel B of Figure 3.2).[93]

[93]There are also cohort-specific spikes in labor force exit between ages 63 and 64 for cohorts born after 1953, driven by the Full Retirement Age. Full Retirement Age slightly differs from the Normal Retirement Age, and it exists only for specific pathways, such as very long-insured workers (with 45 years of social security contributions). These spikes reflect the gradual two-month increases in the Full Retirement Age, which affected workers with very long employment histories (Felder, Geyer, and Haan, 2024). Because this reform varies by birth cohort rather than by peer exposure, I include coworker birth-cohort fixed effects in the empirical specification (see subsection 3.3). These fixed effects absorb cohort-level retirement incentives, ensuring that the spikes observed after age 63 (due to the Full Retirement Age) or after 65 (due to the Normal Retirement Age) do not affect the identification of peer effects.

3.3 Empirical framework and data

This section describes the data and the identification strategy that I use to estimate causal peer effects in employment and retirement decisions at older ages. I begin by outlining the peer effects model, identifying three key issues that can bias the results, and discussing how the literature suggests that one can address these problems using quasi-experiments. Then I illustrate this approach within the reform setting of this study (see subsection 3.2). Next, I provide details on the data and sample construction necessary for applying this method. Finally, I outline the regression equations that I estimate following this approach and assess their validity.

3.3.1 Empirical methodology for identifying peer effects

Consider a workgroup that includes peers and their coworkers, where, throughout this paper, I define *peers* as individuals directly affected by the reform, and *coworkers* as those who may experience indirect effects through their peers. There is no overlap between the two groups. The focus of this study is on peer effects on coworkers, which can be modeled as follows:

$$Y_{ig} = \alpha + \tau \overline{Y_{-ig}} + X_{ig}'\beta + \overline{X_{-ig}}'\gamma + W_g'\eta + \epsilon_{ig} \tag{24}$$

where Y_{ig} represents coworker i's outcome (e.g., an employment indicator at age 62). The main coefficient of interest, τ, captures the effect of $\overline{Y_{-ig}}$, the mean employment outcome of peers (excluding the coworker i) in group g, on coworker i's retirement outcome, Y_{ig}. Additionally, the literature typically accounts for coworkers' individual characteristics (X_{ig}), peer characteristics ($\overline{X_{-ig}}$), and workgroup-level factors (W_g).

Identification challenges. Estimating peer effects may lead to biased results. Manski (1993) lists three challenges in estimating τ: correlated unobservables, endogenous group formation, and simultaneity. The problem of correlated unobservables arises when, even after controlling for the coworker (X_{ig}), peer ($\overline{X_{-ig}}$), and workgroup (W_g) characteristics, some unobserved factors remain correlated with the outcomes of the peers, leading to a bias in the estimated τ. For example, contextual factors or common shocks, such as workplace conditions or industry- and occupation-specific demand fluctuations, may lead both coworkers and their respective peers to delay retirement regardless of each other's decisions. Endogenous group membership occurs when peers and their coworkers self-select into specific occupations or establishments based on their work (or leisure) preferences and their attraction to particular peer groups. This selection process complicates causal inference because the composition of peer groups is not random. Simultaneity in interactions leads to the reflection problem, where the observed coworker outcome, Y_{ig}, may be both a cause and a consequence of peer retirement behavior, $\overline{Y_{-ig}}$. This bidirectional influence makes it difficult to disentangle the true effect of peer outcomes on coworker decisions.

The peer effects literature addresses identification challenges by exploiting instruments that shift peers' outcomes, $\overline{Y_{-ig}}$, without directly affecting individual outcomes Y_{ig}. One approach relies on partially overlapping networks (De Giorgi, Pellizzari, and Redaelli, 2010), while another leverages quasi-experimental variation from reforms that alter peers' employment incentives (Dahl, Løken, and Mogstad, 2014). I adopt the latter, using the increase in ERA for the 1952 cohort as a source of exogenous variation (see subsection 3.2). This choice is motivated by the stronger predictive power of reform-based instruments. Consistent with this, Geyer and Welteke (2021) and Badalyan (2025) document sharp employment responses to the reform.[94] In addition, the reform-based strategy also permits the calculation of a social multiplier, relevant given ongoing global increases in retirement ages.

Overcoming identification challenges through a quasi-experiment.[95] To identify peer effects, I exploit a reform that has heterogeneous implications for peers and homogeneous implications for coworkers. The 1999 pension reform raised the ERA starting with the 1952 birth cohort. For intuition, consider two women in two different workgroups: one born in 1951, still eligible to retire at 60, and another born in 1952, who now faces a higher cost of retiring early. Their retirement choices create exogenous variation in peer behavior.

Coworkers are then defined as women born 1953–1957, i.e., too young to make retirement decisions but close enough in age to be approaching retirement themselves.[96] This design ensures that coworkers differ only in the peer exposure they receive. To avoid reflection, the peer and coworker samples are mutually exclusive: I use a "leave-out group" approach, excluding all peers from the coworker sample.

Figure 3.3 illustrates the idea. Peers are the reform-affected (1952 cohort) and unaffected (1951 cohort) women, whose retirement behavior may indirectly shape that of their coworkers. By aligning groups this way, the strategy addresses the three core challenges in the peer effects literature. First, correlated unobservables are mitigated by exploiting exogenous policy variation across birth cohorts. Second, endogenous group membership is limited since peer groups are defined before members reach retirement age. Third, simultaneity is avoided by focusing on younger coworkers who have the same rules among each other; therefore, any differences displayed should be attributed to their peers' influence.[97]

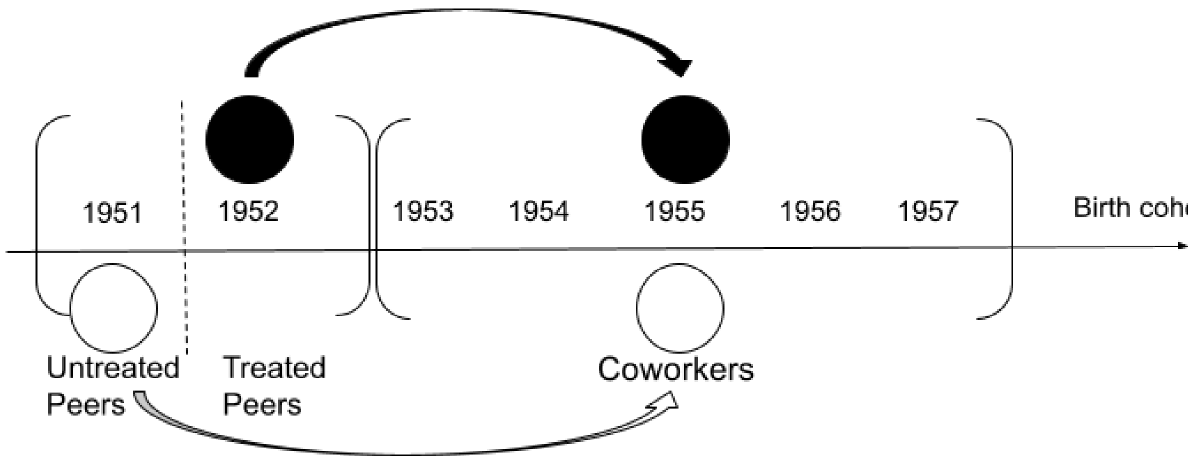
[94] Another reason is data limitations—German social security records lack household identifiers needed for overlapping network designs.

[95] The design and the followed identification strategy were used by Welteke and Wrohlich (2019) to estimate peer effects in parental leave.

[96] Coworkers (women of birth cohorts 1953–1957) do not determine peer exposure and therefore do not affect treatment assignment directly. However, they may still influence the broader workplace environment through interactions among themselves. For example, a larger number of coworkers may amplify the transmission of peer behavior. To account for this, I control for detailed workgroup composition—including the number of coworkers and peers (measured when the peer was 57)—and perform robustness checks using alternative peer bandwidths and workgroup definitions (see subsubsection 3.4.2). The estimated peer effects remain stable across these specifications.

[97] One concern is that coworkers are themselves affected by the reform and may adjust their retirement expectations earlier, potentially transmitting these expectations to peers when workgroups are defined. While this possibility cannot be entirely ruled out, focusing on younger cohorts is a common strategy used to mitigate reflection in the peer effects literature (Dahl, Løken, and Mogstad, 2014; Welteke and Wrohlich,

Figure 3.3: Simplified identification strategy



Notes: The figure illustrates the assignment of workers into groups. *Peers* are women born around the 1952 cutoff: those born in 1951 could still retire at 60, while those born in 1952 faced a higher ERA. Their retirement behavior may, in turn, influence their *coworkers*, who are younger cohorts in the same workgroup. Treated peers (1952) and their coworkers are shaded black; untreated peers (1951) are shaded white. A non-simplified graph is shown in Figure 3.A.3.

If I had a single peer (e.g., spouses or siblings) within a group, a regression discontinuity design may suffice (Dahl, Løken, and Mogstad, 2014). In establishments, however, peer groups often contain multiple treated and untreated members. Restricting to just one peer would reduce external validity and make it impossible to study mechanisms across subgroups. I therefore extend the strategy to allow for multiple peers within workgroups, as long as they are born to one side of the cutoff, as detailed in the next subsection.

3.3.2 Social security data and sample construction

Data source for the baseline sample: Integrated Employment Biographies. I use German social security data from the Integrated Employment Biographies (IEB), provided by the Data- and IT-Management (DIM) at the Institute for Employment Research (IAB).[98] These data cover all workers subject to social security until 2022, with records starting in 1975

2019). Moreover, existing studies of this reform find little evidence of labor supply responses before individuals approach retirement ages (Badalyan, 2025; Geyer and Welteke, 2021), suggesting that anticipatory adjustments among younger coworkers are likely limited. Because administrative data do not include expectations or conversations among coworkers, the analysis relies on observed labor supply behavior.

[98]I use data from the full universe of German employment records *(IEB, version 17_00_00_202212)* of the IAB. Due to its administrative origin, these data are confidential and can only be accessed on-site at IAB. Access for guest researchers requires clearance from the German Federal Ministry of Labour and Social Affairs.

for West Germany and 1991 for East Germany. Employers are required to report information on the full workforce as of June 30$^{\text{th}}$ each year, as well as any changes in employment spells, including job entry, exit, or contract modifications. The reported information includes details such as workplace variables (start and end dates of spells, occupation, contract type, detailed daily wages, industry, and location of establishments), and some basic demographic information (birth year and birth month, gender, education, and place of residence). This dataset is well-suited for the study because its detailed employment records of all workers within the workplace and occupational codes allow for the precise assignment of peer groups across the entire German workforce, while gender and birth dates allow for assigning the treatment groups according to the reform. Throughout this paper, I use the first four digits of *Klassifikation der Berufe (KldB) 2010* (see Paulus, Matthes et al. [2013] for an overview). The rest of the variables, such as tenure, location, occupations, industries, etc, are useful for testing peer effects mechanisms.

Due to data access restrictions, the full universe of German social security records cannot be extracted at once. Instead, the data were constructed by my request to the IT department, selecting the universe of private-sector establishments that employed at least one woman born in the 1950–1953 cohorts in 2008, together with the full employment histories of workers within those establishments. Workers with a history of employment as sailors or miners are excluded because their retirement rules differ. The year 2008 was chosen because it is the last year in which all four cohorts were younger than age 59. In the baseline specification, I focus on peers born in 1951–1952, while the two-year bandwidth allows additional robustness checks. To ensure comparable workplace environments and limit excessive worker turnover, I further restrict the sample to establishments employing between 5 and 500 workers. The resulting dataset includes 190,228 establishments and 26.5 million workers who were ever employed in them, of whom about 9.1 million were employed in 2008. Table 3.B.2 summarizes the sample construction.

Data source for additional samples: Sample of Integrated Employer–Employee Data. For the placebo tests, I use a smaller dataset constructed from the Sample of Integrated Employer–Employee Data (SIEED7518), which contains a random 1.5% sample of establishments in Germany. Because placebo tests only require verifying whether spurious discontinuities appear around artificial cutoffs, and it is not necessary to perform detailed subsample analyses for mechanisms, a representative random sample is sufficient for this purpose. The SIEED dataset allows these tests to be implemented efficiently while preserving the relevant variation in the data. Similar approaches have been used in studies based on German administrative data, for example Welteke and Wrohlich (2019). More details on the placebo tests and the corresponding data sources are provided in subsubsection 3.4.2 and section 3.7.

To conclude, the baseline analysis uses IEB, a large administrative dataset constructed from the universe of establishments satisfying the extraction criteria, while the placebo tests use the SIEED, a representative sample from the IEB. The construction of samples from these two sources is similar and is described below.

Sample construction. Following the empirical methodology outlined in the previous subsection, I proceed with the definition of peer groups.

Workgroup definitions. To avoid misattributing spillover effects across broader organizational units, I restrict the peer exposure to within-establishment and same-occupation groups to ensure that the estimated peer effect is not diluted by non-overlapping work environments. Workgroups are defined at the 4-digit occupation level to capture coworkers who are likely to interact regularly and perform similar tasks. This classification improves the relevance of peer effects by reflecting shared work environments, comparable responsibilities, and similar exposure to firm-level shocks (Jäger and Heining, 2022). Previous literature has used 4-digit occupations to define peer groups, taking advantage of the job cells available in social security records (Fietz and Schmeißer, 2024; Messina, Sanz-de Galdeano, and Terskaya, 2023).

Peers: women who transfer the peer effects. I begin by marking peer women employed at the ages of 56 and 57 and belonging to either the 1951 birth cohort (treatment group) or the 1952 birth cohort (control group). I focus on women employed at ages 56-57 to address the reflection problem: all peers and coworkers are below the pre-reform retirement age of 60 for women. This restriction also ensures that the peer women have at least 2 years of tenure at the given establishment. [99] Conditioning on employment at a certain age would be problematic if there was discontinuity in employment rates among the treated and control peers. However, Geyer and Welteke (2021) show that the employment statuses do not change due to the reform until the workers reach age 60.[100] This restriction results in 214,435 peer women, 110,796 of whom are treated, and 103,639 are in the control group (see the first row in Table 3.B.3).

Before the reform, women could retire at age 60 only if they had accumulated at least 15 years of contributions (see subsection 3.2). While one might consider restricting the sample to those who meet this threshold, I deliberately avoid doing so for three reasons. First, the peer-effects question in this paper concerns *how workplace exposure to treated peers shifts the employment behavior of all coworkers at the policy threshold*, not only those who would have qualified for the discontinued "women's pathway". Conditioning on 15 contribution years would change the estimand from a network-level spillover among the full coworker population to a narrower effect among a selected, highly attached subgroup. This is at odds with how peer interactions and social norms operate in workgroups that include coworkers with heterogeneous attachment histories, and norms diffuse along that full network. Second, the 15-year rule is no longer a relevant decision margin for the younger cohorts who serve as coworkers in my design. Post-1952 cohorts did not face incentives to accumulate 15 years for early claiming; restricting to that legacy threshold would therefore remove precisely those coworkers for whom peer influence is most policy-relevant (e.g., women who might otherwise

[99] This restriction also helps overcome another problem. Since the originally provided data include establishments with at least one peer employed in 2008, it potentially introduces selection on tenure and job attachment, because the older cohorts (treated peers, born in 1952) satisfy more stringent tenure criteria than the younger ones (control peers, born in 1951). By aligning the peers as those employed in the sampled establishments at the ages of 56-57, I thereby avoid asymmetry in tenure-based selection.

[100] I am unable to confirm this result, as that would require requesting entirely new social security data from the data provider.

be inactive, self-employed outside the system, or on DI at 62), and would understate the potential reach of norm transmission inside firms. For example, Geyer and Welteke (2021) report that 32% of all the women in the 1952 birth cohort are non-employed at the age of 62-63. Hence, employment at 62 provides a good measure of a new norm and compliance with it. Third, the outcome I study—labor-market activity at age 62—directly targets the policy's compliance margin (delaying exits to at least age 63). Whether a coworker meets a defunct eligibility rule is orthogonal to this compliance margin: what matters for peer spillovers is whether exposure to treated peers keeps coworkers *active* at age 62. Conditioning on a legacy eligibility criterion would reduce external validity for the policy audience, who care about aggregate compliance and spillovers across *all* workers in affected workgroups.

Next, I group all women who belong to the same workgroup, defined as an occupation group within an establishment. I count the number of peer women born in the 1951-1952 interval and drop any groups where peer women from both cohorts are present or where no women from either cohort are in the workgroup (see the visual walkthrough Figure 3.A.2). This restriction ensures that the peer groups include at least one woman who could claim pensions at age 60 control group) or 63 (treatment group), resulting in 153,647 peer women, 80,114 of whom are treated and 73,533 are in the control group (see the second row in Table 3.B.3). Thus, the simplified model from Figure 3.3 is extended to allow multiple peers, all of whom share the same retirement rules. Because I dropped the groups with peers from both sides of the cutoff, all the peers in each workgroup are subject to the same rules.

Coworkers: women who receive the peer effects. I proceed to mark the coworkers as women in the 1953-1957 birth cohorts, who were employed in the year when the peers were at the age of 57; therefore, the coworkers were 51-56 years old at the date of exposure. I do not analyze the younger cohorts because they are right-censored beyond age 65 in my data.[101] I do not include older cohorts (e.g., born before 1951) in my coworkers group because peers could observe the retirement decisions of older cohorts of coworkers, thereby leading to the reflection problem. I further drop the coworkers who were exposed to more than one workgroup with peers employed at 57, for example, due to a switch of employment or simultaneous employment.[102] In total, there are 503 such workers (368 women and 135 men) who are removed. I assign the resulting coworkers as "treated" if their peer(s) were born in 1952 (and had to wait an additional three years to claim pensions) and "control" if their peer(s) were born in 1951. I keep the workgroups that had at least one coworker in the year when the peers were 57. The roadmap for sample construction is depicted in Figure 3.3. To perform peer effect analyses, I compute the average observable and outcome characteristics over all the peers in the workgroup and transfer this information to all the coworkers. Therefore, my final data consists of coworkers. The final sample includes 64,324 workgroups with exposure to 86,593 peers (40,627 treated and 36,828 control) and their 182,584 female coworkers (see the last row in Table 3.B.3). A coworker has, on average, 59 workgroup colleagues, and the median number of colleagues is 34, which motivates allowing several peers to influence coworkers, rather than restricting them to small workgroups.

[101] At the time when this paper was written, social security data were available up until the year 2022.

[102] Note that I do not perform a similar restriction for peers; therefore, they can affect several workplaces.

Putting these steps together, coworkers are assigned to treatment or control groups based on the birth cohorts of their peers. The treatment group consists of coworkers in workgroups that employed at least one woman born in 1952, no women born in 1951, and at least one coworker born between 1953 and 1957 when the 1952 cohort was age 57. The control group consists of coworkers in workgroups that employed at least one woman born in 1951, no women born in 1952, and at least one coworker born between 1953 and 1957 when the 1951 cohort was age 57. This construction ensures that coworkers are exposed either to treated peers (1952) or untreated peers (1951), but never both.

The sample construction details for alternative samples (such as those used in placebo tests), together with respective sample sizes (including the number of coworkers, peers, and workgroups), can be found in section 3.7 and Table 3.C.1.

Outcome variables. The main outcome variable is employment at 62 years of age. Because the reform raised the earliest retirement age from 60 to at least 63, focusing on an indicator for employment at 62 constitutes a clean test of whether the reform induced delayed labor market exit through continued activity. The binary nature of this variable allows for clear predictions of changes in employment shares and enhances interpretability. It also provides a direct measure of compliance with the new earliest retirement age, which was the reform's explicit target.[103]

As described above, the administrative data used in this paper contain complete employment histories only within the establishments included in the original extraction (i.e., the firms where workers were employed in 2008). For labor market activity outside these establishments, the data provide only the biographical variable "last labor market activity", which aggregates employment, unemployment insurance receipt, and welfare spells. As a result, it is not possible to fully distinguish employment from unemployment across the entire labor market. To address this limitation, I additionally analyze employment at the same establishment as a complementary outcome, which confirms that the main results are driven by continued employment rather than transitions into unemployment. Job mobility at these ages is relatively low, so labor market activity at 62 largely reflects continued employment. I discuss this further in subsection 3.6.

Given that, post-1952, bridging the gaps between employment and retirement was limited (see subsection 3.2), the vast majority of those who are active in the labor market at the age of 62 are employed; therefore, for simplicity, I use the terms interchangeably. Women who are not counted as active in the labor market at 62 therefore fall into the complement category, which primarily consists of: early retirees (rare post-reform, but possible because of DI), inactive women (e.g., caregiving, household work), those receiving disability benefits, self-employed outside the social security system, or those who emigrated or otherwise left the labor force. Large peer effects on this measure thus reflect peers shifting women from these alternative non-covered or inactive states into active labor market participation at the policy-relevant threshold. Job mobility at these ages is relatively low, so labor market activity

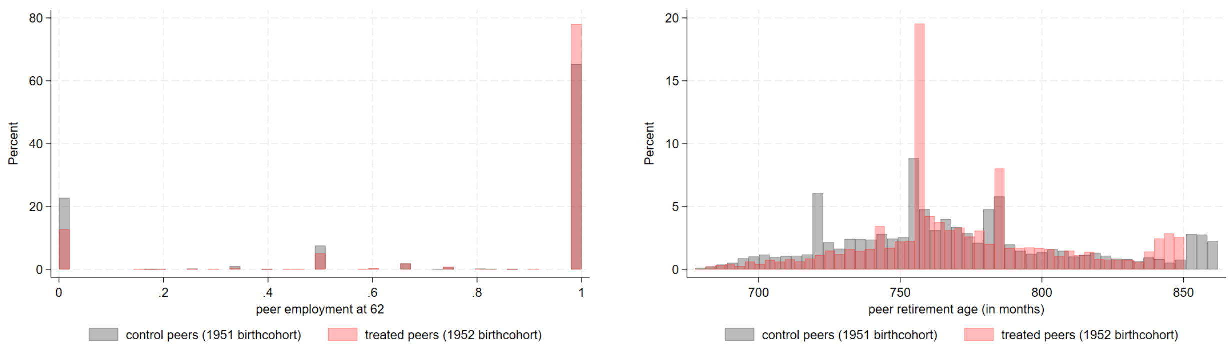
[103] Similarly, (Welteke and Wrohlich, 2019) define the main outcome variable as "parental leave for 10 months" to test for the compliance with the new parental leave duration set by the reform.

at 62 largely reflects continued employment very closely.

Studying a binary threshold such as “employment at 62” is particularly relevant for evaluating whether the reform achieved its intended goal. Even modest changes in continuous retirement age may have limited policy significance if they do not move individuals across this threshold, whereas changes in the share active at 62 directly capture compliance with the higher earliest retirement age.

Figure 3.4: Distribution of peer outcome variables by treatment status

Panel A: Employment at 62 | **Panel B:** Retirement age (in months)



Notes: These graphs show the distributions of peer outcome variables (dummy for employment at 62 in **Panel A** and continuous retirement age in months in **Panel B**) by treatment groups. The peer outcomes are computed as averages over peers. The groups are treated if all the peers are born in 1952 (whose retirement age has been raised), and untreated if all the peers are born in 1951 (eligible to retire at 60 through women’s pathway to early pension claiming).

In a complement analysis, I use indicators for employment between 58 and 64 years of age, mean employment between 60 and 62 years of age, mean employment in the same workgroup, and continuous retirement age (in months). The continuous retirement age captures the extensive margin and longer-term adjustments, offering a fuller picture of peer influence. However, it is noisier due to right-censoring of data in 2022 and because it is proxied from the last spell of labor market activity rather than actual pension claiming. Panel B of Figure 3.A.1 shows the distributions of employment at 62 and retirement age in months in my sample, relative to the distribution of the 2% random sample of workers born in 1951-1957 (in Panel A). The distributions of these outcome variables display similar patterns.

Figure 3.4 displays the distributions of the average peer (1) employment at 61 and (2) retirement ages (in months) by workgroup treatment status in my final sample. The likelihood of being employed at 60 is shifting towards 62 for treated peers (see Panel A in Figure 3.4), while the continuous measure of retirement ages in months shifts towards 756 months, which corresponds to 63 years old (see Panel B in Figure 3.4).

3.3.3 Identification strategy: instrumental variable approach

To address challenges in identifying coworker peer effects, I employ an instrumental variable (IV) approach that leverages quasi-random variation in pensionable retirement ages induced by the January 1952 reform cutoff. Let Z_g denote the workgroup treatment status (eligibility of peers to claim old-age pensions at 60). All peers within a group share the same binary treatment status. The sample includes multiple peers per group, allowing estimation of peer effects on retirement decisions using a two-stage least squares (2SLS) framework.

First stage: I instrument the average peer outcome, $\overline{y_{1g}}$ (employment at age 62 or retirement age), with Z_g:

$$\overline{y_{1g}} = \alpha_0 + \lambda Z_g + W_g'\alpha_1 + X_{ig}'\alpha_2 + \overline{X_{-ig}}'\alpha_3 + u_{1g}, \tag{25}$$

In baseline specifications, I show the results with no controls, and with controls, where the individual controls include wages and education categories, because previous literature confirms that education is an important determinant of employment at an older age (Geyer et al., 2022). The workgroup controls include establishment size, workgroup size, number of peers, number of coworkers (all measured at the time of the workgroup construction, i.e., when the peer was 57), and a dummy for the locations of the establishments in a region of the former East Germany. I additionally include fixed effects for the coworkers' birth cohorts to control for their rises in the normal retirement age (see Panel B of Figure 3.2 for the rule).[104] Controlling nonparametrically for birth cohorts mitigates concerns that age composition drives peer group formation.

Interpretation: λ measures how much peers' retirement outcomes respond, on average, to the reform-induced treatment.

Reduced form: The effect of the workgroup treatment on coworkers' outcomes is estimated as

$$y_{2ig} = \beta_0 + \delta Z_g + W_g'\beta_1 + X_{ig}'\beta_2 + \overline{X_{-ig}}'\beta_3 + v_{2ig}. \tag{26}$$

Interpretation: δ represents the overall change in coworkers' outcomes that occurs when the peers are exposed to the reform, i.e., intention-to-treat (ITT) effect.

Peer effects: The peer effect is obtained by dividing the reduced-form coefficient δ by the first-stage coefficient λ. It is obtained through an instrumental variable approach, where the average peer outcomes $\overline{y_{1g}}$ are instrumented by the workgroup treatment status Z_g:

$$y_{2ig} = \gamma_0 + \tau\widehat{\overline{y_{1g}}} + W_g'\gamma_1 + X_{ig}'\gamma_2 + \overline{X_{-ig}}'\gamma_3 + w_{2ig}. \tag{27}$$

Interpretation: τ captures the effect on a coworker's retirement outcome of a full-scale increase in the average peer retirement outcome, from its minimum to maximum, induced by the

[104] I also include the same set of controls for ITT and 2SLS regressions.

reform.[105]

Standard errors: All regressions cluster standard errors at the workgroup level to account for the correlation within the group.

Identification assumptions. The IV strategy rests on four assumptions. First, *relevance* requires that the reform significantly shifts peers' retirement behavior. Second, *conditional independence* requires that treatment assignment is as good as random around the cutoff, conditional on covariates. Birth dates are exogenous, and strategic manipulation is implausible, as parents could not anticipate the reform.[106] Because eligibility for women's pension claiming at 60 required long contribution histories and the reform was unanticipated, selective sorting into workgroups is unlikely. Third, the *conditional exclusion restriction* requires that the reform affects coworkers' outcomes only through their peers, conditional on observables. Since coworkers are younger and face the same retirement rules across groups, the only systematic channel of influence is peer behavior. Finally, *structural invariance* requires that the relationship between predicted peer outcomes and coworkers' outcomes is stable across individuals. This condition ensures that the 2SLS estimate recovers the causal peer effect.

The F-test on the first-stage regression assesses relevance. Summary statistics (both individual and workplace variables) for peer and coworker samples are displayed in Table 3.B.4 and Table 3.B.5. Coworker, peer, and workgroup characteristics are measured in the year when the peer cohort is 57 years old, ensuring that peer groups are defined before retirement decisions.[107]

Because the groups are defined by adjacent birth cohorts around the reform cutoff, large systematic differences are not expected. The table confirms that the groups are broadly comparable along most dimensions. In particular, vocational education, tenure, foreign status, and managerial status are nearly identical across groups. Some statistically significant differences appear in variables including university education, earnings, establishment size, and workgroup size. For example, treated peers have slightly higher average earnings and are employed in somewhat larger establishments and workgroups. However, these differences are economically small. Expressed in standardized units, the differences are generally below 0.1 standard deviations, suggesting limited practical imbalance despite statistical significance driven by the large sample size. The largest differences occur in workplace composition variables, including workgroup size, number of peers, and the share of older workers, which mechanically vary with cohort composition around the cutoff. Importantly, the empirical

[105] Since peer exposure is measured as the average share of treated peers, a one-unit change corresponds to a switch from none of the peers being treated (0) to all peers being treated (1). Hence, the peer effect captures how the employment of co-workers responds when peer retirement behavior changes throughout this entire range.

[106] Birth dates are widely used for identification; in this context, see Geyer and Welteke (2021).

[107] Defining coworker exposure relative to the peer's age is standard in peer-effects designs because it ensures that peer groups are formed before retirement decisions occur and avoids simultaneity between peers and coworkers. This approach has been used in related quasi-experimental peer-effects settings, for example, in studies of parental leave and fertility behavior (Dahl, Løken, and Mogstad, 2014; Welteke and Wrohlich, 2019), where the childbirth date before and after the reform is used.

specification conditions on these observable characteristics at the individual and workgroup level. The IV estimates are therefore identified under a conditional independence assumption rather than requiring exact balance in the raw data. To assess the robustness of the results, the next section presents specifications with extensive covariate controls and performs a sensitivity check to alter the set of these controls. Across these specifications, the estimated peer effects remain stable, supporting the validity of the identifying assumptions.

3.4 Results

In this section, I proceed in three steps. First, I show the direct effects of the reform on peers at ages 58-64. Next, I estimate the first stage, ITT, and peer effects of the reform on employment at 62 and a continuous measure of retirement age (in months). Finally, I perform comprehensive sensitivity and robustness checks to confirm these results.

3.4.1 Baseline results

First, I estimate direct effects of the reform on employment of peers at ages 58-64 in Figure 3.A.4.[108] Panel A shows a sharp increase in employment at 60-62 in the first stage regressions. There are also significant adjustments at ages 58-59, but at least three times smaller in magnitude. Having shown that the first stage is the largest, around 60-62, I focus on the employment indicator at 62 as the main outcome in peer effects.[109]

The baseline results for employment at 62 are presented in Figure 3.5, and reveal positive and significant effects in the first stage, reduced form, and peer effects.

First stage and instrument validity. In the first stage (first column in Table 3.1), the instrument—whether a worker's peers were subject to the higher ERA—significantly predicts peer employment at 62, i.e., an 11.6 p.p. increase in probability to be employed at 62 (16.2% increase relative to the control mean in the workgroup where peers could retire with old rules). This effect is slightly smaller than in Badalyan (2025) and Geyer and Welteke (2021), who report that a rise in the ERA led to a 13.5 p.p. (30.6%) and 17.1 p.p. (22.1%) increase in employment at the ages of 60-62, respectively. Such a difference stems from the specifics of my sample—the peers in this paper are employed at the ages of 56-57, and in such workgroups where all the peers are born on one side of the cutoff. The corresponding first-stage F-statistic is approximately 768.52, indicating a strong instrument. The confidence intervals of the specification that controls for workgroup, coworker, and peer characteristics (second column of Table 3.1) include the original estimate of 11.6 p.p.

[108]Note that I start analyses from the age of 58 because I restricted the sample of peers to those employed at 56-57, and define coworkers as those employed at this age.

[109]I also provide additional analyses for ages 60-62, continuous measures of retirement age, etc. in additional analyses below.

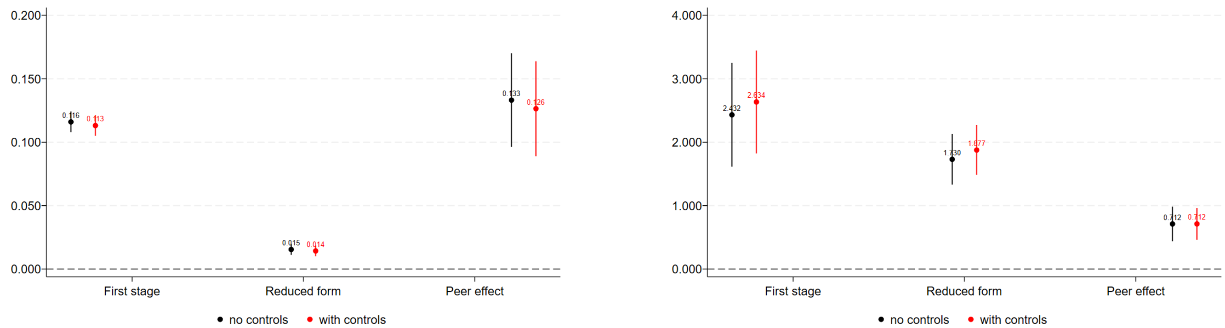
Table 3.1: First stage, reduced form, and peer effect regressions for employment at 62

	First stage		Reduced form		Peer effect	
Z	0.116***	0.113***	0.015***	0.014***	0.133***	0.126***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
firm size		-0.000		0.000		0.000
		(0.000)		(0.000)		(0.000)
peer group size		0.000		-0.000***		-0.000***
		(0.000)		(0.000)		(0.000)
N peers		-0.011***		-0.001		0.001
		(0.003)		(0.002)		(0.002)
N female coworkers		0.001		-0.000		-0.000
		(0.001)		(0.000)		(0.000)
East Germany		-0.018***		0.024***		0.026***
		(0.005)		(0.003)		(0.003)
peer vocational education		0.019**		0.017***		0.015***
		(0.007)		(0.004)		(0.004)
peer university education		0.044***		0.028***		0.023***
		(0.011)		(0.006)		(0.006)
peer earnings		0.000***		-0.000***		-0.000***
		(0.000)		(0.000)		(0.000)
vocational education		0.004		0.033***		0.033***
		(0.004)		(0.004)		(0.004)
university education		0.017***		0.050***		0.048***
		(0.006)		(0.005)		(0.005)
earnings		-0.000***		0.000***		0.000***
		(0.000)		(0.000)		(0.000)
Control mean	0.716		0.785			
Observations	182584	182584	182584	182584	182584	182584
N workgroups	64324	64324	64324	64324	64324	64324
Controls incl. cohort FE	No	Yes	No	Yes	No	Yes
R squared	0.023	0.040	0.000	0.012	-0.002	0.011

Notes: The outcome variable is *employment at 62*. This table shows the effect of the rise in ERA on peers' own employment: 1st stage (Equation 25), ITT (Equation 26), and IV regressions (Equation 27). Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers, establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. The control means are the average values of the outcomes when I limit the sample to the workgroups with peers born in 1951, i.e., who could claim pensions at 60, conditional on contribution years. Robust standard errors in parentheses are clustered at the workgroup level.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Figure 3.5: Baseline reduced form and peer effect regressions

Panel A: Employment at 62 **Panel B:** Retirement age (in months)



Notes: Coefficient plots. The columns correspond to the first stage, reduced form (ITT), and 2SLS (peer effect) regressions for employment at 62 (**Panel A**) and retirement age in months (**Panel B**). The points represent the estimated coefficients, and the bars represent the 95% confidence intervals. The black estimates represent the regressions with no controls, while the red estimates show regressions with controls. Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers (all measured when the peer was 57), establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. Corresponding tables with more details can be found in Table 3.1 and the first row in Table 3.B.6.

Reduced form. Having established a strong first-stage effect, I now estimate the effect of being exposed to peer women who experienced a rise in the ERA on coworkers' probability of being employed at the age of 62. In the reduced form (third and fourth columns in Table 3.1), workers in treated peer groups are 1.5 p.p. (1.9% increase relative to the control mean) more likely to be employed at 62, and this result is robust to the inclusion of covariates. This ITT effect constitutes around 10% of the first stage effect. This means that when the early retirement age of a peer increases, about 0.1–0.2 coworkers respond, which is a plausible effect. Using the same identification as I do for a different setting of parental leave, Welteke and Wrohlich (2019) find a 6.8 p.p. ITT effect, meanwhile, Dahl, Løken, and Mogstad (2014) find a 3.5 p.p. ITT effect for coworkers, which is closer to my estimate.

This estimate captures the ITT effect of peer exposure to the reform: it reflects the change in a worker's probability of remaining employed at 62 due solely to being surrounded by affected peers, without accounting for the degree of peer response. Importantly, this estimate does not rely on strong assumptions about the exclusion restriction and therefore offers a transparent, policy-relevant summary of peer effects from the ERA reform within workgroups.

Peer effects. The last two columns in Table 3.1 show that the peer effects in employment at 62 are around 13.3 p.p. This result implies that coworkers are around 13.3 p.p. (17% increase relative to the control mean) more likely to be employed at the age of 62 if their peers decide to do so in response to the rise in their pensionable ages. Given the instrument's binary nature, the peer effects are obtained by scaling the reduced form effect (columns 3-4) by the first stage estimates (columns 1-2).

Because the first stage captures only the fraction of peers whose behavior changes in response to the reform, the IV estimate rescales the reduced-form effect to reflect the impact of a full change in peer behavior. In this sense, the magnitude of the peer effect partly reflects the incomplete compliance of peers with the reform, rather than implying that coworker responses exceed the direct reform effect.[110] Given that coworkers are exposed to about 1.5 peers on average, this corresponds to roughly a 9 percentage point increase in coworker employment per additional peer remaining employed. This indicates substantial but not implausibly large peer effects.

Moreover, coworkers are typically exposed to multiple peers within the same workgroup, so the estimate reflects the cumulative influence of peer behavior rather than a one-to-one interaction. This helps explain why the estimated peer effects are sizable. At the same time, the magnitude is broadly in line with existing evidence in the literature: it is smaller than estimates reported by Welteke (2015), but comparable to the 11 p.p. effect found by Dahl, Løken, and Mogstad (2014) in the context of paternity leave. Overall, the results indicate economically meaningful, but not implausibly large, spillovers in late-career employment decisions.

3.4.2 Robustness checks

In the following subsection, I perform several robustness and specification tests. The first set of these tests corrects for further potential imbalances between the treated and control groups by including a set of additional covariates. The further four specifications alter the definitions of treatment, peers, workgroups, and coworkers. All the results confirm the existence of peer effects and the robustness of the estimates presented above. Finally, I re-estimate the peer effects on a placebo sample with no change in the retirement regime and show no effects, confirming that the effects studied in the baseline specification can be attributed to peer effects. These robustness checks help ensure that my estimated peer effects reflect behavioral responses to the continued presence of older coworkers, rather than institutional frictions, selection, or unrelated workplace trends.

Sensitivity to the inclusion of covariates. I test the validity of my findings by including additional coworker variables (see Panel B in Table 3.B.7). These covariates include coworkers' full-time status, experience, a dummy for being foreign, management status, and estimated AKM worker fixed effects. To rule out that changes in peer characteristics—rather than their continued presence—drive my results, I also re-estimate peer effects controlling for additional cohort-specific average peer characteristics (same variables as for coworkers

[110] Let y_2 denote coworker employment and y_1 peer employment. The IV estimate identifies $\tau = \frac{\text{Cov}(Z,Y_2)}{\text{Cov}(Z,Y_1)} = \frac{\text{reduced form}}{\text{first stage}}$, where Z is the reform-induced instrument. Because the first stage $cov(Z, Y_1) < 1$ reflects incomplete peer responses to the reform, the IV estimate rescales the reduced-form effect to capture the effect of a one-unit increase in peer employment. As discussed by Welteke and Wrohlich (2019), this scaling can lead to relatively large coefficients even when the underlying reduced-form effects are moderate, because it captures the effect of *all* the peers deciding to retire vs. all deciding to stay (a change from zero to one).

in Panel B) at the workgroup level, and confirm that the results are robust (see Panel C in Table 3.B.7).

Additionally, the coworkers employed with treated peers may differ from those employed with the untreated peers due to different workgroup characteristics. Since the reform was pre-announced, I test whether the peers could also select into different workgroups at the age of 56-57 (the ages when I define the workgroups) based on their treatment status. Adding additional workgroup-level characteristics at the year when peers were 57, such as share of women, share of older workers above 55 years old, share of full-time workers, median wages, number of coworkers, joint tenure of peers and coworkers, establishment AKM fixed effects, number of coworkers, number of female coworkers of each birth cohort (see Panel D in Table 3.B.7) does not alter the baseline results because the confidence intervals still include the coefficients from the baseline specification. Thus, such a selection is unlikely to drive my results because defining the workgroups at the ages when the peer was 56-57 is early enough; moreover, Geyer and Welteke (2021) and Badalyan (2025) do not find anticipatory responses before these ages.

The results are also robust to controlling for aggregated occupation and industry-fixed effects (see Panel E in Table 3.B.7); therefore, I exclude that the effects are confounded by the global financial crisis of 2008-2009 hitting different industries by treatment status. Finally, to show that the effects are not driven by local policies, regional unemployment, or kindergarten availability (as older women might help with care-taking for their grandchildren), I include administrative district ("*kreis*") fixed effects, as such indicators, and corresponding reforms for unemployed and kindergarten expansions vary at the administrative district level. The results are similar to those in the baseline, confirming that the baseline results are not driven by local policies (see Panel F in Table 3.B.7).

Extensive vs. intensive margin of treatment. In Panel B of Table 3.B.8, I explore whether peer effects vary with the number of treated colleagues. While the main specification (Panel A) compares workgroups where all peers are treated versus not treated, it does not account for differences in the number of peers. Since both a group with one peer and one with five peers are coded as "treated", this specification may miss variation in treatment intensity.

To capture this, I interact the treatment indicator (dummy whether all the peers are treated) with the number of peers in the group. The resulting coefficients can be interpreted as the effect of an additional treated peer. As expected, the estimates are smaller in magnitude compared to the binary treatment, since they represent marginal rather than total effects. The direction and significance of the coefficients, however, confirm that peer effects are stronger with an additional peer, consistent with the reinforcement of influence through multiple treated colleagues. This supports the robustness of the main findings while highlighting variation in peer exposure strength.

Group size and linear-in-means assumption. As a robustness check, I examine whether the estimated peer effect varies with the number of peers a coworker has. The identification strategy relies on a linear-in-means structure, i.e., the marginal effect of a change in average peer employment should not depend on group size. If effects differed systematically between coworkers with few versus many peers, the structural invariance assumption would be violated. To test this, I re-estimate the peer effect across peer group size categories and interact the first-stage prediction with indicators for the number of peers. The estimates are stable across specifications, supporting the validity of the linear-in-means model.

To further account for workgroup composition, I conduct subsample analyses by establishment size, workgroup size, and number of coworkers (Panel A-C in Figure 3.A.5). The effects of the peer are somewhat larger in larger establishments and workgroups, but the differences are not statistically significant.[111] Varying the number of coworkers slightly varies the results. In particular, the effects are larger with a higher number of coworkers. This result could be explained by coworkers also influencing each other, and this is the reason why I control for the total number of coworkers in the main specification. Nevertheless, the differences are not significantly different from one another. Examining heterogeneity not only by the number of peers but also by workgroup and establishment size is important, as peer effects could be diluted in larger social contexts or confounded with institutional differences. The stability of estimates across these dimensions strengthens the interpretation of the results as genuine peer effects rather than artifacts of group composition.

Peer definition. In the baseline specification, I define peers as women employed at 56-57 years old, born either one year before or after the reform cutoff, i.e., 1951 (control) and 1952 (treatment) cohorts. I test whether losing this restriction leads to different peer effects. Table 3.B.9 shows the results from several alternative definitions of peers.

One standard assumption in settings that exploit discontinuity is manipulating the running variable. I exclude December 1951 and January 1952 cohorts from my peers from my regressions ("donut hole") and re-estimate the effects. The results in the baseline specification are robust to excluding the peers in the donut hole.

Next, in Panel C of Table 3.B.9, I expand the peer group definition to include peers born up to two years before and after the reform cutoff (1950–1953). This allows for broader peer exposure while maintaining a clear separation from coworkers (born 1953–1957). The estimated effects are similar to the baseline across the first stage, reduced form, and IV specifications. This specification is less preferred than the baseline specification, because the 1951 and 1952 cohort peers are more comparable to one another than peers born in a wider range around the reform cutoff; nevertheless, the results are robust to a wider peer group definition.

Finally, Panel D of Table 3.B.9 relaxes the restriction that all peers must lie on the same side of the reform cutoff. Instead, I allow mixed peer groups and instrument peer employment using the number of treated peers within the workgroup. Because this specification relies on variation in the intensive margin of peer exposure, it requires controlling for the total number

[111] Some confidence intervals (for example, for workgroups with more than 5 peers) are large because the number of observations in these subsamples becomes too small.

of peers and workgroup size to isolate reform-driven variation in peer composition. Without these controls, the instrument partly captures differences in group composition rather than policy exposure, leading to a weak first stage and unstable IV estimates. Once the relevant composition controls are included, the first stage becomes strong, and the estimated peer effects remain statistically significant, although smaller in magnitude than in the baseline specification. This attenuation is expected because variation in partial treatment generates weaker contrasts than the baseline design comparing fully treated and untreated peer groups. Overall, the results confirm that peer effects persist even when peer groups include workers on both sides of the reform cutoff.

Workgroup definition. In the baseline specification, I defined workgroups as 4-digit occupations interacted with establishment, because the workers in more precisely defined occupations are likely to interact more with each other. In this set of robustness checks, I define workgroups as 3-digit occupations interacted with establishments (Panel B in Table 3.B.10), and 2-digit occupations interacted with establishments (Panel C in Table 3.B.10). Neither of these workgroup definitions leads to statistically different results from the baseline specification.

One might expect stronger peer effects within more granular workgroups, where workers interact more closely and information flows are denser, i.e., in 4- vs. 3- vs. or 2-digit occupation level within establishments. Two explanations are plausible. First, more granular workgroups may be precisely those where delayed retirements of peers generate offsetting crowd-out effects, as discussed in subsection 3.6. In such settings, extended employment of some workers could reduce opportunities for others to adjust their own employment trajectories. Second, my baseline sample construction excludes workgroups in which peers fall on both sides of the reform cutoff to avoid contamination from mixed treatment status (see subsubsection 3.3.2). Such restrictions disproportionately exclude larger and more heterogeneous establishments, leaving smaller and more homogeneous workplaces in the sample.[112] Consistent with this, descriptive statistics in Table 3.B.5 show that workgroups are, on average, nearly as large as establishments—about 57 versus 71 workers—indicating substantial overlap between the two levels of aggregation. Thus, the absence of stronger peer effects at more granular definitions likely reflects a combination of crowd-out mechanisms and limited within-firm variation rather than a lack of true peer influence.

Coworkers' definition. The baseline specification defines coworkers as women in birth cohorts 1953-1957, employed in the year when their peer was 57. In this set of robustness tests, I restrict the sample of coworkers to those born starting from 1955, because these coworkers likely observe the retirement behavior between 60 and 62 fully, being at least two years younger than them (Panel B in Table 3.B.11). I further test whether the effects are stronger in the subsample of coworkers whose joint tenure with the peers is larger than two years in Panel C of Table 3.B.11. Finally, to exclude a concern that establishment closures or coworker deaths are wrongly attributed to peer effects, I estimate the analyses on coworkers whose establishments survived at least until 2019, and who did not exit (Panel D of Table 3.B.11) the

[112]A similar argument is used by Welteke and Wrohlich (2019) who uses the same identification for studying parental leave peer effects in Germany.

labor market due to death (Panel E of Table 3.B.11). All three estimates—first stage, reduced form, and peer effects—are positive and significant, and the peer effects confidence intervals of models in both Panels D and E include the original peer effects estimate, highlighting that the effects are not explained by poorly performing establishments or unhealthy coworkers.[113]

To test whether peer effects are driven by a specific cohort of coworkers, Figure 3.A.6 presents subsample analyses by coworker birth cohort. While the peer effects are not significantly different across cohorts, the point estimate for the 1956 cohort is notably larger. One possible explanation is that younger coworkers in this cohort were able to fully observe the retirement behavior of their peers before making their own retirement decisions. In contrast, the 1957 cohort may be more distanced—both temporally and socially—from the peer group, leading to weaker identification with peers and thus smaller peer effects.

Falsification tests. I perform two types of placebo tests: using a placebo cutoff of 1951, i.e., adjacent non-treated cohorts, and using a placebo gender for peers (males).

First, I compare younger coworkers exposed to the 1950 vs. 1951 cohort of females, where both older cohorts were eligible to retire at age 60. For sample construction, see subsubsection 3.3.2.[114] This test shows whether any observed differences are unique to the 1952 reform-treated cohort and not driven by broader age-structure trends in the firm. If the results provided above are attributed to peer effects, the first stage and reduced form effects should be insignificant when I center the reform window around a false cutoff, when there was no change in the retirement age rules. If the first stage estimates turn out to be significant, then there is a difference between a year older vs. a year younger cohort (or workers born in 1951 vs. 1952), which is attributed to the birth cohort effects rather than the reform effects. A significant positive reduced form estimate would imply that the coworkers exposed to older (or 1952) birth cohorts are reacting to the birth cohort's attributes rather than the reform. Sample E (i) in Table 3.B.9 displays the results. The ITT estimates are significant at the 10% level, but they are not robust to the inclusion of covariates. The insignificant peer effects have large coefficients (above the probability of 1) because the confidence intervals are too large, and include 0. Therefore, the placebo test confirms that my main peer estimates are not driven by mechanical correlations or group-level shocks.

Second, I re-estimate the effects using placebo peer gender, i.e., male peers. Opposed to the placebo cohorts used in the test above, this placebo gender test is not as good a candidate for a placebo test, because the statutory retirement ages for males have also been raised

[113] In the baseline specification, I do not make these restrictions because that would lead to conditioning on post-treatment variables, potentially leading to bias. Here, such analyses are performed just to argue that the effects are not driven by bankruptcies of the establishment.

[114] As candidates for placebo cohorts, I take fully untreated cohorts (1950-1951) but not the fully treated cohorts (1952-1953) for several reasons. First, for a good placebo test, I need the two placebo birth cohorts to have the same retirement rules. As shown in Figure 3.2, the normal retirement age of the treated cohorts increased with monthly increments, making them invalid for placebo tests. For a small set of these coworkers, the full retirement age also increased. Second, taking younger cohorts as peers limits the scope to define coworkers, because they have to be even younger to rule out the reflection problem, but the data used for placebo tests is right-censored in 2018.

over these consecutive cohorts, even though to a smaller extent. For example, the reform also removed the unemployment pathway to retirement starting from the 1952 cohort, which mostly affected men (Gudgeon et al., 2023). Sample E (ii) in Table 3.B.9 displays the results. The first stage is somewhat significant when controlling for covariates. There is a small 10% significance for the ITT effect, which is not robust to the excluded covariates. Nevertheless, as explained above, having a small first stage or ITT effect is plausible because the retirement rules were not uniform for these 2 cohorts for men as well. Since both the first stage and the ITT effects are only marginally significant, they result in an insignificant peer effect with very vide confidence intervals, which confirms that my main peer estimates are not driven by some attributes specific to the 1951 vs. 1952 cohorts (which are fixed by gender), but rather by the reform that closed routes to early retirement for women.

3.5 Mechanisms

In the section above, I established the presence of sizable peer effects on old-age employment and retirement delay. Although all coworkers had the same retirement rule, having a peer born before or after the reform cutoff altered their retirement decisions. In this section, I aim to understand the underlying mechanisms behind these peer effects. To explain the presence of peer effects, I test for three main potential mechanisms of peer effects in old-age employment: (1) conformity and social norms, (2) information transmission, and (3) work complementarities.[115] These mechanisms may help understand the relevant policy implications. For example, if information transmission about the reform is the main channel, then education intervention would be a relevant policy instrument. In contrast, an information fair may not help if the main channel is social norms in old-age employment.

3.5.1 Conformity and social norms

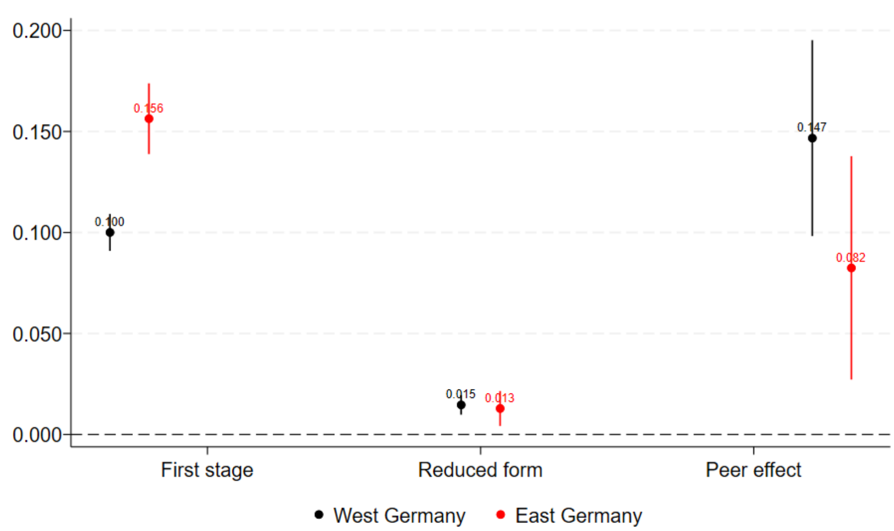
Gender role attitudes are particularly shaped through interactions with peers (Bramoullé, Djebbari, and Fortin, 2020). Theoretical frameworks from sociology and psychology emphasize that individuals are more likely to form connections with—and be influenced by—those who share similar characteristics. Theories of homophily (McPherson, Smith-Lovin, and Cook, 2001) and similarity-based attraction (Byrne, 1971) suggest that people gravitate toward

[115] Peer effects literature in other settings, such as spousal retirement coordination, often discusses leisure complementarities as a channel García-Miralles and Leganza (2024). Leisure complementarities imply that coworkers prefer to retire in the same calendar month to spend more time together outside the firm. I do not analyze this channel, because leisure complementarities are less likely among coworkers than family members, and it is difficult to test for them in social security data without accompanying survey evidence on leisure activities. Testing for this mechanism through the "same date" as in (García-Miralles and Leganza, 2024) who analyzed the couples coordination of retirement ages is difficult in the context of coworker peer effects, as the joint exit does not necessarily imply joint leisure in the context of coworkers, and could, in contrast, imply a different channel- work complementarities, which I discuss later in the text. I could test for this mechanism by checking if, after the reform, the workplace retirees move closer to each other; however, the variables of the place of living are not detailed enough to allow me sufficient heterogeneity in this dimension.

peers who resemble them in age, gender, or background. In parallel, social identity and categorization theories propose that individuals adopt behavioral norms consistent with the social groups they identify with (Tajfel, 1981; Oakes et al., 1987). Applied to the workplace, these insights imply that women nearing retirement age may look to the behavior of similarly aged female coworkers when making their own employment decisions.

Geographical segregation in norms. East and West Germany provide a natural setting to test this conformity and social norms channel in old-age employment. Historical labor market institutions and cultural expectations diverged sharply before reunification: East German women had higher lifetime labor force participation, more continuous work histories, and stronger norms of full-time employment. If peer effects operate partly through shifting norms, they should be more pronounced where norms are *in flux*—namely, in West Germany. According to Welteke and Wrohlich (2019), conformity is expected in places with changing social norms. I therefore split my sample by the current workplace location in East versus West federal states (classifying Berlin as East).[116]

Figure 3.6: Subsample analyses by conformity measures (East vs. West Germany)



Notes: Coefficient plots. The columns correspond to the first stage, reduced form (ITT) and 2SLS (peer effect) regressions for employment at 62. Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers (all measured when the peer was 57), establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. The points represent the estimated coefficients, and the bars represent the 95% confidence intervals. The subsample analyses are performed by the workgroup locations in East and West Germany.

Figure 3.6 shows that the first stage estimates are larger for East German women. These findings are in line with the findings in Geyer and Welteke (2021), who argue that East German women are more likely to fulfill the requirement for old-age pension (15 years of

[116]Given data limitations, I cannot observe workers' location prior to 1991; thus, classification is based on workplace location in the observation window rather than origins.

contribution years, ten of which are accumulated after the age of 40, see subsection 3.2 for details). My results are significantly higher for East German women by a 5.6 p.p. difference. The figure also confirms that where the reform substantially changed the social norms regarding women's old-age employment, the peer effects were larger. I find that in West Germany, there are 14.7 p.p. peer effects, compared to only 8.2 p.p. in the East. The East–West comparison likely reflects differences in social norm evolution: West German women, historically more aligned with the breadwinner model, appear to shift more toward peer-based coordination in late-career employment, whereas East German women had long experienced higher female labor force participation. This interpretation, however, warrants caution. Historical differences in social networks—such as the lower prevalence of "weak ties" in the former GDR due to surveillance concerns (Völker and Flap, 2001)—may have also limited the diffusion of peer effects in East Germany.[117] Nevertheless, because peer groups are constructed in 2008–2009 (when the women were 57), nearly two decades after reunification, much of the observed difference is plausibly driven by social norms in old-age employment. Moreover, the findings remain robust when controlling for detailed occupation and industry fixed effects, suggesting they are not merely a reflection of industrial segregation.

3.5.2 Information channel

Information about the implications of retirement decisions play an important role for people because their retirement is irreversible (Boeri and van Ours, 2021). The actions of peers may be a valuable source of information. I distinguish between two types of information that are important in the coworker-peer effects context: (1) information about the (new) retirement rules, and (2) the information provision channels about the costs and benefits of delaying retirement.

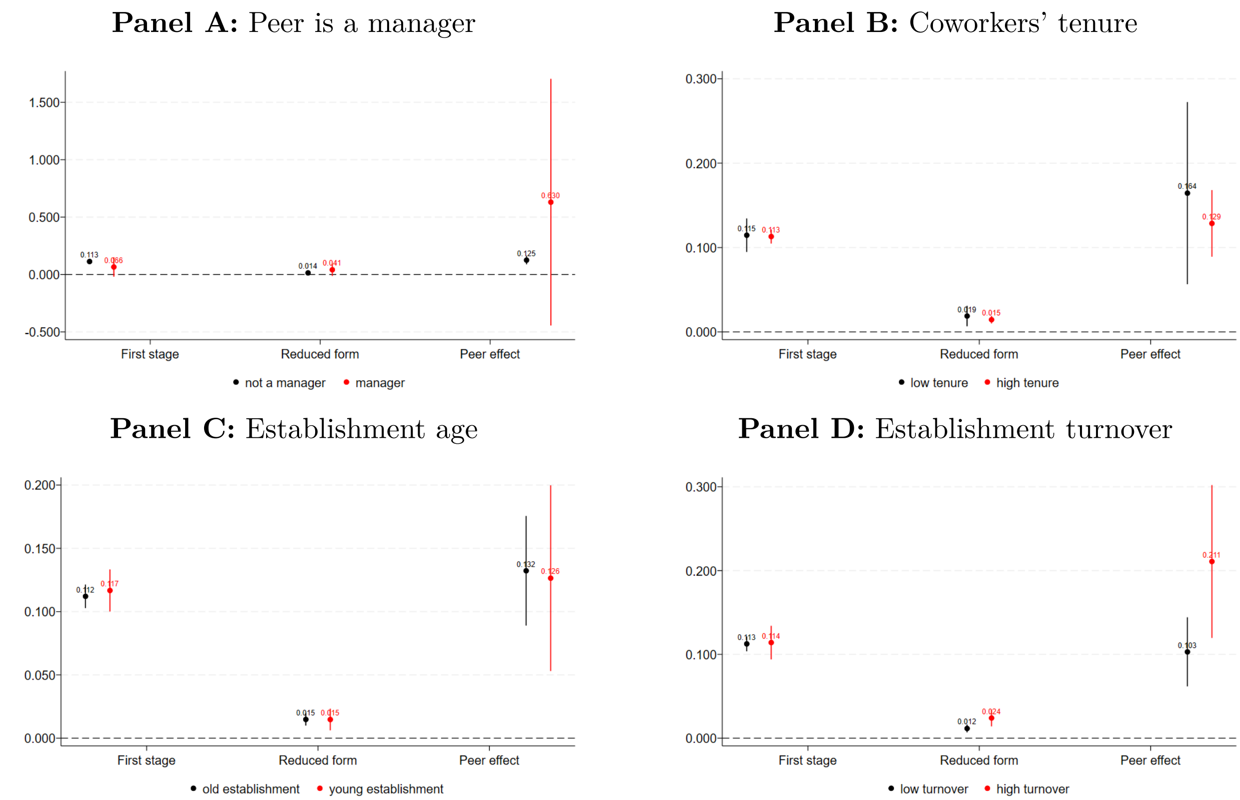
Information transmission about the program. Bottazzi, Jappelli, and Padula (2006) argue that adapting to the new retirement rules and learning about them is a slow process. For example, in the US, there is bunching at the pre-reform retirement ages (Deshpande, Fadlon, and Gray, 2024). Literature has shown that providing information about future pensions or pension accrual due to postponed retirement (Chan and Stevens, 2008), and social security provisions (Liebman and Luttmer, 2015) can lead to an increase in old-age employment. Because the retirement rules in Germany are universal, and there is an online platform to compute the retirement rules for individual cases, it is unlikely that coworkers learn from peers about the retirement pathways and changes to them. Therefore, I abstract from attributing the peer effects to such an information channel.

Information transmission about the costs and benefits of delaying retirement. Providing information about the costs and benefits of retiring at a certain age differs from providing information about the program itself. Even if the retirement rules are known to the coworkers, decisions on when to retire could be influenced by information about the consequences of retiring at a certain age. For example, some career-related uncertainties,

[117] I thank Kerstin Ostermann for suggesting this perspective.

such as uncertainties related to employer reactions, could be decreased by observing the peers' retirement decisions. According to social learning in networks (Goyal, 2011), and empirical findings on peer effects in other contexts (Dahl, Løken, and Mogstad, 2014; Welteke and Wrohlich, 2019), the peer effects are expected to be larger for the coworkers who experience greater career-related uncertainty.

Figure 3.7: Subsample analyses by information channels



Notes: Coefficient plots. The columns correspond to the first stage, reduced form (ITT) and 2SLS (peer effect) regressions for employment at 62. Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers (all measured when the peer was 57), establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. The points represent the estimated coefficients, and the bars represent the 95% confidence intervals. The subsample analyses are performed by hierarchical position in the job (**Panel A**), coworkers' tenure (**Panel B**), establishment age, where young establishments are defined as those with at most ten years of existence (**Panel C**), and establishment turnover defined as turnover rate (share of the difference in hiring and separations in the establishment) being above or below 30% (**Panel D**).

I test for this channel by performing heterogeneity analyses on several measures, showing career uncertainty: the tenure of coworkers, the hierarchical position of the peer, the age of the establishment, and establishment turnover. High career uncertainty is expected to be highly correlated with the information transmission channel; therefore, if this channel dominates, short-term tenured coworkers, and managerial positions of peers should be associated with higher peer effects, because managers often possess deep institutional knowledge, leadership experience, and relationship-specific investments that are difficult to transfer across firms, and are valuable particularly to less tenured workers. In addition, young or high turnover establishments, where there is larger uncertainty and job stability, should be associated with larger peer effects. I follow existing literature on peer effects in various contexts (see Dahl, Løken, and Mogstad (2014) and Welteke and Wrohlich (2019) among others) to create these measures of uncertainty to test for the information channel.[118]

I define a coworker's tenure as low if she has been employed for less than two years (730 days). Managerial occupations are defined based on the last two digits of 5-digit occupational codes.[119] I define young establishments as those that have operated for less than ten years. The establishments with high turnover are those that experienced over 30% turnover rate, defined as the difference in the number of hires and separations in the total workforce. All the variables are defined in the year when the peers were 57 years old.

Results, displayed in Figure 3.7, reveal that there are larger peer effects in establishments with larger turnover, while the rest of the measures do not display significant differences. Therefore, there is limited support for this channel. In combination with the strong suggestive evidence for the social norms mechanism, this finding could reveal that there is a social learning model, where information provided by peers reduces social, and to some extent, career-related uncertainty, similar to the findings on parental leave by Welteke and Wrohlich (2019).

3.5.3 Work complementarities

Beyond the standard peer effects channels of conformity and information transmission, I consider a novel mechanism relevant in the context of old-age employment: work complementarities, which capture productive complementarities between older and slightly younger coworkers. Opposed to leisure complementarities, work complementarities occur if the retirement delay of one worker prevents her coworkers from leaving the workforce. In settings where job tasks are interdependent, older workers may possess firm- or task-specific knowledge that enhances the productivity of their younger peers. When pension reforms delay the retirement of older peers, these complementarities can generate positive spillovers on the employment of their younger coworkers, either by increasing job satisfaction, team output, or by inducing firms to retain matched coworker pairs longer. This mechanism implies that peer exposure can increase coworker employment not through imitation but through improved workplace viability.

[118]It is important to note that these measures are only proxies for the uncertainty channels. One could argue that some variables related to uncertainty, such as tenure or establishment age, relate to familiarity of coworkers and proximity rather than the information channel. Nevertheless, these proxies are used in other papers on peer effects (Dahl, Løken, and Mogstad, 2014; Welteke, 2015), and I am cautious in my interpretations of these subsample analyses as a test of specific mechanisms. I discuss this further in subsection 3.6.

[119]I create a variable showing managerial or supervisory status based on the last two digits of the 5-digit occupations. I pool the supervisors and managers into the dummy variable *manager*. I thank Philipp vom Berge for information on how to define managers in social security records.

The literature on peer effects on performance highlights the influence of connections and peers in exerting effort. Kato and Shu (2009) provide evidence for peer effects in exerting effort within the hierarchies, while Bandiera, Barankay, and Rasul (2009) show how social connections in the workplace can operate across hierarchies. Such peer effects could also operate within the context of delayed retirements. In particular, this is likely to hold in connected teams, where the peers share workgroup-specific human capital (Bartel et al., 2014; Jaravel, Petkova, and Bell, 2018; Jäger and Heining, 2022); therefore, the delay of retirement of one worker would lead to the delayed retirement of his coworkers.

To explore this mechanism empirically, I exploit variation across occupations in the degree of within-establishment task connectedness, i.e., interaction dependence between coworkers. In addition, I test whether the workgroups that are performing occupations that are not easily substitutable externally from the commuting zones are more likely to lead to higher peer effects due to the loss of productivity because of the interruption of peer connections. Finally, I classify the occupations as "bottleneck", and industries as tradable vs. untradable, because employees in such environments that are difficult to substitute for could lead to larger peer effects due to the disruption of work.

Measure group 1: main tasks of the workgroup. In my main specification, I define peer groups at the four-digit occupation-by-establishment level, capturing settings in which workers likely perform closely related tasks. Under the assumption that task complementarity is higher within narrowly defined occupations, stronger peer effects in these cells would be consistent with a complementarity channel. Although direct measures of teamwork intensity are unavailable in administrative data, the heterogeneity in peer effects is tested by the main tasks implemented in a given workgroup occupation.

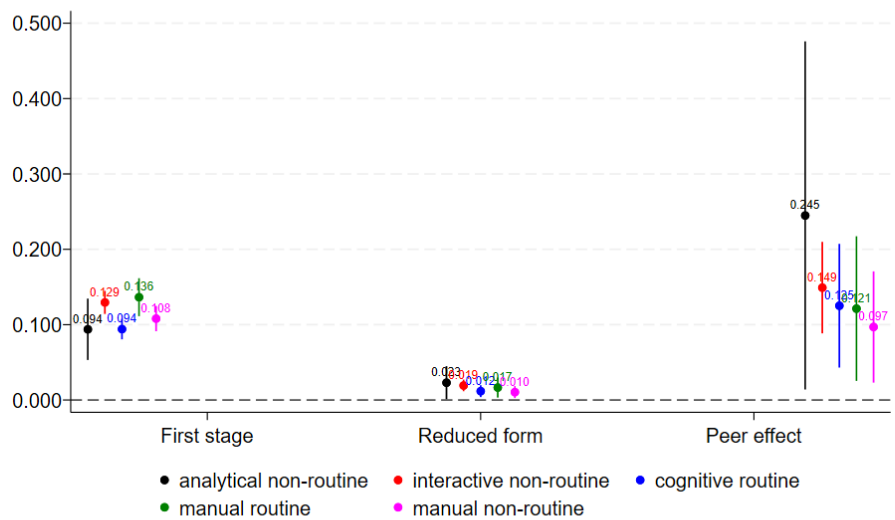
Following the classification in Dengler, Matthes, and Paulus (2014), I assign occupations to categories along two key dimensions: (i) the predominant skill content—analytical, interactive, cognitive, or manual tasks; and (ii) the degree of routineness—routine versus nonroutine.[120] Occupations differ in the extent to which coworkers' actions affect individual employment decisions. Following the task classification in Dengler, Matthes, and Paulus (2014), I expect complementarities to be strongest in *interactive non-routine* and *analytical non-routine* jobs, where output relies on frequent information exchange, coordination, and mutual learning. In such settings, the continued employment of peers may directly influence one's own decision to remain, by sustaining team productivity and preserving workgroup-specific knowledge. Even certain *cognitive routine* tasks, when embedded in coordinated workflows, can exhibit similar spillovers. In contrast, *manual routine* and *manual non-routine* occupations—particularly those performed independently—are less likely to generate strong complementarities.

Figure 3.8 displays the results. I find no statistically significant difference in peer effects across the task types in workgroups. Although point estimates suggest slightly larger effects among analytical non-routine, interactive non-routine, and cognitive routine tasks, the standard

[120] The task classification is matched to the main dataset using three-digit occupation codes. Since baseline workgroups are defined at the four-digit occupation level, the main task classification is, therefore, more aggregated than the workgroup definition.

errors are large, and differences are not statistically significant. These findings suggest that the influence of peer retirement behavior can operate broadly in demographic and workplace environments.

Figure 3.8: Subsample analyses by main tasks in the workgroups



Notes: Coefficient plots. The columns correspond to the first stage, reduced form (ITT) and 2SLS (peer effect) regressions for employment at 62. Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers (all measured when the peer was 57), establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. The points represent the estimated coefficients, and the bars represent the 95% confidence intervals. The subsample analyses are performed by five main task categories for a given occupation.

Measure group 2: external substitutability. In addition, I examine heterogeneity in peer effects using three proxies for job substitutability: external labor market thickness (ELMT), industry tradability, and bottleneck occupations. These measures capture the degree of work complementarity versus substitutability. In thin external labor markets, bottleneck occupations, and non-tradable sectors, internal workers are less easily replaced by external hires. Consequently, disruptions within such groups are more likely to entail substantial productivity and profit losses through higher inter-dependence of such workers, amplifying the potential for peer effects. These measures and hypotheses are motivated by the worker substitutability literature (Badalyan, 2025; Ginja, Karimi, and Xiao, 2023; Jäger and Heining, 2022; Huebener et al., 2024; Schmutte and Skira, 2023), and are described below.

External labor market thickness captures how easily workers in a given industry or occupation can be substituted from the local commuting zone. I create 141 local labor markets based on high within-region and low between-region commuting for work (Kosfeld and Werner, 2012). Next, I create an index that captures the local share of occupation (or industry) employment over the national share of occupation (or industry) employment, following Ginja, Karimi, and Xiao (2023). In the main specification, I compute the index by counting female employment only, to take into account the industry and occupation-based segregation by gender in Germany.[121] Figure 3.A.7 displays the map of Germany by ELMT index values for an example industry and occupation. For example, workgroups specializing in the industry of monetary intermediation in Bamberg would have more difficulties replacing their workers than those in Göttingen.

Bottleneck occupation indicators. I merge my data with the bottleneck occupation indicators from the Bundesagentur für Arbeit, mapped to 4-digit KldB 2010 codes.[122] I utilize the earliest available classifications at the national level; that is, in 2011-2012, and classify an occupation as a bottleneck if it was a bottleneck occupation in 2011 or 2012.

Industry tradability. In tradable industries, it is easier to substitute workers through outsourcing than in non-tradable industries (Drenik et al., 2023). I classify the industries by tradability following Gregory, Salomons, and Zierahn (2022).[123]

Figure 3.9) displays the results. While industry tradability and occupational ELMT categories do not produce significantly different peer effects, the effects are somewhat larger in thin ELMT industries, where external substitutability is lower. The sample size is small for bottleneck occupations because there were very few in 2011-2012, and the standard errors are too large to draw firm conclusions.

Measure group 3: old-age employment in the same workgroup. To further explore the role of workplace interactions, I re-estimate the effects focusing on employment within the same workgroup as in the exposure year (when the peer was 57). This restriction likely captures settings with more frequent interaction and task interdependence, where peer influences may be stronger.

The results (last row of Table 3.B.6) show a smaller first stage compared to the full sample—consistent with a more selected subset—but a substantially larger reduced-form effect. As a result, the implied peer effect estimates are more than twice as large as in the baseline specification. This pattern suggests that sustained exposure within the same workgroup amplifies peer influence on retirement behavior. While this finding is consistent with the presence of work complementarities, it should be interpreted as suggestive rather than conclusive evidence of this channel.

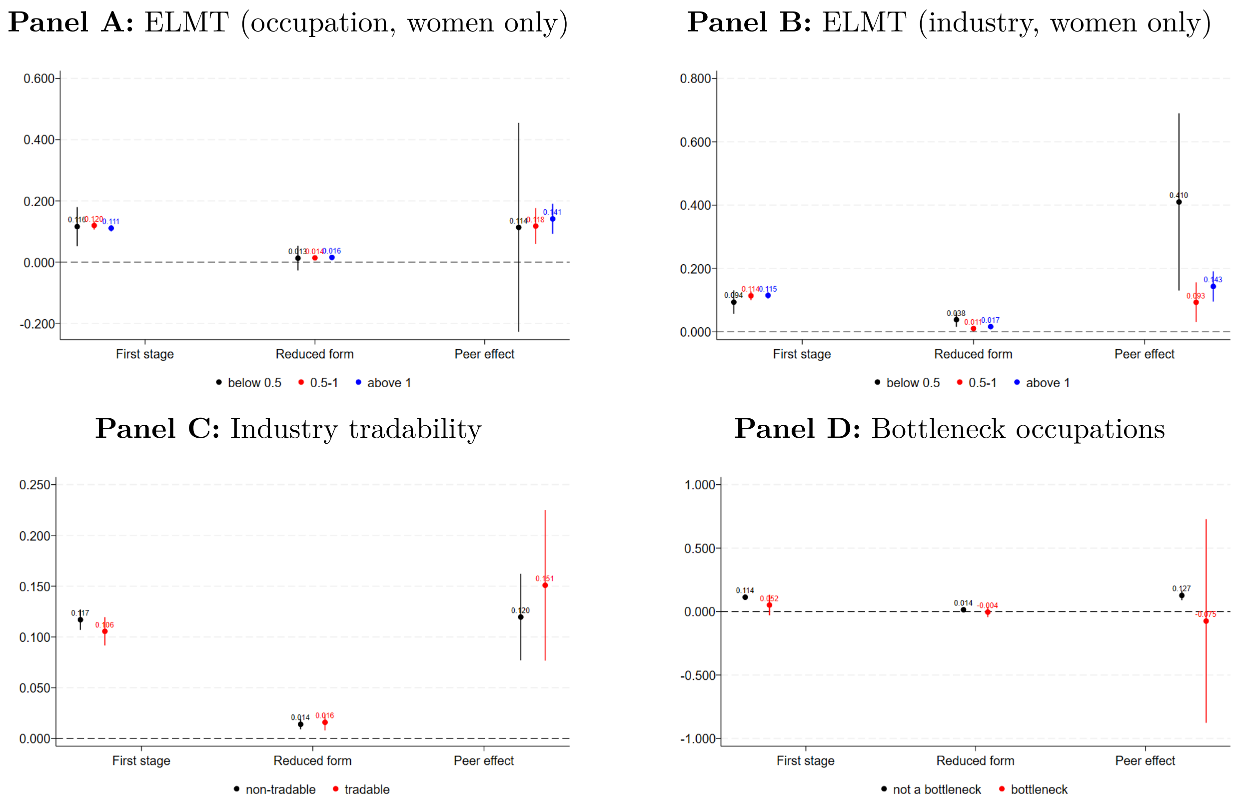
Overall, while I cannot fully rule out work complementarity mechanisms, I find no consistent pattern suggesting they systematically drive the results. Some proxies for this mechanism do not display significant differences in ITT or peer effects, while the others, such as employment

[121]For industry or occupation k and commuting zone c (141 local labor markets), the index is computed as follows: $ELMT_{kc} = \dfrac{(NW_{kc}/NW_c)}{(NW_k/NW)}$, where NW is the number of women.

[122]The classification is based on six indicators, including vacancy duration, job seeker ratios, unemployment specialization, and wage dynamics. I thank Hannah Illing and Anton Klaus for their help with the data.

[123]*Tradable industries* are: Mining (WZ08: B); Manufacturing (WZ08: C); Electricity, water supply (WZ08: D, E); Transport, storage (WZ08: H); Financial services (WZ08: K); Real estate (WZ08: L); Agriculture (WZ08: A); Information and communication (WZ08: J); Scientific and technical services (WZ08: M). I thank Duncan Roth for the help with the data.

Figure 3.9: Subsample analyses by complementarities and substitutabilities



Notes: Coefficient plots. The columns correspond to the first stage, reduced form (ITT) and 2SLS (peer effect) regressions for employment at 62. Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers (all measured when the peer was 57), establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. The points represent the estimated coefficients, and the bars represent the 95% confidence intervals. The subsample analyses are performed by *external labor market thickness (ELMT)*, based on the commuting zone at most half as concentrated in a given occupation (**Panel A**) or industry (**Panel B**) relative to the country-level ($ELMT < 0.5$, see Figure 3.A.7 for examples of ELMT-s in 141 local labor markets), or at least half as concentrated but less concentrated than the country-level ($0.5 < ELMT < 1$), and at least as concentrated as the country-level concentration ($ELMT > 1$). **Panel C** displays results by tradable sectors, and **Panel D** by bottleneck occupations.

in the same workgroup, do. Using the same reform, Badalyan (2025) shows that work substitutability and complementarity mechanisms influence the decision of who to employ in response to the retirement reform (direct effect, similar to the first stage in this paper). However, these measures are of less importance for peer effects behavior. The reform thus operated through a dual mechanism—firms retained workers due to substitutability incentives (Badalyan, 2025), while peers reinforced employment through conformity and social norms regarding old-age employment among women. I do not find strong evidence of worker complementarities explaining peer effects for this reform. My approach for testing for the work complementarities mechanism is novel in this literature; specifically, I follow (Badalyan, 2025,?; Jäger and Heining, 2022; Huebener et al., 2024) to create detailed proxies for work complementarities. This mechanism could also be more important for non-gender-specific reform, but in the context of the gender-specific retirement reform that I study in this paper, it is natural that social norms and conformity play a greater role in explaining the peer effects than do work complementarities.

3.6 Discussion

This section addresses three limitations, reconciles the findings with the literature, and quantifies the social multiplier of the reform.

Crowd-out effects. A key concern is that delayed retirement may reduce hiring or promotion opportunities for coworkers (Badalyan, 2025; Bianchi et al., 2023; Ferrari, Kabátek, and Morris, 2023), introducing a labor demand channel that may bias peer effect estimates downward or affect the exclusion restriction (see subsection 3.3). This channel operates alongside the peer-driven labor supply response and works in the opposite direction.

Some patterns in the data are consistent with the coexistence of these channels. In particular, defining workgroups at more granular occupational levels does not increase estimated peer effects, despite closer interactions. This is consistent with stronger substitutability within occupations (Jäger and Heining, 2022), where demand and supply forces may partially offset each other. Evidence from Badalyan (2025) supports this interpretation: promotions decline by 0.109 per additional older worker within 3-digit occupations, with no effects across job cells, indicating that crowding-out operates within narrowly defined groups. In contrast, this paper identifies positive peer effects using 4-digit workgroups for the reform under study. Although the outcomes differ, both papers document sizable effects within similar occupational settings, with opposite signs, suggesting that the two channels may partially offset each other in more granular workgroups.

Several features of the empirical design mitigate this concern. The analysis focuses on coworkers already employed when peers are age 57, which limits hiring margins, and includes firm, industry, and regional controls that absorb common labor demand shocks. Results are also stable when the sample is restricted to coworkers who remain in the same workgroup. Because the crowd-out channel would reduce employment while peer effects increase it, the estimates likely capture their net effect and represent a conservative lower bound on peer-driven labor supply responses.

Outcome variable and decomposition. The main outcome—employment at age 62—captures whether the last observed labor market activity occurred at or after 62, and may include unemployment. Since the data track only establishment employment histories, a full decomposition is not possible. However, Table 3.B.6 shows that about 75% of this variable reflects continued employment at the same firm, with the remainder spread across other workgroups or unemployment. Importantly, unemployment is unlikely to drive results, as Geyer and Welteke (2021) find little substitution into unemployment. Peer effects are even stronger when restricting outcomes to same-firm employment, which reinforces that the measure primarily captures genuine employment spillovers.

Descriptive evidence further supports this. In the pre-reform 1951 cohort, women aged 60–62 were distributed across unemployment (6.2%), disability (8.3%), and inactivity (4.8%). After the 1952 cutoff, unemployment and inactivity rose modestly due to the passive continuation of prior states, while disability remained unchanged. My ITT results suggest that unemployment at ages 60–62 decreased by 1.6 percentage points when peers delayed retirement, corresponding to an 18.8 percentage point rise in labor force activity—a substantial reduction in inactivity. Compared with other social insurance contexts, such as peer spillovers in parental leave (Welteke and Wrohlich, 2019), the effects are smaller but still substantial given the smaller ITT.

In Table 3.B.12, to clarify what my baseline outcome—labor market (LM) activity—captures, I compare it to a narrower measure of employment at the same firm. The LM activity variable records whether an individual remains active in the labor market, encompassing both continued employment with the 2007 employer and other labor market states such as re-employment elsewhere or short unemployment. Among peers, the control mean of LM activity at ages 60–62 is 79.7%, compared to only 23.5% for same-firm employment, implying that many individuals remain active but not necessarily with their original employer. Yet the reform effects are of similar absolute magnitude across both definitions (8.3 vs. 6.2 percentage points), corresponding to relative increases of about 10% and 26%, respectively. This suggests that the policy primarily delays labor market exit rather than merely reallocating workers across firms or into brief periods of non-employment.

For coworkers, a similar pattern emerges. Although the control mean is higher for LM activity (82.8%) than for same-firm employment (24.2%), the treatment effects at ages 60–62 are 1.6 and 2.7 percentage points, respectively—equivalent to relative increases of about 2% and 11%. This pattern indicates that peers' delayed retirement primarily strengthens coworkers' attachment to the labor market—and particularly to the same firm—rather than inducing broader job mobility. Peer effects are thus more pronounced when coworkers observe peers' retirement decisions directly within the same firm. Taken together, these results suggest that my main outcome variable, LM activity, captures extended employment—whether at the same or another firm—rather than transitions into non-employment or inactivity. Peer effects, therefore, reflect genuine delays in labor market exit, even when measured through this broader definition of activity.

Mechanisms and social multiplier. Peer effects may operate through information channels, work complementarities, and social norms. While the subsample analyses do not allow for a clean separation of these mechanisms, the results are most consistent with a role for social interactions in shaping behavior. In particular, the comparison between East and West Germany suggests that norms matter: peer effects are substantially larger in West Germany, where traditional gender roles have historically been more persistent and where the reform represented a more pronounced shift in expectations about women's employment at older ages.

At the same time, the relevant norms in this setting are likely not limited to gender roles, but also reflect broader retirement norms. Over the past decades, labor force participation at ages 60–65 has increased markedly for both men and women, indicating a general shift in expectations about late-career employment. In this context, peer effects may capture how individuals update their beliefs about appropriate retirement timing based on the behavior of coworkers, rather than solely changes in gender-specific roles.

Compared to settings such as parental leave, where peer effects have been shown to strongly reflect social norms, such as in Welteke and Wrohlich (2019), the norm component in retirement decisions may be less salient or slower to adjust, as these decisions occur later in the life cycle and are shaped by longer employment histories and institutional constraints. Accordingly, the evidence in this paper should be interpreted as suggestive of social and retirement norm channels, rather than definitive proof of a specific mechanism.

To assess aggregate importance, I compute a behavioral social multiplier: the ratio of total (direct plus peer) observed employment effects to direct effects alone, which could lower the marginal value of public funds (MVPF) of retirement age increase (Hendren and Sprung-Keyser, 2020). With a direct reform effect of 11.6 p.p. and a peer effect of 1.5 p.p., the multiplier equals 1.13. Thus, for every woman employed at an older age, an additional 0.13 coworker women remained employed due to peers.

Regional heterogeneity. Effects differ across regions. Direct responses are larger in the East (15.6 p.p.) than in the West (10.0 p.p.), consistent with East Germany's longer history of female employment. Yet peer effects are stronger in the West (14.7 vs. 8.2 p.p.), suggesting that West German workplaces are more susceptible to norm-based change. This pattern points to peer influences as a mechanism for convergence: even if direct responses are weaker, norms can shift collectively, allowing slower-responding regions to catch up.

Taken together, the discussion highlights that retirement reforms do not operate in isolation. Peer interactions amplify their reach, help shift workplace norms, and may facilitate regional convergence in old-age employment behaviors.

3.7 Conclusion

Institutional rules and financial incentives alone cannot fully explain retirement behavior. Even when reforms raise retirement ages, some women remain inactive at the targeted age of the reform, due to prior pension eligibility, health, caregiving, self-employment, or joint leisure with spouses. A small fraction may also be on UI or DI benefits. This paper highlights a behavioral factor: peer effects in retirement decisions. Using a German reform that changed early retirement eligibility, I show that coworkers' retirement choices—especially within close workgroups—substantially influence individual employment at older ages. Evidence points mainly to conformity and social norms, rather than information transmission or work complementarities.

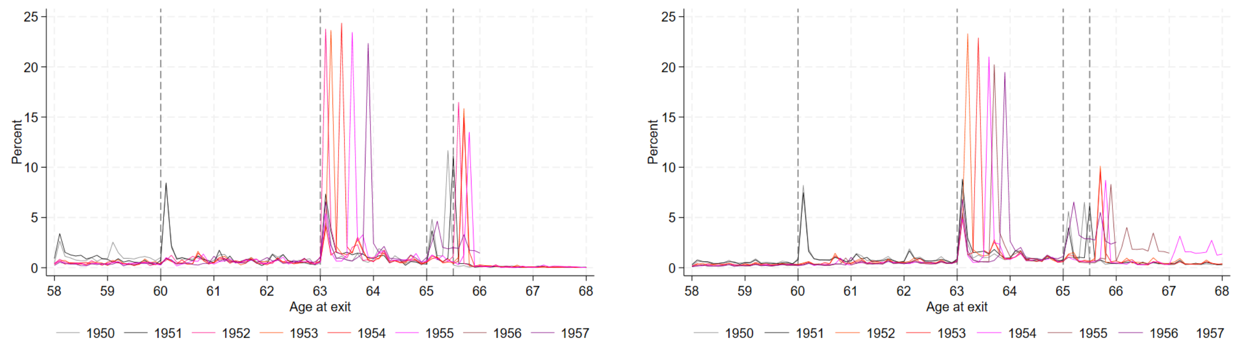
Retirement decisions are shaped by the workplace social context. As women's labor force participation gradually extends beyond age 60 in many countries, peer influences regarding old-age employment can shift norms from non-employment to employment at the reform-targeted age. Regional differences between East and West Germany suggest that preexisting social norms condition how new behaviors spread, with peers potentially contributing to convergence—or persistence—of regional patterns.

The findings have two key implications. First, policy success depends not only on financial incentives but also on shaping or leveraging workplace norms. Second, group-level interventions—such as team-based retention programs or role-model strategies—may achieve behavioral change more efficiently than individual incentives. Future research could explore similar reforms in other countries and examine how the shift to gender-neutral retirement rules spills over into women's education decisions, intra-family bargaining, and household gender norms, such as division of caregiving responsibilities. It would also be interesting to test for the work complementarities channels introduced in this paper in a male-specific retirement reform, as they might be more prominent there.

3.A Appendix: Figures

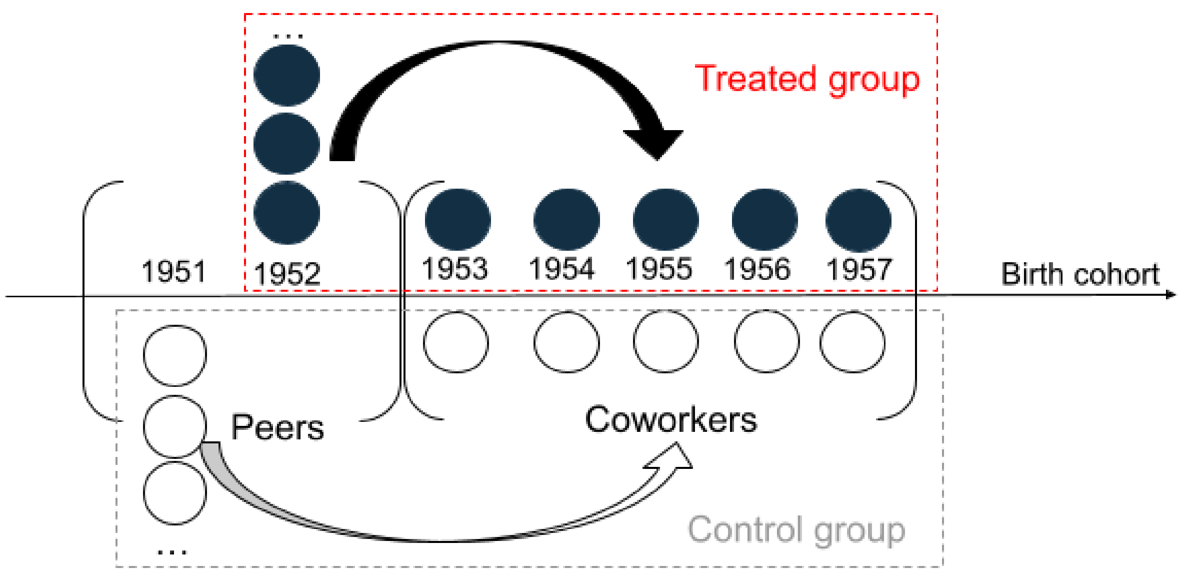
Figure 3.A.1: Retirement age distribution by birth cohorts

Panel A: 2% random sample | **Panel B:** Sample in this study



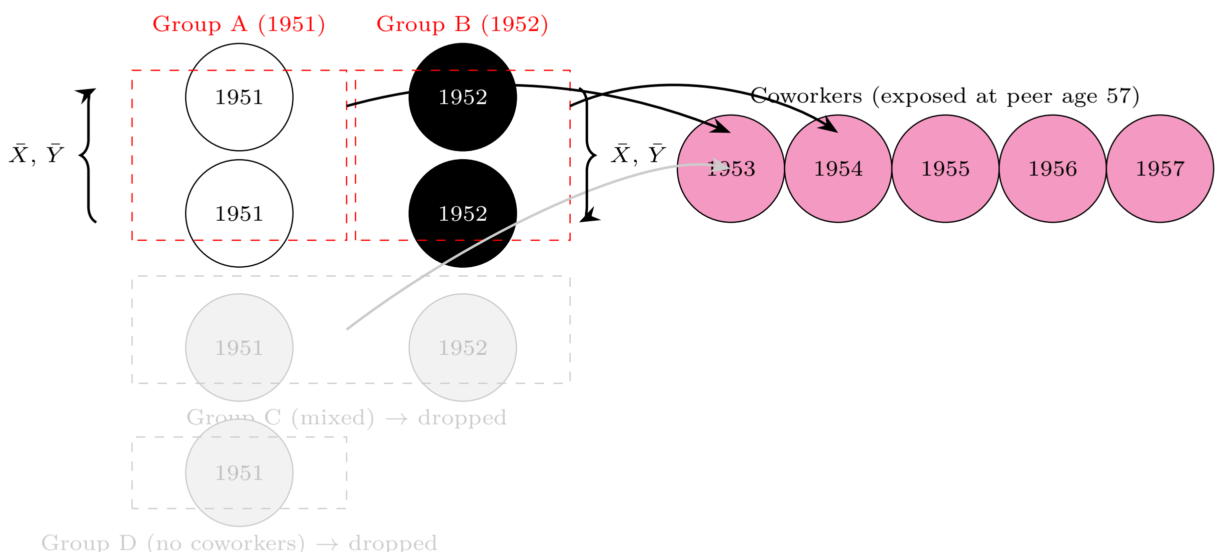
Notes: This graph shows the distribution of retirement ages of the cohort 1950-1957, employed at 57 years old. The graph is generated from the 2% random sample of the population of IEB records (**Panel A**), and the sample used in this paper (**Panel B**). The gray and black lines correspond to the 1950 and 1951 cohorts of women– the cohorts that allowed claiming pensions as early as 60. The rest of the lines demonstrated the retirement age distribution for the cohorts 1952-1957, for whom the retirement age was raised.

Figure 3.A.3: Identification strategy



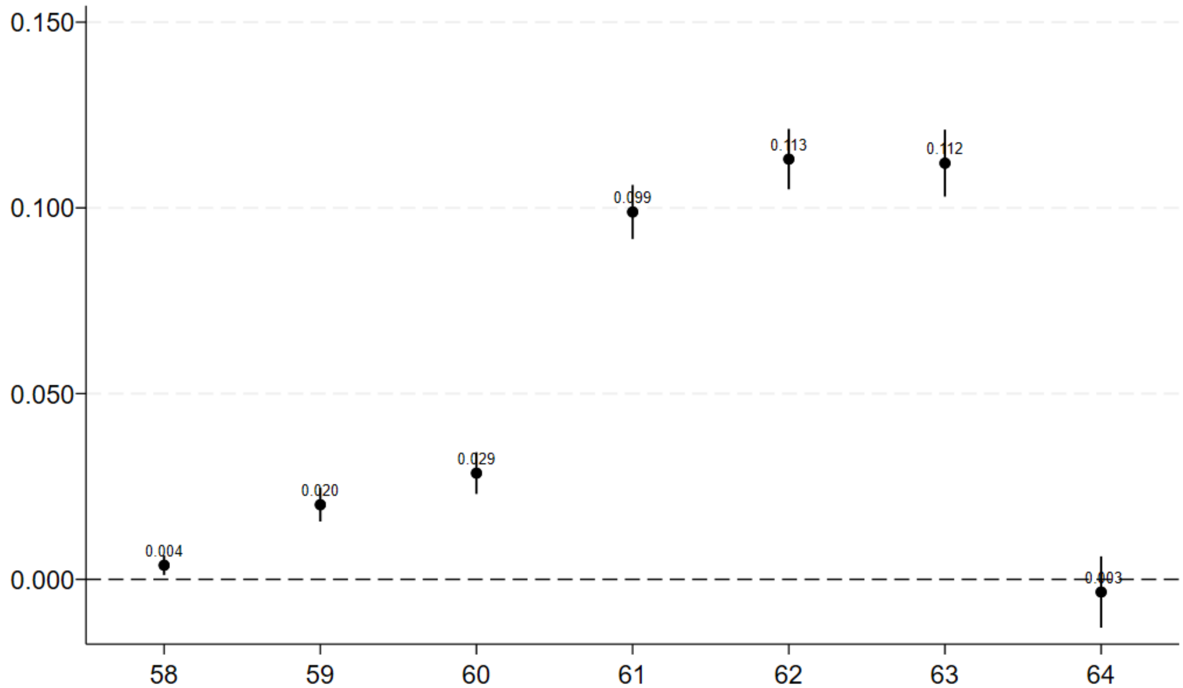
Notes: The figure illustrates the assignment of workers into groups. Peers are women born around the 1952 cutoff: those born in 1951 could still retire at 60, while those born in 1952 faced a higher ERA. Their retirement behavior may, in turn, influence their coworkers, who are younger cohorts in the same workgroup. Treated peers (1952) and their coworkers are shaded black; untreated peers (1951) are shaded white.

Figure 3.A.2: Sample construction: visual walkthrough



Notes: This diagram illustrates the sample construction for baseline regressions. Peers are defined as those employed at the age of 56-57, belonging to the 1951 (control group, colored in black) or 1952 (treatment group, colored in white). I allow peers to be in either the control group (**Group A**) or the treatment group (**Group B**). I exclude the mixed workgroups, where peers appear on both sides of the cutoff (**Group C**, shaded out in gray). From the remaining workgroups, I remove those workgroups (**Group D**, shaded out in gray) where the peers have no coworkers (defined as women of birth cohorts 1953-1957 who were employed in the same workgroup when the peer was 57 years old, colored in pink). The sample size after each restriction can be found in Table 3.B.3.

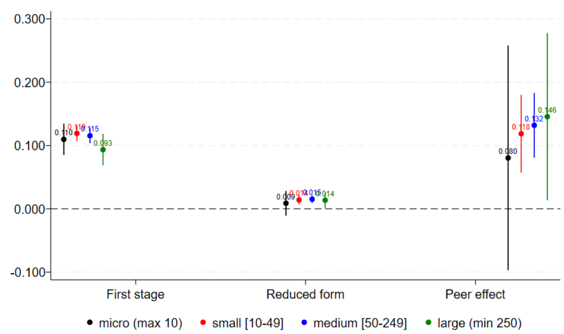
Figure 3.A.4: First stage regressions by age



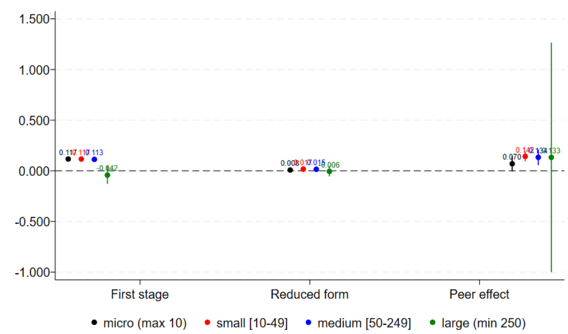
Notes: Coefficient plots. The columns correspond to the first stage regressions for employment by age of peers. Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers (all measured when the peer was 57), establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. The points represent the estimated coefficients at each age, and the bars represent the 95% confidence intervals. I control for the coworker and average peer education, wages, West residence, establishment, and workgroup sizes, number of peers, and coworkers, all measured when the peer was 57.

Figure 3.A.5: Subsample analyses by establishment and workgroup size, number of peers and coworkers

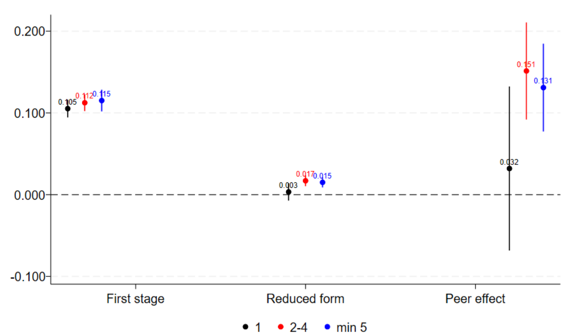
Panel A: Establishment size categories



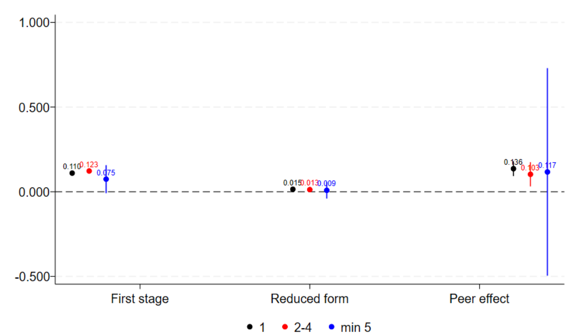
Panel B: Workgroup size categories



Panel C: Number of coworkers

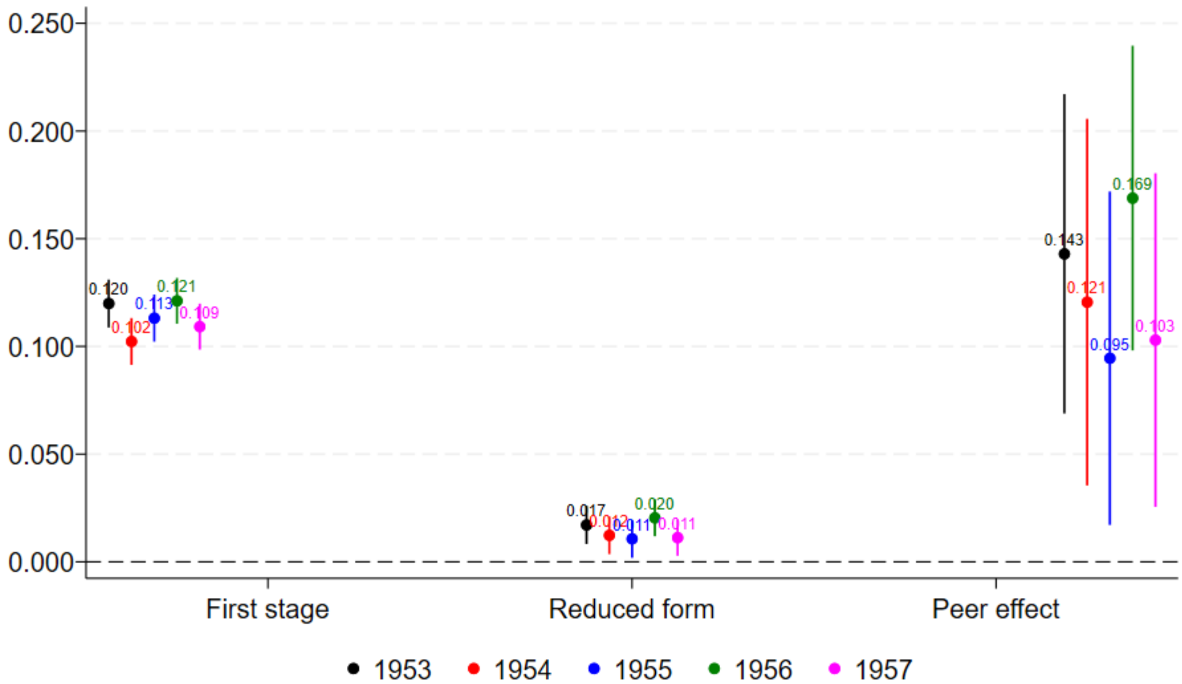


Panel D: Number of peers



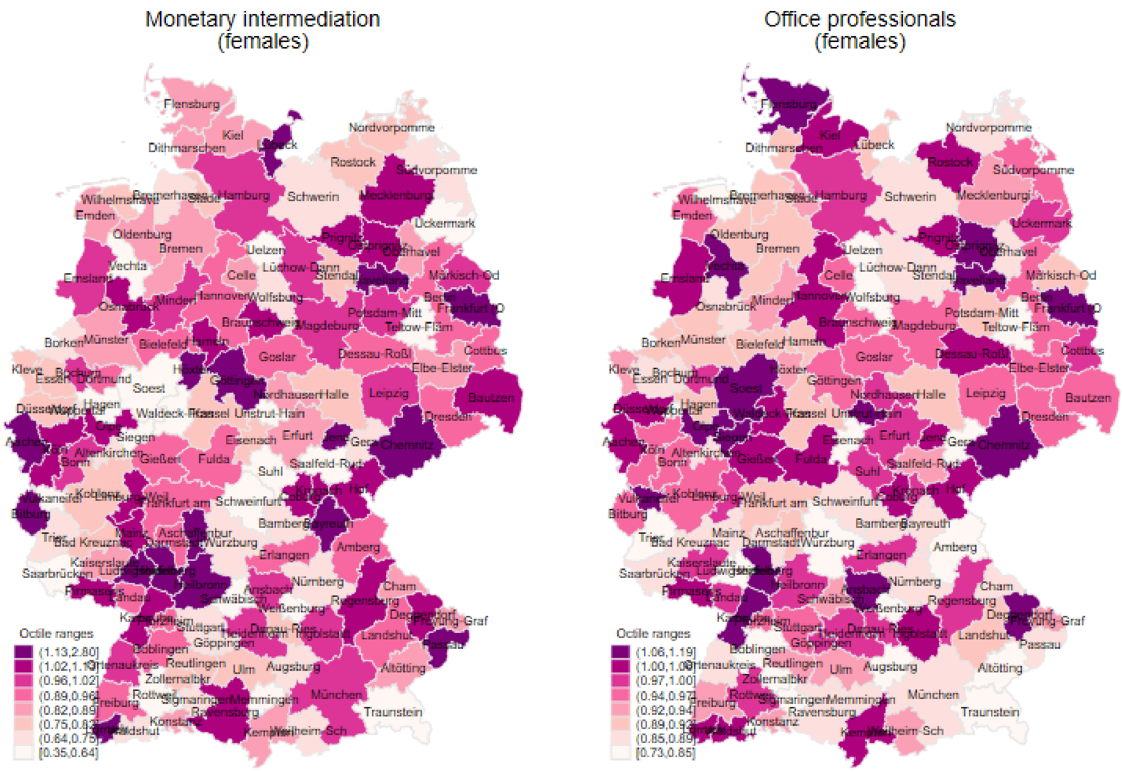
Notes: Coefficient plots. The columns correspond to the first stage, reduced form (ITT) and 2SLS (peer effect) regressions for employment at 62. Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers (all measured when the peer was 57), establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. The points represent the estimated coefficients, and the bars represent the 95% confidence intervals. The subsample analyses are performed by categories of establishment size (**Panel A**), and workgroup size (**Panel B**), where micro establishments are those with less than 10 workers, small in the range 10-49, medium in the range 50-249, and large above 250. The bottom **Panels C and D** show subsample analyses by number of coworkers and peers: 1, 2-4, or more than 5. The estimates are based on the models that control for baseline characteristics (see Table 3.1 for the list.)

Figure 3.A.6: Subsample analyses by cohorts of coworkers



Notes: Coefficient plots. The columns correspond to the first stage, reduced form (ITT) and 2SLS (peer effect) regressions for employment at 62. Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers (all measured when the peer was 57), establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. The points represent the estimated coefficients, and the bars represent the 95% confidence intervals. The subsample analyses are performed by the cohorts of coworkers, 1953-1957.

Figure 3.A.7: External labor market thickness by German industry and occupation in 2007



Notes: This map shows the computed external labor market thicknesses (ELMT) for each of the 141 local labor markets based on the Kosfeld and Werner (2012) classifications, which are constructed based on high within-region and low between-region commuting. I compute ELMT for the industry and occupation with large shares of female employees: “Monetary intermediation” (**left Panel**) and “Office professionals” (**right Panel**). I plot the ELMT indices on the map using the eight quantile ranges (octiles) shown in the left corner of each graph.

3.B Appendix: Tables

Table 3.B.1: Retirement rules by birthcohorts and pathways

Birth cohorts	Pathway					
	women's		long-insured		regular	
	ERA	NRA	ERA	NRA	ERA	NRA
1950	60	65	63	65 y 4 m	65 y 4 m	
1951	60	65	63	65 y 5 m	65 y 5 m	
1952	-	-	63	65 y 6 m	65 y 6 m	
1953	-	-	63	65 y 7 m	65 y 7 m	
1954	-	-	63	65 y 8 m	65 y 8 m	
1955	-	-	63	65 y 9 m	65 y 9 m	
1956	-	-	63	65 y 10 m	65 y 10 m	
1957	-	-	63	65 y 11 m	65 y 11 m	

Notes: This table demonstrates the statutory retirement ages by cohorts and pathways. Starting from the 1952 cohort, the women's pathway to retirement was abolished. NRA stands for Normal Retirement Age, while ERA stands for the Early Retirement Age. "y" is an abbreviation for Year, while "m" is an abbreviation for months. For example, a person born in 1952 is not eligible for women's pathway to retirement (marked by "-"), but can claim pensions at 63 (by long-insurance pathway), or, for the full benefit amount, can retire at 65 years and 6 months (the NRA of long insured and regular pathways). The visual representation of these rules can be found in Figure 3.2

Table 3.B.2: Sample restrictions to obtain the original data extract

Restriction	N establishments	N workers
Universe of establishments and workers in 1995-2019	8,241,529	69,296,143
+ observed in 2008	1,958,754	23,798,218
+ employed at least 1 focal worker in 2008	352,836	15,364,408
+ private sector	317,912	13,388,108
+ at least 5 employees in 2008	193,612	13,059,745
+ at most 500 employees in 2008	190,228	9,117,917
Universe of affected establishments, and their employed workers in 1995-2019	190,228	26,593,003

Notes: This table shows the number of establishments and workers after each restriction in the data extract requested.

Table 3.B.3: Baseline sample size after each restriction in German social security data

	N workgroups	N treated peers (1952)	N control peers (1951)	N coworkers (1953-1957)
Unrestricted sample	**286,046**	**110,796**	**103,639**	
Restriction 1	139,833	80,114	73,533	184,986
Restriction 2	64,324	40,627	36,828	182,584

Notes: This table records the sample size after each of the restrictions in German social security data. The first column names the restrictions. Workgroups are defined as establishments interacted with 4-digit occupations. **"Unrestricted sample"** stands for the sample of employees in workgroups with at least 1 "peer", defined as a woman born in 1951-1953, who never worked as a miner or sailor, and was employed in given workgroups at the age 56-57. The **"Restriction 1"** restricts the workgroups to those where the peers are born either before (in 1951) or after (in 1952) reform cutoff. **"Restriction 2"** represents the baseline sample. It further restricts the sample to the workgroups which had at least 1 coworker employed at the year when the workgroup peer was 57. The visual representation of sample construction can be found in Figure 3.A.2.

Table 3.B.4: Descriptive statistics and balance test for peers born before and after the reform cutoff

	Before the reform			After the reform				
Variable	Mean	SD	N	Mean	SD	N	Difference	t
vocational education	0.806	0.395	36357	0.807	0.395	40191	.001	-0.317
university education	0.073	0.261	36357	0.082	0.275	40191	-.009	-4.555
earnings	25978.191	13464.784	36828	26607.134	13921.251	40627	-628.943	-6.378
full-time	0.587	0.492	36828	0.571	0.495	40627	0.016	4.536
tenure	4401.402	2973.67	36828	4397.013	3045.511	40627	4.389	0.2026
foreign	0.043	0.203	36828	0.043	0.204	40627	-0.000	-0.186
manager	0.005	0.072	36828	0.006	0.08	40627	-.001	-2.354
worker AKM effect (2000-2006)	4.328	0.389	25196	4.341	0.39	27662	-.013	-3.757
establishment size	41.064	50.121	36828	44.965	56.282	40627	-3.901	-10.146
workgroup size	28.312	38.002	36828	31.511	44	40627	-3.198	-10.774
N peers	1.423	0.794	36828	1.481	0.93	40627	-.058	-9.296
East Germany	0.219	0.413	36828	0.233	0.423	40627	-.014	-4.635
share 55+ in workgroup	0.248	0.166	36828	0.265	0.184	40627	-.017	-13.341
share full-time in workgroup	0.629	0.337	36828	0.615	0.338	40627	.014	5.649
share of women in workgroup	0.805	0.219	36828	0.799	0.218	40627	.006	3.623
median workgroup wage in workgroup	25973.615	11680.799	36828	26684.72	12124.672	40627	-711.105	-8.294
N coworkers in workgroup	4.279	5.189	36828	4.68	6.113	40627	-.400	-9.773
establishment AKM effect (2000-2006)	0.087	0.224	35525	0.095	0.225	37768	-0.007	-4.396
N coworkers (1953 birth cohort)	0.595	0.937	36828	0.629	1.043	40627	-0.034	-4.809
N coworkers (1954 birth cohort)	0.625	0.964	36828	0.665	1.08	40627	-0.040	-5.376
N coworkers (1955 birth cohort)	0.633	0.992	36828	0.698	1.136	40627	-0.065	-8.479
N coworkers (1956 birth cohort)	0.679	1.049	36828	0.738	1.199	40627	-0.059	-7.253
N coworkers (1957 birth cohort)	0.687	1.056	36828	0.759	1.203	40627	-0.071	-8.718

Notes: This table shows the summary statistics and balance test for observable characteristics by treatment status (coworkers in workgroups with peers whose ERA raised, vs. those whose ERA was 60.)

Table 3.B.5: Descriptive statistics and balance test for coworkers exposed to peers born before and after the reform cutoff

	Before the reform			After the reform				
Variable	Mean	SD	N	Mean	SD	N	Difference	t
vocational education	0.792	0.406	113413	0.789	0.408	129509	.002	1.4094
university education	0.104	0.305	113413	0.114	0.318	129509	-.010	-7.906
earnings	29543.22	16055.18	114464	30614.501	16687.902	130756	-1071.281	-16.1422
full-time	0.652	0.476	114464	0.637	0.481	130756	0.016	8.187
tenure	4042.044	3096.933	114464	4072.381	3170.994	130756	-30.337	-2.389
foreign	0.048	0.213	114464	0.046	0.21	130756	0.001	1.406
manager	0.004	0.062	114464	0.005	0.07	130756	-0.001	-3.646
worker AKM effect (2000-2006)	4.401	0.412	84905	4.423	0.413	96410	-.023	-11.622
establishment size	71.108	74.068	114464	79.079	83.328	130756	-7.970	-24.879
workgroup size	57.08	66.680	114464	63.828	75.044	130756	-6.748	-23.393
N peers	1.377	0.776	114464	1.454	0.935	130756	-0.077	-22.050
East Germany	0.213	0.409	114464	0.234	0.423	130756	-0.021	-12.531
share 55+ in workgroup	0.195	0.133	114464	0.211	0.145	130756	-.0157768	-27.9140
share full-time in workgroup	0.664	0.325	114464	0.649	0.325	130756	.0147866	11.2550
share of women in workgroup	0.713	0.26	114464	0.705	0.257	130756	.0073516	7.0252
median workgroup wage in workgroup	28189.586	12866.3	114464	29311.531	13528.143	130756	-1121.945	-20.9613
N coworkers in workgroup	8.984	9.543	114464	10.018	10.969	130756	-1.033763	-24.7290
establishment AKM effect (2000-2006)	0.113	0.216	111144	0.123	0.215	121978	-.010	-10.825
1953 birth cohort	0.185	0.388	114464	0.179	0.383	130756	.005401	3.4609
1954 birth cohort	0.194	0.395	114464	0.19	0.392	130756	.0037983	2.3835
1955 birth cohort	0.197	0.398	114464	0.202	0.401	130756	-.0051422	-3.1794
1956 birth cohort	0.208	0.406	114464	0.211	0.408	130756	-.0025493	-1.5480
1957 birth cohort	0.217	0.412	114464	0.218	0.413	130756	-.0015079	-0.9026

Notes: This table shows the summary statistics and balance test for observable characteristics by treatment status (coworkers in workgroups with peers whose ERA raised, vs. those whose ERA was 60.)

Table 3.B.6: First stage, reduced form, and peer effect regressions for indicators for old-age employment measures

	First stage		Reduced form		Peer effect	
Retirement age (in months)	2.432***	2.634***	1.730***	1.877***	0.712***	0.712***
	(0.417)	(0.414)	(0.204)	(0.200)	(0.139)	(0.127)
Control mean	769.753		760.848			
E at 62 (baseline)	0.116***	0.113***	0.015***	0.014***	0.133***	0.126***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Control mean	0.716		0.785			
E at 60-62	0.083***	0.080***	0.016***	0.015***	0.188***	0.181***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.023)	(0.023)
Control mean	0.797		0.828			
E at 60-62 (same firm)	0.062***	0.065***	0.027***	0.029***	0.444***	0.447***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.046)	(0.044)
Control mean	0.235		0.242			
Observations	182584	182584	182584	182584	182584	182584
N workgroups	64324	64324	64324	64324	64324	64324
Controls	No	Yes	No	Yes	No	Yes

Notes: This table shows the effect of the rise in ERA on labor market outcomes (first column): 1st stage (Equation 25), ITT (Equation 26), and IV regressions (Equation 27). Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers, establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. The control means are the average values of the outcomes when I limit the sample to the workgroups with peers whose ERA was fixed at 60. Robust standard errors in parentheses are clustered at the workgroup level.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 3.B.7: Robustness and sensitivity checks. First stage, reduced form, and peer effect regressions by included covariates

	First stage		Reduced form		Peer effect	
Panel A: baseline controls						
E at 62	0.116***	0.113***	0.015***	0.014***	0.133***	0.126***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Panel B: baseline + additional coworker controls						
E at 62	0.116***	0.111***	0.015***	0.016***	0.133***	0.146***
	(0.004)	(0.005)	(0.002)	(0.003)	(0.019)	(0.023)
Panel C: baseline + additional peer controls						
E at 62	0.116***	0.104***	0.015***	0.015***	0.133***	0.143***
	(0.004)	(0.005)	(0.002)	(0.003)	(0.019)	(0.025)
Panel D: baseline + firm and workgroup controls						
E at 62	0.116***	0.114***	0.015***	0.015***	0.133***	0.133***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Panel E: baseline + occupation and industry FE						
E at 62	0.116***	0.111***	0.015***	0.014***	0.133***	0.127***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Panel F: baseline + municipality FE						
E at 62	0.116***	0.113***	0.015***	0.015***	0.133***	0.130***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Control mean	0.716		0.785			
Observations	182584	182584	182584	182584	182584	182584
N workgroups	64324	64324	64324	64324	64324	64324
Controls	No	Yes	No	Yes	No	Yes

Notes: This table shows the effect of the rise in ERA on *employment at 62*: 1st stage (Equation 25), ITT (Equation 26), and IV regressions (Equation 27). **Panel A** includes baseline control variables: coworker and average peer education, wages, dummy for the location of establishments in the regions of the former West Germany, establishment and workgroup sizes, number of peers, and coworkers. **Panel B** adds coworker full-time status, experience, foreign dummy, managerial position, estimated AKM worker fixed effect, while **Panel C** adds the same variables for peers, averaged over peers. **Panel D** adds share of women, share of older workers above 55 years old, share of full-time workers, median wages, number of coworkers, joint tenure of peers and coworkers, establishment AKM fixed effects, number of coworkers, number of female coworkers of each birth cohort. **Panel E** adds fixed effects for 2-digit industries and occupations (classification by Blossfeld (1985)). **Panel F** adds fixed effects for the administrative districts ("*kreis*"). The control means are the average values of the outcomes when I limit the sample to the workgroups with peers whose ERA was fixed at 60. Robust standard errors in parentheses are clustered at the workgroup level.

* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 3.B.8: Robustness and sensitivity checks. First stage, reduced form, and peer effect regressions by treatment definition

	First stage		Reduced form		Peer effect	
Panel A: discrete treatment (baseline)						
Treatment	0.116***	0.113***	0.015***	0.014***	0.133***	0.126***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Panel B: continuous treatment: Treatment × N peers						
N treated peers	0.039***	0.061***	0.002*	0.006***	0.060*	0.105***
	(0.002)	(0.003)	(0.001)	(0.001)	(0.034)	(0.023)
Control mean	0.716		0.785			
Observations	182584	182584	182584	182584	182584	182584
N workgroups	64324	64324	64324	64324	64324	64324
Controls	No	Yes	No	Yes	No	Yes

Notes: This table shows the effect of the rise in ERA on *employment at 62*: 1st stage (Equation 25), ITT (Equation 26), and IV regressions (Equation 27). Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers, establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. The control means are the average values of the outcomes when I limit the sample to the workgroups with peers whose ERA was fixed at 60. Robust standard errors in parentheses are clustered at the workgroup level.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 3.B.9: Robustness and sensitivity checks. First stage, reduced form, and peer effect regressions by peer definitions

	First stage		Reduced form		Peer effect	
Sample A: 1-year peer bandwidth (baseline)						
Employment at 62	0.116***	0.113***	0.015***	0.014***	0.133***	0.126***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Control mean	0.716		0.785			
Observations	182584	182584	182584	182584	182584	182584
N workgroups	64324	64324	64324	64324	64324	64324
Sample B: 1-year peer bandwidth, excluding Dec. 1951 and Jan. 1952						
Employment at 62	0.114***	0.111***	0.016***	0.015***	0.141***	0.132***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.020)
Control mean	0.715		0.784			
Observations	180154	180154	180154	180154	180154	180154
N workgroups	61365	61365	61365	61365	61365	61365
Sample C: 2-year peer bandwidth						
Employment at 62	0.115***	0.112***	0.016***	0.014***	0.138***	0.130***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Control mean	0.716		0.784			
Observations	190784	190784	190784	190784	190784	190784
N workgroups	65690	65690	65690	65690	65690	65690
Sample D: 1-year peer bandwidth, peers on both sides						
Employment at 62	-0.000	0.041***	-0.002***	0.008***	4.418	0.188***
	(0.000)	(0.002)	(0.000)	(0.001)	(2.741)	(0.028)
Control mean	0.717		0.778			
Observations	327156	327156	327156	327156	327156	327156
N workgroups	108696	108696	108696	108696	108696	108696
Sample E (i): 1-year bandwidth around 1951 cohort (false sample)						
Employment at 62	0.010	-0.008	0.030*	0.025	3.081	-3.345
	(0.042)	(0.041)	(0.018)	(0.017)	(13.378)	(18.414)
Control mean	0.697		0.768			
Observations	2626	2626	2626	2626	2626	2626
N workgroups	895	895	895	895	895	895
Sample E (ii): 1-year bandwidth, males (false sample)						
Employment at 62	0.089	0.094**	0.015	0.045*	0.174	0.477
	(0.067)	(0.044)	(0.029)	(0.025)	(0.365)	(0.343)
Control mean	0.802		0.769			
Observations	1288	1288	1288	1288	1288	1288
N workgroups	419	419	419	419	419	419
Controls	No	Yes	No	Yes	No	Yes

Notes: This table shows the effect of the rise in ERA on *employment at 62*: 1^{st} stage (Equation 25), ITT (Equation 26), and IV regressions (Equation 27). Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers, establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. Robust standard errors in parentheses are clustered at the workgroup level. Details about sample definitions can be found in subsubsection 3.3.2 and section 3.7. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 3.B.10: Robustness and sensitivity checks. First stage, reduced form, and peer effect regressions by workgroup definitions.

	First stage		Reduced form		Peer effect	
Sample A: establishments × 4-digit occupations (baseline)						
Employment at 62	0.116***	0.113***	0.015***	0.014***	0.133***	0.126***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Control mean	0.716		0.785			
Observations	182584	182584	182584	182584	182584	182584
N workgroups	64324	64324	64324	64324	64324	64324
Sample F: establishments × 3-digit occupations						
Employment at 62	0.117***	0.115***	0.016***	0.015***	0.138***	0.131***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Control mean	0.714		0.783			
Observations	181854	181854	181854	181854	181854	181854
N workgroups	65410	65410	65410	65410	65410	65410
Sample G: establishments × 2-digit occupations						
Employment at 62	0.115***	0.112***	0.016***	0.014***	0.138***	0.130***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Control mean	0.716		0.784			
Observations	190784	190784	190784	190784	190784	190784
N workgroups	65690	65690	65690	65690	65690	65690
Controls	No	Yes	No	Yes	No	Yes

Notes: This table shows the effect of the rise in ERA on *employment at 62*: 1st stage (Equation 25), ITT (Equation 26), and IV regressions (Equation 27). Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers, establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. The control means are the average values of the outcomes when I limit the sample to the workgroups with peers whose ERA was fixed at 60. Robust standard errors in parentheses are clustered at the workgroup level. Details about sample definitions can be found in subsubsection 3.3.2 and section 3.7.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 3.B.11: Robustness and sensitivity checks. First stage, reduced form, and peer effect regressions by coworker definitions

	First stage		Reduced form		Peer effect	
Panel A: all the coworkers (baseline)						
Employment at 62	0.116***	0.113***	0.015***	0.014***	0.133***	0.126***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Control mean	0.716		0.785			
Observations	182584	182584	182584	182584	182584	182584
N workgroups	64324	64324	64324	64324	64324	64324
Panel B: coworkers born after 1955						
Employment at 62	0.117***	0.114***	0.016***	0.014***	0.133***	0.124***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.023)	(0.023)
Control mean	0.716		0.779			
Observations	114230	114230	114230	114230	114230	114230
N workgroups	51241	51241	51241	51241	51241	51241
Panel C: coworkers with joint experience over 2 years						
Employment at 62	0.116***	0.113***	0.016***	0.015***	0.140***	0.129***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.020)	(0.020)
Control mean	0.719		0.792			
Observations	155358	155358	155358	155358	155358	155358
N workgroups	57875	57875	57875	57875	57875	57875
Panel D: coworkers in firms that survived until 2019						
Employment at 62	0.105***	0.103***	0.009***	0.009***	0.084***	0.086***
	(0.005)	(0.004)	(0.002)	(0.002)	(0.022)	(0.022)
Control mean	0.739		0.804			
Observations	146949	146949	146949	146949	146949	146949
N workgroups	51293	51293	51293	51293	51293	51293
Panel E: coworkers who did not die						
Employment at 62	0.116***	0.113***	0.016***	0.015***	0.135***	0.129***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.019)	(0.019)
Control mean	0.716		0.787			
Observations	182149	182149	182149	182149	182149	182149
N workgroups	64241	64241	64241	64241	64241	64241
Controls	No	Yes	No	Yes	No	Yes

Notes: This table shows the effect of the rise in ERA on *employment at 62*: 1st stage (Equation 25), ITT (Equation 26), and IV regressions (Equation 27). Control variables include coworker and average peer wages and education categories, establishment size, workgroup size, number of peers, number of coworkers, establishment location (West vs. East Germany), and fixed effects for the coworkers' birth cohorts. The control means are the average values of the outcomes when I limit the sample to the workgroups with peers whose ERA was fixed at 60. Robust standard errors in parentheses are clustered at the workgroup level.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table 3.B.12: Post-reform labor market shares and treatment effects at ages 60–62

Panel A: results from Geyer and Welteke (2021)			
	share (cohort 1951)	RDD around 1952	
Employment	44.1	13.5	
Unemployment (UI)	6.2	5.2	
Employment+UI	50.3	17.5	
Disability (DI)	8.3	–	
Inactivity	4.8	6.2	
DI + Inactivity	19.3	6.2	
Panel B: results from this paper			
	share (cohort 1952)	1st stage	
LM activity of peers	79.7	11.6	
	share (exposed to 1951 peers)	ITT effect	peer effect
LM activity of coworkers	82.8	1.6	18.8

Notes: All the numbers are presented in percentage points. **Panel A** displays the shares based on Geyer and Welteke (2021), which include more detailed labor market activity states for peers (birth cohorts 1951-1952), but no workgroup information for defining their coworkers (birth cohorts 1953-1957). **Panel B** displays the results in this paper, which cannot disentangle between the "active states" but includes detailed workgroup data. "LM activity" includes any social security contributions in IEB records, such as regular employment and UI receipt.

3.C Appendix: Alternative Sample Definitions

I create five main alternative samples to perform robustness and sensitivity checks in subsection 3.4. They rely on altering two dimensions of dataset construction: (1) the bandwidth, (2) the occupation group. The first group consists of four samples:

1. **Sample B.** redefining peers by excluding those in a "donut hole" (December 1951, and January 1952).
2. **Sample C.** allowing for the peers to be born two years (opposed to one year) around the reform cutoff. I redefine the peers as those born in 1950-1953 and coworkers as those born after 1953. For such samples where the peers can belong to two (or more) cohorts (Samples C and D), I define coworkers employed in the earliest year when the peer turns 57. For example, in a 2-year bandwidth specification with peers born in 1950 and 1951, the coworkers would be defined in 2007, i.e., when the oldest peer turns 57. The main observables and controls are also defined this year. Additionally, coworkers in these samples belong to 1954-1957 (1953 is excluded) cohorts, such that they do not overlap with peers, which now can include the 1953 cohorts.
3. **Sample D.** allowing for mixed peers. This sample relaxes the restriction that drops the workgroups where peers were born on both sides of the cutoff.
4. **Sample E (i).** Falsification test sample (false cutoff). In this sample, women of the 1951 cohort are falsely attributed to the treated group, while the women of the 1950 cohort are attributed to the control group. I keep workgroups which have peers either in the 1950 cohort (control) or the 1951 cohort (false treatment group). The coworkers are defined as those born in the 1952-1956 cohorts, so that they are younger than the peers, thus circumventing the reflection problem. For details about the reflection problem, see subsection 3.3. I use the Sample of Integrated Employer-Employee Data (SIEED7518), a random 1.5% sample of all establishments in Germany, because of data sensitivity.[124] Since the data is right-censored, unlike the baseline sample, I exclude the 1957 cohort from the coworkers: their employment status at 62 is not observed.
5. **Sample E (ii).** Falsification test sample (false gender). For the second falsification test, I repeat the data creation of the baseline sample on the SIEED7518 data, except that the peers are now defined as men, as opposed to women. The 1957 cohort is removed from the coworkers due to the right-censored data (see the point above).

The second group relates to the workgroup definitions (and the corresponding reform window) and consists of two samples:

1. **Sample F.** workgroups defined as 3-digit occupations in establishments

[124] Due to data sensitivity, I observe the universe of affected workgroups only for the baseline sample, and the false samples are created based on a random sample of establishments. Overall, the data resemble the data used in the baseline specifications, but include only employment spells, and are smaller.

2. **Sample G.** workgroups defined as 2-digit occupations in establishments

These alternative samples have a similar baseline sample definition, described in subsubsection 3.3.2. The sample sizes, including the number of coworkers, peers (by treatment), and workgroups, are recorded in Table 3.C.1.

Table 3.C.1: Sample sizes in baseline and alternative samples

	N workgroups	N treated peers (1952)	N control peers (1951)	N coworkers (1953-1957)
Panel A: baseline sample				
Sample A	64,324	40,627	36,828	182,584
Panel B: alternative samples				
Sample B	61,365	38,518	34,754	180,154
Sample C	76,228	54,664	45,072	159,103
Sample D	108,696	128,977	114,592	327,156
Sample E (i)	895	440	643	2626
Sample E (ii)	419	295	207	1288
Sample F	63,444	41,011	37,282	181,854
Sample G	87,259	41,598	37,872	190,784

Notes: This table describes the number of workgroups, peers (by treatment), and coworkers in the alternative samples. For the details on baseline and alternative sample definitions, see subsubsection 3.3.2 and section 3.7.

Summary: Policy Implications and Synthesis

This dissertation provides a comprehensive evaluation of pension reforms that increase statutory retirement ages, combining evidence across three chapters into a coherent policy message. While such reforms achieve their primary goal—substantially increasing employment among older workers—their broader consequences are more nuanced and extend beyond aggregate labor market outcomes.[125]

First, the results show that the gains from extended employment are unevenly distributed within firms. Benefits accrue primarily to employers and to workers who are difficult to replace, reflecting the value of firm-specific human capital. At the same time, important costs arise within the workforce. Middle-aged employees, particularly women, tend to experience reduced promotion opportunities and weaker hiring prospects due to delayed job turnover. Older workers themselves may bear part of the adjustment through lower wages and weakened bargaining power as their outside options decline. Because total wage bills remain approximately constant, these effects reflect redistribution within firms rather than costs borne by employers.

Second, the findings imply that policy evaluations that are narrowly focused on fiscal savings from delayed pension claims are incomplete. Ignoring intra-firm distributional effects risks overlooking key equity implications of retirement reforms.

Third, the dissertation documents positive spillovers through workplace peer networks. Increased employment among older workers induces additional labor supply responses from slightly younger coworkers, amplifying the direct effects of the reform. As a result, conventional cost-benefit analyses that neglect such peer effects are likely to underestimate total employment gains.

Taken together, the dissertation provides a framework for a more complete policy evaluation that integrates aggregate effects, spillovers, and distributional consequences. Importantly, recent institutional changes suggest that these mechanisms may evolve over time. In particular, the liberalization of rules governing combinations of pension receipt and continued employment strengthens older workers' outside options, which may attenuate the wage suppression documented in Chapter 2 and alter bargaining dynamics within firms. At the same time, increased flexibility for employers to retain workers beyond the statutory retirement age can help to relax position constraints, potentially reducing the crowd-out effects on middle-aged coworkers identified in earlier reforms (Chapter 1). Finally, as working beyond retirement age becomes more common and socially accepted, the peer effects highlighted in Chapter 3 may operate differently, with changing norms influencing both their magnitude and direction. Whether the mechanisms identified in the context of earlier reforms remain operative under evolving institutional arrangements is an open and policy-relevant question. Addressing these questions represents a natural and important direction for future research

[125]I am deeply grateful to Johannes Geyer for his insightful dissertation review comments and for shaping the articulation of these policy implications.

List of References

Acemoglu, D. and **Pischke, J.-S.** (1999). The Structure of Wages and Investment in General Training. Journal of Political Economy, 107(3):539–572.

Acemoglu, D. and **Restrepo, P.** (2022). Demographics and Automation. The Review of Economic Studies, 89(1):1–44.

Addison, J. T. and **Portugal, P.** (1989). Job Displacement, Relative Wage Changes, and Duration of Unemployment. Journal of Labor Economics, 7(3):281–302.

Atalay, K. and **Barrett, G. F.** (2015). The Impact of Pension Eligibility Age on Retirement and Program Dependence: Evidence from an Australian Experiment. Review of Economics and Statistics, 97(1):71–87.

Atalay, K., **Barrett, G. F.**, and **Siminski, P.** (2019). Pension Incentives and the Joint Retirement of Couples: Evidence from two Natural Experiments. Journal of Population Economics, 32:735–767.

Badalyan, S. (2024a). Employer Responses to Raising the Retirement Age: Spillovers on Coworkers and External Hiring. In The Seventeenth Young Economists' Seminar at The Thirtieth Dubrovnik Economic Conference.

Badalyan, S. (2024b). Firm Responses to Raising Women's Retirement Age. Available at SSRN 5788182.

Badalyan, S. (2025a). Crowded Career Ladders? Intra-Firm Spillovers of Raised Retirement Age. CERGE-EI Working Paper Series No 810.

Badalyan, S. (2025b). Peer Effects in Old-Age Employment Among Women. CERGE-EI Working Paper Series No 800.

Badalyan, S. (2025c). Retirement Age Reforms and Worker Substitutability: Implications for Employment of Older Workers. CERGE-EI Working Paper Series No 794.

Baker, G., **Gibbs, M.**, and **Holmstrom, B.** (1994). The Internal Economics of the Firm: Evidence from Personnel Data. The Quarterly Journal of Economics, 109(4):881–919.

Bandiera, O., **Barankay, I.**, and **Rasul, I.** (2009). Social Connections and Incentives in the Workplace: Evidence from Personnel Data. Econometrica, 77(4):1047–1094.

Bartel, A. P., **Beaulieu, N. D.**, **Phibbs, C. S.**, and **Stone, P. W.** (2014). Human Capital and Productivity in a team Environment: Evidence from the Healthcare Sector. American Economic Journal: Applied Economics, 6(2):231–259.

Bayer, C. and **Kuhn, M.** (2018). Which Ladder to Climb? Wages of Workers by Job, Plant, and Education. Technical report, CESifo Working Paper.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. Journal of Political Economy, 70(5, Part 2):9–49.

Becker, S. O. and **Hvide, H. K.** (2022). Entrepreneur Death and Startup Performance. Review of Finance, 26(1):163–185.

Behaghel, L. and **Blau, D. M.** (2012). Framing Social Security Reform: Behavioral Responses to Changes in the Full Retirement Age. American Economic Journal: Economic Policy, 4(4):41–67.

Bennedsen, M., **Pérez-González, F.**, and **Wolfenzon, D.** (2020). Do CEOs Matter? Evidence from Hospitalization Events. The Journal of Finance, 75(4):1877–1911.

Berg, P., **Eckrote-Nordland, M.**, **Hamman, M.**, **Hochfellner, D.**, **Piszczek, M. M.**, and **Ruhm, C. J.** (2025). Pension Reforms and Personnel Decisions. LABOUR, 39(2):89–100.

Bernheim, B. D. (1994). A Theory of Conformity. Journal of Political Economy, 102(5):841–877.

Bertheau, A. (2021). Employer Search Behavior: Reasons for Internal Hiring. Labour Economics, 73:102064.

Bertheau, A., **Cahuc, P.**, **Jäger, S.**, and **Vejlin, R.** (2022). Turnover Costs: Evidence from Unexpected Worker Separations. Unpublished Manuscript, 2.

Bertoni, M. and **Brunello, G.** (2021). Does a Higher Retirement Age Reduce Youth Employment? Economic Policy, 36(106):325–372.

Bertrand, M., **Luttmer, E. F.**, and **Mullainathan, S.** (2000). Network Effects and Welfare Cultures. The Quarterly Journal of Economics, 115(3):1019–1055.

Bianchi, N., **Bovini, G.**, **Li, J.**, **Paradisi, M.**, and **Powell, M.** (2023). Career Spillovers in Internal Labour Markets. The Review of Economic Studies, 90(4):1800–1831.

Blau, D. M. and **Goodstein, R. M.** (2010). Can Social Security Explain Trends in Labor Force Participation of Older Men in the United States? Journal of Human Resources, 45(2):328–363.

Blau, F. D. and **Kahn, L. M.** (2017). The Gender Wage Gap: Extent, Trends, and Explanations. Journal of Economic Literature, 55(3):789–865.

Bleakley, H. and **Lin, J.** (2012). Thick-market Effects and Churning in the Labor Market: Evidence from US Cities. Journal of Urban Economics, 72(2-3):87–103.

Bloemen, H., **Hochguertel, S.**, and **Zweerink, J.** (2019). The Effect of Incentive-Induced Retirement on Spousal Retirement Rates: Evidence from a Natural Experiment. Economic Inquiry, 57(2):910–930.

Blossfeld, H. (1985). Bildungsexpansion und Berufschancen. Frankfurt/Main, New York.

Blume, L. E., **Brock, W. A.**, **Durlauf, S. N.**, and **Ioannides, Y. M.** (2011). Identification of Social Interactions. In Handbook of Social Economics, volume 1, pages 853–964. Elsevier.

Blundell, R., **French, E.**, and **Tetlow, G.** (2016). Retirement Incentives and Labor Supply. In Handbook of the Economics of Population Aging, volume 1, pages 457–566. Elsevier.

Boelmann, B., **Raute, A.**, and **Schönberg, U.** (2025). Wind of Change? Cultural Determinants of Maternal Labor Supply. American Economic Journal: Applied Economics, 17(2):41–74.

Boeri, T., **Garibaldi, P.**, and **Moen, E. R.** (2022). In medio stat victus: Labor Demand Effects of an Increase in the Retirement Age. Journal of Population Economics, 35(2):519–556.

Boeri, T. and **van Ours, J. C.** (2021). The Economics of Imperfect Labor Markets. Princeton University Press, 3rd edition.

Bonney, J., **Pistaferri, L.**, and **Voena, A.** (2025). Childbirth and Firm Performance: Evidence from Norwegian Entrepreneurs. Technical report, National Bureau of Economic Research.

Boockmann, B., **Kroczek, M.**, and **Laub, N.** (2023). Tightening Access to Early Retirement: Who can Adapt? IZA Discussion Paper No. 16292.

Börsch-Supan, A. H. and **Coile, C.** (2018). Social Security Programs and Retirement around the World: Reforms and Retirement Incentives–Introduction and Summary. Technical report, National Bureau of Economic Research.

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

Bramoullé, Y., **Djebbari, H.**, and **Fortin, B.** (2020). Peer Effects in Networks: A Survey. Annual Review of Economics, 12(1):603–629.

Brenøe, A. A., **Canaan, S.**, **Harmon, N. A.**, and **Royer, H. N.** (2024). Is Parental Leave Costly for Firms and Coworkers? Journal of Labor Economics, 42(4):1135–1174.

Brinch, C. N., **Vestad, O. L.**, and **Zweimüller, J.** (2015). Excess Early Retirement? Evidence from the Norwegian 2011 Pension Reform. Mimeo.

Bronson, M. A. and **Thoursie, P. S.** (2019). The Wage Growth and Within-firm Mobility of Men and Women: New Evidence and Theory. Unpublished. https://economicdynamics.org/meetpapers/2018/paper_923. pdf. Accessed March, 25:2021.

Brown, K. M. and **Laschever, R. A.** (2012). When They're Sixty-four: Peer Effects and the Timing of Retirement. American Economic Journal: Applied Economics, 4(3):90–115.

Bulmahn, T. (1998). Rette sich, wer kann? Die Krise der gesetzlichen Rentenversicherung und die Privatisierung der Altersvorsorge. Technical report, WZB Discussion Paper.

Byrne, D. (1971). The Attraction Paradigm, New York: Ac.

Cahuc, P., **Carcillo, S.**, and **Le Barbanchon, T.** (2019). The Effectiveness of Hiring Credits. The Review of Economic Studies, 86(2):593–626.

Cahuc, P., **Marque, F.**, and **Wasmer, E.** (2008). A Theory of Wages and Labor Demand with Intra-firm Bargaining and Matching Frictions. International Economic Review, 49(3):943–972.

Card, D., **Heining, J.**, and **Kline, P.** (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. The Quarterly Journal of Economics, 128(3):967–1015.

Carlsson, M. and **Reshid, A. A.** (2022). Co-worker Peer Effects on Parental Leave Take-up. The Scandinavian Journal of Economics, 124(4):930–957.

Carta, F., **Casarico, A.**, **De Philippis, M.**, and **Lattanzio, S.** (2024). Mom's out: Employment after Childbirth and Firm-level Responses. Bank of Italy Temi di Discussione (Working Paper) No, 1458.

Carta, F. and **De Philippis, M.** (2024). The Forward-looking Effect of Increasing the Full Retirement Age. The Economic Journal, 134(657):165–192.

Carta, F., **D'Amuri, F.**, and **Von Wachter, T.** (2024). Older Workers, Pension Reforms and Firm Outcomes. Pension Reforms and Firm Outcomes.

Carta, F., **D'Amuri, F.**, and **Von Wachter, T.** (2025). Older workers, pension reforms and firm outcomes. Labour Economics, page 102823.

Casarico, A., **Di Porto, E.**, **Kopinska, J.**, and **Lattanzio, S.** (2025). Leave and Let Leave: Workplace Peer Effects in Fathers' Take-up of Parental Leave. Technical report, CESifo Working Paper.

Cavapozzi, D., **Francesconi, M.**, and **Nicoletti, C.** (2021). The Impact of Gender Role Norms on Mothers' Labor Supply. Journal of Economic Behavior & Organization, 186:113–134.

Chan, S. and **Stevens, A. H.** (2008). What you don't know can't help you: Pension knowledge and retirement decision-making. The Review of Economics and Statistics, 90(2):253–266.

Chan, W. (1996). External Recruitment versus Internal Promotion. Journal of Labor Economics, 14(4):555–570.

Ciani, E. (2016). Retirement, pension eligibility and home production. Labour Economics, 38:106–120.

Corekcioglu, G., **Francesconi, M.**, and **Kunze, A.** (2025). Parental Leave from the Firm's Perspective. Institutt for samfunnsøkonomi.

Cornelissen, T., **Dustmann, C.**, and **Schönberg, U.** (2017). Peer Effects in the Workplace. American Economic Review, 107(2):425–456.

Cortes, G. M. and **Salvatori, A.** (2019). Delving into the Demand Side: Changes in Workplace Specialization and Job Polarization. Labour economics, 57:164–176.

Dahl, G. B., **Løken, K. V.**, and **Mogstad, M.** (2014). Peer Effects in Program Participation. American Economic Review, 104(7):2049–2074.

Dauth, W. and **Eppelsheimer, J.** (2020). Preparing the Sample of Integrated Labour Market Biographies (SIAB) for Scientific Analysis: A Guide. Journal for Labour Market Research, 54(1):1–14.

De Giorgi, G., **Pellizzari, M.**, and **Redaelli, S.** (2010). Identification of Social Interactions through Partially Overlapping Peer Groups. American Economic Journal: Applied Economics, 2(2):241–275.

Dengler, K., **Matthes, B.**, and **Paulus, W.** (2014). Occupational Tasks in the German Labour Market. FDZ Methodenreport, 12.

Deshpande, M., **Fadlon, I.**, and **Gray, C.** (2024). How Sticky is Retirement Behavior in the United States? Review of Economics and Statistics, 106(2):370–383.

Dicarlo, E. (2022). How do Firms Adjust to Negative Labor Supply Shocks? Evidence from Migration Outflows. JSTOR.

Doeringer, P. B. and **Piore, M. J.** (1971). Internal Labor Markets and Manpower Analysis. Lexington, Massachusetts: D.C. Heath and Company.

Drechsler, J., **Ludsteck, J.**, and **Moczall, A.** (2023). Imputation der rechtszensierten Tagesentgelte für die BeH. Technical report, Institut für Arbeitsmarkt-und Berufsforschung (IAB), Nürnberg [Institute for . . . .

Drenik, A., **Jäger, S.**, **Plotkin, P.**, and **Schoefer, B.** (2023). Paying Outsourced Labor: Direct Evidence from Linked Temp agency-worker-client Data. Review of Economics and Statistics, 105(1):206–216.

Duflo, E. and **Saez, E.** (2002). Participation and Investment Decisions in a Retirement Plan: The Influence of Colleagues' Choices. Journal of Public Economics, 85(1):121–148.

Duflo, E. and **Saez, E.** (2003). The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment. The Quarterly Journal of Economics, 118(3):815–842.

Dustmann, C., **Fitzenberger, B.**, **Schönberg, U.**, and **Spitz-Oener, A.** (2014). From Sick Man of Europe to Economic Superstar: Germany's Resurgent Economy. Journal of Economic Perspectives, 28(1):167–188.

Dustmann, C., **Glitz, A.**, **Schönberg, U.**, and **Brücker, H.** (2016). Referral-based Job Search Networks. The Review of Economic Studies, 83(2):514–546.

Dustmann, C., **Lindner, A.**, **Schönberg, U.**, **Umkehrer, M.**, and **Vom Berge, P.** (2022). Reallocation Effects of the Minimum Wage. The Quarterly Journal of Economics, 137(1):267–328.

Dustmann, C., **Ludsteck, J.**, and **Schönberg, U.** (2009). Revisiting the German Wage Structure. The Quarterly Journal of Economics, 124(2):843–881.

Etgeton, S., **Fischer, B.**, and **Ye, H.** (2023). The Effect of Increasing Retirement Age on Households' Savings and Consumption Expenditure. Journal of Public Economics, 221:104845.

Felder, L., **Geyer, J.**, and **Haan, P.** (2024). Early Retirement for Early Starters-A Well-targeted Policy for People with High Job Demand? Technical report, Kiel, Hamburg: ZBW-Leibniz Information Centre for Economics.

Ferrari, I., **Kabátek, J.**, and **Morris, T.** (2023). Longer Careers: A Barrier to Hiring and Coworker Advancement? University Ca'Foscari of Venice, Dept. of Economics Research Paper Series No, 6.

Fietz, K. and **Schmeißer, A.** (2024). Racial Peer Effects at Work. World Bank.

Fitzenberger, B. and **Seidlitz, A.** (2020). The 2011 Break in the Part-time Indicator and the Evolution of Wage Inequality in Germany. Journal for Labour Market Research, 54(1):1.

Franz, W. and **Pfeiffer, F.** (2006). Reasons for Wage Rigidity in Germany. Labour, 20(2):255–284.

Friedrich, B. U. and **Hackmann, M. B.** (2021). The Returns to Nursing: Evidence from a Parental-leave Program. The Review of Economic Studies, 88(5):2308–2343.

Frimmel, W., **Horvath, T.**, **Schnalzenberger, M.**, and **Winter-Ebmer, R.** (2018). Seniority wages and the role of firms in retirement. Journal of Public Economics, 164:19–32.

Gallen, Y. (2019). The Effect of Parental Leave Extensions on Firms and Coworkers. Technical report, working paper.

García-Miralles, E. and **Leganza, J. M.** (2024). Joint Retirement of Couples: Evidence from Discontinuities in Denmark. Journal of Public Economics, 230:105036.

Geyer, J., **Haan, P.**, **Lorenz, S.**, **Zwick, T.**, and **Bruns, M.** (2022). Role of Labor Demand in the Labor Market Effects of a Pension Reform. Industrial Relations: A Journal of Economy and Society, 61(2):152–192.

Geyer, J. and **Welteke, C.** (2021). Closing Routes to Retirement for Women: How do they Respond? Journal of Human Resources, 56(1):311–341.

Ginja, R., **Karimi, A.**, and **Xiao, P.** (2023). Employer Responses to Family Leave Programs. American Economic Journal: Applied Economics, 15(1):107–35.

Ginn, J. and **Arber, S.** (1999). Changing Patterns of Pension Inequality: the Shift from State to Private Sources. Ageing & Society, 19(3):319–342.

Glaeser, E. L., **Sacerdote, B. I.**, and **Scheinkman, J. A.** (2003). The Social Multiplier. Journal of the European Economic Association, 1(2-3):345–353.

Glitz, A. (2017). Coworker Networks in the Labour Market. Labour Economics, 44:218–230.

Gohl, N. (2023). Working Longer, Working Stronger? The Forward-Looking Effects of Increasing the Retirement Age on (Un) employment Behaviour. Berlin School of Economics Discussion Paper.

Gohl, N., **Haan, P.**, **Kurz, E.**, and **Weinhardt, F.** (2020). Working Life and Human Capital Investment: Causal Evidence from Pension Reform. IZA Discussion Paper.

Goldin, C., **Kerr, S. P.**, and **Olivetti, C.** (2021). The Other Side of the Mountain: Women's Employment and Earnings over the Family Cycle. IFS Deaton Review of Inequalities.

Goyal, S. (2011). Learning in Networks. In Handbook of Social Economics, volume 1, pages 679–727. Elsevier.

Gregory, T., **Salomons, A.**, and **Zierahn, U.** (2022). Racing with or against the Machine? Evidence on the Role of Trade in Europe. Journal of the European Economic Association, 20(2):869–906.

Grodner, A. and **Kniesner, T. J.** (2008). Labor Supply with Social Interactions: Econometric Estimates and Their Tax Policy Implications. In Work, earnings and other aspects of the employment relation, pages 1–23. Emerald Group Publishing Limited.

Gruber, J. (1997). The Incidence of Payroll Taxation: Evidence from Chile. Journal of Labor Economics, 15(S3):S72–S101.

Gruber, J. and **Wise, D. A.** (2008). Social Security and Retirement around the World. University of Chicago Press.

Gruber, J. and **Wise, D. A.** (2010). Social Security Programs and Retirement around the World: The Relationship to Youth Employment. University of Chicago Press.

Gudgeon, M., **Guzman, P.**, **Schmieder, J. F.**, **Trenkle, S.**, and **Ye, H.** (2023). When Institutions Interact: How the Effects of Unemployment Insurance are Shaped by Retirement Policies. Technical report, National Bureau of Economic Research.

Guvenen, F., **Kuruscu, B.**, **Tanaka, S.**, and **Wiczer, D.** (2020). Multidimensional Skill Mismatch. American Economic Journal: Macroeconomics, 12(1):210–244.

Hanel, B. and **Riphahn, R. T.** (2012). The Timing of Retirement—New Evidence from Swiss Female Workers. Labour Economics, 19(5):718–728.

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

Hensvik, L. and **Rosenqvist, O.** (2019). Keeping the Production Line Running: Internal Substitution and Employee Absence. Journal of Human Resources, 54(1):200–224.

Herbst, D. and **Mas, A.** (2015). Peer Effects on Worker Output in the Laboratory Generalize to the Field. Science, 350(6260):545–549.

Hernæs, E., **Kornstad, T.**, **Markussen, S.**, and **Røed, K.** (2023). Ageing and Labor Productivity. Labour Economics, 82:102347.

Hernæs, E., **Markussen, S.**, **Piggott, J.**, and **Røed, K.** (2016). Pension Reform and Labor Supply. Journal of Public Economics, 142:39–55.

Herrmann, M. A. and **Rockoff, J. E.** (2012). Worker Absence and Productivity: Evidence from Teaching. Journal of Labor Economics, 30(4):749–782.

Holzer, H. J. (1987). Hiring Procedures in the Firm: Their Economic Determinants and Outcomes.

Hsieh, C.-T., **Hurst, E.**, **Jones, C. I.**, and **Klenow, P. J.** (2019). The Allocation of Talent and US Economic Growth. Econometrica, 87(5):1439–1474.

Huebener, M., **Jessen, J.**, **Kuehnle, D.**, and **Oberfichtner, M.** (2024). Parental Leave, Worker Substitutability, and Firms' Employment. The Economic Journal, page ueae114.

Huitfeldt, I., **Kostøl, A. R.**, **Nimczik, J.**, and **Weber, A.** (2023). Internal Labor Markets: A Worker Flow Approach. Journal of Econometrics, 233(2):661–688.

Hut, S. (2024). Impact of Raising the Retirement Age on Firms. Journal of Human Resources.

Hutchens, R. (1986). Delayed Payment Contracts and a Firm's Propensity to Hire Older Workers. Journal of Labor Economics, 4(4):439–457.

Illing, H., **Schwank, H.**, and **Tô, L. T.** (2024). Hiring and the Dynamics of the Gender Gap. ECONtribute.

Imbens, G. and **Kalyanaraman, K.** (2012). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. The Review of Economic Studies, 79(3):933–959.

Isen, A. (2013). Dying to Know: Are Workers Paid Their Marginal Product? Technical report, University of Pennsylvania working paper.

Jacobebbinghaus, P. and **Seth, S.** (2007). The German Integrated Employment Biographies Sample IEBS. Journal of Contextual Economics–Schmollers Jahrbuch, pages 335–342.

Jacobson, L. S., **LaLonde, R. J.**, and **Sullivan, D. G.** (1993). Earnings Losses of Displaced Workers. The American Economic Review, pages 685–709.

Jäger, S. and **Heining, J.** (2022). How Substitutable are Workers? Evidence from Worker Deaths. Technical report, National Bureau of Economic Research.

Jaravel, X., **Petkova, N.**, and **Bell, A.** (2018). Team-specific Capital and Innovation. American Economic Review, 108(4-5):1034–1073.

Jefferson, T. (2009). Women and Retirement Pensions: A Research Review. Feminist Economics, 15(4):115–145.

Johnsen, J. V., **Ku, H.**, and **Salvanes, K. G.** (2023). Competition and Career Advancement. Review of Economic Studies, pages 2954–2980.

Johnsen, J. V., **Ku, H.**, and **Salvanes, K. G.** (2024). Competition and Career Advancement. Review of Economic Studies, 91(5):2954–2980.

Johnsen, J. V., **Vaage, K.**, and **Willén, A.** (2022). Interactions in Public Policies: Spousal Responses and Program Spillovers of Welfare Reforms. The Economic Journal, 132(642):834–864.

Kalwij, A., **Kapteyn, A.**, and **De Vos, K.** (2010). Retirement of Older Workers and Employment of the Young. De Economist, 158(4):341–359.

Kato, T. and **Shu, P.** (2009). Peer Effects, Social Networks, and Intergroup Competition in the Workplace. Technical report, Aarhus University.

Kleven, H., **Landais, C.**, and **Søgaard, J. E.** (2019). Children and Gender Inequality: Evidence from Denmark. American Economic Journal: Applied Economics, 11(4):181–209.

Kosfeld, R. and **Werner, A.** (2012). Deutsche Arbeitsmarktregionen–neuabgrenzung nach den Kreisgebietsreformen 2007–2011. Raumforschung und Raumordnung, 70(1):49–64.

Krstic, A. and **Hideg, I.** (2019). The Effect of Taking a Paternity Leave on Men's Career Outcomes: The Role of Communality Perceptions. In Academy of Management Proceedings, page 13912. Academy of Management Briarcliff Manor, NY 10510. Paper no. 1.

Krueger, A. B. and **Pischke, J.-S.** (1992). The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation. Journal of Labor Economics, 10(4):412–437.

Kuhn, P. and **Yu, L.** (2021). How Costly is Turnover? Evidence from Retail. Journal of Labor Economics, 39(2):461–496.

Lalive, R., **Magesan, A.**, and **Staubli, S.** (2023). How Social Security Reform Affects Retirement and Pension Claiming. American Economic Journal: Economic Policy, 15(3):115–150.

Lalive, R. and **Parrotta, P.** (2017). How does Pension Eligibility Affect Labor Supply in Couples? Labour Economics, 46:177–188.

Lalive, R. and **Staubli, S.** (2015). How does Raising Women's Full Retirement Age Affect Labor Supply, Income, and Mortality. NBER Working Paper, 18660:17.

Lancee, B. and **Radl, J.** (2012). Social Connectedness and the Transition from Work to Retirement. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 67(4):481–490.

Lazear, E. P. (1979). Why is there Mandatory Retirement? Journal of Political Economy, 87(6):1261–1284.

Lazear, E. P. (2009). Firm-specific Human Capital: A Skill-weights Approach. Journal of Political Economy, 117(5):914–940.

Lazear, E. P. and **Oyer, P.** (2004). Internal and External Labor Markets: A Personnel Economics Approach. Labour Economics, 11(5):527–554.

Liebman, J. B. and **Luttmer, E. F.** (2015). Would people behave differently if they better understood social security? evidence from a field experiment. American Economic Journal: Economic Policy, 7(1):275–299.

Lorenz, S., **Pfister, M.**, **Zwick, T.**, and **others** (2018). Identification of the Statutory Retirement Dates in the Sample of Integrated Labor Market Biographies (SIAB). FDZ Methodenreport, 8:2018.

Manoli, D. S. and **Weber, A.** (2016). The Effects of the Early Retirement Age on Retirement Decisions. Technical report, National Bureau of Economic Research.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. The Review of Economic Studies, 60(3):531–542.

Martins, P. S., **Novo, Á. A.**, and **Portugal, P.** (2009). Increasing the Legal Retirement Age: The Impact on Wages, Worker Flows and Firm Performance. Technical report, IZA Discussion Papers.

Mas, A. and **Moretti, E.** (2009). Peers at Work. American Economic Review, 99(1):112–145.

Mastrobuoni, G. (2009). Labor Supply Effects of the Recent Social Security Benefit Cuts: Empirical Estimates using Cohort Discontinuities. Journal of Public Economics, 93(11-12):1224–1233.

McPherson, M., **Smith-Lovin, L.**, and **Cook, J. M.** (2001). Birds of a Feather: Homophily in Social Networks. Annual Review of Sociology, 27(1):415–444.

Meekes, J. and **Hassink, W. H.** (2022). Gender Differences in Job Flexibility: Commutes and Working Hours after Job Loss. Journal of Urban Economics, 129:103425.

Meekes, J. and **Lent, M.** (2025). The impact of neighbour, colleague, and family peers on parental labour supply. JSTOR.

Messina, J., **Sanz-de Galdeano, A.**, and **Terskaya, A.** (2023). Birds of a Feather Earn Together. Gender and Peer Effects at the Workplace. IEB Working Paper 2023/10.

Mohnen, P. (2025). The Impact of the Retirement Slowdown on the US Youth Labor Market. Journal of Labor Economics, 43(1):203–246.

Muehlemann, S. and **Pfeifer, H.** (2016). The Structure of Hiring Costs in Germany: Evidence from Firm-Level Data. Industrial Relations: A Journal of Economy and Society, 55(2):193–218.

Neal, D. (1995). Industry-specific Human Capital: Evidence from Displaced Workers. Journal of Labor Economics, 13(4):653–677.

Nicoletti, C., **Salvanes, K. G.**, and **Tominey, E.** (2018). The Family Peer Effect on Mothers' Labor Supply. American Economic Journal: Applied Economics, 10(3):206–234.

Oakes, P., **Turner, J. C.**, **Hogg, M. A.**, **Reicher, S.**, and **Wetherell, M.** (1987). Rediscovering the Social Group: A Self-Categorization Theory.

Oberfichtner, M. and **Schnabel, C.** (2019). The German Model of Industrial Relations:(Where) Does it Still Exist? Jahrbücher für Nationalökonomie und Statistik, 239(1):5–37.

Oral, E., **Rabaté, S.**, and **Seibold, A.** (2024). The Social Multiplier of Pension Reform. Technical report, CESifo Working Paper.

Paulus, W., **Matthes, B.**, and **others** (2013). The German Classification of Occupations 2010–Structure, Coding and Conversion Table. FDZ-Methodenreport, 8:2013.

Pink, S., **Leopold, T.**, and **Engelhardt, H.** (2014). Fertility and Social Interaction at the Workplace: Does Childbearing Spread among Colleagues? Advances in Life Course Research, 21:113–122.

Pissarides, C. A. (2000). Equilibrium Unemployment Theory. MIT press.

Plöger, D. (2024). How Do Establishments Choose Their Location? Taxes, Monopsony, and Productivity. IZA Discussion Paper No. 17742.

Poege, F., **Gaessler, F.**, **Hoisl, K.**, **Harhoff, D.**, and **Dorner, M.** (2025). Filling the Gap: The Consequences of Collaborator Loss in Corporate R&D. Management Science.

Queisser, M. and **Whitehouse, E.** (2006). Neutral or Fair?: Actuarial Concepts and Pension-system Design. OECD.

Rabaté, S., **Jongen, E.**, and **Atav, T.** (2024). Increasing the Retirement Age: Policy Effects and Underlying Mechanisms. American Economic Journal: Economic Policy, 16(1):259–291.

Rege, M., **Telle, K.**, and **Votruba, M.** (2012). Social Interaction Effects in Disability Pension Participation: Evidence from Plant Downsizing. The Scandinavian Journal of Economics, 114(4):1208–1239.

Riphahn, R. T. and **Schrader, R.** (2021). Reforms of an Early Retirement Pathway in Germany and their Labor Market Effects. Available at SSRN 3982024.

Rosenfeld, R. A., **Trappe, H.**, and **Gornick, J. C.** (2004). Gender and work in Germany: Before and after reunification. Annu. Rev. Sociol., 30:103–124.

Sauvagnat, J. and **Schivardi, F.** (2024). Are Executives in Short Supply? Evidence from Death Events. Review of Economic Studies, 91(1):519–559.

Saygin, P. O., **Weber, A.**, and **Weynandt, M. A.** (2021). Coworkers, Networks, and Job-search Outcomes among Displaced Workers. ILR Review, 74(1):95–130.

Schmidtlein, L., **Seth, S.**, and **Vom Berge, P.** (2020). Sample of Integrated Employer Employee Data (SIEED) 1975 2018. Technical report, Institut für Arbeitsmarkt-und Berufsforschung (IAB), Nürnberg [Institute for . . . .

Schmieder, J. F., **von Wachter, T.**, and **Bender, S.** (2016). The Effect of Unemployment Benefits and Nonemployment Durations on Wages. American Economic Review, 106(3):739–777.

Schmutte, I. M. and **Skira, M. M.** (2023). The Response of Firms to Maternity Leave and Sickness Absence. Journal of Human Resources.

Seibold, A. (2021). Reference Points for Retirement Behavior: Evidence from German Pension Discontinuities. American Economic Review, 111(4):1126–1165.

Selin, H. (2017). What Happens to the Husband's Retirement Decision When the Wife's Retirement Incentives Change? International Tax and Public Finance, 24:432–458.

Smith, K. U. (1965). Behavior Organization and Work: A New Approach to Industrial Behavior Science. College Printing and Typing Company.

Staubli, S. and **Zweimüller, J.** (2013). Does Raising the Early Retirement Age Increase Employment of Older Workers? Journal of Public Economics, 108:17–32.

Stole, L. A. and **Zwiebel, J.** (1996a). Intra-firm Bargaining under Non-binding Contracts. The Review of Economic Studies, 63(3):375–410.

Stole, L. A. and **Zwiebel, J.** (1996b). Organizational Design and Technology Choice under Intrafirm Bargaining. The American Economic Review, pages 195–222.

Stutzer, A. and **Lalive, R.** (2004). The Role of Social Work Norms in Job Searching and Subjective Well-being. Journal of the European Economic Association, 2(4):696–719.

Tajfel, H. (1981). Human Groups and Social Categories: Studies in Social Psychology. Cup Archive.

Topel, R. H. and **Ward, M. P.** (1992). Job Mobility and the Careers of Young Men. The Quarterly Journal of Economics, 107(2):439–479.

Tophoven, S., **Du Prel, J.-B.**, **Peter, R.**, and **Kretschmer, V.** (2015). Working in Gender-dominated Occupations and Depressive Symptoms: Findings from the Two Age Cohorts of the lidA Study. Journal for Labour Market Research, 48(3):247–262.

Vermeer, N., **van Rooij, M.**, and **van Vuuren, D.** (2019). Retirement Age Preferences: The Role of Social Interactions and Anchoring at the Statutory Retirement Age. De Economist, 167(4):307–345.

Vestad, O. L. (2013). Labour Supply Effects of Early Retirement Provision. Labour Economics, 25:98–109.

Völker, B. and **Flap, H.** (2001). Weak Ties as a Liability: The Case of East Germany. Rationality and Society, 13(4):397–428.

Waldman, M. (2003). Ex ante versus ex post Optimal Promotion Rules: The Case of Internal Promotion. Economic Inquiry, 41(1):27–41.

Welteke, C. (2015). Peers at Work-a Brief Overview of the Literature on Peer Effects at the Workplace and the Policy Implications. DIW Roundup: Politik im Fokus.

Welteke, C. and **Wrohlich, K.** (2019). Peer Effects in Parental Leave Decisions. Labour Economics, 57:146–163.

Ye, H. (2020). The Effect of Pension Subsidies on the Retirement Timing of Older Women. Journal of the European Economic Association.

Yi, M., **Müller, S.**, and **Stegmaier, J.** (2024). Industry Mix, Local Labor Markets, and the Incidence of Trade Shocks. Journal of Labor Economics, 42(3):000–000.

Zweimüller, J., **Winter-Ebmer, R.**, and **Falkinger, J.** (1996). Retirement of Spouses and Social Security Reform. European Economic Review, 40(2):449–472.

Zwick, T., **Bruns, M.**, **Geyer, J.**, and **Lorenz, S.** (2022). Early Retirement of Employees in Demanding Jobs: Evidence from a German Pension Reform. The Journal of the Economics of Ageing, 22:100387.