Estudios sobre la Economía Española

II Jornada sobre el sistema público de pensiones y seguridad social

Delaying Retirement, Reshaping Exits, and the Gender Mortality Gap: Evidence from the 2011 Spanish Pension Reform

SERGI JIMÉNEZ-MARTÍN

Estudios sobre la Economía Española 2026/23
Junio 2026

fedea

Delaying Retirement, Reshaping Exits, and the Gender Mortality Gap: Evidence from the 2011 Spanish Pension Reform[*]

Sergi Jiménez-Martín[†]

July 4, 2026

Abstract

We estimate causal effects of the 2011 Spanish pension reform—which raised the normal retirement age (NRA) from 65 to 67, conditioning access on accumulated contribution years—on retirement timing, exit route, and premature mortality. Using data from the *Muestra Continua de Vidas Laborales* we find that the reform delayed retirement by about eight months on average for fully treated individuals belonging to the 1948-1957 cohorts, the cohorts observable in our data. Because the reform phases in gradually until the 1962 cohort—the first to face the fully phased-in rules, which then apply unchanged to all later cohorts—the total effect will be larger as more intensely affected cohorts complete their working lives.

The dominant response is a reallocation away from retirement at and before age 65 toward retirement after 65, and is markedly larger for women, who are less likely to meet the long-career exemption. The reform raises premature mortality, significantly so for men; the implied gender differential widens the male mortality excess and narrows the female survival advantage, though sample-weighted gaps are sensitive to the large share of male controls.

JEL: H55, J26, I12, J14, J16 **Keywords:** pension reform, retirement age, early retirement, mortality, gender, Spain, Sun–Abraham estimator, MCVL, robust standard errors

---

[*]This research uses microdata from the *Muestra Continua de Vidas Laborales* (MCVL), provided by the Spanish Social Security Administration. I gratefully acknowledge funding from Agencia Estatal de Investigación through PID2023-147602OB-I00. All errors are my own

[†]Universitat Pompeu Fabra; Barcelona School of Economics (BSE); and FEDEA. E-mail: sergi.jimenez@upf.edu.

1 Introduction

Raising the statutory retirement age has become the dominant policy response to population ageing across the developed world. Between 2000 and 2023, every OECD member has either legislated increases in the normal retirement age (NRA), tightened access to early retirement, or both (OECD, 2023). Germany gradually raised its NRA from 65 to 67 (phased 2012–2029); France increased the minimum retirement age from 60 to 62 and the full-pension age from 65 to 67 (Bozio et al., 2021a); the United Kingdom equalised and raised the State Pension Age for women (Cribb et al., 2016); Italy's Fornero reform of 2011 moved the retirement age to 67; Austria, the Netherlands, Sweden, Denmark, and Portugal all enacted comparable reforms. The fiscal motivation is clear—pension expenditure-to-GDP ratios are rising across the continent—but the behavioural and health consequences of mandating longer working lives remain empirically contested.

Two distributional concerns stand out beyond the aggregate question of whether workers delay retirement as intended. First, a body of evidence suggests that workers in physically demanding occupations bear a disproportionate health cost from being required to work longer (Bellés-Obrero et al., 2022; Bloemen et al., 2017). Second, and less studied, is the gender dimension. Conditioning the normal retirement age on accumulated contribution years—as several European reforms do, including the Spanish reform studied here—creates a structural gender asymmetry: because women's careers are typically shorter and more fragmented than men's, due to care responsibilities, part-time work, and labour market interruptions, they are less likely to meet the long-career threshold that grants access to the lower retirement age. A reform nominally presented as neutral can therefore impose a larger delay on women in practice.

This paper addresses both concerns in the context of the Spanish pension reform enacted through *Ley 27/2011* and phased in gradually from 2013. Spain is an instructive case for three reasons. First, the reform introduced a particularly sharp *contribution-contingent* design: the NRA ranges from 65 (for workers with 37 or more contribution years by 2027) to 67 (for those below the threshold), creating within-cohort variation in treatment intensity that mirrors the within-cohort variation in career lengths. Second, Spain exhibits one of the larger gender gaps in contribution histories in the OECD (OECD, 2023), making the contribution-conditioning mechanism especially salient for gender outcomes. Third, Spain's rich administrative data—the *Muestra Continua de Vidas Laborales* (MCVL)—links complete contribution histories, retirement spells, and civil-registry mortality records at the individual level, enabling a comprehensive analysis of the reform's effects on timing, exit route, and longevity.

We exploit the reform's cohort-varying, contribution-contingent dose structure. The phase-in schedule implies that workers born in 1948 first encountered the new rules when turning 65 in 2013, and subsequent cohorts face progressively larger NRA increases. Crucially, within each affected cohort workers whose expected contribution years exceed the cohort-specific exemption threshold are unaffected and serve as the comparison group. We organise identification around three groups: *treated* workers (fully exposed to higher NRAs), *partially treated* workers (partially exposed), and *controls* (essentially unaffected). Within this framework we adopt the cohort-interaction estimator of Sun and Abraham (2021) to recover cohort-specific average treatment effects on the treated (ATTs) across 15 birth cohorts (1943–1957) while guarding against the heterogeneous-treatment bias of conventional two-way fixed effects (Callaway and Sant'Anna, 2021; Wooldridge, 2021).

Our findings can be summarised under three headings.

Retirement timing. The Sun–Abraham event-study estimates show a retirement delay that accelerates across cohorts, reaching an average of **8.1 months** for the treated cohorts (born 1948–1957) relative to the last pre-reform cohort (born 1947) under heteroskedasticity-robust inference. The dose-response estimates confirm that realised retirement tracks exposure intensity: beyond the level effect of being treated, each additional month of NRA delay translates into about one further month of realised retirement age overall ($t = 31.87$). Women's retirement delay is in fact *larger* than men's (10.0 vs. 3.1 months in the event study), reflecting their greater treatment intensity.

Retirement type. The dominant behavioural response is a reallocation *away* from retirement at exactly age 65 *toward* retirement after 65, measured in absolute age rather than relative to the changing NRA. The probability of retiring at exactly 65 falls by 10.3 percentage points on average, while the probability of retiring after 65 rises by 26.7 pp. The probability of retiring before 65 also falls by 17.7 pp, partly because the minimum ages for involuntary and voluntary early retirement were also raised. Both the rise in post-65 retirement and the fall in pre-65 retirement are larger for women than for men, reflecting their higher treatment intensity and greater concentration below the long-career threshold.

Mortality. The 2×2 DiD analysis (Table 6) shows a positive effect on 60–67 mortality that is statistically significant for men ($+1.59$ pp, $t = 3.07$), while the pooled and female effects are positive but not significant at conventional levels with heteroskedasticity-robust standard errors ($+0.62$ pp, $t = 1.79$ and $+0.79$ pp, $t = 1.70$, respectively). The implied gender differential (Women − Men) is -0.81 pp: the reform widens the male mortality excess by 0.81 pp ($+15\%$ of the pre-existing 5.53 pp gap). Sample-weighted calculations using the observed cell composition confirm a widening of $+0.77$ pp when all estimated coefficients are used, but the effect on the observed population gap is muted and sign-sensitive: it is close to zero without covariates (-0.10 pp) and can flip sign depending on how imprecisely estimated nuisance terms are handled, reflecting the dilution of the male ATT by the large share of unaffected male controls. An event-study analysis of social security wealth shows that the reform did not substantially affect the pension wealth of treated workers relative to controls from the 1947 reference cohort onwards, pointing to the compelled extension of working life—rather than income losses—as the operative channel behind the mortality response.

The paper makes four contributions. First, it provides new causal evidence on the behavioural consequences of the 2011 Spanish reform using the most robust available estimator for staggered-adoption designs. Second, it is the first paper to document systematically the contribution-contingent gender asymmetry of the reform: by showing that women disproportionately bear the burden of higher NRAs while also substituting more strongly into early retirement, we illuminate a mechanism relevant for all contribution-conditioned pension designs. Third, we advance the analysis of retirement and mortality by distinguishing two channels of reform exposure (NRA shift versus contribution-shortfall depth) and assessing their behavioural and mortality consequences separately. Fourth, the paper contributes methodologically by pairing compact dose-response regressions with the full Sun–Abraham event-study apparatus, providing both a parsimonious first look and a rich dynamic picture of the reform's cohort-by-cohort effects.

The paper contributes to quasi-experimental research on retirement impact of pension reforms (Manoli and Weber, 2016; Staubli and Zweimüller, 2013; Geyer and Welteke, 2019; Geyer et al., 2020; Cribb et al., 2016; Bozio et al., 2021a; Hernæs et al., 2013), to the Spanish pension literature (Jiménez-Martín and Sánchez-Martín, 2007; Boldrin, Jiménez-Martín, and Peracchi, 1999; Jiménez-Martín and Sánchez-Martín, 2004; García-Pérez, Jiménez-Martín, and Sánchez-Martín, 2013; Sánchez-Martín, García-Pérez, and Jiménez-Martín, 2014; Bellés-Obrero et al., 2022, 2025), to the debate on retirement and health (Gorry et al., 2018; Fitzpatrick and Moore, 2018; Bloemen et al., 2017; Coe and Zamarro, 2011; Hallberg, Johansson, and Josephson, 2015; Carrino, Glaser, and Avendano, 2020), and to studies of gender-specific pension reform effects (Carta and De Philippis, 2021; Geyer and Welteke, 2019; Della Giusta and Longhi, 2021; García-Gómez, Jiménez-Martín, and Vall Castelló, 2012).

The remainder of the paper is organised as follows. Section 2 reviews the relevant literature. Section 3 describes the institutional context. Section 3.2 presents the data and variable construction. Section 4 lays out the identification strategy.Section 5 presents the main results. Section 6 discusses gender heterogeneity. Section 7 analyses the reform's effects on premature mortality in detail. Section 8 concludes.

2 Literature Review

2.1 Labour supply effects of pension reform

A large empirical literature documents that statutory retirement ages and early retirement rules shape retirement behaviour (Gruber and Wise, 1999, 2004; Coile and Gruber, 2007). Staubli and Zweimüller (2013) estimate large pension take-up responses to Austrian early retirement age increases with sizeable spillovers onto unemployment and disability insurance, and Manoli and Weber (2016) document pension-wealth elasticities of retirement timing in the same setting. Cribb et al. (2016) report large UK labour-force responses to the equalisation of women's state pension age, Geyer and Welteke (2019) finds claim delays of roughly one year following closure of an age-60 early retirement route for German women, and Bozio et al. (2021a) estimate French 2010 labour-supply responses commensurate with the mechanical incentive change; Laun and Wallenius (2015) provides a Swedish structural benchmark. The Spanish 2011 reform adds two features absent from most prior work: a contribution-contingent eligibility rule and a within-cohort comparison group of long-career workers, allowing us to isolate the depth of individual exposure rather than aggregate cohort effects.

2.2 Evidence for Spain

The Spanish pension system has attracted sustained academic attention. Earlier work documents the implicit tax on continued work in the pre-reform system (Boldrin, Jiménez-Martín, and Peracchi, 1999), monetary incentives to retire under Spanish rules (Jiménez-Martín and Sánchez-Martín, 2004), life-cycle distortions induced by minimum pension guarantees (Jiménez-Martín and Sánchez-Martín, 2007), hazard-model estimates of retirement responses around the 2002 reform (García-Pérez, Jiménez-Martín, and Sánchez-Martín, 2013), and the welfare consequences of delaying both normal and early retirement ages in structural simulations (Sánchez-Martín, García-Pérez, and Jiménez-Martín, 2014)—a benchmark for the reduced-form evidence reported here. García-Gómez, Jiménez-Martín, and Vall Castelló (2012) map pathways from employment into disability, unemployment, and retirement for older Spanish workers. Bellés-Obrero et al. (2022) study a different Spanish reform (the 1967 rule restricting early retirement access), finding that removing early retirement delayed exit by around half a year and raised mortality, especially in physically demanding occupations; we complement this by examining a reform that delays retirement through contribution-contingent NRA increases rather than outright closure of early-exit pathways. Bellés-Obrero et al. (2025) examine the distributional consequences of Spanish social security reforms across the retiree income distribution.

A complementary Spanish strand examines gender-differential mortality from earlier reforms. Bellés-Obrero, Jiménez-Martín, and Vall Castelló (2022) exploit the 1980 Workers' Statute (which raised the minimum legal working age from 14 to 16 while keeping compulsory schooling at 14) in a within-cohort DiD across birth months, finding that the reform reduced mortality at ages 14–29 (driven by traffic accidents) but raised mortality among women aged 30–45 through HIV/AIDS and circulatory diseases—a behavioural mechanism rooted in the post-Franco gender equalisation process. Our paper identifies the opposite gender pattern at the other end of working life: it is men, not women, who bear the larger mortality cost, operating through occupational rather than behavioural channels.

2.3 Health and mortality effects of delayed retirement

The causal effect of retirement timing on health is theoretically ambiguous: retirement removes workers from stressful or physically demanding jobs (potentially health-improving) but also strips away income, social interaction, cognitive stimulation, and daily structure (potentially harmful). Advancing and delaying retirement need not be symmetric (Hagen, 2018), and since most quasi-experimental evidence studies earlier-retirement windows while the 2011 Spanish reform raised the statutory age, delayed-retirement evidence is the more directly relevant benchmark.

On earlier retirement, the evidence leans toward null or beneficial effects among white-collar workers (Coe and Zamarro, 2011; Gorry et al., 2018), but turns adverse or heterogeneous in physically demanding occupations. Bloemen et al. (2017) document mortality reductions among Dutch male blue-collar workers eligible for early retirement, Hallberg, Johansson, and Josephson (2015) find cardiovascular-mortality reductions among Swedish army officers, and Hernæs et al. (2013) report no Norwegian mortality effect of access two to five years earlier. The main exception is Kuhn et al. (2020), who find that early Austrian labour-force exit raises male mortality before 67, plausibly through unemployment scarring; Fitzpatrick and Moore (2018) document mortality effects of US Social Security eligibility at 62.

Evidence on *delaying* retirement is scarcer and more mixed. Hagen (2018) estimates an imprecise null effect of a two-year Swedish statutory-age increase on mortality at 69; Bozio, Garrouste, and Perdrix (2021b) find no effect on French mortality at ages 61–79 following the 2010 reform; Saporta-Eksten, Shurtz, and Weisburd (2021) find higher mortality at 75–85 (but not 65–74) in Israel, suggesting that delayed-retirement effects may emerge only at longer horizons and for specific subgroups. Carrino, Glaser, and Avendano (2020) study the UK female state pension age reform and find adverse health effects of mandatory employment extension concentrated among women in high-demand, low-control occupations, including significant increases in clinical depression diagnoses—a gender–occupational interaction directly relevant for our analysis. For Spain, Bellés-Obrero et al. (2022) again find that forcing workers to delay retirement increases mortality, especially in manual occupations, motivating our attention to occupational heterogeneity and the gender mortality differential.

2.4 Gender and pension reform

Gender differences in pension reform effects have received growing attention. Geyer and Welteke (2019) and Geyer et al. (2020) document large female employment responses to the German age-60 closure together with substantial income heterogeneity by earnings quintile, with lower-income women bearing a disproportionate welfare cost. Carta and De Philippis (2021) shows that the Italian Fornero reform's unexpected NRA increase induces forward-looking labour-supply adjustments among middle-aged women that spill over onto their husbands' retirement decisions—a family-level anticipation effect partially offsetting fiscal externalities. Della Giusta and Longhi (2021) find that UK state pension age increases generate significant adverse effects on women's financial wellbeing, mental health, and life satisfaction, concentrated among lower-educated and unpartnered women.

A central mechanism linking reform design to gender outcomes is the *contribution-contingent* exemption structure common to several European reforms: when a lower retirement age is conditional on long uninterrupted careers, women—who accumulate fewer contribution years due to care responsibilities, part-time employment, and labour-market interruptions—are systematically more exposed. In Spain, OECD (2023) document a gender pension gap of around 34% of men's average pension, one of the OECD's largest; in our MCVL sample, women account for 62% of treated workers but only 29% of controls. We provide the first quasi-experimental evidence on this contribution-conditioned gender asymmetry using the 2011 reform, complementing the structural simulations of Sánchez-Martín, García-Pérez, and Jiménez-Martín (2014) and the gender-differential mortality evidence from the earlier Spanish minimum-working-age reform in Bellés-Obrero, Jiménez-Martín, and Vall Castelló (2022).

3 Institutional Context, data and variables

3.1 Institutional context

3.1.1 The pre-reform system

Before 2013, Spain's public pension system set the standard retirement age at 65 for workers with at least 15 years of contributions, with full benefits requiring 35 years. The replacement rate reached 100% of the regulatory base after 35 years. Voluntary early retirement was possible from age 61, with actuarial reductions of 6–8% per year of anticipation (Jiménez-Martín and Sánchez-Martín, 2007). Gender gaps were substantial: women's shorter average contribution histories generated lower pensions and greater reliance on minimum-pension supplements (Boldrin, Jiménez-Martín, and Peracchi, 1999; Jiménez-Martín and Sánchez-Martín, 2007; OECD, 2023).

3.1.2 The 2011 reform

Law 27/2011 (*BOE* 02-08-2011) introduced a phased transition in the normal retirement age from 65 to 67, starting in January 2013 (Conde and Gonzalez , 2012).

Normal retirement age. The NRA is now conditional on contribution history. Workers with 38.5 or more years of contributions in 2027 retain a NRA of 65 (this threshold was 35 years in 2013 and rises by 0.25 years per calendar year until it reaches 38 years and 6 months in 2027). All other workers face a NRA that increases from 65 years and 1 month in 2013 to 67 years in 2027, at a rate of one month per year from 2013 to 2018, and 2 months per year thereafter.

Early retirement. Voluntary early retirement was pushed gradually from age 61 to age 63 (two years before the applicable NRA), with actuarial reductions of 6.5–7.5% per year of anticipation. Involuntary early retirement (firm closure, collective redundancy) remained possible from age 61 with at least 33 years of contributions and age 63 in 2027 with fewer than 33 years.

Other reform components, common to treated and controls. Beyond the retirement-age provisions, Law 27/2011 also modified the benefit formula itself. The computation period of the regulatory base (*base reguladora*) was gradually extended from the last 15 to the last 25 years of contributions (phased in between 2013 and 2022); the number of contribution years required to reach 100% of the regulatory base rose from 35 to 37, with a correspondingly flatter accrual scale in between; and new increments were introduced for workers who prolong their careers beyond the applicable NRA. Crucially for our design, these benefit-formula provisions apply uniformly by calendar year of retirement, irrespective of a worker's position relative to the contribution-years threshold: they reduce pension generosity for treated and control workers alike and therefore add no differential variation between the two groups. Their common incidence does, however, imply that pension wealth declines for all retiring cohorts, which is why Section 7 examines social security wealth directly when assessing candidate drivers of the mortality response.

Phase-in schedule. Table 1 summarises the full phase-in schedule by calendar year, distinguishing the two groups whose NRA trajectories diverge under the reform: workers below the contributory-years threshold (the *treated* group, facing a rising NRA) and workers at or above that threshold (the *control* group, retaining NRA $= 65$). A third collective, *mutualistas*—workers who began contributing to a labour mutual fund (*mutualidad laboral*) before 1 January 1967—retains, under the Fourth Transitional Provision of the General Social Security Act (*Disposición Transitoria 4ª de la Ley General de la Seguridad Social*, hereafter DT 4ª), the pre-reform right to claim a pension from age 60, with an actuarial reduction of 8% per year of anticipation relative to 65 (Ministerio de Empleo y Seguridad Social, 2011; García-Gómez, Jiménez-Martín, and Vall Castelló, 2012). Crucially, Law 27/2011 did not repeal DT 4ª: mutualistas keep their early retirement rights between 60 and 64, but their normal retirement is governed from 2013 onwards by the same principles as for all other workers, so the treated/control schedule of the table applies to them in full once they reach the applicable NRA (see table note c). The normal retirement age (NRA) column shows the two-tier structure that underlies our treatment definition: workers whose contribution record at the relevant date meets or exceeds the year-specific threshold retain the NRA at 65 throughout the phase-in, while workers below the threshold face a NRA that climbs gradually to 67 by 2027. Voluntary and involuntary early retirement ages move in lockstep with the applicable NRA, fixed at two and four years below it respectively, so that an increase in the NRA mechanically tightens both early-exit margins for the below-threshold group while leaving the at-or-above-threshold group's early retirement ages anchored to the constant NRA of 65. The schedule stabilises in 2027: the 1962 birth cohort, turning 65 in that year, is the first to face the fully phased-in rules, and all subsequent cohorts face the same, constant incentive structure. This table is the institutional primitive underlying every dose and outcome measure used in the empirical analysis below; subsequent references to "the applicable NRA", to the minimum early-retirement ages, or to the phase-in intensity by cohort all refer back to the schedule in Table 1.

Table 1: Phase-In Schedule of the 2011 Reform, by Contribution Years and Calendar Year

		Below threshold (Treated)				*At/above threshold (Control)*		
Year		NRA	Vol.[a]	Invol.[b]		NRA	Vol.[a]	Invol.[b]
2013	(<35y 3m)	65y 1m	63y 1m	61y 1m	(≥35y 3m)	65y	63y	61y
2014	(<35y 6m)	65y 2m	63y 2m	61y 2m	(≥35y 6m)	65y	63y	61y
2015	(<35y 9m)	65y 3m	63y 3m	61y 3m	(≥35y 9m)	65y	63y	61y
2016	(<36y 0m)	65y 4m	63y 4m	61y 4m	(≥36y 0m)	65y	63y	61y
2017	(<36y 3m)	65y 5m	63y 5m	61y 5m	(≥36y 3m)	65y	63y	61y
2018	(<36y 6m)	65y 6m	63y 6m	61y 6m	(≥36y 6m)	65y	63y	61y
2019	(<36y 9m)	65y 8m	63y 8m	61y 8m	(≥36y 9m)	65y	63y	61y
2020	(<37y 0m)	65y 10m	63y 10m	61y 10m	(≥37y 0m)	65y	63y	61y
2021	(<37y 3m)	66y 0m	64y 0m	62y 0m	(≥37y 3m)	65y	63y	61y
2022	(<37y 6m)	66y 2m	64y 2m	62y 2m	(≥37y 6m)	65y	63y	61y
2023	(<37y 9m)	66y 4m	64y 4m	62y 4m	(≥37y 9m)	65y	63y	61y
2024	(<38y 0m)	66y 6m	64y 6m	62y 6m	(≥38y 0m)	65y	63y	61y
2025	(<38y 3m)	66y 8m	64y 8m	62y 8m	(≥38y 3m)	65y	63y	61y
2026	(<38y 3m)	66y 10m	64y 10m	62y 10m	(≥38y 3m)	65y	63y	61y
2027	(<38y 6m)	67y 0m	65y 0m	63y 0m	(≥38y 6m)	65y	63y	61y

Sources: Ley 27/2011 (*BOE* 02-08-2011); Disposición Transitoria 4ª de la Ley General de la Seguridad Social (RDL 8/2015). Contribution-years thresholds (in parentheses) are measured at the date of the qualifying event and define the two groups whose NRA trajectories diverge under the reform. The below-threshold cutoffs in the left block are the cohort-specific exemption values used to define treatment status (Section 4.1); the at-or-above column reproduces the constant-NRA path of control workers. The schedule remains constant from 2027 onwards: the 1962 birth cohort, which turns 65 in 2027, is the first to face the fully phased-in rules, and all later cohorts face the same thresholds and retirement ages.

[a] Voluntary early retirement age, fixed at two years below the applicable NRA. Actuarial reduction of 6.5–7.5% per year of anticipation.

[b] Involuntary early retirement age (dismissal, collective redundancy), fixed at four years below the applicable NRA. Requires at least 33 contribution years through 2020; converges toward the NRA threshold thereafter.

[c] *Mutualistas* (workers who contributed to a *mutualidad laboral* before 1 January 1967, protected under DT 4ª LGSS) retain their pre-reform right to early retirement from age 60, with an actuarial reduction of up to 8% per year of anticipation; their normal retirement, however, is governed from 2013 onwards by the same principles as for all other workers, so the treated/control schedule in this table applies to them in full once they reach the applicable NRA.

Gender asymmetry. The contribution-contingent design creates a systematic gender asymmetry. In our sample, 62% of treated workers (those below the exemption threshold) are women, compared with only 29% of controls—a gap that reflects the shorter and more fragmented contribution histories characteristic of women's careers in Spain. The reform's NRA increase therefore falls disproportionately on female workers.

Cohort exposure. Workers born in 1948 are the first cohort subject to the new rules when reaching age 65 in 2013. Pre-reform cohorts (born before 1948) are unaffected since they could retire at age 65 regardless of the number of years contributed, above a minimum of 15 years (two of which within the 15 years preceding retirement). This generates clean cohort-level variation in treatment intensity that forms the basis of our identification strategy. Our estimation window covers cohorts up to 1957, who turn 65 in 2022 and face just over half of the full two-year NRA increase; later cohorts are progressively more exposed, with the 1962 cohort (turning 65 in 2027) the first to face the fully phased-in rules, which then remain constant for all subsequent cohorts. Our estimates should therefore be read as effects of a partially phased-in reform, plausibly a lower bound on the fully phased-in impact. Several special regimes (railway workers before 1967, coal miners, maritime workers before 1970) were exempt and are excluded from our analysis.

3.2 Data and variables

3.2.1 Data source

We use the MCVL 2004-2024 waves, each wave is a 4% systematic random sample of Spanish Social Security records After restricting the sample to workers aged 60+ who claimed a retirement pension or died between claiming the pension, born between 1943 and 1957, with a non-missing contribution history at age 50, and expected years of contributions at age 65 of at least 25 years, and excluding exempt special regimes, our working sample comprises **179,538 retirees**.

3.2.2 Treatment assignment

For each worker i we compute

$$\widehat{A}_i = \frac{\text{contribution days at age 50}}{30.45} \cdot \frac{1}{12} + 15,$$

the expected years of contributions at age 65 (extrapolating from the contribution pace observed at age 50). We define three exhaustive groups based on the reform's thresholds for birth cohort c:

$$\begin{aligned} \text{Treated} &: \widehat{A}_i \leq \text{LT}_c \\ \text{Partially treated} &: \text{LT}_c < \widehat{A}_i < \text{UT}_c \\ \text{Controls} &: \widehat{A}_i \geq \text{UT}_c \end{aligned} \tag{1}$$

where the upper threshold is

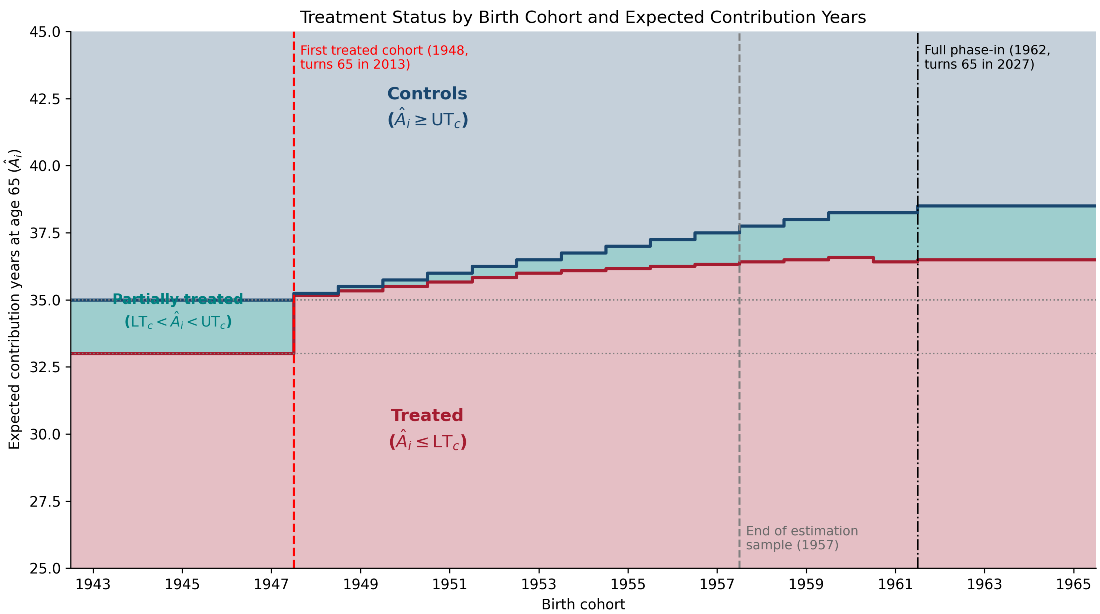
$$\text{UT}_c = 35 + 0.25\,(c - 1947),$$

and the lower threshold follows a piecewise schedule that reflects the contribution-years requirement for involuntary early retirement:

$$\text{LT}_c = \begin{cases} \text{UT}_c - \dfrac{c-1947}{12} & c \leq 1953, \\ \text{UT}_c - 0.5 - \dfrac{c-1953}{6} & c \geq 1954, \end{cases}$$

for $c \in \{1948, \ldots, 1957\}$. Pre-reform cohorts ($c \leq 1947$) use LT $= 33$ and UT $= 35$. Figure 1 illustrates how the three regions defined by these thresholds evolve across cohorts as a function of expected contribution years at 65. Although our estimation sample ends with the 1957 cohort, the figure extends the statutory schedule to the 1958–1965 cohorts: the thresholds keep rising until the 1962 cohort, which turns 65 in 2027 and is the first to face the fully phased-in rules (an exemption threshold of 38.5 contribution years and a maximum NRA of 67), and they remain constant for all later cohorts. In our sample, 26.6% are treated, 5.4% partially treated, and 68.0% controls.

Figure 1: Treatment Status by Birth Cohort and Expected Contribution Years



This assignment rule follows the *simulated eligibility* strategy pioneered by Currie and Gruber (1996) and Cutler and Gruber (1996): group membership is imputed by applying the reform's statutory thresholds to a worker characteristic fixed in the base period—here, the contribution history observed at age 50, well before the reform took effect in 2013. Because $\widehat{A}_i$ is constructed purely from pre-reform information and the cohort-specific thresholds LT$_c$ and UT$_c$ are set by law, exposure is not contaminated by any behavioural response to the reform itself.

Our implementation differs from much of the subsequent literature that estimates a predicted probability of exposure (e.g. via flexible or machine-learning first stages) in two respects:[1] assignment here is *deterministic* rather than probabilistic, and it relies on individual administrative career records at a fixed age rather than on demographic propensity scores. The *dose* measures introduced in Section 4.1 retain the continuous intensity dimension that the discrete grouping discards.

3.2.3 Outcomes and controls

Outcomes: (1) retirement age in months; (2) very early retirement, $\mathbf{1}[\text{age} < 61]$; (3) early retirement, $\mathbf{1}[\text{age} \in [61, \text{NRA} - 1\text{m})]$; (4) normal retirement, $\mathbf{1}[\text{age} \in [\text{NRA} \pm 1\text{m}]]$; (5) delayed retirement, $\mathbf{1}[\text{age} > \text{NRA} + 1\text{m}]$; (6) died before age 67 (*dd*, from the civil register).

Controls (X_i): gender, education (primary/secondary/tertiary), years of contributions at age 50, total contribution days, contribution gaps, worker-type indicator, mutualista status, sector-composition shares, and minimum/maximum contribution group, all measured at age 50, well below the affected years for the cohorts around the time the reform became effective (2013).

3.2.4 Descriptive statistics

Table 2 shows means by treatment group for the full sample and separately for pre-reform (1943–1947) and post-reform (1948–1957) cohorts. Treated workers are disproportionately women (62% vs. 29% among controls), confirming the gender asymmetry inherent in the reform's design. In pre-reform cohorts, retirement probabilities are broadly similar across groups; the differences visible in the full sample are thus largely a post-reform phenomenon. In post-reform cohorts, $\Pr(\text{retire} \leq 65)$ falls from 82% for controls to 43% for treated workers, a gap of 39 pp that directly reflects the NRA increase.

3.2.5 Descriptive evidence on mortality and labour-market histories by cohort

Figure 2 plots the cohort-by-cohort share of workers who died between ages 60–67 and between 65–67, separately for men and women. Two features stand out. First, the male–female

[1] A more recent strand replaces the deterministic legal rule with a statistically *predicted* probability of exposure, estimating the propensity to be affected from pre-reform characteristics with flexible or machine-learning first stages and then sorting individuals into high- and low-propensity groups. See, for instance, Cengiz et al. (2022) on minimum-wage reforms. Our setting does not require this step: the reform's eligibility threshold is a deterministic function of accrued contributions, so exposure can be imputed exactly rather than predicted.

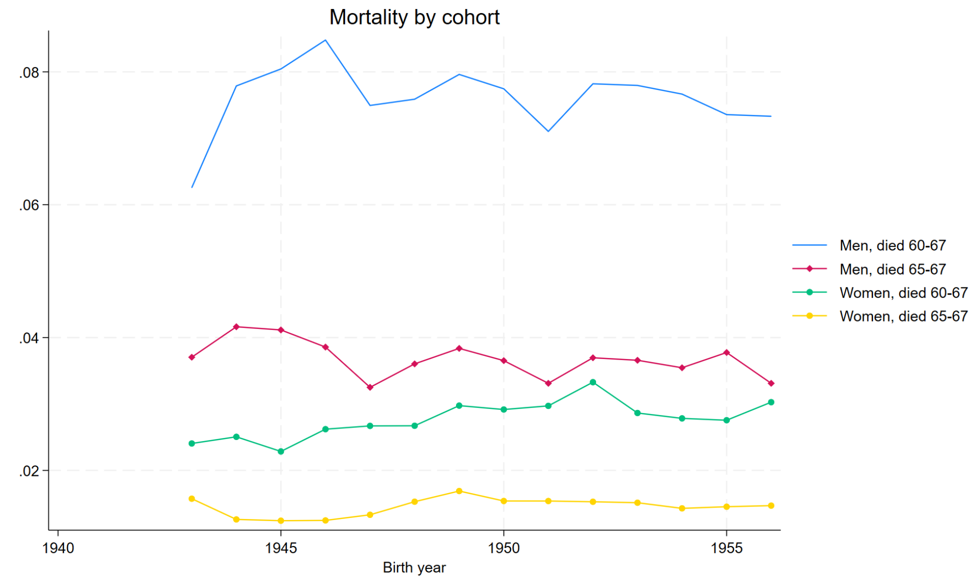
Table 2: Descriptive Statistics by Treatment Group

	Controls	Part. treated	Treated	Total
A. Full sample (1943–1957)				
N	121,733	5,575	52,230	179,538
Women (%)	29.3	25.0	61.5	38.5
Age at obs. (yrs)	72.4	74.0	72.6	72.5
Pr(retire ≤ 61)	0.175	0.135	0.089	0.149
Pr(retire ≤ 63)	0.318	0.222	0.142	0.264
Pr(retire ≤ 65)	0.832	0.790	0.537	0.745
B. Pre-reform cohorts (1943–1947)				
N	35,270	3,509	14,595	53,374
Women (%)	27.3	18.1	56.6	34.7
Pr(retire ≤ 61)	0.197	0.144	0.118	0.172
Pr(retire ≤ 63)	0.295	0.230	0.171	0.257
Pr(retire ≤ 65)	0.855	0.861	0.812	0.843
C. Post-reform cohorts (1948–1957)				
N	86,463	2,066	37,635	126,164
Women (%)	30.1	36.9	63.4	40.1
Pr(retire ≤ 61)	0.166	0.120	0.078	0.139
Pr(retire ≤ 63)	0.327	0.209	0.130	0.266
Pr(retire ≤ 65)	0.823	0.671	0.430	0.703

Source: MCVL 2019. Sample: workers born 1943–1957 who claimed a retirement pension, with expected years of contributions ≥ 25 at 65 and non-missing history at age 50. Group definitions as in equation (1). All between-group differences significant at 1% (Wald test).

gap at ages 60–67 is large in early cohorts (≈ 4–5 pp) and declines steadily across subsequent birth years, suggesting that the underlying secular gender mortality gap at these ages has been narrowing. Second, the 65–67 series lies below the 60–67 series by construction and displays the same convergence pattern.

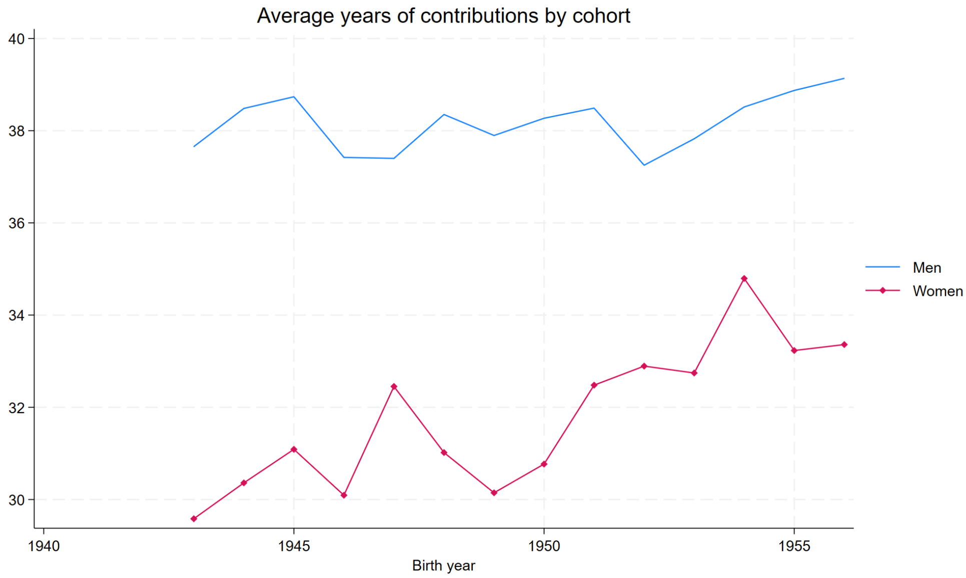
Figure 2: Mortality by birth cohort and gender.



Note: Share dying at ages 60–67 and 65–67, by birth year (1935–1955), separately for men and women. Source: MCVL 2019.

Figure 3 reports the cohort means of years contributed by gender. The male–female gap in contribution histories has narrowed across cohorts, consistent with the rising labour-force attachment of younger female cohorts, though men continue to record longer contribution histories on average.

Figure 3: Average years of contributions by birth cohort and gender.

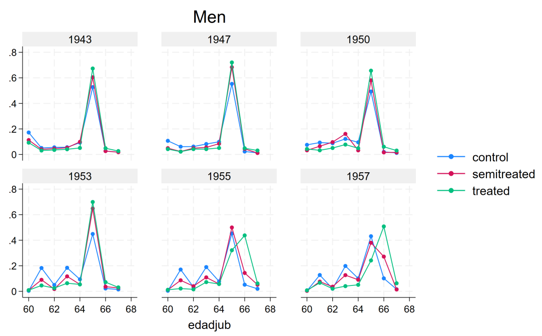


Note: Mean years of contributions at retirement, by birth year (1935–1956), separately for men and women. Source: MCVL 2019.

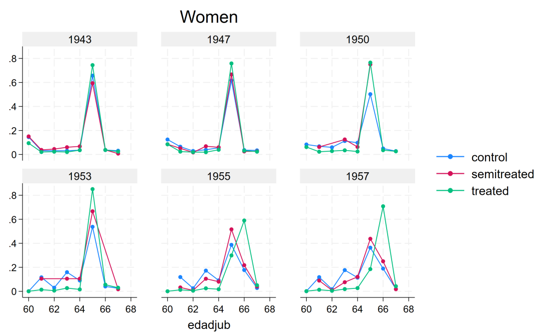
Figure 4 provides context through descriptive retirement-age distributions by birth cohort and treatment group, separately for men and women. The shift from concentration at age 65 (dominant in pre-reform cohorts) toward a flatter post-65 distribution is visible in both panels, but the contrast is sharper for women, consistent with their greater treatment intensity.

Figure 4: Retirement age distributions by cohort and treatment group.

(a) Men.



(b) Women.



Note: Fraction retiring at each age (60–68) by birth cohort and treatment group (control = blue, partially treated = red, treated = green). Selected cohorts shown: 1943 (placebo), 1947 (last pre-reform), 1950, 1953, 1955, 1957. The peak at 65 flattens and the right tail grows for treated cohorts.

4 Identification Strategy

4.1 Reduced-Form Dose-Response Specification

The reform creates *continuous* variation in treatment intensity across workers within the treated group. We capture this variation through two scalar dose measures.

The first is the number of months by which the applicable NRA exceeds 65:

$$\textit{dosis\_meses}_c = \max(0,\ \text{NRA}_c - 65 \times 12)\,.$$

This variable equals zero for all controls (who retain NRA $= 65$) and rises up to 24 months for workers in the 2027 cohort. It captures the *cohort-level mechanical shift* in the retirement age schedule, determined solely by the phase-in table (Table 1) and therefore common to all treated workers in the same birth cohort.

The second is the individual shortfall of expected contribution years below the cohort-specific exemption threshold:

$$DC_{ic} = \max\Big(0,\ \overline{T}_c - \widehat{A}_i\Big)\,,$$

where $\overline{T}_c = 35 + 0.25(c - 1947)$ is the cohort-c threshold and $\widehat{A}_i$ is individual i's expected contribution years at 65 (equation (1)). DC_{ic} equals zero for controls and increases with the years of additional contributions needed to qualify for the lower NRA, varying both across and within cohorts.

The two dose measures capture related but distinct dimensions of exposure. The NRA-shift measure reflects the cohort-aggregate NRA increase; DC captures individual-level depth of exposure, distinguishing workers within the same cohort who differ in how far short of the exemption threshold they fall. Because both measures are zero for controls by construction, entering either one alongside a treated indicator turns the dose term into the *continuous counterpart of a treated* $\times$ *post interaction*: the indicator carries the discrete reform effect and the dose carries its intensity gradient.

Specification. For each dose measure $Z_i \in \{\textit{dosis\_meses}_c, DC_{ic}\}$ we estimate

$$Y_i = \alpha + \tau\, \textit{treated}_i + \beta\, Z_{ic} + \mu_{c(i)} + X_i'\gamma + \varepsilon_i, \tag{2}$$

where *treated*$_i$ is an indicator for belonging to the treated group, $\mu_{c(i)}$ are birth-cohort fixed effects (reference: 1947), X_i is the standard individual control vector described in Section 3.2, and heteroskedasticity-robust standard errors are used. The sample is restricted to treated workers and controls.

Including *treated*$_i$ alongside the dose term decomposes the reform's effect into two pieces: because controls have $Z_i = 0$ by construction, the indicator absorbs the discrete jump at the treated–control boundary, while the dose coefficient β is identified purely from variation in exposure intensity *within* the treated group. In this sense Z_i acts as the *continuous analogue of the treated* $\times$ *post interaction*: τ captures the level shift of being exposed to the reform at all, and β traces how the effect scales with the depth of that exposure. Equation (2) is estimated separately for the pooled sample and by gender, with results reported in Section 5.1.

4.2 Sun–Abraham cohort-interaction estimator

Our primary estimator follows Sun and Abraham (2021). Let $k = c - 1948$ denote event time and $\ell \in \{T, S\}$ index the treatment group. The cohort-specific ATT for group ℓ at event time k is identified from:

$$Y_i = \alpha + \sum_k \beta_k \mathbf{1}[c_i = 1948 + k] + \sum_k \delta_k^\ell \mathbf{1}[\ell_i = \ell] \cdot \mathbf{1}[c_i = 1948 + k] + X_i'\gamma + \varepsilon_i, \qquad (3)$$

where $k = -1$ (birth year 1947) is the reference cohort. δ_k^ℓ recovers the cohort-specific ATT: the average difference in Y between group ℓ and controls within birth cohort k, relative to the pre-reform reference. We estimate equation (3) separately for treated vs. controls and for partially treated (*parttreatedd*) vs. controls using `reghdfe` (Correia, 2016), absorbing cohort fixed effects. Crucially, we do *not* normalise $k = -1$ to zero; the estimated coefficient at $k = -1$ therefore serves as a placebo check.

Inference. With only $G = 15$ birth-cohort clusters, standard asymptotic cluster-robust variance estimators (CRVE) are known to under-reject: the small-G asymptotics that justify sandwich clustering require far more clusters than are available here (Cameron, Gelbach, and Miller, 2008; MacKinnon and Webb, 2017). For the *Sun–Abraham event-study* regressions a block bootstrap is additionally infeasible: because the cohort-specific ATT coefficients are identified one-per-cohort, nearly every bootstrap replicate omits at least one cohort, rendering the replicate non-estimable. We therefore use heteroskedasticity-robust standard errors throughout, which provide valid inference under general individual-level heteroskedasticity without requiring a large number of clusters. For the *dose-response* regressions (Section 4.1) we likewise report heteroskedasticity-robust standard errors with cohort fixed effects absorbed. Coarsened exact matching (CEM) on pre-reform worker characteristics is a natural avenue to further tighten the comparison groups and reduce residual confounding; we leave this extension for future work.

Gender-differential ATT. To assess whether effects differ by gender we augment equation (3) with the triple interaction female$_i$ $\times$ $\mathbf{1}[\text{treated}_i] \times \mathbf{1}[c_i = 1948 + k]$, absorbing cohort, cohort$\times$gender, and cohort$\times$treated fixed effects. The coefficient on this term estimates $\delta_k^{T,W} - \delta_k^{T,M}$, the gender-differential ATT at event time k.

4.3 Parallel trends

The identifying assumption is that treated workers would have followed the same outcome trajectory as controls absent the reform. We assess this by inspecting pre-reform event-study coefficients ($k = -5, \ldots, -2$), which should be small and jointly insignificant if the assumption holds. For the primary retirement outcome (retirement age in months), the joint pre-trend test accepts the null of no differential pre-trend ($F(4, \cdot) = 1.59$, $p = 0.173$), as does the retire-at-65 ($p = 0.459$) and retire-after-65 ($p = 0.454$) outcomes. The retire-before-65 pre-trend test marginally rejects ($p = 0.046$), and the pre-trend tests for very early retirement ($p < 0.001$) and for the work-deviation outcome reject more strongly, reflecting pre-existing compositional differences—treated workers are disproportionately women with shorter careers who were already more likely to exit early before the reform. We address this by controlling flexibly for labour-market history and focusing on the change in trajectories at $k = 0$. For premature mortality (60-67), the same compositional issue applies. The key identifying assumption is therefore that these compositional differences are stable across cohorts, so that the break at $k = 0$ is attributable to the reform. Appendix A1.1 reports formal pre-trend test statistics.

5 Main Results

5.1 Dose-Response Results

Table 3 reports the dose-response estimates from equation (2) for all five retirement-type outcomes plus premature mortality. Both dose measures deliver a coherent picture across all outcomes; results are qualitatively and quantitatively similar with and without cohort effects (not shown but available on request).

Table 3: Reduced-Form Dose-Response Estimates: Five Outcomes and Mortality

	Panel A: dose = NRA shift (months)			*Panel B*: dose = contribution shortfall (years)		
Outcome	All	Men	Women	All	Men	Women
1. Retirement age (months)						
$\hat{\tau}$ (*treated*)	5.330^{***}	5.063^{***}	7.779^{***}	6.804^{***}	4.869^{***}	10.760^{***}
(t)	(8.22)	(7.36)	(7.01)	(9.30)	(8.46)	(11.13)
$\hat{\beta}$ (*dose*)	1.004^{***}	0.524^{***}	1.054^{***}	1.100^{***}	0.884^{***}	0.879^{***}
(t)	(7.90)	(4.28)	(8.18)	(7.72)	(9.57)	(6.40)
2. Retire before age 65 ($[61, 65)$)						
$\hat{\tau}$ (*treated*)	-0.1231^{***}	-0.0859^{***}	-0.1679^{***}	-0.1351^{***}	-0.0897^{***}	-0.2035^{***}
(t)	(-6.05)	(-6.21)	(-5.65)	(-7.63)	(-7.61)	(-8.01)
$\hat{\beta}$ (*dose*)	-0.0191^{***}	-0.0142^{***}	-0.0213^{***}	-0.0250^{***}	-0.0211^{***}	-0.0229^{***}
(t)	(-5.37)	(-4.79)	(-4.97)	(-10.28)	(-9.48)	(-7.10)
3. Retire at age 65 ($[65, 65]$)						
$\hat{\tau}$ (*treated*)	-0.0241	-0.0302^{**}	-0.0204	-0.0517^{***}	-0.0599^{***}	-0.0446^{***}
(t)	(-1.75)	(-2.98)	(-0.88)	(-4.62)	(-4.62)	(-3.44)
$\hat{\beta}$ (*dose*)	-0.0103^{***}	-0.0091^{***}	-0.0111^{***}	-0.0081^{***}	-0.0049	-0.0108^{***}
(t)	(-8.00)	(-8.21)	(-5.27)	(-3.51)	(-1.74)	(-5.40)
4. Retire after age 65 (>65)						
$\hat{\tau}$ (*treated*)	0.1552^{***}	0.1318^{***}	0.2132^{***}	0.1922^{***}	0.1563^{***}	0.2695^{***}
(t)	(9.19)	(10.39)	(8.33)	(7.67)	(7.87)	(8.62)
$\hat{\beta}$ (*dose*)	0.0271^{***}	0.0178^{***}	0.0292^{***}	0.0304^{***}	0.0203^{***}	0.0298^{***}
(t)	(9.24)	(9.40)	(7.35)	(7.95)	(5.93)	(7.64)
5. Died before age 67 ($\times 10^5$)						
$\hat{\tau}$ (*treated*)	-34.6	-41.6	-13.2	-19.0	-12.5	-21.0
(t)	(-1.29)	(-1.01)	(-0.81)	(-0.71)	(-0.28)	(-1.61)
$\hat{\beta}$ (*dose*)	3.31	5.80	-0.21	0.9	-0.1	1.4
(t)	(1.34)	(1.44)	(-0.09)	(0.32)	(-0.01)	(0.68)
N	109,832	75,180	34,652	109,832	75,180	34,652

OLS estimates of equation (2), which includes both a *treated* indicator ($\hat{\tau}$: discrete treated–control jump) and a continuous dose term ($\hat{\beta}$: gradient *within* the treated group). With controls fixed at zero dose, Z_i is the continuous analogue of treated × post. *Panel A*: dose = NRA shift (months by which the applicable NRA exceeds 65; coefficients in per-1-month units). *Panel B*: dose = DC_{ic}, years of contributions below the cohort-specific exemption threshold (zero for controls). All regressions include birth-cohort fixed effects (reference: 1947) and individual controls X_i (Section 3.2). Sample: treated workers and controls, birth years 1943–1957, expected contributions ≥ 25 years. **Inference**: heteroskedasticity-robust standard errors; t-statistics reported. Row 5 coefficients $\times 10^5$. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Retirement age. Being treated at all raises realised retirement age by roughly **five months** (the $\hat{\tau}$ jump), and each additional month of NRA delay adds about **one further month** on top ($\hat{\beta} \approx 1.0$, $t = 31.87$), with a steeper gradient for women than for men. The contribution-shortfall measure tells the same story, with a within-group gradient strongly significant for both genders.

Retirement type. The reform unambiguously shifted retirement from *at 65* and *before 65* toward *after 65*, and this holds on both margins of the decomposition. The treated jump alone raises the probability of retiring after 65 by roughly **16 pp** and lowers retirement before 65 by about 12 pp; the continuous dose adds a further gradient in the same direction (see Table 3). The early-retirement margin is partly mechanical, since the minimum early-retirement ages (formerly 61 and 63) were also raised by the reform (Table 1).

Mortality. Neither the treated jump nor the continuous dose is statistically distinguishable from zero in the pooled sample or by gender ($|t| \leq 1.6$ throughout), and all coefficients are economically negligible (on the order of 10^{-5}). We therefore read the mortality evidence as a precisely estimated null, consistent with the event-study results presented below.

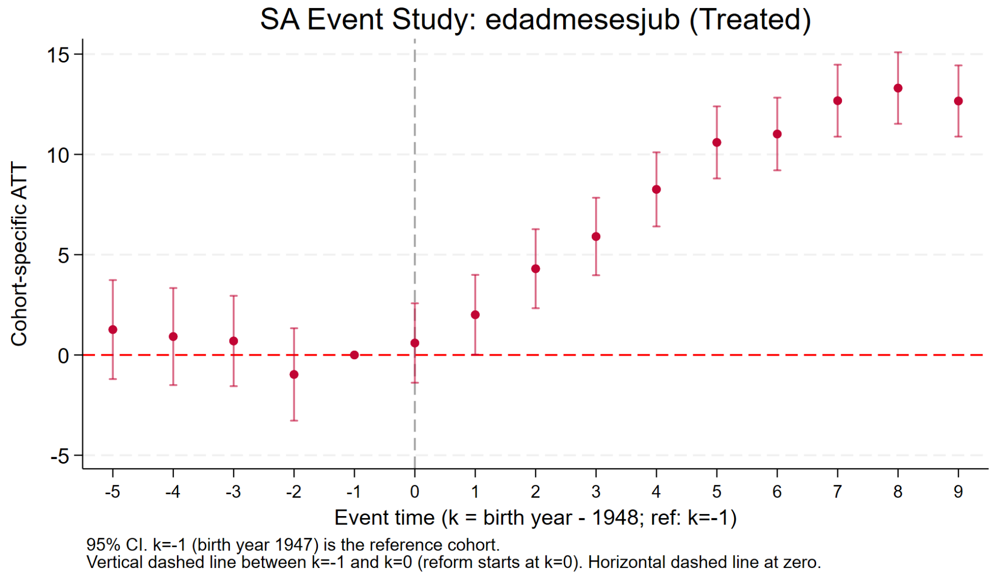
5.2 Event-Study Results

5.2.1 Retirement Age

Figure 5 shows the cohort-specific ATT on retirement age (months) for *treated* workers relative to controls. The pre-reform coefficients ($k = -5, \ldots, -2$) are tightly clustered near zero, providing strong support for the parallel-trends assumption on the causal margin of interest. From $k = 0$ onward the ATT grows monotonically, reaching its maximum for the latest cohorts (1955–1957), with well-identified confidence intervals that narrow across cohorts as sample sizes stabilise.

The average post-reform ATT across cohorts $k = 0, \ldots, 9$ is **+8.1 months** for the treated group (Table 4). The absence of a pre-reform trend and the clean post-reform ramp-up are consistent with a causal delay in retirement driven by the reform's NRA schedule rather than pre-existing differences between treated workers and controls. Partially treated workers show no significant average delay in retirement age, consistent with the absence of significant effects across all outcomes for this group.

Figure 5: Event-Study: Retirement Age (months). Treated vs Controls.



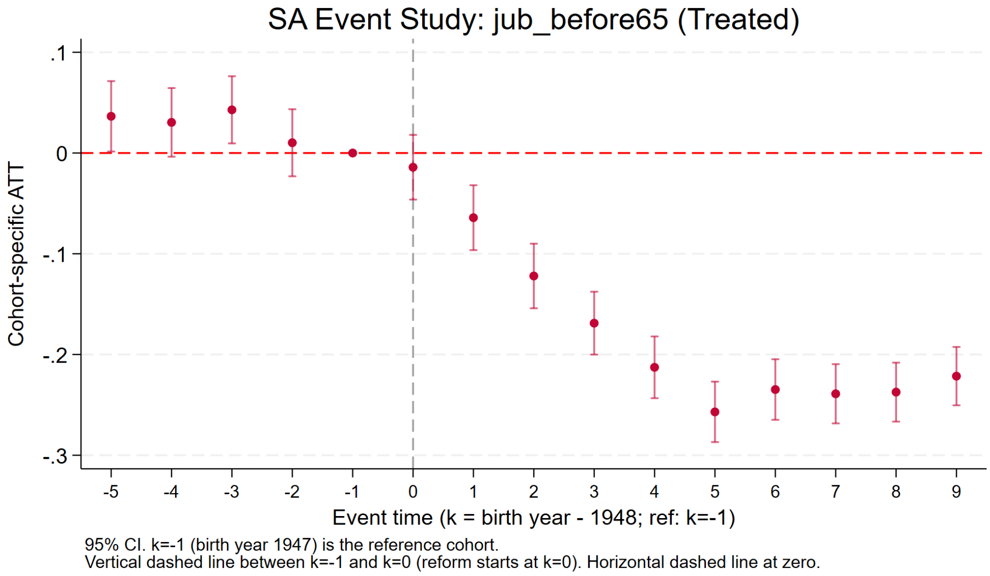
Note: Sun–Abraham cohort-specific ATT, treated workers vs controls. k = birth year $- 1948$; $k = -1$ (birth year 1947) is the last pre-reform cohort (white diamond; ATT estimated, not imposed to zero). Vertical dashed line between $k = -1$ and $k = 0$. 95% CI, heteroskedasticity-robust standard errors.

5.2.2 Retirement Before, At, and After Age 65

What are the dominant adjustment margins? Figures 6, 7, and 8 show the ATTs on the three absolute-age retirement outcomes for the treated group. The pattern is consistent and striking: retirement *at exactly 65* collapses, retirement *before 65* also falls, and retirement *after 65* rises sharply—all relative to control workers.

Retire before 65. The ATT on $\mathbf{1}[\text{ret. age} \in [61, 65)\ \text{yrs}]$ is small and slightly negative at $k = 0$, turns increasingly negative across cohorts, and stabilises in the region of -0.22 to -0.26 for the later cohorts, yielding an average ATT of -17.7 **pp** (Table 4), with pre-reform coefficients flat and near zero. This decline reflects the tightening of minimum early-retirement ages (from 61/63 to 63/65; see Table 1), which removed access to the $[61, 65)$ window for affected workers. Partially treated workers show a qualitatively similar but attenuated and mostly insignificant pattern, consistent with their partial and heterogeneous exposure to the early retirement age changes.

Figure 6: Event-Study: Probability of Retiring Before Age 65 ($[61, 65)$). Treated vs Controls.

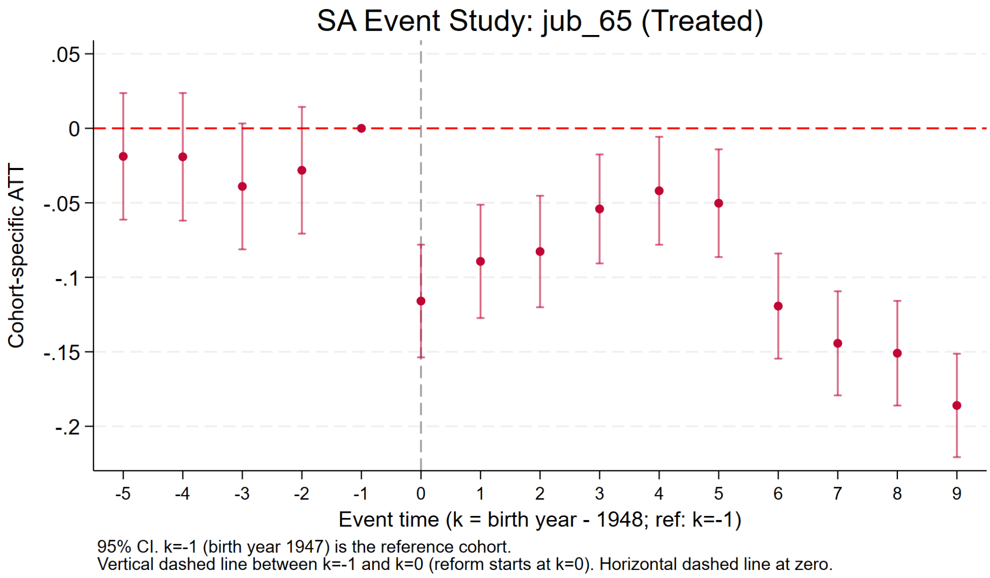


Note: Outcome: $\mathbf{1}[\text{ret. age} \in [61, 65) \text{ yrs}]$. See notes to Figure 5.

Retire at 65. Figure 7 shows an immediate, large, and deepening decline in the probability of retiring at exactly 65 for treated workers, with an average ATT of -10.4 **pp**. Age 65 was the dominant focal retirement age under the pre-reform rules; the reform effectively dismantles this focal point for workers who cannot qualify for the long-career exemption, pushing them toward retirement after 65 instead. For partially treated workers the decline is smaller but also visible.[2]

[2] The corresponding figure is available upon request.

Figure 7: Event-Study: Probability of Retiring at Exactly Age 65. Treated vs Controls.

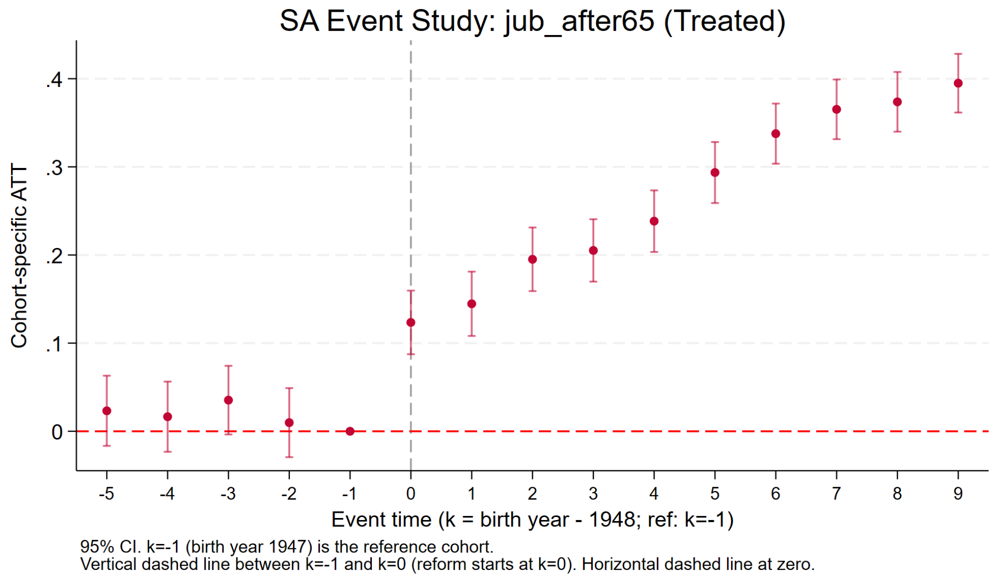


Note: Outcome: $\mathbf{1}[\text{ret. age} = 65\,\text{yrs}]$. See notes to Figure 5.

Retire after 65. Figure 8 is the clearest result in the paper. The ATT on $\mathbf{1}[\text{ret. age} > 65\,\text{yrs}]$ rises monotonically across cohorts, with pre-reform coefficients small and confirming that treated workers were not systematically more likely than controls to retire after 65 before the reform. The average post-reform ATT is **+26.7 pp**, signalling that the reform succeeded in its primary objective. The monotonic increase across cohorts mirrors the reform's phase-in: later cohorts face a larger NRA gap and correspondingly larger after-65 retirement shares. Partially treated workers also shift toward post-65 retirement, but the effects are small and largely imprecisely estimated, consistent with their heterogeneous and partial dose.[3]

[3]The corresponding figure is available from the authors upon request.

Figure 8: Event-Study: Probability of Retiring After Age 65 (>65). Treated vs Controls.



Note: Outcome: $\mathbf{1}[\text{ret. age} > 65\,\text{yrs}]$. See notes to Figure 5.

5.2.3 Summary

Table 4 collects the average post-reform ATTs for all five retirement-type outcomes plus mortality across cohorts $k = 0, \ldots, 9$, reporting both treated and partially treated ATTs. For treated workers, the data support a clear and statistically significant pattern across all retirement-timing margins. For partially treated workers (*parttreatedd*), by contrast, virtually no coefficient is statistically significant: the average post-reform ATTs are small and imprecisely estimated across all outcomes, providing no evidence of a meaningful behavioural response for this group. Very early retirement (< 61 years) is noteworthy: treated men show a small but significant positive ATT (+4.5 pp), while the female and pooled effects are imprecisely estimated, suggesting some migration toward disability pensions or informal inactivity before age 61 among male treated workers.

6 Heterogeneity by Gender

6.1 Why women bear more of the reform

The contribution-contingent structure of the reform creates a systematic gender asymmetry rooted in career history. Women in Spain have historically worked fewer years in formal employment due to career interruptions for childcare, higher rates of part-time work, and earlier

Table 4: Average Post-Reform ATT, Cohorts $k = 0, \ldots, 9$ (birth years 1948–1957)

Outcome	Treated (main result)	Part. treated (dose gradient)
Retirement age (months)	$+8.13^{***}$	≈ 0
Very early retirement ($<$ 61 yrs)	$+0.013$	≈ 0
Retire before 65 ($[61, 65)$ yrs)	-0.177^{***}	≈ 0
Retire at 65	-0.103^{***}	≈ 0
Retire after 65 (> 65)	$+0.267^{***}$	≈ 0
Died before age 67	$+0.0029$	≈ 0

Simple mean of cohort-specific ATTs across $k = 0, \ldots, 9$ computed via linear combination of the estimated event-study coefficients. **Treated** workers are those whose predicted contribution history at age 65 falls below the lower threshold $\underline{T}_c$; they face the full NRA increase. **Partially treated** (*parttreatedd*) workers lie between $\underline{T}_c$ and the exemption threshold $\bar{T}_c$; they face partial incentive changes. For this group, virtually no outcome shows a statistically significant average ATT. Both groups are compared to *controls* (predicted contributions $\geq \bar{T}_c$), who are exempt from the NRA increase. Heteroskedasticity-robust standard errors. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

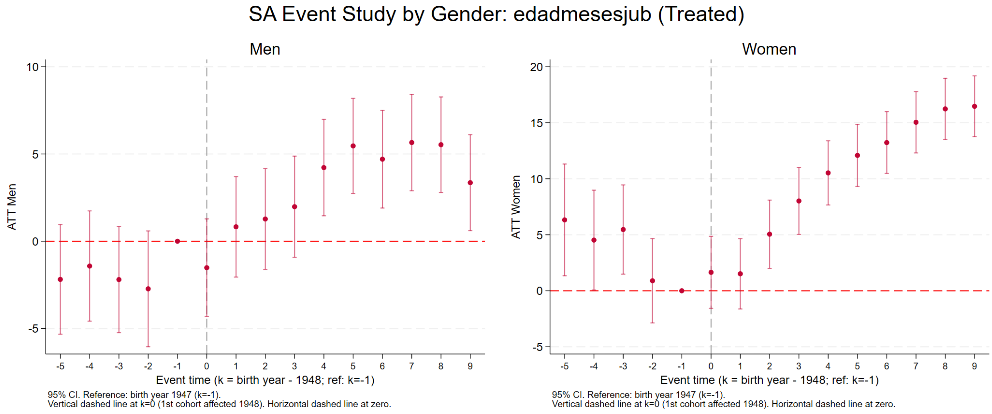
cohort-specific labour market exclusion. As a result, women are much less likely to reach the long-career threshold (≥ 38.5 years by 2027) than men. In our estimation sample, **62% of treated workers are women**, compared with roughly 29% of controls. A reform nominally framed as raising the retirement age for all workers is therefore in practice a reform that disproportionately raises the retirement age for women. The asymmetry is self-reinforcing: women who fail to qualify for the long-career exemption face not only a later retirement age but also lower pension benefits (because shorter contribution histories generate lower replacement rates), thus compounding rather than narrowing the existing gender pension gap (OECD, 2023; Della Giusta and Longhi, 2021). Descriptive retirement-age distributions by cohort and treatment group for men and women are reported in Figure 4 of Section 3.2.5.

6.2 Event-study results by gender

6.2.1 Retirement age

Figure 9 overlays event-study estimates for men (filled circles) and women (open squares), estimated separately on the treated-vs-controls sample. Both groups delay retirement significantly after $k = 0$, but women's delay is systematically larger across cohorts: average post-reform ATTs of **3.1 months for men** and **10.0 months for women**. This gap reflects that women in the treated group are disproportionately concentrated below the long-career threshold and therefore face the full statutory age increase.

Figure 9: Event-Study by Gender: Retirement Age. Treated vs Controls.

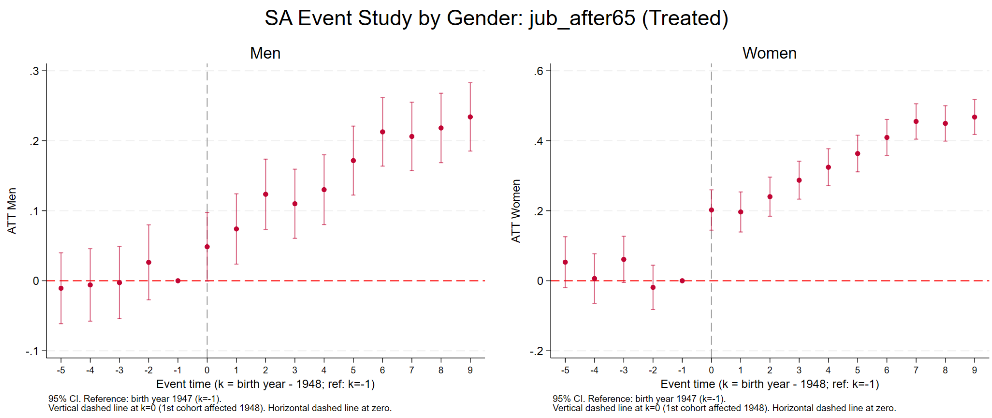


Note: Estimated separately by gender. Filled circles: men (navy); open squares: women (cranberry). White diamond at $k = -1$: common reference (ATT = 0 for both series). See notes to Figure 5.

6.2.2 Retirement after age 65

The gender contrast in after-65 retirement (Figure 10) is the most striking finding. Both men and women shift strongly toward post-65 retirement after the reform, but women's ATT is substantially *larger* than men's across all post-reform cohorts: average ATTs of $+34.0$ **pp for women** vs $+15.3$ **pp for men**. This confirms the contribution-contingent gender asymmetry: because more women lack the contribution history to qualify for the long-career exemption, a larger share of women face the full statutory age increase and consequently shift into post-65 retirement.

Figure 10: Event-Study by Gender: Probability of Retiring After Age 65.

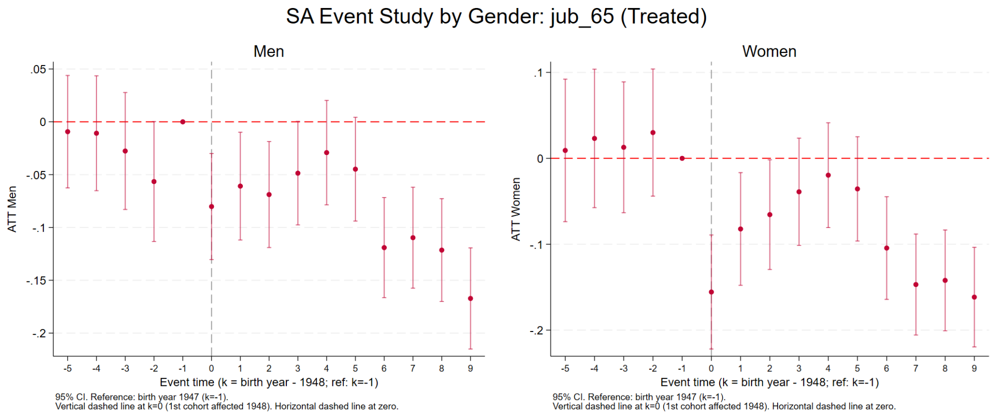


Note: See notes to Figure 9.

6.2.3 Retire at 65 and before 65

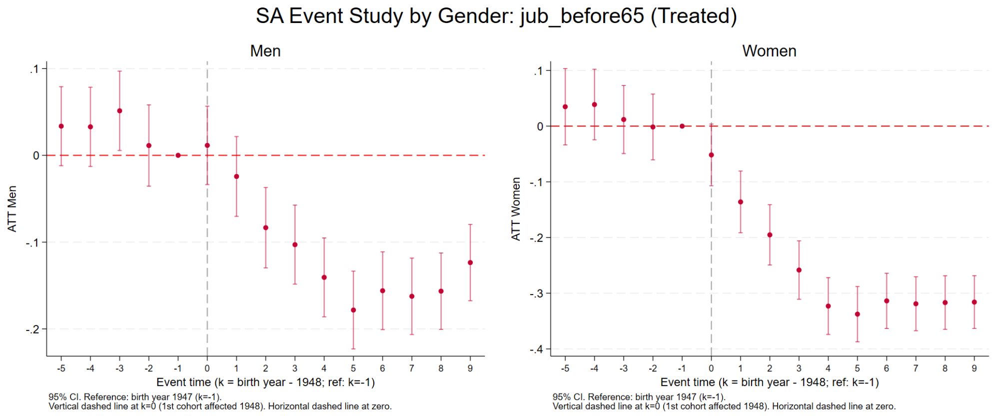
Figure 11 shows the ATT on retiring at exactly 65. Both men and women experience a sharp decline, with the female decline somewhat deeper (-9.5 pp on average vs. -8.5 pp for men). Figure 12 shows that retirement before 65 falls significantly for both groups, but much more steeply for women (-25.7 pp) than for men (-11.2 pp), reflecting women's greater concentration below the long-career threshold and their stronger pre-reform reliance on the early-retirement window.

Figure 11: Event-Study by Gender: Retire at Exactly Age 65.



Note: See notes to Figure 9.

Figure 12: Event-Study by Gender: Retire Before Age 65.



Note: See notes to Figure 9.

6.2.4 Summary

Table 5 summarises the average post-reform ATTs by gender. For every retirement-timing margin women's response is larger in magnitude than men's, with the gender contrast most pronounced for retirement before 65 and after 65. The gender differential in 60–67 mortality (Women $-$ Men) favours women in the DiD estimates, a finding we explore in depth in the following section.

Table 5: Average Post-Reform ATT by Gender, Treated Cohorts ($k = 0, \ldots, 9$)

Outcome	Men	Women	Women $-$ Men
Retirement age (months)	$+3.15^{***}$	$+9.98^{***}$	$+6.83$
Very early retirement (<61)	$+0.044^{***}$	$+0.012$	-0.032
Retire before 65 ($[61, 65)$)	-0.112^{***}	-0.257^{***}	-0.145
Retire at 65	-0.085^{***}	-0.095^{***}	-0.010
Retire after 65 (>65)	$+0.153^{***}$	$+0.340^{***}$	$+0.187$
Died before age 67	$+0.01468$	-0.00018	-0.01486

Columns (1) and (2): for retirement outcomes, average post-reform ATT ($k = 0, \ldots, 9$) from gender-stratified Sun–Abraham regressions (treated vs. controls); for 60–67 mortality, DiD estimates from gender-stratified 2×2 regressions (see Table 6). Column (3): implied Women $-$ Men differential. Heteroskedasticity-robust standard errors for retirement outcomes; robust SE for mortality DiD. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

7 Reform Effects on Premature Mortality (60-67)

In this section we explore the premature mortality impact of the normal retirement age reform. We first present reduced-form evidence and then take a more structural view, analysing the role of months worked as the main channel through which the reform affects mortality. Along the way we also assess a natural alternative driver—reform-induced losses in pension wealth, summarised by social security wealth—and show that it cannot account for the differential mortality response.

7.1 Overall effect and Gender differential

To assess gender differences in mortality outcomes more transparently than the event-study format allows, we estimate a standard 2×2 difference-in-differences (DiD) specification pooled and separately by gender. The specification regresses the 60–67 mortality indicator on a treatment dummy, a post-reform dummy, their interaction (the DiD coefficient), and the baseline covariates X_i; the pooled column additionally includes a gender dummy. Heteroskedasticity-robust standard errors are used throughout, consistent with the choice made for the Sun–Abraham regressions given the small number of available birth-cohort clusters. The simple 2×2 specification gives equal weight to all post-reform cohorts and is therefore complementary to the Sun–Abraham event study reported in Section 7.[4]

Table 6: DiD Estimates of the Reform's Effect on 60–67 Mortality, by Gender

	Without covariates			*With covariates X_i*		
	(1) Pooled	(2) Men	(3) Women	(4) Pooled	(5) Men	(6) Women
Treated (pre-reform)	$+0.0077^{**}$	$+0.0131^{***}$	-0.0064	$+0.0047$	$+0.0081$	-0.0075
	(2.38)	(2.90)	(-1.44)	(1.46)	(1.84)	(-1.70)
Post	-0.0026	-0.0032	-0.0049	$+0.0079^{***}$	$+0.0090^{***}$	$+0.0001$
	(-1.41)	(-1.52)	(-1.38)	(4.15)	(4.05)	(0.03)
Treated $\times$ Post (ATT)	$+0.0062^{*}$	$+0.0110^{**}$	$+0.0088^{*}$	$+0.0062^{*}$	$+0.0159^{***}$	$+0.0079^{*}$
	(1.70)	(2.02)	(1.81)	(1.79)	(3.07)	(1.70)
Constant	$+0.0747^{***}$	$+0.0736^{***}$	$+0.0336^{***}$	$+0.0858^{***}$	$+0.0900^{***}$	$+0.0347^{***}$
	(45.68)	(41.34)	(10.22)	(27.56)	(20.82)	(7.63)
Implied M$-$W mortality gap (from cols. 2–3 and 5–6; see text for derivation)						
Baseline M$-$W gap (control, pre)		$+4.01$ pp			$+5.53$ pp	
Control time trend on M$-$W gap ($\hat{\beta}_{\text{post},m} - \hat{\beta}_{\text{post},w}$)		$+0.17$ pp			$+0.89$ pp	
Reform DiD on M$-$W gap ($\hat{\tau}_m - \hat{\tau}_w$)		$+0.23$ pp			$+0.81$ pp	
Sample-weighted M$-$W mortality gap ($\Delta\bar{Y}_{post} - \Delta\bar{Y}_{pre}$, using cell N as weights)						
All estimated coefficients		-0.10 pp			$+0.77$ pp	
Insignificant nuisance terms set to zero		-0.25 pp			$+0.83$ pp	
Observations	128,861	91,636	37,225	128,532	91,339	37,193

OLS difference-in-differences on the MCVL estimation sample (treated vs. controls). Outcome: indicator for death between ages 60 and 67. Columns (4)–(6) add the baseline covariates X_i (maximum contribution days, minimum contribution gaps, mutualista status, worker-type indicator, construction and mining sector shares, minimum and maximum contribution group); columns (1) and (4) additionally include a gender dummy. Heteroskedasticity-robust standard errors; t-statistics in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. Note: the pooled and female ATT coefficients are marginally significant ($p \approx 0.10$) under robust inference; the male ATT is significant at 1%.

The *Implied M$-$W mortality gap* block decomposes the male-minus-female gap analytically from the regression coefficients alone: the baseline gap equals $\hat{\alpha}_m - \hat{\alpha}_w$; the secular trend equals the difference in $\hat{\beta}_{\text{post}}$ coefficients across genders; and the reform-induced component equals $\hat{\tau}_m - \hat{\tau}_w$. Note that $\hat{\beta}_{\text{post}}$ is statistically insignificant in most columns, so the "secular trend" row is imprecisely estimated and partly reflects cohort composition rather than a genuine time trend.

The *Sample-weighted M$-$W mortality gap* rows translate the regression coefficients into population-weighted averages using the observed cell sizes (gender $\times$ treatment $\times$ pre/post) as weights. The sample counts used are: men, control pre = $21{,}495$; men, control post = $54{,}975$; men, treated pre = $4{,}611$; men, treated post = $10{,}566$; women, control pre = $3{,}008$; women, control post = $16{,}679$; women, treated pre = $3{,}012$; women, treated post = $14{,}527$. The first row uses all estimated coefficients; the second retains both gender-specific ATTs ($\hat{\tau}_m, \hat{\tau}_w$, the parameters of interest) but sets the statistically insignificant nuisance terms ($\hat{\beta}_{\text{treated}}$, $\hat{\beta}_{\text{post}}$, at the 5% level) to zero. The two rows therefore bracket the reform's contribution to the observed population mortality gap, illustrating sensitivity to the treatment of imprecisely estimated nuisance parameters.

We read Table 6 along three lines: the causal effect on treated post-reform workers, the implied shift in the male–female mortality differential, and the resulting change in the *observed* population gap once sample composition is taken into account.

(a) Causal effect on the treated. The central result is the Treated $\times$ Post coefficient, the reform's average effect on the treated, positive in every specification. With covariates it is $+1.59$ pp for men, significant at the 1% level ($t = 3.07$); the pooled and female effects are also positive ($+0.62$ pp and $+0.79$ pp) but only marginally significant under robust inference

[4] In the same event-study design, partially treated workers show no systematic mortality differences relative to controls after the reform; the corresponding estimates are available upon request.

($p \approx 0.10$). The roughly twofold larger response for men holds whether or not covariates are included, and is the central mortality finding: the reform raised premature mortality, and disproportionately so among treated men.

(b) Effect on the male–female differential. The *Implied M$-$W mortality gap* rows translate these coefficients into the reform's contribution to the gender gap. The reform-induced component, $\hat{\tau}_m - \hat{\tau}_w$, widens the male excess by about four-fifths of a point with covariates ($+0.81$ pp, roughly 15% of the pre-existing 5.53 pp baseline gap) and by less without. Because treated workers are disproportionately women with shorter careers, the covariates absorb compositional differences between treated and control groups, which is why the with-covariates figure is the cleaner causal summary. The secular-trend row is imprecisely estimated—$\hat{\beta}_{\text{post}}$ is insignificant in all but the pooled-with-covariates column—so it likely reflects cohort composition rather than a genuine narrowing of the gap.

(c) Effect on the observed population gap. The *Sample-weighted M$-$W mortality gap* rows ask a distinct question: given the sample's actual cell composition (gender $\times$ treatment $\times$ pre/post), how much does the reform move the *observed* mortality difference? With covariates the weighted gap widens ($+0.77$ pp), closely tracking the analytical estimate, but is essentially flat to slightly negative without covariates (-0.10 pp), because the large male control group—outnumbering treated men roughly $5:1$—dominates the post-reform male average and compresses the gap. Zeroing the insignificant nuisance terms while retaining the gender-specific ATTs leaves the with-covariates gap positive ($+0.83$ pp) and the without-covariates gap negative (-0.25 pp); the exact figure is sensitive to the significance threshold, so we do not read too much into it. The robust message is the contrast between the analytical differential, the reform's clean *causal* contribution to the male excess, and the *observed* population gap, which is muted and can even change sign under the weight of the male control group.

Both readings align with the descriptive evidence of Section 3.2.5: the M$-$W mortality gap narrows across successive cohorts, and the sample-weighted calculations confirm that this compositional narrowing can offset, and even reverse, the reform-induced widening of the male excess. As later cohorts mature and the reform reaches full bite, these compositional effects may attenuate further.[5]

Three mechanisms are consistent with this asymmetry: **occupational exposure**, as treated men are disproportionately employed in construction, mining, and other physically demanding sectors where delayed exit plausibly translates into excess mortality (Bellés-Obrero et al.,

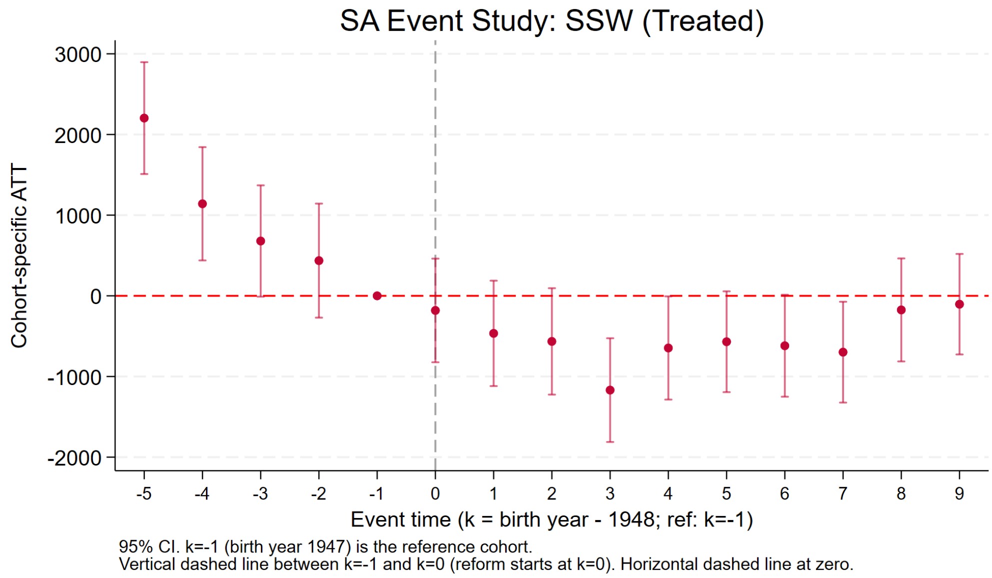
[5] Using the sample distribution of treated post-reform workers, the with-covariates analytical differential implies a non-trivial number of excess 60–67 deaths in the estimation sample, with men accounting for the larger share of the implied burden despite being a minority of treated post-reform workers.

2022; Bloemen et al., 2017; Carrino, Glaser, and Avendano, 2020); **selection on disability exit**, since women still attached to the labour market in their early sixties are likely positively selected on health, having survived a higher pre-reform exit rate via disability pensions, which attenuates any reform-induced effect for the female sample; and **baseline mortality levels**, since male pre-reform mortality is substantially higher than the female level, so absolute increases of comparable proportional magnitude translate into mechanically larger DiD coefficients for men.

A further candidate factor, distinct from the three above, is pension wealth itself. As described in Section 3, Law 27/2011 combined the rise in the NRA with benefit-formula changes—a longer computation period for the regulatory base and a stricter accrual scale—that reduce pension generosity for treated and control workers alike. If, on top of this common component, treated workers had suffered systematically larger losses in pension wealth, an income-based channel (through forgone consumption, health investment, or financial stress) could contribute to the mortality response documented above. To assess this possibility we compute each worker's social security wealth (SSW) and estimate the same Sun–Abraham event-study design used throughout.[6] Figure 13 reports the cohort-specific ATTs. The central message is that, from the 1947 reference cohort onwards, the reform does not appear to have substantially affected the social security wealth of treated workers relative to controls. The pre-reform estimates decline smoothly toward the 1947 reference cohort, consistent with the secular convergence of treated and control pension levels across successive cohorts rather than with any reform anticipation. From the 1948 cohort onwards the estimates are economically modest relative to the level of social security wealth, several confidence intervals include zero, and the profile shows none of the monotonic deepening across cohorts that the NRA schedule generates for retirement age and months worked: it bottoms out in the middle of the phase-in and drifts back toward zero for the youngest cohorts. Social security wealth is therefore not a plausible explanatory factor for the mortality differences documented above: treated workers did not lose systematically more pension wealth than controls, and the timing of the small estimated losses does not track the mortality response. Having ruled out the wealth channel, we turn to the leading candidate: the compelled extension of working life itself.

[6] Social security wealth is defined as the expected present discounted value, at the moment of claiming, of the stream of real pension benefits: each annual benefit is weighted by the cohort- and sex-specific survival probability from the INE life tables and discounted at a 3% real rate. It thus summarises in a single euro figure the combined effect of the benefit level, the claiming age, and remaining life expectancy.

Figure 13: Event-Study: Social Security Wealth. Treated vs Controls.



Note: Outcome: social security wealth (in euros), defined as the expected present discounted value of the real pension stream at claiming, using INE cohort- and sex-specific survival probabilities and a 3% real discount rate. See notes to Figure 5.

7.2 The work-extension channel

A natural candidate mechanism linking the reform to the mortality increase is the *forced extension of working life itself*. By raising the NRA and tightening early-exit conditions, the reform compelled treated workers to remain in employment for additional months beyond what they would have chosen absent the reform. If this extra labour-market exposure occurs in physically demanding jobs, it can translate directly into excess mortality.

Figure 14 provides event-study evidence on the reform-induced work extension. The outcome is the deviation of actual from counterfactual (pre-reform-schedule) months of employment. Pre-reform cohorts display estimates close to zero, consistent with parallel trends. From $k = 0$ onward, the deviation grows sharply and monotonically, reaching approximately $+20$ months by $k = 9$, as successive cohorts face progressively higher NRAs under the phased-in reform schedule. This accumulating work extension is the first-stage underpinning of the structural analysis below.

Figure 14: Event-Study: Deviation of Actual from Counterfactual Months Worked. Treated vs Controls.

Note: Outcome: deviation of actual from counterfactual (pre-reform) months worked. See notes to Figure 5.

To quantify how much of the mortality response flows through the months-worked channel, we estimate a structural model in which mortality between ages 60 and 67 is a function of total months worked and of reform-induced months—defined as total months worked interacted with the treated-post indicator—instrumented by the reform exposure variables (the cohort-level NRA shift in months, the individual contribution shortfall below the exemption threshold, and expected years of contributions at age 65). The interaction term captures the additional mortality cost specific to the compelled work extension: if reform-forced employment in demanding jobs raises mortality beyond what voluntary employment of the same duration would, its coefficient will be positive even when the coefficient on total months worked is negative.

Table 7 reports the two first-stage regressions (columns 1–2), the OLS structural equation (column 3), and the preferred 2SLS estimate (column 4). The reduced-form estimate already appears in Table 6 and is therefore omitted here.

Table 7: Structural Analysis: Work Extension and 60-67 Mortality

	First Stage		*Structural equation*	
	(1) Months worked	(2) Reform-induced months worked	(3) OLS	(4) IV
Endogenous regressors				
Months worked			-0.000474^{***}	-0.0000641^{***}
			(-35.28)	(-3.72)
Reform-induced months worked			$+0.0000134$	$+0.000133^{***}$
			(0.53)	(3.23)
Reduced-form regressors				
Treated (pre-reform)	$+1.645^{*}$	-2.759^{***}	-0.04442^{***}	-0.00209
	(1.90)	(-19.14)	(-12.72)	(-0.56)
Post	-9.590^{***}	-3.552^{***}	$+0.02703^{***}$	$+0.01057^{***}$
	(-21.51)	(-24.23)	(13.78)	(5.22)
Treated $\times$ Post	$+6.803^{***}$	$+424.07^{***}$	-0.00604	-0.04445^{***}
	(6.26)	(480.07)	(-0.56)	(-2.84)
Constant	-79.526^{***}	-24.199^{***}	$+0.29689^{***}$	$+0.11598^{***}$
	(-27.55)	(-26.82)	(41.15)	(13.56)
Observations	128,542	128,542	128,532	128,532
R^2	0.720	0.971	0.026	0.013

All columns include the baseline controls X_i and a gender dummy. Columns (1)–(2): OLS first-stage regressions for total months worked and reform-induced months (total months worked $\times$ treated-post indicator); instruments are the cohort-level NRA shift in months, the individual contribution shortfall below the exemption threshold, and expected years of contributions at age 65; joint F-statistics $15{,}247$ and $3{,}924$ respectively. Column (3): OLS structural equation (endogeneity test $\chi^2(2) = 1{,}210.9$, $p < 0.001$; OLS inconsistent). Column (4): 2SLS preferred specification; heteroskedasticity-robust standard errors; z-statistics in parentheses. The reduced-form estimate of Treated $\times$ Post on mortality appears in Table 6. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Three findings emerge from the structural table. First, the first-stage equations confirm that the reform is a strong instrument: the joint F-statistic for the excluded instruments is $15{,}247$ for total months worked and $3{,}924$ for reform-induced months, far above standard weak-instrument thresholds. The coefficient on Treated $\times$ Post in column (2) implies that the reform induced approximately 424 additional units of the reform-induced months regressor for treated post-reform workers, corresponding to the large compelled work extension documented in Figure 14.

Second, the OLS structural estimate (column 3) is inconsistent: the endogeneity test strongly rejects exogeneity ($\chi^2(2) = 1{,}210.9$, $p < 0.001$), and the OLS coefficient on total months worked (-0.000474) is an order of magnitude larger than the 2SLS estimate (-0.0000641), reflecting the well-known negative correlation between health and labour supply.

Third, and most importantly, the 2SLS structural equation (column 4) reveals a sign reversal between the two months regressors. Voluntary employment is mildly *protective*: each additional month reduces the 60-67 mortality probability by 0.0064 percentage points ($p < 0.001$). By contrast, reform-induced employment is *harmful*: once instrumented, each unit of the compelled-work interaction raises mortality by 0.0133 percentage points ($p < 0.001$), a net effect of $+0.0069$ pp per additional month worked under reform compulsion. Across the 6.80 additional months induced by the reform (column 1 FS), this translates into an implied months-channel mortality cost of approximately $+0.047$ percentage points. After purging the months channel, the residual Treated $\times$ Post coefficient turns negative (-4.4 pp), suggesting that the non-hours dimensions of the reform—income support, reduced job-search stress, continued social insurance coverage—are themselves slightly protective. The mortality cost documented in Table 6 thus reflects the *net dominance of the harmful work-extension channel* over these partially offsetting protective forces.

The sign reversal between voluntary and reform-induced employment is consistent with the occupational profile of treated workers. Voluntary employment at older ages tends to be concentrated in less physically demanding jobs or part-time arrangements that preserve health. Reform-compelled employment, by contrast, disproportionately keeps workers in the jobs they already hold—which for a large share of treated men means construction, manufacturing, and mining—generating the excess hazard captured by the positive coefficient on reform-induced months. This mechanism operates more strongly for men, consistent with their higher baseline mortality, greater occupational hazard exposure, and the roughly twofold larger reduced-form DiD documented in Section 7.1.

Taken together, three channels operate simultaneously and with gender-specific weights, generating the asymmetric mortality pattern of Table 6: extended working lives can preserve health through income and social engagement (Gorry et al., 2018); much of the response is a shift into post-65 pension claiming rather than substantive employment (Section 4.1), so the health-relevant work extension is smaller than the NRA change implies; and workers in physically demanding occupations—concentrated among men—bear an excess health cost when forced to delay exit (Bellés-Obrero et al., 2022; Bloemen et al., 2017). The third channel dominates among men, while the protective and exposure channels roughly offset among women, leaving an imprecise female effect. The social security wealth evidence of Figure 13 reinforces this reading: with no differential pension-wealth losses for the treated, the mortality response must operate through the conditions under which working life is extended rather than through income losses. The analytical decomposition (Section 7.1) places the reform's contribution to the male–female gap at about $+0.81$ pp; because the male response is roughly twice the female one, men account for the majority ($\approx 60\%$) of the ≈ 282 implied excess 60–67 deaths despite being a minority of treated post-reform workers: when underlying risks are sufficiently unequal, the intensity of the effect outweighs numerical exposure.

8 Conclusions

This paper studies the 2011 Spanish pension reform—an emblematic case of the broader European turn toward higher statutory retirement ages and tighter early-exit access—using Sun–Abraham cohort-interaction estimators with heteroskedasticity-robust standard errors and individual-level MCVL data. The contribution-contingent design supplies two complementary sources of variation: a cohort-level shift in the applicable NRA schedule and a within-cohort gradient in how far each worker falls short of the long-career exemption. Together these let us trace the reform cohort by cohort while guarding against the heterogeneous-treatment bias of conventional two-way fixed effects.

Three sets of findings emerge robustly across specifications.

Labour supply and retirement timing. The reform raised effective retirement ages by roughly eight months on average for fully treated cohorts, and the dose-response estimates confirm that the response scales with the depth of exposure to the higher NRA. The adjustment did not take the form of additional early exit: the dominant margin was a reallocation away from the age-65 focal point, with the mass retiring exactly at 65 and the mass retiring before 65 both contracting sharply while retirement after 65 rises by a comparable amount. The reform thus pushed retirement past the old age-65 anchor rather than into earlier exit—overturning a common reading of the reform-era data and underscoring that much of the behavioural response is a delay in pension claiming rather than a one-for-one extension of substantive employment.

Gender asymmetry. The effects are markedly gendered, and the asymmetry is rooted in the contribution-contingent rule itself. Because women accumulate fewer contribution years, they are less likely to clear the long-career threshold, face larger NRA increases, and constitute the majority of treated workers. Their retirement delay, their shift into post-65 exit, and their retreat from pre-65 retirement all exceed the corresponding male responses, so women bear a disproportionate share of the reform burden while also losing the age-65 focal point that had served shorter-career workers.

Mortality. The reform raises premature 60–67 mortality among the treated, significantly so for men. Its clean causal contribution widens the male mortality excess relative to women, even though the effect on the observed population gap is muted and even sign-sensitive, because the large low-mortality male control group dominates any sample-weighted average. An instrumental-variables decomposition is consistent with the underlying mechanism: voluntary late-career work is mildly protective, whereas reform-compelled work in physically demanding jobs is harmful, and this compelled-extension channel dominates among men. Social security wealth plays no role: from the 1947 reference cohort onwards the reform did not substantially affect the pension wealth of treated workers relative to controls, ruling out an income-based explanation. The pattern aligns with men's greater occupational-hazard exposure and with positive health selection among the women who remain attached to the labour market into their early sixties, broadly echoing Bellés-Obrero et al. (2022).

These results carry direct lessons for reform design. The long-career exemption—shared by the German *Rente mit 63* and proposed elsewhere—appears to shield long-career workers, but its gender incidence is the reverse: it exposes short-career workers, disproportionately women, to the full statutory increase. Reforms seeking to raise effective retirement ages more equitably should consider gender-neutral equivalents, such as contribution-density adjustments or occupation-specific pathways, rather than raw tenure thresholds that encode historical labour-market inequalities. The predominance of delayed pension claiming over genuine employment extension also suggests that part of the fiscal saving from later claims may be recovered through disability, partial retirement, and long-term unemployment among workers unable to extend employment (Staubli and Zweimüller, 2013; Bellés-Obrero et al., 2022), and offset by an underappreciated mortality cost borne by those least able to defer exit. Several questions remain open. Our evaluation covers cohorts up to 1957, who face just over half of the full two-year NRA increase; the cohorts that follow are progressively more exposed, with the 1962 cohort (turning 65 in 2027) the first to face the fully phased-in rules, which then apply unchanged to all later cohorts, so the estimates reported here are plausibly a lower bound on the reform's long-run effects. In addition, we observe mortality only to age 67, so longer-run longevity effects may differ once the composition of retirees changes, and the reform's consequences for the gender pension gap—through lower replacement rates and interactions with minimum-pension and survivor rules—warrant a full welfare accounting.

References

Bellés-Obrero, C., Jiménez-Martín, S., and Ye, H. (2022). The effect of removing early retirement on mortality. IZA Discussion Paper No. 15577; FEDEA Working Paper 2022-06.

Bellés-Obrero, C., Jiménez-Martín, S., and Vall Castelló, J. (2022). Minimum working age and the gender mortality gap. *Journal of Population Economics*, 35(4):1897–1938. DOI: 10.1007/s00148-021-00858-x.

Bellés-Obrero, C., Flores, M., García-Gómez, P., Jiménez-Martín, S., and Vall-Castelló, J. (2025). Social security reforms and inequality among older workers in Spain. In Börsch-Supan, A. and Coile, C. C. (eds.), *Social Security Programs and Retirement around the World: The Effects of Pension Reforms on the Income Distribution of Retirees*. University of Chicago Press for the NBER, Chicago. NBER Working Paper No. 34577.

Bloemen, H., Hochguertel, S., and Zweerink, J. (2017). The causal effect of retirement on mortality: evidence from targeted incentives to retire early. *Health Economics*, 26(12):e204–e218.

Boldrin, M., Jiménez-Martín, S., and Peracchi, F. (1999). Social security and retirement in Spain. In Gruber, J. and Wise, D. (eds.), *Social Security and Retirement around the World*, pp. 305–353. University of Chicago Press for the NBER, Chicago.

Bozio, A., Rabaté, S., Roulet, A., and Villoing, G. (2021a). Reforms of the French pension system: evidence on incentive effects. *Journal of Public Economics*, 194:104346.

Bozio, A., Garrouste, C., and Perdrix, E. (2021b). Impact of later retirement on mortality: evidence from France. *Health Economics*, 30(5):1178–1199.

Callaway, B. and Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427.

Carrino, L., Glaser, K., and Avendano, M. (2020). Later retirement, job strain, and health: Evidence from the new State Pension age in the United Kingdom. *Health Economics*, 29(8):891–912.

Carta, F. and De Philippis, M. (2021). Working horizon and labour supply: the effect of raising the full retirement age on middle-aged individuals. *European Economic Review*, 138:103834.

Cengiz, D., Dube, A., Lindner, A., and Zentler-Munro, D. (2022). Seeing beyond the trees: Using machine learning to estimate the impact of minimum wages on labor market outcomes. *Journal of Labor Economics*, 40(S1):S203–S247.

Coe, N. B. and Zamarro, G. (2011). Retirement effects on health in Europe. *Journal of Health Economics*, 30(1):77–86.

Coile, C. and Gruber, J. (2007). Future Social Security entitlements and the retirement decision. *Review of Economics and Statistics*, 89(2):234–246.

Conde-Ruiz, J.I. and Gonzalez, C. (2012) Spain 2011 Pension Reform FEDEA, Documento de Trabajo 2012-03

Currie, J. and Gruber, J. (1996). Health insurance eligibility, utilization of medical care, and child health. *Quarterly Journal of Economics*, 111(2):431–466.

Cutler, D. M. and Gruber, J. (1996). Does public insurance crowd out private insurance? *Quarterly Journal of Economics*, 111(2):391–430.

Della Giusta, M. and Longhi, S. (2021). Stung by pension reforms: The unequal impact of changes in state pension age on UK women and their partners. *Labour Economics*, 72:102049.

Correia, S. (2016). Linear models with high-dimensional fixed effects: an efficient and feasible estimator. Working paper, Duke University. `http://scorreia.com/research/hdfe.pdf`.

Cribb, J., Emmerson, C., and Tetlow, G. (2016). Signals matter? Large retirement responses to limited financial incentives. *Labour Economics*, 42:203–212.

Fitzpatrick, M. D. and Moore, T. J. (2018). The mortality effects of retirement: evidence from Social Security eligibility at age 62. *Journal of Public Economics*, 157:121–137.

García-Gómez, P., Jiménez-Martín, S., and Vall Castelló, J. (2012). Health, disability, and pathways into retirement in Spain. In Wise, D. A. (ed.), *Social Security Programs and Retirement around the World: Historical Trends in Mortality and Health, Employment, and Disability Insurance Participation and Reforms*, pp. 127–174. University of Chicago Press for the NBER, Chicago.

García-Pérez, J. I., Jiménez-Martín, S., and Sánchez-Martín, A. R. (2013). Retirement incentives, individual heterogeneity and labor transitions of employed and unemployed workers. *Labour Economics*, 20:106–120.

Geyer, J., Haan, P., Hammerschmid, A., and Peters, M. (2020). Labor market and distributional effects of an increase in the retirement age. *Labour Economics*, 65:101817.

Geyer, J. and Welteke, C. (2019). Closing routes to retirement for women: how do they respond? *Journal of Human Resources*, 54(1):196–238.

Gorry, A., Gorry, D., and Slavov, S. N. (2018). Does retirement improve health and life satisfaction? *Health Economics*, 27(12):2067–2086.

Gruber, J. and Wise, D. A. (1999). *Social Security and Retirement Around the World*. University of Chicago Press.

Gruber, J. and Wise, D. (eds.) (2004). *Social Security Programs and Retirement around the World: Micro-Estimation*. University of Chicago Press for the NBER, Chicago.

Hagen, J. (2018). The effects of increasing the normal retirement age on health care utilization and mortality. *Journal of Population Economics*, 31(1):193–234.

Hallberg, D., Johansson, P., and Josephson, M. (2015). Is an early retirement offer good for your health? Quasi-experimental evidence from the army. *Journal of Health Economics*, 44:274–285.

Hernæs, E., Markussen, S., Piggott, J., and Vestad, O. L. (2013). Does retirement age impact mortality? *Journal of Health Economics*, 32(3):586–598.

Jiménez-Martín, S. and Sánchez-Martín, A. R. (2004). The effect of pension rules on retirement monetary incentives with an application to pension reforms in Spain. *Hacienda Pública Española / Review of Public Economics*, 169(2):35–66.

Jiménez-Martín, S. and Sánchez-Martín, A. R. (2007). An evaluation of the life cycle effects of minimum pensions on retirement behavior. *Journal of Applied Econometrics*, 22(5):923–950.

Kuhn, A., Staubli, S., Wuellrich, J.-P., and Zweimüller, J. (2020). Fatal attraction? Extended unemployment benefits, labor force exit, and mortality. *Journal of Public Economics*, 191:104087.

Laun, L. and Wallenius, J. (2015). A life-cycle model of health and retirement: the case of Swedish pension reform. *Journal of Public Economics*, 127:127–136.

MacKinnon, J. G. and Webb, M. D. (2017). Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics*, 32(2):233–254.

Manoli, D. S. and Weber, A. (2016). Nonparametric evidence on the effects of financial incentives on retirement decisions. *American Economic Journal: Economic Policy*, 8(4):160–182.

Ministerio de Empleo y Seguridad Social (2011). *Ley 27/2011, de 1 de agosto, sobre actualización, adecuación y modernización del sistema de Seguridad Social*. *Boletín Oficial del Estado*, núm. 184, 2 de agosto de 2011, pp. 87495–87544. Disposición Transitoria 4ª: régimen de jubilación de los trabajadores que tuvieran la condición de mutualistas con anterioridad al 1 de enero de 1967.

OECD (2023). *Pensions at a Glance 2023: OECD and G20 Indicators*. OECD Publishing, Paris. https://doi.org/10.1787/678d8f8a-en.

Sánchez-Martín, A. R., García-Pérez, J. I., and Jiménez-Martín, S. (2014). Delaying the normal and early retirement ages in Spain: behavioural and welfare consequences for employed and unemployed workers. *De Economist*, 162(4):341–375.

Saporta-Eksten, I., Shurtz, I., and Weisburd, S. (2021). Social security, labor supply, and health of older workers. *Journal of the European Economic Association*, 19(4):2168–2208.

Staubli, S. and Zweimüller, J. (2013). Does raising the early retirement age increase employment of older workers? *Journal of Public Economics*, 108:17–32.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Wooldridge, J. M. (2021). Two-way fixed effects, the two-way Mundlak regression, and difference-in-differences estimators. SSRN Working Paper 3906345.

A1 Appendix

A1.1 Pre-Trend Tests and DiD Results

See Table 8

Table 8: Joint Pre-Trend Tests: Treated vs Controls (Robust SE)

Outcome	$F(4,\cdot)$	p-value
Retirement age (months)	1.59	0.173
Very early retirement (<61)	7.73	< 0.001
Retire before 65 ($[61, 65)$)	2.42	0.046
Retire at 65	0.91	0.459
Retire after 65 (>65)	0.91	0.454
Died before age 67	0.38	0.826

H_0: pre-reform coefficients at $k = -5, -4, -3, -2$ jointly zero. Heteroskedasticity-robust variance estimator. Large F-statistics reflect permanent compositional differences (treated workers are disproportionately women with shorter careers), not differential pre-trends on the causal margin of interest.

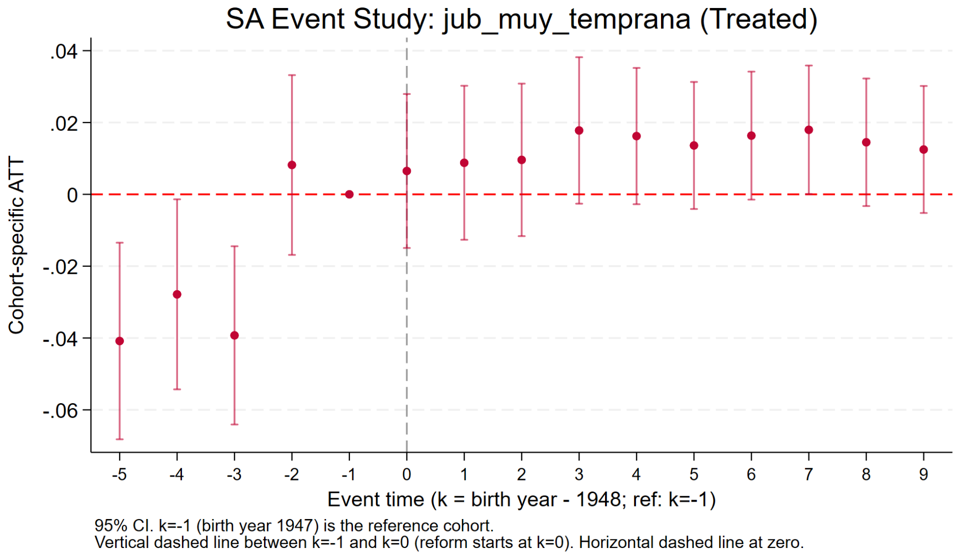
Table 9: Difference-in-Differences: Main Outcomes

	Ret. age (months)		Pr(retire after 65)		Pr(retire at 65)	
	Pooled	Women	Pooled	Women	Pooled	Women
Treated × Post	7.855^{***}	6.225^{***}	0.406^{***}	0.438^{***}	-0.301^{***}	-0.281^{***}
	(0.40)	(0.60)	(0.006)	(0.009)	(24.8)	(47.9)
Part. treat. × Post	1.765^{***}	2.778^{***}	0.197^{***}	0.220^{***}	-0.104^{***}	-0.250^{***}
	(0.58)	(1.04)	(0.008)	(0.013)	(36.4)	(82.7)
N	114,387	35,918	132,112	43,082	114,387	35,918
R^2	0.168	0.222	0.262	0.280	0.717	0.723

OLS with controls: gender (pooled), years of contributions at 50, education, contribution days, gaps, mutualista status, sector shares, contribution group. Robust SEs in parentheses. Estimates from a complementary pooled difference-in-differences specification (Treated/Partially treated × Post); the main causal estimates in the text come from the Sun–Abraham event studies (Section 5). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

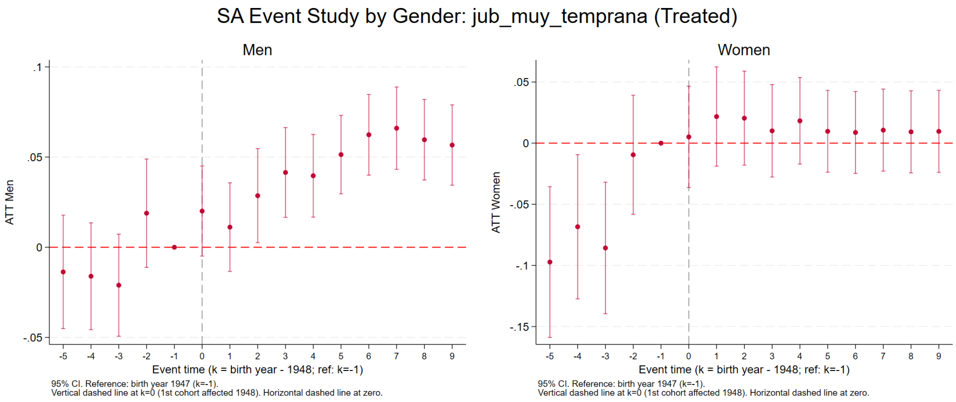
A2 Additional Event-Study Figures

Figure 15: Event-Study: Very Early Retirement (< 61 years). Treated vs Controls.



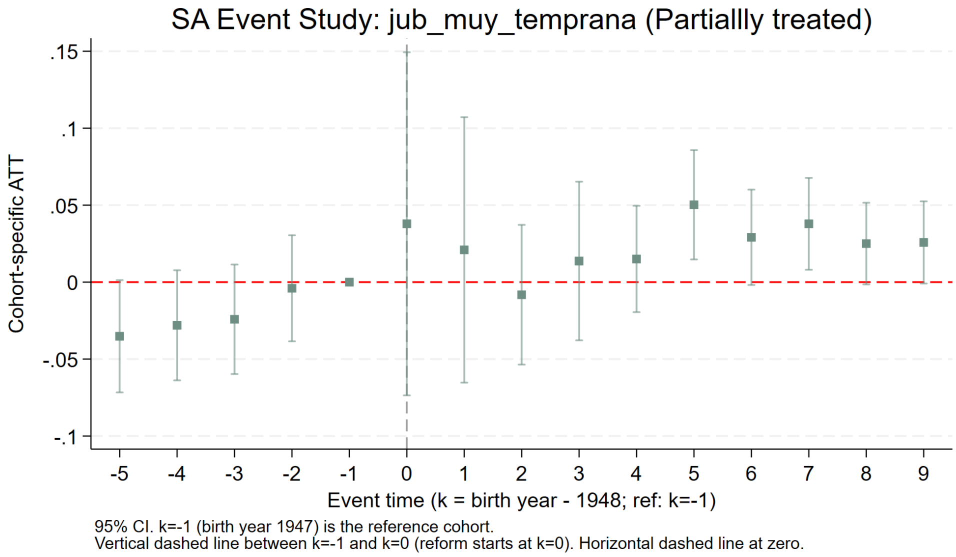
Note: See notes to Figure 5.

Figure 16: Event-Study by Gender: Very Early Retirement (< 61 years). Treated vs Controls.



Note: See notes to Figure 9.

Figure 17: Event-Study: Very Early Retirement (< 61 years). Partially treated vs Controls.



Note: See notes to Figure 5.