KU LEUVEN

FACULTY OF ECONOMICS AND BUSINESS

How Raising the Statutory Retirement Age Affects Household Savings

Evidence from the 2007 German Pension Reform

Anqi XU

R0960207

Thesis submitted to obtain the degree of

Master of Advanced Studies in Economics (Leuven)

Promoter: Prof. Hannah Zillessen

Academic year 2024-2025

KU LEUVEN

FACULTY OF ECONOMICS AND BUSINESS

How Raising the Statutory Retirement Age Affects Household Savings

Evidence from the 2007 German Pension Reform

Population ageing poses a major challenge to the sustainability of pay-as-you-go pension systems. In response, Germany's 2007 pension reform gradually raised the statutory retirement age (SRA) from 65 to 67. This thesis examines how that policy change affected private household saving behaviour. Using the German Socio-Economic Panel (SOEP) and the Sparen und Altersvorsorge in Deutschland panel (SAVE), I implement a difference-in-differences design that compares employees, who are bound by the new rule, with self-employed workers, who are exempt. To disentangle mechanisms, I match realised saving flows with forward-looking expectations on retirement age and life expectancy available in SAVE.

Extensive-margin analysis using Probit and Heckman-Probit models reveals no statistically significant effect on the probability of saving, while the intensive-margin results for "always-savers" show that the reform reduced logged monthly savings by 0.37. To uncover the underlying channel, I leverage SAVE's forward-looking modules: the reform increased expected retirement age by around 1.7 years, and shortened expected years in retirement by around 3.5 years. Under the Life-Cycle Hypothesis, a shorter anticipated retirement horizon rationalises the observed drop in savings. Subgroup analysis further indicates that this adjustment was more pronounced among women.

These findings contribute to the literature by providing causal evidence on the effects of pension-age reforms on private saving, identifying retirement expectations as a key mechanism, and highlighting gender-specific responses. The results also suggests policymakers to pair reforms delaying retirement eligibility with targeted tax-advantaged pension schemes or financial-literary initiatives to vulnerable groups to prevent new distributional gaps in retirement security.

Acknowledgements

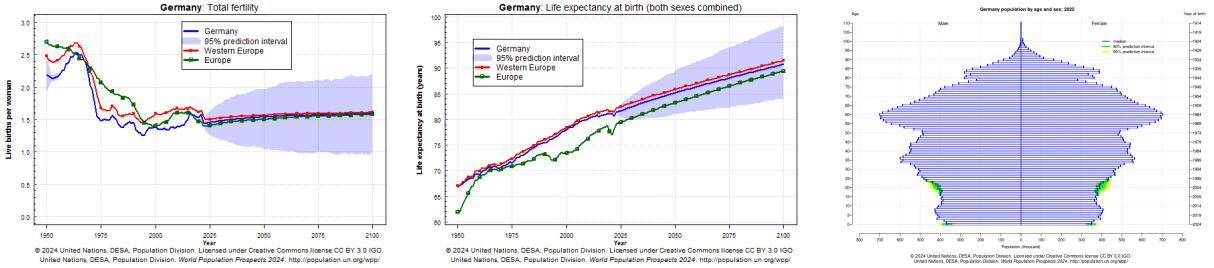
I would like to express my deepest gratitude to my thesis supervisor, Prof. Hannah Zillessen, for her continuous guidance, constructive feedback, and support throughout this research process. Her encouragement and suggestions have been invaluable to the completion of this thesis.

Table of Contents

Chapter 1

Introduction

Over the past few decades, population aging has become one of the most pressing challenges for many countries worldwide. This issue is driven by both increasing life expectancy and low fertility rates, which is constantly below the replacement level. Germany, as one of the most rapidly aging societies in Europe, exhibits a population pyramid with a narrow base, while a significant portion of its population falls into older age groups. This demographic shift is particularly pronounced as the baby boomer generation, born in the 1950s and 1960s, is reaching retirement age. As a result, Germany faces growing pressure on its pension system, as its core ”pay-as-you-go” (PAYG) scheme relies directly on current workers financing the pensions of retirees. With fewer working-age individuals supporting an increasing number of retirees—i.e., a rising dependency ratio—the sustainability of the system is increasingly at risk.



(a) Total Fertility (b) Life Expectancy at Birth (c) Population Pyramid: 2025

Figure 1.1: Demographic Trends. Source: United Nations, DESA, Population Division (United Nations, DESA, Population Division, 2024)

One natural policy response to this demographic challenge is to extend the working life, given the increase in life expectancy. In line with this approach, the 2007 German pension reform raised the statutory retirement age (SRA) from 65 to 67. While the effectiveness of the reform in prolonging employment has been well studied, its social implications remain underexplored. This thesis aims to evaluate how the reform affects household saving behavior. Household saving is not only a critical component of economic stability and financial security, but also serves as an indicator of retirement preparedness and may reflect heterogeneous effects across individuals with different socio-economic backgrounds. Understanding these dynamics is essential for assessing the broader impact of pension reforms on financial well-being.

I hypothesize that postponing pension eligibility to age 67 will lead households to lower their saving targets, since the reform may lead to a compressed expected retirement period. The delaying retirement directly shifts the expected working span longer, however, the subjective life expectancy typically adjust more gradually and does not

necessarily respond to the reform. Moreover, some studies suggest that raising the statutory retirement age (SRA) can worsen health, potentially leading individuals to revise their expected lifespan downward. (Barschkett et al., 2022) These factors together may shorten the perceived retirement horizon.

Under life-cycle hypothesis, rational households will save less during working years to smooth the consumption. Although the reform was primarily designed to react to the rising dependency ratio and strengthen pension finances by keeping people employed longer, it also alters households' intertemporal budget constraints. The private-saving response interacts directly with the reform's core aim of strengthening pension finances through later retirement. The resulting substitution effect between public pension and private savings for old age is therefore crucial for assessing the net effect on household long-run welfare, such as retirement preparedness. Understanding this interplay and the distribution of impact is therefore key both to evaluating the reform's true impact and to designing complementary measures.

This thesis is structured as follows. Chapter 2 reviews the existing literature on household saving behavior and the effects of pension reforms, identifying the research gaps this study addresses; Chapter 3 gives a historical overview of the German pension system and reforms, with a special focus on the reform implemented in 2007; Chapter 4 describes the data and outlines the empirical methodology used in the analysis; Chapter 5 begins with descriptive statistics and presents the results of the econometric models; Chapter 6 conducts robustness checks to assess the validity of the main findings; Chapter 7 explores potential mechanisms underlying the observed effects; Chapter 8 provides a broader discussion of the results, explores policy implications, gives suggestions for future research, and finally concludes the thesis.

Chapter 2

Literature Review

This thesis relates to three strands of literature. First, empirical evidence of the implications of pension reforms on labor market outcome and well-being. Second, more directly related to our research question, how pension reforms affect household savings. Third, theoretical frameworks on the relationship between retirement and savings. Building on these foundations, this section is concluded by identifying gaps in the existing literature.

2.1 How pension reforms affect labor outcome and well-being

A significant body of research examines how raising the retirement age and other pension reforms influence labor supply. Pension reforms are generally effective in shaping people's expected retirement age and reversing the trend of early retirement (Buchholz et al., 2013; De Grip et al., 2013; Etgeton et al., 2023; Nivalainen, 2023). The key mechanism driving changes in retirement behavior is a shift in reference points rather than active adjustments based on life-cycle optimization (Lalive et al., 2017; Seibold, 2021). As a result, employment rates tend to increase among affected populations, leading to an overall rise in labor market supply (Cribb et al., 2016; Lalive & Staubli, 2015; Pilipiec et al., 2021b; Soosaar et al., 2021).

Despite increased labor force participation, job satisfaction remains largely unchanged (Pilipiec et al., 2021a). Some studies suggest that pension reforms enhance career advancement opportunities, yet they do not significantly affect wages (Gohl et al., 2023). However, extending the working life can have adverse health consequences (Barschkett et al., 2022).

The impact of raising the statutory or early pensionable age is often asymmetrical across different socio-economic groups. The magnitude and direction of these effects vary significantly depending on individuals' education levels, wages, and health status (Ardito, 2021; Buchholz et al., 2013; Honekamp & Schwarze, 2010; Staubli & Zweimüller, 2013). Generally, better-educated, higher-wage, and healthier individuals are more responsive to these changes. (Bucher-Koenen & Lusardi, 2011; Oehler & Werner, 2008)

2.2 How Pension Reforms Affect Household Savings

Several studies analyze the effects of raising retirement age on household savings. For instance, an increase in the early retirement age in Germany led to a decline in savings, with better-educated women being more likely to reduce their savings rate (Etgeton et al., 2023). Conversely, an increase in the pensionable age in Japan resulted in higher savings rates (Okumura & Usui, 2014).

Merely providing individuals with information about their expected pension payments has been found to boost tax-deductible private savings in Germany (Dolls et al., 2018).

Another type of pension reform, the introduction of the "Riester pension"—a government-subsidized private pension scheme—has been linked to a displacement effect, where individuals reallocate savings from other purposes (e.g., real estate) toward retirement savings (A. H. Börsch-Supan et al., 2007; Coppola & Wilke, 2010). However, findings on this issue remain contradictory, as some studies report no crowding-out effect and suggest that national aggregate savings increased instead (A. Börsch-Supan et al., 2012). The Riester pension also disproportionately affected certain groups, with women and part-time employees experiencing greater difficulty predicting their future replacement rates (Honekamp & Schwarze, 2010).

There are empirical evidence showing the reforms impact saving decisions through changes in pension wealth. A study in Poland finds that for every unit increase in pension wealth, household savings decrease by 0.3 units (Lachowska & Myck, 2018). Similarly, a transition to a defined contribution scheme in Chile reduced precautionary savings by 1.4% (Cifuentes, n.d.). In the UK, A study also finds a negative relationship between earnings-related pension wealth and private savings, with additional evidence that when pension benefits are uniform and not linked to earnings, households' saving behaviours remain largely unchanged. (Attanasio & Rohwedder, 2003)

2.3 Theoretical Frameworks on Retirement and Savings

A widely used framework to understand household saving behaviour is the Life Cycle Hypothesis (LCH), first proposed by Modigliani and Brumberg in 1954 (Modigliani & Brumberg, 2013). The LCH posits that rational individuals aim to maximize their utility from consumption by smoothing their consumption path over their entire finite lifespan. This implies a sequential decision-making process where individuals allocate their lifetime resources – including current wealth and anticipated future income – across different periods of their life. The model typically distinguishes between a productive working phase and a subsequent non-productive retirement phase. During the working years, individuals are expected to save to accumulate sufficient wealth to maintain their desired consumption level during retirement.

Under the life-cycle hypothesis (LCH), households smooth consumption over the working and retirement phases by allocating lifetime resources accordingly. When a reform raises the statutory retirement age, the expected retirement spell shortens; holding life expectancy fixed, the LCH predicts a lower desired stock of retirement wealth and thus lower saving during working years (Attanasio & Rohwedder, 2003). At the same time, rising life expectancy—one motivation for the reform—pushes in the

opposite direction by lengthening the retirement horizon and increasing desired wealth, thereby raising savings (OECD, 2006). Hence, the net effect of such reforms on savings is a priori ambiguous and empirical.

However, the standard life-cycle model also has difficulties explaining the heterogeneity in retirement wealth and inconsistency of discount rate over time. (Browning & Crossley, 2001) Moreover, uncertainty and anxiety may play a more prominent role in retirement saving decisions when individuals are confronted with policy reforms. These psychological factors can lead to an increase in precautionary savings. (Okumura & Usui, 2014)

Building on this framework, this thesis investigates whether the 2007 German pension reform—which gradually increased the statutory retirement age—led to adjustments in household saving behaviour consistent with life cycle predictions.

2.4 Gaps in the Literature

While previous studies have extensively examined the labor market and well-being consequences of pension reforms, the effect on household saving behavior—particularly in the context of the 2007 German pension reform—remains relatively unexplored.

Moreover, evidence on heterogeneity in saving responses is scarce. Some studies have pointed out that pension reforms often affect different socio-economic groups in different ways, and these groups may also exhibit varying behavioral responses. Understanding the distribution of the consequences of the 2007 pension reform is crucial for future policy design.

In addition, the mechanisms underlying saving adjustments, such as changes in retirement expectations, precautionary motives, or financial constraints, are not well identified in the literature. Finally, most studies focus on short-term responses, leaving the longer-term behavioral adjustments largely unexamined due to data limitations or identification challenges.

By addressing these gaps, this study contributes to the literature by investigating the causal effect of the 2007 German pension reform on saving behavior.

Chapter 3

Institutional Background

3.1 The German Pension System

Initially established as a fully funded scheme, the German pension system was later converted to a pay-as-you-go system (PAYG), and eventually transitioned to a multi-pillar system through various reforms(Holzmann et al., 2006). To best describe today's complex structure, I examine it in the context of the World Bank definition of the five pillared pension system. (Holzmann & Hinz, 2012)

1. The "zero pillar" provide basic social assistance to protect retirees from old-age poverty. In Germany, this basic pension only applies to a small portion of retirees who are in the bottom of the income-percentile.

2. The "first pillar" is publicly managed pension plan which can be defined benefit, defined contribution or hybrid. In Germany, the pure pay-as-you-go system (PAYG) emerged in 1957 and it still accounts for the largest share of the aggregate retirement income today. (Wilke, 2009)

3. The "second pillar" refers to the mandated occupational or personal pension plans. In the context of Germany, this pillar corresponds to voluntary company pension provided by employers funded by financial assets.

4. The "third pillar" constitute of voluntary personal pension plans. With the introduction of the 2001 Riester Reform, the third pillar along with the second pillar plays an increasing role of German pension system while the PAYG system's role has diminished. (Bonin, 2009)

5. The "fourth pillar" consists of support from family, access to health care system, and other financial and non-financial assets, such as real-estate property. It plays a supplementary role in retirement security.

Building on the five-pillar framework just outlined, the structure of Germany's pension system directly shapes the questions this thesis seeks to answer. The reform directly impacted the PAYG pillar. Increasing the statutory retirement age from 65 to 67 anchors expectations about when pension income anchors expectations about when pension income will begin and alters households' lifetime budget constraints. The impact is then absorbed by the third and fourth pillar. Household savings, which form part of the fourth pillar and effectively substitute for third-pillar voluntary pension products, become the adjustable margin through which consumption is smoothed. Understanding how much of the overall adjustment falls within these accounts is essential for gauging the reform's net effect on aggregate capital formation and welfare.

The dynamics between and within the five pillars has changed through several reforms to ensure the sustainability and stability of the pension system. In the next section, I will introduce the major German pension reforms after 1990.

3.2 Recent Pension Reforms in Germany

The pay-as-you-go (PAYG) public pension is the main pillar of the German system. Historically generous—high replacement rates and relatively early retirement—it has come under pressure from demographic change. Persistently low fertility reduces the inflow of contributors, while rising life expectancy lengthens the average duration of benefit receipt, straining PAYG finances. Recent reforms can be grouped into (i) *parameter* reforms within the PAYG pillar and (ii) *structural* reforms that strengthen the other pillars to offset the PAYG pillar's declining role.

3.2.1 Parameter reforms within the PAYG system

Within the PAYG system, the contribution rate–a percentage of individual income allocated to the public pension fund–is related to the system dependency ratio quantifying the relationship between the number of beneficiaries and contributors, the benefit ratio which reflects the relationship between average pension and average wage, and finally, the proportion of funding derived from sources outside the pension scheme, such as tax revenue or withdrawals from other reserve funds. (Hinrichs, 2021)

To address the rising contribution rate driven by demographic changes, several reforms aimed to adjust the parameters were taken place since the early 1990s.

Firstly, the dependency ratio can be lowered by shifting upwards the statutory and early retirement ages which simultaneously prolongs working lives contributing to the system and limits access to the retirement benefits. The 1992 and 1999 reforms were the first responses to the over generosity of the past pension scheme. Efforts were made especially to restrict early retirement, where the early retirement age for women and unemployed were abolished, practically raising the eligible age of pension from 60 to 65. Meanwhile, the actuarial adjustment for early retirement and late retirement were introduced to financially discourage early retirement and motivate late retirement. **The 2007 reform** aimed to raise normal retirement age from 65 to 67 with a gradual phase-in period, starting in 2012 and expected to end in 2030. The 2017 reform introduced the Flexirente initiative, designed to further incentivize employees to work beyond normal retirement age. (Bonin, 2009; Wilke, 2009)

Secondly, the benefit ratio were adjusted by modifying the indexation of pension level. The 1992 reform marked the first shift from indexing pensions based on gross wages to net wages. It was further modified in 2001 by the Riester Reform to keep both the contribution rate and replacement stable and adequate by accounting for the lagged change in gross wage, contribution rate to PAYG and private pension accounts, and a sensitivity factor related to demographic changes. In the 2004 reform, the modification of indexation was extended with a "sustainability factor" which could be seen as an upgraded version of sensitivity factor. It captures not only the dependency rate, but also the entire demographic evolvement including life expectancy, birth rate, and migration.

This shift also signifies the transition of German PAYG system from defined benefit towards the quasi-notional defined contribution mode. (Wilke, 2009)

3.2.2 Structural reforms among other pillars

To compensate for the deliberately declined pension provided by the first pillar, the 2001 reform first introduced the Riester pension to integrate private pension to be an essential and growing part of German pension system. This reform encourages individuals to contribute to certified private pension accounts by offering direct savings subsidies and child allowances. Alternatively, the contribution to the Riester pension can be tax-deductible with a specific limit, targeting high-income workers. (Hinrichs, 2021)

Later in the 2004 reform, the pillars of occupational pensions and private pensions were further reinforced through various administrative changes such as expanded eligibility, dynamic pension benefit, and enhanced transparency.

3.3 The German Pension Reform 2007

As part of the move toward a multi-pillar, sustainable pension system, the 2007 reform raised the statutory retirement age (SRA) for employees from 65 to 67, while leaving the self-employed unaffected. Starting from 2012, with a gentle cohort-oriented phased-in stage of 18 years, the process is planned to complete in 2030 (Coppola & Wilke, 2010). Exceptionally, for employees with a service life at least 45 years, these long-time insured workers could retire at the pre-reform age of 65, without facing any reduction in benefits. (See Table 3.1)

[1] if born on 1.1

Birth Year	Ret. Age	Ret. Year[1]	Age (Long-Insured)
1945	65y 0m	2010	65
1946	65y 0m	2011	65
1947	65y 1m	2012	65
1948	65y 2m	2013	65
1949	65y 3m	2014	65
1950	65y 4m	2015	65
1951	65y 5m	2016	65
1952	65y 6m	2017	65
1953	65y 7m	2018	65
1954	65y 8m	2019	65
1955	65y 9m	2020	65
1956	65y 10m	2021	65
1957	65y 11m	2022	65
1958	66y 0m	2024	65
1959	66y 2m	2025	65
1960	66y 4m	2026	65
1961	66y 6m	2027	65
1962	66y 8m	2028	65
1963	66y 10m	2029	65
1964	67y 0m	2031	65
1965	67y 0m	2032	65

Table 3.1: Legal Retirement Ages by Birth Year. Source: Bucher-Koenen & Wilke (Bucher-Koenen & Wilke, 2009)

Chapter 4

Data and Methods

4.1 Data

To exploit both breadth and depth in measuring households' saving responses, this thesis combines two complementary German surveys. The German Socio-Economic Panel (GSOEP) offers a large sample spanning more than three decades, which delivers the statistical power and long-window perspective needed to identify both short-term adjustments and longer-run shifts in saving behavior around pension reforms. On the other hand, the Sparen und Altersvorsorge in Deutschland – Saving and Old-Age Provision in Germany (SAVE) survey—although much smaller—provides rich, dedicated modules on saving, expectations and satisfactions, allowing me to unpack the mechanisms driving those aggregate patterns. Drawing on GSOEP's scale to establish the reform's overall impact and SAVE's granularity to illuminate the expectation channel thus ensures a more robust and nuanced analysis than either dataset alone could deliver.

4.1.1 Data Source

The SAVE dataset is a rich panel survey designed to collect detailed information on household saving behavior in Germany. Initiated in 2001 and ended in 2013, the SAVE study was created to fill critical gaps in data needed to analyze private saving patterns, old-age provision, and the interplay between economic, psychological, and demographic factors affecting financial decisions. The survey includes multiple waves, from 2001 to 2013, using random and quota samples drawn via stratified multistage procedures to ensure representativeness of the German population. The questionnaire is divided into several parts, covering socio-economic status, health, financial literacy, income and asset details, saving motives, and psychological traits (Schunk, 2006). The data is further enhanced by rigorous handling of item non-response through a multiple imputation method (MIMS), allowing analysts to work with complete datasets that preserve the underlying data structure and variability (A. Börsch-Supan et al., 2009).

As the survey collects both conventional financial variables-such as income, wealth, and savings-and non-economic factors including expectations, satisfactions, and attitudes toward the future. This dual focus enables me to explore not only how the 2007 pension reform altered savings but also probe the mechanism through which individuals re-anchor retirement expectations and adjust saving behaviours.

The second dataset I used is the German Socio-Economic Panel (SOEP), the largest and longest-running nationally representative longitudinal study of private households in Germany. Comprising nearly 15,000 households and around 30,000 individ-

uals, the SOEP collects detailed information on a wide range of subjective variables, such as employment status, income, and saving behavior, as well as objective indicators like life satisfaction. Initiated in 1984, the SOEP has evolved over more than 40 years from a basic cross-sectional survey into a multidimensional longitudinal dataset, incorporating innovations, experiments, and collaborations with other studies (Goebel et al., 2019).

For this analysis, The data used is from the SOEP-Core dataset, the central component of the SOEP collection. Based primarily on questionnaire surveys conducted through random sampling, SOEP-Core gathers comprehensive information on household demographics, education, employment biographies, earnings, health, and satisfaction measures (Selin Kara & Zimmermann, 2023). The version used in this study (v38) covers the years 1984 to 2022.

Its large sample size and nearly four decades of data afford the statistical power to detect policy-driven shifts in amount of savings and net saving rates, while its extended time horizon enables analysis of both short-term adjustments (two years around the 2007 reform) and longer-run trends (up to ten years post-reform). Exploiting GSOEP's panel structure, I implement a difference-in-differences design with individual fixed effects and year controls to isolate the causal impact of raising the statutory retirement age on private saving behavior.

4.1.2 Sample Selection

4.1.2.1 For the GSOEP dataset

To construct the analysis sample, I applied several selection criteria and data cleaning steps.

First, the sample is restricted to individuals surveyed from 1992 onward, as questions related to savings behavior were first introduced in the SOEP in 1992. To match the SAVE-based analysis, I then further restrict the survey waves to 2003–2010. To focus on economically active individuals, only those who are employed either full-time or part-time are included; individuals outside the labor force, such as retirees, students, and unemployed persons, are excluded. Furthermore, the analysis focuses on individuals aged 17 and older at the time of the survey, as very young respondents are typically unlikely to have established stable saving behavior.

In line with the research design, the sample is limited to cohorts fully affected by the 2007 pension reform, which gradually raised the statutory retirement age from 65 to 67. Accordingly, only individuals born after 1964 are retained. To maintain consistent treatment and control group over time, individuals who switch between employment types (from employee to self-employed or vice versa) are excluded. A graphical summary of employment type transitions is provided in Appendix Figure 1.

For data quality issues, observations with missing or irregular information on key variables—such as age, sex, employment status, migration background, marital status, and household size—are dropped from the sample. All monetary variables, including savings amounts and household income, are deflated to real euros using the Consumer Price Index (CPI) to account for inflation over time.

To account for compositional changes, the sample is restricted to the households that have been recorded both before and after the reform.

Finally, a subsample restricted to household heads is constructed to better capture saving behavior at the primary decision-making level within households. Moreover, because the savings behavior variables are collected at the household level rather than the individual level, restricting the analysis to household heads prevents inflating the sample size by the number of household members interviewed.

4.1.2.2 For the SAVE dataset

The second analysis sample is constructed from the German SAVE panel (waves 2003–2010). I begin by appending the individual-level data across these seven waves. Wave 2011–12 and 2013 are excluded because the key outcome—realized savings in the previous year—is missing in those waves.

Although the SAVE datasets provide imputed values based on a Markov Chain Monte Carlo (MCMC) method to address item nonresponse, I revert the outcome variable to missing whenever it was originally unobserved. This adjustment improves the accuracy of our inference by preserving the true pattern of missingness in the dependent variable. In addition, I generate a separate "indicator" dataset that flags missing values in the forward-looking expectation variables. These indicators are used as exclusion criteria in Heckman selection model.

Consistent with the approach used in the GSOEP dataset, I restrict the sample to individuals born after 1964 to ensure full exposure to the 2007 pension reform, and to those actively participating in the labor force. I also deflate all monetary variables—such as savings and income—using the Consumer Price Index (CPI) (Statistisches Bundesamt (Destatis), 2024), and exclude individuals under age 17.

Finally, given the relatively small sample size of the SAVE panel, I do not impose strict restrictions to control for compositional change over time. Instead, I construct panel-continuation indicators (`s03`–`s09`) that equal one if the respondent appears in wave t and also in wave $t+1$, for the years 2003 through 2010. The final SAVE sample includes all respondent-year observations from 2003 to 2010 with valid information on realized savings, forward-looking expectations, and key demographic and economic covariates.

4.2 Empirical Strategy

This study employs a Difference-in-Differences (DiD) approach to estimate the causal effect of the 2007 German pension reform on household saving behavior. Employees, who were directly affected by the reform, serve as the treatment group, while self-employed individuals, who were not subject to the reform, constitute the control group. DiD is used to isolate the effect of the policy change (Angrist & Pischke, 2008).

The canonical parallel trend assumption can be relaxed to holding conditional on covariates. The identification hinges on employees and self-employed forming valid counterfactuals after conditioning on a rich set of pre-determined characteristics (Roth et al., 2023). Substantively, the 2007 reform changed the statutory retirement age only for employees, while self-employed workers were not bound by the new rule, yet both groups are drawn from the same birth cohorts, live in the same macroeconomic, informational and cultural environment. Intuitively, the self-employed could provide the

counterfactual path of savings that employees would have followed in the absence of the statutory retirement-age hike, after conditioning and fixed effects, after controlling for a rich set of observables capturing demographics, household composition, human capital and labor-market attachment. [1]

However, there might be unobserved and time-varying differences between the two groups, and that might lead to a different pathway pre-reform. Therefore, I formally tested the (conditional) parallel trend assumption in Section 6.3.1, and found no evidence of pre-treatment divergence, supporting the design.

4.2.1 Extensive Margin

Given the two-part structure of the savings outcome, I treat the extensive and intensive margins as distinct behavioral outcomes. (Chen & Roth, 2024; McKenzie, 2023) The decision to save and the decision on how much to save, conditional on saving, are conceptually different, potentially driven by different determinants. Therefore, I first estimate and interpret the treatment effects on these two margins separately without aggregating them into a single total effect.

First, a probit DiD model is used to analyze the extensive margin, that is, the probability of saving. Here I estimate the treatment effect on the decision to save at all.

I model the binary outcome $D_{it} = 1\{\text{if\_savings}_{it} = 1\}$ using a probit model:

$$\Pr(D_{it} = 1 \mid X_{it}) \;=\; \Phi\Big(\theta\, T_i \;+\; \delta_t \;+\; \beta_{\text{DID}}^{(E)}(T_i \times Post_t) \;+\; X_{it}'\gamma\Big)\,. \tag{4.1}$$

via a latent index:

$$D_{it}^{*} \;=\; \theta\, T_i \;+\; \delta_t \;+\; \beta_{\text{DID}}^{(E)}(T_i \times Post_t) \;+\; X_{it}'\gamma \;+\; \varepsilon_{it}, \qquad D_{it} = \mathbf{1}\{D_{it}^{*} > 0\}, \quad \varepsilon_{it} \sim \mathcal{N}(0,1). \tag{4.2}$$

where $\Phi(\cdot)$ is the standard normal CDF; T_i is the treatment group indicator, $Post_t$ denotes post-treatment periods, δ_t is the year fixed effects, and X_{it} is a vector of controls, which include gender, age, cohabiting partner, migration background, household size, number of children, full-/part-time employment, employment history, education, vocational training, panel-attrition control, and a constant. The variance $\mathrm{Var}(\varepsilon_i)$ is normalized to 1, so the coefficients are on latent scale. Therefore, I also report average marginal effects (AMEs) in probability units for interpretability with standard errors clustered at the household level.

4.2.2 Intensive Margin

For households that save in both pre- and post-reform periods ("always-savers"), I estimate a linear DiD model for the log amount saved:

$$\ln S_{it} = \alpha_i + \delta_t + \beta_{\text{DID}}^{(I)}\big(T_i \times Post_t\big) + X_{it}'\pi + \varepsilon_{it}, \quad i \in \mathcal{A}, \tag{4.3}$$

[1] Concretely, I condition on: age and age^2; sex; marital/partner status; migration background; household size and number of children; education level and vocational training; full-time indicator, years employed.

where $\mathcal{A} \equiv \{\, i : D_{it} = 1$ in both pre and post periods $\}$, T_i is the treatment indicator, $Post_t$ is a post-reform indicator, α_i and δ_t are household and year fixed effects (respectively), and X_{it} is a vector of controls (demographics, employment, and education). Standard errors are clustered at the household level.

Restricting to always-savers ensures the composition of savers remains stable across periods, isolating the within-saver treatment effect from the composition effect of those newly induced to save.

4.2.3 Full-Sample Estimator (for Robustness Check)

4.2.3.1 Tobit

Although the decision to save and the amount saved can be modeled separately, the observed zeros could also be interpreted as the result of latent negative savings—reflecting, for example, dissaving, debt accumulation, or financial reliance on others. To account for this censoring, I estimate a pooled Tobit model:

$$S_{it}^{*} = \alpha + \delta_t + \beta_{\text{DID}}\,(T_i \times Post_t) + X_{it}\theta + u_{it}, \quad u_{it} \sim N(0, \sigma^2),$$

$$S_{it} = \begin{cases} S_{it}^{*}, & S_{it}^{*} > 0, \\ 0, & S_{it}^{*} \le 0. \end{cases}$$

Here the observed savings S_{it} are left-censored at zero. This specification captures the idea that some households may want to save a negative amount but are constrained at zero. The Tobit model relies on strong assumptions, homoskedastic and normally distributed errors, both of which were examined in Chapter 6. Furthermore, incorporating household fixed effects in nonlinear models such as the Tobit is generally problematic due to the incidental parameters problem, so I restrict to pooled specifications.

4.2.3.2 Poisson Pseudo-Maximum-Likelihood (PPML)

As a more flexible alternative, I estimate a Poisson Pseudo-Maximum Likelihood (PPML) model, which accommodates both the presence of zeros and heteroskedasticity without imposing distributional assumptions on the error term. The conditional mean specification is:

$$E[S_{it} \mid \cdot] = \exp\big(\alpha + \delta_t + \beta_{\text{DID}}\,(T_i \times Post_t) + X_{it}\phi\big).$$

PPML does not require the dependent variable to be a count and is valid for nonnegative continuous outcomes. It provides consistent estimates of the parameters under correct specification of the conditional mean, even if the variance is misspecified. This robustness makes PPML particularly attractive for modeling saving amounts, which are highly skewed and zero-inflated.

Chapter 5

Results

5.1 Descriptive Statistics

My sample focuses on individuals of working age around the period of the 2007 pension reform, drawing on variables related to savings behavior, employment type, and demographic characteristics, using the SOEP core data (v.38) and SAVE data (year 2003-2010).

Table 5.1 and 5.2 presents descriptive statistics for the main variables used in the analysis, separately for employees and self-employed individuals in GSOEP and SAVE dataset respectively.

Overall, self-employed individuals report a higher average monthly savings amounts, and higher labor and household incomes compared to employees. The self-employed are also slightly older on average, less likely to be women and more likely to work full-time. Regarding educational attainment, self-employed individuals have slightly more years of schooling on average or higher education level, and a marginally lower share holding vocational degrees compared to employees.

These descriptive patterns highlight notable baseline differences between employees and self-employed individuals, underlining the importance of controlling for covariates and using appropriate empirical strategies. In particular, for the outcome variable, the self-employed group reports substantially higher average monthly amounts of savings compared to employees. While this difference does not necessarily violate the Difference-in-Differences identification strategy, it underscores the importance of carefully considering the parallel trends assumption in the subsequent analysis (Wooldridge, 2021).

Table 5.1: Descriptive Statistics by Employment Type (GSOEP)

Variable	**Self-employed**	**Employee**
Savings and Demographics		
Probability of saving	0.64	0.69
	(0.48)	(0.46)
Savings amount (EUR)	304.36	238.44
	(457.12)	(359.14)
Female	0.19	0.44
	(0.40)	(0.50)
Age	36.56	34.91
	(4.58)	(5.32)
Migration Background		
no migration background	0.85	0.83
	(0.36)	(0.38)
direct migration background	0.10	0.10
	(0.30)	(0.30)
indirect migration background	0.05	0.07
	(0.23)	(0.25)
Household and Labor		
Married	0.51	0.49
	(0.50)	(0.50)
Household size	2.75	2.58
	(1.39)	(1.25)
Number of children	0.95	0.82
	(1.08)	(1.00)
Full-time	0.96	0.83
	(0.18)	(0.38)
Years employed	14.30	12.21
	(6.20)	(5.67)
Education (years)	13.29	12.80
	(2.62)	(2.59)
Vocational degree	0.78	0.80
	(0.42)	(0.40)
Labor income (EUR)	2456.27	1570.14
	(1762.90)	(770.94)
Household income (EUR)	3139.39	2416.05
	(1638.23)	(1094.73)
No. of observations	667	18163

Table 5.2: Descriptive Statistics by Treatment Status (SAVE)

Variable	**Self-employed**	**Employee**
Savings Behavior		
Probability of saving	0.66	0.63
	(0.48)	(0.48)
Savings amount (EUR)	328.39	200.56
	(938.74)	(444.29)
Demographics		
Female	0.27	0.43
	(0.45)	(0.50)
Age	37.19	35.69
	(5.45)	(6.64)
Lives with partner	0.72	0.68
	(0.45)	(0.47)
East Germany	0.32	0.28
	(0.47)	(0.45)
German citizenship	0.99	0.98
	(0.10)	(0.14)
Number of children	1.05	1.24
	(1.45)	(1.15)
Household size	2.73	2.92
	(1.37)	(1.31)
Education		
Low	0.05	0.24
Lower-intermediate	0.19	0.28
Intermediate	0.08	0.18
Higher-intermediate	0.09	0.10
High	0.60	0.19
Employment History		
Currently unemployed	0.05	0.00
	(0.21)	(0.06)
No past unemployment	0.46	0.41
Past unemp. $<$ 6 months	0.31	0.37
Past unemp. 6m–2y	0.18	0.18
Past unemp. $>$ 2 years	0.05	0.04
No. of observations	111	1242

5.2 Extensive Margin

The extensive margin captures whether individuals choose to save. Using both GSOEP and SAVE dataset, I find a negative but statistically insignificant treatment effect, suggesting no clear evidence of behavioral change.

In the GSOEP Probit models (Table 5.3), interaction coefficients range from 0.009 to -0.058 with standard errors of 0.141–0.144. The corresponding average marginal effects (Table 5.4) lie between 0.003 and -0.020, i.e. roughly $+0.3$ to -2.0 percentage points, and are not statistically different from zero. Model fit improves as more controls and year fixed effects are added: the log-likelihood increases from -7646.06 to -6990.78.

Table 5.3: Estimated Treatment Effects on Savings: Summary of Probit DiD Models (GSOEP)

	GSOEP Probit			
	(1)	(2)	(3)	(4)
Treatment effect	0.009	-0.058	-0.018	-0.021
Standard error	(0.141)	(0.143)	(0.144)	(0.144)
Demographic controls	No	Yes	Yes	Yes
Employment & Education	No	No	Yes	Yes
Year fixed effects	No	No	No	Yes
Log-likelihood	-7646.06	-7476.19	-6999.18	-6990.78
Pseudo R-squared	0.000	0.022	0.074	0.076
Number of observations	12401	12397	12272	12272

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table 5.4: Average Marginal Effects of Treatment (GSOEP)

	(1)	(2)	(3)	(4)
AME of Treatment (interaction)	0.003	-0.020	-0.006	-0.007
Standard error	(0.049)	(0.049)	(0.047)	(0.047)
Number of observations	12401	12397	12272	12272

The SAVE Heckman–Probit results point in the same qualitative direction but are less precisely estimated, , reflecting the much smaller sample size. Point estimates for the treatment effect range from -0.328 to -0.373 with standard errors 0.274–0.286 (Table 5.5); the AMEs are -0.117 to -0.131 with standard errors around 0.10 (Table 5.6). While point estimates are negative, the confidence intervals are wide and include zero.

The estimated selection correlation $\rho < 0$ indicates that unobserved factors increasing sample selection are associated with a lower propensity to save; the $\rho = -1.000$ boundary in column (1) flags a fragile selection correction when controls are sparse, but ρ stabilizes to around -0.7 as richer covariates are added. Importantly, the treatment estimates remain statistically insignificant throughout.

Table 5.5: Estimated Treatment Effects on Savings: Summary of Heckman Probit Models (SAVE)

	SAVE Heckman Probit			
	(1)	(2)	(3)	(4)
Treatment effect	-0.328	-0.321	-0.333	-0.373
Standard error	(0.274)	(0.280)	(0.282)	(0.286)
Demographic controls	Yes	Yes	Yes	Yes
Employment & Education	No	Yes	Yes	Yes
Year fixed effects	No	No	Yes	Yes
Attrition control	No	No	No	Yes
Log-likelihood	-1052.16	-1043.30	-1040.13	-1039.01
Wald chi^2	30.52	43.92	48.61	50.06
ρ	-1.000	-0.735	-0.721	-0.714
Number of observations	1204	1204	1204	1204

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table 5.6: Average Marginal Effects of Treatment (SAVE)

	(1)	(2)	(3)	(4)
AME of Treatment (interaction)	-0.117	-0.113	-0.117	-0.131
Standard error	(0.098)	(0.099)	(0.099)	(0.100)
Number of observations	1204	1204	1204	1204

Taken together, these findings imply that the reform did not alter *whether* households save. A natural interpretation is that participation is sticky in the short run—due to fixed costs of setting up/maintaining saving arrangements or habit formation. Behavioral adjustments might occur primarily on the intensive margin, which is what I examine next.

5.3 Intensive Margin

This section assesses the reform's effect on *how much* established savers set aside, restricting to individuals who report saving in every observed wave ("always-savers"). Across GSOEP specifications, the estimated treatment effect is consistently negative, indicating that employees saved less relative to self-employed individuals after the reform.

In the OLS DiD models with the log of the amount saved as the dependent variable (Table 5.7), the interaction coefficient ranges from about -0.25 to -0.38 log points. The effect is statistically significant in columns (2)–(4) after rich controls and year fixed effects are included, but not in the baseline (column (1)) or when household fixed effects

are added (column (5)). Translating logs to percentages of savings, the results implies a roughly 22.2% to 31.5% decline.[1]

The loss of statistical significance in column (5) after adding household fixed effects suggests that part of the cross-sectional difference is explained by time-invariant household heterogeneity; at the same time, the point estimate remains negative and sizable (≈ -0.28), indicating that the sign and economic magnitude are robust even as precision drops when identification relies on within-household variation.

For the SAVE dataset, the sample of always-savers is too small—particularly in the control group, which includes fewer than 20 observations per year. This extreme sparsity causes high sensitivity of outliers and low statistic power. Therefore, the results are not reported for reliability concern.

Table 5.7: Estimated Treatment Effects on Savings: Intensive Margin (GSOEP)

	GSOEP OLS (Log Amount Saved)				
	(1)	(2)	(3)	(4)	(5)
Treatment effect	-0.251	-0.378**	-0.375*	-0.373*	-0.280
Standard error	(0.171)	(0.188)	(0.200)	(0.201)	(0.198)
Demographic controls	No	Yes	Yes	Yes	Yes
Employment & Education	No	No	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes
Household fixed effect	No	No	No	No	Yes
R-squared	0.010	0.079	0.173	0.176	0.048
Number of observations	4808	4806	4767	4767	4767

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

I also examine heterogeneity by gender via a pooled DDD model (details see Appendix 1). I find no gender gap on the extensive margin, while on the intensive margin, men reduce amounts more than women (marginal at 10%); see Appendix Table 1.

5.4 Discussion of Key Findings

First, on the *extensive margin*—the decision to save at all—the estimated treatment effects are small and statistically insignificant across Probit (GSOEP) and Heckman-Probit (SAVE) specifications, as are the corresponding average marginal effects. In other words, the reform did not measurably change whether households participated in saving. A plausible interpretation is that participation is sticky in the short run due to fixed costs and habits (e.g., standing orders, account set-up, contractual saving product), so most adjustment does not occur via entry/exit.

[1] Percent change computed as $\exp(\hat{\beta}) - 1$: $\exp(-0.251) - 1 \approx -0.222$, $\exp(-0.378) - 1 \approx -0.315$, $\exp(-0.280) - 1 \approx -0.244$.

Second, on the *intensive margin* among always-savers, the difference-in-differences estimates for the log amount saved are negative and economically meaningful. Across specifications, the interaction coefficient lies between about -0.25 and -0.38 log points, and is statistically significant in the richer models. Translating logs to percentages, these magnitudes imply roughly a 22–32% decline in monthly saving for treated households—an economically large change.[2] Using the employee's mean of saving which is about €238 per month as a reference, the estimates corresponds to roughly €52-€76 less saved per month.

Taken together, these findings suggest that the 2007 SRA increase primarily operated on *how much* established savers set aside, rather than on whether households saved at all. This pattern is consistent with an expectations-driven adjustment: if treated individuals anticipate a later retirement start while holding perceived longevity roughly fixed in the short run, the expected duration of retirement falls, lowering the perceived need for current saving flows. While the detailed mechanism analysis is developed in Chapter 7 Mechanisms, this interpretation aligns with the sign and magnitude of the main estimates here.

Restricting the intensive-margin sample to always-savers ensures that the within-saver response is not confounded by composition changes in who saves across periods. The absence of an extensive-margin effect reinforces this reading: the reform appears to have negligible effect on entry or exit from saving, but a sizable reduction of saving *levels* among those who were already participating.

[2]Percent change is computed as $\exp(\hat{\beta}) - 1$. For example, $\exp(-0.378) - 1 \approx -0.315$.

Chapter 6

Robustness Check

6.1 Full-Sample Estimator

6.1.1 Tobit

For the Tobit model, the coefficients on the interaction term are both negative but statistically insignificant for saving amount. However, for saving rate, the treatment effect is negative and significant on 10% level. (See Table 6.1 and 6.2)

Table 6.1: Estimated Treatment Effects on Savings Amount: Summary of Tobit Models

	GSOEP Tobit – Savings Amount			
Treatment effect	-78.937	-100.512	-91.625	-91.761
Standard error	(71.623)	(72.170)	(70.388)	(75.885)
Demographic controls	No	Yes	Yes	Yes
Employment & Education	No	No	Yes	Yes
Year fixed effects	No	No	No	Yes
Log-likelihood	-65695.90	-65440.24	-64108.31	-64102.46
Pseudo R-squared	0.000	0.004	0.015	0.015
Number of observations	12037	12035	11911	11911

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table 6.2: Estimated Treatment Effects on Saving Rate: Summary of Tobit Models

	GSOEP Tobit – Saving Rate			
Treatment effect	-0.033*	-0.035*	-0.032*	-0.032*
Standard error	(0.020)	(0.020)	(0.019)	(0.019)
Demographic controls	No	Yes	Yes	Yes
Employment & Education	No	No	Yes	Yes
Year fixed effects	No	No	No	Yes
Log-likelihood	1374.77	1652.80	2158.14	2164.85
Pseudo R-squared	-0.001	-0.206	-0.560	-0.565
Number of observations	11691	11689	11566	11566

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table 6.3 reports both unconditional and conditional average marginal effects from Tobit models for savings amount and saving rate. The unconditional effect reflects the impact of the pension reform on savings outcomes across the full sample, including non-savers, while the conditional effect captures the marginal effect among households with positive savings.

All estimates are negative, while the average marginal effect of on the saving rate is statistically significant at the 10% level. This suggests that, following the reform, employees (treated group) experienced a modest reduction in their saving rate compared to the self-employed (control group). The estimated average marginal effect is approximately -3.2% for the full sample, and -2.1% among savers.

Table 6.3: Tobit Model: Marginal Effects of the Pension Reform (Interaction Term)

	Savings Amount	**Saving Rate**
Unconditional AME	-117.49 (97.16)	-0.032* (0.025)
Conditional AME ($>$ 0)	– 71.66 (59.26)	-0.021* (0.013)

6.1.2 PPML

Table 6.4 and Table 6.5 report the estimated treatment effects of the 2007 pension reform on the amount saved and the saving rate, respectively, using Poisson Pseudo Maximum Likelihood (PPML) models. Across all specifications, the estimated coefficients on the interaction term (the treatment effect) are consistently negative, statistically significant in Table 6.5, suggesting that employees—who were directly affected by the increase in statutory retirement age—saved less than self-employed individuals following the reform.

For the savings amount (Table 6.4), the treatment effect estimates range from 0.182 to 0.070. While none of the specifications yield statistically significant results, the negative coefficients persist across models with increasing sets of controls.

Table 6.4: Estimated Treatment Effects on Savings Amount: Summary of PPML Models

	(1)	(2)	(3)	(4)	(5)
Treatment effect (interaction)	-0.182	-0.245	-0.201	-0.199	-0.070
Standard error	(0.172)	(0.174)	(0.175)	(0.175)	(0.173)
Demographic controls	No	Yes	Yes	Yes	Yes
Employment & Education	No	No	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes
Household fixed effects	No	No	No	No	Yes
Pseudo R-squared	0.002	0.051	0.182	0.183	0.670
Number of observations	12037	12035	11911	11911	10769

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The results for the saving rate (Table 6.5) exhibit a similar pattern. The treatment effect is statistically significant and negative in the baseline and intermediate models

(columns 1 to 4). These estimates imply that the reform was associated with a 25–28% decrease in the saving rate for treated individuals, relative to the control group. However, With household fixed effects in the TWFE specification (column 5), the estimate attenuates to 0.114 and is no longer statistically significant, suggesting earlier estimates partly reflected unobserved household-level heterogeneity. This pattern aligns with the insignificance of our main intensive-margin results.

Table 6.5: Estimated Treatment Effects on Saving Rate: Summary of PPML Models

	(1)	(2)	(3)	(4)	(5)
Treatment effect (interaction)	-0.276*	-0.292*	-0.260*	-0.259*	-0.114
Standard error	(0.156)	(0.156)	(0.154)	(0.154)	(0.126)
Demographic controls	No	Yes	Yes	Yes	Yes
Employment & Education	No	No	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes
Household fixed effects	No	No	No	No	Yes
Pseudo R-squared	0.000	0.011	0.026	0.026	0.111
Number of observations	11691	11689	11566	11566	10467

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Overall, the PPML estimates provide consistent directional evidence with the main results that the 2007 pension reform may have discouraged saving behavior among affected employees.

6.2 Alternative specifications of the model

6.2.1 Alternative Outcome Variable

As a robustness check for the intensive margin, Table 6.6 reports results using the saving rate as the dependent variable instead of the absolute amount saved. Across specifications, the treatment effect remains consistently negative and statistically significant at the 10% or 5% level.

These findings are consistent with the main results based on the savings amount (see Table 5.3), where the treatment effect was also negative and statistically significant across comparable specifications (ranging from 0.251 to 0.373). The consistency in sign and significance across both measures of saving behavior reinforces the conclusion that the 2007 pension reform reduced savings among employees relative to self-employed individuals.

Table 6.6: Robustness Check: Estimated Treatment Effects on Saving Rate

	(1)	(2)	(3)	(4)	(5)
Treatment effect (interaction)	-0.051**	-0.048**	-0.048*	-0.048*	-0.046*
Standard error	(0.026)	(0.024)	(0.025)	(0.025)	(0.027)
Demographic controls	No	Yes	Yes	Yes	Yes
Employment & Education	No	No	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes
Household fixed effects	No	No	No	No	Yes
R-squared	0.004	0.048	0.084	0.088	0.028
Number of observations	4716	4714	4675	4675	4675

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.2.2 Compositional Selection Bias

As a robustness check, I re-estimate the models using the full sample of households, without restricting it to those observed both before and after the reform. The main analysis was based on a balanced panel to control for potential compositional changes in the sample across time. However, this restriction may introduce selection bias if the decision to remain in the panel (i.e., to continue participating in the survey) is correlated with the household's saving behavior — for example, if savers are more likely to remain in the panel than non-savers.

By using the full sample in this robustness check, I relax that restriction and assume that sample entry and exit are random with respect to treatment status and saving decisions. For the extensive margin, the results are reported in Tables 6.7 and 6.8. The treatment effect remains consistently negative across specifications, with coefficients ranging from 0.096 to 0.148. While the estimates are not statistically significant, the direction and magnitude are broadly consistent with the main results. The average marginal effects (AMEs) also support this pattern.

Table 6.7: Robustness Check: Estimated Treatment Effects on Savings (Extensive Margin)

	GSOEP Probit – Full Sample (Composition)			
	(1)	(2)	(3)	(4)
Treatment effect	-0.096	-0.162	-0.143	-0.148
Standard error	(0.110)	(0.111)	(0.115)	(0.115)
Demographic controls	No	Yes	Yes	Yes
Employment & Education	No	No	Yes	Yes
Year fixed effects	No	No	No	Yes
Log-likelihood	-12344.93	-12040.01	-10803.96	-10794.07
Pseudo R-squared	0.000	0.024	0.077	0.077
Number of observations	19321	19310	18379	18379

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6.8: Average Marginal Effects of Treatment (Extensive Margin)

	(1)	(2)	(3)	(4)
AME of Treatment (interaction)	-0.035	-0.057	-0.048	-0.050
Standard error	(0.040)	(0.039)	(0.038)	(0.038)
Number of observations	19321	19310	18379	18379

For the intensive margin, the results, reported in Tables 6.9 and 6.10, show that the estimated treatment effects remain negative across all specifications, consistent with the main findings. While statistical significance is somewhat weaker, the direction and magnitude of the effects remain stable, particularly in the two-way fixed effects models (column 5).

These findings suggest that the main results are not driven by selection into the balanced sample, and lend further credibility to the conclusion that the 2007 pension reform reduced savings among treated households.

Table 6.9: Robustness Check: Estimated Treatment Effects on Savings Amount

	(1)	(2)	(3)	(4)	(5)
Treatment effect (interaction)	-0.041	-0.144	-0.238	-0.255	-0.238
Standard error	(0.168)	(0.171)	(0.183)	(0.182)	(0.187)
Demographic controls	No	Yes	Yes	Yes	Yes
Employment & Education	No	No	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes
Household fixed effects	No	No	No	No	Yes
R-squared	0.008	0.084	0.178	0.182	0.038
Number of observations	8227	8221	7782	7782	7782

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6.10: Robustness Check: Estimated Treatment Effects on Saving Rate

	(1)	(2)	(3)	(4)	(5)
Treatment effect (interaction)	-0.007	-0.013	-0.024	-0.025	-0.045*
Standard error	(0.025)	(0.023)	(0.024)	(0.024)	(0.025)
Demographic controls	No	Yes	Yes	Yes	Yes
Employment & Education	No	No	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes
Household fixed effects	No	No	No	No	Yes
R-squared	0.007	0.051	0.088	0.090	0.022
Number of observations	8041	8036	7616	7616	7616

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.2.3 Extended Time Window

To test the robustness of the results to the choice of time window, I extend the sample period of the GSOEP dataset from 2003–2010 (used in the main analysis) to 1992–2022.

As shown in Table 6.11 and Table 6.12, the estimated treatment effects for the extensive margin using Probit models become positive in the extended window, although statistically insignificant across all specifications. The average marginal effects similarly suggest a modest increase in the probability of saving among treated individuals, but these effects also lack statistical significance. These findings contrast with the negative but insignificant estimates found in the main analysis using the 2003–2010 sample. Taken together, the results imply that the impact of the reform on the extensive margin is ambiguous, with no clear evidence of either a short-term or long-term effect on the probability of saving.

Table 6.11: Robustness Check: Estimated Treatment Effects on Savings (Probit DiD Models, Extended Window)

	GSOEP Probit – 1992–2022 Sample			
	(1)	(2)	(3)	(4)
Treatment effect	0.048	0.082	0.101	0.083
Standard error	(0.169)	(0.174)	(0.175)	(0.174)
Demographic controls	No	Yes	Yes	Yes
Employment & Education	No	No	Yes	Yes
Year fixed effects	No	No	No	Yes
Log-likelihood	-16997.54	-16564.86	-15647.22	-15551.43
Pseudo R-squared	0.002	0.020	0.061	0.067
Number of observations	28881	28669	28301	28301

Table 6.12: Average Marginal Effects of Treatment (Probit Models, Extended Window)

	(1)	(2)	(3)	(4)
AME of Treatment (interaction)	0.016	0.027	0.031	0.026
Standard error	(0.056)	(0.057)	(0.055)	(0.054)
Number of observations	28881	28669	28301	28301

As shown in Table 6.13, the treatment effect estimates remain negative across all specifications, consistent with the main results. However, the coefficients are smaller in magnitude and not statistically significant, with estimates ranging from 0.097 to 0.107. This attenuation may reflect greater variability in household saving behavior over a 30-year window or the influence of overlapping reforms and macroeconomic events,or a diminishing long-term effect of the 2007 reform. In contrast to the sharper and statistically significant negative impact observed in the shorter 2003–2010 window, the extended period suggests that the initial response may have been acute but did not persist over time.

Table 6.13: Robustness Check: Estimated Treatment Effects on Log Amount Saved (Extended Window)

	(1)	(2)	(3)	(4)
Treatment effect (interaction)	-0.097	-0.154	-0.129	-0.107
Standard error	(0.243)	(0.287)	(0.287)	(0.281)
Demographic controls	No	Yes	Yes	Yes
Employment & Education	No	No	Yes	Yes
Year fixed effects	No	No	No	Yes
R-squared	0.030	0.107	0.210	0.216
Number of observations	7104	7074	6986	6986

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.3 Model Assumptions and Diagnostic Checks

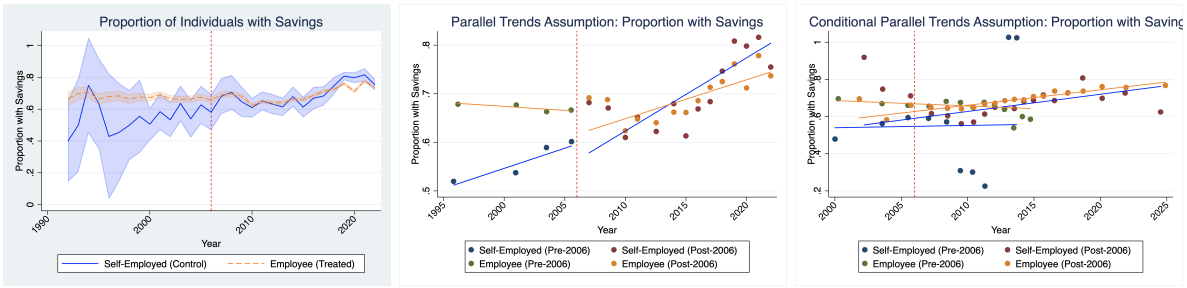
To further assess the validity of the estimated treatment effects, I examine key assumptions underlying the empirical models used. This includes testing for parallel trends in the Difference-in-Differences framework, and checking for heteroskedasticity and normality in nonlinear specifications. These diagnostic checks help to ensure that the reported results are not driven by violations of core modeling assumptions.

6.3.1 Parallel trends assumption

A key identifying assumption of the Difference-in-Differences framework is that, absent the reform, the treatment and control groups would have followed parallel trends in saving behavior. To assess the the parallel trends assumption underlying the Difference-in-Differences (DiD) approach, three outcome variables are tested: a binary indicator of whether the household saves, the log of monthly savings amount, and the saving rate. For each outcome, three panels are shown: raw trends with confidence intervals, a comparison of linear trends before and after the reform, and conditional trends that account for observed covariates.

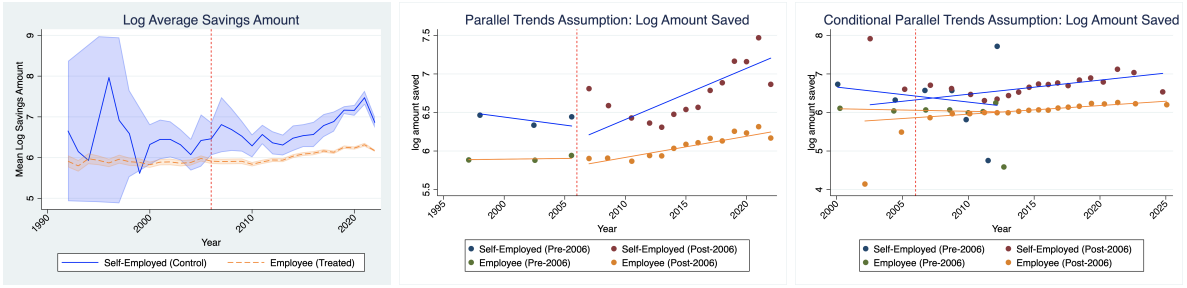
As shown in figure 6.1, the trends in the proportion of savers appear relatively stable and similar between the treatment and control groups prior to the reform, particularly after controlling for observable charasteristics. Similar patterns are observed for the logarithm of amount of savings and saving rate shown in 6.2 and 6.3, where pre-reform trajectories are largely parallel across groups.

While visual inspection of the pre-reform trends suggests broadly parallel movements between the treatment and control groups for the binary saving indicator and the logarithm of monthly savings, the patterns are less clear for the saving rate. To further assess the plausibility of the parallel trends assumption, I estimate pre-reform regressions that interact the treatment indicator with year fixed effects and jointly test whether the interaction terms are significantly different from zero. For the binary saving indicator, the test yields $\chi^2(14) = 8.56$ (p = 0.86); for the log of savings amount, F(14, 12,615) = 0.54 (p = 0.91); and for the saving rate, F(14, 12,391) = 0.91 (p = 0.54). In



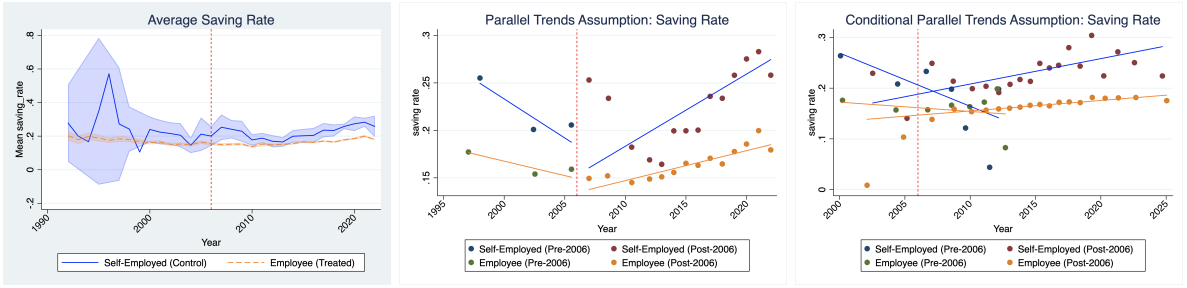
(a) (b) (c)

Figure 6.1: Parallel-trend checks for proportion of savers.



(a) (b) (c)

Figure 6.2: Parallel-trend checks for log savings.



(a) (b) (c)

Figure 6.3: Parallel-trend checks for saving rate.

all three cases, the tests fail to reject the null hypothesis of no differential pre-trends, providing statistical support for the common trend assumption.

6.3.2 Homoskedasticity and Normality

To examine the validity of the homoskedasticity assumption in the Probit model, I estimate a heteroskedastic Probit specification using the `hetprobit` command in Stata. The variance of the error term is modeled as a function of several covariates, including *female*, *age*, *household size*, *number of children*, *full-time employment*, and *years employed*, while clustering standard errors at the household level. I then conduct a Wald test of the joint significance of these variables in the variance equation. The test fails to reject the null hypothesis of homoskedasticity, with a test statistic of $\chi^2(6) = 4.84$ and a p-value of 0.56. This indicates that there is no statistically significant evidence of heteroskedasticity, and the assumption of constant variance in the Probit model is not rejected.

The probit model also relies on the assumption that the latent error term follows a normal distribution. To assess the validity of this assumption, I implement a score test for probit model proposed by Davidson. The procedure involves augmenting the probit model with squared and cubic terms of the linear index function $X\hat{\beta}$. Under the null hypothesis of normality, these higher-order terms should not be jointly significant (Davidson & MacKinnon, 1984; Silvapulle & Silvapulle, 1995).

I estimate the augmented probit model and conduct a joint test of the coefficients on the added terms. The test yields a test statistic of $\chi^2(2) = 2.97$ with a corresponding p-value of 0.227. This result implies that I fail to reject the null hypothesis of normality, suggesting that the standard probit model is appropriately specified and the normality assumption does not appear to be violated in our setting.

Chapter 7

Mechanisms

7.1 Expected Retirement Age

This section examines how the 2007 German pension reform affected individuals' expectations regarding their retirement age. Using the SAVE dataset, which contains detailed questions about retirement expectations, I estimate the impact of the reform on the expected retirement age.

The results indicate a negative and statistically significant treatment effect on the expected retirement age. This suggests that the reform was effective in shifting individuals' retirement expectations, consistent with the policy's objective of encouraging delayed retirement.

Table 7.1: Estimated Treatment Effect on Expected Retirement Age (SAVE)

	(1) Heckman	(2) Heckman	(3) Heckman	(4) OLS
Treatment Effect	0.216	1.718**	1.706**	1.202
	(0.470)	(0.832)	(0.831)	(0.837)
Controls	Yes	Yes	Yes	Yes
Wald χ^2	176	184	208	/
p-value	0.000	0.000	0.000	/
(pseudo) R-squared	/	/	/	0.0576
Sample size	2,933	2,933	2,933	2,776

Notes: Standard errors in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All models include socio-demographic controls.

Column (1) uses a baseline post-2007 interaction; columns (2)–(4) use 2006 as a reference year.

7.2 Expected Years in Retirement

The SAVE dataset also provides individual-level information on subjective life expectancy. I combine this with expected retirement age to compute the expected years in retirement for each respondent.

Understanding this measure is crucial, as it directly relates to saving behavior: the longer individuals expect to spend in retirement, the greater their incentive to save. The analysis reveals a negative and statistically significant treatment effect on expected years in retirement. This implies that, following the reform, individuals—particularly the

treated group—did not significantly revise their expectations about life expectancy but did anticipate retiring later. As a result, their expected retirement duration shrank.

This reduction in expected retirement years may partially explain the observed decline in savings. If individuals believe they will spend fewer years in retirement, they may feel less need to accumulate savings.

Table 7.2: Estimated Treatment Effect on Expected Years in Retirement (SAVE)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment Effect	3.219*	0.582	-2.372	-2.535	-2.151	-2.934	-3.527*
	(1.841)	(1.731)	(1.834)	(1.778)	(1.747)	(1.818)	(1.893)
Controls							
Demographics	✗	✓	✓	✓	✓	✓	✓
Education & Employment	✗	✗	✓	✓	✓	✓	✓
Wealth & Income	✗	✗	✗	✓	✓	✓	✓
Expectation Variables	✗	✗	✗	✗	✓	✓	✓
Attrition Controls	✗	✗	✗	✗	✗	✓	✓
Year Fixed Effects	✗	✗	✗	✗	✗	✗	✓
Observations	3,302	3,302	3,302	3,302	3,302	3,302	3,302
Wald χ^2	64	126	145	163	237	243	253
p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Notes: Standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. ✓ = included; ✗ = not included.

Subgroup analysis by gender reveals heterogeneity: the effect is predominantly driven by women. While the female subgroup shows a statistically significant decline in expected retirement years, the male subgroup does not display a significant change. This gender-specific response suggests that women may be more responsive—or vulnerable—to the reform in terms of adjusting retirement expectations.

Table 7.3: Treatment Effect on Expected Years in Retirement – Female Subsample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment Effect	2.065	1.756	-4.125	-4.277	-3.661	-4.227	-5.449*
	(2.493)	(2.467)	(2.627)	(2.604)	(2.531)	(2.701)	(2.957)
Controls							
Demographics	✗	✓	✓	✓	✓	✓	✓
Education & Employment	✗	✗	✓	✓	✓	✓	✓
Wealth & Income	✗	✗	✗	✓	✓	✓	✓
Expectation Variables	✗	✗	✗	✗	✓	✓	✓
Attrition Controls	✗	✗	✗	✗	✗	✓	✓
Year Fixed Effects	✗	✗	✗	✗	✗	✗	✓
Observations	1,607	1,607	1,607	1,607	1,607	1,607	1,607
Wald χ^2	48	50	64	78	110	111	119
p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Notes: Robust standard errors in parentheses. ✓ = included; ✗ = not included.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7.4: Treatment Effect on Expected Years in Retirement – Male Subsample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment Effect	-0.555	-1.770	-2.009	-2.365	-1.392	-2.692	-1.942
	(2.476)	(2.427)	(2.517)	(2.347)	(2.239)	(2.406)	(2.339)
Controls							
Demographics	✗	✓	✓	✓	✓	✓	✓
Education & Employment	✗	✗	✓	✓	✓	✓	✓
Wealth & Income	✗	✗	✗	✓	✓	✓	✓
Expectation Variables	✗	✗	✗	✗	✓	✓	✓
Attrition Controls	✗	✗	✗	✗	✗	✓	✓
Year Fixed Effects	✗	✗	✗	✗	✗	✗	✓
Observations	1,695	1,695	1,695	1,695	1,695	1,695	1,695
Wald χ^2	27	42	61	77	127	127	145
p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Notes: Robust standard errors in parentheses. ✓ = included; ✗ = not included.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

7.3 Interpretation of Results

The main findings suggest a robust and negative treatment effect on household savings following the 2007 pension reform. One plausible mechanism behind this behavioral response lies in individuals' expectations about retirement. The reform raised the statutory retirement age, which many treated individuals incorporated into their expectations. However, life expectancy perceptions remained largely unchanged. As a result, expected years in retirement declined.

This decline may have reduced the perceived need to accumulate savings, as individuals anticipated a shorter retirement period. The results are consistent with the life-cycle hypothesis, which posits that individuals plan consumption and savings over their expected lifetime. Under the LCH framework, saving behavior is determined by the expected duration of retirement, expected lifetime income and the desire to smooth consumption over time.

In this case, if households continued to expect a similar lifespan but anticipated working for more years, the number of years in retirement effectively declined. This reduced the financial burden of sustaining consumption post-retirement, leading to lower saving needs during the working period. Thus, the decline in savings reflects a rational adjustment to a changing intertemporal consumption-savings trade-off. This mechanism reinforces the explanatory power of the life cycle hypothesis in understanding saving responses to pension reforms.

7.4 Alternative Channels and Competing Explanations

While the preceding evidence highlights an expectations–based mechanism, several additional channels could jointly produce the observed decline in liquid saving and in expected years spent in retirement. For clarity, Figure 7.1 schematically summarizes these pathways.

Figure 7.1: Proposed Mechanism and Alternative Channels

A first possibility is endogenous retirement timing. Although the statutory retirement age (SRA) increased, individuals retain discretion to retire earlier or later than the legal threshold. Forward-looking households may therefore adjust saving to implement their preferred timing: some may increase saving to finance an early exit despite a higher SRA, whereas others who plan to work longer may rationally reduce saving because a prolonged career lowers the need for retirement resources. This heterogeneity in chosen retirement timing can offset, reinforce, or even reverse the average expectation channel documented above.

Second, baseline differences between employees and the self-employed may shape both the timing and magnitude of responses. In our data, the self-employed are, on average, older, more educated, and have higher labor and household incomes than employees. These compositional differences plausibly affect information acquisition, salience, and planning horizons. If the self-employed anticipated the reform earlier–for example, during the discussion stage of the reform, adjustments in expectations and

saving could occur sooner or more smoothly in the control group than the treatment group–for example, react at the announcement of the reform.

Third, part of the decline in measured saving may reflect displacement across asset "pillars" rather than a reduction in overall retirement preparation. Households might reallocate from liquid saving to tax-advantaged or employer-sponsored retirement products (e.g., Riester or occupational pensions) or toward housing investment (amortization and home purchase), which serve as retirement wealth in Germany. Such reallocation would cause recorded liquid saving to fall even if lifetime retirement wealth might not.

Taken together, these considerations suggest that the main mechanism—shorter expected retirement horizons lowering desired saving—should be interpreted alongside endogenous retirement timing, compositional differences between treatment and control groups, and potential substitution across saving vehicles.

Chapter 8

Discussion and Conclusion

8.1 Summary of findings

This thesis set out to investigate how the 2007 German pension reform—in which the statutory retirement age (SRA) was gradually raised from 65 to 67—impacted household saving behaviour. Exploiting a Difference-in-Differences design that compares employees (treated) to self-employed individuals (control), and drawing on two rich panel datasets (SOEP and SAVE) covering the period immediately before and after the reform, the main results can be summarised as follows:

For the extensive margin, across Probit and Heckman-Probit specifications, the treatment effect on the likelihood of saving is negative but small and never statistically significant. This suggests the reform did not meaningfully alter whether households chose to save.

For the intensive margin, focusing on the subsample of ”always-savers”, OLS DiD estimates indicate that employees exposed to the reform reduced their logged savings amounts by approximately 37% relative to the self-employed. These effects remain negative and statistically significant after robustness checks with alternative outcome variable (savings rate) and with larger sample without controlling for compositional changes.

In the full sample, Tobit model estimates point to an average decline of around 3.2% in the savings rate. Complementary estimates from the Poisson Pseudo-Maximum Likelihood (PPML) model corroborate this negative treatment effect, indicating a reduction in saving rate of approximately 28%.

I proposed a mechanisms via retirement expectations. Analysis of the SAVE survey's forward-looking modules shows that the reform successfully shifted expected retirement ages upward by about 1–1.7 years. Because subjective life expectancy remained unchanged, the result was a reduction in expected years spent in retirement (by roughly 2–3.5 years). This contraction in anticipated retirement duration provides a direct channel: under the Life-Cycle Hypothesis, shorter expected retirement reduces the need to accumulate assets, thus rationalizing the observed drop in savings amounts.

Subgroup analysis further suggests that the decline in expected retirement years—and hence the indirect effect on saving—was more pronounced among women, pointing to potential gender-specific vulnerabilities in response to pension policy changes.

Taken together, these findings paint a coherent picture: the reform led savers to adjust the scale of their saving downward by recalibrating their life-cycle plans to a shorter retirement horizon.

8.2 Contribution to existing literature

Whereas much prior work has focused on labour supply and retirement timing, this thesis offers one of the first causal estimates of how raising the statutory retirement age influences private saving behaviour in Germany, thus filling a key gap in the empirical literature.

By combining traditional DiD methods with detailed expectation modules in the SAVE dataset, the analysis identifies a clear psychological and economic channel—via anticipated retirement duration—bridging policy change and saving outcomes. The evidence on gender differences underscores that reform effects are not uniform across socio-demographic groups, highlighting the importance of considering distributional impacts when designing pension policy.

Methodologically, the use of complementary datasets (SOEP for broad macro-representativeness; SAVE for rich forward-looking and more saving-specific variables) and multiple econometric approaches (linear and nonlinear DiD, Tobit, PPML, selection correction) strengthens the credibility and generality of the results.

8.3 Policy Implications

Policymakers should not assume that later retirement automatically boosts household wealth at retirement; indeed, people may recalibrate their saving targets downward, leaving some households under-prepared. To mitigate such risks, reforms that delay retirement eligibility should be complemented by measures that actively encourage private saving—such as tax-advantaged pension schemes (e.g., the Riester Plan).

The pronounced effect among women suggests that the reform might be gender-sensitive and may have unintended distributional consequences in other subgroups and could exacerbate existing inequalities in retirement security. Policymakers should therefore consider tailored interventions—such as targeted financial literacy programs or enhanced access to supplementary pension schemes—aimed at vulnerable groups, including women and other socioeconomically disadvantaged populations, to ensure that the long-term benefits of pension sustainability.

8.4 Limitations of the Study

Several limitations should be acknowledged. First, the analysis relies on self-reported survey data from the SAVE and GSOEP dataset, which may be subject to recall bias or measurement error, particularly in variables related to expected retirement age, life expectancy, and savings behavior. These subjective measures, while valuable, may not perfectly reflect actual behavior or expectations.

Second, while the SAVE dataset offers rich information on savings and expectations, it has a smaller sample size compared to the GSOEP dataset, which may limit the statistical power and reliability, particularly in subgroup analyses.

Third, unobserved confounding factors and overlapping pension reforms and major macroeconomic event-such as the Riester incentives and the 2008 financial crisis-may

have influenced the saving behaviours differently, making it harder to isolate the effect of the 2007 pension reform

Fourth, the heterogenous treatment effects are not fully explored. Life-cycle considerations suggest differential responses by distance to retirement. And the treatment effects may vary across not only gender, but also education, income quintiles, regions, and occupations.

Fifth, the savings concept may not fully capture reallocation across pillars. Contributions to occupational or third-pillar schemes, mortgage amortisation, and housing investment are only partially observed; if households shift *where* they save rather than *how much*, measured flows could understate total retirement preparation (see Section 3.1 and 3.2.2).

Moreover, the analysis focuses on the short- to medium-term response to the reform. Longer-term behavioral adjustments may take time and are not fully captured in the available data. Finally, external validity is also limited: the results are embedded in Germany's institutional environment, and generalisation to other settings should be made with caution.

8.5 Suggestion for future research

Building on the findings of this thesis, several suggestions for future research are given. First, to assess the persistence and long-term trajectory of saving responses, future studies could exploit more recent waves of the SOEP or draw on administrative records that extend further beyond the reform period.

Second, given the predictions of the Life-Cycle Hypothesis, which posits heterogeneity in saving behaviour across different stages of the life course, further work should investigate how the reform's effects vary by age and proximity to retirement.

Third, more detailed heterogeneity analyses—beyond gender—are warranted to identify which vulnerable populations are most sensitive to pension reforms, thereby informing the design of more equitable policies. Lastly, cross-country comparative studies examining similar increases in statutory retirement age could test the external validity of these findings and help to generalize the observed behavioural responses.

In conclusion, this thesis demonstrates that pension reforms exert a nuanced influence on household finance: by altering retirement horizons, they reshape saving behaviour in ways that extend beyond labour-market participation. A careful balancing of statutory parameters and private-saving incentives will therefore be essential to ensure that later retirement translates into adequate financial well-being in old age.

Bibliography

Angrist, J. D., & Pischke, J.-S. (2008, December). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press. https://doi.org/10.2307/j.ctvcm4j72

Ardito, C. (2021). The unequal impact of raising the retirement age: Employment response and program substitution. *IZA Journal of Labor Economics*, *10*(1), 20210003. https://doi.org/10.2478/izajole-2021-0003

Attanasio, O. P., & Rohwedder, S. (2003). Pension Wealth and Household Saving: Evidence from Pension Reforms in the United Kingdom. *American Economic Review*, *93*(5), 1499–1521. https://doi.org/10.1257/000282803322655419

Barschkett, M., Geyer, J., Haan, P., & Hammerschmid, A. (2022). The effects of an increase in the retirement age on health — Evidence from administrative data. *The Journal of the Economics of Ageing*, *23*, 100403. https://doi.org/10.1016/j.jeoa.2022.100403

Bonin, H. (2009). 15 Years of Pension Reform in Germany: Old Successes and New Threats. *The Geneva Papers on Risk and Insurance - Issues and Practice*, *34*(4), 548–560. https://doi.org/10.1057/gpp.2009.20

Börsch-Supan, A., Coppola, M., Essig, L., Eymann, A., & Schunk, D. (2009). The German save study Design and Results.

Börsch-Supan, A., Coppola, M., & Reil-Held, A. (2012, April). *Riester Pensions in Germany: Design, Dynamics, Targetting Success and Crowding-In* (tech. rep. No. w18014). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w18014

Börsch-Supan, A. H., Reil-Held, A., & Schunk, D. (2007). The Savings Behaviour of German Households: First Experiences with State Promoted Private Pensions. *MEA Discussion Paper*, (No. 136-07).

Browning, M., & Crossley, T. F. (2001). The Life-Cycle Model of Consumption and Saving. *Journal of Economic Perspectives*, *15*(3), 3–22. https://doi.org/10.1257/jep.15.3.3

Bucher-Koenen, T., & Wilke, C. (2009). Zur Anhebung der Altersgrenzen. In *Sozialer Fortschritt* (pp. 69–79).

Bucher-Koenen, T., & Lusardi, A. (2011). Financial literacy and retirement planning in Germany. *Journal of Pension Economics and Finance*, *10*(4), 565–584. https://doi.org/10.1017/S1474747211000485

Buchholz, S., Rinklake, A., & Blossfeld, H.-P. (2013). Reversing Early Retirement in GermanyA Longitudinal Analysis of the Effects of Recent Pension Reforms on the Timing of the Transition to Retirement and on Pension Incomes. *Comparative Population Studies*, *38*(4). https://doi.org/10.12765/CPoS-2013-23

Chen, J., & Roth, J. (2024). Logs with Zeros? Some Problems and Solutions. *The Quarterly Journal of Economics*, *139*(2), 891–936. https://doi.org/10.1093/qje/qjad054

Cifuentes, R. (n.d.). How Does Pension Reform Affect Savings and Welfare.

Coppola, M., & Wilke, C. B. (2010). How Sensitive are Subjective Retirement Expectations to Increases in the Statutory Retirement Age? The German Case. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1701630

Cribb, J., Emmerson, C., & Tetlow, G. (2016). Signals matter? Large retirement responses to limited financial incentives. *Labour Economics*, *42*, 203–212. https://doi.org/10.1016/j.labeco.2016.09.005

Davidson, R., & MacKinnon, J. G. (1984). Convenient specification tests for logit and probit models. *Journal of Econometrics*, *25*(3), 241–262. https://doi.org/10.1016/0304-4076(84)90001-0

De Grip, A., Fouarge, D., & Montizaan, R. (2013). How Sensitive are Individual Retirement Expectations to Raising the Retirement Age? *De Economist*, *161*(3), 225–251. https://doi.org/10.1007/s10645-013-9205-5

Dolls, M., Doerrenberg, P., Peichl, A., & Stichnoth, H. (2018). Do retirement savings increase in response to information about retirement and expected pensions? *Journal of Public Economics*, *158*, 168–179. https://doi.org/10.1016/j.jpubeco.2017.12.014

Etgeton, S., Fischer, B., & Ye, H. (2023). The effect of increasing retirement age on households' savings and consumption expenditure. *Journal of Public Economics*, *221*, 104845. https://doi.org/10.1016/j.jpubeco.2023.104845

Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C., & Schupp, J. (2019). The German Socio-Economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik*, *239*(2), 345–360. https://doi.org/10.1515/jbnst-2018-0022

Gohl, N., Schrenker, A., & Steuernagel, A. (2023). Working Longer: Causal Effects on Career Trajectories of Raising the Statutory Retirement Age. *Mimeo*.

Hinrichs, K. (2021). Recent pension reforms in Europe: More challenges, new directions. An overview. *Social Policy & Administration*, *55*(3), 409–422. https://doi.org/10.1111/spol.12712

Holzmann, R., & Hinz, R. (2012). *Old Age Income Support in the 21st century: An International Perspective on Pension Systems and Reform*. World Bank.

Holzmann, R., Palmer, E. E., & World Bank (Eds.). (2006). *Pension reform: Issues and prospect for non-financial defined contribution (NDC) schemes*. World Bank.

Honekamp, I., & Schwarze, J. (2010). Pension reforms in Germany: Have they changed savings behaviour? *Pensions: An International Journal*, *15*(3), 214–225. https://doi.org/10.1057/pm.2010.20

Lachowska, M., & Myck, M. (2018). The Effect of Public Pension Wealth on Saving and Expenditure. *American Economic Journal: Economic Policy*, *10*(3), 284–308. https://doi.org/10.1257/pol.20150154

Lalive, R., Magesan, A., & Staubli, S. (2017). Raising the Full Retirement Age: Defaults vs Incentives. *NBER Working Paper orrc17–12*.

Lalive, R., & Staubli, S. (2015). How Does Raising Women's Full Retirement Age Affect Labor Supply, Income, and Mortality? *NBER working paper, 18660, 17*.

McKenzie, D. (2023). Interpreting treatment effects on an inverse hyperbolic sine outcome variable and alternatives. Retrieved November 26, 2024, from https://blogs.worldbank.org/en/impactevaluations/interpreting-treatment-effects-inverse-hyperbolic-sine-outcome-variable-and

Modigliani, F., & Brumberg, R. (2013). Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data 1. In *Post-keynesian economics* (pp. 388–436). Routledge.

Nivalainen, S. (2023). Retirement Intentions and Increase in Statutory Retirement Age: 2017 Pension Reform and Intended Retirement Age in Finland. *Journal of Aging & Social Policy*, 1–22. https://doi.org/10.1080/08959420.2023.2195787

OECD. (2006, February). *Live Longer, Work Longer*. https://doi.org/10.1787/9789264035881-en

Oehler, A., & Werner, C. (2008). Saving for Retirement—A Case for Financial Education in Germany and UK? An Economic Perspective. *Journal of Consumer Policy*, *31*(3), 253–283. https://doi.org/10.1007/s10603-008-9074-5

Okumura, T., & Usui, E. (2014). The effect of pension reform on pension-benefit expectations and savings decisions in Japan. *Applied Economics*, *46*(14), 1677–1691. https://doi.org/10.1080/00036846.2013.870654

Pilipiec, P., Groot, W., & Pavlova, M. (2021a). The causal influence of increasing the statutory retirement age on job satisfaction among older workers in the Netherlands. *Applied Economics*, *53*(13), 1498–1527. https://doi.org/10.1080/00036846.2020.1827136

Pilipiec, P., Groot, W., & Pavlova, M. (2021b). The Effect of an Increase of the Retirement Age on the Health, Well-Being, and Labor Force Participation of Older Workers: A Systematic Literature Review. *Journal of Population Ageing*, *14*(2), 271–315. https://doi.org/10.1007/s12062-020-09280-9

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, *235*(2), 2218–2244. https://doi.org/10.1016/j.jeconom.2023.03.008

Schunk, D. (2006). *The German SAVE Survey: Documentation and Methodology* (tech. rep.). Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy.

Seibold, A. (2021). Reference Points for Retirement Behavior: Evidence from German Pension Discontinuities. *American Economic Review*, *111*(4), 1126–1165. https://doi.org/10.1257/aer.20191136

Selin Kara & Zimmermann, S. (2023). *SOEPcompanion (v38)* (tech. rep. No. 1261). , Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.

Silvapulle, M. J., & Silvapulle, P. (1995). A Score Test against One-Sided Alternatives. *Journal of the American Statistical Association*, *90*(429), 342–349. https://doi.org/10.1080/01621459.1995.10476518

Soosaar, O., Puur, A., & Leppik, L. (2021). Does raising the pension age prolong working life? Evidence from pension age reform in Estonia. *Journal of Pension Economics and Finance*, *20*(2), 317–335. https://doi.org/10.1017/S1474747220000244

Statistisches Bundesamt (Destatis). (2024). Consumer price index: Germany, years. Retrieved November 26, 2024, from https://www-genesis.destatis.de/datenbank/online/statistic/61111/table/61111-0001

Staubli, S., & Zweimüller, J. (2013). Does raising the early retirement age increase employment of older workers? *Journal of Public Economics*, *108*, 17–32. https://doi.org/10.1016/j.jpubeco.2013.09.003

United Nations, DESA, Population Division. (2024). *World Population Prospects 2024* (tech. rep.).

Wilke, C. (2009, January). *German Pension Reform*. Peter Lang D. https://doi.org/10.3726/b13605

Wooldridge, J. M. (2021). Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3906345

List of Figures

List of Tables

Appendix

A.1 Heterogeneity Analysis by Gender

To assess whether treatment effects differ by gender, I estimate pooled specifications that fully interact the post indicator, treatment status, and a female dummy. The triple interaction $Post_t \times Treated_i \times Female_i$ is a difference-in-difference-in-differences (DDD): it equals the DiD for women minus the DiD for men.

Table 1: Heterogeneity by Gender: Extensive and Intensive Margins

	Panel A. Extensive margin (Participation)	
	SAVE	GSOEP
$Post \times Treated \times Female$	-0.493	-0.166
	(0.599)	(0.517)
DiD (men, index)	-0.123	-0.001
	(0.327)	(0.172)
DiD (women, index)	-0.615	-0.167
	(0.507)	(0.487)
AME (men)	-0.044	-0.000
	(0.121)	(0.061)
AME (women)	-0.244	-0.057
	(0.200)	(0.163)
AME (women$-$men)	-0.200	-0.057
	(0.233)	(0.174)
Log-likelihood	-1168.56	-6997.22
Pseudo R^2	/	0.075
Observations	1353	12272

Panel B. Intensive margin (Log amount), GSOEP	
$Post \times Treated \times Female$	0.554*
	(0.323)
DiD (men, log points)	-0.459*
	(0.228)
DiD (women, log points)	0.095
	(0.230)
R^2	0.174
Observations	4767

Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors clustered at the household level.

On the extensive margin, the triple interaction is small and statistically indistinguishable from zero across both datasets. Post-estimation average marginal effects likewise yield gender-specific DiDs that are close to zero with wide confidence intervals. Hence, I find no detectable gender gap in the probability of saving.

On the intensive margin, for households that save in both pre- and post-reform periods, the triple interaction is positive and marginally significant at the 10% level, implying that the reduction in amounts saved is larger for men than for women. Equivalently, the DiD for women is less negative (close to zero) than the DiD for men, suggesting evidence of a gender gap on the intensive margin.

While this result doesn't support the proposed mechanism, I argue it does not negate the mechanism either with the three reasons below.

1. Different margins can move differently: Mechanisms about retirement expectations can affect participation versus amounts in distinct ways. On the extensive margin, I found negative but insignificant gender gap between men and women, suggesting it is possible that the women exit the participation with higher probability. But among persistent savers, men reduced amounts more. This could reflect margin substitution: women may be stickier on amounts once they already save, while men adjust intensive margins more.

2. Selection into the intensive-margin sample: The intensive margin uses always-savers only. If the composition of always-savers differs by gender (age, income, risk preferences), the conditional response can differ even when the unconditional mechanism is the same. Conditional effects may therefore diverge from unconditional mechanisms without contradicting them.

3. Timing and macro shocks: The post window overlaps the 2008–2009 crisis. If male-dominated sectors had larger income volatility during those years, men's amounts may fall more independently of expectation-reanchoring. That would attenuate or flip the gender gap on the intensive margin without saying the expectations channel is wrong.

A.2 Transitions in Employment Type Over Time

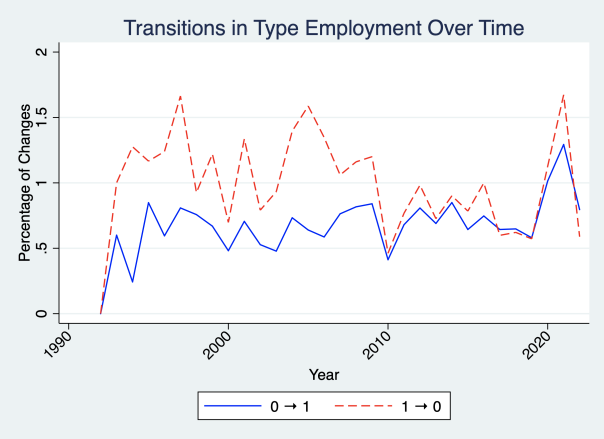


Figure 1: Transitions in Employment Type Over Time

FACULTY OF BUSINESS AND ECONOMICS
www.feb.kuleuven.be