Norges Bank

14 | 2023

Working Paper

Spousal Spillovers in the Labor Market: A Structural Assessment

Norges Bank Research

Authors:
Sigurd Mølster Galaasen
Herman Kruse

Keywords:
Joint retirement, couples, life cycle, pension reform, leisure complementarity, spousal spillover

Working papers fra Norges Bank, fra 1992/1 til 2009/2 kan bestilles på e-post: servicesenter@norges-bank.no

Fra 1999 og senere er publikasjonene tilgjengelige på www.norges-bank.no

Working papers inneholder forskningsarbeider og utredninger som vanligvis ikke har fått sin endelige form. Hensikten er blant annet at forfatteren kan motta kommentarer fra kolleger og andre interesserte. Synspunkter og konklusjoner i arbeidene står for forfatternes regning.

Working papers from Norges Bank, from 1992/1 to 2009/2 can be ordered by e-mail:
servicesenter@norges-bank.no

Working papers from 1999 onwards are available on www.norges-bank.no

Norges Bank's working papers present research projects and reports (not usually in their final form) and are intended inter alia to enable the author to benefit from the comments of colleagues and other interested parties. Views and conclusions expressed in working papers are the responsibility of the authors alone.

ISSN 1502-8143 (online)

ISBN 978-82-8379-304-8 (online)

Spousal Spillovers in the Labor Market: A Structural Assessment*

Sigurd M. Galaasen† Herman Kruse‡

December 12, 2023

Abstract

We explore the importance and nature of elderly couples' labor market interlinkages, and how such linkages shape the response to welfare reforms. To this end, we build a life cycle model with dual-earner households, featuring heterogeneous age gaps, non-separable leisure preferences, and endogenous retirement. To inform key preference parameters, our calibration exploits quasi-experimental evidence of spousal retirement spillovers from a pension reform in Norway. We show that the experimental evidence is highly informative about the degree of non-separability of leisure and that a substantial level of complementarity is required to match the data. Using our calibrated model, we find that the commonly observed tendency of couples to retire together, despite considerable age-gap heterogeneity, can be entirely explained by leisure complementarities. Moreover, comparing to a model with leisure separability reveals that one-third of the long-run labor supply impact of the pension reform is attributed to complementarity. This illustrates the importance of accounting for interdependent decisions when evaluating policy reforms.

Keywords: Joint retirement, couples, life cycle, pension reform, leisure complementarity, spousal spillover
JEL codes: D15, E2, E6, E65, H55, J08, J2

*The authors thank Kjetil Storesletten, Eric French, Fang Yang, seminar participants at Norges Bank, EEA-ESEM 2021, and IIPF 2021 for invaluable comments and suggestions. All data belong to Statistics Norway and are used in compliance with the rules given by the Norwegian Data Protection Authority.

†Norges Bank. E-mail: sigurd-molster.galaasen@norges-bank.no. This paper should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank.

‡Statistics Norway. E-mail: herman.kruse@ssb.no. This paper should not be reported as representing the views of Statistics Norway. The views expressed are those of the authors and do not necessarily reflect those of Statistics Norway. I would like to thank the Research Council of Norway for financial support (The evaluation of the pension reform, grant number 248868). I would also like to thank NORAM for travel funds.

1 Introduction

There has been a shift in the composition of households participating in the labor market, from traditional "breadwinner" male providers to dual-earner households, and only recently are these household compositions commonly approaching retirement age. This change became particularly noticeable starting with individuals born in the 1940s, who belong to the first cohorts where women have frequently had full-length careers. This leads to the narrow model-perspective of households as single units making a solitary labor market decision being potentially outdated.[1] To the extent that couples' labor choices are interlinked, the solitary view may result in erroneous conclusions when evaluating policy changes aimed at, for instance, incentivizing employment. Our goal in this paper is to explore the importance and nature of such spousal labor market interlinkages.

Individuals make a sharp and usually absorbing extensive margin labor market decision around retirement age (Rogerson and Wallenius, 2013). This decision may therefore be a fruitful route for explore the jointness of labor market activity within couples. Specifically, joint retirement is a common phenomenon in dual-earner households, despite significant age gaps between partners.[2] Several structural mechanisms have been proposed as potential explanations of this synchronized behavior, but the literature has not concluded on their relative importance (Michaud et al., 2020). Recent quasi-experimental evidence has, however, revealed a causal channel operating through preferences: due to complementarity in leisure, the retirement of one spouse increases the likelihood of the other also retiring (Kruse, 2021; Lalive and Parrotta, 2017; Stancanelli and van Soest, 2012a,b). Failing to consider the causal spillover between partners' choices when evaluating policy reforms can introduce biases in estimating the impacts of such reforms. Therefore, it is crucial to account for these interdependent decisions in policy evaluations (Coile, 2003).

In this paper we connect these empirical findings, i.e. the prevalence of joint retirement and the causal spillover effect, using a dual-earner life cycle model. Our fully specified structural model with endogenous retirement provides two key payoffs compared with the empirical approach. First, it enables quantification of the contribution of causal spillover effects to overall synchronized retirement behavior, thereby isolating the role of leisure complementarity in determining labor market outcomes. Second, the quasi-experimental approach usually only offers partial identification of complementarities. This is because it lacks the ability to separately identify the underlying income and substitution effects that determine spousal spillover due to having only one source of exogenous variation.[3] To make accurate policy predictions, it is necessary to disentangle these two forces. Our structural framework facilitates such separation, enabling a comprehensive evaluation of the significance of spousal spillover effects for the outcomes of labor market reforms.

Our framework is based on a standard life cycle model, which we enhance by incorporating

[1] See for example Doepke and Tertilt (2016) and Borella et al. (2018) for quantitative evaluations of the consequence of ignoring the family dimension in macroeconomic modelling.

[2] See Hurd (1990), Banks et al. (2010) and Hospido (2015) among others.

[3] When one spouse exits the labor market, theoretically, two opposing effects are generated on the labor supply incentives of the other spouse. First, as total household income is reduced, the standard income effect increases the work incentives of the other spouse. On the other hand, complementarity in leisure implies that the other spouse now face higher marginal utility of leisure, thereby decreasing their incentive to work.

dual-earner households with varying age gaps and non-separable leisure preferences as well as endogenous retirement. At the outset, individuals are modeled as married couples. Within each period, households make decisions on consumption and saving as well as the labor supply of both spouses on the intensive and extensive margins. These decisions are influenced by idiosyncratic shocks to productivity, marital status, and mortality.[4] At the age of 62, workers become eligible for old-age pension through an early retirement program. Following Heathcote et al. (2010), we adopt a unitary model of household decision-making. Under this framework, the household collectively determines the allocation of each spouse's time and pools their resources into a common budget constraint.

The model is calibrated to the Norwegian economy. Key preference parameters are estimated using a combination of simulated method of moments and indirect inference. We have access to individual level administrative data spanning the period 1993–2017. Importantly, this time period coincides with the implementation of the 2011 Norwegian pension reform, which aimed to enhance work incentives. In our calibration, we leverage this reform to quantify the extent and direction of non-separability in leisure. Crucially, the reform eliminated an implicit tax on labor income for an identifiable subset of older workers, leading to quasi-experimental variation in work incentives. We leverage this variation in our calibration. In particular, we first reproduce on our data the empirical findings in Kruse (2021), and show that workers are on average 9 percentage points more likely to postpone retirement if their spouse postpones retirement. In doing so, we utilize a panel of cohorts (1944–1952) covering the time period 2007–2015.[5] We then target this moment in our calibration. Specifically, to determine the degree of complementarity required for the model to align with the observed spousal spillover effects, we replicate the empirical design on a similar simulated panel. We find that a large degree of complementarity is required for the model to be consistent with the reduced-form spillover evidence. To further illustrate the identification process, we also consider an alternative calibration assuming separable leisure preferences, where we omit the reduced-form target. The implied spillover effect is then highly negative, illustrating the power of the reduced-form moment in identifying the complementarity parameter.

A crucial aspect of our calibrated model is its ability to accurately replicate the frequency of joint retirement observed in the data. In fact, we show that the causal link between spouses' labor market decisions is highly informative about the jointness of retirement. Our estimated model, featuring the empirical age gap distribution,[6] successfully generates the pronounced joint retirement patterns observed in the data. Consequently, the degree of complementarity consistent with the causal spillover effects of the Norwegian pension reform is simultaneously able to account for the observed patterns of joint retirement. In contrast, in the model with separable leisure, couples tend to not retire at the same time. The importance of leisure complementarity for explaining the incidence of joint retirement is consistent with other structural models of couples labor supply decisions (e.g.

[4]Upon separation, either through divorce or death, individuals remain single for the remainder of their lives.

[5]In line with the sample restrictions made by Kruse (2021).

[6]It is crucial to incorporate age gaps in the model in order to compare to the data along this dimension. In models with only same-age couples it is difficult to separate between preferences and age-effects as explanations for synchronized decisions.

Gustman and Steinmeier (2004); Casanova (2010); Michaud et al. (2020)). Our contribution to this literature lies in demonstrating that leisure complementarity can simultaneously account for the quantitative magnitude of spillover effects and the prevalence of couples retiring together. In other words, the causal spillover evidence implies a degree of complementarity which entirely accounts for prevalence of couples retiring together.

We then quantify the aggregate importance of leisure complementarity for old-age labor supply elasticities. To this end, we simulate the long-run impact of the 2011 Norwegian pension reform. The importance of complementarity is derived by comparing the long-run employment response in the model with and without leisure complementarity. Our quantitative model exercise demonstrates that, in the context of the Norwegian pension reform which aimed to enhance work incentives, approximately 28 percent of the long-run increase in labor supply can be attributed to leisure complementarity.

The importance of leisure complementarity for aggregate employment is dependent on whether either or both spouses are directly targeted by the policy change. If we consider only couples in which both spouses are targeted, meaning that both spouses are subject to some policy change that affects the individual incentives to work, leisure complementarity accounts for 12 percent of the aggregate employment response. On the other hand, in couples where only one spouse is targeted directly, leisure complementarity accounts for 37 percent of the aggregate employment response. Consequently, our results indicate that spillover effects are nearly three times as important in targeted versus universal reforms.[7] Many labor market reforms are effectively targeted as opposed to universal (e.g. policies aimed at reducing early retirement through welfare benefits such as disability pension). Our findings indicate that such reforms may have substantial positive indirect effects, through the spousal spillover channel.

1.1 Related Literature

Our paper relates to a growing body of structural models in macroeconomics and public finance that study the role of females and families for a range of macroeconomic questions, in which the canonical model of one-earner households is extended with a secondary earner. Borella et al. (2018) use this type of model to argue that without the family dimension, structural models will have a difficult time matching macroeconomic data. This is demonstrated in contributions by e.g. Eckstein and Lifshitz (2011) and Fernández and Wong (2014), both seeking to explain the large rise in female labor force participation over the 21st century. Similarly, Fukui et al. (2023) study whether the same secular trend could explain the rapid recovery of employment after recessions, and Heathcote et al. (2010) study role of families shaping wage inequality in the US. Another strand of the literature study the role of families for risk and social insurance. Cubeddu and Rios-Rull (1997) demonstrate how families may contribute to risk, and Blundell et al. (2016b) quantify the role of families as a risk sharing device, while Guner et al. (2012) and Blundell et al. (2016a) study the implication for tax and welfare reforms. Common to most of this literature is to impose separability or even

[7] The intuition is that without leisure complementarity, the non-targeted spouse reduces labor supply due to a positive income effect caused by the targeted spouse increasing their labor supply.

substitutability in spousal leisure.[8] We contribute to this literature by quantifying the degree of leisure complementarity among older workers and its implication for pension reforms.[9]

More directly related to our paper is the literature on structural models with joint retirement. Early contributions are those of Gustman and Steinmeier (2000), Gustman and Steinmeier (2004) and van der Klaauw and Wolpin (2008), who by maximum likelihood and indirect inference fit their structural models to non-experimental data on labor supply in dual-earner households. The estimated parameters imply a contribution of both correlated preferences and leisure complementarity for older US households' labor market outcomes. The complementary channel is further emphasized as being important in Michaud and Vermeulen (2011) and Michaud et al. (2020) using either spousal death or survey data to inform the model parameters, whereas the estimated model in Casanova (2010) puts more weight on interlinkage via the US spousal benefit program to explain joint exit. In stark contrast to these papers, Merkurieva (2023) find a high degree of coordination even with leisure being substitutes, but where couples' leisure preferences are interlinked via common increase in the desire for leisure as they age.

The distinguishing feature of our analysis is that we rely on quasi-experimental evidence of synchronized retirement behavior to structurally pin down the degree of complementarity. Its experimental design means that the reduced-form moment we target is entirely driven by the causal complementarity channel of joint exit, and thus not influenced by other channels such as correlated income shocks or preferences. We show that the causal spillover evidence is indeed very informative for our structural preference parameters, and implies a substantial desire for shared leisure which quantitatively explains the overall degree of synchronized retirement observed in the data. Our paper thus quantitatively connects the quasi-experimental literature on spillover effects (Kruse, 2021; Lalive and Parrotta, 2017; Stancanelli and van Soest, 2012a,b) with studies documenting couples' retirement synchronization (Hurd, 1990; Banks et al., 2010; Hospido, 2015).

Finally, our paper also relates to a growing literature that utilizes quasi-experimental evidence in the estimation of dynamic structural models, as discussed in Low and Meghir (2017).[10] There is a potential two-way benefit from this approach. On the one hand, using experimental data in the estimation provides sharper and more transparent identification of structural parameters. On the other hand, the structural model allows us to ask questions beyond those that can be answered in a purely empirical framework, by disentangling the underlying mechanisms and performing counterfactuals.

The remainder of this paper is organized as follows. First, in Section 2 we document the presence of joint exit in the context of Norway, and explain the Norwegian social security institutional setting with emphasis on the 2011 pension reform. In Section 3, we describe our life cycle model. Section 4 explains the model calibration, while Section 5 presents the results. Finally, we conclude in Section 6.

[8]Important exceptions are Blundell et al. (2016b) and Blundell et al. (2016a), whose estimated models imply the presence of leisure complementarity.

[9]Martín and Marcos (2010) models dual-earners in an OLG model with endogenous retirement, imposing that couples can only retire at the same time, and show that traditional single-earner households models are result in underestimation of the future financial burden of pension systems.

[10]For example, in labor (Attanasio et al., 2012), household finance (Briggs et al., 2021) and growth (Chen et al., 2021).

2 Data and Institutional Background

Before describing our model in Section 3, we first provide an overview of the data and document that the prevalence of joint labor market exit among older couples in Norway is substantial, in line with patterns observed elsewhere. We also present the main elements of the Norwegian pension system before and after 2011 reform and explain how the reform gave rise to quasi-experimental cross-sectional variation in work incentives. This variation enables us to uncover a causal explanation of joint retirement, driven by a desire to enjoy leisure together, which we exploit in the calibration of the model to pin down the degree of leisure complementarity.[11]

2.1 Data

Our data come from administrative records for the universe of Norwegians residents in the period 1993–2017. From these data we observe individuals' income and wealth from tax returns, workplace information such as contractual hours and firm attributes from an employer-employee register, as well as demographic characteristics such as age, gender and marital status from various population registers. Individuals are identified through unique personal identifiers which can be linked to their spouse.

In Section 3 we develop a life cycle model for couples and separated individuals which is calibrated to match retirement behavior before the 2011 Norwegian pension reform and the spillover impact of the reform. To align our data with the sample in Kruse (2021), we therefore restrict our empirical sample to married and separated individuals covered by the Norwegian early retirement program at age 60,[12] where at least one of the partners are born in the period 1944–1952. The restriction to individuals eligible for early retirement is related to the quasi-experimental design exploited in the reduced form analyse, explained in more detail in Section 2.3.

Because of its significance to our analysis, it is worth highlighting a central aspect of the data. Since tax returns in Norway are filed individually we observe each spouse's labor income. Using individual level income, a person is defined as retired if annual labor income is below the social security basic amount (abbreviated G in Norwegian). This threshold, which we throughout will denote by G, is indexed to average wage growth and is roughly USD 10,500 in 2023.[13]

2.2 Prevalence of Joint Retirement

As emphasized in Section 1, a large literature documents that joint retirement is common. We confirm that this is also true for Norway. Even though age gaps in couples are rather disperse, there is substantial bunching around same-year exits from the labor market within couples (Figure 1). As the figure shows, about 35 percent of couples in Norway retire within a year of each other, while

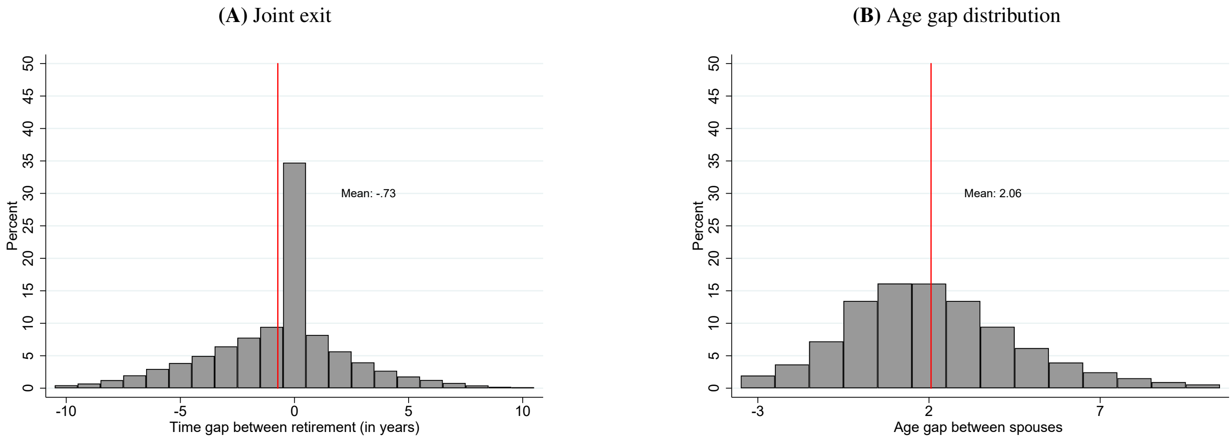
[11] The calibration procedure is explained in detail in Section 4. In short, we include the empirical spillover estimate as model target by replicating it on model simulated data.

[12] Workers' eligibility for early retirement is contingent on the type of firm they are employed with prior to retirement.

[13] This threshold is less than half of the minimum pension benefit in Norway, and it is reasonable to assume that individuals below this threshold have exited the labor market.

only around 15 percent of couples are of same age.[14] This is consistent with evidence from the US, where around 30–40 percent of couples exit the labor force within a year of each other (Hurd, 1990; Blau, 1998; Michaud, 2003).[15]

Figure 1: Joint exit and age gap distribution in couples.



Notes: Each bar in panel (A) represents the fraction of couples in which both spouses retire within a specific time gap. Negative values refer to female retiring first. Panel (B) displays the distribution of age gaps among couples. Negative values refer to female being older. The red lines show the mean of each distribution. Data source: Norwegian register data on all married couples where at least one partner is born in the period 1944–1952. Both partners must work at age 60.

However, the stylized fact does not imply that there is a causal relationship giving rise to the observed coordination. To identify causality, plausibly exogenous variation in the spouse's retirement decision is needed. This is precisely what the 2011 Norwegian pension reform provided by removing work disincentives embedded in the existing system. Importantly, the reform applied only to a subset of workers, thus generating quasi-experimental variation in incentives to work. We next explain the key aspects of the Norwegian pension system and the implementation of the 2011 reform.

2.3 The 2011 Norwegian Pension Reform

Since 1973, the statutory retirement age in Norway has been 67 years. However, from 1989, workers in the public sector and those covered by a collective agreement in the private sector have had the option to retire earlier, through a contractual early retirement program referred to by its Norwegian abbreviation "AFP".[16] The earliest retirement age under the AFP program was initially 66, but subsequently reduced to 62 by 1997. To be eligible for early retirement benefits, a worker had to be

[14]Figure 1 shows somewhat lager share of joint exit compared with Kruse (2021), because the latter defines joint exit as retirement within the same calendar year, as opposed to within a year of each other.

[15]Hospido (2015) find similar patterns using the Survey of Health, Ageing, and Retirement in Europe, and Banks et al. (2010) using both UK and US data.

[16]AFP is the initialism for the Norwegian term "*Avtalefestet pensjon*" and is similar to a standard early retirement program.

employed in a AFP firm for at least three out of the last five years leading up to retirement. Roughly 60 percent of the workforce satisfied this requirement at age 61 (Holmøy and Stensnes, 2008).

Under the AFP program, workers who were covered by the agreement could retire at age 62 on the same terms as those who retired at age 67. It was not possible to carry over the benefits from ages 62 to 66 if a worker chose to retire later. Additionally, if a worker claimed AFP benefits and continued working, an earnings-test gave a proportional reduction in benefits, effectively resulting in a marginal tax rate of one. This design of the system created strong incentives for workers to fully retire from the labor market and receive AFP benefits.

In line with recent policy advice throughout the OECD, Norway reformed its pension system in 2011. The AFP program was one of the major components that were subject to changes in the reform. Because the design of the AFP program gave strong incentives to retire at age 62, the earnings-test on continued work was abolished. Although the intention was to reform the public and private sector simultaneously, the negotiations in the public sector stranded. This resulted in a quasi-experimental setup, utilized by several researchers (e.g. Hernæs et al. (2016) and Kruse (2021) among others). Essentially, the workers in the private sector with a collective agreement could now claim early retirement benefits from age 62 and face *no earnings test*, while workers in the public sector still faced the same earnings test as before. By comparing private and public sector workers before and after 2011, researchers are therefore able to exploit exogenous variation in work incentives among the elderly. In particular the estimates in Kruse (2021) imply that in response to the spouse's retirement, individuals are roughly 10 percentage points more likely to themselves retire, indicating a significant degree of leisure complementarity.

In the life cycle model outlined in Section 3, we leverage the Norwegian reform to structurally assess the extent of leisure complementarity. To achieve this, we first integrate sectoral heterogeneity into the model: individuals are born as either covered by public sector or private sector AFP, and this characteristic remains fixed over the life cycle.[17] Our model thus generates the same form of reform induced exogenous variation in work incentives as in Kruse (2021).

In the model calibration in Section 4 we then simulate the introduction of the 2011 reform. Using the both our actual and simulated data we proceed by replicating the empirical design in that study, and calibrate parameters by matching the simulated reduced form estimates with estimates obtained from actual data.

3 Model

We consider a standard consumption-saving life cycle model, extended to include multi-person households and endogenous retirement. There are J overlapping generations, and each model period corresponds to a calendar year (age). Individuals enter the model at $j_g = 22$, born with gender g

[17] The model thus ignores sectoral job switches by assuming perfect foresight about AFP coverage to ease the computational burden. However, sectoral affiliation is very stable during the latter half of working careers in Norway. Hernæs (2017) shows roughly two-thirds of workers are covered by the AFP program at age 54. For individuals with private (public) sector AFP coverage at that age, only 3.5 (1.6) percent switches sector at age 55. At age 59 the persistence is even higher with only 2.8 (0.9) switching sectors.

(either male m or female f), and either married or single. If born as single, the person foretells future marriage at a pre-determined year when the spouse turns into age 22. Over the life cycle, marriages may dissolve either due to divorce or death. Households derive utility from consumption and leisure, and make decisions on consumption, saving, and labor supply. Financial markets are incomplete, and agents face three sources of idiosyncratic risk: labor earnings, marital status (divorce), and mortality. The household can save in a risk-free asset subject to a borrowing constraint.

In addition to heterogeneity generated by the realization of idiosyncratic risk over the life cycle, individuals are ex-ante heterogeneous along two dimensions. First, upon reaching age 62, agents are eligible for an old-age pension, either from the public or private sector early retirement as in the Norwegian AFP system. We define an individual ex-ante as being born either a private or public sector worker. The distinction between private and public employment is required for the model to incorporate the heterogeneous 2011 pension reform exposure, described in Section 2.3. In addition to heterogeneous reform exposure, private and public sector types differ in their labor earnings process, in line with the observed patterns in the data. Second, couples are ex-ante heterogeneous in their within-couple age disparity. In the data, most individuals live in marriages in which one spouse is older than the other. In the presence of leisure complementarity, such age gaps may have important implications for individual retirement patterns.[18]

We now turn to the household decision problem. First we present the couples' environment before we describe how the environment changes when when becoming single, in our model interpreted as being a divorcé(e) or a widow(er).

3.1 Couples

When explicitly modeling households as dual-earners, one must take a position on how household members coordinate their decisions. We follow Heathcote et al. (2010) and adopt a unitary model of multi-person households. In this framework, the household unanimously decides the allocation of each individual's time, and they pool their income into a common budget constraint. One interpretation of this approach is that both consumption and leisure are public goods within the household, or that spouses are perfectly altruistic towards each other. In contrast to households in Heathcote et al. (2010), our multi-person household also cares about future non-shared states in the event of marriage dissolution. To this end, we assume a Pareto problem with fixed Pareto weights, as in Hong and Ríos-Rull (2012).[19]

In the following sections, we introduce the household's choices and preferences, shock processes, and budget constraints. Since each household member can be one of two types (private or public sector worker), a couple household can be one of four types. To ease the notation, we suppress the subscript denoting which sector (type) the individual belongs to (private or public), but

[18] Age gaps mean that the older spouse will enter the model before the younger spouse. When solving the model, we make the (innocuous) assumption that the older spouse knows that marriage will be formed once the younger spouse enters the model at age 22.

[19] An alternative approach is to assume that each household member maximizes its own utility, with household allocations determined through bargaining. In this paper, we choose our approach due to its tractability and computational simplicity.

highlight the difference when introducing parameters which are sector dependent. In general, all parameters are common to all individual types, except for parameters related to the labor earnings process and post-reform pension scheme, which are type specific.

3.1.1 Preferences, Choices and Constraints

A couple household consists of two members $g \in \{m, f\}$, a male ($g = m$) and a female ($g = f$) spouse, of ages $j = \{j_m, j_f\}$. The couple enjoys utility from joint consumption, denoted by C, and male and female leisure denoted by L_m and L_f. We adopt an iso-elastic instantaneous joint-utility function, separable in consumption and household total leisure, but non-separable in spousal leisure. Instantaneous joint utility is given by the weighted sum of consumption and leisure utility:

$$U(C, L_f, L_m) = U_c(C) + \kappa U_l(L_f, L_m) \tag{1}$$

with κ representing the relative weight on leisure. The consumption component is given by:

$$U_c(C) = \frac{\left(\frac{C}{\zeta}\right)^{1-\gamma}}{1-\gamma} \tag{2}$$

where γ is the inverse of the intertemporal elasticity of substitution (IES) of consumption, and ζ is the equivalence scale of shared consumption determining the returns to scale in joint consumption. The leisure component is given by:

$$U_l(L_f, L_m) = \frac{1}{1-\phi}\left[\left(\eta L_f^{\rho} + (1-\eta) L_m^{\rho}\right)^{\frac{1}{\rho}}\right]^{1-\phi} \tag{3}$$

where ϕ is the inverse IES of total leisure, η the weight on female leisure, ρ the substitution elasticity between the male and female leisure.

The household chooses how much to consume and how much each member works each period, denoted H_g, for $g \in \{m, f\}$. The labor supply choice is discrete, and each member can either not work or work part-time or full-time. Full-time work consumes 0.4 units of time, and part-time is two-thirds of full-time work, i.e. $H_g \in \{0, \frac{2}{3} \times 0.4, 0.4\}$. Total time endowment, which can be allocated to leisure and work, is weakly decreasing in age and normalized to unity at birth:

$$E_{j_g} = \begin{cases} 1 & \text{if} \quad j_g \leq j_\chi \\ \exp(-\chi(j_g - j_\chi)) & \text{if} \quad j_g > j_\chi \end{cases} \tag{4}$$

where j_χ is the age at which the decline starts. This effectively makes work more costly as the agent approaches older ages, as less effective time is available. This is consistent with life cycle models in French (2005), French and Jones (2011) and Capatina (2015), in which individual time endowment is on average convexly decreasing with age, due to deteriorating health. The latter study also shows that an individual's expected time endowment decreases non-linearly with age. We do not model health explicitly and instead adopt a reduced-form approach.[20]

The labor supply choices maps onto leisure consumption at age j_g as follows:

$$L_{g,j_g} = E_{j_g} - H_{g,j_g} - \theta_g \mathbb{I}^{PT}_{g,j_g} \tag{5}$$

where the last term represents a gender-specific time cost of working part time. If working part time $\mathbb{I}^{PT}_{g,j_g} = 1$ the household lose θ_g units of time over and above the time spent working part-time. Introducing such non-convexities in the intensive margin labor supply is in line with e.g. French (2005) and Rogerson and Wallenius (2013) and is needed to explain the wide usage of partial retirement in Norway, in particular among women.[21]

Non-work (i.e $H_{g,j_g} = 0$) is assumed to be an absorbing state, which means that once retired, the worker cannot return to the labor market.[22] Formally, let $\mathbb{I}^{Ret}_{g,j_g}$ denote a 0-1 indicator function taking the value 1 if the household member g did not work in the previous period. The current period labor supply choice set is then defined as:

$$H_{g,j_g} \in \begin{cases} \{0\} & \text{if} \quad \mathbb{I}^{Ret}_{g,j_g} = 1 \\ \{0, \frac{2}{3} \times 0.4, 0.4\} & \text{if} \quad \mathbb{I}^{Ret}_{g,j_g} = 0 \end{cases} \tag{6}$$

The household evolution of joint assets follows from the sequence of period-by-period resource constraints:

$$\begin{aligned} A_{j+1} &= (1 + r(1-\tau_a))A_j + T(Y_{f,j_f}, B_{f,j_f}) + T(Y_{m,j_m}, B_{m,j_m}) - C_j \\ A_{j+1} &\geq 0, \end{aligned} \tag{7}$$

where the last constraint rules out borrowing. The asset A_j yields an exogenous safe return r taxed at rate τ_a and the function T maps pre-tax labor and pension income, Y_{g,j_g} and B_{g,j_g} onto after-tax income.

In the parameterization of the model explained in Section 4, we take values for (γ, ϕ, ζ) from standard estimates in the literature, and calibrate $(\rho, \eta, \kappa, \chi, \theta_m, \theta_f)$ internally by matching data moments reflecting retirement patterns in the data. Although all parameters influence all moments, intuitively, κ determines the average retirement age in the model, η the difference in average retirement between males and females, χ the age profile of retirement rates, and θ_g the average fraction of part-time workers. Finally, the value of ρ determines the degree of spousal spillover in retirement. Importantly, a key feature of the instantaneous utility function is the tractable interpretation

[20] More generally, this modeling assumption is broadly in line with several studies that adopt a convex (in age) time cost of working, to address the sharp rise in non-participation among old-age workers, see e.g. Kitao (2014); Cooley and Henriksen (2018); Cooley et al. (2022).

[21] In French (2005) the hourly wage rate is increasing in hours worked.

[22] Our primary focus is on the retirement pattern of older workers and our main calibration targets are labor supply patterns of workers between the ages of 60 and 69, who have not yet retired at age 60. The assumption of absorbing non-participation is consistent with the very low transition rate back to work among these individuals.

of leisure complementarity. Specifically, if:

$$1-\phi-\rho \begin{cases} >0 & \\ =0 & \implies \\ <0 & \end{cases} \begin{array}{l} \text{leisure is complementary} \\ \text{separable utility of leisure} \\ \text{leisure is a substitute} \end{array}$$

This simple relationship comes from the fact that the sign of the cross-partial derivative of utility with respect to leisure, e.g. how marginal utility of female leisure changes when male leisure increase, depends only on the sign of $1-\phi-\rho$. Complementarity in leisure implies that the marginal utility of leisure is increasing in the spouse's leisure. As we explain in Section 4, we identify the degree of complementarity by calibrating ρ such that the model matches the reduced-form spousal spillover impact of the 2011 Norwegian pension reform.[23]

3.1.2 Demographics

Since our focus is mainly on the later stages of the life cycle, we adopt a simple process for the family structure. Individuals are born into a couple, and remain married until divorce or widowhood. We abstract from remarriage, assuming that once the relationship ends, individuals remain single for the remainder of their lives.

Couples have five possible next-period demographic outcomes: they survive and remain married, they survive and divorce, the female spouse dies, the male spouse dies, or both die. Ignoring the state where both spouses die, the probability of transitioning from currently married at age composition $j=\{j_m,j_f\}$ to next-period states is denote by $\xi^z_{j_m,j_f}$, where $z\in\{rm,d,mw,fw\}$ denotes "remain married", "divorce", "(male) widower" and "(female) widow", respectively. These probabilities are derived from the exogenous survival probabilities p_{g,j_g} and divorce probabilities $d(j_m,j_w)$.[24] Hence, divorce is an exogenous shock similar to e.g. Cubeddu and Ríos-Rull (2003) and Hong and Ríos-Rull (2012).

3.1.3 Labor Earnings Process

At the beginning of a period, each household member $g\in\{m,f\}$ is endowed with ω_{g,j_g} units of log labor efficiency per unit of time, generating a labor income of:

$$Y_{g,j_g}=\exp(\omega_{g,j_g})H_{g,j_g} \tag{8}$$

Endowments $\omega_{g,j_g}=q_{g,j_g}+e_{g,j_g}$ consist of two parts. The first term, q, captures the deterministic age-component of the labor earnings profile, common to all agents. The second term, e, captures

[23]Note that the degree of complementarity depends on both θ and ρ. In Appendix B, we show that the estimated degree of complementarity remains quantitatively unchanged when re-calibrating the model under alternative ϕ values.

[24]For example, the probability of becoming divorced in the next period is $\xi^d_{j_m,j_f}=s_{m,j_m+1}p_{f,j_f+1}d(j_m,j_f)$.

idiosyncratic shocks arising from an AR(1) process:

$$e_{g,j_g} = \nu_g e_{g,j_g-1} + \varepsilon_{g,j_g}, \qquad (9)$$
$$\varepsilon_{g,j_g} \sim N(0, \sigma^2_{\varepsilon_g})$$

Earnings shocks are independent across spouses. Note that correlated earnings processes could generate correlated retirement behavior across spouses, in much the same way as leisure complementarity. However, since we pin down the degree of complementarity from exogenous variation in retirement incentives embedded in the Norwegian pension reform, our estimate is robust to potential correlated income processes in the data.[25] When taking the labor earnings relationship to the data, we estimate both the age profile and the shock process separately for the different worker types (public or private sector)

3.1.4 Pension System

To keep the model tractable, we adopt a simplified pension system in the model. Pension benefits are pre-determined and assumed to be 60 percent of the deterministic part of full-time labor earnings at age 64:

$$\hat{B}_g = 0.6 \exp(q_{g64}) 0.4 \qquad (10)$$

Although highly stylized, the fraction 0.6 is roughly in line with the average gross replacement rate in the Norwegian system. Note that despite the fact that the pension benefit level is homogeneous (within worker type/gender), heterogeneous labor earnings generate progressive replacement rates.

Workers are eligible for pensions when they reach age 62. However, the pension benefit is tested against labor earnings. Importantly, consistent with the Norwegian pension system prior to 2011, the lost pension benefit does not lead to an upward readjustment of future pensions, causing a substantial implicit tax on labor. The per-period net pension benefit becomes:

$$B_{g,j_g} = f_{j_g}(\hat{B}_g, Y_{g,j_g}), \qquad (11)$$

where the function f is zero until age 62, and decreasing in Y.

In particular, we differentiate between two types of earnings-tests broadly consistent with the pre-reform pension system. Between ages 62 abd 66 (i.e. the early retirement scheme), pension benefits are reduced in proportion to labor earnings relative to full-time labor earnings. Between ages 67 and 69, pension is reduced by 40 percent of labor earnings above $2G$, where G denotes the basic amount in the Norwegian pension system.[26]

[25] In Appendix A we also present evidence of small spousal earnings correlation in the data and robustness exercises allowing for a wide range of correlation coefficients.

[26] The basic amount G is calibrated to be 18 percent of the average labor earnings of individuals between age 40 and 44, consistent with the data

$$
f_{j_g}(\hat{B}_g, Y_{g,j_g}) = \begin{cases} 0 & \text{if} \quad j_g < 62 \\ \hat{B}_g(1 - \frac{Y_{g,jg}}{Y_{g,jg|fulltime}}) & \text{if} \quad 62 \geq j_g < 67 \\ \max(0, \hat{B}_g - 0.4\max(0, Y_{g,j_g} - 2G)) & \text{if} \quad 67 \geq j_g < 70 \\ \hat{B}_g & \text{if} \quad 70 \geq j_g \end{cases} \tag{12}
$$

As we explain in the internal model calibration in Section 4.2, we implement the 2011 pension reform as a relaxation of the earnings test in (12), but only for private sector workers.

3.2 Divorced or Widowed Individuals

When separated, individuals become permanent single-person households with flow utility:

$$
U^s(C_g, L_g) = \frac{C_g^{1-\gamma}}{1-\gamma} + \frac{\kappa^s}{1-\phi} L_g^{1-\phi} \tag{13}
$$

and time endowment:

$$
E_{j_g}^s = \begin{cases} 1 & \text{if} \quad j_g \leq j_\chi \\ \exp(-\chi^s(j_g - j_\chi)) & \text{if} \quad j_g > j_\chi, \end{cases} \tag{14}
$$

where the superscript s denotes single household. We allow the leisure weight κ^s and the rate of decline in time endowment χ^s to potentially differ to that of a couple household. When calibrating the model, these parameters will target the retirement patterns of separated individuals.

The evolution of assets is given by:

$$
\begin{aligned} A_{g,j_g+1} &= (1 + r(1-\tau_a))A_{g,j_g} + T(Y_{g,j_g}, B_{g,j_g}) - C_{g,j_g} \\ A_{g,j_g+1} &\geq 0. \end{aligned} \tag{15}
$$

Separated agents face the same individual mortality risk, earnings process, and pension system as when in a couple. Labor earnings thus evolve according to (8)–(9) and pension benefits are determined by (10)–(11). Work hours maps onto leisure following (5) and labor supply remains potentially constrained by absorbing retirement following (6).

3.3 Recursive Formulation

We formulate the maximization problem in recursive form. Let the vector of state variables for a couple be:

$$
\Omega = \left\{ A, e_m, e_f, j_m, j_f, \mathbb{I}_m^{Ret}, \mathbb{I}_f^{Ret} \right\}, \tag{16}
$$

where m denotes the male spouse and f denotes the female spouse, and a is the shared assets, e is the current income shock, j is the current age. The corresponding state vector for a single person of gender g is:

$$\Omega_g = \left\{A_g, e_g, j_g, \mathbb{I}_g^{Ret}\right\}. \tag{17}$$

Let V and V_g^s denote couples' and singles' value functions. To ease the notation, we let X' represent the next period's value of a vector X, and suppress the age subscript. Then in recursive form, the couples' problem is given by:

$$\begin{aligned} V(\Omega) = \max_{C,H_f,H_m} &\left\{u(C,L_f,L_m) + \right. \\ &\beta\left[\xi^{rm}\mathbb{E}V(\Omega'|\Omega) + (1-\lambda)(\xi^d + \xi^{fw})\mathbb{E}V_f^s(\Omega_f'|\Omega) + \right. \\ &\left.\lambda(\xi^d + \xi^{mw})\mathbb{E}V_m^s(\Omega_m'|\Omega))\right] \end{aligned} \tag{18}$$

subject to (4)–(11), where β is the discount factor. The first term in the square brackets represents the continuation value of remaining married, which happens with probability ξ^{rm}. The next four continuation values represent states that are not shared. The first two belong to the female spouse, who may either get divorced or become a widow, occurring with probability ξ^d or ξ^{fw}. The final two are the continuation values for the male spouse who may become divorced or a widower, similarly occurring with probability ξ^d or ξ^{mw}. In the event of both spouses dying, the continuation value is zero, implying no bequest motive. The parameter λ reflects the weight the couple puts on the male continuation value in future states that are not shared. We set this parameter to $\lambda = 0.5$, implying equal weight on future non-shared states, consistent with the altruistic interpretation of the instantaneous utility function. The expectation operator $\mathbb{E}$ is over the next period earnings shock realization, conditional on current period state.

Upon death, the joint asset A' is transferred to the widow(er). Upon divorce, joint assets are divided among the spouses and a share π is allocated to the female, implying $A_f' = \pi A'$ and $A_m' = (1-\pi)A'$.

For singles, the corresponding problem is:

$$V_g^s(\Omega_g) = \max_{C_g,H_g}\left\{U^s(C_g,L_g) + \beta p_g \mathbb{E}\left[V_g^s(\Omega_g'|\Omega_g)\right]\right\} \tag{19}$$

subject to (5)–(6), (8)–(11) and (14)–(15).

For the older spouse in a couple, we also need to solve the decision problem prior to the younger spouse entering the model. We do so by assuming that the older spouse adopts the single household decision problem, with perfect foresight about the transition into a couple when the younger spouse enters the model at age 22.

3.4 Solution Method

The model is solved by backward recursion. We use a straightforward discretization of the state and control space. The auto-regressive part of the earnings process is approximated by a two-state first-order Markov process, following Tauchen and Hussey (1991). We consider a partial equilibrium framework, meaning that interest rates, wages, and tax rates are constant and exogenous.

4 Calibration

We estimate the model using a combination of external and internal calibrations. In the external calibration presented in Section 4.1, we quantify the mortality, divorce and labor earnings processes using Norwegian register data. The tax and pension system rules are set to mimic key features of the Norwegian system, both before and after the 2011 pension reform. Finally, a set of preference parameters are taken from the literature. Table 1 provides an overview of the external parameters, values, and data sources. In the internal calibration, presented in Section 4.2, we pin down parameters by matching simulated moments to corresponding data moments. In particular, the remaining parameters $\{\rho, \eta, \beta, \kappa, \kappa^S, \chi, \chi^S, \theta_m, \theta_f\}$ target the retirement patterns of men and women, wealth to income ratios, as well as spousal spillover effects in retirement. The target empirical estimate for the spousal spillover effect is found through a replication of Kruse (2021) which we present in Section 4.2 below. Table 3 provides an overview of the parameters that are internally calibrated and their targets.

4.1 External Calibration

Mortality and Divorce Survival probabilities are taken from the cross-sectional gender-specific mortality table from 2007 and we cap the maximum age to $J = 102$ (which corresponds to 81 model periods).

The divorce data is taken from a longitudinal panel of all Norwegian individuals from the period 1993–2015. The only initial restriction we make when computing the demographic statistics is that the individuals must be above age 22 and be Norwegian residents, which means that we exclude emigrants (but include immigrants). Divorce probabilities are then estimated for every age between 22 and 70. The divorce risk is the (weighted) average of married individuals who divorce at age $j+1$ to the number of individuals married at age j.[27]

Age Gap and Sector Distributions The distributions are computed on the sample of the 1944–1952 cohorts of married individuals who are working at age 60.[28] We use the same sample as in the empirical estimation of spousal spillover. On average, men are about two years older than their wives, and the distribution is skewed heavily towards men being the older spouse.

[27] The weights reflect the relative sample size for each year.

[28] Since the spillover coefficient in Kruse (2021) is a key target in the internal calibration, we choose the same sample of individuals as in that paper. See Section 4.2.

Table 1: Externally calibrated model parameters.

Parameter	Description	Value	Source
Demographics			
p_{g,j_g}	Survival probabilities		Statistics Norway
$d(j_m, j_w)$	Divorce probabilities		Statistics Norway
J	Maximum age	102	
j_χ	Age when time endowment starts declining	61	
$\max(j_m - j_f)$ and $\min(j_m - j_f)$	Age gap extrema	$\in \{-3, 7\}$	
$j_m - j_f$	Age gap distribution		Statistics Norway
Preferences			
ϕ	Shape parameter total leisure	3	Heathcote et al. (2010)
γ	Shape parameter consumption	1.5	Low and Pistaferri (2015)
ζ	Equivalence scale	1.3	Hong and Ríos-Rull (2012)
π	Sharing rule, joint assets	0.5	Hong and Ríos-Rull (2012)
λ	Weight on male continuation value	0.5	
Labor earnings process			
ν_m	Wage persistency, men, private sector	0.947	Statistics Norway
ν_f	Wage persistency, women, private sector	0.918	– –
$\sigma_{\varepsilon_m^2}$	Wage innovation, men, private sector	0.013	– –
$\sigma_{\varepsilon_f^2}$	Wage innovation, women, private sector	0.014	– –
ν_m	Wage persistency, men, public sector	0.893	– –
ν_f	Wage persistency, women, public sector	0.794	– –
$\sigma_{\varepsilon_m^2}$	Wage innovation, men, public sector	0.019	– –
$\sigma_{\varepsilon_f^2}$	Wage innovation, women, public sector	0.026	– –
q_{g,j_g}	Gender- and sector- specific age-earnings profiles	Table 2	– –
Pension system			
$f_{j_g}()$	Earnings test schedule on pension benefits		Norwegian social security rules
Government			
r	Pre-tax interest rate	0.04	
τ_a	Tax on interest income	0.28	Norwegian tax rules

Source: Statistics Norway: authors' own calculations using administrative register data.

We calibrate the fraction of men and women in each sector to the corresponding sample distribution. Since all individuals in the model are covered by AFP (the early retirement option) from either the private or the public sector, we also exclude non-covered individuals from the sample when we find the ratios. We then get an average of the four different couple compositions: both have AFP in private sector, both have AFP in public sector and one spouse is covered by private AFP while the other is covered by public AFP. There is a significant difference in the fraction of men and women who work in the two sectors, which means that the group where the husband is covered by private AFP and the wife by public AFP is larger than vice versa.[29]

Labor Earnings Process The labor earnings process is estimated on an annual panel data of monthly earnings for Norwegian workers above age 22 in the period 1997–2017.[30] The data also contains information about registered contracted hours, from which we obtain a measure of hourly wages.[31] The equation map to the data is the empirical analog to the endowment ω in (8), where we have approximated the age effect by a fourth degree polynomial:

$$\ln(w_{i,j,t}) = \alpha_i + \beta_1 j + \beta_2 j^2 + \beta_3 j^3 + \beta_4 j^4 + \psi_t \mathbb{D}_t + e_{i,j,t}. \tag{20}$$

The subscripts (i, j, t) refer to an individual i of age j in year t. The dependent variable is hourly wages, α_i an individual fixed effect, $\mathbb{D}_t$ a set of year t dummies and $e_{i,j,t}$ represents the idiosyncratic shock component of earnings. We estimate the coefficients by running OLS on (20), separately for men and women, and for public and private sector workers with access to AFP.[32] Our deterministic age-earnings profiles in (8) then follow:

$$q_{g,j_g} = \hat{\alpha} + \hat{\beta}_{g1} j + \hat{\beta}_{g2} j^2 + \hat{\beta}_{g3} j^3 + \hat{\beta}_{g4} j^4, \tag{21}$$

where $\hat{\alpha}$ represents the average of the individual specific intercepts. The estimated age-coefficients from these four regressions are reported in Table 2.

As is common in life cycle model with endogenous retirement, e.g. see Imrohoroğlu and Kitao (2012), we set labor efficiency $\omega_{g,j_g} = 0$ for individuals older than 69, implicitly assuming that agents do not work past age 70. In the data, very few work after reaching that age.[33] The resulting polynomial coefficients from estimating (20) using OLS are reported in Table 2.

The stochastic AR(1) earnings process in (9) is estimated on the residuals, $\hat{e}_{ij}$, obtained from

[29] In the data, the group where the husband has private AFP, while the wife does not, comprises 16 percent of the couples, while the group where the wife has private AFP and the husband does not comprises about 10 percent.

[30] We use both married and unmarried individuals to maximize individual time horizons on which we infer age profiles of earnings.

[31] Statistics Norway (SSB) produces (annually in September/October) salary statistics, following Eurostat regulations. Private sector data comes from sample surveys, excluding primary industries and covers around 70 percent of workers, while public sector statistics rely on registry information and covers the entire sector.

[32] We also rescale age so that age = 22 is considered zero in the estimation, while age = 23 is considered one, and so on. This ensures consistent mapping to the model, where the individuals enter at age 22.

[33] Due to selection bias, it would be difficult to estimate the labor productivity past age 70. Additionally, under the Norwegian Working Environment Act, workers are protected from unfair dismissal, e.g. due to age. However, during the time period we calibrate the model to, workers are not protected by this act when reaching age 70.

Table 2: Income profile polynomial regression. Coefficients from OLS on (20).

	α	β_1	β_2	β_3	β_4
Private sector, men	5.293	0.0197	-0.0005	$3.5\cdot10^{-6}$	$-3.0\cdot10^{-8}$
Public sector, men	5.082	0.0532	-0.0024	0.00005	$-3.5\cdot10^{-7}$
Private sector, women	5.139	0.0145	-0.0003	$-3.5\cdot10^{-6}$	$-4.9\cdot10^{-8}$
Public sector, women	5.253	0.0264	-0.0017	0.00004	$-3.6\cdot10^{-7}$

the OLS regressions on (20). By OLS estimation of:

$$\hat{e}_{ij} = \nu\hat{e}_{ij-1} + \varepsilon_{ij}, \tag{22}$$

we obtain values for the persistency and standard deviation of the shock. The results reported in Table 1, imply that men and private sector workers face higher lifetime earnings risk compared with women and public sector workers.[34]

Tax System We set $\tau_a = 0.28$, in line with the tax rate on capital income during our sample period. The tax function T takes into account the progressivity of the Norwegian tax schedule. In particular, it consists of a national insurance contribution, a general income tax, and a surtax on high income earners. The national insurance contribution amounts to 7.8 and 3 percent of labor earnings and pension income, respectively. The general income tax is computed based on taxable income, which is total labor and pension income net of a deductible amount. The deductible amount varies between $0.55G$ and $1.55G$, and increases as the share of income derived from labor increases. The taxable amount is then taxed at 28 percent. Finally, the surtax on high income earners follows a three-bracket schedule, with a marginal tax rate of 0 (up to $6G$), 9 (up to $10G$) and 21 percent (above $10G$).[35]

Interest Rate and Wealth Partition The interest rate is set to $r = 0.04$. We follow Hong and Ríos-Rull (2012) and assume that, upon divorce, household wealth is divided equally between the spouses, $\pi = 0.5$.[36]

External Preference Calibration We set $\gamma = 1.5$ following a long tradition in the literature, e.g. as in Low and Pistaferri (2015), Low et al. (2018) and Heathcote et al. (2010).[37] For the leisure curvature parameter we set $\phi = 3$, which is the value used in Heathcote et al. (2010), in a model

[34] This is consistent with the fact that the public sector wages are more tightly linked to the Norwegian centralized wage negotiations compared with private sector wages.

[35] In the post-pension reform regime, workers combine full time work with untested pension income. To avoid placing these individuals in the surtax schedule, we levy this only on labor income. This is broadly in line with adjustments made in the actual Norwegian tax system.

[36] This partition rule is in line with the Marriage Act in Norway.

[37] $\gamma = 1.5$ is within the range of commonly estimated values; e.g. estimates vary between 1.35 in Attanasio et al. (1999) to 2.0 in Banks et al. (2001). In Appendix B we show that the degree of complementary remains quantitatively large when reducing (increasing) γ to 1.35 (2.0).

with separable spousal leisure in utility.[38]

The value for the equivalence scale in the couples' utility function is set to $\zeta = 1.3$, in line with the estimate in Hong and Ríos-Rull (2012). This implies that a couple would need to spend only 30 percent more than a single individual to draw the same utility from consumption.

4.2 Internal Calibration

The remaining eight parameters $(\beta, \eta, \kappa, \chi, \theta_m, \theta_f, \rho, \kappa^S, \chi^S)$ are calibrated jointly by matching simulated model moments to corresponding data moments. The first six parameter targets data for couples, whereas the remaining two single-specific parameters (κ^S, χ^S) target data for separated individuals.

Except for the parameter governing the degree of leisure complementarity (ρ), all parameters target pre-2011 data. For ρ we target the spillover effect of the 2011 pension reform. Before explaining the details of the calibration, we first describe our implementation of the pension reform and how we generate simulated data.

Simulated Data and Pension Reform Implementation Pre-reform simulated data is generated from a stationary model distribution with the pre-2011 pension system in place. We simulate several cohorts, such that at least one spouse is born in the period 1944–1952 given the empirical age-gap distribution. We then invoke the pension reform, i.e. removal of the earnings test in (12), as a shock occurring at the beginning of 2011. The pension reform applies only to a subset of workers. In particular, it applies only to workers in the private sector. In addition, private sector workers who take out an early retirement pension before the reform still face the pre-reform early retirement earnings test (i.e. equation (12) for workers younger than 67). Overall, our model implementation captures the key changes in labor supply incentives embedded in the actual 2011 reform. In Appendix C we argue that our implementation also implicitly captures other features of the reform.

Couple Parameters The parameters to be estimated are $\beta, \eta, \kappa, \chi, \theta_m, \theta_f$ and ρ. The discount factor β targets the average wealth-to-labor income ratio at age 60. The weights on the wife's and total leisure, η and κ target average retirement rate among males and females in age group 63–65. The rate of decline in time endowment χ targets the average growth rate in retirement rates between age groups 63–65 and 66–68.[39] In the data we classify an individual as retired when annual labor income is below $1G$.[40] Finally, the part-time penalty cost parameters θ_m and θ_f targets the average share of older (age 60–69) employed workers who work part-time in the data, where part-time is defined as working less than 30 hours per week.

[38] The value for ϕ is in line with commonly used values, referring to the Frisch elasticity of labor supply. In a preference specification such as in (1) with $\phi = 3$, $\rho = 1 - \phi$ (no complementarity) and $\chi = 0$, assuming continuous labor supply choice (h) , the Frisch elasticity is $\frac{1-h}{h}\frac{1}{\phi}$. The elasticity of an individual working full time hours would be 0.5. In B we show that the degree of complementarity remains quantitatively unchanged when varying $\phi = 3$.

[39] Since we focus on retirement patterns for individuals who still work at age 60, we assume that the decline starts at age 61 by setting $j_\chi = 61$.

[40] This is consistent with the definition of retirement used in Kruse (2021).

Calibrating Complementarity The remaining parameter, ρ, governing the degree of complementarity in leisure, targets the reduced form spillover effects of the Norwegian pension reform. Identification relies on the fact that public sector workers were not affected by the reform as explained in Section 2.3. We then estimate spousal spillover effects by comparing the employment of an individual married to spouse who works in the private sector with AFP (i.e. a treated spouse) to those married to a spouse who works in the public sector (i.e. a non-treated spouse). When calibrating ρ we implement a reduced form method on simulated data from our model. In particular we estimate the following linear probability model as in Kruse (2021):

$$D_{ij} = a_0 + a_1 D^s_{ij} + a_2 X_{i,j} + \varepsilon_{i,t} \tag{23}$$

where D_{it} is a 0-1 indicator of retirement of individual i in period at age j, and D^s_{it} is a similar indicator for individual i's spouse. Additional control variables are included in the vector $X_{i,j}$, including a dummy variable for whether the spouse works in the private sector. The coefficient of interest is a_1 which measure the impact of spousal retirement on the likelihood of own retirement. We estimate the relation in (23) by instrumenting the endogenous spousal retirement state with the spouse's reform exposure.

The estimation is performed using the same sample restriction, instrumental variable definition, and controls as in Kruse (2021). In particular, we use a simulated sample of individuals where the spouse of individual i is born in the periods 1944–1947 or 1949–1952 and works at age 60, and both the individual and the spouse are between ages 63 and 66. The instrument for D^s_{ij} is a dummy variable D^{ref}_i taking the value one if worker i's spouse is born between 1949 and 1952 and works in the private sector.[41] The estimation of (23) is conducted for employment outcomes of individuals i who themselves work in the public sector, thus not directly affected by the reform, and performed separately for men and women. The simulated spillover moment is constructed as the equally-weighted average of the gender-specific estimates of a_1.

The corresponding targeted moment is the equally-weighted average of the gender-specific estimates in the data. Following the procedure in Kruse (2021), we replicate the exact setting of that paper using our updated data which contains more precise information on individuals' AFP affiliation. The resulting spillover data moment which we target in our calibration is 0.09,[42] which is interpreted as a 9 percentage point increase in the probability of postponing retirement if ones spouse postpones retirement.

[41] Defining the instrument based on the spouse's year of birth is exactly in line with Kruse (2021). However, the reform applied to all private sector AFP-affiliated workers as of 2011, who had not taken up early retirement pension prior to the reform. Hence, spouses born in the period 1945-1947 who are still working full time when the reform is implemented in 2011 face the new pension scheme, both in the data and in the model. This implies that a subset of observations with $D^{ref}_i = 0$ may in fact respond to the reform. However, our results are not sensitive to this. When using the calibrated parameters, and re-estimate (23) restricting the sample to cohorts born in 1944 and between 1949 and 1952, the simulated $\hat{a}_1$ barely moves (it goes from 0.09 to 0.093).

[42] Compared with 0.115 in Kruse (2021).

Single Parameters The parameters κ^S and χ^S target average retirement for the age group 63–65 and the growth in retirement rate between ages 63–65 and 66–68. We use individuals born in the 1944–1952 cohorts who separate through divorce or the death of their spouse between ages 50 and 59.

5 Quantitative Results

In this section, we present our main results, and we rely on the structural model for three primary purposes: (i) to identify the degree of leisure complementarity, leveraging quasi-experimental data on spousal spillover effects, (ii) to provide an explanation for the observed synchronization of retirement choices in the empirical data, and (iii) to quantify the importance of joint retirement for aggregate labor supply elasticities. Sections 5.2–5.4 are organized around these three issues. Before we proceed, we present the calibrated model in Section 5.1.

5.1 Discussion of the Estimated Model

Table 3: Internally calibrated model parameters.

Parameter	Value	Target moment	Data value*	Model value
Couple parameters complementarity				
ρ	-3.495	Retirement spillover estimate in Kruse (2021)	0.09	0.09
κ	0.368	Mean retirement at age 63–65, women	0.38	0.38
η	0.5	Level difference of retirement between 63–65 year old men and 63–65 year old women	0.05	0.05
χ	0.0306	Growth in retirement from 63–65 to 66–68	1.89	1.89
Single parameters				
κ^S	0.324	Mean retirement at age 63–65, single	0.28	0.28
χ^S	0.0444	Growth in retirement from 63–65 to 66–68, single	2.42	2.42
Shared parameters				
β	0.961	Wealth-to-income ratio (at age 60)	1.87	1.87
θ_m	0.089	Part-time share of men aged 60–69	0.20	0.20
θ_f	0.074	Part-time share of women aged 60–69	0.50	0.50

*Source: Norwegian register data.
Notes: Data values for couple parameters are computed on a sample consisting of married couples where at least one spouse is born in the period 1944–1952, both work at age 60. For singles, we use the same birth cohorts and the single individual must work at age 60.

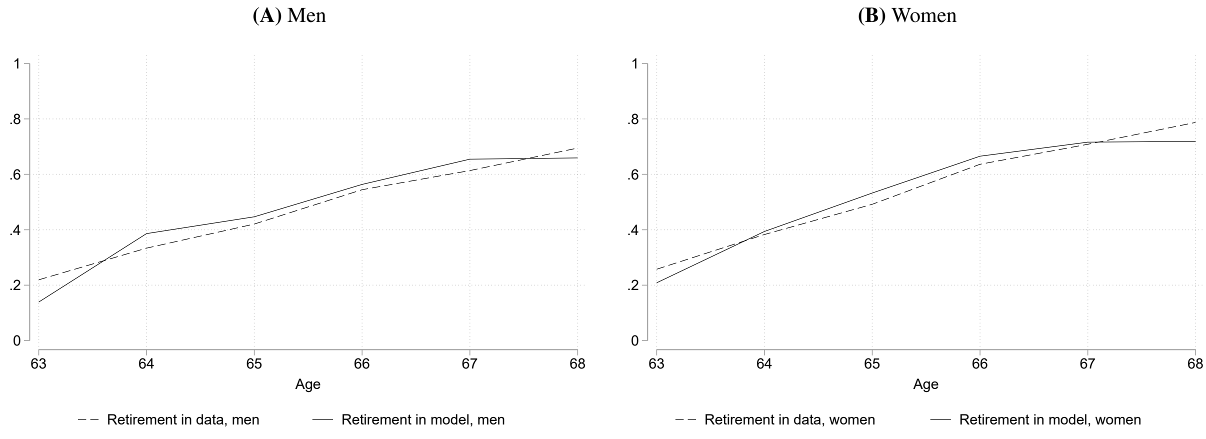
In Table 3, we report the parameter values and model simulated moments, obtained through our calibration process. The average (across gender) empirical point estimate suggests that individuals are 9 percentage point more likely to work if their spouse is also working. Our model calibration, which replicates the reduced form data sample and experimental setting, yields a similar spillover effect. Specifically, we find that the preference parameter governing non-separability is $\rho \approx -3.5$. This implies that leisure complementarity is a crucial factor in explaining the observed spousal spillover effect in the data.[43]

In our calibration, we target the mean retirement at age 63–65 for both men and women separately. We observe in the empirical data that, on average, women retire earlier than men. To account for this difference, we calibrate two key parameters: the weight on household total leisure, denoted as κ, and the relative weight on the wife's leisure component, denoted as η. We set the value of κ to match the mean retirement of women in the age group 63–65. Then we adjust η to account for the relative difference in retirement ages between men and women. We precisely match the empirical moments using these parameters, with $\kappa = 0.368$ and $\eta = 0.5$.

We then target the growth in retirement from 63–65, which is an equally weighted average of growth for both men and women. This growth rate serves as the basis for calibrating our time endowment parameter, denoted as χ, governing the increase in labor utility cost among elderly workers as they age. We obtain a value of $\chi = 0.0306$, which can be interpreted as a percentage loss of ≈ 3.1 percent of time endowment for each year surpassing age 61. Figure 2 shows that our model retirement profile performs well across all ages and both genders, even though only targeting average retirement rates between 63–65 and 66–68.

Finally, we target the average wealth-to-income ratio in our empirical sample using the parameter β. We match the empirical wealth-to-income ratio with a reasonable parameter value of $\beta \approx 0.96$.

Figure 2: Average retirement, data versus model.



Notes: Retirement in data is calculated on the couples where at least one partner is born in the period 1944–1952. Both spouses must be observed working at age 60. Retirement in data is defined as having annual labor income less than the social security basic amount (G). Similar restrictions are made to the model simulated individuals.

Parameters specific to separated individuals are calibrated to be somewhat smaller (κ^S) and larger (χ^S) compared with the corresponding couple parameters. These parameter values imply that separated individuals value leisure relatively less, but face a steeper decline in time endow-

[43] As explained in Section 3.1.1 spousal leisure are complements if $1 - \phi - \rho > 0$.

ment in old age. The parameter adjustment is needed in order to rationalize the fact that separated individuals on average postpone retirement relative to couples.

5.2 The Estimated Complementarity Parameter

As demonstrated by our benchmark model, we are required to incorporate substantial leisure complementarity to align with the observed reduced-form spillover effects. To explore the role of the parameter ρ in shaping this result, we conduct an alternative calibration in which we impose separability in the couple's utility function. In this case, we set $\rho = -2$ and re-calibrate the model without targeting the reduced form coefficient. The resulting parameter values are reported in Table 4.

Table 4: Internally calibrated model parameters (separable leisure).

Parameter	Value	Target moment	Data value*	Model value
Couple parameters separability				
κ	0.303	Mean retirement at age 63–65, women	0.38	0.38
η	0.504	Level difference of retirement between 63–65 year old men and 63–65 year old women	0.05	0.05
χ	0.0438	Growth in retirement from 63–65 to 66–68	1.91	1.92
Single parameters				
κ^S	0.331	Mean retirement at age 63–65, single	0.26	0.26
χ^S	0.0459	Growth in retirement from 63–65 to 66–68, single	2.25	2.25
Shared parameters				
β	0.96	Wealth-to-income ratio (at age 60)	1.87	1.87
θ_m	0.073	Part-time share of men aged 60–69	0.20	0.20
θ_f	0.068	Part-time share of women aged 60–69	0.50	0.50

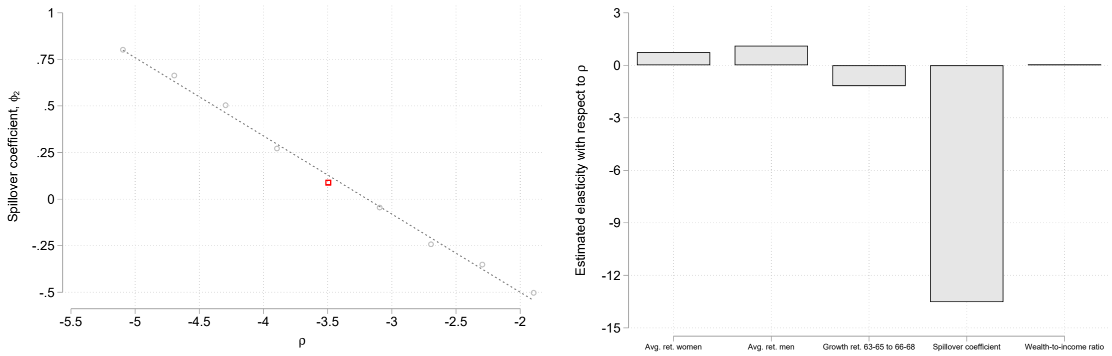
*Source: Norwegian register data.
Notes: Data values for couple parameters are computed on a sample consisting of married couples where at least one spouse is born in the period 1944–1952, both work at age 60. For singles, we use the same birth cohorts and the single individual must work at age 60.

In the case of leisure separability, the model generates a negative spillover effect of -0.36, in contrast to the positive spillover effect of 0.09 with non-separable leisure. This implies that the probability of retirement is decreased by 36 percentage points when your partner is retired. This negative spillover comes entirely from the income effect, as household income is higher when the partner works. Moreover, this highlights that zero spillover coefficient in the reduced-form setting is not necessarily evidence of separability. Substantial complementarity is required to counteract the income effect.

To provide further insight into how ρ is pinned down by the spillover data moment, we show in Figure 3A the sensitivity of the corresponding model moment with respect to ρ. The dots represent the simulated spillover coefficient for different values of ρ, fixing all other parameters at their benchmark values in Table 3. Clearly, the parameter is locally well identified. The figure reveals a

Figure 3: Interaction between ρ and key target moments in the calibration.

(A) Spillover coefficient **(B)** Elasticity of target moments



Notes: Panel 3A shows the model simulated spillover effect for different levels of ρ. The red dot displays our calibrated value of ρ. Panel 3B shows the implied elasticity of each target moment with respect to a one percent change in ρ.

nearly linear relationship between the parameter and the simulated moment. Furthermore, Figure 3B shows the elasticity of all moments with respect to ρ, revealing a particularly tight link between ρ and spousal spillover. The elasticity of the spillover moment is more than twelve times higher than elasticity of other moments.

5.3 Joint Retirement

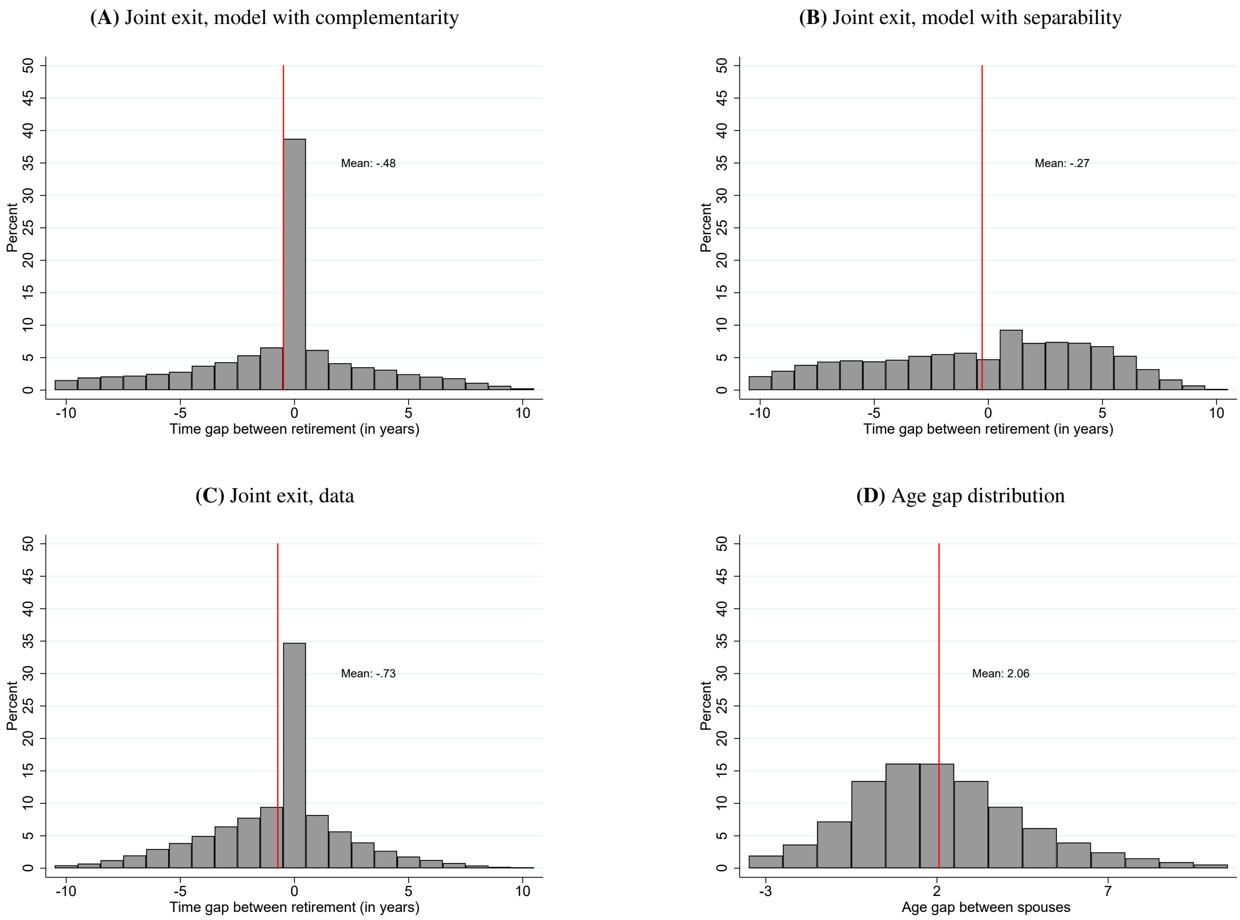
The observation that spouses tend to synchronize their retirement decisions, often leading to a bunching of retirement dates within a year of each other, is a phenomenon that has been noted in empirical studies not only in Norway but also in other countries. For instance Hospido (2015) showed that the same pattern can be found using the Survey of Health, Ageing, and Retirement in Europe, and Banks et al. (2010) showed similar patterns using both UK and US data. We now interpret this stylized fact through the lens of our model.

In Figure 4, we present a comparison between simulated joint retirement from our baseline model, which incorporates leisure complementarity, and the data. Notably, this analysis serves as an out-of-sample test. Remarkably, our model with leisure complementarity successfully replicates the observed phenomenon of synchronized retirement events. This result is particularly consequential because it demonstrates that the model, which quantitatively accounts for the causal spillover effect, also closely aligns with the observed joint retirement behavior in Norway.

Furthermore, in Figure 4 we also display the joint retirement distribution generated by our model with leisure separability. By doing so we gain valuable insights into the critical role of leisure complementarity. It becomes evident that without the inclusion of leisure complementarity, the model fails to replicate data. In fact, due to negative income effects, the model with separability predicts that couples tend to avoid retiring within the same year.[44] Consequently, through the lens

[44]The corresponding figure for the model with separable leisure, when allowing for correlated earnings shocks, is shown in Appendix A. The joint retirement pattern is essentially unchanged when incorporating a correlation of 0.05

Figure 4: Joint exit in the model and data.



Notes: Model with leisure complementarity (panel 4A) and model with leisure separability (panel 4B). Joint exit in data (panel 4C) and the age gap distribution (panel 4D). The red lines show the mean in each distribution. Norwegian register data on all married couples where at least one partner is born in the period 1944–1952. Both spouses must work at age 60.

of the model, the observed joint retirement is entirely driven by leisure complementary.

5.4 Aggregate Implications

In this section, we aim to evaluate the significance of leisure complementarity and joint retirement in the context of the aggregate labor supply elasticity of elderly workers, specifically within the framework of the Norwegian pension reform. In our main results, we assumed that the cohorts born between 1944 and 1952 were part of the pre-reform pension system and introduced the pension reform as a shock in 2011. This approach allowed us to identify the degree of complementarity in leisure. However, to better understand the role of leisure complementarity for overall labor supply elasticity, we now adopt an alternative approach.

We perform this evaluation in two steps. In the first step, we generate two simulated panels of

(in line with the empirical evidence). We also show that even at a counterfactually perfect correlation, the model with separability vastly under-predicts the degree of synchronization.

the entire life cycle of individuals using our baseline model. In the first panel we do not implement the 2011 pension reform, while in the second panel we assume that individuals are born into the post-reform pension system. The long-run reform-induced increase in retirement age can then be assessed by comparing the average retirement rates in the two panels. In the second step of the evaluation, we redo the same exercise for the model with leisure separability.

We find that the increase in average retirement age of 1.5 years in the baseline model is 36 percent larger compared with the model with separability, while the increase in total labor supply in age group 62–69 is 28 percent larger. Consequently, roughly one-third of the labor response is due synchronization of retirement behavior arising from leisure complementarity. To derive these results we simulate an economy in which the share of treated individuals (private sector workers) are as in the 2011 reform.[45]

However, the influence of joint retirement for the aggregate employment response is dependent on whether either or both of the spouses are directly affected by the reform. In the former case, the unaffected individual is only indirectly exposed to the reform. Intuitively, the relative importance of complementarity is larger in these couples compared to couples in which both are directly affected. Our simulation results confirm this intuition. Among couples in which both are directly exposed, complementarity accounts for 12 percent of the labor supply response. By contrast, in couples with only one directly exposed spouse, the importance of complementarity is three times larger (37 percent). The reason is that without complementarity, the indirectly exposed spouse faces only a negative income effect, caused by the increased labor supply of the affected spouse.

6 Conclusion

During the latter part of the 20th century there was a structural change in the composition of households participating in the labor market towards dual-earner households, and only recently have these couples begun to reach retirement age. Both the prevalence of joint retirement in couples, and evidence of causal retirement spillovers, imply that this structural change may have implications for the overall labor supply behavior of the older working population.

In this paper we have assessed this conjecture by interpreting the empirical labor supply interlinkages among couples through the lens of dynamic structural model. The framework is based on a life cycle theory of consumption, savings, and labor supply, incorporating dual-earner households and non-separability in preference for leisure. Central to our approach is that we leverage quasi-experimental evidence of spousal spillovers from a Norwegian pension reform to inform key preference parameters.

Our first finding is that the quasi-experimental evidence is highly informative for the degree of non-separability and suggestive of high levels complementarity. We then use our model to show that the stark correlation in retirement behavior among couples, evident in the prevalence of joint retirement, is critically and causally shaped by leisure complementarities. We do so by comparing

[45]These results are based on couples. Complementarity is less important for the overall labor supply effect if we also include separated individuals (it then accounts for 25 percent of the response).

our baseline model to a model in which we impose leisure separability. The comparison reveals that the baseline model accounts for nearly the entire frequency of joint labor market exit among couples observed in the data. Finally, we assess what these findings imply for the aggregate labor supply response to a pension reform which was designed to incentives workers to postpone retirement. We show that, in the absence of complementarities, the reform is predicted to be less able to stimulate longer working careers. Hence, correctly accounting for interdependent labor supply behavior among couples seems to be important when evaluating policy reforms.

We have highlighted the importance of jointness of labor market activity among couples, motivated by the compositional shift from single to dual-earner households. However, there is another compositional trend known as the "second demographic transition" which may change future family structures (Lesthaeghe, 2010). While the first demographic transition, related to reduced mortality and fertility rates, increases the old-age dependency ratio, the second wave, partly related to rapidly rising divorce and lower marriage rates unfolding in the later part of the 20th century, is predicted to increase the share of single elderly workers and retirees. By explicitly modelling separated individuals, our framework can contribute to the discussion on the potential implications of these changes in family formation. Our calculations show that single individuals respond less than couples to the reform: in the long run, couples increase their labor supply by 16 percentage points more than singles. This result suggest that, due to changing patterns of family formation, aggregate old-age labor supply may become less responsive to policy reforms in the future.

References

Attanasio, O. P., Banks, J., Meghir, C., and Weber, G. (1999). Humps and Bumps in Lifetime Consumption. *Journal of Business & Economic Statistics*, 17(1):22–35.

Attanasio, O. P., Meghir, C., and Santiago, A. (2012). Education choices in mexico: using a structural model and a randomized experiment to evaluate progresa. *The Review of Economic Studies*, 79(1):37–66.

Banks, J., Blundell, R., and Agar, B. (2001). Risk pooling, precautionary saving and consumption growth. *The Review of Economic Studies*, 68(4):757–779.

Banks, J., Blundell, R., and Casanova Rivas, M. (2010). The dynamics of retirement behavior in couples: Reduced-form evidence from England and the US. *UCLA Working paper (July 1, 2010)*.

Blau, D. M. (1998). Labor Force Dynamics of Older Married Couples. *Journal of Labor Economics*, 16(3):595–629.

Blundell, R., Costa Dias, M., Meghir, C., and Shaw, J. (2016a). Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5):1705–1753.

Blundell, R., Pistaferri, L., and Saporta-Eksten, I. (2016b). Consumption inequality and family labor supply. *American Economic Review*, 106(2):387–435.

Borella, M., De Nardi, M., and Yang, F. (2018). The aggregate implications of gender and marriage. *The Journal of Economics of Ageing*, 11:6–26.

Briggs, J., Cesarini, D., Lindqvist, E., and Östling, R. (2021). Windfall gains and stock market participation. *Journal of Financial Economics*, 139(1):57–83.

Capatina, E. (2015). Life-cycle effects of health risk. *Journal of Monetary Economics*, 74(September):67–88.

Casanova, M. (2010). Happy Together: A Structural Model of Couples' Joint Retirement Choices. *UCLA Working Paper*.

Chen, Z., Liu, Z., Suárez Serrato, J. C., and Xu, D. Y. (2021). Notching r&d investment with corporate income tax cuts in china. *American Economic Review*, 111(7):2065–2100.

Coile, C. (2003). Retirement incentives and couples' retirement decisions. Technical report, National Bureau of Economic Research.

Cooley, T. F. and Henriksen, E. (2018). The demographic deficit. *Journal of Monetary Economics*, 93(C):45–62.

Cooley, T. F., Henriksen, E., and Nusbaum, C. (2022). Demographic Obstacles to European Growth.

Cubeddu, L. and Rios-Rull, J.-V. (1997). Marital risk and capital accumulation. Staff Report 235, Federal Reserve Bank of Minneapolis.

Cubeddu, L. and Ríos-Rull, J.-V. (2003). Families as shocks. *Journal of the European Economic Association*, 1(2-3):671–682.

Doepke, M. and Tertilt, M. (2016). Families in macroeconomics. In *Handbook of macroeconomics*, volume 2, pages 1789–1891. Elsevier.

Eckstein, Z. and Lifshitz, O. (2011). Dynamic Female Labor Supply. *Econometrica*, 79(6):1675–1726.

Fernández, R. and Wong, J. C. (2014). Divorce Risk, Wages and Working Wives: A Quantitative Lifecycle Analysis of Female Labour Force Participation. *The Economic Journal*, 124:319–358.

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

French, E. and Jones, J. B. (2011). The Effects of Health Insurance and Self-insurance on Retirement Behavior. *Econometrica*, 79(3):693–732.

Fukui, M., Nakamura, E., and Steinsson, J. (2023). Women, wealth effects, and slow recoveries. *American Economic Journal: Macroeconomics*, 15(1):269–313.

Guner, N., Kaygusuz, R., and Ventura, G. (2012). Taxation and Household Labour Supply. *The Review of Economic Studies*, 79(3):1113–49.

Gustman, A. L. and Steinmeier, T. L. (2000). Retirement in dual-career families: a structural model. *Journal of Labor economics*, 18(3):503–545.

Gustman, A. L. and Steinmeier, T. L. (2004). Social security, pensions and retirement behaviour within the family. *Journal of Applied Econometrics*, 19(6):723–737.

Heathcote, J., Storesletten, K., and Violante, G. L. (2010). The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy*, 118(4):681–722.

Hernæs, E. (2017). Pensjonsordninger og mobilitet. *Frisch rapport*, 2:2017.

Hernæs, E., Markussen, S., Piggot, J., and Røed, K. (2016). Pension Reform and Labor Supply. *Journal of Public Economics*, 142(October):39–55.

Holmøy, E. and Stensnes, K. (2008). Will the norwegian pension reform reach its goals?: an integrated micro-macro assessment.

Hong, J. H. and Ríos-Rull, J.-V. (2012). Life Insurance and Household Consumption. *American Economic Review*, 102(7):3701–30.

Hospido, L. (2015). Pension reform and couples' joint retirement decision. *IZA World of Labor*, 142.

Hurd, M. D. (1990). The Joint Retirement Decision of Husbands and Wives. In Wise, D. A., editor, *Issues in the Economics of Aging*, pages 231–258. University of Chicago Press.

Imrohoroğlu, S. and Kitao, S. (2012). Social security reforms: Benefit claiming, labor force participation, and long-run sustainability. *American Economic Journal: Macroeconomics*, 4(3):96–127.

Kitao, S. (2014). Sustainable social security: Four options. *Review of Economic Dynamics*, 17(4):756–779.

Kruse, H. (2021). Joint retirement in couples: Evidence of complementarity in leisure. *The Scandinavian Journal of Economics*, 123(3):995–1024.

Lalive, R. and Parrotta, P. (2017). How does pension eligibility affect labor supply in couples? *Labour Economics*, 46:177–188.

Lesthaeghe, R. (2010). The unfolding story of the second demographic transition. *Population and development review*, 36(2):211–251.

Low, H. and Meghir, C. (2017). The use of structural models in econometrics. *Journal of Economic Perspectives*, 31(2):33–58.

Low, H., Meghir, C., Pistaferri, L., and Voena, A. (2018). Marriage, Labor Supply and the Dynamics of the Social Safety Net. NBER Working Papers 24356, National Bureau of Economic Research, Inc.

Low, H. and Pistaferri, L. (2015). Disability Insurance and the Dynamics of the Incentive Insurance Trade-Off. *American Economic Review*, 105(10):2986–3029.

Martín, A. R. S. and Marcos, V. S. (2010). Demographic Change and Pension Reform in Spain: An Assessment in a Two-Earner, OLG Model. *Fiscal Studies*, 31(3):405–452.

Merkurieva, I. (2023). Retirement coordination and leisure complementarity. *Labour Economics*, 85:102431.

Michaud, P.-C. (2003). Joint Labour Supply Dynamics of Older Couples. IZA Discussion Papers 832, Institute of Labor Economics (IZA).

Michaud, P.-C., Van Soest, A., and Bissonnette, L. (2020). Understanding joint retirement. *Journal of economic behavior & organization*, 173:386–401.

Michaud, P.-C. and Vermeulen, F. (2011). A collective labor supply model with complementarities in leisure: Identification and estimation by means of panel data. *Labour Economics*, 18(2):159–167.

Rogerson, R. and Wallenius, J. (2013). Nonconvexities, retirement, and the elasticity of labor supply. *American Economic Review*, 103(4):1445–1462.

Stancanelli, E. and van Soest, A. (2012a). Joint Leisure Before and After Retirement: A Double Regression Discontinuity Approach. IZA Discussion Papers 6698, Institute of Labor Economics (IZA).

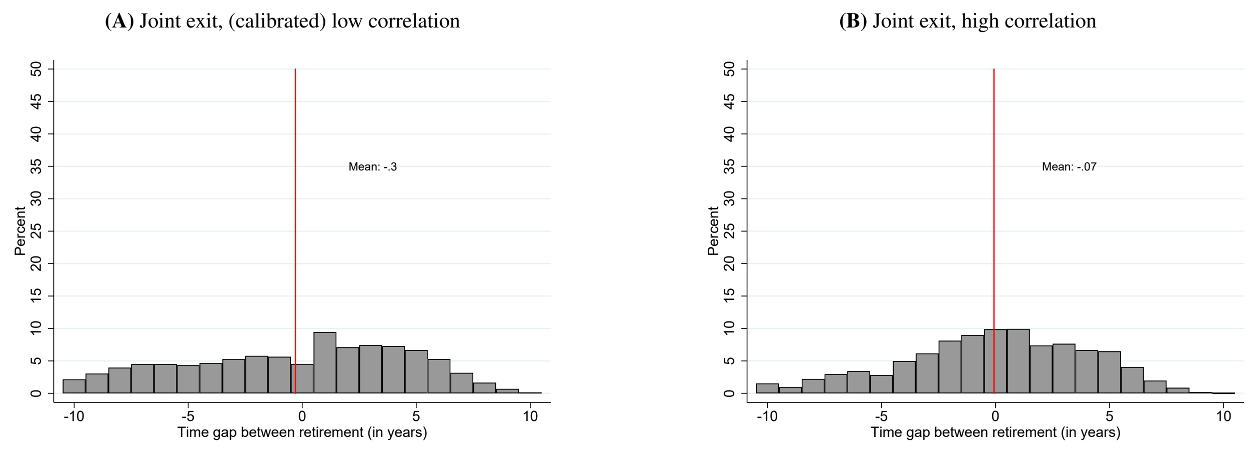
Stancanelli, E. and van Soest, A. (2012b). Retirement and Home Production: A Regression Discontinuity Approach. *American Economic Review, American Economic Association*, 102(3):600–605.

Tauchen, G. and Hussey, R. (1991). Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models. *Econometrica*, 59(2):371–396.

van der Klaauw, W. and Wolpin, K. I. (2008). Social security and the retirement and savings behavior of low-income households. *Econometrica*, 79(6):1675–1726.

A Correlated Shocks

Figure A.1: Joint exit in model with correlated earnings shocks and separable leisure.



Notes: In panel A, the labor earnings shocks have a correlation coefficient of 0.05. In panel B, the couple face perfectly correlated earnings shocks.

In Figure A.1, we illustrate that correlated earnings shocks across the spouses likely cannot explain the characteristic "spike" in retirement synchronization. In panel A.1A we impose a correlation coefficient of 0.05, in line with the empirical correlation in labor earnings fluctuations observed among Norwegian couples around retirement age. The empirical correlation estimate is obtained from a sample of all individuals who are aged 55 to 60 and their respective spouses. We normalize employment at age 55, meaning that everyone starts out as employed. To avoid bias from zero-earnings individuals due to take-up of disability insurance (DI), we disregard couples where either partner receives DI within the time-period. We then compute the following correlations:

$$\rho_{60,55} = \frac{cov(Y_{\Delta_{60,55},w}, Y_{\Delta_{60,55},m})}{\sigma_{\Delta_{60,55},w}\sigma_{\Delta_{60,55},m}} \tag{24}$$

where $Y_{\Delta_{60,55},g}$ is the change in earnings for spouse g from age 55 to 60 and σ is the standard deviation of this, and:

$$\rho_{a+1,a} = \frac{cov(Y_{\Delta_{a+1,a},w}, Y_{\Delta_{a+1,a},m})}{\sigma_{\Delta_{a+1,a},w}\sigma_{\Delta_{a+1,a},m}} \tag{25}$$

where $Y_{\Delta_{a+1,a},g}$ is the change in earnings for spouse g from age a to $a+1$ and σ is the standard deviation of this. We obtain $\rho_{60,55} = 0.054$ and $\rho_{a+1,a} = 0.039$. Re-simulating the model with this degree of correlation, results in hardly any change in the observed degree of joint exit, as depicted in Figure A.1.

In Figure A.1B we show that if we impose perfectly correlated labor earnings shocks, the model generates moderate degree of joint exit, but still far below the pattern in the data.

B Parameter Robustness

In this section, we conduct robustness checks by exploring alternative values for certain exogenous parameters in our economic model. Specifically, we focus on varying the parameters related to leisure (ϕ) and consumption utility curvature (γ). We calibrate our models, both with and without non-separable leisure, using the same procedure outlined in Section 4.2. For the leisure curvature ϕ we consider two alternative values, by adjusting the baseline value up or down by 0.5. For the consumption curvature we vary the parameter across the range of empirical estimates reported in the literature, by setting γ to either 1.35 or 2.

First, in all the four robustness exercises Table B.1 demonstrates that a substantial degree of leisure complementarity is required to align the model's predictions with the observed data. Second, when we vary the leisure curvature parameter (ϕ), we find that the calibrated parameter ρ closely tracks the changes in ϕ. This implies that the overall degree of complementarity, defined as $1-\phi-\rho$, remains relatively stable across different values of ϕ. Notably, Figure B.2 also shows that when assuming separability in leisure the model still fails to capture joint exit. Third, when increasing (decreasing) the consumption curvature parameter γ, we implicitly increase (decrease) the income effect, implying that a stronger (weaker) degree of complementarity is required to match the reduced-form spillover effect of the reform. However, within the range of empirically plausible values for γ, the implications of complementarity for coordinated retirement behavior remain consistent, as depicted in Figure B.2.

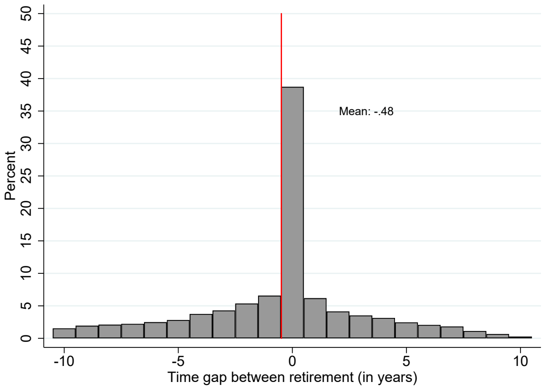
Table B.1: Robustness: degree of complementarity and spillover.

Parameter	$\hat{\rho}$		Spillover	
	Non-separable	Separable	Non-separable	Separable
Baseline	−3.495	−2.0	0.09	−0.36
$\phi=3.5$	−4.072	−2.5	0.09	−0.37
$\phi=2.5$	−2.898	−1.5	0.09	−0.37
$\gamma=2.0$	−3.785	−2.0	0.09	−0.52
$\gamma=1.35$	−3.376	−2.0	0.09	−0.32

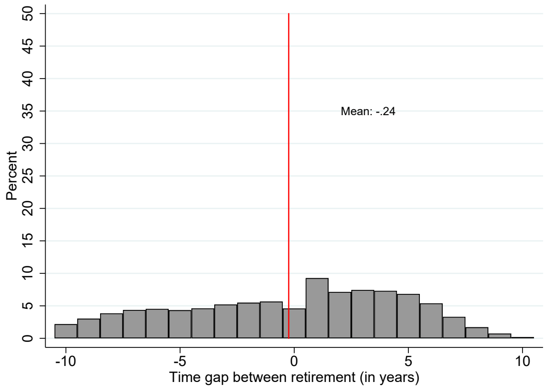
Notes: the table reports the estimated parameter for ρ and the simulated reduced form spillover moment for different calibrations of the model. The first row reports the baseline model, while the next four rows report the results from re-calibrating the models, both with and without separability in leisure, after changing an exogenously set parameter to the value in the first column.

Figure B.2: Joint exit.

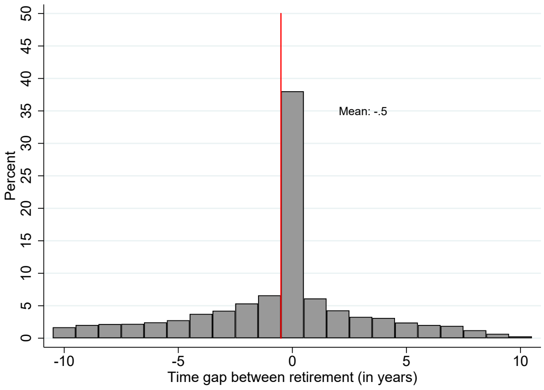
(A) Joint exit, w/ complementarity and $\phi = 3.5$



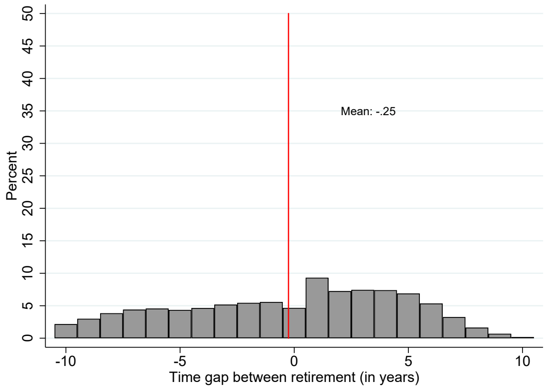
(B) Joint exit, w/ separability and $\phi = 3.5$



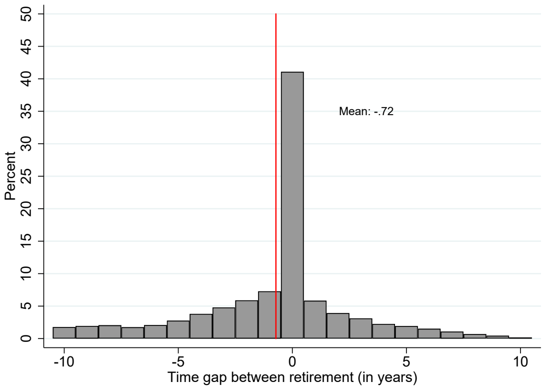
(C) Joint exit, w/ complementarity and $\phi = 2.5$



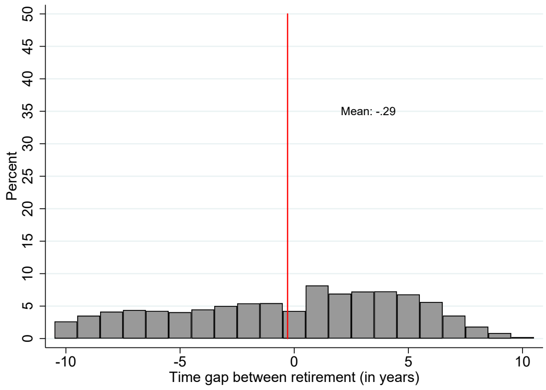
(D) Joint exit, w/ separability and $\phi = 2.5$



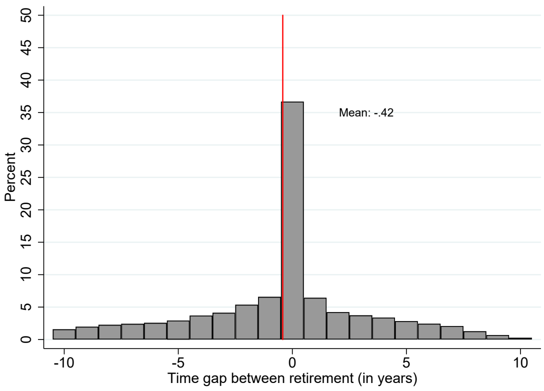
(E) Joint exit, w/ complementarity and $\gamma = 2.0$



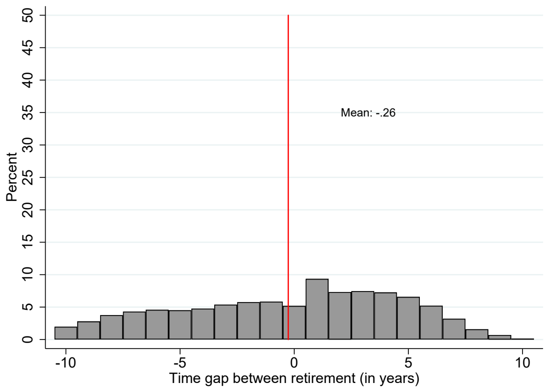
(F) Joint exit, w/ separability and $\gamma = 2.0$



(G) Joint exit, w/ complementarity and $\gamma = 1.35$



(H) Joint exit, w/ separability and $\gamma = 1.35$



C Pension Reform in the Model and Data

In this section we explain the how the central features of the actual 2011 Norwegian pension reforms maps to the reform implementation in the model.

Before 2011, the retirement pension scheme in Norway was based on two main pillars, the public pension from the National Insurance Scheme (NIS) accessible from age 67 and an early retirement pension accessible between age 62-66 for workers affiliated with the AFP program. The workers considered in this paper could then retire with AFP pension from age 62, and then transfer to the NIS pension at age 67. As the AFP and NIS pension were based on the same rules, the pension benefit did not change when transferred to the NIS. Both the AFP and the NIS pension benefits where subject to earnings testing, and pension benefits were not readjusted upwards if retirement is postponed.

The 2011 reform involved changes to both the NIS and AFP pensions. However, the changes to the AFP pension applied only to workers in the private sector (those who had not yet claimed AFP), while public sector workers faced the old AFP system. Under the new system, both the NIS and AFP earnings test was completely abolished. Both the NIS and AFP could be claimed already from age 62, and the AFP pension went from being a temporary benefit between ages 62 and 66 to a life-long pension. Postponement of pension claiming beyond age 62, implied that future annual pensions were readjusted upwards in an actuarial fashion.

As explained in Section 4.2 our model implementation of the reform is performed as an abolishment of the the earnings tests in equation (12) for private sector workers. There are, however, four other key elements in the actual 2011 pension reform implicitly affecting the level of pre-earnings tested pension benefits $\hat{B}_g$, which we abstract from in our model. Under the new pension system, (i) pension benefits are longevity-adjusted, (ii) delayed claiming increases future annual pension benefit, (iii) the AFP pension went from being a temporary benefit until age 67, to becoming a life-long annuity, and (iv) the NIS pension became accessible from age 62.

Regarding (i), the actual pension reform was calibrated such that individuals who were 67 in 2011 (the 1944 cohort) received the same level of life-time benefits as in the pre-reform system. Later cohorts will receive less, roughly proportional to the growth in remaining life expectancy at age 62. Since we focus on the 1944–1952 cohorts our assumption that the pre-earnings tested pension benefit remains unchanged is therefore broadly consistent with the actual reform.

Regarding (ii), delayed pension uptake in the new pension system leads to a roughly actuarial fair readjustment of future annual pension benefit, leaving the expected lifetime pension approximately neutral with respect to the timing of pension claiming. We therefore abstract from a pension claiming decision, assuming that individuals exposed to the reform never delay claiming beyond age 62.

When claiming pension at age 62, the pension benefits in the new system are the sum of actuarial adjusted life-long NIS and AFP pensions. Given (i) and (ii), the post-reform pension benefit level for the cohorts considered in this paper is then essentially identical to what they would have received in the pre-reform system.

Regarding (iv), the lowering of the NIS statutory retirement age mean that public sector workers, whose AFP pension was not reformed, can in principle combine a full NIS pension and full-time work from age 62. However, this is a very costly option for public sector workers, because early claiming of the new NIS pension implies losing AFP pension eligibility. We abstract from the option in the model, implying that public sector workers cannot claim an NIS pension before age 67. In reality, the early claiming of an NIS pension is only a relevant option for workers who do not plan to retire before 67 and value the additional liquidity. Thus it will be of second-order importance when deciding when to retire.

Finally, in the model we assume that public sector workers face the pre-reform NIS pension earnings test between ages 67 and 69 also after 2011. In reality, they could claim an untested NIS pension. However, in addition to the NIS pension, public sector workers receive an occupational pension benefit (OTP) from age 67. This occupational benefit was not reformed, and as a result, it remained costly for public sector workers to postpone retirement beyond age 67 also after 2011.