UNIVERSITY OF COPENHAGEN
DEPARTMENT OF ECONOMICS

Master's Thesis

Alexander Krikonis

Employment and Retirement Decisions in a Couples Context: Evidence from raising the Danish Early Retirement Age

Supervisor: Jakob Egholt Søgaard
Submitted on: May 14, 2021
Keystrokes: 143,676 (60 pages)

Employment and Retirement Decisions in a Couples Context: Evidence from raising the Danish Early Retirement Age*

Alexander Krikonis†

May 14, 2021

Abstract

This thesis investigates if leisure complementarities cause couples to retire jointly by exploiting two Danish reforms that raised the early retirement age. Utilizing a regression discontinuity design, I find that raising the focal partners' retirement age increases the spouses' employment by 2.6 to 3.7 percentage points in reduced-form. This corresponds to a spillover effect from the change in the focal partners' employment to the change in the spouses' employment of 16.5 to 18.3 percentage points. Hence, the estimates suggest that the spouse's employment decision depends on the focal partner's employment decision due to complementarities in leisure. The conclusions of this thesis emphasize that policymakers have to account for leisure complementarities, as ignoring the spillover effect from the focal partners' employment to the spouses' employment underestimates the positive effect on the government budget of raising the retirement age. Calculations indicate that ignoring leisure complementarities may underestimate the positive impact on the government budget of raising the retirement age by 100 million DKK for the first group of couples affected by the reforms.

*I would like to thank my supervisor, Jakob Egholt Søgaard, for valuable comments and feedback. Furthermore, I am grateful to the Danish Ministry of Finance for providing access to the administrative data. All errors are my own.

†Department of Economics, University of Copenhagen. Email: brz451@alumni.ku.dk.

Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Name	Definition
ERP scheme	early retirement pension scheme	a scheme where an individual can retire before the national retirement age if being a member
ERA	early retirement age	age at which an individual is eligible to claim retirement benefits if being a member of the early retirement pension scheme
ERB	early retirement benefits	benefits an individual can claim upon reaching the early retirement age if being a member of the early retirement pension scheme
NRP scheme	national retirement pension scheme	a scheme where every individual who has lived in Denmark for more than three years can retire on
NRA	national retirement age	age at which an individual can claim pension benefits
NRB	national retirement benefits	benefits an individual can claim upon reaching the national retirement age
RDD	regression discontinuity design	a research design where the assignment to treatment is based on a cutoff and a running variable
SRDD	sharp regression discontinuity design	the regression discontinuity design where the assignment to treatment is a deterministic function of the running variable
FRDD	fuzzy regression discontinuity design	the regression discontinuity design where the explanatory variable is instrumented by a variable where the assignment to treatment is a deterministic function of the running variable

1 Introduction

Longer life expectancy and broader access to pension benefits have caused an increased number of individuals to rely on pension benefits in most Western countries, including Denmark. As a result, multiple Western countries have raised the retirement age[1] in recent years or will do so soon to improve the government budget by reducing public expenditures to pension schemes and increasing the tax revenue by increasing employment (OECD, 2011). Estimates of how raising the retirement age affects employment rates, public expenditures, and tax revenue have mainly concerned the direct effect but ignored possible second-order effects (Jørgensen, 2014). A potential second-order effect includes the spillover effect of employment in couples, as raising the focal partners' retirement age may also increase the spouses' employment if couples prefer to retire jointly.[2] In this case, the positive effect on the government budget of raising the retirement age is underestimated as spillover effects are currently ignored, implying that policymakers may not allocate resources efficiently.

Several studies have investigated whether such spillover effects exist. Hurd (1990) was the first to emphasize that spouses value joint leisure, causing couples to prefer to retire jointly. The literature refers to this as *leisure complementarities*, as the spouse's utility of leisure increases when the focal partner retires and enjoys more leisure time, causing the spouse's leisure and the focal partner's leisure to be complements. Studies have reached different conclusions on whether couples retire jointly due to leisure complementarities, and the findings have proven to depend on the type of pension scheme, suggesting that the estimates are country-specific (Coile, 2004; Legendre et al., 2018). The aim of this thesis is to investigate whether couples retire jointly because of complementarities in leisure and whether raising the focal partners' retirement age thus increases the spouses' employment.

Studies in the literature of joint retirement often fail to obtain causal estimates, as they exploit within-country variation or cross-country variation where the identifying assumptions are likely to be violated, implying that the estimates cannot be given a causal interpretation (Atalay et al., 2019; Banks et al., 2010; Coile, 2004). I overcome these concerns by considering a quasi-experimental approach where I exploit two Danish reforms that raised the early retirement age (ERA) four times, each time by six months. Hence, individuals born just after four different cutoff dates were able to claim early retirement benefits (ERB) six months later than if they were born just the day before. If the spouse's leisure and the focal partner's leisure are complements, the spouse's utility of leisure is lower if the focal partner postpones retirement, creating an

[1] In this thesis, the term "the retirement age" refers to the age at which individuals are eligible to claim pension benefits, and not the age at which they actually retire.

[2] Each couple consists of a *focal partner* and a *spouse*. If spillover effects exist, policy interventions aimed to affect the focal partner will also affect the spouse.

incentive of the spouse to postpone retirement as well. In this case, the spouses' employment will increase if the focal partners' employment increases as a result of the increase in the ERA since working can be seen as the opposite of retiring.[3]

To investigate the implications of raising the retirement age, I define a theoretical model, which reveals that the spouses' employment responds to the focal partners' employment if spouses value joint leisure. Furthermore, the model shows that, although all spouses value joint leisure similarly, only older spouses respond to the focal partner's employment decision, as younger spouses experience high economic costs of retiring prior to the retirement age.

Utilizing a regression discontinuity design (RDD) to test the predictions of the theoretical model, couples are assigned to treatment if the focal partners' ERA increases and to the control group otherwise. Using high-quality administrative data in which the employment of each individual in the population is observed monthly, I compare the spouses' employment along the extensive margin in each of the six months in which the focal partners in the control group and the treatment groups are the same age in months, but where only focal partners in the control group are eligible to claim ERB.[4]

I find that raising the retirement age increases the spouses' employment by 2.6 to 3.7 percentage points in reduced-form. This corresponds to a spillover effect from the focal partners' employment to the spouses' employment of 16.5 to 18.3 percentage points. Ignoring leisure complementarities thus underestimates the increase in employment by 14 to 15 percent by raising the retirement age. The estimates are robust to different specifications, confirming the validity of the results. I note that couples retire jointly by waiting until both partners are eligible to claim pension benefits. Therefore, couples' age structure explains why men are more likely to respond to their wife's employment decision, as men are, on average, older than their wife. This suggests that even if spouses value joint leisure similarly, economic incentives cause the couple's age structure to influence how much the spouse responds to the focal partner's employment decision following the predictions from my theoretical model.

Calculations indicate that ignoring leisure complementarities underestimates the effect of raising the retirement age by 18.9 to 99.6 million DKK for the first group of couples affected by the reforms. Hence, the conclusions of this thesis suggest that policymakers have to account for leisure complementarities to ensure that resources are allocated efficiently.

This thesis contributes to the literature of joint retirement by providing causal evidence that complementarities in leisure cause couples to retire jointly, which has been called for in

[3] Although studies investigate whether couples retire jointly, being employed or in the labor force usually serves as the outcome variable in the literature of joint retirement. See, e.g., Lalive & Parrotta (2017), Atalay et al. (2019), Baker (2002), or Cribb et al. (2014).

[4] Changes in the employment can be considered along the extensive and intensive margin. The former relates to the decision to work or not, while the latter relates to the choice of how many hours to work.

recent years (Atalay et al., 2019). The eligibility rules of the early retirement pension (ERP) scheme ensure that the focal partners who can claim ERB at an older age remain employed. Furthermore, the focal partners' ERB are not means-tested based on the spouses' income, and the spouses do not receive pension benefits when the focal partners retire. Therefore, only complementarities in leisure can explain if the spouses postpone retirement and remain employed due to the increase in the ERA. Hence, it is possible to distinguish between the impact of economic incentives and complementarities in leisure to explain why couples retire jointly, which has been a concern in other studies.[5] Both partners in the majority of older Danish couples are members of the ERP scheme, implying that the estimates may be extrapolated to the general population. The estimation strategy and the data used are ideal for investigating whether leisure complementarities cause couples to retire jointly. The large sample, the similar retirement age for men and women, and the small gender gap in employment also make it possible to properly investigate whether the genders value joint leisure differently conditional on the couple's age structure.[6] My work is likely to be built upon in future research to investigate other possible second-order effects to obtain a broader picture of the benefits and consequences of raising the retirement age.

The rest of this thesis is structured as follows: Section 2 reviews the related literature, while the institutional setting is described in section 3. Section 4 presents the theoretical framework, explaining why spouses may respond to the focal partner's employment decision. The empirical strategy is outlined in section 5, and a description of the data including the sample restrictions and summary statistics is presented in section 6. I present the results in section 7, which is followed by calculations of how the government budget is affected by ignoring leisure complementarities in section 8. Finally, section 9 contains a discussion of the results, while section 10 concludes on the findings of this thesis.

2 Related literature

Hurd (1990) was the first to stress that complementarities in leisure cause couples to retire jointly. He emphasized the importance of this conclusion, as an increased number of individuals claimed pension benefits when the paper was published. A growing number of studies have since investigated if leisure complementarities cause couples to retire jointly by different empirical approaches. In the 1990s and 2000s, structural models such as dynamic programming models

[5] See, e.g., Blau & Gilleskie (2006), Kapur & Rogowski (2007), Gustman & Steinmeier (2004), or van der Klaauw & Wolpin (2008).

[6] The gender gap in employment is calculated by the use of data from OECD (2021). Few women in the labor force is a concern in, e.g., Hurd (1990), Baker (2002), and Sand et al. (2019), as researches cannot investigate whether men and women value joint leisure differently.

and discrete choice models gained traction and were highly applied. In recent years, researchers have utilized reduced-form models and exploited within-country variation caused by changes in eligibility rules.

This thesis utilizes a reduced-form model, exploiting two Danish reforms that raised the early retirement age of both men and women, causing the new retirement ages to depend on the date of birth. The spouses do not receive pension benefits when the focal partners retire, nor are the focal partners' pension benefits means-tested based on the spouses' income. I exploit that the assignment to treatment is exogenous, and the lack of economic incentives of the spouses to retire jointly with the focal partners ensure that only complementarities in leisure can explain if couples retire jointly. This has been a concern in various studies as argued in this section. Similarly, I estimate the response conditional on gender, which has also been impossible in multiple studies because only a small number of women were in the labor force or since only women's retirement age increased.

2.1 Empirical approaches

Structural models

Structural models have been utilized, as the models allow the researcher to estimate utility parameters. The structural models allow distinction between different factors, such as leisure complementarities, economic incentives, health, etc. (Coile, 2004). Furthermore, changes in eligibility rules can be simulated in the models, providing an idea of how policy interventions influence the propensity to retire jointly. The approach enables Zweimüller et al. (1996), Blau (1998), and Gustman & Steinmeier (2004) to conclude that leisure complementarities cause couples to retire jointly. The former paper finds that raising the retirement age of women in Austria postpones their spouses' retirement and that the spillover effect constitutes half of the direct effect. The other latter two papers obtain estimates within the same range, both considering data from the United States.

Coile (2004) and Bloemen et al. (2015) argue that structural models require strong assumptions of the model's parameters, and that it is challenging to ensure that these are correct, which is a drawback. Furthermore, they state that structural models predict the retirement behavior assuming the absence of uncertainty and that many of the models fail to include all channels in which the focal partner and spouse are linked.

Reduced-form models

Coile (2004) and Bloemen et al. (2015) stress that reduced-form models are desirable, as they are more transparent, and as it is in most cases possible to examine whether the identifying assumptions are satisfied. Three different sources of variation have been used to investigate if complementarities in leisure cause couples to retire jointly: Cross-country variation resulting from differences in eligibility rules between countries, within-country variation caused by changes in retirement ages or incentives to retire, and the regression discontinuity design where the retirement age serves as the cutoff, and the focal partners' age is the running variable.

Reduced-form models: cross-country variation

Banks et al. (2010) consider men's transition out of work by exploiting the differences in women's eligibility rules between the United Kingdom, where the retirement age is 60 years, and the United States, where the retirement age is 62 years. They obtain an insignificant estimate by considering the entire sample. Upon restricting the sample to couples in which the wife is younger than the husband, they state that British men are 15 percentage points more likely to retire if their wife is eligible to claim pension benefits. Using an IV strategy, they find that British men are 46 to 64 percentage points more likely to retire if their wife has retired.

Hospido & Zamarro (2014) and Legendre et al. (2018) utilize the bivariate probit model in combination with the regression discontinuity design in a multi-country setting. Both papers find that leisure complementarities cause couples to retire jointly. The former paper finds that only women respond to their husband's eligibility rules, while the latter paper reaches the opposite conclusion.

A limitation of utilizing cross-country variation is that only the retirement age or the incentives to retire are allowed to differ between countries. This assumption is likely to fail, as couples in different countries may value (joint) leisure differently or face different social norms.

Reduced-form models: within-country variation

Within-country variation is exploited by Johnsen et al. (2015) and Kruse (2020) who utilize the difference-in-difference design both considering Norwegian data. The studies use different reforms, but both conclude that complementarities in leisure cause couples to retire jointly. They observe asymmetric responses, as women respond to their husband's eligibility rules, while men do not. Furthermore, women retire jointly with their husband on pension benefits and disability pension benefits. The estimated spillover effect is approximately 20 percentage points in both papers. Bloemen et al. (2015) consider a Dutch reform where individuals who had worked more than ten years in the public sector were eligible to claim pension benefits at a younger age. Utilizing a difference-in-difference design, they find that men are 25 percentage points more likely to retire if their wife retires.

All three papers exploit reforms that lowered the retirement age of individuals working in specific firms or sectors, and the papers consider the spouses' employment conditional on whether the focal partners are eligible to claim pension benefits at a younger age. Individuals who value (joint) leisure may choose to work in the firms or sectors where the retirement age was lowered, implying that couples may be different conditional on whether the focal partners work in these firms or sectors. This is a threat to identification, questioning the causality of the estimates in all three studies.

Cribb et al. (2014) consider a British reform, which gradually raised women's retirement age from 60 to 65 years. They conclude that the husbands' employment increases by 4 to 5 percentage points if their wife's retirement age increases. Atalay et al. (2019) exploit that the same type of reform was introduced in the 1990s in Australia and estimate the spillover effect to the spouses' labor force participation to be between 7 and 17 percentage points. They find that men are more likely to be a part of the labor force if their wife's retirement age increases and that the response is particularly strong for men who are older than their wife. Due to the reforms, the papers consider only the response by men. Should the genders respond asymmetrically, e.g., due to different utilities of joint leisure, these results cannot be extrapolated to the general population.

Coile (2004) considers couples in the United States where the focal partners have different economic incentives to retire and investigates how likely the spouses are to retire jointly with the focal partners. She stresses that ignoring the spillover effect from the focal partners' employment to the spouses' employment underestimates the impact of raising the retirement age by 13 to 20 percent. The spouses also receive pension benefits when the focal partners retire, meaning that the spouses have an economic incentive to retire jointly with the focal partners. Hence, couples' joint retirement cannot solely be attributed to complementarities in leisure.

Reduced-form models: the regression discontinuity design with a fixed retirement age

Lalive & Staubli (2014), Lalive & Parrotta (2017), and Stancanelli (2017) utilize a regression discontinuity design as they observe a sharp discontinuity in the focal partners' employment at the retirement age. The former paper uses the baseline regression discontinuity design, while the two latter papers utilize a double regression discontinuity design. All studies find evidence that leisure complementarities cause couples to retire jointly. However, Lalive & Staubli (2014) note that the estimates are susceptible to the functional form when using a regression discontinuity design based on a fixed retirement age. This implies that the true estimates may not have been obtained since the functional form must properly approximate the true conditional expectation function. Although the double regression discontinuity design is attractive in this context, much work remains to be done by imposing bandwidth and density tests when considering two discontinuities (Lalive & Parrotta, 2017).

Heterogeneous responses

Economic motives have a substantial impact on couples' incentives to retire jointly, and the design of pension schemes is therefore essential. Legendre et al. (2018) conclude that joint retirement is usually achieved by the older partner postponing retirement until the younger partner is eligible to claim pension benefits since it is more costly to retire before the retirement age. If pension benefits are means-tested based on the couple's total income, the spouse's incentive to retire jointly with the focal partner increases, causing couples to be more likely to retire jointly (Gustman & Steinmeier, 2004; van der Klaauw & Wolpin, 2008). The healthcare system's design also influences the motivation to retire jointly; if the focal partners' healthcare insurance schemes also cover younger spouses who retire although they have not reached the retirement age, the incentive to retire jointly is increased (Blau & Gilleskie, 2006; Kapur & Rogowski, 2007). The differences in pension benefits and healthcare systems constitute yet another argument that estimates are country-specific.

Michaud et al. (2019) argue that the partner with the highest income has the greatest bargaining power, meaning that spouses are more likely to respond to the focal partners' retirement decision if the spouses earn less than the focal partners. Furthermore, they stress that couples' identical preferences are more likely than leisure complementarities to explain why couples retire jointly. Individuals tend to marry someone who has the same preferences as themselves, which may imply that both partners prefer to retire within the same period independently of their spouse's retirement decision. Hamermesh (2002) and Stancanelli & Van Soest (2016) conclude that couples coordinate their work calendar prior to retirement and spend more time together as retired, arguing that couples value joint leisure and thus retire jointly due to complementarities in leisure.

Bhatt (2017) stresses that younger cohorts are less likely to retire jointly due to socio-economic, employment, and health-related factors. An important insight is that the social norms are different for younger cohorts, as women are more likely to have built their careers, and men are more likely to do domestic work than older cohorts. This implies that spouses in younger couples depend less on the focal partners, causing younger cohorts to be less likely to retire jointly.

Summary

Most papers conclude that couples retire jointly and that leisure complementarities explain these findings. The estimates differ considerably, and it is likely to be a consequence of different designs of the pension system, as economic incentives also affect the incentive to retire jointly. Recent studies have emphasized that the couples' age structure is essential, as older spouses are more likely than younger spouses to retire jointly with the focal partners. Furthermore, the literature is inconsistent on whether or not genders value joint leisure similarly. Many of the papers have failed to obtain causal estimates because of identifying assumptions that are likely to be violated or pension systems where economic incentives induce couples to retire jointly.

2.2 Studies using Danish data

High-quality administrative data covering a large part of the Danish population have implied that several studies have investigated if leisure complementarities cause Danish couples to retire jointly (An et al., 2004; García-Miralles et al., 2021; Jørgensen, 2014; Kallestrup-Lamb, 2011). An advantage of using Danish data is to avoid the self-report bias that may arise if applying surveys, as well as the possibility to follow the same individual in multiple periods (Kallestrup-Lamb, 2011).

An et al. (2004) utilize a multivariate mixed proportional hazard model, considering annual employment data from 1980 to 1990. They consider a small number of couples, as the initial sample covers only 0.5 percent of the Danish population. Furthermore, the problem of right-censoring causes them to exclude many couples, as they only include surviving couples. Concluding that couples tend to retire jointly, they cannot distinguish between the impact of leisure complementarities and identical preferences. The paper finds that husbands and wives respond in the same way to the focal partner's retirement decision. Kallestrup-Lamb (2011) reaches the opposite conclusion, as she argues that men are more likely to respond to their wife's retirement decision. She utilizes a semi-parametric single risk grouped duration proportional hazard model, considering data from 1985 to 2001.

Applying a dynamic structural model using data from 1996 to 2008, Jørgensen (2014) estimates the spouses' value of leisure to be twice as large if the focal partners have retired. Simulating the Retirement Reform of 2011, he finds that the reform underestimates the impact on the government budget by 7 billion DKK by ignoring leisure complementarities, as spouses also tend to retire later due to the reform. This increases the reform's positive effect by 40 percent more than predicted by the Danish Ministry of Finance.

García-Miralles et al. (2021) utilize a regression discontinuity design where they consider couples in which both partners were born in 1953 at the latest and are members of the ERP scheme such that the retirement age remains stable at 60 years. Observing the focal partners' and the spouses' propensity to retire when the focal partners reach the ERA, they find a spillover effect of 7.5 percent. Exploiting two reforms that raised the focal partners' ERA (the same reforms as considered in this thesis) and utilizing a difference-in-difference design, they find a spillover effect of 9 percent. In both settings, they conclude that joint retirement is achieved by couples postponing retirement until both partners are eligible to claim pension benefits.

Summary

All four studies find that couples retire jointly due to leisure complementarities. However, while Jørgensen (2014) finds a large effect, the estimates by García-Miralles et al. (2021) are within a much smaller range. Furthermore, the papers disagree on whether or not the genders value joint leisure similarly.

3 Institutional setting

The Danish pension system consists of five pillars: public pensions, labor market pensions, private pension savings, savings through real estate, and a residual group (Hastrup, 2020). In recent years, policymakers have encouraged the Danes to increase payments to labor market pensions and private pension savings to reduce government expenditures and increase the economic well-being of the elderly. The elderly of today, however, still rely on public pensions, implying that the retirement age has a substantial impact on the decision to work.

3.1 Raising the retirement ages

Denmark has two different public pension schemes: the early retirement pension (ERP) and the national retirement pension (NRP). The early retirement age (ERA) and the national retirement age (NRA) was 60 and 65 years, respectively until 2005, both not changed since the beginning of the millennium. The Welfare Agreement of 2006 resulted in gradual increases in the ERA and the NRA beginning in 2014 and 2019, respectively. People born in 1959 would be the first to be affected by this (The Danish Ministry of Finance, 2006). As a result of greater increases in life expectancy than estimated, the Welfare Agreement of 2006 was followed up by the Retirement Reform of 2011. This reform brought the increases in both the ERA and the NRA forward by five years, meaning that individuals born in 1954 would be the first to be affected by the increase in the ERA and NRA (The Danish Ministry of Finance, 2011).

The Parliament agreed to raise the ERA and the NRA gradually by six months based on the date of birth, as depicted in table 3.1. In recent years, most individuals have retired on ERB, causing me to focus on the increase in the ERB in this thesis (The Agency of Labor Market and Recruitment, 2020). Furthermore, I consider individuals born in 1955 at the latest, as individuals born after 1955 are eligible to claim fewer years of ERB and will be paid a lower ERB, as all types of pensions are deducted. This implies that the incentive to claim ERB is greater if born between 1953 and 1955.

The Welfare Agreement of 2006 and the Retirement Reform of 2011 generate exogenous variation in the retirement age, as individuals cannot manipulate their date of birth. The

Table 3.1: Retirement ages conditional on the date of birth

Cohort (month)	Early retirement age	Two-year rule	National retirement age
-1953 (12)	60	62	65
1954 (1) - 1954 (6)	60.5	62.5	65.5
1954 (7) - 1954 (12)	61	63	66
1955 (1) - 1955 (6)	61.5	63	66.5
1955 (7) - 1955 (12)	62	64	67
1956 (1) - 1956 (6)	62.5	64	67
1956 (7) - 1958 (12)	63	64	67
1959 (1) - 1959 (6)	63.5	64	67
1959 (7) - 1962 (12)	64	✗	67
1963 (1) - 1966 (12)	65	✗	68

Note: The early retirement age and the national retirement age are not reported for cohorts born after 1966, as they are not passed by the Parliament yet. The two-year rule is explained in section 3.2. The two-year rule cannot be applied by people born after July 1, 1959. The retirement ages are retrieved from The Dane Age Association (2020).

unanticipated advance in the ERA of individuals born between 1954 and 1958 (who would be able to claim ERB at age 60 in the absence of the Retirement Reform of 2011) leaves little room for optimizing labor and pension planning since the cohorts were advised of their new retirement age few years before reaching the ERA.

3.2 The early retirement pension scheme

The ERP scheme was introduced in 1979 as a path to retirement prior to the NRA for individuals from mainly the lower class and the lower-middle class suffering from poor health caused by hard physical work (Petersen, 2011). The ERP scheme quickly became popular, and few years after the implementation, more than three times as many as expected claimed ERB (Kallestrup et al., 2000). As a result, the requirements to be eligible to claim ERB have been tightened several times.

Since I only consider individuals born from 1953 to 1955, I describe the rules that apply to these cohorts. One has to have been a member of an unemployment insurance fund for at least 20 of the last 25 years before one's 60th birthday. Furthermore, one must have made payments to the ERP scheme within the period to be eligible to claim ERB. One was automatically assigned to the ERP scheme if being a member of an unemployment insurance fund, which may explain why the scheme became so popular.[7]

One is eligible to claim ERB on the day where (s)he reaches the ERA. It is impossible

[7] People had the possibility to opt-out of the scheme, but several studies suggest that individuals take their pension plan as given although they would increase utility by following another plan. See, e.g., Chetty et al. (2014), Gustman & Steinmeier (2002), Gustman & Steinmeier (1989), or Mitchell (1988).

to receive the full benefit in the month where (s)he reaches the ERA, and wage earners are usually required to wait until the end of the month if they want to leave their job, causing most individuals to wait until the month after they reach the ERA to retire. I define an individual to be eligible to claim ERB if (s)he can claim ERB every day of the month, i.e., if (s)he has reached the ERA in earlier months. Therefore, I define the month after one turns 60 years as the first month where individuals born before 1954 are eligible to claim ERB. To be eligible to claim ERB, one has to be in the labor force upon reaching the ERA, implying that it is very costly to retire before reaching the ERA, as one will waive five years of ERB.

The ERP scheme is not seen as a savings account since the ERB does not depend on the amount paid to the scheme through life (as long as people have paid in the required period). Instead, the ERP scheme functions as a sort of membership scheme allowing for people who have made payments before reaching the ERA to receive ERB upon reaching the ERA. The maximum monthly take-up of ERB is 17,400 DKK in 2020-prices, and the ERB is means-tested based on labor income and pension savings regardless of whether the latter has been paid out. The focal partners' ERB are not means-tested based on the spouses' income, causing the focal partners' ERB to be independent of whether the spouses work. Similarly, spouses do not receive any pension benefits when the focal partners retire.

If one does not claim their ERB within the first two years upon reaching the ERA, and subsequently works at least 30 hours a week in this period, the monthly take-up of ERB increases to 19,000 DKK, known as the *two-year rule*. Furthermore, this rule states that only paid out pension savings are deducted from the ERB. Therefore, the rule creates an incentive to postpone retirement and remain employed. The increase in the ERA also raised the age at which one is eligible to apply the two-year rule as reported in table 3.1.

4 Theoretical framework

Upon outlining the eligibility rules of the ERP scheme in section 3, it is evident why raising the ERA is expected to increase the focal partners' employment. However, the eligibility rules cannot explain if the increase in the focal partners' employment causes the spouses' employment to increase as well. This section defines a theoretical model that demonstrates why leisure complementarities cause couples to retire jointly, implying that the spouses' employment responds to the focal partners' employment. Furthermore, I investigate how the couples' age structure influences the decision to retire jointly.

4.1 The theoretical model

A couple has often been assumed to behave as one agent in the economic literature. This assumption cannot be imposed to explain why couples retire jointly, as it is not reasonable to assume that the partners have identical preferences (Browning et al., 1994; Chiappori, 1988). Instead, I consider a couple that consists of a focal partner (i) and a spouse (j) where the spouse gets utility from consumption (C) and leisure (L_j). Chiappori (1988) argues that couples consist of either egoistic or altruistic partners. In the former case, partners do not cooperate, while they cooperate to maximize the couple's utility in the latter case. I consider a hybrid between the two following Lalive & Parrotta (2017) and Hiedemann et al. (1998). Therefore, I define a simplified model in which the spouse maximizes utility, holding the focal partner's behavior fixed, as the spouse cannot influence whether the focal partner works.

The spouse's utility depends on whether the focal partner and spouse work. Define $y_i = 1$ if the focal partner is employed and 0 otherwise. Similarly, define $y_j = 1$ if the spouse is employed and 0 otherwise. The spouse can choose whether or not to work but cannot choose the number of working hours following Selin (2014). There are no frictions on the labor market in the model, implying that the spouse cannot be involuntarily unemployed. The utility of the spouse is given by

$$U_j(y_i, y_j) = C + L_j \tag{4.1}$$

Assuming that the couple cannot loan or save money, the consumption equals the couple's total income.[8] The spouse receives labor income (w_j) if working and a benefit (b_j) if not working. I assume that labor income and benefits are taxed by the same average rate (τ) where $0 < \tau < 1$. Hence, the spouse's utility of consumption equals

$$C = y_i(1-\tau)w_i + (1-y_i)(1-\tau)b_i + y_j(1-\tau)w_j + (1-y_j)(1-\tau)b_j \tag{4.2}$$

Assuming that the spouse is compensated by α_j relative to the wage if not working, the benefit is given by $b_j = \alpha_j w_j$ where $0 \leq \alpha_j < 1$ such that the benefit is lower than the wage, but cannot be negative. Inserting this into 4.2 yields

$$C = y_i(1-\tau)w_i + (1-y_i)(1-\tau)\alpha_i w_i + y_j(1-\tau)w_j + (1-y_j)(1-\tau)\alpha_j w_j \tag{4.3}$$

The spouse does only get utility of leisure if not working.[9] In this case, the spouse's utility of

[8] This assumption does not have an impact on the conclusions, meaning that the assumption is not critical.

[9] Considering the ordinal utility approach in which the utility can only be used to rank alternatives, it is equivalent to assuming that the spouse gets disutility from working. This assumption is frequently imposed in the literature of labor economics and public finance and is based on economic theory, as people in the absence of labor market work can do domestic work or enjoy leisure time. Examples include Cahuc et al. (2014) and Hindriks & Myles (2013).

individual leisure is given by ϕ_j where $0 < \phi_j < 1$. If the focal partner does not work either, the spouse's utility of leisure is given by $\phi_j + \sigma_{ij}\phi_j$ where $\sigma_{ij} \geq 0$, and the latter term thus denotes the spouse's utility of joint leisure with the focal partner. In the simple case where $\sigma_{ij} = 0$, the spouse does not value joint leisure, meaning that the utility of leisure does not depend on whether the focal partner works. If $\sigma_{ij} > 0$, the spouse values leisure more if the focal partner does not work either, causing the spouse's leisure and focal partner's leisure to be complements. The spouse's utility of leisure can thus be expressed by

$$L_j = (1 - y_j)\phi_j + (1 - y_j)(1 - y_i)\sigma_{ij}\phi_j \tag{4.4}$$

Inserting 4.3 and 4.4 into 4.1, the spouse's utility can be expressed by

$$\begin{aligned} U_j(y_i, y_j) &= C + L_j \\ &= [y_i(1 - \tau)w_i + (1 - y_i)(1 - \tau)\alpha_i w_i + y_j(1 - \tau)w_j + (1 - y_j)(1 - \tau)\alpha_j w_j] \\ &+ [(1 - y_j)\phi_j + (1 - y_j)(1 - y_i)\sigma_{ij}\phi_j] \end{aligned} \tag{4.5}$$

4.1.1 Solving the model

There are four possible outcomes:

1. Both the focal partner and the spouse work ($y_i = 1$, $y_j = 1$).
2. The focal partner works, and the spouse does not work ($y_i = 1$, $y_j = 0$).
3. The focal partner does not work, and the spouse works ($y_i = 0$, $y_j = 1$).
4. Both the focal partner and the spouse do not work ($y_i = 0$, $y_j = 0$).

The spouse's utility in the four outcomes is given by

1. $U_j(1, 1) = (1 - \tau)(w_i + w_j)$
2. $U_j(1, 0) = (1 - \tau)(w_i + \alpha_j w_j) + \phi_j$
3. $U_j(0, 1) = (1 - \tau)(\alpha_i w_i + w_j)$
4. $U_j(0, 0) = (1 - \tau)(\alpha_i w_i + \alpha_j w_j) + (1 + \sigma_{ij})\phi_j$

As the spouse cannot influence the behavior of the focal partner, the spouse will choose to work if the focal partner works if $U_j(1, 1) > U_j(1, 0)$. If the focal partner does not work, the spouse will choose to work if $U_j(0, 1) > U_j(0, 0)$.

Focal partner works

$$
\begin{aligned}
U_j(1,1) > U_j(1,0) &\iff (1-\tau)(w_i + w_j) > (1-\tau)(w_i + \alpha_j w_j) + \phi_j \\
&\iff \phi_j < (1-\tau)(1-\alpha_j)w_j
\end{aligned}
\tag{4.6}
$$

The threshold utility of individual leisure is given by $\phi_j^* = (1-\tau)(1-\alpha_j)w_j$, and the spouse will only work if the utility of individual leisure is lower than ϕ_j^*.

Focal partner does not work

$$
\begin{aligned}
U_j(0,1) > U(0,0) &\iff (1-\tau)(\alpha_i w_i + w_j) > (1-\tau)(\alpha_i w_i + \alpha_j w_j) + (1+\sigma_{ij})\phi_j \\
&\iff \phi_j < \frac{(1-\tau)(1-\alpha_j)w_j}{1+\sigma_{ij}}
\end{aligned}
\tag{4.7}
$$

The threshold utility of individual leisure is given by $\phi_j^{**} = \frac{(1-\tau)(1-\alpha_j)w_j}{1+\sigma_{ij}}$, and the spouse will only work if the utility of individual leisure is lower than ϕ_j^{**}.

4.1.2 The incentive to retire jointly

The optimal strategy of the spouse can be expressed as follows:

- a: Always work if $\phi_j \in (0, \phi_j^{**})$.
- b: Work if the focal partner works and do not work if the focal partner does not work if $\phi_j \in (\phi_j^{**}, \phi_j^*)$.
- c: Never work if $\phi_j \in (\phi_j^*, 1)$.

Based on ϕ_j^* and ϕ_j^{**}, the spouse will be more likely to work if the focal partner works if $\sigma_{ij} > 0$, as it implies that $\phi_j^{**} < \phi_j^*$. The reason is that if $\sigma_{ij} > 0$, the spouse values joint leisure, implying that complementarities in leisure cause the spouse to respond to the focal partner's employment decision. If $\sigma_{ij} \to \infty$ then $\phi_j^{**} \to 0$, meaning that if the spouse values joint leisure infinitely much more than individual leisure, (s)he will never work if the focal partner does not work.

Considering the economic incentives, a lower rate of compensation increases the incentive to work everything else equal. A larger wage increases the incentive to work, while a higher average tax rate reduces the motivation to work.

The aggregate employment of the spouses can be expressed as follows where $f(\phi_j)$ is the density function of ϕ_j:

A $= \int_0^{\phi_j^{**}} f(\phi_j)\, d\phi_j$ is the share of spouses that always work.

B $= \int_{\phi_j^{**}}^{\phi_j^{*}} f(\phi_j)\, d\phi_j$ is the share of spouses that only work if the focal partners work.

C $= \int_{\phi_j^{*}}^{1} f(\phi_j)\, d\phi_j$ is the share of spouses that never work.

The larger the value of B, the more does the spouses' employment respond to the focal partners' employment as a result of complementarities in leisure.

4.1.3 The impact of the couples' age structure on the spouses' employment

ϕ_j^* and ϕ_j^{**} confirm that, besides leisure complementarities, economic incentives also affect how likely the spouse is to work if the focal partner works. While it is reasonable to assume that the average tax rate does not differ substantially between spouses, the wage and compensation rate may vary greatly.

If the spouse is older than the focal partner, (s)he can retire on ERB. In this case, $0 < \alpha_j^{age_j \geq ERA} < 1$ such that the spouse receives a positive amount in benefits, which is lower than the potential wage. If the spouse retires and is younger than the focal partner, (s)he will not receive any benefits and will have to waive five years of ERB.[10] Hence, $\alpha_j^{age_j < ERA} < \alpha_j^{age_j \geq ERA}$. In both situations, the spouse will be more likely to work if the focal partner works if $\sigma_{ij} > 0$. ϕ_j^* and ϕ_j^{**} are largest if the spouse is younger than the focal partner because $\alpha_j^{age_j < ERA} < \alpha_j^{age_j \geq ERA}$, implying that the spouse is more likely to work if (s)he is younger than the focal partner. It is ambiguous whether the older or younger spouses' employment is most affected by whether the focal partners work, as this depends on the distribution of the utility of individual leisure in the population.

4.2 Simulations of the model

Multiple studies have estimated the utility of leisure, or equivalently, the disutility of working. Rätzel (2012) and Kaplan & Schulhofer-Wohl (2018) stress that the disutility of working is low and only occurs, as people cannot choose to work their preferred number of hours. They state that the utility of people enabled to work their preferred number of hours is greater if working than if not working, even if income is held fixed. This suggests that the disutility of working is rather low for the majority of the population. Assuming that the utility of individual leisure is low for most of the population, the distribution of the utility of individual leisure in the population may be described by a Beta distribution in which most of the density is close

[10] The spouse will have to actively search for a job if claiming unemployment benefits or social assistance. Assuming that searching for a job requires the same effort as working, (s)he will never retire on benefits before the ERA, as the benefits will always be lower than the wage.

to zero. Considering a population of 1,000,000 spouses, I draw 1,000,000 values of ϕ_j where $\phi_j \sim \beta(1, 10)$. As the Beta distribution has only support for $\phi_j \in (0, 1)$, the assumption that $0 < \phi_j < 1$ is satisfied by definition.

Alternatively, one may argue that the utility of individual leisure increases with age (Gustman & Steinmeier, 2002). In this case, the utility of individual leisure follows different distributions conditional on whether the spouses are younger or older than the focal partners. However, if limiting the sample to spouses that are just slightly older or younger than the focal partners, it is reasonable to believe that the utility of individual leisure follows the same distribution.

To simulate the model, I impose assumptions about the parameters of the model. I assume that the wage is homogeneous and normalize the wage to 1 ($w_j = 1$). The average tax rate of the average Danish tax payer is 35 percent ($\tau = 0.35$) c.f. The Danish Ministry of Taxation (2016), while the ERB constitutes 80 percent of the maximum take-up of unemployment benefits ($\alpha_j^{age_j \geq ERA} = 0.80$) if not applying the two-year rule. If retiring prior to the ERA, spouses are not entitled to benefits ($\alpha_j^{age<ERA} = 0$). In accordance with Jørgensen (2014), I assume that spouses value leisure twice as much if the focal partners do not work ($\sigma_{ij} = 1$). The model's predictions concerning the spouses' employment conditional on whether the focal partners work and the couple's age structure are depicted in figure 4.1. In figure A.1 and figure A.2 in appendix A.3, I simulate the model with different values of the utility of joint leisure to investigate how it affects the difference between the spouses' employment. Furthermore, one can benefit from considering appendix A.1 to investigate how changing the assumptions about the model's parameters or the distribution of the spouses' utility of individual leisure in the population affects how the spouses' employment responds to the focal partners' employment.

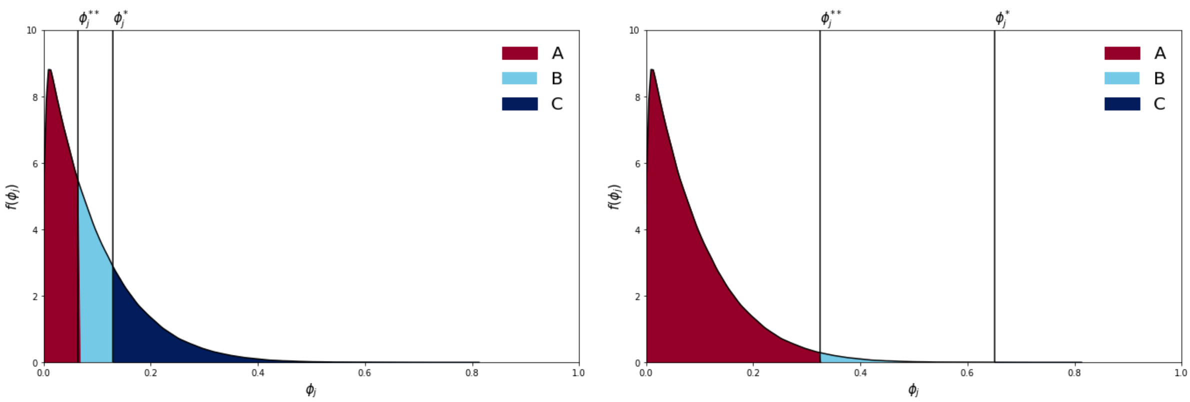
Figure 4.1a shows that the spouses' employment, in general, is low if they are older than the focal partners, as they have lower economic incentives to work. The assumption about the parameters of the model and the distribution of the utility of individual leisure imply that $B = \int_{\phi_j^{**}}^{\phi_j^*} f(\phi_j)\, d\phi_j = 0.20$. Hence, leisure complementarities cause the spouses' employment to be 20 percentage points larger if the focal partners work provided that they are older than the focal partners.

The economic costs of not working are considerable if spouses are younger than the focal partners, causing almost every spouse to be employed as depicted in figure 4.1b. As $B = \int_{\phi_j^{**}}^{\phi_j^*} f(\phi_j)\, d\phi_j = 0.02$, leisure complementarities do not cause the difference between the spouses' employment conditional on the focal partners' employment to be substantial provided that the spouses are younger than the focal partners.

The theoretical model reveals that if spouses value joint leisure, their employment increases if the focal partners' employment increases. The assumption about the model's parameters and the distribution of the spouses' utility of individual leisure in the population implies that only

Figure 4.1: The spouses' employment conditional on the utility of individual leisure

(a) Older spouses (b) Younger spouses



Note: The utility of individual leisure is increasing in ϕ_j, and $\phi_j \sim \beta(1, 10)$. ϕ_j^* denotes the threshold utility of individual leisure where the spouse is indifferent between working and not working if the focal partner works. ϕ_j^{**} denotes the threshold utility of individual leisure where the spouse is indifferent between working and not working if the focal partner does not work. $A = \int_0^{\phi_j^{**}} f(\phi_j)\, d\phi_j$ is the share of spouses that always work. $B = \int_{\phi_j^{**}}^{\phi_j^{*}} f(\phi_j)\, d\phi_j$ is the share of spouses that only work if the focal partners work. $C = \int_{\phi_j^{*}}^{1} f(\phi_j)\, d\phi_j$ is the share of spouses that never work. One can benefit from considering appendix A.1 to investigate how changing the assumptions about the model's parameters and the distribution of the spouses' utility of individual leisure in the population affect how the spouses' employment responds to the focal partners' employment.

the employment of spouses that are older than the focal partners responds considerably. This emphasizes the importance of considering economic incentives, as older and younger spouses value joint leisure identically in the model. The assumptions about the model's parameters and the distribution of the spouses' utility of individual leisure in the population are critical, as changing these may generate completely different conclusions. If assuming that the spouses' disutility of working is substantial, only the younger spouses' employment responds to the focal partners' employment. Therefore, the predictions from the theoretical model must be supported by empirical evidence.

5 Empirical strategy

This section describes the empirical strategy utilized to investigate whether leisure complementarities cause couples to retire jointly. Following the notation from section 4, I define the focal partner as individual i and the spouse as individual j. Let $Y_j(1)$ denote the spouse's employment in the case where the focal partner's ERA increases. Identically, let $Y_j(0)$ denote the spouse's employment if the focal partner's ERA does not increase. The causal effect of raising the ERA on the spouse's employment equals $\iota_j = Y_j(1) - Y_j(0)$. However, I never observe the same spouse's employment in a scenario where the focal partner is exposed to the increase in the ERA and in a scenario where (s)he is not. This is known as *the fundamental problem of evaluation*, stating that it is not possible to observe the same individual in two different outcomes simultaneously. To overcome this problem, one must rely on natural experiments or quasi-natural experiments.

It is ideal to consider a natural experiment if a researcher can find one that works in the same way as a randomized controlled trial, which is viewed as the gold standard in economic research (Lee & Lemieux, 2010). Such a natural experiment could be if couples are randomly assigned to treatment, ensuring that the variation in the focal partners' retirement age is entirely exogenous. Randomly assigning different retirement ages to focal partners, or randomly matching a large set of focal partners who face different retirement ages to a set of spouses, the law of large numbers states that the difference in the spouses' employment can be attributed to the difference in the focal partners' retirement age. In this case, the estimates yield the average treatment effect. However, the focal partners' retirement age is based on the date of birth, and people are only randomly matched in TV shows. Therefore, I rely on a quasi-experimental approach by exploiting the exogenous variation in the retirement age caused by the increase in the ERA. If the estimates are to be given a causal interpretation when considering the quasi-experimental approach, one must ensure that a suited research design is utilized where the identifying assumptions are not violated. I utilize the regression discontinuity design (RDD), as this research design is the most suitable in this context. I argue why the RDD is favorable compared to other research designs in section 9.2.

5.1 The sharp regression discontinuity design

The fundamental idea in the sharp regression discontinuity design (SRDD) is that the treatment rule creates a local randomized experiment. The estimates in the SRDD can be given a causal interpretation if couples, in which the focal partners were born just before the cutoff, are identical to couples, in which the focal partners were born just after the cutoff, which is the main identifying assumption. In this case, only the difference in the focal partners' ERA can explain a potential difference in the spouses' employment. Couples are assigned to the treatment group if the focal partners were born after the cutoff, and couples are assigned to the control group if the focal partners were born before the cutoff. Therefore, the control group serves as the counterfactual scenario if the ERA is not raised. As the treatment rule is completely deterministic, the assignment to treatment can be expressed by

$$D_{ij} = \begin{cases} 1 \text{ if } x_i \geq 0 \\ 0 \text{ if } x_i < 0 \end{cases}$$

where x_i is the running variable, and D_{ij} defines whether a couple is assigned to treatment or not. Hence, no matter how close x_i comes to zero (from below), treatment is unchanged until

$x_i = 0$ (Angrist & Pischke, 2009). Since D_{ij} is a deterministic function of x_i, it is impossible to observe the same value of x_i for both treated and non-treated couples, implying that the overlap condition is violated by definition (G. W. Imbens & Lemieux, 2008).[11] It is the main difference between the RDD and many other research designs in which covariates are balanced. Instead of matching covariates, the SRDD uses an interval near the cutoff where treated and non-treated couples are expected to be identical. Following G. W. Imbens & Lemieux (2008) and Lee & Lemieux (2010), the average causal effect of the treatment at the cutoff can be expressed by

$$\tau_{SRDD} = E[Y_j(1) - Y_j(0)|x_i = 0] = \lim_{x \downarrow 0} E[Y_j|x_i = x] - \lim_{x \uparrow 0} E[Y_j|x_i = x]$$

I observe four cutoffs of interest, as the ERA increased four times, each time by six months as argued in section 3. I utilize the number of days the focal partner is born after the cutoff as the running variable (x_i), implying that, at the first cutoff, focal partners born on December 31, 1953, and January 1, 1954, are given a value of -1 and 0 respectively. Identically, at the second cutoff, focal partners born on June 30, 1954, and July 1, 1954, are given a value of -1 and 0, respectively. The same applies to the third and fourth cutoff as depicted in table 5.1. By normalizing the running variable at a cutoff of 0, I pool the four cutoffs such that observations at all four cutoffs are included in the same estimation. The approach of pooling cutoffs is advantageous, as the number of observations is larger, increasing the statistical power by reducing the estimated standard errors. A pitfall, however, arises if the effects are substantially different between the separate cutoffs. In this case, the pooled estimates are less robust and cannot be interpreted as the overall effects of raising the ERA by six months since the effects depend on the specific cutoff (Cattaneo et al., 2016). I investigate the potential presence of heterogeneity between the cutoffs in section 7.5.2.

The increase in the ERA implies that, although the focal partners are the same age in months, focal partners born just before the cutoff are eligible to claim ERB for six months, within which the focal partners born just after the cutoff are ineligible. Thus, at the first cutoff, I measure the spouses' employment in the month after the focal partners' 60th birthday and in the following five months. Identically, at the second cutoff, I measure the spouses' employment in the month after the focal partners reach 60.5 years and in the following five months. I do the same at the third and fourth cutoff, however measured in the month after the focal partners reach 61 years and 61.5 years, respectively. Hence, I estimate the models in each of the six months where the focal partners in the control group and the treatment groups are the same age in months, but where only focal partners in the control group are eligible to claim ERB as depicted in figure 5.1.

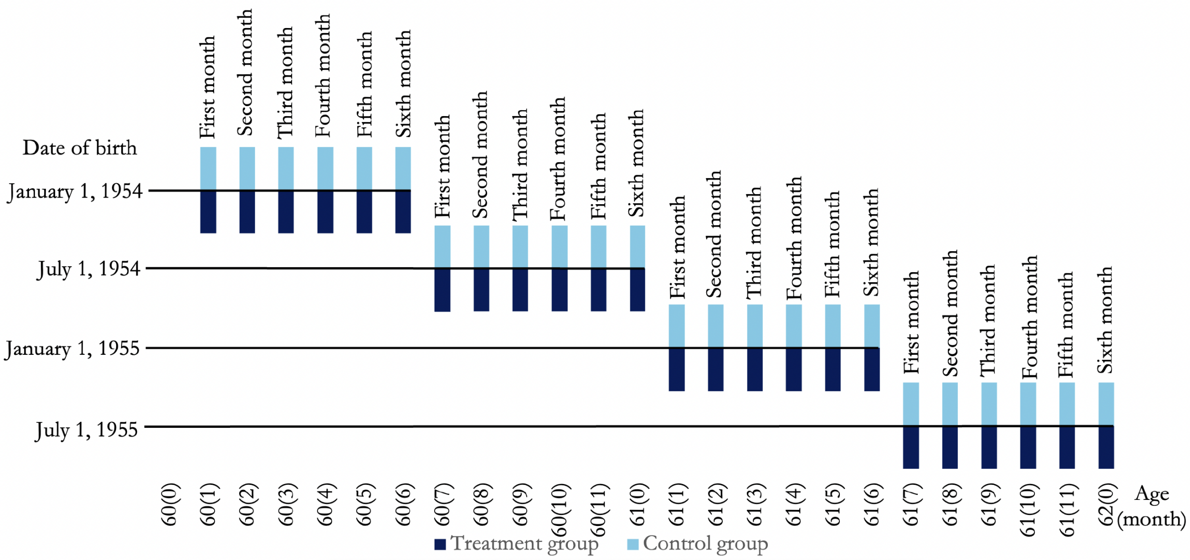
[11] The overlap condition states that the treatment group's and control group's covariates must overlap, i.e. that the same values of the covariates are observed for observations in both the treatment group and control group.

Table 5.1: Constructing the pooled cutoff

Cutoff	Date of birth	x_i	D_{ij}	ERA	Observed at age (month)	Observed at point of time
(I)	December 31, 1953	-1	0	60	60(1) - 60(6)	January, 2014 - June, 2014
(I)	January 1, 1954	0	1	60.5	60(1) - 60(6)	February, 2014 - July, 2014
(II)	June 30, 1954	-1	0	60.5	60(7) - 61(0)	January, 2015 - June, 2015
(II)	July 1, 1954	0	1	61	60(7) - 61(0)	February, 2015 - July, 2015
(III)	December 31, 1954	-1	0	61	61(1) - 61(6)	January, 2016 - June, 2016
(III)	January 1, 1955	0	1	61.5	61(1) - 61(6)	February, 2016 - July, 2016
(IV)	June 30, 1955	-1	0	61.5	61(7) - 62(0)	January, 2017 - June, 2017
(IV)	July 1, 1955	0	1	62	61(7) - 62(0)	February, 2017 - July, 2017

Note: The second column depicts the focal partner's date of birth. The third column shows the normalized value of the running variable (x_i). The fourth column shows whether the couple is in the treatment group ($D_{ij} = 1$) or the control group ($D_{ij} = 0$). In the sixth column, the number in parentheses denotes the focal partner's age. E.g., "60(1) - 60(6)" states that the spouse's employment is observed between the month after the focal partner's 60th birthday and the month in which the focal partner turns 60.5 years.

Figure 5.1: Overview of the estimation strategy



Note: The bandwidth applied in the estimations is 100 days as argued in section 5.3.3. Hence, couples in which the focal partners are born within 100 days before or after the pooled cutoff are included in the estimation sample.

As I measure the spouses' employment in months based on the focal partners' age, I observe the spouses' employment in different months of different years. This is not a threat to identification, as the macroeconomic conditions, measured by the employment and the GDP growth, have been stable within the period as depicted in figure A.3 in appendix A.3. Furthermore, the employment of the elderly is less affected by economic fluctuations, as they are less likely to be employed within the service and construction industries (Hoynes et al., 2012).

Certain assumptions must be satisfied to ensure that couples in the neighborhood of the cutoff are identical such that the estimates obtained in the SRDD are valid (G. W. Imbens & Lemieux, 2008; Lee & Lemieux, 2010). The identifying assumptions (RDD.1, RDD.2, RDD.3) are stated below. I test and discuss whether the assumptions are satisfied in section 7.5.1.

Assumption RDD.1: Continuity of the conditional expectation function
If treating a potential discontinuity at the cutoff as the causal effect of treatment, the conditional expectation function must be continuous. Hence, it is necessary that $E[Y_j(1)|x_i]$ and $E[Y_j(0)|x_i]$ are continuous in x_i.

Assumption RDD.2: No manipulation of the running variable
The estimates are only valid if couples have imprecise control over the running variable. If couples can manipulate their value of the running variable, it is possible to self-select in or out of treatment, meaning that the assignment at the cutoff is not random. In this thesis, the assumption implies that focal partners cannot manipulate their date of birth and that spouses are no more likely to marry focal partners born just before the cutoff dates.

Assumption RDD.3: Continuity of the predetermined covariates
This assumption states that the predetermined covariates must evolve smoothly at the cutoff. Two factors usually cause discontinuities in the covariates; either the couples can manipulate their value of the running variable to select in or out of treatment, or the cutoff correlates with other variables than employment. The former relates to the assumption of no manipulation of the running variable, while the latter may occur if, e.g., focal partners born just after the cutoff dates are more likely to be highly skilled.[12]

5.2 The fuzzy regression discontinuity design

Utilizing the SRDD makes it possible to identify whether raising the retirement age increases the spouses' employment. The reduced-form estimates are interesting from a policy perspective, as they yield the intention-to-treat effect. However, the estimates cannot necessarily identify the spillover effect of future changes in the retirement age, as the change in the spouses' employment depends on the change in the focal partners' employment since not all focal partners respond to the increase in the ERA. Hence, I investigate the spillover effect from the change in the focal partners' employment to the change in the spouses' employment.

Focusing on whether the focal partner works (y_i following the notation from section 4) instead

[12] This may occur if people born in the beginning of the year obtain more years of completed schooling due to an older age at school entry as documented by Meng (2018).

of whether (s)he is affected by the increase in the ERA (D_{ij}), however, is not straightforward, as the spouse is more likely to be employed if the focal partner is employed independently of the increase in the ERA. This may partially be explained by homogeneous couples, as people tend to find a partner with the same characteristics.[13] Therefore, the difference in the spouses' employment conditional on whether the focal partners are employed cannot solely be attributed to leisure complementarities. By attributing the difference in the spouses' employment exclusively to leisure complementarities, the importance of leisure complementarities is overestimated, as y_i is correlated with the error term. To overcome this problem, I utilize the fuzzy regression discontinuity design (FRDD). The FRDD combines the RDD and the instrumental variables (IV) approach, as I use D_{ij} as an instrument for y_i. As in the SRDD, the estimates can only be given a causal interpretation if couples, in which the focal partners were born just before the cutoff, are identical to couples, in which the focal partners were born just after the cutoff. Lee & Lemieux (2010) state that the local average treatment effect (LATE) at the cutoff can be expressed by

$$\rho_{FRDD} = \frac{\lim_{x \downarrow 0} E[Y_j | x_i = x] - \lim_{x \uparrow 0} E[Y_j | x_i = x]}{\lim_{x \downarrow 0} E[D_{ij} | x_i = x] - \lim_{x \uparrow 0} E[D_{ij} | x_i = x]}$$

The focal partners' employment is not a deterministic function of x_i since the focal partners may work regardless of the value of x_i. However, the probability that the focal partners work in the six months of interest is larger if they are affected by the increase in the ERA. This can be expressed by

$$Pr[y_i = 1 | x_i] = \begin{cases} g_1(x_i) \text{ if } x_i \geq 0 \\ g_0(x_i) \text{ if } x_i < 0 \end{cases}$$

where $g_1(x_i) > g_0(x_i)$.

The identifying assumptions in the IV approach and assumption RDD.1[14], RDD.2, and RDD.3 from section 5.1 have to be satisfied to obtain valid estimates. The additional identifying assumptions in the FRDD (RDD.4, RDD.5, RDD.6) are stated below. I test and discuss whether the assumptions are satisfied in section 7.5.1.

[13] E.g., Goux & Maurin (2003) conclude that people marry partners that have the same socioeconomic characteristics as themselves. This implies that highly-skilled spouses will be more likely to be married to highly-skilled focal partners than low-skilled spouses. At the same time, highly-skilled focal partners and spouses are more likely to work than low-skilled focal partners and spouses because of a lower demand for low-skilled workers (Wolcott, 2021). Hence, the highly-skilled spouses may be more likely to work because the focal partners work, but also because of a larger labor demand. Therefore, the difference in the spouses' employment conditional on whether the focal partners work cannot entirely be attributed to complementarities in leisure.

[14] Y_j is to be replaced by Y_i in assumption RDD.1 in the FRDD.

Assumption RDD.4: The first stage exists

If $g_1(x_i) > g_0(x_i)$, the first stage exists by definition. The larger the difference between $g_1(x_i)$ and $g_0(x_i)$, the greater the exogenous variation, suggesting a stronger first stage.

Assumption RDD.5: The exclusion restriction

The instrument and the error term must be uncorrelated such that the instrument only influences the outcome variable through the instrumented variable. In other words, the spouses' employment is only to be affected by whether the focal partners were born before or after the cutoff through the focal partners' employment.

Assumption RDD.6: The monotonicity assumption

Focal partners can be divided into four groups: *always takers*, *never takers*, *compliers*, and *defiers*.

- Always takers always take the treatment, meaning that the focal partners in this group always work regardless of the eligibility rules. This implies that $y_i = 1$ regardless of the value of D_{ij}.
- Never takers never take the treatment and therefore never work regardless of the eligibility rules. This implies that $y_i = 0$ regardless of the value of D_{ij}.
- Compliers only take the treatment if the instrument is switched on, meaning that focal partners in this group only work if they are affected by the increase in the ERA. This implies that $y_i = 1$ if $D_{ij} = 1$ and $y_i = 0$ if $D_{ij} = 0$.
- Defiers only take the treatment if the instrument is switched off, meaning that the focal partners in this group only work if they are not affected by the increase in the ERA. This implies that $y_i = 1$ if $D_{ij} = 0$ and $y_i = 0$ if $D_{ij} = 1$.

Always takers and never takers do not respond to the instrument, causing LATE to be non-informative about the spillover effects of couples in which the focal partners are always takers or never takers. The monotonicity assumption requires the absence of defiers such that the effects are only obtained for the compliant subpopulation, as the presence of defiers cancels out the effect on compliers. If assumption RDD.6 is satisfied, I obtain the estimates for spouses married to focal partners that decide to work if they are affected by the increase in the ERA and not to work if they are not affected by the increase in the ERA.

5.3 Estimation

Having outlined the SRDD and the FRDD and the identifying assumptions in both research designs, I present the models' specifications. This includes the functional form, the choice of model, and the applied standard errors.

5.3.1 The model specification

Before defining the models for estimation, I must consider whether to utilize a parametric or a nonparametric approach. If utilizing parametric estimation, all observations in the sample are used in the estimation after specifying a bandwidth, i.e., the number of days the focal partners were born relative to the cutoff to be included in the estimation. If observations far away from the cutoff are included in the estimation, more observations are used with the implication of enhanced statistical power. At the same time, including observations far from the cutoff in the estimation implies that the misspecification bias may be larger, as the approximation of the conditional expectation function will probably deviate more from the true conditional expectation function. Therefore, utilizing the parametric approach requires that the functional form is correctly specified.

To simplify, Jacob et al. (2012) stress that the difference between the parametric and nonparametric approach is that the former chooses a model that fits the data, while the latter selects the data that fit the model. The nonparametric approach, such as local linear regression, is favored by Hahn et al. (2001), as they argue that the misspecification bias is smaller. The fundamental idea in the nonparametric approach is to only include observations within a small range of the cutoff where a local linear regression can properly approximate the true conditional expectation function. As only observations within a small neighborhood of the cutoff are used for estimation, the statistical power is lower as a result of the smaller number of observations. To select the range where a local linear regression can properly approximate the conditional expectation function, i.e. the optimal bandwidth, Calonico et al. (2020), G. Imbens & Kalyanaraman (2012), and G. W. Imbens & Lemieux (2008) have developed data-driven rules.

The trade-off between bias and precision is vital when choosing whether to utilize the parametric or nonparametric approach. The parametric approach is associated with a greater precision, but comes with the risk of a larger misspecification bias. In the nonparametric approach, the misspecification bias may be more negligible, but it is followed by a larger variance. Neither the parametric approach nor the nonparametric approach is superior, as it is impossible to know with certainty which one has a smaller bias unless knowing the true conditional expectation function (Lee & Lemieux, 2010). Lee & Lemieux (2010) argue that instead of relying on only one of the approaches, the estimates have to be stable across both approaches to verify the robustness of the estimates.

To my knowledge, no papers have used the exact date of birth as the running variable in the literature of (joint) retirement. However, the date of birth has been used by, e.g., Dobkin & Ferreira (2010) and McCrary & Royer (2011), who both utilize the FRDD, to investigate how the age at school entry affects future outcomes. These papers exploit how several countries require all children born in the same year to enroll in school at the same time, creating a cutoff between December 31 and January 1. Dobkin & Ferreira (2010) utilize the parametric approach, as they argue that the date of birth is a discrete variable and specify a bandwidth of 180 days. McCrary & Royer (2011) utilize the nonparametric approach, but show that the estimates are stable if using the parametric approach instead. As the date of birth may be seen as a discrete variable and because of the small number of observations in the near vicinity of the cutoff, I utilize the parametric approach. However, I also investigate whether the estimates remain stable if I consider the nonparametric approach instead.

5.3.2 The models

I estimate three different models. First, I estimate the change in the focal partners' employment due to the increase in the ERA to ensure that the focal partners' employment responds to the reforms. Second, I estimate the change in the spouses' employment caused by the increase in the ERA. Third, I estimate the spillover effect from the change in the focal partners' employment to the change in the spouses' employment.

The change in the focal partners' employment

$$y_i = \beta_1 + \beta_2\theta_i + \beta_3\theta_j + f(x_i) + \gamma D_{ij} + \epsilon_{ij} \tag{5.1}$$

The change in the spouses' employment

$$y_j = \beta_1 + \beta_2\theta_i + \beta_3\theta_j + f(x_i) + \tau D_{ij} + \epsilon_{ij} \tag{5.2}$$

y_i, y_j and D_{ij} follow the same notation as earlier defined. β_1 denotes the intercept, and β_2 and β_3 are parameters to be estimated. θ_i is a vector of the focal partner's covariates, while θ_j is a vector of the spouse's covariates. The inclusion of covariates is not necessary to obtain consistent estimates since the assignment to treatment does not depend on the covariates. However, including covariates is usually an advantage to increase the statistical power (Lee & Lemieux, 2010). ϵ_{ij} is the error term in both model 5.1 and model 5.2, and $f(x_i)$ denotes the functional form, which I define in section 5.3.3.

γ is the parameter of interest in model 5.1 where $\hat{\gamma}$ denotes the change in the focal partners' employment in percentage points. τ is the parameter of interest in model 5.2 where $\hat{\tau}$ denotes the change in the spouses' employment in percentage points. Hence, $\hat{\tau}$ is the reduced-form estimate that yields the intention-to-treat effect of raising the focal partners' ERA.

The spillover effect from the change in the focal partners' employment to the change in the spouses' employment

I utilize the FRDD by considering a two-stage least squares (2SLS) procedure. This procedure divides the estimation into two stages: the first stage and the second stage. I insert the fitted values of y_i ($\hat{y}_i$) from the first stage (model 5.1) into the second stage (model 5.3) such that the second stage is given by

$$y_j = \beta_1 + \beta_2\theta_i + \beta_3\theta_j + f(x_i) + \rho\hat{y}_i + v_{ij} \tag{5.3}$$

where all parameters and variables except ρ and v_{ij} are to be interpreted in the same way as in model 5.1 and model 5.2. $v_{ij} = \eta_{ij} + \rho(y_i - \hat{y}_i)$ is the OLS residual variance whereas η_{ij} is the correct 2SLS residual variance (Angrist & Pischke, 2009). If ignoring that $\hat{y}_i$ is estimated, the estimated standard errors will usually be too low. In Stata, the correct 2SLS standard errors are reported by default.

ρ is the parameter of interest where $\hat{\rho}$ denotes the difference in the spouses' employment in percentage points conditional on whether the focal partners work as a result of the increase in the ERA. The estimator is consistent for ρ as the covariates and the first stage fitted values are uncorrelated with η_{ij} and $(y_i - \hat{y}_i)$ (Angrist & Pischke, 2009). Therefore, ρ captures the causal marginal effect of the difference in the probability that the spouse works conditional on whether the focal partner works, which is interpreted as the spillover effect. The spillover effect is expected to be larger than the intention-to-treat effect ($\rho > \tau$) as not all focal partners respond to the increase in the ERA.

5.3.3 The functional form and choice of bandwidth

The key to ensuring consistent estimates when utilizing the parametric approach is to specify a functional form to serve as a sufficient approximation of the true conditional expectation function to reduce the misspecification bias. Therefore, I must impose assumptions of the functional form of $f(x_i)$. Lee & Lemieux (2010) and Jacob et al. (2012) stress that the functional form has to be flexible to approximate the conditional expectation function. They recommend the usage of two different methods to choose the correct functional form: the Akaike information criterion (AIC) and an F-test.[15] The former approach evaluates the bias-precision trade-off of applying a

[15] A more detailed description of the two approaches are provided in Lee & Lemieux (2010) and Jacob et al. (2012).

more complex model such that it measures the goodness-of-fit of a model (Jacob et al., 2012).[16] While the approach can be used to rank a set of models, it is not informative on whether the models fit the data poorly. The two papers favor the F-test approach, as this approach can be used to rank models, while it is also informative on whether the preferred model fits the data well.[17] Both approaches often imply that high-order polynomials are preferred (Lee & Lemieux, 2010).

Gelman & Imbens (2019) and Gelman & Zelizer (2015) argue that high-order polynomials are overused in the RDD, as it is a flawed approach in most cases. They argue that using high-order polynomials leads to noisy estimates, estimates that are not robust, and poor coverage of confidence intervals. Researchers tend to include high-order polynomials to show that the model fits the data well, although they are unable to argue why a high-order polynomial is a good approximation of the conditional expectation function. Using high-order polynomials can cause bias, suggesting that applying high-order polynomials may do more harm than good. The papers recommend to use a linear or quadratic polynomial instead of relying on formal tests to specify the functional form. Hence, I apply a linear specification, meaning that the functional form is represented by a linear polynomial which is allowed to differ on each side of the cutoff. This implies that $f(x_i) = \delta_1 x_i + \delta_2 x_i D_{ij}$ where δ_1 and δ_2 are parameters to be estimated. Furthermore, I investigate whether the estimates are robust to the functional form by including second- and third-order polynomials.[18] Should the estimates remain stable, this would suggest that the misspecification bias is not worrying.

It is tempting to use all or at least many of the running variable's values to increase the statistical power. However, using a larger bandwidth usually increases the difference between the approximation of the conditional expectation function and the true conditional expectation function, resulting in inconsistent estimates. The misspecification bias can usually be ignored if approximating the conditional expectation function with a narrow bandwidth. However, it means that fewer observations are included in the estimation, resulting in larger standard errors. Therefore, the choice of bandwidth reflects the trade-off between bias and precision. A pragmatic approach is to consider the data graphically and try different bandwidths to investigate whether

[16] The AIC is given by $Nln(\hat{\sigma}_b^2) + 2p$ where $\hat{\sigma}_b^2$ is the estimated residual variance, and p is the number of parameters in the model (Jacob et al., 2012).

[17] Using this approach, one includes a set of $K-2$ bin dummies in the estimation, where K denotes the total number of bins. The functional form is given by a linear polynomial which is allowed to differ on each side of the cutoff. Performing an F-test, if the bin dummies are jointly significant, the functional form is instead represented by a quadratic polynomial. One keeps applying polynomials of a higher order until the bin dummies are jointly insignificant.

[18] Including a second-order polynomial, $f(x_i) = \mu_1 x_i + \mu_2 x_i^2 + \mu_3 x_i D_{ij} + \mu_4 x_i^2 D_{ij}$ where μ_1, μ_2, μ_3 and μ_4 are parameters to be estimated. Including a third-order polynomial, $f(x_i) = \zeta_1 x_i + \zeta_2 x_i^2 + \zeta_3 x_i^3 + \zeta_4 x_i D_{ij} + \zeta_5 x_i^2 D_{ij} + \zeta_6 x_i^3 D_{ij}$ where ζ_1, ζ_2, ζ_3, ζ_4, ζ_5 and ζ_6 are parameters to be estimated.

the estimates change substantially. As my baseline specification, I consider a bandwidth of 100 days, almost half of the bandwidth used by Dobkin & Ferreira (2010). As a robustness check, I consider a bandwidth of 200 days and 50 days to investigate whether the estimates are robust to the choice of bandwidth.

5.3.4 The choice of model

When investigating the change in the spouses' employment along the extensive margin, a spouse can choose to work or not, meaning that the outcome variable is binary. Therefore, the predicted probability that the spouse is employed ($\hat{y_j}$) has to be between zero and one by definition. However, the linear probability model does not account for that, implying that predicted probabilities can be outside of this range. A possible solution is to utilize nonlinear models such as the probit model, where the predicted probabilities are restricted to be between zero and one. However, the probit model does not produce marginal effects by default, as the marginal effects depend on the covariates of the couple. Marginal effects are obtained by estimating the marginal effect of each spouse and calculating the average, and the corresponding standard errors are obtained by the Delta method.[19] Cameron & Trivedi (2005) argue that nonlinear models are favorable, as they correctly allow for nonlinearities. Angrist & Pischke (2009) stress that marginal effects are in most cases identical in linear and nonlinear models if a considerable part of the predicted probabilities in the linear probability model lie within the range of zero and one. They emphasize that the linear probability model is preferable since the approach is more standardized and straightforward. Due to the arguments of Angrist & Pischke (2009), I apply the linear probability model as my baseline model, but subsequently investigate whether the marginal effects change substantially when utilizing the probit model instead in section 7.5.2.

5.3.5 The standard errors

The choice of standard errors is essential, as the estimated standard errors' magnitude determines the statistical significance of the estimates. The outcome variable is binary, causing the residuals to be heteroskedastic by definition since the residuals are larger outside of the interval of zero and one. A possible solution is to use White robust standard errors, as they are favorable compared to the classical standard errors when the regression residuals are heteroskedastic (Angrist & Pischke, 2009). Furthermore, Kolesár & Rothe (2018) argue that White robust standard errors combined with a more narrow bandwidth can be attractive to correct for the misspecification bias.

[19] Cameron & Trivedi (2005) provide an overview of the probit model, including properties, the link function, etc. The book furthermore offers a detailed explanation of the Delta method.

I pool different cutoffs, implying that the same individual may be the focal partner at one of the cutoffs, and the spouse at another. Some covariates, including gender and the region of residence, are perfectly correlated within couples, meaning that the error terms are correlated within the clusters. Hence, the estimated standard errors will probably be too low if using the classical standard errors. To overcome this problem, the standard errors can be clustered on the couple level. Lee & Card (2008) propose to cluster the standard errors on the running variable to correct for the bias of the classical standard errors, which arises if observations that take on the same value of the running variable are correlated. Angrist & Pischke (2009) stress that, as a rule of thumb, the number of clusters must exceed 50. Therefore, too few clusters are not a concern in this thesis regardless of whether the standard errors are clustered on the couple level or the running variable.

In the light of the different arguments, none of the standard errors are superior, leading me to adopt a relatively conservative approach where I use the largest standard errors to ensure that the statistical significance is not overestimated. In my baseline specification, I cluster the standard errors on the running variable. Furthermore, I estimate the models with White robust standard errors and standard errors clustered on the couple level to investigate whether this changes the conclusions concerning the significance of the estimates.

6 Data

I consider Danish administrative data containing information about each individual in the population. The data include information about personal characteristics, employment, income, transfers, and tax payments. It is possible to match data from different sources, as each individual is identified with a unique ID number. The potential spouse of each individual is identified with a unique ID number too, allowing me to match the focal partner to his/her spouse such that I observe the variables of both partners. As the data are recorded weekly, monthly, or annually, it is possible to create an unbalanced panel data set in which the same individual appears in multiple periods. The data are third-party reported by the Danish authorities, ensuring that the variables are consistently measured for everyone to avoid the self-reported bias that may arise if considering surveys (Gustman & Steinmeier, 2002; Kallestrup-Lamb, 2011). Another advantage of using this data is the access to a large number of covariates, which has often been missing in the literature of joint retirement (Lumsdaine & Mitchell, 1999).

6.1 Variables

The outcome variable

In the literature of joint retirement, the date of retirement, being employed, or participating in the labor force have mainly served as the outcome variables. Most papers interpret their findings with respect to the retirement decision, although the outcome variable in most cases is whether one works or participates in the labor force (Lumsdaine & Mitchell, 1999; Michaud, 2003). These outcome variables are advantageous compared to the date of retirement for several reasons. The date of retirement is observed either by the date on which the last employment spell ended in the administrative data, or by the self-reported retirement date in surveys. Using the last employment spell can be challenging, as the researcher cannot be certain that the individual actually retired and not just began an unemployment spell and thus was involuntarily out of work. Using the self-reported retirement date in surveys is problematic, as individuals may define retirement differently. Furthermore, between 26 and 53 percent of all individuals who retire re-enter the labor force at a later point, suggesting that the date of retirement may not be final (Maestas, 2010). Another obstacle is the problem of right-censoring. The researcher has to consider how to handle couples in which one of the partners is still working at the time of the observation or died prior to retirement. They can either be excluded from the estimation sample or defined as retired at the time of the observation, but neither of these solutions are optimal. Honoré & de Paula (2018) state that the estimates may change substantially depending on how the problem of right-censoring is handled, implying that the estimates are less robust.

I define the outcome variable as a binary variable taking the value of one if employed and zero otherwise. The findings of this thesis can be interpreted with respect to the retirement decision if assuming that all individuals who do not work are retired, requiring the absence of involuntary unemployment. The employment in the administrative data is observed monthly as a ratio where a full-time wage earner is given the value of one. As the number of paid working hours is reported by the tax authorities, I avoid systematical measurement errors and ensure that employment is defined similarly for each individual. Concerning individuals who work more than 37 hours a week, the number of actual working hours can deviate from the number of observed working hours, as these individuals may receive an annual or biannual bonus instead of overtime payment. This implies that this type of data is not suitable for analyzing intensive margin responses (the change in working hours) but is convenient to use to investigate extensive margin responses. I define an individual to be employed in a given month if (s)he works more than eight hours a week on average, as (s)he, in this case, is covered by the Danish Salaried Employees Act. Robustness checks are conducted in section 7.5.2 where other thresholds determine when individuals are defined as employed.

The running variable

I consider the date of birth as the running variable in the RDD. The date of birth is automatically registered by the authorities if born in Denmark, as the date of birth serves as the first six digits in each individual's social security number. Therefore, I observe the actual date of birth without (systematical) measurement errors, which is very important in this context.[20]

The covariates

I include educational dummies, a dummy for being a parent, and a dummy for being a grandparent of both the focal partner and the spouse. I include age-in-months dummies of the focal partner, as the focal partners' age differs between the four separate cutoffs. Furthermore, I include age-in-month dummies of the spouse to control for age-specific effects concerning labor market behavior. To control for spouses eligible to claim ERB at an older age due to the increase in the ERA, I include month-of-birth dummies of the spouse. I include dummies of the age-in-month difference between the partners since the couple's age structure has proven to be relevant for the retirement patterns of couples (Banks et al., 2010; García-Miralles et al., 2021; Legendre et al., 2018). Finally, I include regional dummies and a gender dummy, both only of the focal partner, since the variables are perfectly correlated between the focal partner and the spouse, as I observe only cohabiting couples and no same-gender couples.

6.2 Sample restrictions

I restrict the sample to obtain a meaningful estimate, as not all couples are affected by the increase in the ERA. Following the empirical strategy presented in section 5, I include all couples in which the focal partner is born within 100 days before or after the pooled cutoff in the baseline sample. The baseline sample contains approximately 84,000 observations in each of the six months where the focal partners in the control group and the treatment groups are the same age in months, but where only focal partners in the control group are eligible to claim ERB as shown in table 6.1.

In order to ensure that only individuals affected by the increase in the ERA are included in the estimation sample, I must identify individuals who are members of the ERP scheme. I employ a unique data set received from The Danish Ministry of Finance in which individuals are defined as members of the ERP scheme based on contributions since 1980 and possible repayments. This approach is more detailed than earlier studies on Danish data.[21] If individuals are wrongfully identified as members of the ERP scheme, the estimates will be downward biased, as some individuals in the control group cannot claim ERB upon reaching the ERA.

[20] To avoid systematical measurement errors, I exclude couples in which the focal partners were born outside of Denmark, as they are more likely to be given the birthday of January 1 or July 1. This procedure is described in section 6.2.

[21] Jørgensen (2014) defines individuals as members of the ERP scheme if they have been members of an unemployment insurance fund for at least one year. García-Miralles et al. (2021) define individuals as members of the ERP scheme if they have contributed to the ERP scheme at least once between the ages of 50 and 59. Both papers might identify individuals as members of the ERP scheme wrongfully, which may affect the estimates.

The Danish authorities use foreign authorities' information to register the date of birth if individuals migrate to Denmark. If an individual enters the country without a birth certificate, (s)he would, until recently, have been given the birthday of January 1 or July 1. This is a threat to identification since the running variable will show signs of bunching just after the cutoff, causing the assumption of no manipulation of the running variable (assumption RDD.2 from section 5.1) to be violated. To overcome this problem, I exclude couples in which the focal partners were born outside of Denmark.

Only the employment of wage earners is observed in the administrative data, and thus self-employed appear unemployed. To overcome this challenge, I exclude self-employed by excluding couples in which either the focal partner or the spouse had a surplus in a company of their own in the year of the observation.

In line with Bloemen et al. (2015) and Jørgensen (2014), I exclude couples in which either the focal partner or the spouse has left the labor force permanently prior to the ERA due to health issues, as these couples will not respond to the increase in the ERA. Therefore, I exclude couples in which either the focal partner or the spouse claimed disability pension benefits three years before the first of the six months of interest.[22] Individuals are evaluated physically and mentally by the authorities before they are awarded disability pension benefits, meaning that it is implausible that one will re-enter the labor force upon claiming disability pension benefits.

I restrict the sample to couples with approximately the same age structure following, e.g., Jørgensen (2014), García-Miralles et al. (2021) and Banks et al. (2010). This ensures that spouses are more likely to respond to the focal partners' employment decisions. It is very costly to retire for spouses that are much younger than the focal partners, and thus, this is not very likely. Identically, spouses will probably have retired if they are much older than the focal partners, e.g., due to health issues. Hence, only couples in which the age difference between the focal partner and the spouse is six years or less are included in the sample.

Only couples married prior to the Retirement Reform of 2011 are included in the sample. This ensures that individuals cannot respond to the changes in the ERA by choosing a partner that is born just before the cutoff, which may be optimal if spouses prefer to retire jointly with the focal partners at a younger age. Imposing the restriction of being married to the current spouse in 2010 at the latest furthermore increases the probability of observing stable marriages where couples are more likely to coordinate their retirement decision (Bloemen et al., 2015).

Finally, I restrict the sample to include couples where all relevant covariates of the focal partner and the spouse are observed. The final sample is an unbalanced panel containing

[22] This implies that I exclude couples if the focal partner or the spouse claimed disability pension benefits when the focal partner was 57 years at the first cutoff, 57.5 years at the second cutoff, 58 years at the third cutoff, and 58.5 years at the fourth cutoff.

Table 6.1: The number of observations conditional on the sample restrictions

	First	Second	Third	Fourth	Fifth	Sixth
Baseline	84,075	83,965	83,830	83,790	83,790	83,790
Members of ERP	44,220	44,175	44,131	44,125	44,125	44,125
Born in Denmark	43,202	43,163	43,126	43,120	43,120	43,120
No self-employed	36,077	36,068	36,080	36,102	36,102	36,102
No disability pension	33,823	33,778	33,747	33,747	33,747	33,747
Age difference	28,528	28,498	28,470	28,471	28,471	28,471
Married before reform	27,396	27,369	27,346	27,348	27,348	27,348
Variables must exist	27,260	27,228	27,198	27,200	27,200	27,200
Final (obs.)	27,260	27,228	27,198	27,200	27,200	27,200
Final (% of baseline)	32.4	32.4	32.4	32.5	32.5	32.5

Note: Table 6.1 depicts the number of observations conditional on the sample restrictions in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. Only surviving couples that remain married are included, implying that the number of observations differs between the months. If couples are self-employed during the first months, but subsequently cease to be, they are only part of the sample in the following months.

approximately 27,200 observations in each month of interest, corresponding to about 32.5 percent of the baseline sample.

6.3 Summary statistics

It is only an identifying assumption in the RDD that couples in the vicinity of the cutoff are identical. However, it is reassuring if the observable characteristics are similar between couples in the entire treatment group and control group. Only unobservable characteristics, e.g., health, are expected to differ between the couples in the treatment group and the control group following the arguments presented in section 9.2.

I confirm in table 6.2 that the observable covariates are almost identical between couples in the treatment group and the control group. In the table, I show the means of the focal partners' and spouses' covariates conditional on being in the treatment group or control group.

More focal partners are women in both the treatment group and control group, and the spouses are on average 0.3 years older than the focal partners. The educational levels are almost identical between the treatment group and control group, and the same applies to the region of residence. The focal partners and spouses in the control group are, on average, one percentage point more likely to be parents and two percentage points more likely to be grandparents.

Table 6.2: Summary statistics of the couples in the estimation sample

	Treatment group		Control group	
	Focal partner	Spouse	Focal partner	Spouse
Man	0.47	0.53	0.47	0.53
Age	61.0	61.3	61.0	61.3
Unskilled	0.25	0.21	0.25	0.21
Secondary education	0.03	0.03	0.03	0.03
Skilled	0.42	0.47	0.43	0.47
Short-cycle higher education	0.04	0.05	0.04	0.04
Medium-cycle higher education	0.21	0.20	0.21	0.20
Long-cycle higher education	0.06	0.06	0.06	0.06
Copenhagen	0.20	0.20	0.21	0.21
Zealand	0.14	0.14	0.15	0.15
Southern Denmark	0.26	0.26	0.26	0.26
Northern Denmark	0.13	0.13	0.13	0.13
Central Jutland	0.26	0.26	0.25	0.25
Parent	0.92	0.92	0.93	0.93
Grandparent	0.66	0.66	0.68	0.68

Note: Table 6.2 depicts the means of the covariates of the couples in the estimation sample conditional on whether they are in the treatment group or the control group. All covariates are measured in the first of the six possible months where the focal partners in the control group and the treatment group are the same age in months, but where only the focal partners in the control group are eligible to claim ERB.

7 Results

Following the empirical strategy outlined in section 5, I estimate whether leisure complementarities cause couples to retire jointly and thus if raising the focal partners' ERA increases the spouses' employment. I exploit the exogenous variation in the ERA caused by the Welfare Agreement of 2006 and the Retirement Reform of 2011. I present visual evidence of the retirement behavior of couples that are not affected by the reforms, and figures that graphically show how the focal partners' and spouses' employment responds to the increase in the ERA. Thereupon, I estimate how raising the ERA affects the focal partners' and spouses' employment. Finally, I investigate the robustness of the results.

7.1 Graphical evidence

I investigate whether couples tend to retire jointly by showing the difference in the year of retirement of partners in older couples that are not affected by the reforms. Even if couples tend to retire jointly, other factors than complementarities in leisure may explain this, implying that the evidence is not causal. To overcome this problem, I exploit the exogenous variation in the ERA. A graph is presented to examine a possible discontinuity visually and to choose a suitable functional form in the estimations.

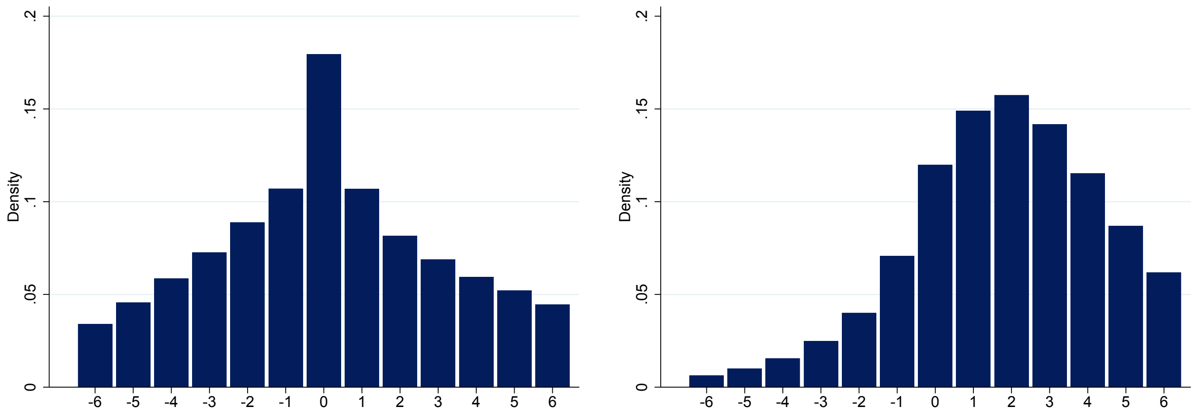
7.1.1 Couples with a stable retirement age

It is only reasonable to believe that the spouse's leisure and the focal partner's leisure are complements if couples retired jointly prior to the reforms. Thus, I consider couples in which both partners were born in 1953 at the latest, ensuring that the ERA remained stable at 60 years for everyone.[23] I consider the year in which the last employment spell ended as the year of retirement since the employment data prior to 2008 are only recorded annually, although I am aware of the problems of using the last employment spell to define the year of retirement as outlined in section 6. There is a breakpoint in 1986, causing me to consider a sample with yearly observations from 1986 to 2018. To deal with the problem of right-censoring, I exclude couples if one of the partners died while still being employed or was still employed in 2018. This is not ideal as described in section 6, but since the graph is only applied as an indication of whether couples retire jointly, the problem of right-censoring is ignored.

Figure 7.1: The propensity to retire jointly prior to the reforms

(a) Difference in the year of retirement (man - woman)

(b) Age difference between the partners (man - woman)



Note: Only couples in which the difference between the year of retirement is six years or less are included in the sample. Similarly, only couples with an age difference of six years or less are included in the sample. The age difference is based on the year of birth.

Figure 7.1a shows that couples tend to retire within the same year. As depicted in figure 7.1b, this cannot be explained by a similar age structure of the couple, as men, on average, are older than their wives. Leisure complementarities can explain why couples tend to retire jointly, but figure 7.1 does not provide causal evidence of this, as couples may retire jointly because of similar preferences, as argued in section 2. If couples retire jointly because of identical preferences and not due to leisure complementarities, spouses do not respond to the focal partners'

[23] The ERA is unchanged in the period, but several other changes were introduced, e.g., the two-year rule. As figure 7.1 is only used to show signs of joint retirement, I ignore this.

retirement decision. In this case, raising the retirement age does not affect the spouses' employment. Therefore, final conclusions cannot be drawn based on figure 7.1, but the figure sets the groundwork for the rest of this section.

7.1.2 Couples affected by the increase in the ERA

An advantage of the RDD is that a graph can be presented to investigate the effects of a policy change visually. The graph sums up the purpose of the research design since a discontinuity at the cutoff suggests that the spouses' employment responds to the increase in the ERA. Another advantage of presenting such a graph is to visually investigate the data, ensuring that I choose the correct functional form, i.e., the right specification of $f(x_i)$ in the estimations. Finally, a graph shows if discontinuities are present at other values than the cutoff, which will violate the identifying assumption of continuity in the conditional expectation function (RDD.1 from section 5.1). Lee & Lemieux (2010) emphasize that the visual evidence cannot stand alone and is only meant to serve as an indication of whether an effect is present. As standard errors are not reported in the graph, no conclusions can be drawn based on the figure, as it is impossible to determine the statistical significance of a potential discontinuity.

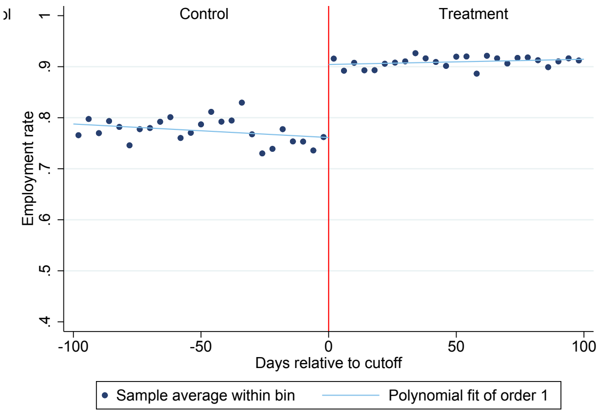
I divide the running variable into equal-sized intervals, known as *bins*. The focal partners' average employment within each bin summarizes the relationship between the focal partners' employment and the number of days they were born relative to the cutoff. Similarly, the spouses' average employment within each bin summarizes the relationship between the spouses' employment and the number of days the focal partners were born relative to the cutoff. The size of the bins is thus rather important; too narrow bins make the graph too noisy, while too large bins may erase a discontinuity. I choose the bin size based on a rule proposed by Lee & Lemieux (2010) and Jacob et al. (2012).[24] Their approach is to compare two different bin sizes, and if the smaller bin size does not ensure a better fit of the model, the larger bin size is preferred. In this case, the approach implies that I consider a bin size of 4 days, meaning that I consider 25 bins on each side of the cutoff as depicted in figure 7.2 and figure 7.3.

A linear polynomial that differs on each side of the cutoff looks to properly fit the data in figure 7.2. It is tempting to use high-order polynomials in figure 7.3 to ensure a better fit of the model, but this will probably overfit the model following the arguments by Gelman & Imbens (2019) and Gelman & Zelizer (2015). When using a relatively narrow bandwidth of 100 days, there are no sound arguments for using high-order polynomials. However, I investigate whether the estimates change substantially if applying high-order polynomials to examine the robustness of the estimates in section 7.3.

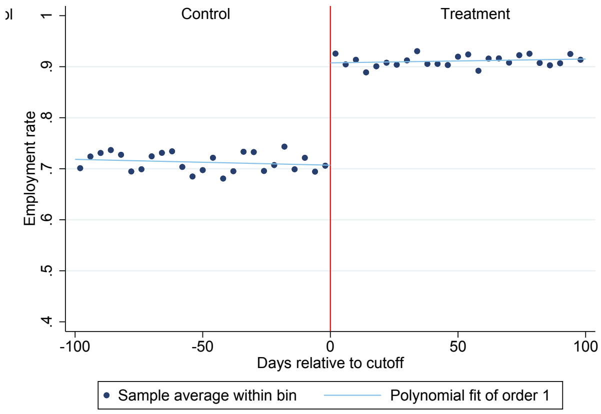
[24] The rule is implemented as a Stata routine by Calonico et al. (2017) and is given by *rdrobust*.

Figure 7.2: The focal partners' employment as a function of the number of days the focal partners are born relative to the cutoff

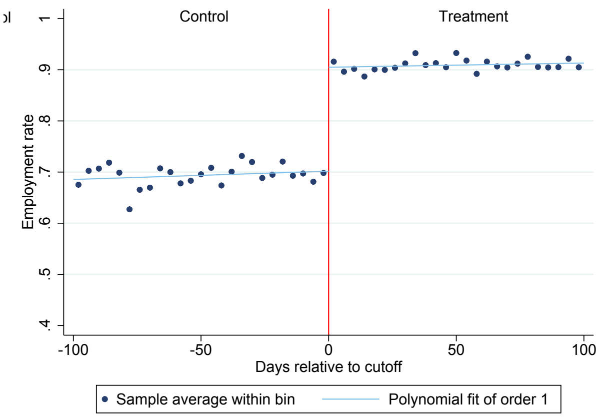
(a) First month



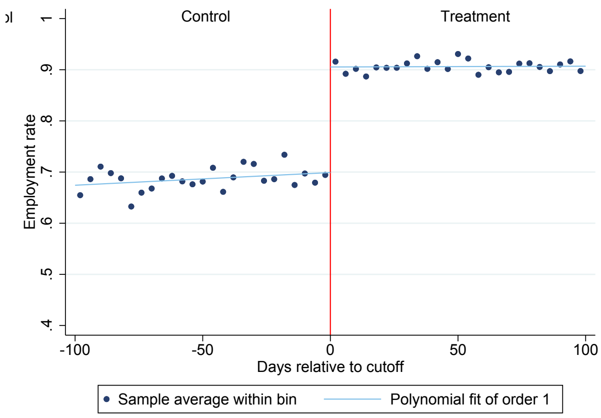
(b) Second month



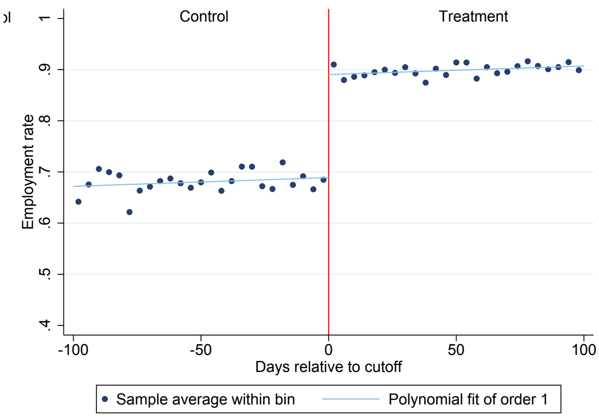
(c) Third month



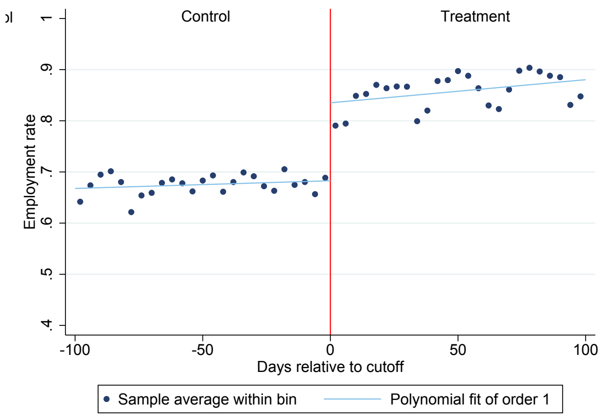
(d) Fourth month



(e) Fifth month



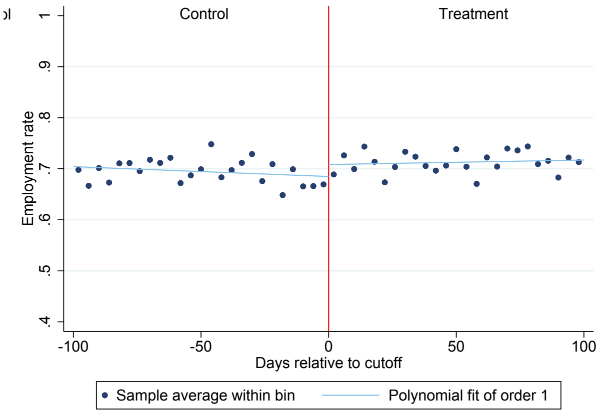
(f) Sixth month



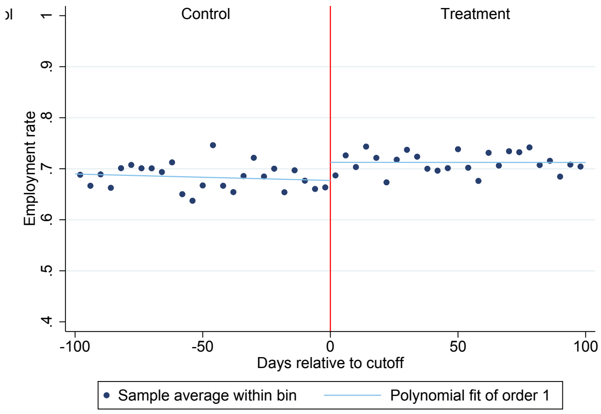
Note: Figure 7.2 shows the focal partners' employment as a function of the number of days the focal partners are born relative to the cutoff in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. A bin size of 4 days and a first-order polynomial are used in all figures.

Figure 7.3: The spouses' employment as a function of the number of days the focal partners are born relative to the cutoff

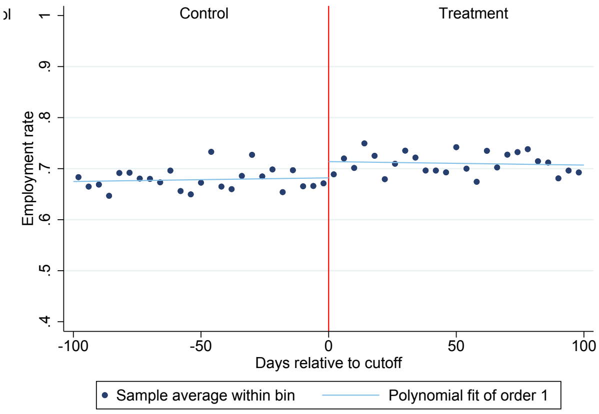
(a) First month



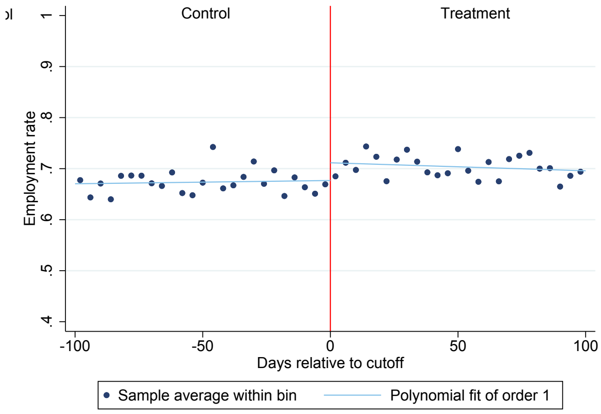
(b) Second month



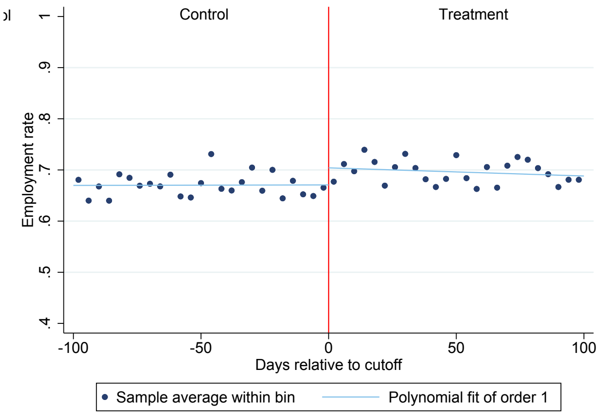
(c) Third month



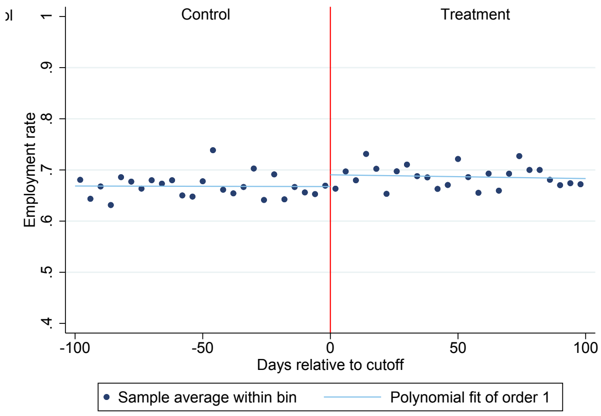
(d) Fourth month



(e) Fifth month



(f) Sixth month



Note: Figure 7.3 shows the spouses' employment as a function of the number of days the focal partners are born relative to the cutoff in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. A bin size of 4 days and a first-order polynomial are used in all figures.

The focal partners respond to the eligibility rules, as raising the ERA causes the focal partners' employment to increase considerably, like depicted in figure 7.2. The change in the spouses' employment is less visible, but a small discontinuity looks to be present in all of the six months as depicted in figure 7.3. This indicates that raising the retirement age also increases the spouses' employment. Still, as no standard errors are reported, it is impossible to conclude if the effect is statistically significant.

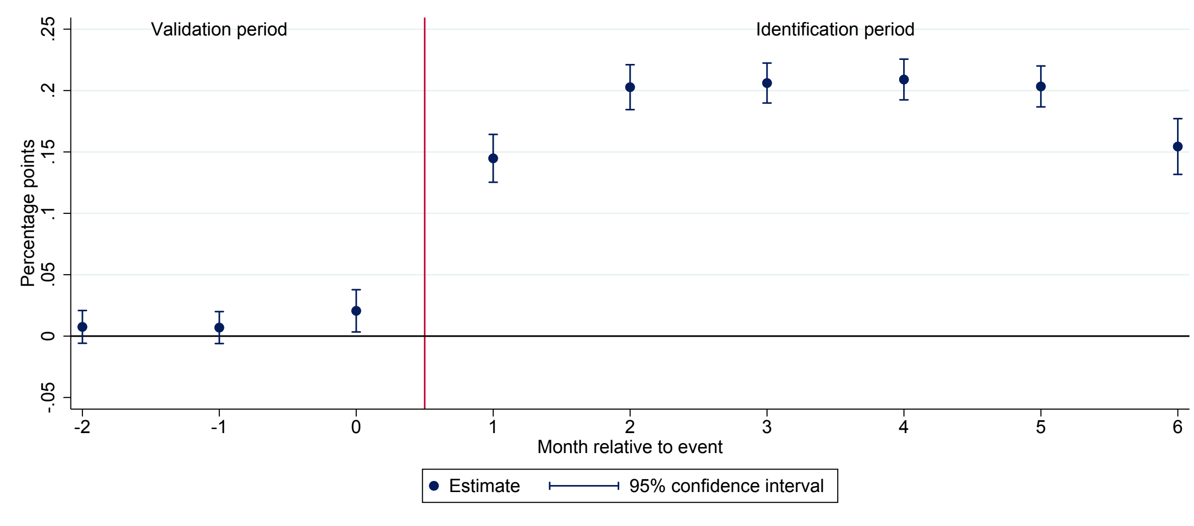
7.2 The change in the focal partners' employment

The visual evidence cannot stand alone, as a large discontinuity can be combined with large standard errors, leaving the estimate statistically insignificant. Hence, I estimate model 5.1 from section 5.3.2 to investigate the change in the focal partners' employment caused by the increase in the ERA. The estimates are depicted in figure 7.4 three months before and six months after the focal partners in the control group become eligible to claim ERB, while the focal partners in the treatment group are ineligible in the entire period although they are the same age in months. I denote the months prior to treatment, i.e., the months in which neither the focal partners in the control group nor in the treatment group are eligible to claim ERB, as *the validation period*. Similarly, I denote the months after treatment, i.e., the months in which the focal partners in the control group are eligible to claim ERB, while the focal partners in the treatment group are ineligible as *the identification period*. If the estimates are to be given a causal interpretation, the difference between the employment of the focal partners in the control group and the treatment group has to be insignificant in the validation period. Otherwise, the difference in the focal partners' employment in the identification period cannot solely be attributed to the increase in the ERA.

I find that raising the ERA increases the focal partners' employment by approximately 20 percentage points in the second to the fourth month after the focal partners in the control group become eligible to claim ERB. The focal partners' employment increases by approximately 15 percentage points in the first and sixth month, as depicted in figure 7.4. The lower estimate in the first month is likely caused by focal partners in the control group that use more than one month to transition into retirement. Equivalently, the lower estimate in the sixth month may result from focal partners in the treatment group transitioning into retirement, although they cannot claim ERB for the entire month.

The estimates are insignificant at a five percent level in the validation period except for the month in which the focal partners in the control group are only eligible to claim ERB for less than the entire month. The estimate in this month is relatively small compared to the estimates in the following months. Hence, the employment is not substantially different between the focal partners in the treatment group and the control group when both groups are ineligible to

Figure 7.4: The change in the focal partners' employment



Note: Figure 7.4 shows the estimated values of $\hat{\gamma}$ from model 5.1 three months before and six months after the focal partners in the control group become eligible to claim ERB, while the focal partners in the treatment group are ineligible in the entire period although they are the same age in months. The estimations are conducted with a bandwidth of 100 days, a linear polynomial that is allowed to differ on each side of the cutoff, and standard errors clustered on the running variable. The validation period consists of the last three months in which neither the focal partners in the control group nor the treatment group are eligible to claim ERB. The identification period consists of the six months in which the focal partners in the control group are eligible to claim ERB, while the focal partners in the treatment group are ineligible although they are the same age in months. The estimates and standard errors are reported in table A.1 in appendix A.2.

claim ERB. This ensures that the estimates in the identification period can be given a causal interpretation.

7.3 The change in the spouses' employment

Section 7.1.1 revealed that couples retired jointly prior to the reforms, while section 7.2 proved that raising the ERA increases the focal partners' employment. If couples retired jointly prior to the reforms due to leisure complementarities, raising the ERA may increase the spouses' employment. Hence, I estimate model 5.2 from section 5.3.2 to obtain the reduced-form estimates of the change in the spouses' employment caused by the increase in the ERA.

7.3.1 The average change in the spouses' employment

Table 7.1 depicts the change in the spouses' employment caused by the increase in the ERA. The estimates are shown in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. In the table, I investigate the importance of including controls, the choice of the functional form, and the choice of bandwidth.

Table 7.1: The change in the spouses' employment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First month	0.023**	0.024**	0.024**	0.026**	0.047***	0.039*	0.013*	0.042**
	(0.012)	(0.011)	(0.010)	(0.010)	(0.017)	(0.024)	(0.007)	(0.016)
	27,260	27,260	27,260	27,260	27,260	27,260	53,967	13,322
Second month	0.036***	0.035***	0.035***	0.037***	0.038**	0.034	0.026***	0.038**
	(0.012)	(0.011)	(0.010)	(0.010)	(0.017)	(0.023)	(0.007)	(0.016)
	27,228	27,228	27,228	27,228	27,228	27,228	53,935	13,301
Third month	0.032***	0.031***	0.031***	0.034***	0.032**	0.038*	0.028***	0.036**
	(0.011)	(0.011)	(0.009)	(0.010)	(0.015)	(0.022)	(0.006)	(0.015)
	27,198	27,198	27,198	27,198	27,198	27,198	53,905	13,301
Fourth month	0.035***	0.034***	0.034***	0.037***	0.045***	0.039*	0.028***	0.045***
	(0.012)	(0.011)	(0.010)	(0.010)	(0.017)	(0.023)	(0.007)	(0.016)
	27,200	27,200	27,200	27,200	27,200	27,200	53,877	13,301
Fifth month	0.033***	0.033***	0.033***	0.036***	0.048***	0.046**	0.024***	0.049***
	(0.011)	(0.011)	(0.010)	(0.010)	(0.016)	(0.023)	(0.007)	(0.015)
	27,200	27,200	27,200	27,200	27,200	27,200	53,834	13,301
Sixth month	0.023**	0.023**	0.023**	0.026**	0.034**	0.027	0.015**	0.036**
	(0.012)	(0.011)	(0.010)	(0.010)	(0.017)	(0.025)	(0.007)	(0.016)
	27,200	27,200	27,200	27,200	27,200	27,200	53,779	13,301
Controls (focal)		✓		✓	✓	✓	✓	✓
Controls (spouse)			✓	✓	✓	✓	✓	✓
Polynomial	Linear	Linear	Linear	Linear	Quadratic	Cubic	Linear	Linear
Bandwidth	100	100	100	100	100	100	200	50

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table 7.1 reports the estimate of interest ($\hat{\tau}$) from model 5.2 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. In column (1), I set $\beta_2 = \beta_3 = 0$ from model 5.2 prior to estimation. In column (2), I set $\beta_3 = 0$ from model 5.2 prior to estimation. In column (3), I set $\beta_2 = 0$ from model 5.2 prior to estimation. Column (4) shows the baseline specification with controls of both the focal partner and the spouse, a linear polynomial that is allowed to differ on each side of the cutoff, and a bandwidth of 100 days. Column (5) shows the estimates if using a quadratic polynomial, while column (6) shows the estimates if using a cubic polynomial, both where the slopes are allowed to differ on each side of the cutoff. Column (7) and (8) shows the estimates if using a bandwidth of 200 and 50 days respectively. See appendix A.1 to obtain the estimates of all parameters.

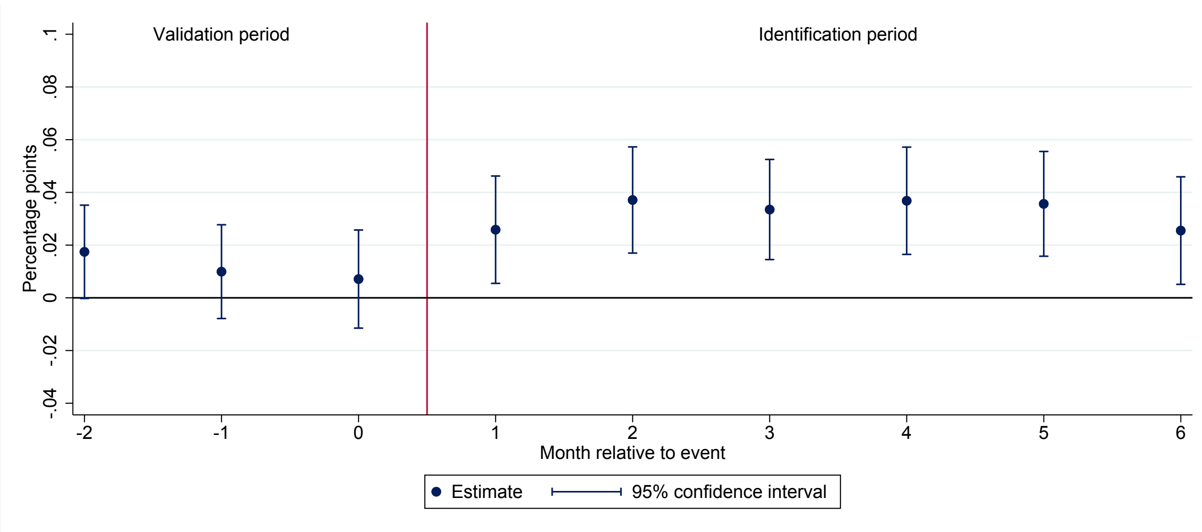
In my preferred specification, I include controls of both the focal partner and the spouse, use a linear polynomial that is allowed to differ on each side of the cutoff, and consider a bandwidth of 100 days (column (4) in the table 7.1). In this specification, the estimates are between 0.034 and 0.037 in the second to fifth month and significant at a one percent level. The estimates are 0.026 and significant at a five percent level in the first and sixth month. This implies that raising the ERA increases the spouses' employment by 2.6 to 3.7 percentage points in reduced-form. The spouses' employment increases the most in the second to the fifth month. When borne in mind that this corresponds with the estimates provided concerning the focal partner, these lower estimates in the first and sixth month prove a link between the change in the focal partners' employment and the change in the spouses' employment.

The estimates are almost identical when including controls, but the standard errors are smaller, emphasizing the importance of including controls to increase the statistical power. Using a quadratic or cubic polynomial only affects the estimates substantially in the first month but increases the standard errors considerably in all months. As high-order polynomials do not change the estimates substantially, the misspecification bias is not worrying, as argued in section 5.3.3. A smaller bandwidth does not affect the estimates considerably. Equivalently, the broader bandwidth of 200 days reduces the estimates slightly but does not affect the significance of the estimates.

The estimates in table 7.1 reveal that neither the inclusion of controls, the choice of functional form, nor the choice of bandwidth has a substantial impact on the estimates. This suggests that the estimates are robust to the different specifications, confirming that the research design is suitable. The standard errors appear larger if clustered on the running variable relative to clustered on the couple level or using White robust standard errors, as depicted in table A.2 in appendix A.2. The statistical significance is not affected by the type of applied standard errors, as the difference in the estimated standard errors is small. I follow a conservative approach to ensure that I do not overestimate the significance as argued in section 5.3.5, and thus cluster the standard errors on the running variable.

If the estimates are to be given a causal interpretation, the difference between the employment of the spouses in the treatment group and the control group has to be insignificant in the validation period. Figure 7.5 shows the change in the spouses' employment with a corresponding 95 percent confidence interval in the validation period and in the identification period where both periods are defined in the same way as in section 7.2. The estimates in the identification period are identical to those in my preferred specification in table 7.1. All estimates in the validation period are insignificant at a 5 percent level. This means that the difference between the employment of the spouses in the control group and the treatment group is insignificant in the months in which neither the focal partners in the control group nor the treatment group are eligible to claim ERB. The estimates support the hypothesis that the spouse's leisure and the focal partner's leisure are complements and thus that leisure complementarities cause spouses to be more likely to work if the focal partners work. Equivalently, the results can be given the interpretation that spouses are more likely to retire if the focal partners retire, implying that leisure complementarities cause couples to retire jointly.

Figure 7.5: The change in the spouses' employment



Note: Figure 7.5 shows the estimated values of $\hat{\tau}$ from model 5.2 three months before and six months after the focal partners in the control group become eligible to claim ERB. The estimations are conducted with a bandwidth of 100 days, a linear polynomial that is allowed to differ on each side of the cutoff, and standard errors clustered on the running variable. The validation period consists of the last three months in which neither the focal partners in the control group nor the treatment group are eligible to claim ERB. The identification period consists of the six months in which the focal partners in the control group are eligible to claim ERB, while the focal partners in the treatment group are ineligible although they are the same age in months. The estimates and standard errors are reported in table A.1 in appendix A.2.

It is possible to utilize the same research design but exploit the increase in the age at which individuals can apply the two-year rule or the increase in the NRA. Figure A.4 in appendix A.3 reveals that raising the age at which individuals are eligible to apply the two-year rule increases the focal partners' employment modestly, while the increase in the spouses' employment is insignificant as a result. Raising the NRA increases the focal partners' employment considerably, while the change in the spouses' employment is insignificant. This results from large standard errors caused by few observations, as I only observe few couples in which neither the focal partner nor the spouse is a member of the ERP scheme. Furthermore, only focal partners at the first cutoffs have reached the NRA in the data, causing the number of observations in the sample to be even lower. As a result, I limit the focus to the effects of raising the ERA in the rest of this thesis.

7.3.2 The change in the spouses' propensity to claim benefits

In the previous section, I concluded that raising the ERA increases the spouses' employment. An increase in the spouses' employment is likely to result in fewer spouses claiming benefits. In column (5) in table 7.2, I show that spouses in the treatment group are less likely to claim ERB, suggesting that spouses in the treatment group substitute from claiming ERB to working due to the increase in the ERA. This implies that raising the ERA affects the government budget through two channels: the tax revenue increases since spouses are more likely to work, while the public expenditures decrease since fewer spouses claim benefits.

Table 7.2: The change in the spouses' propensity to claim benefits

	(1)	(2)	(3)	(4)	(5)
First month	0.004	0.009*	0.000	-0.000	-0.019**
	(0.005)	(0.005)	(0.000)	(0.001)	(0.008)
	27,260	27,260	27,260	27,260	27,260
Second month	0.006	0.007	0.000	0.000	-0.021**
	(0.005)	(0.006)	(0.000)	(0.001)	(0.008)
	27,228	27,228	27,228	27,228	27,228
Third month	0.002	0.003	0.000	0.000	-0.020**
	(0.005)	(0.006)	(0.000)	(0.001)	(0.008)
	27,198	27,198	27,198	27,198	27,198
Fourth month	0.002	0.002	0.000	-0.000	-0.022***
	(0.005)	(0.005)	(0.000)	(0.001)	(0.008)
	27,200	27,200	27,200	27,200	27,200
Fifth month	0.002	-0.001	0.000	-0.000	-0.025***
	(0.005)	(0.005)	(0.000)	(0.001)	(0.008)
	27,200	27,200	27,200	27,200	27,200
Sixth month	-0.004	0.000	0.000	-0.000	-0.020**
	(0.005)	(0.005)	(0.000)	(0.001)	(0.008)
	27,200	27,200	27,200	27,200	27,200
Outcome variable	SS	UB	SA	DBP	ERB

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table 7.2 reports the estimate of interest ($\hat{\tau}$) from model 5.2 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. "SS" denotes self-support without working, "UB" denotes unemployment benefits, "SA" denotes social assistance, "DBP" denotes disability pension, and "ERB" denotes early retirement benefits. See appendix A.1 to obtain the estimates of all parameters.

Spouses are not less likely to claim other benefits or more likely to be self-supporting without working, which is evident based on the rules of the ERP scheme. If the spouse is older than the focal partner, (s)he can only retire on pension benefits. If the spouse is younger than the focal partner, retiring on unemployment benefits or social assistance is not ideal, as the spouse will have to waive five years of ERB. Johnsen et al. (2015) find that younger spouses retire jointly with the focal partners by claiming disability pension benefits. I do not reach the same conclusion, as spouses in the treatment group are not less likely to claim disability pension benefits. Individuals are evaluated physically and mentally by the authorities upon their request to be awarded disability pension benefits. The conclusion that spouses do not retire jointly on disability pension benefits suggests that, although spouses may prefer to retire jointly with the focal partners on disability pension benefits, they cannot credibly lie about their health, indicating that the problem of moral hazard is not present in this context.[25]

7.3.3 The change in the spouses' employment conditional on gender and age structure of the couple

Whether men or women value joint leisure the most has been investigated in several papers as described in section 2. However, the importance of the couple's age structure has been highlighted to a limited extent only. Legendre et al. (2018) stress that couples' age structure explains the retirement pattern of couples. They argue that spouses who are younger than the focal partners do not respond to the focal partner's retirement decision, as it is very costly to retire prior to the retirement age. Table 7.3 confirms this, as raising the ERA solely increases the employment of older spouses.

The finding that only the employment of older spouses increases corresponds with the results from table 7.2, which showed that spouses in the treatment group are less likely to claim ERB, while the propensity to claim other benefits remains unaffected. The conclusion that raising the ERA does not affect the younger spouses' employment must be understood in connection with the ERP scheme's eligibility rules, on the grounds of which individuals will waive five years of ERB if retiring prior to the ERA. As it is thus very costly to retire if the spouse is younger than the focal partner, couples who want to retire jointly postpone retirement until both partners are eligible to claim pension benefits.

Following Coile (2004), I find that men are more likely than women to postpone retirement if the focal partner is affected by the increase in the retirement age. This suggests that men value joint leisure more than women, or that the genders differ in terms of other characteristics. Investigating the employment response conditional on both gender and the couple's age structure, I find that neither men's nor women's employment increases as a result of raising the ERA provided that they are younger than the focal partners. The employment of men and women increases with the same magnitude by raising the ERA provided that they are older than the focal partner. However, the estimates are only significant for men due to large standard errors concerning women resulting from a smaller number of observations. This suggests that men and women do not value joint leisure differently but that the couples' age structure causes men to be more likely to respond to the focal partner's retirement decision in reduced-form.

[25] The spouses' health is private information, causing the presence of information asymmetry. If spouses can credibly lie about their health to retire jointly with the focal partners on disability pension benefits, the problem of moral hazard would arise, as spouses do not bear the full costs of retiring early.

Table 7.3: The change in the spouses' employment conditional on gender and age structure of the couple

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First month	0.036**	-0.001	0.011	0.025	0.041	0.036*	0.001	-0.011
	(0.017)	(0.011)	(0.013)	(0.016)	(0.036)	(0.019)	(0.013)	(0.021)
	14,244	12,123	12,593	14,088	3,095	11,149	9,338	2,785
Second month	0.048***	0.007	0.014	0.042***	0.037	0.052***	0.007	0.008
	(0.017)	(0.011)	(0.013)	(0.016)	(0.037)	(0.018)	(0.013)	(0.022)
	14,242	12,090	12,569	14,079	3,099	11,143	9,309	2,781
Third month	0.045***	0.003	0.009	0.040**	0.041	0.046**	-0.002	0.019
	(0.017)	(0.010)	(0.013)	(0.016)	(0.036)	(0.019)	(0.012)	(0.021)
	14,239	12,062	12,543	14,068	3,100	11,139	9,285	2,777
Fourth month	0.046***	0.008	0.018	0.037**	0.037	0.047**	0.011	-0.004
	(0.017)	(0.010)	(0.013)	(0.017)	(0.038)	(0.020)	(0.012)	(0.021)
	14,237	12,068	12,551	14,064	3,103	11,134	9,290	2,778
Fifth month	0.046***	0.007	0.017	0.038**	0.049	0.045**	0.005	0.010
	(0.018)	(0.010)	(0.012)	(0.017)	(0.036)	(0.021)	(0.011)	(0.020)
	14,237	12,068	12,551	14,064	3,103	11,134	9,290	2,778
Sixth month	0.041**	-0.003	0.005	0.031*	0.024	0.044**	-0.000	-0.011
	(0.018)	(0.010)	(0.012)	(0.017)	(0.037)	(0.021)	(0.012)	(0.020)
	14,237	12,068	12,551	14,064	3,103	11,134	9,290	2,778
Gender (spouse)			Woman	Man	Woman	Man	Woman	Man
Age (spouse)	Oldest	Youngest			Oldest	Oldest	Youngest	Youngest

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table 7.3 reports the estimate of interest ($\hat{\tau}$) from model 5.2 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. See appendix A.1 to obtain the estimates of all parameters.

7.4 The spillover effect

The change in the spouses' employment depends on the change in the focal partners' employment. Hence, I estimate the spillover effect from the focal partners' employment to the spouses' employment since it is essential for the effects of raising the retirement age in the future. Thus, I utilize the FRDD and estimate model 5.3 from section 5.3.2. Recalling the definition of LATE, the effects are only obtained for the compliant subpopulation.

7.4.1 The average spillover effect

I find that the spouses' employment is 16.5 to 18.3 percentage points larger if the focal partners are employed as a result of the increase in the ERA as depicted in table 7.4. Hence, the spillover effect is 16.5 to 18.3 percentage points, meaning that when the focal partners' employment increases by 1 percentage point, the spouses' employment increases by 0.165 to 0.183 percentage points.

Table 7.4: The spillover effect

	(1)	(2)
First month	0.178**	0.086**
	(0.071)	(0.035)
	27,260	27,260
Second month	0.183***	0.083***
	(0.051)	(0.032)
	27,228	27,228
Third month	0.163***	0.077***
	(0.046)	(0.030)
	27,198	27,198
Fourth month	0.176***	0.081***
	(0.049)	(0.029)
	27,200	27,200
Fifth month	0.175***	0.091***
	(0.049)	(0.029)
	27,200	27,200
Sixth month	0.165**	0.083***
	(0.065)	(0.032)
	27,200	27,200
Outcome variable	Employment	ERB

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table 7.4 reports the estimate of interest ($\hat{\rho}$) from model 5.3 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. The corresponding F-statistics from the first stage are reported in table A.6.

The spillover effect can be attributed to leisure complementarities, as the spouses do not have economic incentives to retire jointly with the focal partners since the focal partners' ERB are not means-tested based on the spouses' income, and the spouses do not receive pension benefits when the focal partners retire. Ignoring leisure complementarities thus underestimates the positive effect of raising the retirement age on the employment by 14 to 15 percent.[26] Neither the inclusion of controls, the choice of functional form, nor the choice of bandwidth has a substantial impact on the estimates as shown in table A.3 in appendix A.2, emphasizing the robustness of the estimates.

Table 7.4 reveals that spouses also respond to the focal partner's decision to claim benefits. If the share of focal partners who claim ERB drops by 1 percentage point, the share of spouses who claim ERB drops by 0.08 to 0.09 percentage points. Hence, ignoring leisure complementarities underestimates the decrease in the number of individuals who claim ERB by 7 to 8 percent. This suggests that a spillover effect from the focal partners to the spouses is observed regardless of whether the decision to work or claim pension benefits is considered the outcome variable.

The rules of the ERP scheme may explain why spouses respond more to the focal partners' decision to work than to claim ERB. Spouses can work a limited number of hours and still claim ERB, suggesting that approximately half of the spouses who choose to work if the focal partners work still claim ERB. This indicates that the increase in the spouses' employment is larger along the extensive margin relative to along the intensive margin, as these spouses are not full-time employed.

7.4.2 The spillover effect conditional on gender and age structure of the couple

Utilizing the FRDD is attractive when estimating the spillover effect by subgroups, as the change in the spouses' employment is measured relative to the change in the focal partners' employment. This is ideal since a low estimate of the change in the spouses' employment in reduced-form can be caused by a small spillover effect, but also by a less considerable response from the focal partners.

Table 7.5 confirms that the spillover effect depends on the age structure of the couple. The spouses are 21 to 26 percentage points more likely to be employed if the focal partners are employed provided that the spouses are the oldest in the couple. The estimates are highly significant in each of the six months of interest. The estimates in column (2) reveal that spouses are not significantly more likely to be employed if the focal partners are employed provided that the spouses are the youngest in the couple. This confirms the conclusions from section 7.3.3, which stated that couples who want to retire jointly do so by postponing retirement until both partners are eligible to claim pension benefits. While only a positive spillover effect is found for spouses that are older than the focal partners, this does not necessarily imply that younger spouses value joint leisure less. In the theoretical model from section 4, only older spouses respond to the focal partner's employment decision, although all spouses value joint

[26] This is calculated in the following way: $\frac{0.16}{0.16+1} \cdot 100 = 13.8$ and $\frac{0.18}{0.18+1} \cdot 100 = 15.3$.

leisure identically. Hence, the conclusion that only older spouses respond to the focal partner's employment decision is a result of economic incentives, as retiring prior to the ERA is associated with substantial economic costs.

Table 7.5: The spillover effect conditional on gender and age structure of the couple

	(1)	(2)	(3)	(4)	(5)	(6)
First month	0.260***	0.022	0.180	0.181**	0.394	0.243***
	(0.090)	(0.104)	(0.142)	(0.080)	(0.345)	(0.093)
	14,550	12,387	12,868	14,392	3,159	11,391
Second month	0.235***	0.075	0.154	0.198***	0.237	0.237***
	(0.062)	(0.080)	(0.102)	(0.056)	(0.220)	(0.063)
	14,549	12,354	12,844	14,384	3,162	11,387
Third month	0.214***	0.043	0.108	0.184***	0.238	0.211***
	(0.060)	(0.074)	(0.099)	(0.054)	(0.207)	(0.064)
	14,549	12,330	12,820	14,378	3,163	11,386
Fourth month	0.213***	0.090	0.200**	0.164***	0.247	0.204***
	(0.062)	(0.073)	(0.100)	(0.057)	(0.223)	(0.066)
	14,546	12,335	12,827	14,373	3,166	11,380
Fifth month	0.212***	0.083	0.218**	0.157***	0.438	0.184***
	(0.064)	(0.072)	(0.108)	(0.056)	(0.271)	(0.067)
	14,546	12,335	12,827	14,373	3,166	11,380
Sixth month	0.238***	0.007	0.157	0.162**	0.330	0.221***
	(0.082)	(0.103)	(0.169)	(0.068)	(0.441)	(0.083)
	14,546	12,335	12,827	14,373	3,166	11,380
Gender (spouse)			Woman	Man	Woman	Man
Age (spouse)	Oldest	Youngest			Oldest	Oldest

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table 7.5 reports the estimate of interest ($\hat{\rho}$) from model 5.3 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. See appendix A.1 to obtain the estimates of all parameters.

The spouses' economic incentives to retire jointly with the focal partners do not deviate between the genders if accounting for the couples' age structure. Hence, the estimates in column (3) and (4) indicate that the genders value joint leisure similarly, as the spillover effect is within the same range for men and women. However, the estimates are only significant for men in all months since the estimated standard errors are larger for women. Restricting the sample to spouses that are older than the focal partners suggests that women are more likely to respond to the focal partner's employment decision. The estimates for women are, once again, insignificant as a result of large standard errors due to fewer observations. I find no evidence that men and women value joint leisure differently. When taking the couples' age structure into account, the magnitude of the estimates suggests that women value joint leisure the most, while the opposite conclusion is reached if considering the statistical significance of the estimates.

7.4.3 The spillover effect by different subgroups

The predictions from the theoretical model derived in section 4 indicate that spouses with lower incomes are more likely to respond to the focal partner's employment decision, everything else equal. Table 7.6 confirms that spouses respond more to the focal partner's employment decision if they earn less than the focal partner.[27] This indicates that the partner with the highest income has the greatest bargaining power, implying that the partner with the highest income is more likely to determine when the couple retires jointly. This finding corresponds with the conclusions from Michaud et al. (2019).

For the time being, I have not been able to reject the simplified assumption from the theoretical model that spouses value joint leisure identically. In order to understand the differences in the retirement patterns of couples, it is crucial to investigate whether specific characteristics of couples are associated with a higher utility of joint leisure. Lalive & Parrotta (2017) and Kruse (2020) stress that spouses value joint leisure more if they are married to focal partners who have the same characteristics as themselves. Hence, spouses in homogeneous couples are expected to respond to the focal partner's employment decision to a greater extent, as their utility of leisure is greater if retiring jointly with the focal partner. Based on the estimates reported in table 7.6, I confirm that spouses in homogeneous couples, defined as couples in which the partners have the same educational level, respond more to the focal partner's employment decision.[28] This indicates that homogeneous couples value joint leisure more than heterogeneous couples. This proves the importance of incorporating heterogeneity in the utility of joint leisure in structural models provided that the aim is to investigate the (joint) retirement behavior.

The presence of grandchildren may influence the utility of joint leisure as well. In table 7.6, I note that spouses respond less to the focal partner's employment decision provided that they are grandparents. This conclusion suggests that couples value joint leisure more if they are not grandparents in accordance with Jørgensen (2014). The greater utility of joint leisure of

[27] The focal partners' and spouses' incomes are measured two years before the first month in which the focal partners are eligible to claim ERB if being in the control group. This implies that I measure the incomes when the focal partners are 58 years at the first cutoff, 58.5 years at the second cutoff, 59 years at the third cutoff, and 59.5 years at the fourth cutoff. Hence, I avoid that incomes are affected by the focal partners' and spouses' employment decisions in the months of interest.

[28] I define people who have taken further education (short-cycle higher education, medium-cycle higher education, and long-cycle higher education) as highly skilled. In contrast, I define people who have not taken further education (unskilled, skilled, and secondary education) as low-skilled.

Table 7.6: The spillover effect by subgroups

	(1)	(2)	(3)	(4)	(5)	(6)
First month	0.356***	0.122	0.192***	0.064	0.131	0.370**
	(0.158)	(0.101)	(0.074)	(0.175)	(0.081)	(0.160)
	10,479	10,374	20,300	6,960	16,763	8,076
Second month	0.342***	0.160**	0.196***	0.093	0.147**	0.327***
	(0.141)	(0.066)	(0.055)	(0.108)	(0.058)	(0.117)
	10,431	10,354	20,271	6,957	16,908	7,911
Third month	0.331***	0.127*	0.177***	0.067	0.151***	0.244**
	(0.124)	(0.067)	(0.050)	(0.109)	(0.052)	(0.102)
	10,407	10,334	20,249	6,949	17,059	7,748
Fourth month	0.356***	0.137**	0.197***	0.057	0.154***	0.299***
	(0.130)	(0.065)	(0.053)	(0.102)	(0.053)	(0.109)
	10,416	10,324	20,249	6,951	17,103	7,704
Fifth month	0.338***	0.145**	0.182***	0.104	0.143***	0.307***
	(0.128)	(0.066)	(0.055)	(0.100)	(0.055)	(0.106)
	10,416	10,324	20,249	6,951	17,103	7,704
Sixth month	0.295*	0.129	0.147**	0.155	0.156**	0.276**
	(0.164)	(0.082)	(0.070)	(0.130)	(0.074)	(0.136)
	10,416	10,324	20,249	6,951	17,103	7,704
Group	Spouse earns less	Spouse earns more	Homo-geneous	Hetero-geneous	Grand-parent	Not grand-parent

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table 7.6 reports the estimate of interest ($\hat{\rho}$) from model 5.3 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. The focal partners' and spouses' incomes are measured two years before the first month where the focal partners are eligible to claim ERB if they are in the control group. Couples are only included in the estimations in column (1) and column (2) if both the focal partners' and the spouses' incomes are positive at the time of the observation. The focal partners and spouses in homogeneous couples have the same levels of education. See appendix A.1 to obtain the estimates of all parameters.

couples without grandchildren may result from a smaller social set, causing them to value joint leisure more. Another explanation may be that grandparents prefer to retire as soon as possible to spend time with their grandchildren. It eliminates the possibility to retire jointly if both partners retire upon reaching the ERA, as couples only retire jointly if both partners are eligible to claim pension benefits as concluded in the previous sections. This explanation corresponds with the findings by Rupert & Zanella (2018). They stress that grandchildren have a negative impact on the grandparents' labor supply, as grandparents take care of their grandchildren, limiting the possibilities to work.

7.5 Robustness checks

It is not possible to confirm that the estimates are correctly identified, but the robustness of the estimates can be evaluated to investigate how reasonable it is to assume that they are. Testing if the identifying assumptions are violated is therefore necessary.

7.5.1 The identifying assumptions

Multiple identifying assumptions were stated in section 5.1 and 5.2. If one is to give the reduced-form estimates a causal interpretation, RDD.1, RDD.2, and RDD.3 cannot be violated. If the estimated spillover effect is to be given a causal interpretation, assumption RDD.4, RDD.5, and RDD.6 also cannot be violated.

Assumption RDD.1: Continuity of the conditional expectation function
The assumption that the conditional expectation function is continuous is impossible to verify, as one does not know the true conditional expectation function. If $f(x_i)$ serves as a proper approximation of the conditional expectation function, graphical evidence and pseudo cutoffs can be used to investigate whether the assumption is satisfied. Figure 7.3 verifies that the only discontinuity observed is at the cutoff.

Furthermore, I estimate model 5.2 two times where -50 and 50 serve as pseudo cutoffs.[29] The ERA is identical just before and just after the pseudo cutoffs, implying that no discontinuity is to be observed at the new cutoffs. I obtain insignificant estimates at both cutoffs as reported in table A.4. As a result of the visual evidence and the estimations using the pseudo cutoffs, I have no reason to believe that the conditional expectation function is discontinuous.

Assumption RDD.2: No manipulation of the running variable
If spouses prefer to retire jointly with the focal partners at a younger age, it is ideal to marry focal partners born just before the cutoff dates. I limit the sample to couples that were married prior to the announcement of the Retirement Reform of 2011 to avoid that spouses intentionally marry focal partners that were born just before the cutoff. If spouses, however, were more likely to divorce focal partners born just after the cutoff as a result of the increase in the ERA, the

[29] In the former case, couples are assigned to the treatment group if the focal partners were born 50 days or less earlier than the baseline cutoff. This implies that, e.g., couples in which the focal partner is born on September 11, 1953, are assigned to the control group, while couples in which the focal partner is born on September 12, 1953, are assigned to the treatment group. In the latter case, couples are assigned to the control group if the focal partners were born 50 days or less later than the baseline cutoff. E.g., couples in which the focal partner is born on February 20, 1954, are assigned to the control group, while couples in which the focal partner is born on February 21, 1954, are assigned to the treatment group.

estimates may be invalid. The assumption that couples have imprecise control over the running variable is usually tested by presenting histograms and density tests where the presence of bunching just before or after the cutoff suggests that the assumption is violated. In figure A.5, I present histograms with the distribution of the focal partners' date of birth relative to the cutoff. The figure shows no signs of bunching just before or after the cutoff, ensuring that it is impossible to reject the assumption graphically.

The McCrary density test is a widely used tool to investigate whether the density of the running variable is substantially different just before or after the cutoff (McCrary, 2008). Frandsen (2017) argues that the McCrary density test performs poorly provided that the running variable is discrete, as the test will often falsely reject the assumption. Instead, he proposes another test that he argues is more suitable. I follow a similar approach developed by Cattaneo et al. (2018) who have implemented the test as a Stata routine.[30] Obtaining p-values in the range of 0.57 to 0.70 in the six months of interest as reported in column (1) in table A.5, I find no evidence that couples can manipulate their value of the running variable.

Assumption RDD.3: Continuity of the predetermined covariates
A method used frequently to test the assumption of continuity of the covariates is to replace the outcome variable by each covariate and run a regression where each covariate serves as the outcome variable. This implies that if K covariates are included in the estimation, the following K models must be estimated

$$\theta_1 = \xi_1 + f(x_i) + \lambda_1 D_{ij} + \psi_1$$
$$...$$
$$\theta_K = \xi_K + f(x_i) + \lambda_K D_{ij} + \psi_K$$

where θ_k is the covariate, ξ_k is the intercept, and λ_k is a parameter to be estimated for $k = 1, ..., K$. ψ_k is the error term, and $f(x_i)$ and D_{ij} have the same interpretation as in model 5.2. If all estimates of λ_k are individually insignificant, the assumption cannot be rejected. The test may work well with few covariates, but running many regressions resulting from a large set of covariates is likely to provide significant estimates by random chance, leading to a wrong conclusion (Lee & Lemieux, 2010).[31]

Instead, Lee & Lemieux (2010) propose to estimate a Seemingly Unrelated Regression (SUR) where the estimates of λ_k only have to be jointly insignificant. Estimating a SUR, the error

[30] The routine is given by *rddensity* in Stata.

[31] Recall the definition of the 95 percent confidence interval: 95 percent of the times, the interval contains the true value. This implies that 5 percent of the times, the interval does not contain the true value, and the assumption will thus probably wrongfully be rejected by random chance since a large number of covariates is included in the estimations.

terms are allowed to be correlated across the K estimations. Performing a χ^2-test for the hypothesis that $\lambda_1, ..., \lambda_K$ are jointly equal to zero, I obtain estimates way beyond the critical level of 0.05 in all months of interest as shown in column (2) in table A.5. Therefore, I cannot reject that the covariates evolve smoothly at the cutoff.

Assumption RDD.4: The first stage exists
The first stage exists by definition if $g_1(x_i) > g_0(x_i)$ as argued in section 5.2. This implies that raising the ERA has to increase the focal partners' employment. Although I confirmed this in section 7.2, the first stage can be too weak such that the FRDD produces invalid estimates. The assumption that the first stage exists is investigated using an F-test where, as a rule of thumb, the F-statistic of the first stage regression has to be greater than 10 (Stock & Yogo, 2002). I obtain the F-statistic of the baseline first stage regressions (i.e. model 5.1 from section 5.3.2) that include all couples in the estimation sample. The F-statistic is way beyond 10 in each of the six months of interest, as reported in table A.6 in appendix A.2, suggesting a powerful first stage. This was also expected based on the focal partners' employment response depicted in figure 7.2 and figure 7.4.

Assumption RDD.5: The exclusion restriction
The spouses' employment must only be influenced by whether the focal partners were born before or after the cutoff through the focal partners' employment. This assumption cannot be tested but relies on argumentation. Although some spouses can claim ERB at an older age due to the increase in the ERA, there is no systematic difference between the spouses' month of birth at the cutoff, as assumption RDD.3 was not rejected.[32] Hence, the spouses' ERA is not significantly more likely to be raised if they are married to focal partners born just after the cutoff. This is consistent with the conclusion that the change in the spouses' employment is insignificant in the validation period and significant in the identification period. Therefore, I have no reason to suppose that the exclusion restriction is violated.

Assumption RDD.6: The monotonicity assumption
There is no means of testing this assumption, wherefore it must instead be discussed. As the eligibility rules of the ERP scheme require individuals to be in the labor force upon reaching the ERA, it will never be ideal to retire prior to the ERA, but work when being eligible to claim ERB, as individuals will waive five years of ERB if retiring prior to the ERA. Based on the eligibility rules, it is reasonable to assume the absence of defiers, meaning that nothing supports

[32] Following section 6.1, I include month-of-birth dummies of the spouses. Assumption RDD.3 was not rejected, suggesting that the covariates evolve smoothly at the cutoff. Therefore, there is no systematic difference between the spouses' month of birth at the cutoff, meaning that focal partners born just after the cutoff are not more likely to be married to spouses affected by the increase in the ERA.

that the monotonicity assumption is violated.

7.5.2 Additional robustness checks

As argued in section 5, it is necessary to decide whether to utilize parametric or nonparametric estimation and choose whether to apply a linear or nonlinear model. Furthermore, it is essential to investigate if the effects vary substantially between the four separate cutoffs. Finally, it is crucial to examine how methodological choices affect the estimates.

The model specification

Following the arguments by Lee & Lemieux (2010), utilizing parametric estimation is not necessarily favorable relative to nonparametric estimation and vice versa. The estimates need instead to be stable across both approaches to ensure that the estimates are robust. In table A.7 in appendix A.2, I investigate how utilizing the nonparametric approach affects the estimates in order to test the robustness. First, I utilize nonparametric estimation using a bandwidth of 100 days in accordance with my baseline approach. Second, I follow the approach by Calonico et al. (2020) to select the optimal bandwidth. In both cases, I utilize the triangular kernel instead of the uniform kernel. Table A.7 shows that the estimates are not substantially different when utilizing nonparametric estimation instead of parametric estimation, confirming that the estimates are robust. The estimates are within the same range but are insignificant when using the approach by Calonico et al. (2020) to select the optimal bandwidth. This results from the more narrow bandwidth that reduces the number of observations, causing the standard errors to be larger.

The choice of model

The linear probability model can generate predicted probabilities that are smaller than zero or larger than one, which is impossible by definition, as argued in section 5.3.4. As a check of the robustness of the estimates, I apply the probit model where probabilities are restricted to be between zero and one. Marginal effects are not reported by default since the marginal effects depend on the covariates of the couple. Estimating the average marginal effect, I confirm in table A.8 that using the probit instead of the linear probability model does not change the estimates substantially. The conclusion that the average marginal effect is similar in the linear probability model and the probit model corresponds with Angrist & Pischke (2009).

The cutoffs estimated separately

Pooling cutoffs to use more observations is widely applied to reduce the standard errors, but substantially different estimates at each of the four cutoffs questions the robustness of the pooled estimates. In table A.9, I report the estimates at each of the four cutoffs estimated separately. The estimates are largest at the first (I) and fourth (IV) cutoff, but the estimates are not substantially different relative to the other cutoffs. The standard errors are larger when estimating each cutoff separately, emphasizing the importance of pooling the cutoffs to increase the statistical power. As the estimates are within the same range at all four cutoffs, I find no evidence that the estimated effects are attributable to a single cutoff, confirming the robustness and external validity of the estimates.

Definition of employment

I define an individual to be employed if (s)he works more than eight hours a week on average during the month. The threshold of eight hours is chosen, as one is only covered by the Danish Salaried Employees Act if working more than this threshold as argued in section 6.1. However, the threshold could also be zero hours, five hours in accordance with An et al. (2004), or more hours. I confirm in table A.10 that the estimates do not change substantially when considering other thresholds, emphasizing that the estimates are robust to the choice of when an individual is defined as employed.

8 Macroeconomic implications

The results derived in section 7 confirm that raising the ERA increases the spouses' employment. Similarly, raising the ERA causes spouses to be less likely to claim benefits. The Danish Ministry of Finance does not account for leisure complementarities and thus ignores the spillover effect from the focal partners to the spouses.[33] This implies that the positive effect of raising the ERA on employment is underestimated by 14 to 15 percent, while the decrease in the number of people claiming ERB is underestimated by 7 to 8 percent for the in-sample population. Hence, the tax revenue increases by more than expected since more people are employed. The public expenditures furthermore decrease by more than expected since fewer people claim benefits. This section investigates how much the predictions made prior to the reforms underestimate the positive effect on the government budget by ignoring leisure complementarities.

8.1 The individual effect

To investigate to what extent the calculations made prior to the reforms underestimate the positive effect on the government budget by ignoring leisure complementarities, I multiply the number of spouses with the estimated average marginal change in the spouses' individual net

[33] Based on The Danish Ministry of Finance (2011), Jørgensen (2014) argues that leisure complementarities are not included in the calculations of the effects of raising the ERA.

contribution to the government budget. The individual net contribution to the government budget, or simply *the net contribution*, measures how much each individual contributes financially to the government budget. I define this as the tax payment on labor earnings minus the net-of-tax take-up of benefits. This definition of the net contribution deviates from The Danish Ministry of Finance (2017), as I do not include expenditures to, e.g., healthcare, maintenance of roads, fire service, police force, etc. Imposing the assumption that the reforms solely influence couples through employment, applying the simple definition of the net contribution does not have an impact on the estimates, as I consider only the difference between the spouses' net contribution.

I estimate model 5.2 from section 5.3.2, but use each spouse's monthly net contribution as the outcome variable. The estimates are depicted in table 8.1.

Table 8.1: The change in the spouses' net contribution

	(1)	(2)	(3)
First month	690.6*	522.0	-168.6
	(393.0)	(320.9)	(122.6)
	27,260	27,260	27,260
Second month	937.9**	560.9*	-377.0***
	(374.9)	(293.9)	(120.4)
	27,228	27,228	27,228
Third month	1088.5***	741.3**	-347.3***
	(361.3)	(297.0)	(109.1)
	27,198	27,198	27,198
Fourth month	819.8**	488.4	-331.4***
	(397.0)	(330.8)	(104.6)
	27,200	27,200	27,200
Fifth month	678.7*	344.6	-334.1***
	(403.0)	(348.0)	(100.1)
	27,200	27,200	27,200
Sixth month	723.8*	402.6	-321.2***
	(408.6)	(340.6)	(113.0)
	27,200	27,200	27,200
Outcome variable	Net contribution	Tax payment	Benefits

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table 8.1 reports the estimate of interest ($\hat{\tau}$) from model 5.2 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. See appendix A.1 to obtain the estimates of all parameters.

The monthly net contribution is not recorded in the administrative data, implying that I have to approximate this. Furthermore, I consider the tax payment on labor earnings and the net-of-tax take-up of benefits as the outcome variables to investigate whether the increase in the tax payments or the decrease in the take-up of benefits has the largest impact on the government budget. In appendix A.4, I provide a detailed explanation of how the tax payment on labor earnings and the net-of-tax take-up of benefits are computed.

Table 8.1 shows that raising the ERA increases the spouse's net contribution by 678.7 to 1088.5 DKK. The estimates are highly significant in the second to the fourth month, while the estimates are significant at a 10 percent level in the first, the fifth, and the sixth month. The increase in the spouse's net contribution results from both a larger tax payment and a smaller take-up of benefits as expected. However, the estimated increase in the tax payment is only significant in two of the months because of large standard errors due to large variation in the spouses' tax payments. If ignoring the statistical significance, the estimates suggest that the increase in the tax payments has a greater impact on the government budget than the decrease in the take-up of benefits.

8.2 The aggregate effect

The estimates from table 8.1 show that raising the ERA increases the spouses' net contribution. Adding the estimates in all six months of interest, the spouse's net contribution increases by 4,939 DKK, whereas 3,059 DKK comes from a larger tax payment, and 1,880 DKK comes from a smaller take-up of benefits.

For individuals born within the last six months of 1954 or later, the ERA increases by more than six months relative to the counterfactual scenario in which the reforms were not implemented. However, the present research design only allows me to estimate the effects of raising the ERA by six months. As a result, I cannot determine to what extent the calculations made prior to the reforms underestimate the aggregate positive impact on the government budget by ignoring leisure complementarities for couples in which the focal partners were born within the last six months of 1954 or later. For individuals born within the first six months of 1954, the ERA increases only by six months relative to the counterfactual scenario in which the reforms were not implemented. As shown in table 8.2, ignoring leisure complementarities for couples in which the focal partners were born within the first six months of 1954 underestimates the positive effect on the government budget by 18.9 million DKK for the in-sample population, which includes 3,821 couples.

Solely couples that are likely to respond to the increase in the ERA are included in the sample. However, it may be misleading to conclude that raising the ERA increases only the net contribution of the spouses in the sample. Suppose I instead assume that spouses in all married couples in which the focal partners were born within the first six months of 1954 and both partners are members of the ERP scheme, respond similarly to raising the ERA, as the spouses in the sample. In this case, the calculations made prior to the reforms underestimate the positive effect on the government budget by 53.0 million DKK by ignoring leisure complementarities. Extrapolating the estimates to spouses in couples who are married or cohabiting in which the focal partners were born within the first six months of 1954 and both partners are members of the ERP scheme, the calculations made prior to the reforms underestimate the positive effect on the government budget by 68.6 million DKK. Finally, suppose I assume that spouses in all married or cohabiting couples in which the focal partners were born within the first six months of 1954, respond similarly to the in-sample population. In this case, the calculations made prior to the reforms underestimate the positive effect on the government budget by 99.6 million DKK. This assumption is reasonable if spouses respond similarly to raising the NRA as to raising the ERA, which cannot be rejected based on figure A.4d in appendix A.3.

Table 8.2: The aggregate effect on the government budget

	In-sample	Married	Also cohabiting	Also non-ERP
Tax revenue				
Couples	3,821	10,725	13,883	20,162
Sum (DKK)	3,059	3,059	3,059	3,059
Total (DKK)	11,688,439	32,807,775	42,468,097	61,675,558
Public expenditures				
Couples	3,821	10,725	13,883	20,162
Sum (DKK)	-1,880	-1,880	-1,880	-1,880
Total (DKK)	-7,183,480	-20,163,000	-26,100,040	-37,904,560
Total net contribution				
Couples	3,821	10,725	13,883	20,162
Sum (DKK)	4,939	4,939	4,939	4,939
Total (DKK)	**18,871,919**	**52,970,775**	**68,568,137**	**99,580,118**

Note: All amounts are measured in 2019-prices. The sum in DKK is the sum of the estimated change in the net contribution in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB.

A conservative estimate suggests that ignoring leisure complementarities in the calculations made prior to the reforms underestimate the positive effect on the government budget by 18.9 million DKK. An upper limit estimate indicates that this amount is 99.6 million DKK. The actual effect lies within this range and depends on to which degree it is possible to extrapolate the estimates from the in-sample population to the general population, i.e., the degree of external validity. The aggregate effect on the government budget is only estimated for couples in which the focal partners were born within the first six months of 1954, implying that the aggregate effect across cohorts is much larger. The calculations clearly show that accounting for leisure complementarities when raising the retirement age improves the government budget by more than predicted and that ignoring leisure complementarities impacts the difference between the estimated effect and the actual effect substantially. This difference will develop to be even larger in the future, as the ERA and NRA will continue increasing.

9 Discussion

In this section, I discuss the results obtained in the previous sections. I discuss the validity of the estimates and argue why this thesis contributes to the literature of joint retirement. Finally, I discuss the policy implications of the results and suggest improvements for future research.

9.1 Validity

Assessing the validity of the estimates is necessary if the estimates are used for policy recommendations. The validity is divided into two categories: internal validity and external validity. The former relates to the robustness of the estimates and whether the estimates can be given a causal interpretation. The latter describes to which extent the obtained estimates can be extrapolated to other populations than the one considered in this thesis.

The reduced-form estimates and the spillover effect are robust to the inclusion of controls, the choice of functional form, and the choice of bandwidth, and none of the identifying assumptions are rejected. Furthermore, the estimates do not change substantially by utilizing nonparametric estimation, the probit model, estimating each cutoff separately, or changing the definition of when an individual is defined as employed. These findings confirm that the internal validity is high, assuring that this research design is suited for analyzing whether leisure complementarities cause couples to retire jointly.

Utilizing the SRDD, the effects are only obtained for the subpopulation in the vicinity of the cutoff (G. W. Imbens & Lemieux, 2008). When using the FRDD, the effects are obtained for an even smaller subpopulation, as the effects are only obtained for the compliant subpopulation close to the cutoff. This may question the external validity, as it is uncertain to what extent the estimates can be extrapolated to the general population.

The estimation sample is restricted to avoid including couples that are not likely to respond to the increase in the ERA. Excluding couples where not both partners are members of the ERP scheme almost halves the sample. Excluding couples in which either the focal partner or the spouse is self-employed also impacts the number of observations in the sample substantially. The other restrictions do only have a limited impact on the number of observations in the sample, implying that the estimation sample covers a large share of married couples in the population in which the focal partners were born in the years of interest. Hence, the results are valid for couples in which both partners are wage earners and members of the ERP scheme if assuming that the estimates can be extrapolated to couples that are not close to the cutoff. Other studies have often considered reforms or schemes that typically only cover a small part of the working couples, which is thus associated with a lower degree of external validity, as the results cannot be extrapolated to the general population.[34]

The spillover effect is probably larger in countries where the focal partners' pension benefits are means-tested based on the spouses' income, or where the spouses receive pension benefits when the focal partners retire, as spouses have economic incentives to retire jointly with the focal partners.[35] The spillover effect may also be lower in Denmark relative to other countries, as the eligibility rules of the ERP scheme require people to be in the labor force upon reaching the ERA, implying that younger spouses do not respond to the focal partner's employment decision. This suggests that the results are country-specific, indicating that one must be cautious extrapolating the results to other countries.

Bhatt (2017) argues that younger cohorts in the United States value joint leisure less than older cohorts because of socioeconomic, employment, and health-related factors. If this also applies to Danish couples, the spillover effect will be lower for younger cohorts, emphasizing that the spillover effect is ideally to be estimated frequently to ensure that leisure complementarities are accounted for rightfully.

9.2 Comments on the empirical strategy

Other research designs could have been utilized to investigate whether leisure complementarities cause couples to retire jointly. This includes the simple OLS approach (where $f(x_i)$ is restricted to be zero) or the difference-in-difference design. In the former approach, couples in the treatment group and the control group must be identical such that only the ERA is different. In the latter approach, the identifying assumption only states that the employment of the spouses in the treatment group and the control group would evolve identically if the ERA was unchanged. A potential threat to both research designs is that the focal partners and the spouses in the treatment group were born after the focal partners and the spouses in the control group.[36] Bhatt (2017) argues that cohorts differ in terms of preferences and other key characteristics that are

[34] See, e.g., Kruse (2020), Baker (2002), or Bloemen et al. (2015) that consider schemes which only cover a small part of the population.

[35] The impact of economic incentives to retire jointly is documented by, e.g., Hiedemann et al. (1998), Johnson & Favreault (2001), or Michaud (2003).

[36] As the ERA depends on the date of birth, the focal partners in the treatment group are exposed to the increase in the ERA since they were born after the focal partners in the control group. Hence, focal partners in the treatment group are by definition born after focal partners in the control group. Following table 6.2, the average age difference between the focal partners and the spouses is similar in the treatment group and the control group. This implies that spouses in the treatment group are, on average, born after spouses in the control group.

unobservable to the econometrician, meaning that the identifying assumption in both designs is likely to be violated.

Furthermore, The Danish Ministry of Finance (2020) shows that better health and longer life expectancy causes younger cohorts to retire later than older cohorts independently of the increase in the ERA. This implies that the employment of younger cohorts would be larger than the employment of older cohorts even in the absence of the reforms, supporting that couples are not identical across cohorts. Hence, it is a critical assumption that couples are identical or that the spouses' employment would evolve identically in the absence of the reforms. This implies that the estimates will be upward biased if utilizing the simple OLS approach or the difference-in-difference design if not accounting for the differences in health and life expectancy across cohorts.

The main advantage of utilizing the RDD is that only couples in the vicinity of the cutoff have to be identical. This implies that couples in which the focal partners were born just after the cutoff only have to be similar to couples in which the focal partners were born just before the cutoff. Hence, the identifying assumption in the RDD is less strict and more likely to be satisfied.

9.3 Contribution to the literature

This thesis contributes to the literature by providing causal evidence based on within-country variation that leisure complementarities cause couples to retire jointly. I restrict the sample to couples married prior to the announcement of the Retirement Reform of 2011, while the focal partners' date of birth serves as the running variable. Hence, the assignment to treatment is completely exogenous. This essential issue has turned out to be challenging to overcome in other studies. Furthermore, every Dane is entitled to healthcare treatment, the focal partners' pension benefits are not means-tested based on the spouses' income, and spouses do not receive pension benefits when the focal partners retire. This is in opposition to many other studies where it is questionable if leisure complementarities can explain entirely why couples retire jointly, or if couples retire jointly mainly due to economic incentives.[37] By using high-quality administrative data, I avoid the self-reported bias that is likely to cause a threat to identification in other studies in the field of joint retirement.[38]

All this ensures that I obtain causal evidence that leisure complementarities cause couples

[37] See, e.g., Blau & Gilleskie (2006), Kapur & Rogowski (2007), Gustman & Steinmeier (2004), or van der Klaauw & Wolpin (2008).

[38] Gustman & Steinmeier (2002) stress that the conclusion if leisure complementarities cause couples to retire jointly depends on the specific survey. It confirms that the self-reported bias in surveys has a substantial impact on the estimates.

to retire jointly, which has been called for in recent years. Furthermore, the large sample, the identical ERA of men and women, and the small gender gap in employment ensure that I can investigate whether the genders respond differently to the focal partner's employment decision conditional on the couple's age structure.

Raising the ERA increases the spouses' employment by 2.6 to 3.7 percentage points. This corresponds to a spillover effect from the focal partners' employment to the spouses' employment of 16.5 to 18.3 percentage points. This spillover effect is smaller than found by Jørgensen (2014) using a dynamic programming model. However, the spillover effect estimated in this thesis is larger than estimated by García-Miralles et al. (2021) who exploit the same reforms but utilize the difference-in-difference design. Different sample restrictions likely cause the difference in the estimated spillover effect; García-Miralles et al. (2021) are more likely to include individuals who are not members of the ERP scheme, as argued in section 6.2. They also include couples with larger age differences, which, according to their paper, reduces the estimated spillover effect. Finally, they only consider individuals born at the end of 1953 and the beginning of 1954, implying that they use a subset of my sample to estimate the spillover effect. The spillover effect estimated in this thesis is slightly smaller than estimated by Bloemen et al. (2015), Kruse (2020), and Coile (2004) that all consider within-country variation in the Netherlands, Norway, and the United States, respectively.

My estimates suggest that whether the spouses responds to the focal partner's employment decision depends on the couple's age structure. Couples retire jointly by waiting until both partners are eligible to claim pension benefits, as it is very costly to retire prior to the retirement age. This conclusion corresponds with Legendre et al. (2018). This implies that men who are, on average, the oldest in the couple, are more likely to respond to the focal partner's employment decision in reduced-form in accordance with Kallestrup-Lamb (2011). The spillover effect, however, is identical for men and women in accordance with An et al. (2004). Hence, I find no evidence to suggest that men and women value joint leisure differently, as the couple's age structure explains the different responses by the genders in reduced-form. Furthermore, economic incentives probably explain why only the employment of older spouses responds to the focal partners' employment. Therefore, a pitfall arises if using the estimates to conclude that older spouses value joint leisure more than younger spouses, or that men value joint leisure more than women. This emphasizes that researchers cannot use their estimates to conclude who values joint leisure the most without accounting for economic incentives, which is a concern in, e.g., Coile (2004).

9.4 Policy implications

The conclusion that leisure complementarities cause couples to retire jointly emphasizes that policymakers have to account for second-order effects when designing a policy. Ignoring leisure complementarities when raising the retirement age underestimates the increase in employment by 14 to 15 percent and the decrease in the number of individuals who claim pension benefits by 7 to 8 percent for the in-sample population. This implies that the positive effect on the government budget is underestimated by at least 18.9 million DKK for the first group of couples affected by the reforms. If the results can be extrapolated to all the first couples affected by the increase in the ERA, this amount may be 99.6 million DKK.

As the positive effect of raising the ERA on the government budget is underestimated by ignoring leisure complementarities, resources may not be allocated efficiently. Accounting for leisure complementarities will improve the government budget, which may imply that resources are to be allocated differently to increase citizens' utility, e.g., by increasing public expenditures or cut taxes. The ERA and NRA will continue increasing, meaning that accounting for leisure complementarities becomes even more necessary in the future.

Accounting for leisure complementarities and thus allow for spillover effects implies that raising the retirement age is a more powerful tool to improve the government budget than previously thought. If policymakers aim to improve the government budget in the future, it may be advantageous to raise the retirement age relative to other interventions.

9.5 Suggestions for future research

Multiple studies have investigated whether leisure complementarities cause couples to retire jointly. However, much work has yet to be done in order to be able to distinguish between economic incentives and leisure complementarities to explain why couples retire jointly. Hence, obtaining causal estimates is crucial in future research as stressed by Atalay et al. (2019).

When more people reach the NRA, it provides the option to investigate whether the spillover effect when raising the NRA is similar to the spillover effect when raising the ERA. People can retire prior to the NRA without waiving benefits, meaning that it is less costly for younger spouses to retire when the focal partners become eligible to claim national retirement benefits. Furthermore, the focal partners' national retirement benefits are means-tested based on the spouses' income, creating an economic incentive for spouses to retire jointly with the focal partners.[39] Thus, researchers can exploit the increase in the NRA to examine the importance of economic incentives relative to leisure complementarities concerning the decision to retire

[39] The national retirement benefit consists of a basic amount and a supplementary amount. The latter is means-tested based on a potential spouse's income.

jointly.

Raising the ERA was expected to increase the employment of those that are eligible to claim ERB at an older age, but potential second-order effects have not gained much traction. Rodrik (2015) argues that second-order effects are important and necessary to take into account since a policy change often affects additional areas than the one intended. The research design presented in this thesis can be utilized to investigate how raising the retirement age influences other outcomes. This includes couples' savings and consumption behavior, as a higher retirement age decreases the incentive to save for retirement if people prefer to smooth lifetime consumption. Raising the retirement age may also influence health, but the effects remain ambiguous: Eibich (2015) and Hallberg et al. (2014) argue that raising the retirement age weakens health while Coe & Zamarro (2011) reach the opposite conclusion. Based on the findings by Rupert & Zanella (2018), raising the retirement age may decrease parents' labor supply as grandparents will be less able to take care of their grandchildren. These examples show that raising the retirement age potentially affects other areas than the intended, leaving much work to be done for researchers to get a broader picture of the benefits and consequences of raising the retirement age.

10 Conclusion

In this thesis, I investigate whether leisure complementarities cause couples to retire jointly and thus whether raising the focal partners' retirement age increases the spouses' employment. This is highly relevant, as multiple countries have raised the retirement age or will do so soon. While the affected individual is likely to postpone retirement and remain employed, policymakers do usually not account for spillovers to a potential spouse's employment.

I exploit two Danish reforms that raised the early retirement age four times, each time by six months. Utilizing the RDD, I find that raising the retirement age increases the spouses' employment by 2.6 to 3.7 percentage points in reduced-form. This corresponds to a spillover effect of 16.5 to 18.3 percentage points. The estimates are significant and robust to different specifications, confirming the validity of the results. Only older spouses respond to the focal partner's employment decision due to economic incentives, which follows the predictions from my theoretical model.

My estimates suggest that leisure complementarities cause couples to retire jointly and that spouses respond to the focal partner's employment and retirement decisions. Hence, raising the retirement age underestimates the aggregate effect on the government budget by 18.9 to 99.6 million DKK for the first group of affected couples, as the Danish Ministry of Finance does not account for spillover effects. Therefore, the conclusions of this thesis suggest that policymakers have to account for leisure complementarities in the future when estimating the effects of interventions in the pension system.

This thesis contributes to the literature by providing causal evidence that leisure complementarities cause couples to retire jointly. The focal partners' pension benefits are not means-tested based on the spouses' income, and spouses do not receive pension benefits when the focal partners retire, ensuring that only leisure complementarities can explain why raising the retirement age increases the spouses' employment. As the estimates are obtained for a large share of the couples in the population in which the focal partners were born in the years of interest, the results can likely be extrapolated to the general population.

Future research has to focus on obtaining causal estimates, as the lack of causality has been a concern in multiple previous studies. Furthermore, researchers will do well by investigating the importance of economic incentives relative to leisure complementarities concerning the decision to retire jointly. Finally, researchers may examine whether raising the retirement age has spillovers to other areas, e.g., savings and consumption behavior, health, and labor supply of other family members to get a broader picture of the benefits and consequences of raising the retirement age.

References

The Agency of Labor Market and Recruitment. (2020). *Statistik om efterløn.* Retrieved 2020-11-24, from `https://star.dk/viden-og-tal/statistik/statistik-om-efterloen/`

An, M. Y., Christensen, B. J., & Gupta, N. D. (2004). Multivariate Mixed Proportional Hazard Modelling of the Joint Retirement of Married Couples. *Journal of applied econometrics (Chichester, England)*, *19*(6), 687-704.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics, an empiricist's companion.* Princeton: Princeton University Press.

Atalay, K., Barrett, G. F., & Siminski, P. (2019). Pension incentives and the joint retirement of couples: evidence from two natural experiments. *Journal of Population Economics*, *32*(3), 735-767.

Baker, M. (2002). The Retirement Behavior of Married Couples: Evidence from the Spouse's Allowance. *The Journal of human resources*, *37*(1), 1-34.

Banks, J., Blundell, R., & Rivas, M. C. (2010). The dynamics of retirement behavior in couples: Reduced-form evidence from England and the US. *University College London.*

Bhatt, V. (2017). Cohort Differences in Joint Retirement: Evidence from the Health and Retirement Study. *Journal of Labor Research*, *38*(4), 475-495.

Blau, D. M. (1998). Labor Force Dynamics of Older Married Couples. *Journal of labor economics*, *16*(3), 595-629.

Blau, D. M., & Gilleskie, D. B. (2006). Health Insurance and Retirement of Married Couples. *Journal of applied econometrics (Chichester, England)*, *21*(7), 935-953.

Bloemen, H., Hochguertel, S., & Zweerink, J. (2015). *Joint Retirement of Couples: Evidence from a Natural Experiment.* IZA Institute for the Study of Labor.

Browning, M., Bourguignon, F., Chiappori, P., & Lechene, V. (1994). Incomes and outcomes: a structural model of intrahousehold allocation. *The Journal of political economy*, *102*(6), 1067-1096.

Cahuc, P., Carcillo, S., & Zylberberg, A. (2014). *Labor economics.* MIT press.

Calonico, S., Cattaneo, M. D., & Farrell, M. H. (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, *23*(2), 192–210.

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, *17*(2), 372–404.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: methods and applications.* Cambridge university press.

Cattaneo, M. D., Jansson, M., & Ma, X. (2018). Manipulation testing based on density discontinuity. *The Stata Journal*, *18*(1), 234–261.

Cattaneo, M. D., Titiunik, R., Vazquez-Bare, G., & Keele, L. (2016). Interpreting regression discontinuity designs with multiple cutoffs. *The Journal of Politics*, *78*(4), 1229–1248.

Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T. H., & Olsen, T. (2014). Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark. *The Quarterly Journal of Economics*, *129*(3), 1141–1219.

Chiappori, P.-A. (1988). Rational household labor supply. *Econometrica: Journal of the Econometric Society*, 63–90.

Coe, N. B., & Zamarro, G. (2011). Retirement effects on health in Europe. *Journal of health economics*, *30*(1), 77-86.

Coile, C. (2004). Retirement Incentives and Couples' Retirement Decisions. *The B.E. Journal of Economic Analysis & Policy*, *4*(1), 17-.

Cribb, J., Emmerson, C., & Tetlow, G. (2014). *How does increasing the early retirement age for women affect the labour supply of women and their husbands?, Netspar Discussion Papers* (Tech. Rep.). DP 01/2014-003.

The Dane Age Association. (2020). *Efterløn.* Retrieved 2021-03-01, from https://www.aeldresagen.dk/viden-og-raadgivning/penge-og-pension/efterloen

The Danish Ministry of Finance. (2006). *Aftale om fremtidens velstand og velfærd og investeringer i fremtiden .*

The Danish Ministry of Finance. (2011). *Aftaler om Reformpakken 2020.*

The Danish Ministry of Finance. (2017). *Økonomisk Analyse: Indvandreres nettobidrag til de offentlige finanser i 2017.*

The Danish Ministry of Finance. (2020). *Regneprincip for arbejdsudbudsvirkning af ændret pensionsalder og levetid.*

The Danish Ministry of Taxation. (2016). *Gennemsnitsskatten for samtlige skattepligtige i 2016.* Retrieved 2021-02-24, from https://www.skm.dk/skattetal/statistik/arkiv/gennemsnitsskatten-for-samtlige-skattepligtige-i-2016/

Dobkin, C., & Ferreira, F. (2010). Do school entry laws affect educational attainment and labor market outcomes? *Economics of education review*, *29*(1), 40–54.

Eibich, P. (2015). Understanding the effect of retirement on health: Mechanisms and heterogeneity. *Journal of health economics*, *43*, 1–12.

Frandsen, B. R. (2017). Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete. In *Regression discontinuity designs.* Emerald Publishing Limited.

García-Miralles, E., Leganza, J. M., et al. (2021). *Joint retirement of couples: Evidence from discontinuities in Denmark* (Tech. Rep.). University of Copenhagen. Department of Economics. Center for Economic Behaviour and Inequality.

Gelman, A., & Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, *37*(3), 447–456.

Gelman, A., & Zelizer, A. (2015). Evidence on the deleterious impact of sustained use of polynomial regression on causal inference. *Research & Politics*, *2*(1), 2053168015569830.

Goux, D., & Maurin, E. (2003). Who marries whom in France? In *Who marries whom?* (pp. 57–78). Springer.

Gustman, A. L., & Steinmeier, T. L. (1989). An Analysis Of Proposals By The Social Security Administration To Eliminate The Retirement Earnings Test For Those Beyond The Normal Retirement Age, To Accelerate The Scheduled Increase In The Delayed Retirement Credit, And To Alter Procedures For Benefit Recomputation. *Hanover, NH: Final Report to the Department of Health and Human Services, August.*

Gustman, A. L., & Steinmeier, T. L. (2002, December). *Retirement and the Stock Market Bubble* (Working Paper No. 9404). National Bureau of Economic Research. Retrieved from `http://www.nber.org/papers/w9404` doi: 10.3386/w9404

Gustman, A. L., & Steinmeier, T. L. (2004). Social Security, Pensions and Retirement Behaviour within the Family. *Journal of applied econometrics (Chichester, England)*, *19*(6), 723-737.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, *69*(1), 201–209.

Hallberg, D., Johansson, P., & Josephson, M. (2014). *Early retirement and post retirement health* (Tech. Rep.). Working Paper.

Hamermesh, D. S. (2002). Timing, togetherness and time windfalls. *Journal of population economics*, *15*(4), 601-623.

Hastrup, B. (2020). *Major Trends*. Multivers.

Hiedemann, B., Suhomlinova, O., & O'Rand, A. M. (1998). Economic independence, economic status, and empty nest in midlife marital disruption. *Journal of Marriage and the Family*, 219–231.

Hindriks, J., & Myles, G. D. (2013). *Intermediate public economics*. MIT press.

Honoré, B. E., & de Paula, A. (2018). A new model for interdependent durations. *Quantitative economics*, *9*(3), 1299-1333.

Hospido, L., & Zamarro, G. (2014). Retirement patterns of couples in Europe. *IZA Journal of European Labor Studies*, *3*(1), 1-18.

Hoynes, H., Miller, D. L., & Schaller, J. (2012). Who suffers during recessions? *Journal of Economic perspectives*, *26*(3), 27–48.

Hurd, M. D. (1990). The joint retirement decision of husbands and wives. In *Issues in the economics of aging* (pp. 231–258). University of Chicago Press, 1990.

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, *79*(3), 933–959.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, *142*(2), 615–635.

Jacob, R., Zhu, P., Somers, M.-A., & Bloom, H. (2012). A practical guide to regression discontinuity. *MDRC*.

Johansen, P. U., & Trier, M. (2010). *Praktisk statistisk metode for økonomer*. Handelshøjskolens Forlag.

Johnsen, J. V., Vaage, K., et al. (2015). *Spouses' retirement and the take-up of disability pension* (Tech. Rep.).

Johnson, R. W., & Favreault, M. (2001). Retiring together or working alone: The impact of spousal employment and disability on retirement decisions.

Jørgensen, T. H. (2014). *Life-Cycle Consumption and Retirement within the Family* (Unpublished doctoral dissertation). University of Copenhagen.

Kallestrup, M., Larsen, C. A., & Kristensen, M. K. (2000). *Velfærdsstatens spor: en historisk komparativ policy-analyse af to efterlønsreformer*. University of Aalborg.

Kallestrup-Lamb, M. (2011). *The Role of the Spouse in Early Retirement Decisions for Older Workers*. Institut for Økonomi, Aarhus Universitet.

Kaplan, G., & Schulhofer-Wohl, S. (2018). The changing (dis-) utility of work. *Journal of Economic Perspectives*, *32*(3), 239–58.

Kapur, K., & Rogowski, J. (2007). The Role of Health Insurance in Joint Retirement among Married Couples. *Industrial & labor relations review*, *60*(3), 397-407.

Kolesár, M., & Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, *108*(8), 2277–2304.

Kruse, H. (2020). Joint Retirement in Couples: Evidence of Complementarity in Leisure. *The Scandinavian journal of economics*.

Lalive, R., & Parrotta, P. (2017). How does pension eligibility affect labor supply in couples? *Labour economics*, *46*, 177-188.

Lalive, R., & Staubli, S. (2014). How does raising women's full retirement age affect labor supply, income, and mortality? Evidence from Switzerland. In *Joint meeting of the retirement research consortium conference volume.*

Lee, D. S., & Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, *142*(2), 655–674.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, *48*(2), 281–355.

Legendre, B., Pedrant, A.-C., & Sabatier, M. (2018). Should I stay or should I go? An econometric analysis of retirement decisions by couples. *Applied Economics*, *50*(53), 5814-5829.

Lumsdaine, R. L., & Mitchell, O. S. (1999). Chapter 49: New developments in the economic analysis of retirement. *Handbook of Labor Economics*, *3*(3), 3261-3307.

Maestas, N. (2010). Back to Work: Expectations and Realizations of Work after Retirement. *The Journal of human resources*, *45*(3), 718-748.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, *142*(2), 698–714.

McCrary, J., & Royer, H. (2011). The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. *American economic review*, *101*(1), 158–95.

Meng, C. (2018). *The Long Term Effect of Age at School Entry on Education, Labor and Health Outcomes-Evidence from China* (Unpublished doctoral dissertation). University of Illinois–Chicago.

Michaud, P.-C. (2003). Joint Labour Supply Dynamics of Older Couples. , *2003-69*.

Michaud, P.-C., Van Soest, A., & Bissonnette, L. (2019). Understanding joint retirement. *Journal of economic behavior & organization*, *173*, 386-401.

Mitchell, O. S. (1988). Worker Knowledge of Pension Provisions. *Journal of labor economics*, *6*(1), 21-39.

OECD. (2011). *Pensions at a Glance 2011.* Retrieved from `https://www.oecd-ilibrary.org/content/publication/pension_glance-2011-en` doi: https://doi.org/https://doi.org/10.1787/pension_glance-2011-en

OECD. (2021). *Employment: Employment/population ratio, by sex and age group.* Retrieved 2021-02-26, from `https://stats.oecd.org/index.aspx?queryid=54742#`

Petersen, K. (2011). *Bastarden: Historien om efterløn. Tilblivelsen 1975-1979.* Syddansk Universitetsforlag.

Rätzel, S. (2012). Labour supply, life satisfaction, and the (dis) utility of work. *The Scandinavian Journal of Economics*, *114*(4), 1160–1181.

Rodrik, D. (2015). *Economics rules: The rights and wrongs of the dismal science.* WW Norton & Company.

Rupert, P., & Zanella, G. (2018). Grandchildren and their grandparents' labor supply. *Journal of Public Economics*, *159*, 89–103.

Sand, E., et al. (2019). The Effects of a Retirement Age Reform on Couples' Labor Supply Decisions. *Shirlee, The Effects of a Retirement Age Reform on Couples' Labor Supply Decisions (May 30, 2019).*

Selin, H. (2014). *The rise in female employment and the role of tax incentives. An empirical analysis of the Swedish individual tax reform of 1971* (Vol. 21) (No. 5).

Stancanelli, E. (2017). Couples' retirement under individual pension design: A regression discontinuity study for France. *Labour economics*, *49*, 14-26.

Stancanelli, E., & Van Soest, A. (2016). Partners' leisure time truly together upon retirement. *IZA Journal of Labor Policy*, *5*(1), 1-19.

Statistics Denmark. (2002). *Sæsonkorrigering*.

Stock, J. H., & Yogo, M. (2002). *Testing for weak instruments in linear IV regression* (Tech. Rep.). National Bureau of Economic Research.

van der Klaauw, W., & Wolpin, K. I. (2008). Social security and the retirement and savings behavior of low-income households. *Journal of econometrics*, *145*(1-2), 21-42.

Wolcott, E. L. (2021). Employment inequality: Why do the low-skilled work less now? *Journal of Monetary Economics*, *118*, 161–177.

Zweimüller, J., Winter-Ebmer, R., & Falkinger, J. (1996). Retirement of spouses and social security reform. *European economic review*, *40*(2), 449-472.

A Appendix

A.1 Supplementary material

The reader can benefit from considering the supplementary material related to this master's thesis provided in the following link:
https://github.com/akriko/Master-thesis

The supplementary material contains the code used to generate the tables and figures of this thesis. This includes the code to simulate the theoretical model from section 4 used to construct figure 4.1, figure A.1 and figure A.2. One can change the values of the model's parameters or the distribution from which the utility of individual leisure is drawn to assess how it impacts the predictions of the model.

It must be possible to replicate the results derived in empirical papers to ensure that the professional standards of economic research are satisfied (Rodrik, 2015). Even a minor error in the code may change the conclusions of a paper. Hence, the supplementary material contains the code to generate the data sets used to construct the tables and figures of this thesis and the code to construct the tables and figures of this thesis.

It is only possible possible to report the estimate of interest in the tables, as the number of covariates is large. To overcome this problem, numerous .csv- or .txt-files are uploaded, containing the estimates of every table. Six files are uploaded to each table, as one file only contains the estimates in one of the six months of interest. This implies that, e.g., "Table_7.1_1" contains the estimates of table 7.1 in the first month where the focal partners in the control group are eligible to claim ERB, while the focal partners in the treatment group are ineligible although they are the same age in months.

A.2 Tables

Table A.1: The change in the spouses' employment (in the validation period and identification period, by different thresholds)

	(1)	(2)	(3)	(4)	(5)	(6)
Three months before	0.008	0.017*	-0.009	0.004	0.044	0.052
	(0.007)	(0.009)	(0.010)	(0.010)	(0.038)	(0.033)
	27,432	27,432	27,190	27,190	2,755	2,755
Two months before	0.007	0.010	-0.012	-0.000	0.053	0.032
	(0.007)	(0.009)	(0.010)	(0.011)	(0.039)	(0.035)
	27,356	27,356	27,323	27,323	2,408	2,408
One month before	0.021**	0.007	-0.012	-0.007	0.039	-0.004
	(0.009)	(0.009)	(0.010)	(0.011)	(0.054)	(0.045)
	27,319	27,319	27,460	27,460	2,093	2,093
First month	0.145***	0.026**	0.057***	0.007	0.063	0.008
	(0.010)	(0.010)	(0.011)	(0.010)	(0.060)	(0.050)
	27,260	27,260	27,645	27,645	1,807	1,807
Second month	0.203***	0.037***	0.079***	0.017	0.232***	0.034
	(0.009)	(0.010)	(0.011)	(0.011)	(0.062)	(0.057)
	27,228	27,228	27,814	27,814	1,476	1,476
Third month	0.206***	0.034***	0.079***	0.016	0.243***	0.023
	(0.008)	(0.010)	(0.010)	(0.010)	(0.062)	(0.056)
	27,198	27,198	27,938	27,938	1,183	1,183
Fourth month	0.209***	0.037***	0.078***	0.014	0.235***	0.032
	(0.008)	(0.010)	(0.010)	(0.010)	(0.063)	(0.051)
	27,200	27,200	27,975	27,975	1,073	1,073
Fifth month	0.203***	0.036***	0.083***	0.010	0.223***	0.095*
	(0.008)	(0.010)	(0.011)	(0.010)	(0.064)	(0.054)
	27,200	27,200	27,975	27,975	1,073	1,073
Sixth month	0.154***	0.026**	0.071***	-0.001	0.215***	0.059
	(0.012)	(0.010)	(0.011)	(0.010)	(0.064)	(0.055)
	27,200	27,200	27,975	27,975	1,073	1,073
Employment of	Focal partner	Spouse	Focal partner	Spouse	Focal partner	Spouse
Threshold	ERA	ERA	Two-year rule	Two-year rule	NRA	NRA

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table A.1 reports the estimate of interest ($\hat{\tau}$) from model 5.2 in all six months where the focal partners in the control group are eligible to claim benefits while the focal partners in the treatment group are ineligible although they are the same age in months. See appendix A.1 to obtain the estimates of all parameters.

Table A.2: The change in the spouses' employment (investigating the importance of the type of standard errors)

	(1)	(2)	(3)
First month	0.026**	0.026***	0.026***
	(0.01033)	(0.00973)	(0.00944)
	27,260	27,260	27,260
Second month	0.037***	0.037***	0.037***
	(0.01022)	(0.00973)	(0.00944)
	27,228	27,228	27,228
Third month	0.034***	0.034***	0.034***
	(0.00963)	(0.00972)	(0.00947)
	27,198	27,198	27,198
Fourth month	0.037***	0.037***	0.037***
	(0.01032)	(0.00972)	(0.00946)
	27,200	27,200	27,200
Fifth month	0.036***	0.036***	0.036***
	(0.01008)	(0.00973)	(0.00947)
	27,200	27,200	27,200
Sixth month	0.026**	0.026***	0.026***
	(0.01035)	(0.00973)	(0.00948)
	27,200	27,200	27,200
Standard errors	Clustered (running variable)	Robust	Clustered (couple level)

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table 7.1 reports the estimate of interest ($\hat{\tau}$) from model 5.2 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. See appendix A.1 to obtain the estimates of all parameters.

Table A.3: The spillover effect (investigating the importance of the inclusion of controls, the functional form, and the bandwidth)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First month	0.164**	0.166**	0.168**	0.178**	0.293***	0.247*	0.111**	0.257***
	(0.082)	(0.079)	(0.072)	(0.071)	(0.106)	(0.148)	(0.056)	(0.098)
	27,260	27,260	27,260	27,260	27,260	27,260	53,967	13,322
Second month	0.177***	0.173***	0.174***	0.183***	0.194**	0.173	0.139***	0.195**
	(0.058)	(0.057)	(0.051)	(0.051)	(0.083)	(0.116)	(0.036)	(0.081)
	27,228	27,228	27,228	27,228	27,228	27,228	53,935	13,301
Third month	0.157***	0.154***	0.154***	0.163***	0.164**	0.184*	0.136***	0.179**
	(0.053)	(0.052)	(0.046)	(0.046)	(0.077)	(0.102)	(0.032)	(0.074)
	27,198	27,198	27,198	27,198	27,198	27,198	53,905	13,301
Fourth month	0.168***	0.165***	0.167***	0.176***	0.221***	0.190*	0.137***	0.221***
	(0.055)	(0.054)	(0.049)	(0.049)	(0.081)	(0.109)	(0.033)	(0.077)
	27,200	27,200	27,200	27,200	27,200	27,200	53,877	13,301
Fifth month	0.166***	0.163***	0.165***	0.175***	0.232***	0.219**	0.119***	0.234***
	(0.054)	(0.053)	(0.049)	(0.049)	(0.077)	(0.106)	(0.033)	(0.072)
	27,200	27,200	27,200	27,200	27,200	27,200	53,834	13,301
Sixth month	0.151**	0.148**	0.149**	0.165**	0.257**	0.208	0.097**	0.254**
	(0.074)	(0.072)	(0.065)	(0.065)	(0.122)	(0.181)	(0.044)	(0.107)
	27,200	27,200	27,200	27,200	27,200	27,200	53,779	13,301
Controls (focal)		✓		✓	✓	✓	✓	✓
Controls (spouse)			✓	✓	✓	✓	✓	✓
Polynomial	Linear	Linear	Linear	Linear	Quadratic	Cubic	Linear	Linear
Bandwidth	100	100	100	100	100	100	200	50

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table A.3 reports the estimate of interest ($\hat{\rho}$) from model 5.3 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. In column (1), I set $\beta_2 = \beta_3 = 0$ from model 5.3 prior to estimation. In column (2), I set $\beta_3 = 0$ from model 5.3 prior to estimation. In column (3), I set $\beta_2 = 0$ from model 5.3 prior to estimation. Column (4) shows the baseline specification with controls of both the focal partner and the spouse, a linear polynomial that is allowed to differ on each side of the cutoff, and a bandwidth of 100 days. Column (5) shows the estimates if using a quadratic polynomial, while column (6) shows the estimates if using a cubic polynomial, both where the slopes are allowed to differ on each side of the cutoff. Column (7) and (8) shows the estimates if using a bandwidth of 200 and 50 days respectively. See appendix A.1 to obtain the estimates of all parameters.

Table A.4: The change in the spouses' employment (investigating the continuity of the conditional expectation function)

	(1)	(2)	(3)
First month	0.026**	0.001	-0.012
	(0.010)	(0.010)	(0.010)
	27,260	27,260	27,260
Second month	0.037***	0.002	0.006
	(0.010)	(0.010)	(0.011)
	27,228	27,228	27,228
Third month	0.034***	0.002	0.010
	(0.010)	(0.011)	(0.010)
	27,198	27,198	27,198
Fourth month	0.037***	-0.007	0.010
	(0.010)	(0.010)	(0.011)
	27,200	27,200	27,200
Fifth month	0.036***	-0.007	0.007
	(0.010)	(0.010)	(0.010)
	27,200	27,200	27,200
Sixth month	0.026**	-0.006	0.004
	(0.010)	(0.010)	(0.010)
	27,200	27,200	27,200
Cutoff value	0	-50	50

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table A.4 reports the estimate of interest ($\hat{\tau}$) from model 5.2 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. Estimates from the baseline specification in column (1), and estimates with pseudo cutoffs in column (2) and (3). See appendix A.1 to obtain the estimates of all parameters.

Table A.5: Test of the identifying assumptions

	(1)	(2)
First month	[0.704]	[0.561]
Second month	[0.653]	[0.553]
Third month	[0.572]	[0.563]
Fourth month	[0.571]	[0.628]
Fifth month	[0.571]	[0.628]
Sixth month	[0.571]	[0.628]
Test of assumption	RDD.2	RDD.3

Note: p-values in square brackets. Table A.5 reports the test of RDD.2 (no manipulation of the running variable) and RDD.3 (continuity of the predetermined covariates) in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB.

Table A.6: F-statistics from the first stage

	(1)	(2)
First month	214.7	1103.4
Second month	478.1	1379.2
Third month	624.0	1391.2
Fourth month	617.3	1577.5
Fifth month	578.7	1619.1
Sixth month	180.1	235.8
Outcome variable	Employment	ERB

Note: Table A.6 reports the F-statistics from the first stage, while the estimates from the second stage are reported in table 7.4.

Table A.7: The change in the spouses' employment (investigating the importance of the type of estimation)

	(1)	(2)	(3)
First month	0.026**	0.032***	0.043*
	(0.010)	(0.011)	(0.024)
	27,260	27,260	7,374
	100	100	28
Second month	0.037***	0.035***	0.039
	(0.010)	(0.011)	(0.026)
	27,228	27,228	6,050
	100	100	23
Third month	0.034***	0.031***	0.026
	(0.010)	(0.011)	(0.026)
	27,198	27,198	6,050
	100	100	23
Fourth month	0.037***	0.038***	0.032
	(0.010)	(0.011)	(0.026)
	27,200	27,200	6,591
	100	100	25
Fifth month	0.036***	0.038***	0.028
	(0.010)	(0.011)	(0.037)
	27,200	27,200	6,050
	100	100	23
Sixth month	0.026**	0.026**	0.008
	(0.010)	(0.011)	(0.026)
	27,200	27,200	6,313
	100	100	24
Approach	Parametric	Nonparametric	Nonparametric
Kernel	Uniform	Triangular	Triangular

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. The bandwidth used in the estimation is reported below the corresponding number of observations. Table A.7 reports the estimate of interest ($\hat{\tau}$) from model 5.2 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB.

Table A.8: The change in the spouses' employment (investigating the importance of the type of model)

	(1)	(2)
First month	0.026**	0.025**
	(0.010)	(0.010)
	27,260	27,260
Second month	0.037***	0.036***
	(0.010)	(0.010)
	27,228	27,228
Third month	0.034***	0.033***
	(0.010)	(0.010)
	27,198	27,198
Fourth month	0.037***	0.037***
	(0.010)	(0.010)
	27,200	27,200
Fifth month	0.036***	0.035***
	(0.010)	(0.010)
	27,200	27,200
Sixth month	0.026**	0.024**
	(0.010)	(0.011)
	27,200	27,200
Model	LPM	Probit

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table A.8 reports the estimate of interest ($\hat{\tau}$) from model 5.2 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. See appendix A.1 to obtain the estimates of all parameters.

Table A.9: The change in the spouses' employment (investigating the importance of estimating each cutoff by itself)

	(1)	(2)	(3)	(4)	(5)
First month	0.026**	0.030*	0.022	0.015	0.028
	(0.010)	(0.017)	(0.017)	(0.023)	(0.019)
	27,260	7,246	7,075	6,407	6,532
Second month	0.037***	0.046***	0.029	0.031	0.034*
	(0.010)	(0.017)	(0.018)	(0.022)	(0.019)
	27,228	7,227	7,064	6,414	6,523
Third month	0.034***	0.036**	0.016	0.042*	0.034*
	(0.010)	(0.016)	(0.017)	(0.022)	(0.019)
	27,198	7,216	7,049	6,411	6,522
Fourth month	0.037***	0.042**	0.021	0.030	0.046**
	(0.010)	(0.017)	(0.019)	(0.022)	(0.018)
	27,200	7,213	7,042	6,412	6,533
Fifth month	0.036***	0.041**	0.022	0.037*	0.036**
	(0.010)	(0.017)	(0.019)	(0.021)	(0.018)
	27,200	7,213	7,042	6,412	6,533
Sixth month	0.026**	0.036**	0.008	0.021	0.032*
	(0.010)	(0.017)	(0.019)	(0.021)	(0.018)
	27,200	7,213	7,042	6,412	6,533
Cutoff	Pooled	(I)	(II)	(III)	(IV)

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table A.9 reports the estimate of interest ($\hat{\tau}$) from model 5.2 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. See appendix A.1 to obtain the estimates of all parameters.

Table A.10: The change in the spouses' employment (investigating the importance of the definition of employment)

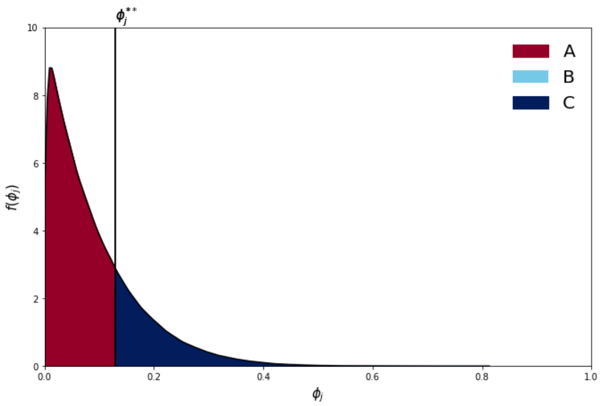
	(1)	(2)	(3)	(4)	(5)	(6)
First month	0.026***	0.024**	0.026**	0.025**	0.027**	0.024**
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)
	27,260	27,260	27,260	27,260	27,260	27,260
Second month	0.034***	0.036***	0.037***	0.036***	0.033***	0.037***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)
	27,228	27,228	27,228	27,228	27,228	27,228
Third month	0.034***	0.034***	0.034***	0.034***	0.031***	0.033***
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)
	27,198	27,198	27,198	27,198	27,198	27,198
Fourth month	0.035***	0.035***	0.037***	0.035***	0.041***	0.041***
	(0.009)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)
	27,200	27,200	27,200	27,200	27,200	27,200
Fifth month	0.028***	0.033***	0.036***	0.035***	0.032***	0.028***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)
	27,200	27,200	27,200	27,200	27,200	27,200
Sixth month	0.021**	0.024**	0.026**	0.023**	0.012	0.002
	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)
	27,200	27,200	27,200	27,200	27,200	27,200
Hours	0	5	8	10	15	20

Note: *p<0.10, **p<0.05, ***p<0.01. Standard errors clustered on the running variable in parentheses. The number of observations in each month is reported below the corresponding standard errors. Table A.10 reports the estimate of interest ($\hat{\tau}$) from model 5.2 in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB. The last row denotes the number of weekly hours that the individual on average has to have worked during the month to be defined as employed. See appendix A.1 to obtain the estimates of all parameters.

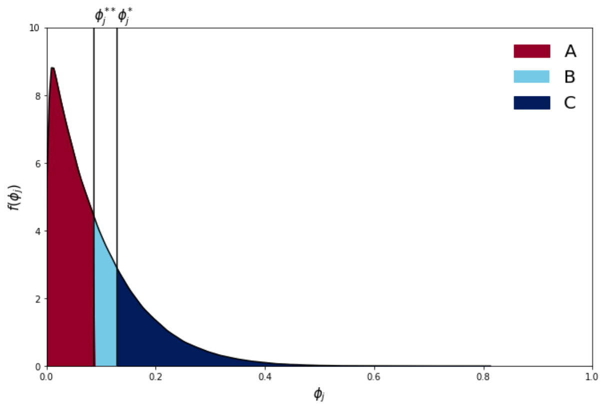
A.3 Figures

Figure A.1: The employment of the older spouses conditional on the utility of individual leisure (different values of the utility of joint leisure)

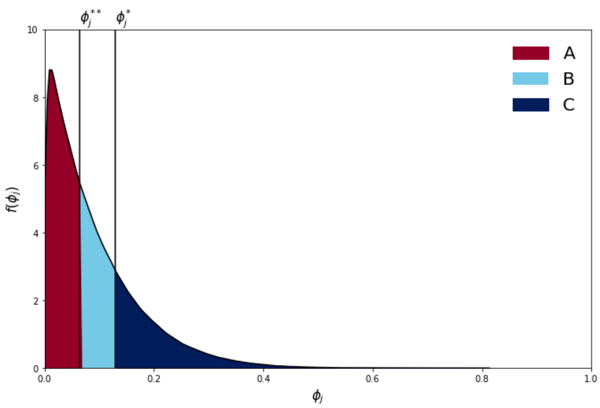
(a) $\sigma_{ij} = 0$



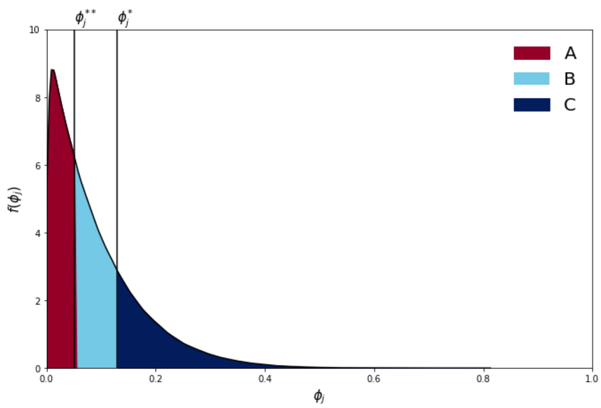
(b) $\sigma_{ij} = 0.5$



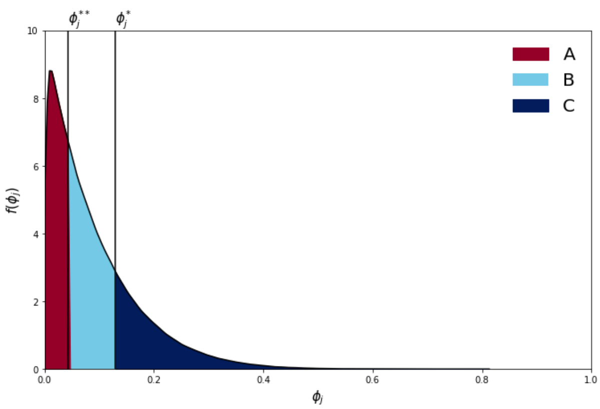
(c) $\sigma_{ij} = 1$



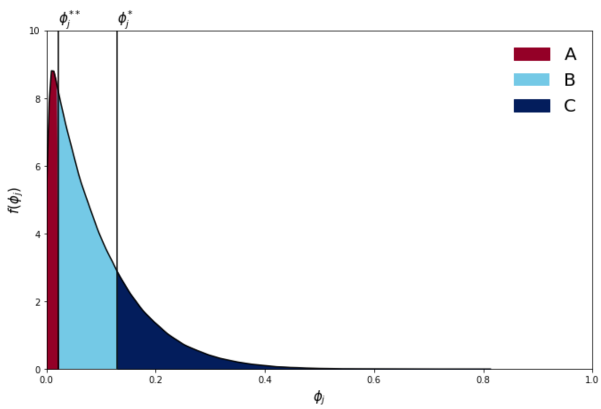
(d) $\sigma_{ij} = 1.5$



(e) $\sigma_{ij} = 2$



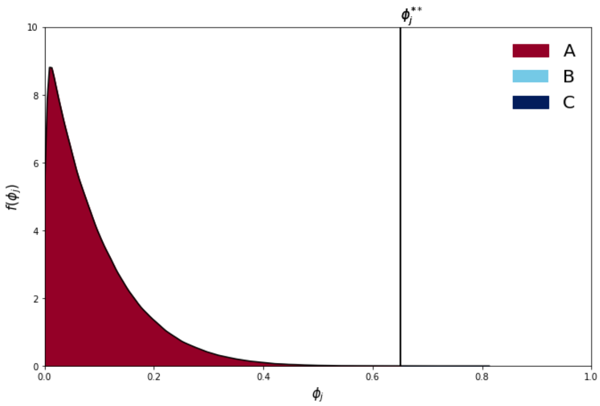
(f) $\sigma_{ij} = 5$



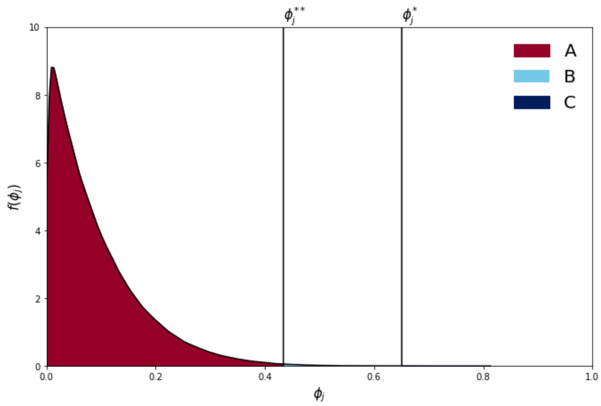
Note: The utility of individual leisure is increasing in ϕ_j, and $\phi_j \sim \beta(1, 10)$. ϕ_j^* denotes the threshold utility of individual leisure where the spouse is indifferent between working and not working if the focal partner works. ϕ_j^{**} denotes the threshold utility of individual leisure where the spouse is indifferent between working and not working if the focal partner does not work. $A = \int_0^{\phi_j^{**}} f(\phi_j)\, d\phi_j$ is the share of spouses that always work. $B = \int_{\phi_j^{**}}^{\phi_j^{*}} f(\phi_j)\, d\phi_j$ is the share of spouses that only work if the focal partners work. $C = \int_{\phi_j^{*}}^{1} f(\phi_j)\, d\phi_j$ is the share of spouses that never work.

Figure A.2: The employment of the younger spouses conditional on the utility of individual leisure (different values of the utility of joint leisure)

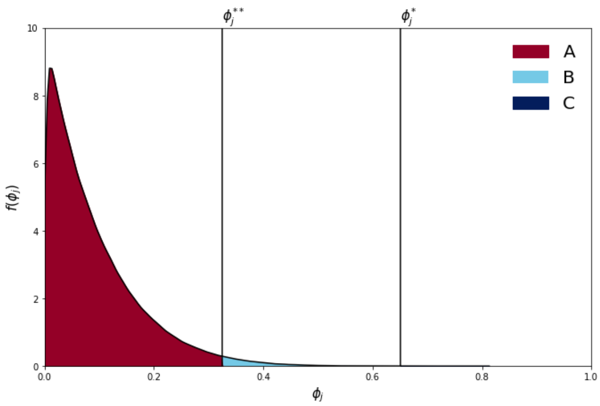
(a) $\sigma_{ij} = 0$



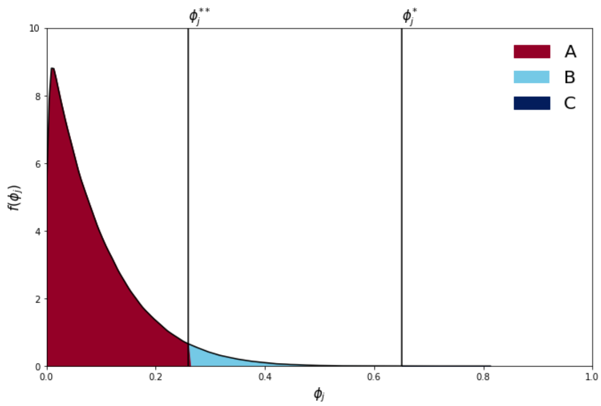
(b) $\sigma_{ij} = 0.5$



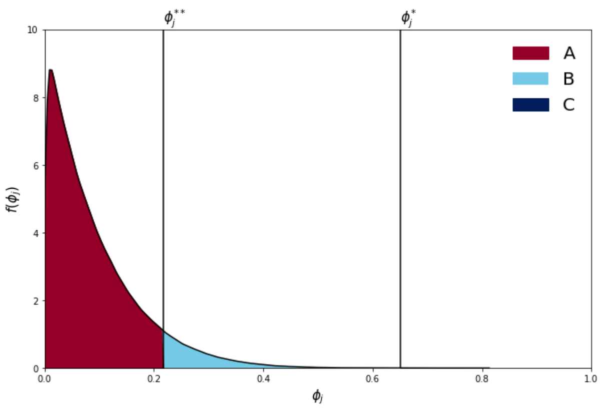
(c) $\sigma_{ij} = 1$



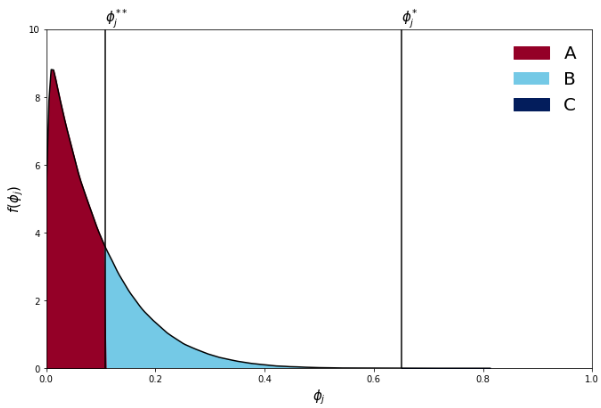
(d) $\sigma_{ij} = 1.5$



(e) $\sigma_{ij} = 2$



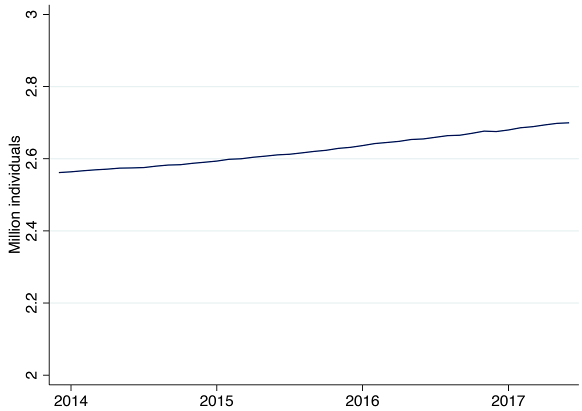
(f) $\sigma_{ij} = 5$



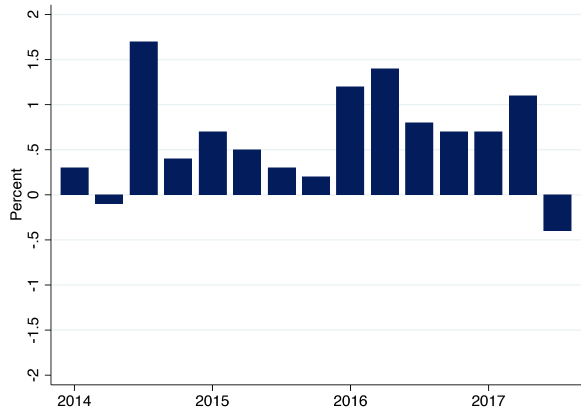
Note: The utility of individual leisure is increasing in ϕ_j, and $\phi_j \sim \beta(1, 10)$. ϕ_j^* denotes the threshold utility of individual leisure where the spouse is indifferent between working and not working if the focal partner works. ϕ_j^{**} denotes the threshold utility of individual leisure where the spouse is indifferent between working and not working if the focal partner does not work. $A = \int_0^{\phi_j^{**}} f(\phi_j)\, d\phi_j$ is the share of spouses that always work. $B = \int_{\phi_j^{**}}^{\phi_j^*} f(\phi_j)\, d\phi_j$ is the share of spouses that only work if the focal partners work. $C = \int_{\phi_j^*}^{1} f(\phi_j)\, d\phi_j$ is the share of spouses that never work.

Figure A.3: Macroeconomic conditions in the estimation period

(a) Employment



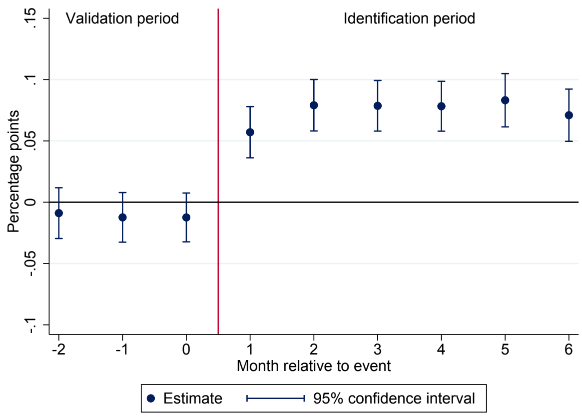
(b) GDP growth



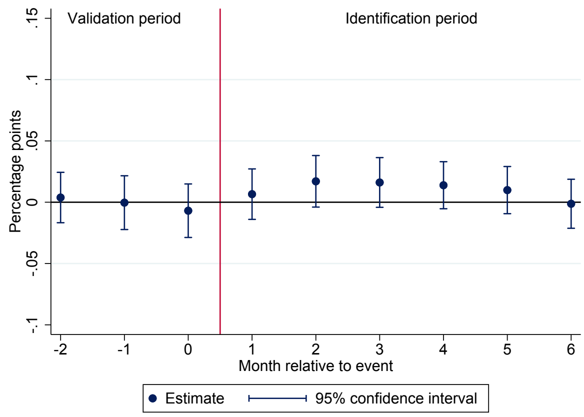
Note: The data is retrieved from Statistics Denmark. Employment is measured monthly, while GDP growth is measured quarterly. Neither employment nor GDP growth is seasonally adjusted since only the raw series is of interest in this paper.

Figure A.4: The employment response caused by changes in other thresholds

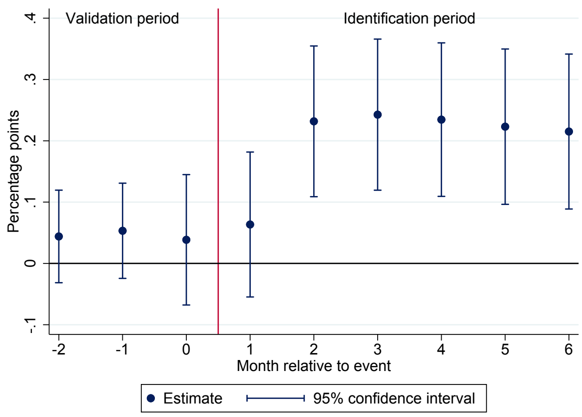
(a) By the focal partners (two-year rule)



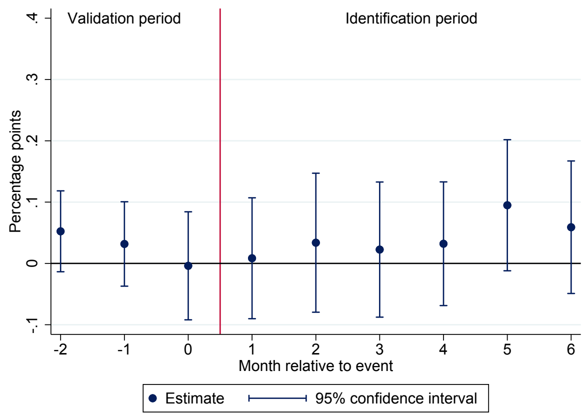
(b) By the spouses (two-year rule)



(c) By the focal partners (NRA)



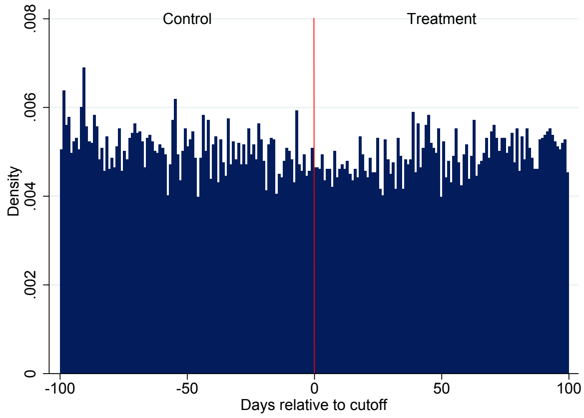
(d) By the spouses (NRA)



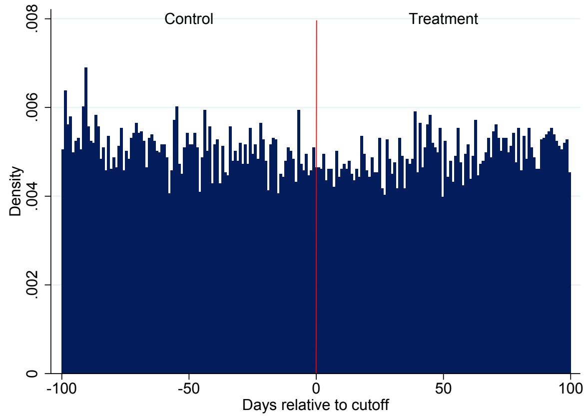
Note: Figure A.4 show the estimated values of $\hat{\gamma}$ and $\hat{\tau}$ from model 5.1 and model 5.2 respectively. In figure A.4a and figure A.4b, the estimates are shown three months before and six months after the focal partners in the control group become eligible to apply the two-year rule, while the focal partners in the treatment group are ineligible in the entire period. In figure A.4c and figure A.4d, the estimates are shown three months before and six months after the focal partners in the control group become eligible to claim national retirement benefits, while the focal partners in the treatment group are ineligible in the entire period. The sample restrictions are different in figure A.4c and figure A.4d, as only couples where neither the focal partner nor the spouse is a member of the ERP scheme are included in the sample. The estimations are conducted with a bandwidth of 100 days, a linear polynomial that differs on each side of the cutoff, and standard errors clustered on the running variable. The validation period consists of the first three where neither focal partners in the control group nor focal partners in the treatment group are eligible. The identification period consists of the six months where the focal partners in the control group are eligible, while focal partners in the treatment group are ineligible although they are the same age in months. The estimates with corresponding standard errors are presented in table A.1 in appendix A.2.

Figure A.5: Density of the running variable

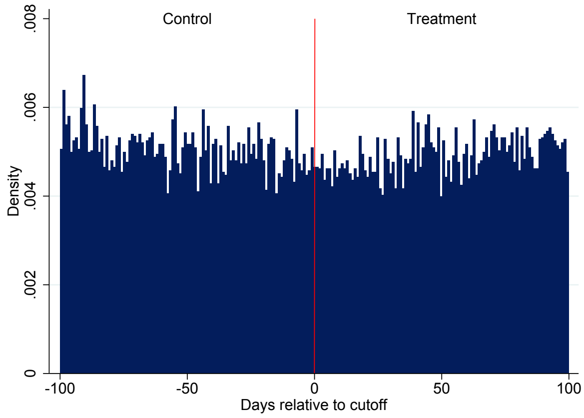
(a) First month



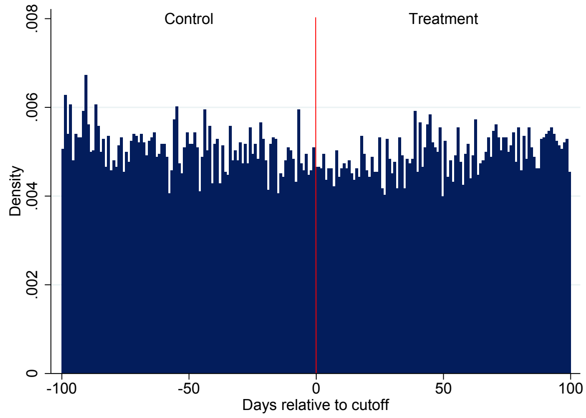
(b) Second month



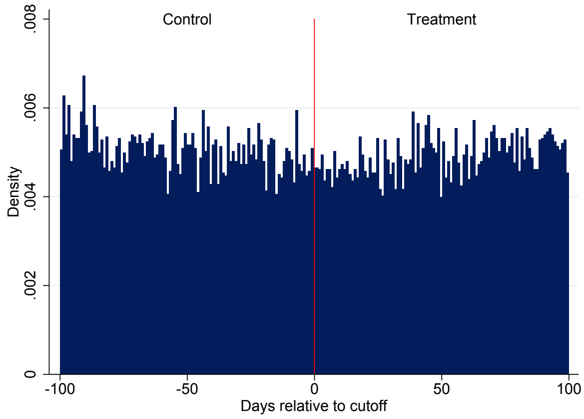
(c) Third month



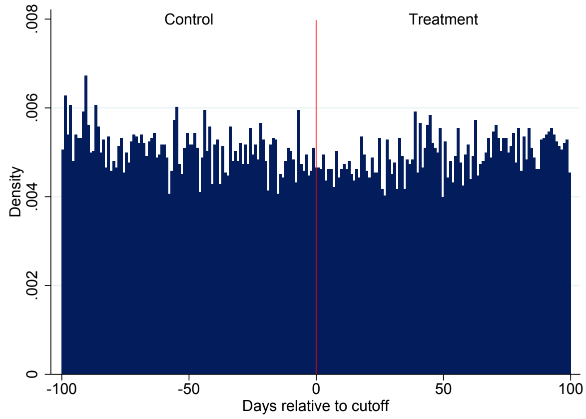
(d) Fourth month



(e) Fifth month



(f) Sixth month



Note: Figure A.5 shows the density of the running variable in each of the six months in which the focal partners in the control group and the treatment group are the same age in months, but where only focal partners in the control group are eligible to claim ERB.

A.4 Computing the net contribution

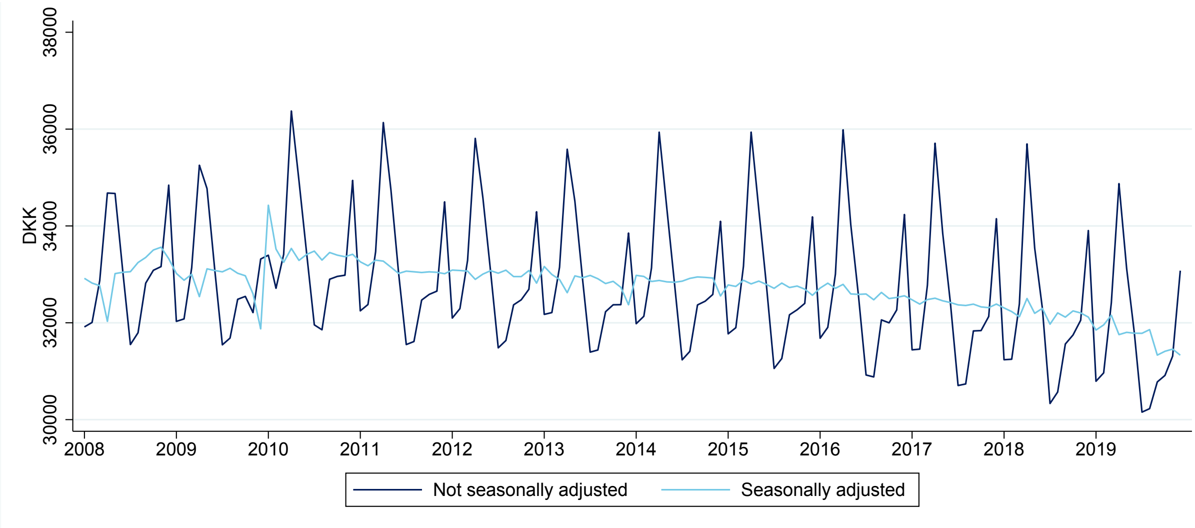
I define the individual net contribution to the government budget as the individual's tax payment on labor income minus the net-of-tax take-up of benefits. These variables are not a part of the administrative data, meaning that they have to be computed. It requires certain assumptions to approximate the net contribution, but the estimation will provide unbiased estimates if there is no systematic deviation between the actual and the estimated net contribution. In Denmark, people receive gross benefits, which are taxed. The tax payment is defined as taxes paid of labor income, while the take-up of benefits is defined net-of-taxes to ensure that raising the ERA does not underestimate the positive effect of larger tax revenue and overestimate the positive impact of reduced public expenditures.

The tax payment

I compute the tax payment on labor income by multiplying the labor income in each month by the yearly average tax rate on labor income. The former is part of the administrative data but is measured in current prices. I adjust the wages by the evolution in the wage levels in the industry to ensure that wages are measured in constant prices. Yet, figure A.6 shows that the average monthly wages follow seasonal patterns where the average monthly wages are greatest in March, April, May, and December. This results from public holidays and the timing of the payments of bonuses, not because people work more or are more productive in these months. As I consider the spouses' labor income when the focal partners are the same age in months, I compare the spouses' monthly wages in different calendar months. Thus, I obtain wrongful conclusions if ignoring the seasonal patterns of the monthly wages. I follow the same procedure as Statistics Denmark (2002) to adjust the average monthly wages by an ARIMA-11 correction where the seasonally adjusted average monthly wages can be seen in figure A.6 relative to the not seasonally adjusted monthly wages. Dividing the not seasonally adjusted average monthly wages by the seasonally adjusted average monthly wages, I estimate the factor in each month of each year since 2008. This is known as *the multiplicative method*, and this method is recommended by Johansen & Trier (2010) when adjusting for seasonally patterns. The estimated factor is multiplied with the spouse's monthly wage in the corresponding year and month to ensure that the monthly wages remain stable across the period.

I calculate the average tax rate on labor income by taking the sum of the social security contribution, the municipality tax payment, the health care contribution, the bottom tax payment, the top tax payment, and the church tax payment and divide this with the taxable income reported by the tax authorities, all measured annually. Finally, the tax payment on labor income in the given month is obtained by multiplying the month's seasonally adjusted labor income by the yearly average tax rate.

Figure A.6: The average monthly wages



Note: All wages are measured in 2019-prices. The wages are adjusted before the seasonally adjustment by the evolution in the wage levels in the industry. An ARIMA-11 correction is applied to obtain the seasonally adjusted monthly wages.

The take-up of benefits

The type of benefit that the individual receives is recorded weekly, while I observe the take-up of benefits annually. Dividing the take-up of each benefit with the number of weeks the individual received the corresponding type of benefit within the year, the weekly take-up of benefits in DKK is obtained. The monthly take-up of benefits is obtained by adding the number of weeks within the month where the individual receive the benefit.

I calculate the sum of the municipality tax payment, the health care contribution, the bottom tax payment, the top tax payment, and the church tax payment. The tax payment on benefits is computed by dividing this sum with the taxable income reported by the tax authorities, all measured annually. I subtract the tax payment on the benefits from the gross benefits to ensure that I compute the actual government expenditures to the individual's benefits. The take-up of benefits is adjusted to account for wage inflation such that the take-up of all benefits, similar to labor income, are measured in 2019-prices.