Joint Retirement in Couples

Evidence of Asymmetric Complementarity in Leisure

Herman Kruse*

June 2018

Abstract

This paper studies the spillover effects in joint retirement decisions within couples. I exploit a major reform of the Norwegian pension system from 2011, which provided a quasi-experimental setting, to obtain causal estimates of spousal spillover effects in labor market decisions. I use highly detailed register data on the entire Norwegian population from 1993-2015. Disentangling spillovers is often difficult due to absence of independent individual variation even in detailed register data, but this quasi-experimental setting offers an opportunity to identify individual variation in work incentives. The results show that spousal spillovers are asymmetric. Wives respond to their husbands' choices, but not vice versa. A lower bound estimate of complementarity of leisure shows that women eligible for retirement are about 20-24 percentage points more likely to retire if their husband is retired. There is no indication of men responding to their wives' choices.

Keywords: Retirement, Policy evaluation, Couples, Leisure complementarity, Spousal spillovers
JEL classifications: J14, C01, C36, D04

*I acknowledge the Frisch Center for access to their data files. All the data belongs to Statistics Norway and are used in compliance with the rules given by the Norwegian Data Inspectorate. I am grateful towards the Norwegian Research Council for the financial support (The evaluation of the pension reform). The invaluable comments and suggestions from Simen Markussen and Kjetil Storesletten are greatly appreciated. I am grateful for all helpful comments and suggestions from Andreas Myhre and Ragnar Nymoen, and from seminar participants at various events at the Department of Economics, UiO.

1 Introduction

Labor supply decisions among individuals in couples are naturally influenced by interdependencies between the partners. In particular in decisions of retirement, the decision of one partner might have spillover effect on the spouse due to complementarity of leisure. When incentives are interlinked between individuals, policy changes affecting certain groups may have wider implications also for those not directly affected. Therefore, estimating spousal spillovers is key to correctly quantify aggregate labor supply effects of policy changes. However, disentangling spousal spillovers usually proves difficult using information even in detailed register data, due to the absence of independent individual variation. Thus, a reform affecting the work incentives of a subset of workers offers an intriguing opportunity to study spousal spillovers in a quasi-experimental setting. I will exploit a major reform of the Norwegian pension system from 2011 to sharply identify work incentives among elderly and the spillover effects between spouses.

The increasing ratio of elderly dependents to the working population puts massive pressure on long-term fiscal sustainability for pension systems across the OECD (OECD, 2017). In Norway, a reform from 2011 targeted this development by increasing work incentives among the elderly. More specifically, it removed a confiscatory earnings test among private sector workers with access to an early retirement scheme[1], which can be interpreted as a removal of an implicit tax on continued work[2]. Since essentially everyone else was unaffected in terms of work incentives and the reform was swift, this reform offers a quasi-experimental set-

[1]Pre-reform, the workers in public sector and approximately half of the private sector had access to an early retirement scheme, offering full pension from 62 (old-age pension could be claimed from 67). However, the early retirement pension was subject to a full confiscatory earnings test, from the first dollar earned after claiming benefits. This implied a very high implicit tax rate on continued work and strong disincentive to work past the age of 62.

[2]This group amounts to approximately 23 percent of the workers.

ting to study labor market responses to work incentives among elderly. As will be described below, this reform offered a clear identification of interlinked work incentives and their implications for aggregate labor market effects of reforms targeting work incentives.

The ultimate aim of economic policy evaluation is to find the aggregate effects of the policy change. However, it might be misleading to rely on models of individual behavior for obtaining aggregate implications, since individual-based models do not incorporate indirect spillover-effect between interlinked individuals. Standard economic theory predicts two indirect effects of postponed retirement in a couple (Mincer, 1962): an income effect and a substitution effect. First, when one partner postpones retirement, the household income increases on the margin. This will lead to a lower marginal utility of consumption, which in turn will lower the work incentives of the spouse. Second, postponing retirement of one partner means that the partner enjoys less leisure than if he or she was retired. This will affect the marginal utility of leisure of the spouse. If leisure is considered a substitute, the work incentives of the spouse are decreased. This is typical for families with responsibility for young children, where one spouse typically specializes in childcare and the other in the labor market (Blundell et. al, 2015). However, among older couples, it may be more plausible to believe that leisure is a complement, which would lead to increased work incentives for the spouse if one partner postpones retirement. We then have two opposing effects; when one partner postpones retirement, the income effect is negative on work incentives of the spouse, while the effect of complementarity in leisure is positive on work incentives of the spouse. Due to these two effects, which cannot be disentangled using a single measure of work incentives, the estimates provided will be *lower bound* estimates of leisure complementarity. The implication of a dominating complementarity effect is that

aggregate effects are downward-biased when spousal spillovers are disregarded.

I use register data on the entire Norwegian population from 1993-2015. I will focus on couples in retirement age and analyze the effect of the pension reform jointly in a couple, using a difference-in-differences method to capture the effect of increased work incentives for one individual on the labor market participation of the spouse. As will be documented below, the reform sharply increased the work incentives for private sector workers with access to an early retirement scheme, while it was neutral in terms of work incentives for everyone else. I exploit this sharp and abrupt change to identify work incentives for individuals. Then I am able to estimate how work incentives for one individual affects the labor participation of the spouse. For instance, if a woman is directly affected by the reform[3], it can have indirect implications for the labor market participation of her husband. The implication is that incentives possibly can account for a larger part of aggregate effects of reforms in the labor market than previous studies are able to account for. It also has substantial implications for how we understand the formation of preferences in couples, in particular the non-separability of leisure between spouses.

The results show that the direct effect of the reform matters for both men and women. When directly affected by the reform, individuals tend to postpone retirement. I find that 63 year old men who are directly affected by the reform are about 13.9 percentage points less likely to retire, while 63 year old women who are directly affected are 14.7 percentage points less likely to retire. There is strong evidence of spillover effects on the wives when their husbands are directly affected by the reform. If the husband of a 63 year old woman is retired, she is

[3]Here, being directly affected by the reform means that an individual belongs to the group who had the earnings test removed, and hence had increased work incentives. An individual is indirectly affected by the reform if the spouse of that individual was directly affected. More details of the reform will be provided in a later section.

approximately 21.5 percentage points more likely to retire as well. This suggests that wives' marginal utility of leisure increases when their husbands are retired, and I interpret this as leisure complementarity. However, men do not respond to their wives being retired. This suggests that spousal spillovers are asymmetric, and that the income and complementarity effects cancel for men. The aggregate labor supply effect among all 63 year old women is a reduction in the probability of retiring of 3 percentage points[4]. However, if we ignore spousal spillovers, the aggregate effect is only 2.2 percentage points. This implies that 25 percent of the aggregate effect of the reform on women comes through spillover-effects from their husbands. This clearly shows that aggregate effects found by aggregating models of individuals are downward-biased when there is complementarity. For men, the aggregate labor supply effect among all 63 year olds is a 4 percentage points reduction in the probability of retiring [5].

This paper contributes to the joint retirement literature in several aspects. First, the paper shows that incentives can explain a much larger part of adaption in the labor market literature than what they have been attributed in previous literature. Because this paper is able to identify large spillover effects, incentives are interlinked between spouses and therefore changes to incentives for one spouse has implications beyond the direct effects, which has a substantial implication on the importance of incentives. Second, this paper shows that leisure should be treated as non-separable between spouses in a household, due to the presence of complementarity in leisure. Third, the implication of complementarity in leisure will be relevant for understanding the aggregate effect of policy changes. If such effects are

[4]I compute this aggregate by summing the direct and indirect effect among all women who are directly and indirectly affected and divide by the number of women.

[5]This is larger in magnitude than in Hernæs et. al (2016) who found an effect on Norwegian men of 1.6 percentage points using the same reform. However, they restrict their attention to single-individual households and abstract from joint retirement.

neglected, one will typically underestimate the aggregate effect of a change in labor supply incentives. Fourth, this paper uses administrative data on the individual level for the entire Norwegian population and is thus not plagued by attrition and mismeasurement due to self-reporting. Fifth, the reform that is exploited as identification was a swift process with little time and possibility to self-select into treated groups. Hence workers could not strategically move to those groups that would be beneficial to them. Sixth, the reform also affected identifiable groups in different ways which gives strong and convincing evidence of a causal estimation.

1.1 Related literature

This study relates to the increasing literature on joint retirement behavior, which emerged from Hurd (1990) and Blau (1998). The most recent Norwegian study related to this is that of Johnsen and Vaage (2017) . They study spousal spillovers in retirement in Norway, and find that male workers affected by changes in work incentives affect the employment rates of their wives. Their methodology is quite similar to this paper. However, they use a different reform with different implications for work incentives. They use a reform from 1989, which introduced early retirement schemes for public sector workers and centrally negotiated private sector firms. Hence, their reform is a negative incentive change and it targets a different worker group than the reform from 2011 did.

Lalive and Parrotta (2016) is another recent study that is related to this. Their approach is not based on a reform, but uses Swiss register data to conduct a double regression discontinuity design using institutional variation in eligibility to identify effects of individuals' own pension eligibility on their partners' work incentives. They find, similar to this paper, that partner pension eligibility matters for labor force exit for women, but no significant effect for men. Stancanelli and Van

Soest (2012 a,b) use the same empirical strategy as Lalive and Parrotta (2016), but instead estimate the effects of partners' retirement on home production or joint leisure on French survey data for 1,000 couples. They do not find significant effects of partners' eligibility on individuals' own retirement, but do find that female partners' retirement increases the hours of joint leisure. The effect is not significant for men.

Another related study is by Gerard and Nekby (2012), who uses a Swedish pension reform from 2001 to conduct a difference-in-difference-in-difference identification strategy to derive the spillover effects of spouses work incentives. They suggest that ignoring the impact of spousal spillover effects underestimates the impact of the pension reform by 14 percent. However, the Swedish pension reform did not incorporate a swift change in work incentives in the same way as the Norwegian pension reform. Hence, while the Norwegian pension reform offers a clear identification of work incentives, the Swedish reform did not. Casanova (2010) presents a structural model of joint retirement, where leisure complementarities are positive and significant for both spouses. Bratsberg and Stancanelli (2018), developed independently of this paper, also study the Norwegian 2011 reform. They impose a highly restrictive sample selection in terms of age and age differences, as well as cohort selection, and therefore end up studying just a fraction of the sample studied in this paper.

The remainder of this paper is organized as follows. Section 2 provides background information on the institutional settings. Then section 3 presents the data and descriptive statistics and figures. Section 4 explains the empirical strategy, and section 5 lays out the results. Finally, section 6 concludes.

2 Institutional setting

In the pre-reform system the earliest access age for public pension was 67 years of age, but around two thirds of workers had access to an early retirement scheme (AFP), offering access to early take-up of benefits from the age of 62. Early take-up of pension benefits was subject to a full confiscatory earnings test, meaning that continued work after retirement resulted in a proportional cut in benefits. This earnings test can be viewed as a strong disincentive to continue working after reaching early retirement age for individuals with access to AFP. For the purpose of this paper, Table 1 sums up all the relevant features of the pre-reform system.

The post-reform system completely removed this earnings test of continued work

Table 1: Pre-reform institutional system

Public sector with AFP	**Private sector with AFP**	**No AFP, private and public**
AFP from 62	AFP from 62	
Earnings test (disincentive)	Earnings test (disincentive)	
Old-age pension from 67	Old-age pension from 67	Old-age pension from 67

for private sector workers with access to AFP. The post-reform system in this sense decouples the timing of pension claims and the timing of labor market participation, since one is now no longer implicitly disincentivized to combine continued work and pension claims. For a comprehensive study on the details of the pension reform, see Fredriksen et. al (2015) or Hernæs et. al (2016). As shown in Hernæs et. al (2016), the incentive changes due to removal of the earnings test are of substantial importance for labor market participation, but the other changes had negligible importance for labor market participation[6]. Thus, this paper abstracts

[6]The actuarial adjustment implies that early takeup of benefits result in a cut in annual pension, based on average life-expectancy within birth-cohorts. This means that individuals with

from those changes and will only focus on the incentive changes due to removal of the earnings test in the private sector with AFP. Table 2 therefore summarizes all the relevant changes from the post-reform system for the purpose of this study.

For private sector workers with access to AFP, the post-reform system was a direct

Table 2: Post-reform institutional system

Public sector with AFP	**Private sector with AFP**	**No AFP, private and public**
Access to actuarially fair adjusted takeup of public pension from 62	Access to actuarially fair adjusted takeup of public pension from 62	Access to actuarially fair adjusted takeup of public pension from 62
No changes to work incentives Still faces earnings test	**Improved work incentives** Removal of earnings test on continued work	**No changes to work incentives**

increase in work incentives. The removal of the confiscatory earnings test introduces strong work incentives since this group can now freely combine continued work and early take-up of benefits. Private sector workers with AFP thus have increased work incentives.

For public sector workers, the full confiscatory earnings test is preserved, meaning that they still face strong disincentives to continue working after the age of 62. They thus have no change in work incentives. Only private sector workers with access to AFP had increased work incentives, while all other groups faced unchanged

life-expectancy lower (higher) than the average of their own cohort may find the actuarial adjustment disadvantageous (advantageous). However, as Hernæs et. al (2016) show, the actuarially fair adjustment only has a very small response on the intensive margin, while maintaining employment status. Hence, the actuarially fair adjustment is purely viewed as improved liquidity and not as incentive changes.

work incentives post-reform[7]. This introduces an identifiable source of variation that can be exploited in the identification of work incentives.

3 Data and sample restrictions

The data is administrative register data on individual level for all individuals in Norway from 1993-2015. Individuals are identified using a unique identification key, and can be linked to their spouse through this key. The data include annual information on labor earnings, wealth, birth cohort, education and access to AFP in private sector or public sector. The selected cohorts are 1942-1952, and I restrict the sample to include only those couples where both work[8] at age 60, both are residing in Norway at age 60 and both partners are alive at age 60. I make these restrictions to make sure that the individuals we include are actually claiming benefits in Norway, and that they have not already exited the labor market for other reasons than retirement (for instance disability insurance). I am also interested in couples, and a widowed spouse in terms of leisure complementarity must be considered a single individual. Hence both must be alive. Note that the spouse of those born in cohorts 1942-1952 can be born outside of those cohorts. For example, when studying the effect of women on men, the wives are not restricted to be born within the 1942-1952 cohorts.

[7]Note that there are exceptions from this, for some few sub-groups who might be affected differently than the group they belong to in this definition, but these very small sub-groups are disregarded here.

[8]Working is defined as receiving labor earnings greater than 2G. 1G is 93,634 NOK (approximately $11,500) in 2017 and is adjusted by the government every year to match the growth rate of the wages. 1G is approximately 17.5 percent of the median wage across all households in Norway, which in 2017 was 535,900 NOK (approximately $66,200).

3.1 Descriptive statistics

Table 3 shows descriptive statistics pre-reform (pre-1949 cohorts) and post-reform (1949-1952 cohorts). The differences in the pre-reform and post-reform groups, in particular the income and wealth, reflects time-trends. The other observed characteristics are quite similar pre- and post-reform. The small increase in the fractions with higher education can be explained by the expansion of higher education that is observed in most Western countries.

Table 4 shows the composition of 61 year old workers in the different sectors pre-reform and post-reform. As Table 4 shows, there is little evidence of self-selection into treatment in the sense that the fraction working with different retirement schemes is fairly stable across the pre-reform and post-reform groups. Changing work place just before retirement will not result in possibility to claim AFP, hence any change from the pre-period to the post-period is a result of time trend and not adaptation to the reform. The fraction not working can be assumed to leave the labor force through the disability insurance channel. There are sizeable gender-differences in the fractions working in each sector. However, since the incentive changes in the reform are not targeted specifically towards sub-groups of workers within each category, this will only result in a level difference in the aggregate effect of men and women, and have no implications for the causality measure. It is clear that a larger fraction of men will be directly affected by the reform than women, since a larger fraction of men work in private sector with an AFP scheme. Therefore, if men and women respond similarly to work incentives, the aggregate effect for men should be higher than for women.

Figure 1 shows six panels with retirement patterns pre- and post-reform. The figure clearly shows that the exit rates among those working in private sector with

access to AFP has declined sharply for all ages. There is a small decrease in exit rates also in the four panels for workers in public sector with AFP and those without access to AFP, which cannot be explained by direct reform effects since these four groups are unaffected in terms of work incentives. Possibly, the decline among those groups could be due to time trends, for instance due to general health improvement or an increase in the average willingness to work longer. However, there could be another explanation, namely spillover effects. If those individuals who postpone retirement post-reform are married to individuals who were affected by the reform, that could explain why they choose to postpone retirement. In order to investigate this, we need to consider joint exit patterns in couples.

Table 3: Descriptive means for men and women in cohorts 1942-51 and their spouse, at age 61

Variable	Men, and their wives		Women, and their husbands	
-	Pre-ref.	Post-ref.	Pre-ref.	Post-ref.
Husb. income	7.04	7.38	5.22	5.52
Wife income	4.62	4.75	4.51	4.75
Husb. wealth	20.79	25.37	20.51	24.66
Wife wealth	7.84	10.59	8.33	11.11
Husb. higher educ	.37	.40	.35	.37
Wife higher educ	.32	.38	.31	.35

Note: Income and wealth measured in basis points (G) and education in the fraction with higher education, defined as bachelors degree or higher. Pre-reform are cohorts born in 1942-1948, post-reform are cohorts born in 1949-1952.

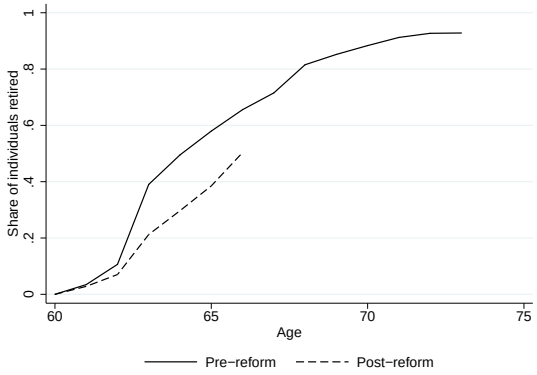
Table 4: Percent of 61 year old workers with different retirement schemes.

Scheme	Male		Female	
-	Pre	Post	Pre	Post
AFP private	26.29 %	28.92 %	13.77 %	14.59 %
AFP public	27.65 %	27.85 %	55.47 %	58.84 %
No AFP	40.52 %	39.30 %	24.89 %	22.22 %
Not working	5.53 %	3.94 %	5.87 %	4.34 %

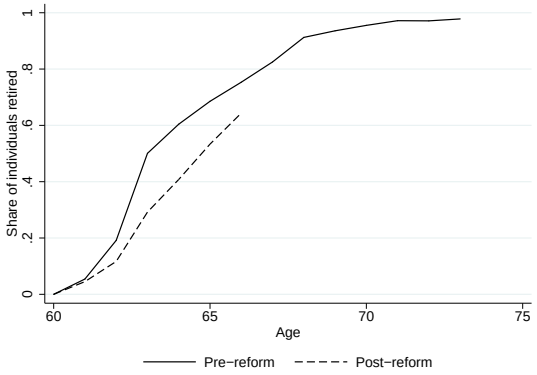
Note: Pre-reform are cohorts born in 1942-1948, post-reform are cohorts born in 1949-1952.

Figure 1: Labor market exit, married male and female individuals

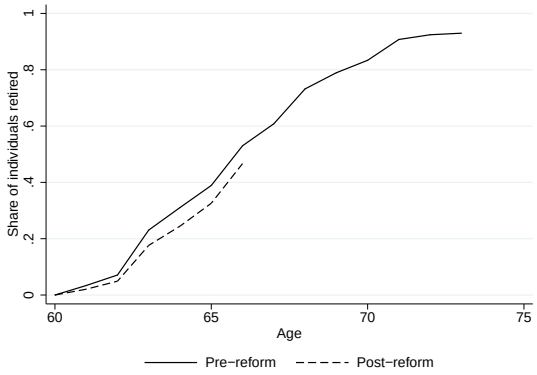
(a) Males, access to private AFP



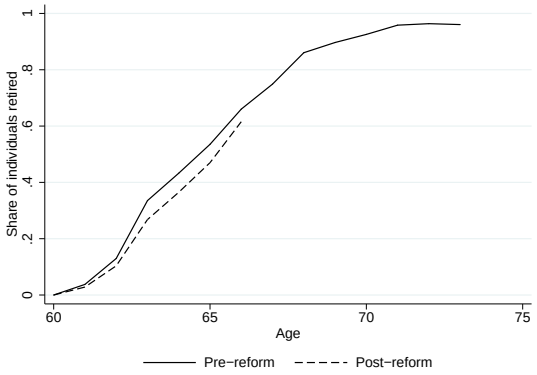
(b) Females, access to private AFP



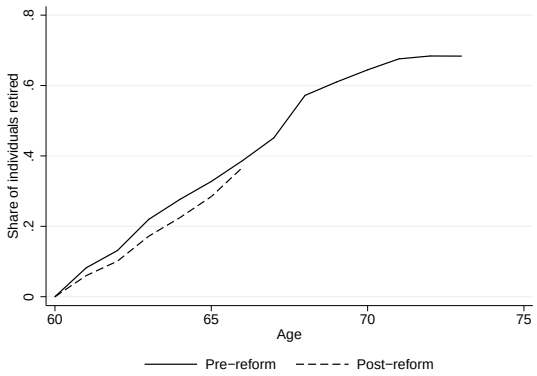
(c) Males, access to public AFP



(d) Females, access to public AFP



(e) Males without access AFP



(f) Females without access to AFP

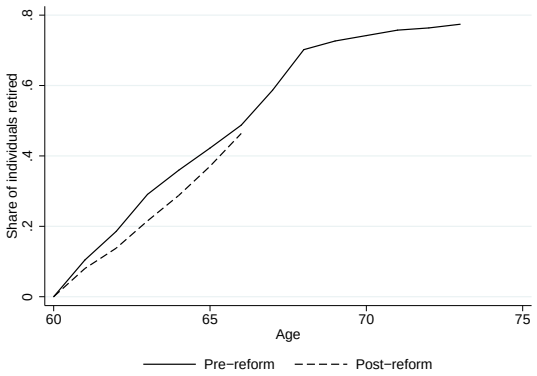


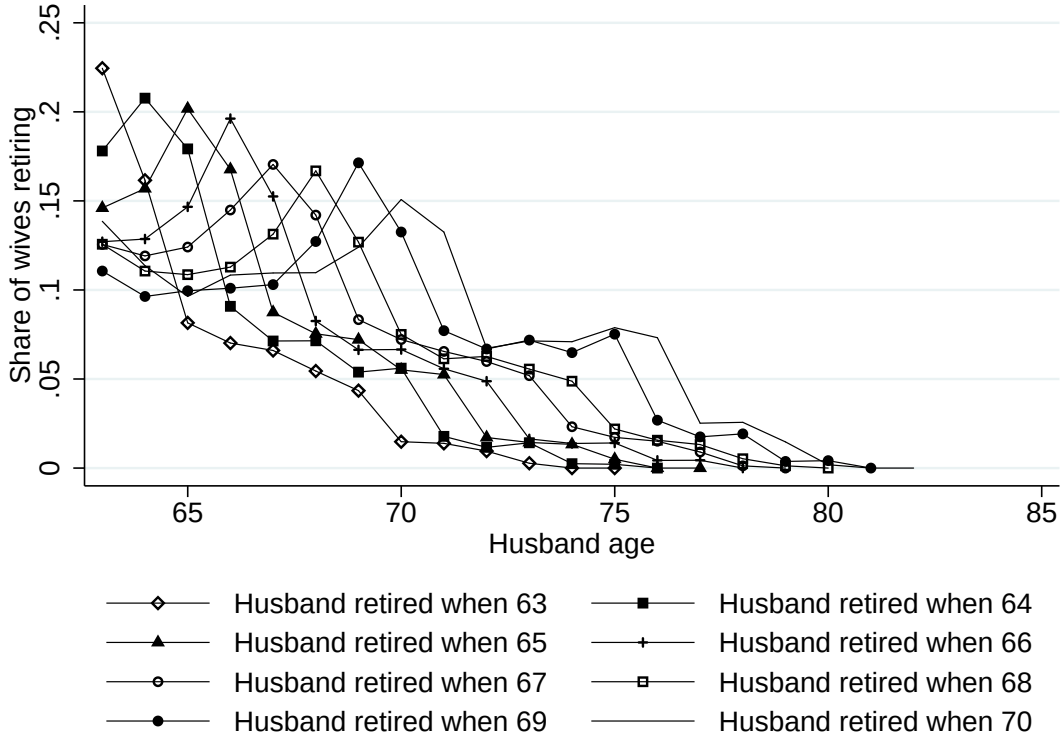
Figure 2a plots the exit rate among women according to the age of their husband when he retired. On the horizontal axis, I plot the age of the husband, while on the vertical axis I plot the share of wives retired. Each line represents the age of when the husband retired. Thus, a peak mass around the same age means that a large mass of the women retire within the same calendar year as their husband. We see that all the lines have peaks that correspond to the age they represent, which means that a large fraction of couples retire within the same calendar year. Figure 2b shows the same pattern for men.

In order to investigate the spillover-mechanism, Figure 3 plots the change in the probability of retirement post-reform by the treatment status of the spouse. In panel 3a, we see that among women in public sector, there is a large gap between those married to men in private sector with AFP and those married to men who are not working in private sector with AFP. Thus, if the husband is treated and thus receives increased work incentives, the probability of full retirement for the wife decreases more than for those women who are married to non-treated men. This is an indication of spillover-effects. However, if we look at panel 3b, the lines are almost overlapping[9], which indicates that the spillover-effects are not present for men.

[9] At least for the ages of interest which is 63 and 64. Age 62 is not interesting due to the data being annual data and the possibility of mis-interpretations due to eligibility-concerns. If an individual is born late in the calendar year, then being observed as a 62 year old and thus "assumed" to be eligible for retirement, we might run into mis-interpretations of results. We cannot estimate on individuals who are older than 64 due to lack of data.

Figure 2: Joint exit patterns

(a) Labor market exit among women, by husband's age



(b) Labor market exit among men, by wife's age

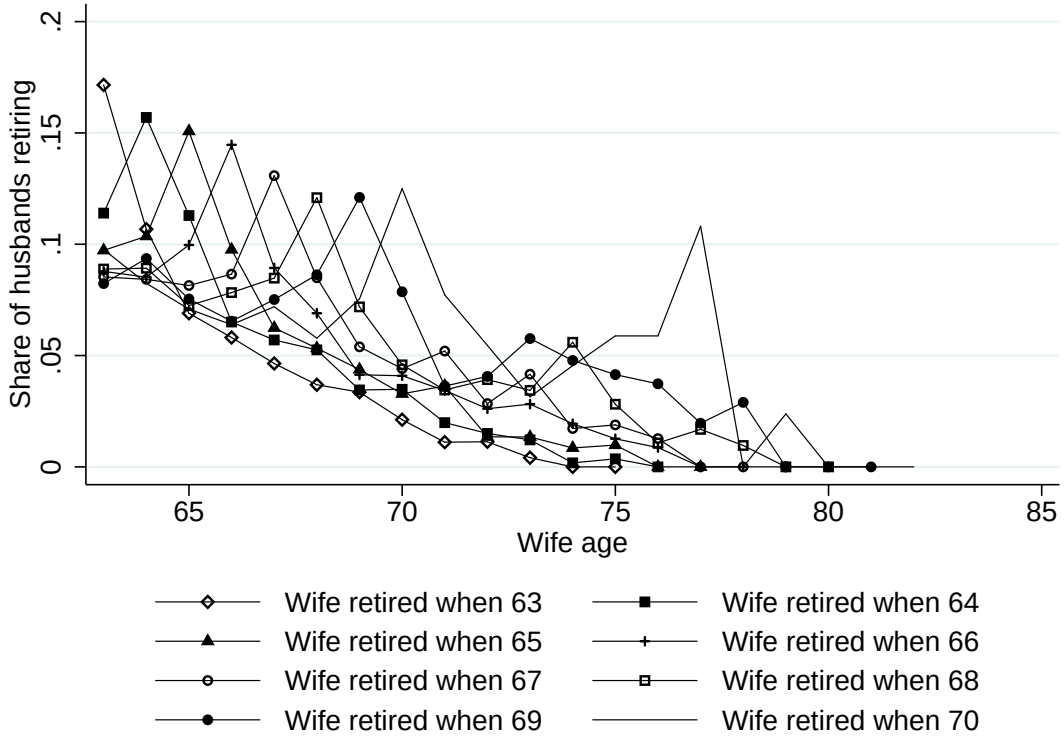
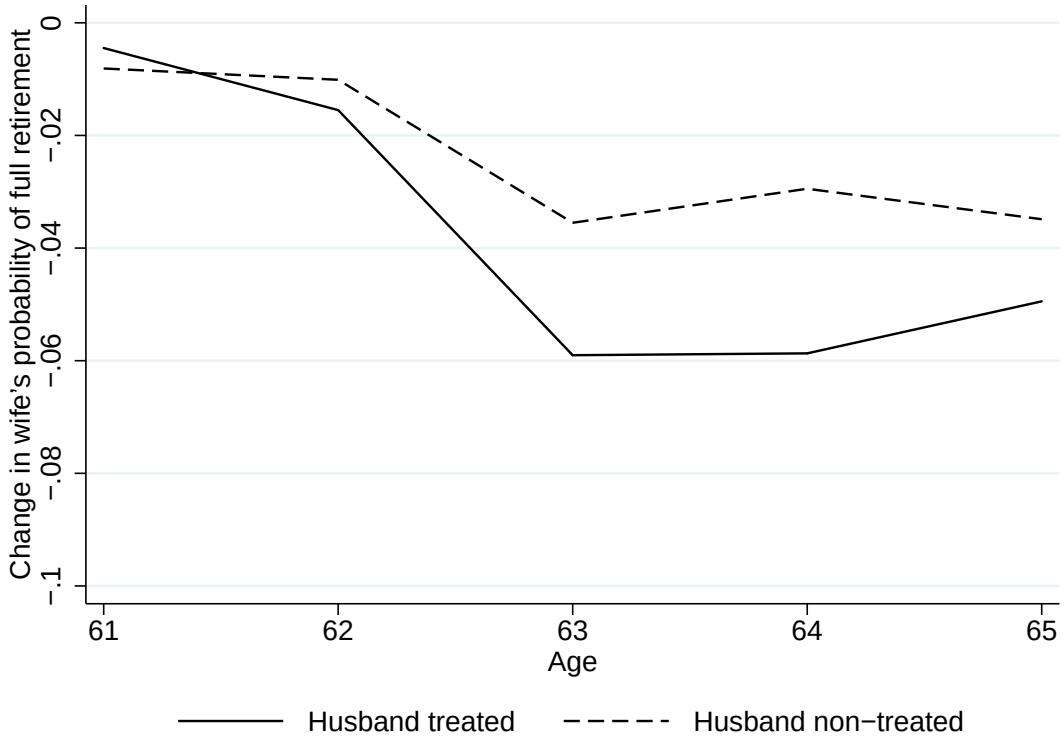
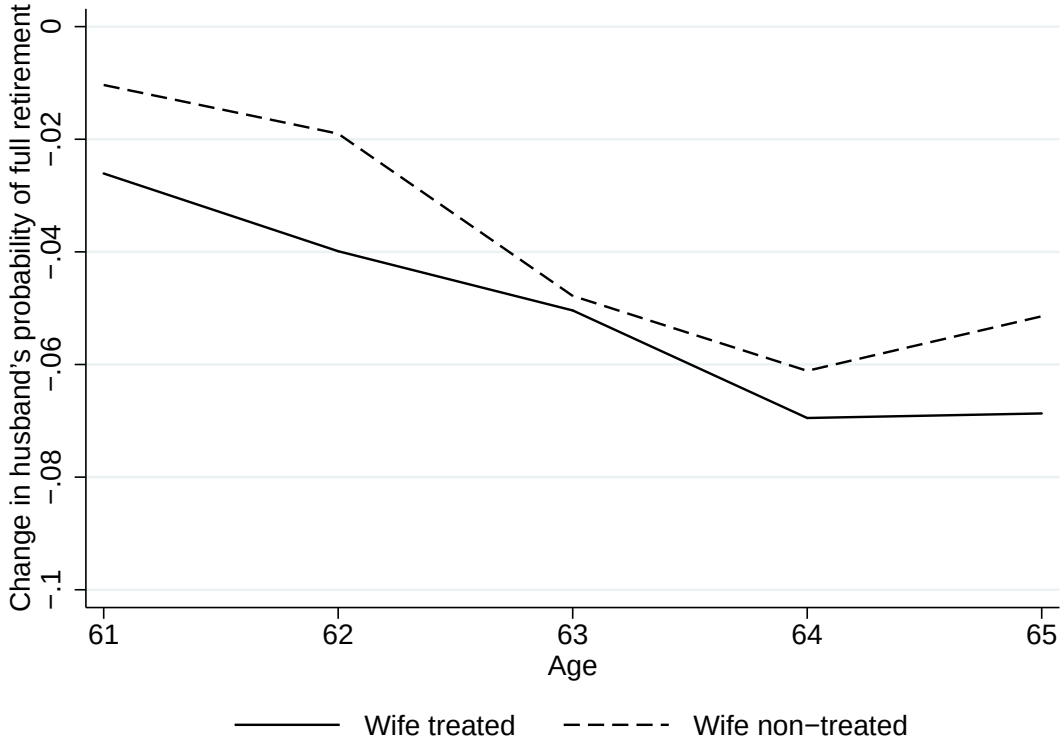


Figure 3: Spillover-effects

(a) Change in probability of full retirement among women in public sector, post-reform



(b) Change in probability of full retirement among men in public sector, post-reform



4 Empirical framework

The econometric method used is difference-in-differences. Since we have to tackle the issue of endogeneity in the labor market participation of the spouse, we need to find an instrument to identify this variation. The identification strategy is to exploit the variation in access to AFP in private sector and the cutoff before the 1949 cohort, by including those without access to private sector AFP (namely those without access to any AFP and those with access to public AFP), I control for all common time trends or changes across cohorts in labor participation among elderly. Throughout, I will refer to individual j in the model as "the focal partner", which means that this is the individual for which we are estimating spillovers onto. I will refer to individual i in the model as "the spouse", which means that this is the individual who is partnered with "the focal partner". The econometric model looks as follows[10]:

$$R_{a_j,i} = \mathbf{X}_{a_j}\mathbf{B}_{a_j,i} + \gamma_{a_j,i}treat_i + \xi_{a_j,i}treat_j + \varepsilon_i \quad (1)$$

$$R_{a_j,j} = \mathbf{X}_{a_j}\mathbf{B}_{a_j,j} + \gamma_{a_j,j}treat_j + \zeta_{a_j,j}R_{a_j,i} + \varepsilon_j \quad (2)$$

for a_j = age of the focal partner, and $i, j = (\underline{\text{h}}\text{usband}, \underline{\text{w}}\text{ife})$ or $i, j = (\underline{\text{w}}\text{ife}, \underline{\text{h}}\text{usband})$

$R_{a_j,k}$ is a binary variable, taking the value 1 if individual $k = i, j$ is retired in the year when the focal partner is of age a. The matrix of control variables $\mathbf{X}_{a_j}$ includes access to AFP in private sector or public sector, income[11] at age 60, wealth, education and birth cohort for both individuals, as well as a constant

[10]Note that due to the endogeneity of $R_{aj,i}$ in (2), the predicted values from (1) are used in place of $R_{a_j,i}$ when estimating the parameters in (2). This means that $treat_i$ is the instrumental variable for $R_{a_j,i}$. I omit a running indicator for the population for notational simplification.

[11]Income is always set to the year when the focal partner is of age 60, since using the same age as the left-hand side R-variable would result in a selection bias.

term, age difference and current year. The timing is set to the calendar year when the focal partner is aged a_j, which implies that the age of the spouse can vary across the couples, and the age of the spouse therefore enters as a control variable through age difference. The variable *treat* takes the values:

$$treat_k = \begin{cases} 1 \quad \text{if} \quad cohort_k \geq 1949 \text{ and } AFP_{private,k} = 1 \\ 0 \qquad\qquad\qquad\qquad \text{otherwise} \end{cases} \tag{3}$$

$$\text{for } k = i, j$$

4.1 Instrument validity

Note that an individual can never have AFP both in private and public sector, but can have neither, as explained in Section 2. The key identifying assumption is that the treatment variable is a a valid instrument for the retirement status of the spouse. Clearly, the spouse receiving increased work incentives will affect the retirement status of the spouse (instrument relevance). However, changing the work incentives of the spouse does not directly affect the focal partner, only indirectly through the spouse. As discussed, the indirect effect is composed of two effects with opposite signs; the income effect and the complementarity effect of couples' leisure. This instrument is not able to discriminate between the two, and therefore the interpretations of the results must be a lower-bound estimate of complementarity. Since the two effects are opposing, a positive sign means that the complementarity effect is stronger than the income effect, but we cannot separate the two, and hence the estimate is a lower-bound. As long as we interpret both the income and complementarity effect as a part of the spillover effect, the treatment variable does not belong to the second stage regression (instrument excludability).

Finally, the individuals held no power over choosing whether to be targeted by the reform or not (instrument exogeneity), since changing work place just before retirement does not result in the possibility to claim AFP as discussed in Section 3.

4.2 Example

Suppose as an example that I am interested in the effect on women of their husbands' labor market status when the women are 63 years old, hence she is the focal partner (individual j) and he is the spouse (individual i). I then estimate the first-stage regression:

$$R_{63,h} = \mathbf{X}_{63}\mathbf{B}_{63,h} + \gamma_{63,h}treat_h + \xi_{63,h}treat_w + \varepsilon_h$$

I obtain the predicted values $\hat{R}_{63,h}$ from this first-stage, which I then insert in the the second-stage regression:

$$R_{63,w} = \mathbf{X}_{63}\mathbf{B}_{63,w} + \gamma_{63,w}treat_w + \zeta_{63,w}\hat{R}_{63,h} + \varepsilon^*_w$$

Then I obtain estimated parameter values for all the coefficients in the second stage, and in particular, the estimated parameter $\hat{\zeta}_{63,w}$ would then be interpreted as the implied percentage change in probability of the wife retiring at age 63 if her husband is retired in that calendar year.

4.3 Common trend

A key identifying assumption for difference-in-differences is that the treatment and control groups have parallel trends prior to the policy change. Figure 4 shows plots for the common trend for both the direct effect and the indirect effect through

spouses. We should expect that all groups are similar in their trend before the reform (prior to 2011). Another method to investigate the common trend assumption, is to run a regression which is reported in Table 4 and explained thoroughly below. The treatment variable in the case of common trend is:

$$T_k = \begin{cases} 1 & \text{if} \quad AFP_{private,k} = 1 \text{ or } AFP_{private,-k} = 1 \text{ or both} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$$\text{for } k = i, j$$

This means that if either spouse has private sector AFP, then in order to check for common trend an individual in that couple is considered treated. We should then expect that year-dummies interacted with this treatment variable should be insignificant pre-reform and significant post-reform when I run the regression:

$$E_{k,t} = \mathbf{X_k} + \delta_t + \mathbf{D_{k,t-v-1}B_{v-1}} + \mathbf{D_{k,t+q}B_q} + T_k + \varepsilon_{i,t} \tag{5}$$

Where $E_{k,t}$ is the employment share at time t, $\mathbf{X}$ is the matrix of individual fixed effects, δ_t is time fixed effects, $\mathbf{D_{k,t-v-1}}$ is the interaction between the treatment and time dummies for the $v-1$ pre-treatment periods and $\mathbf{D_{k,t+q}}$ is the same for the q post-treatment periods.

Table 5 shows that the common trend holds for both men and women[12].

[12]Note that D_{2011} is insignificant for women and has the wrong sign for men. This can be explained in two ways. Those aged 63 in 2011 are the 1948 cohort, for which there were transition rules with choices between the old system and the new system. This can explain why the response among those individuals is small. Another explanation is that reforms involves a period of "learning", causing short-run elasticities to be smaller than long-run elasticities.

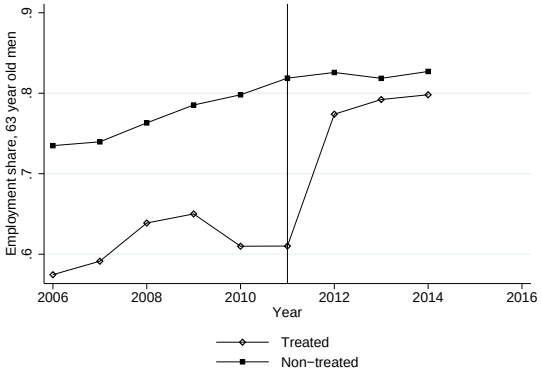
Table 5: Common trend, 63 year old individuals

	Variable	Men	Women
Pre	D_{2007}	.0068	.0024
		(.0107)	(.0117)
	D_{2008}	.0187	.0080
		(.0106)	(.0114)
	D_{2009}	.0129	.0143
		(.0102)	(.0110)
	D_{2010}	−.0079	.0080
		(.0104)	(.0109)
Post	D_{2011}	−.0214**	.0109
		(.0104)	(.0109)
	D_{2012}	.0721***	.0477***
		(.0106)	(.0109)
	D_{2013}	.0855***	.0426***
		(.0108)	(.0109)
	D_{2014}	.0737***	.0440***
		(.0110)	(.0110)
Controls			
Time FE		✓	✓
Individual FE		✓	✓
No. of obs.		87,289	90,834

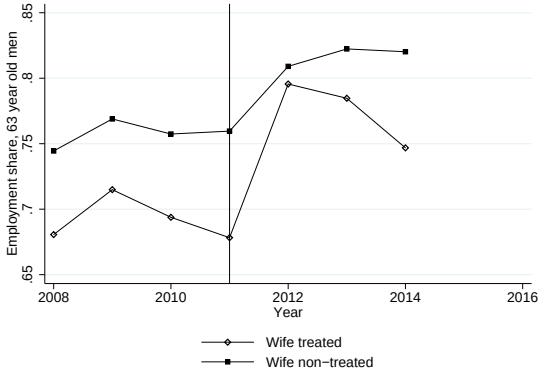
*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

Figure 4: Common trend, employment among 63 year olds

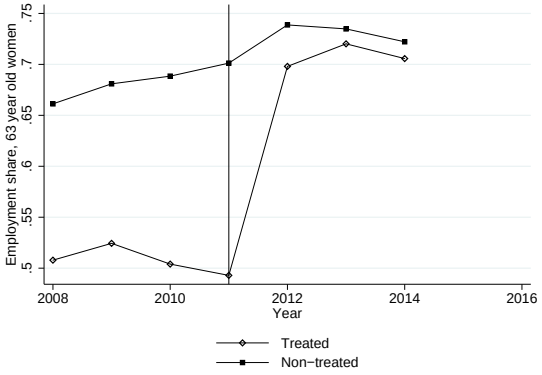
(a) Direct effect on men



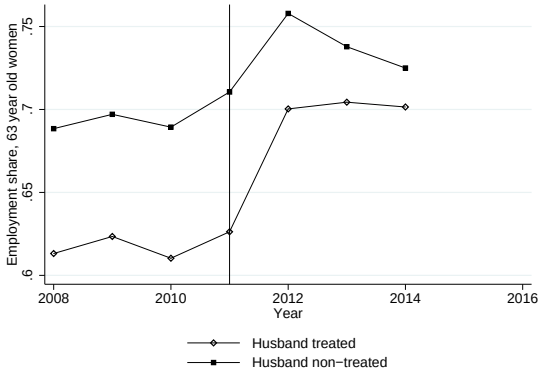
(b) Indirect effect on men



(c) Direct effect on women



(d) Indirect effect on women



The common trend visually seems to hold for all these groups. In particular, the pre-trends are visually parallel in all four groups, with the possible exception of panel a and c. However, the pre-trend differences are at the very least non-shrinking, and therefore does not exhibit any pre-adjustment to the treatment. The indirect effect on men seems to diverge in the years after the reform, which may suggest that the indirect treatment effect from women to men is fading out quickly. The indirect effect on women show a significant gap closing between the two graphs post-reform, and it seems to stabilize rather than diverge in the same

way as the indirect effect on men.

5 Results

The econometric model is solved for three different specifications of age-composition for both women and men as the focal partner. Sample 1 includes every couple, which means that no restriction is put on the age of the spouse. Sample 2 uses the sub-sample of couples where the spouse is the oldest, while sample 3 uses the sub-sample of same-aged couples. The reasoning behind solving for different age compositions is the fact that the average age difference between men and women in Norwegian couples is systematically biased towards older men. On average, men are about 2.6 years older than their wives, which could mean that on average women tend to "follow what their husband chose" because he became eligible for retirement first. Solving for these different age compositions attempts to deal with that issue. When the spouses are both the same age, we completely internalize this effect in the results. When we focus on "spouse older", we isolate the effect on those with an older spouse in order to see if this group is significantly different from the other couples.

Table 6: Overview of samples

(1)	**(2)**	**(3)**
Fix focal partner age at a_j		
All spouses	Spouse older	Spouse same age

5.1 Effect of men on their wives

First I present the results for the women, that is when the women are the focal partner. Table 7 shows the three sample selections when the wife is the focal partner and the husband is the spouse, and I fix the age of the wives at 63 years. Since the first stages[13] are significant for all the different sample selections, we can interpret the coefficients in these tables directly as the implied change in probability measured in percentage points of retiring if the spouse has retired.

When the spouse is retired, a 63 year old female has a 21.5 percentage points higher probability of retiring. The direct own-treatment effect of the reform is large in magnitude. A 63 year old woman who is affected by the reform has a 14.7 percentage points lower probability of retiring compared to the pre-reform system. The aggregated effect is found by multiplying the spillover effect with the number of women with a treated husband, add direct effect multiplied with the number of treated women and divide by the total number of women. The implied aggregate effect on labor supply among 63 year old women is a reduction in the probability of retirement of 3 percentage points. If we ignore the spillover-effects, the aggregate effect is 2.2 percentage points, which is only 75 percent of the implied aggregate effect when spousal spillovers are accounted for.

Comparing sample 1 to sample 2, we see that the coefficient on spousal retirement status is quite stable[14], and the small increase in magnitude is expected since the spouse is always older in sample 2. When the spouse is older, one should expect the spillovers to be larger in magnitude since they reach retirement age first and thus is the "first mover".

[13] For results of the first stage regressions, see Appendix A

[14] However, for sample 3 the coefficient is insignificant. This is likely due to selection bias and a small treatment group.

Table 7: Wife as focal partner when the fixed age is 63

Variable	**(1)**	**(2)**	**(3)**
Spouse retired	.215***	.244***	.101
	(.079)	(.087)	(.118)
Own treatment	−.147***	−.145***	−.140***
	(.009)	(.012)	(.026)
Controls			
Year	✓	✓	✓
Cohort	✓	✓	✓
Income	✓	✓	✓
Wealth	✓	✓	✓
Education	✓	✓	✓
No. of obs.	93,998	78,152	12,291

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

The implication of the results in Table 7 is that when a husband's work incentives are increased, the effect should carry over to his wife. This is evidence of complementarity in leisure, as the wife then postpones her retirement if the husband postpones his retirement. The results also show that 25 percent of the aggregate effect of womens' response to the reform is in fact through their husbands. Table 8 shows the same analysis when the fixed age is 64 years. For 64 year old women, the spillover effect is larger in magnitude than for 63 year old women and the direct treatment effect smaller. The implied aggregate effect on all 64 year old women is a reduction in the probability of retiring of 2.8 percentage points. Ignoring spousal spillovers yields an aggregated effect of 2.1 percentage points. Thus, also for 64 year old women, 25 percent of the aggregate reform effect comes through spousal spillovers.

Table 8: Wife as focal partner when the fixed age is 64

Variable	**(4)**	**(5)**	**(6)**
Spouse retired	.232***	.311***	.167
	(.087)	(.096)	(.122)
Own treatment	−.134***	−.130***	−.137***
	(.009)	(.010)	(.027)
Controls			
Year	✓	✓	✓
Cohort	✓	✓	✓
Income	✓	✓	✓
Wealth	✓	✓	✓
Education	✓	✓	✓
No. of obs.	92,860	77,147	12,179

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

5.2 Effect of women on their husbands

Second I present the results for the men, that is when the men are the focal partner. Similar to the previous section, I fix the age of the men to 63 years and estimate using the three different sample selections. Table 9 shows the results.

Table 9: Husband as focal partner when the fixed age is 63

Variable	**(1)**	**(2)**	**(3)**
Spouse retired	.004	.108	−.153
	(.185)	(.141)	(.186)
Own treatment	−.139***	−.146***	−.167***
	(.006)	(.012)	(.019)
Controls			
Year	✓	✓	✓
Cohort	✓	✓	✓
Income	✓	✓	✓
Wealth	✓	✓	✓
Education	✓	✓	✓
No. of obs.	98,676	27,104	12,447

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

From Table 9 we see that there are no significant effect on husbands when their

wife has retired, in large contrast to the results in Table 7 which showed the effect of men on women. However, the own-treatment effect is quite large and significant, and similar to the direct effect on women in magnitude. The estimation in sample 1 shows that a 63 year old male individual who is affected by the reform will have a 13.9 percentage points lower probability of retiring compared to the pre-reform system. The implied aggregate effect on labor supply among 63 year old men is found by aggregating across all individuals. This yields a reduction in the probability of retirement among 63 year old men of 4 percentage points, which is larger in magnitude than the effect on women, even when we account for spillover-effects. This is not surprising since a larger fraction of men are directly affected by the reform, as discussed in Section 2.

For sample 1 and sample 2 the coefficients also change in magnitude, however interpreting these coefficients is of negligible importance due to their statistical insignificance[15].

Table 10 shows the same analysis when the age of the men is fixed at 64 years. The spillovers are still insignificant, and the direct reform effect on men is larger among 64 year olds than 63 year olds.

5.3 Robustness

To further investigate, I conduct several robustness checks and analyses on different sub-samples. Since the spillover-effects are asymmetric, I check the robustness of the estimates where the women are the focal partner individuals. First, Table

[15] A possible interpretation of this is that women with older men have a preference towards "caring" for their older spouse. Another interpretation is that the oldest spouse has a "first-mover advantage" when it comes to the joint decision of when to retire. Thus if the oldest spouse retires, it will increase the likelihood of the youngest spouse retiring too.

Table 10: Husband as focal partner when the fixed age is 64

Variable	(1)	(2)	(3)
Spouse retired	−.020	−.048	−.345
	(.149)	(.190)	(.236)
Own treatment	−.151***	−.154***	.178***
	(.007)	(.015)	(.024)
Controls			
Year	✓	✓	✓
Cohort	✓	✓	✓
Income	✓	✓	✓
Wealth	✓	✓	✓
Education	✓	✓	✓
No. of obs.	97,824	26,814	12,333

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

11 estimates sample 1 with women as the focal partner at fixed age 63, when excluding some of the controls. As is evident from Table 11, the results are robust to omitting controls. The coefficients change in magnitude as is expected, but they remain significant.

Table 11: Robustness check for sample 1 (all spouses) with wife as focal partner when the fixed age is 63, excluding controls

Variable	(i)	(ii)	(iii)	(iv)	(v)
Spouse retired	.215***	.208***	.211***	.165**	.153**
	(.079)	(.081)	(.077)	(.066)	(.068)
Own treatment	−.147***	−.147***	−.144***	−.147***	−.141***
	(.009)	(.009)	(.009)	(.009)	(.009)
Controls					
Year	✓	✓	✓	✓	✓
Cohort	✓	✓	✓	✓	✓
Income	✓	✓	✓		
Wealth	✓	✓		✓	
Education	✓		✓	✓	

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

A possible caveat of the treatment variable that is used, is the possibility of two-

way causality in the couples where both spouses have access to private AFP and are born in the affected cohorts. For those couples, the endogeneity problem is possibly not solved by the use of the instrument, and this may cause computational problems. To address this issue, Table 12 estimates the samples using wife as focal partner at age 63 when dropping those couples where both work in private sector with AFP. If both work in private sector with AFP, we cannot argue that the instrument captures exogenous variation that is different from R. As we can see in Table 12, the results are still positive and significant, except for the very conservative measure in sample 3 (which still suffers from small treatment group as in the main results). Nevertheless, the magnitudes of the coefficients are not substantially different from those of Table 7, which is a strong indication of robustness. Sample 1 and sample 2 show strong evidence of robustness with regards to the potential two-way causality problem in Table 7, since the coefficients are still significant and do not change in magnitude.

Table 12: Robustness check for the three sample selections with wife as focal partner when the fixed age is 63, two-way causality check

Variable	**(1)**	**(2)**	**(3)**
Spouse retired	.212**	.221**	.096
	(.087)	(.095)	(.114)
Own treatment	−.139***	−.139***	−.139***
	(.011)	(.012)	(.034)
Controls			
Year	✓	✓	✓
Cohort	✓	✓	✓
Income	✓	✓	✓
Wealth	✓	✓	✓
Education	✓	✓	✓
No. of obs.	89,628	74,470	11,663

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

Furthermore, I want to look at different socio-economic subgroups such as couples with large differences in income levels, different compositions of education level and large differences in age. Since the spillover-effects are significant only on women, I focus on sub-sample analyses using women as the focal partner with the age fixed at 63 years.

Table 13: Education groups (Wife/Husband) for sample 1 (all spouses) with wife as focal partner when the fixed age is 63

Variable	**High/High**	**High/Low**	**Low/High**	**Low/Low**
Spouse retired	.404**	.168	.209	.245**
	(.177)	(.186)	(.312)	(.122)
Own treatment	−.128***	−.103**	−.133***	−.148***
	(.032)	(.048)	(.022)	(.011)
Controls				
Year	✓	✓	✓	✓
Cohort	✓	✓	✓	✓
Income	✓	✓	✓	✓
Wealth	✓	✓	✓	✓
No. of obs.	20,701	9,817	13,560	49,920

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

In Table 13, the we see that the spillover-effects are stronger in couples with the same education level. In fact, in couples with different education levels the effect is insignificant. Some of this might be due to small sample of treated individuals. Couples with similar education levels thus seem to plan more together or simply enjoy time spent together to a higher degree, whereas in couples with different education levels the couple have decoupled the retirement timing of the two individuals. The effect is also very large in magnitude among couples where both have high education. A possible interpretation of this is that couples with similar levels of education are simply more homogeneous to one another, and therefore share tastes for leisure activities. The couples with different education levels can

be more heterogeneous in their leisure preferences.

Next, I look at different income levels, in particular large differences in income between the husband and wife. I define a high income earner as 60 year old earning more than the average earnings in their own cohort and a low income earner as those earning less than the average earnings in their own cohort. We can then look at four different compositions of couples, namely two high income earners, two low income earners and the two groups with one high and one low income earner. As we see from Table 14, the group with two high-income individuals has significant spillover effect. This is then also a homogeneous group in terms of composition of individuals in the couples, which supports the interpretation of Table 13. However, the low-low-income group has an insignificant spillover effect. This leads to an interpretation of double-correlation in spousal spillovers; first, correlation between education levels show that homogeneous couples enjoy leisure together while heterogeneous couples do not. Second, the household income has to support leisure activities. A possible interpretation of the low/high-group could be that when the husband is earning above the average, that is sufficient to support joint retirement for the couple. However, when the wife earns above average it is not sufficient to support joint retirement because women earn less than men on average.

Table 14: Income groups groups (Wife/Husband) for sample 1 (all spouses) with wife as focal partner when the fixed age is 63

Variable	High/High	High/Low	Low/High	Low/Low
Spouse retired	.416**	−.138	.387**	.075
	(.177)	(.609)	(.162)	(.307)
Own treatment	−.119***	−.142*	−.159**	−.176**
	(.017)	(.020)	(.021)	(.016)
Controls				
Year	✓	✓	✓	✓
Cohort	✓	✓	✓	✓
Education	✓	✓	✓	✓
Wealth	✓	✓	✓	✓
No. of obs.	23,069	18,218	21,781	30,930

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

Finally, I look specifically at the public sector AFP and the group without access to AFP among the women[16]. The sum of these two groups is equivalent to the estimation in Table 12. We might believe that public sector workers with AFP are different from the workers without access to AFP, and hence I should check if the spillover effects are different in those groups. Note that these groups will never have an own-treatment effect since they by definition can never have AFP in private sector. Table 15 shows the results. We see that the coefficient is significant for the women with access to public AFP, but not for those without access to any AFP. This also supports the interpretation from Table 13 and Table 14 which suggested that more homogeneous couples enjoy leisure together more, since the group without access to any AFP will have the most heterogeneous spouses. The couples where both have access to a form of AFP are possibly more homogeneous with respect to preferences for leisure, and thus we should expect the spillover

[16]I cannot consistently estimate on a subsample of only those with access to private AFP due to the potential two-way causality. There would be no couples left with a different variation in the instrument than the endogenous variable, and hence the estimation would be inconsistent even in large samples. Thus, we can only look at the sub-samples of public sector with AFP and the group without AFP.

effects to be stronger among those spouses.

Table 15: Women in public sector with AFP and women without access to AFP for sample 4 with wife as focal partner when the fixed age is 63

Variable	**Public AFP**	**No AFP**
Spouse retired	.301**	.092
	(.120)	(.125)
Controls		
Year	✓	✓
Cohort	✓	✓
Income	✓	✓
Wealth	✓	✓
Education	✓	✓
No. of obs.	56,010	23,964

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

6 Concluding remarks

This paper has studied the importance of spillover-effects in determining labor market effects of reforms targeting work-incentives. It is evident that the reform had a positive effect on labor participation among the elderly for both women and men. However, incentive effects are stronger and able to explain a larger part of the labor market responses when we account for spousal spillover effects.

Spousal spillovers seem to be a non-negligible driving force behind retirement choices among women in Norway. An approximate 21.5 percentage points increase in the probability of a 63 year old women to retire if the husband has retired is a substantial implication. As much as 25 percent of the aggregate effect of the reform on 63 year old women comes from spillover effects from their husbands, which implies that aggregating single-individual behavior without regarding spillover effects may result in underestimated aggregate effects. This effect may also explain why

there are changes in labor participation in sectors that were subject to no change in work incentives after the reform, in particular among public sector workers.

This estimate is based on a clear-cut identification strategy using highly detailed register data with high precision. In this quasi-experimental setting, I provide strong evidence of spousal spillovers from men to their wives with clear causal interpretation. However, the study does not find evidence of men responding to their wives incentives and thus show that spousal spillovers are asymmetric. The effects are robust to various tests. The effect is larger if the spouses are more homogeneous in their composition, with regards to income, education and choice of work place. This supports that preferences towards joint leisure is a driving force of joint retirement.

There are several implications of these results. First, the utility that wives draw from their own leisure seems to be non-separable from their husbands, which means that whenever the husband increases leisure the marginal utility of leisure for wives increase. This is a sign of complementarity from the viewpoint of the wives, and is a lower-bound estimate since I cannot separate it from the income effect. However, the utility that husbands draw from their own leisure seems to be separable from their wives, or at least the income and complementarity effect cancel out from the viewpoint of the husbands. This has potentially large distributional effects in a structural model based on this finding. An additional effect is that the complementarity has to be linked to the homogeneity of the spouses, so that the non-separability is dependent on the characteristics of the spouses being similar. Third, this finding has implications for how we understand the importance of incentives. Since this suggests that incentives are interlinked, it is not straight forward to account for the aggregate implications of a change to individual in-

centives. One has to take into consideration that changing incentives for some individuals will have effects on individuals that have no direct impact from such a change. The total aggregate effect of a reform targeting work incentives, at least among elderly, is likely to be underestimated in computations based on individual decision making rather than joint decision making, since those studies are unable to account for a substantial spousal spillover effect.

References

Blau, D.M., 1998. Labor force dynamics of older married couples. *Journal of Labor Economics*, 16(3), 595-629

Blundell, R., Pistaferri, L., Saporta-Eksten, I., 2015. Children, time allocation and consumption insurance. IFS Working Paper, March 2015

Bratsberg, B., Stancanelli, E., 2017. Partial retirement and partners' labor supply: Learning from a Norwegian retirement reform. Preliminary draft, The Frisch Center, Oslo, Oct 31 2017

Casanova, M., 201. Happy Together: A Structural Model of Couples' Joint Retirement Choices. UCLA Working Paper

Fredriksen, D., Holmøy, E., Strøm, B., Stølen, N.M., 2015. Fiscal effects of the Norwegian pension reform - A micro-macro assessment. *Statistics Norway*, Discussion Papers no. 821

Gerard, F., Nekby, L., 2012. Spousal Joint Retirement: A Reform Based Approach to Identifying Spillover Effects. The Institute for Labor Market Policy Evaluation, Working Paper No. 2010: 126

Hernæs, E., Markussen, S., Piggott, J., Røed, K., 2016. Pension reform and labor supply. *Journal of Public Economics*, 142, 39-55

Hurd, M.D., 199. The Joint Retirement Decision of Husbands and Wives. Issues in the Economics of Aging, Wise, D.A. (ed.), *University of Chicago Press*, 231-258

Johnsen, Julian V., Vaage, Kjell, 2017. Spouses' retirement and the take-up of disability pension, Working Papers in Economics 03/15, University of Bergen, Department of Economics

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

Mincer, J., 1962. Labor force participation of married women: A study of labor supply, in *Aspects of labor economics*, *Princeton University Press*, 63-105

OECD, 2017. Pensions at Glance 2017: OECD and G20 Indicators, *OECD Publishing*, Paris

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

Stancanelli, E., Van Soest, A., 2012b. Joint Leisure Before and After Retirement: A Double Regression Discontinuity Approach. IZA Discussion Papers 6698, Institute for the Study of Labor (IZA)

Appendix A - First-stage results

Table 16: First-stage regression results

Variable	(1)	(2)	(3)
Table 9			
Own treat.	−.067***	−.130***	−.133***
	(.011)	(.020)	(.026)
Table 7			
Own treat.	−.124***	−.130***	−.164***
	(.009)	(.011)	(.018)
Table 10			
Own treat.	−.079***	−.109***	−.125***
	(.010)	(.021)	(.027)
Table 8			
Own treat.	−.117***	−.128***	−.167***
	(.010)	(.012)	(.020)
Table 12			
Own treat.	−.122***	−.129***	−.161***
	(.010)	(.012)	(.020)

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

Table 17: First-stage regression results for **Table 11**

Variable	(i)	(ii)	(iii)	(iv)	(v)
Own treat.	−.124***	−.121***	−.126***	−.148***	−.146***
	(.009)	(.009)	(.009)	(.009)	(.009)

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

Table 18: First-stage regression results, sub-samples

Variable	High/High	High/Low	Low/High	Low/Low
Table 13				
Own treat.	−.090***	−.151	−.090***	−.109***
	(.021)	(.027)	(.024)	(.012)
Table 14				
Own treat.	−.088***	−.042	−.113***	−.069***
	(.015)	(.028)	(.014)	(.020)

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.

Table 19: First-stage regression results for **Table 15**

Variable	**Public AFP**	**No AFP**
Own treat.	$-.072^{***}$	$-.141^{***}$
	(.010)	(.018)

*** p<.01, ** p<.05, * p<.1. Robust standard errors in parenthesis.