Joint Retirement in Couples*

Evidence of Complementarity in Leisure

Herman Kruse†

June 2020

Abstract

This paper examines leisure complementarities in the joint retirement decisions of couples by quantifying the effect of the 2011 Norwegian pension reform. The reform abolished an earnings test on early retirement benefits for private sector workers, but not for public sector workers. I analyze population-wide registry data on labor market participation and consider couples where the focal partner works in the public sector with a spouse either employed in the private sector (treatment group) or the public sector (control group). I find that spousal spillovers account for over 40 percent of the aggregate employment effect for women. They claim less early retirement benefits, and both men and women pay additional labor income taxes. In total, government budgets improved by approximately $46 million due to the spousal spillover effect between 2012–2015.

Keywords: Policy evaluation, Spousal spillovers
JEL classifications: J14, J26, C36, D04, H31

*I acknowledge the Frisch Center for data access. All data belongs to Statistics Norway and are used in compliance with the rules given by the Norwegian Data Inspectorate. I am grateful for the financial support from the Norwegian Research Council (grant 248-868), and NORAM for travel funds. Excellent comments have been given by Simen Markussen, Kjetil Storesletten, Andreas Myhre and Sigurd M. Galaasen, and from seminar participants at the Department of Economics, UiO. I also thank two anonymous referees for insightful and constructive comments. Finally, thanks to Victoria Pålsson for proof-reading.

†Contact information: herman.kruse@econ.uio.no, University of Oslo, Department of Economics, 0851 Oslo, Norway.

I Introduction

Throughout the OECD, countries are reforming their social security systems to accommodate an aging population and increased fiscal pressure (OECD, 2017, 2019). Most countries are debating prescriptions such as increasing the minimum legal retirement age or cutting benefits. In Norway, emphasis has been put on linking the retirement decision to work incentives, by decoupling the link between exiting the labor market and take-up of benefits. At the same time, there is a long anticipated, yet not well documented trend among the workers now approaching retirement ages; households more often consist of two individuals with two full-length careers and two retirement decisions, as women's labor market participation rates dramatically increased in the 1970s among the women born in the late 40s and early 50s. Before, the retirement decision was usually an individual decision that the husband in the household made, as the wife did not have a full-length career, or often times not a career in the labor market. Therefore, investigation of the spousal dimension and the presumably strong link between the spouses' leisure spending, is important for understanding aggregate policy effects.

In this paper, I investigate and quantify the spousal spillover effects in retirement decisions. In Norway, the pension system was reformed in 2011 to increase work incentives among the elderly. The most radical change in the reform was the removal of a confiscatory earnings test on early retirement benefits for a subset of workers; private sector workers with access to an early retirement scheme. Other workers, that is, public sector workers and private sector workers without an early retirement scheme, were largely unaffected by the reform in terms of work incentives, and were not at all affected by the removal of the earnings test. More details will be presented in Section II below. The implementation of the reform was a swift process and affected a subset of workers, and thus allows for a quasi-experimental approach to identifying work incentives among the elderly. By linking spouses through comprehensive register data, I am able to identify spousal spillover effects in retirement decisions.

The narrowing of the gender-gap in labor force participation leads to questioning whether spouses have a tendency to influence each other's retirement decisions due to

complementarity of leisure. In fact, the gender gap in labor force participation among workers in their 50s has narrowed from approximately 60 percent in the 1970s to 85 percent in the 1980s, and women likely surpassing men within few years.[1] Thus, if spousal spillovers in retirement are quantitatively important, it will influence the aggregate labor force participation effects of policy changes. Using comprehensive Norwegian register data, I find that women tend to be influenced by their husbands' retirement decisions, while the evidence is inconclusive for husbands. The aggregate effects of spousal spillovers are important for women, and account for over 40 percent of women's aggregate increase in labor force participation due the aforementioned policy change in the Norwegian pension system, which will be described in more detail below.

In the analysis, I use the instrumental variable (IV) method in a two-stage least squares (2SLS) setup. The important policy change that I exploit is the removal of a confiscatory earnings test on continued work after claiming early retirement benefits in the private sector, hereafter denoted using its acronym "AFP" (from Norwegian "Avtalefestet pensjon"). The earnings test, which was common to both private and public sectors before the reform, implied that if a person who claimed benefits continued to work, the benefits would be cut *pro rata*, essentially imposing a marginal tax rate of close to one, against the earnings. This imposed a major disincentive to continue working past the minimum eligibility age for AFP, which was (and still is) 62 years of age. In a nutshell, to quantify the effect of leisure complementarity, I measure the labor supply response of individuals who themselves worked in the public sector (and therefore were not directly treated by the reform themselves), who had a spouse working in the private sector with AFP (the treatment group). I compare these couples to those where the spouse works in the public sector (the control group). Since the policy change, as will be described below, greatly increased work incentives for workers in the private sector with AFP and not for other workers, this approach identifies how the labor supply response of the treated partner affects his or her spouse.

The IV approach will give the local average treatment effect (LATE), the average effect

[1]Defined as the ratio of employed women to men. Source: Statistics Norway, open source statistics.

of the spouse postponing retirement on the employment of the focal spouse (Imbens and Angrist, 1994). An important advantage of this approach is the possibility of aggregating the household employment effects, which allows for macro-relevant policy evaluations.

I find that when the husband postpones retirement, his wife is approximately 16 percentage points more likely to postpone labor market exit between ages 63–66. The largest effect occurs when the wife is 64 years old, while the response is stronger at 65 years of age than at 63 years of age. In total, spousal spillovers account for more than 40 percent of the aggregate labor supply effect on elderly women as a result of this reform. The regression evidence is imprecisely estimated when focusing on spillovers from women to their husbands. However, if we ignore the imprecision, the point estimate is still positive, but somewhat smaller in magnitude than for women, suggesting that spillovers may at least be of less importance to men than women. Moreover, if we aggregate the spillover effect onto men (even though it is imprecisely estimated), spillovers from the wives account for less than 5 percent of the aggregated employment effect on men. There is little evidence for wealth or income to be determinant factors, but there are some indications that poorly educated women are more likely to be influenced by their husbands than highly educated women.

Finally, government budgets improved by approximately $46 million only due to spousal spillover effects onto women, through lower take-up of AFP benefits and additional taxes paid between the years 2012–2015.[2] The public sector wives aged 63–66 married to private sector husbands on average decreased their take-up of AFP benefits by $18,100 and paid approximately $3,850 in additional taxes on labor income as a result of postponing retirement. Additionally, men aged 63–66 paid approximately $11,000 in additional taxes on labor income if their private sector wives postponed retirement in response to the reform, which may indicate an intensive margin response by men, as there is no clear indication of an extensive margin response.

[2]In comparison, the total public expenditure on public old-age pension in the period increased from approximately $20 billion in 2012 to approximately $24 billion in 2015. Source: The Norwegian Labour and Welfare Administration, open source statistics.

Related literature

The paper relates to the growing literature on joint retirement, which can broadly be categorized into structural and empirical studies, the latter closest related to this paper. The literature emerged from Hurd (1990) who reported significant correlation of the timing of retirement between spouses, which he attributed to complementarity in leisure. Later, Blau (1998) studied the labor force dynamics of older married couples in the US using a structural model, using data from the Health and Retirement Survey (HRS). He finds significant spousal spillover effects; men are more likely to exit the labor market if the wife is unemployed, but he does not find the opposite. However, he cannot rule out that this is due to women being observed at younger ages on average, and thus may not be eligible for pension benefits.

Gustman and Steinmeier (2000, 2004) develop structural models of joint retirement. Gustman and Steinmeier (2000) find that interdependence of spouses comes through preferences rather than coordination of retirement. Gustman and Steinmeier (2004) quantify how much each spouse values retirement with the other. They find that husbands are more influenced by having a retired wife, than women are of having a retired husband. Zweimüller et al. (1996) and Coile (2004) find similar asymmetry in their reduced form studies, using Austrian data and the HRS respectively. Casanova (2010) presents a structural model of joint retirement, where leisure complementarities are positive and significant for both spouses. However, this list of papers lack an experimental design, which gives reason to question the causality. If there is significant sorting of spouses into firms or assortative mating, the evidence provided in these contributions could in principle be misleading, and may reflect correlations.

Recent contributions are Stancanelli and van Soest (2012a,b), using a double regression discontinuity design and institutional variation in eligibility to identify effects of individuals' own pension eligibility on their partners' work incentives, they 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. Lalive and Parrotta (2017) use a similar empirical strategy on Swiss register data, and find that partner pension eligibility matters for labor force exit for women, but no significant effect for men. Neither of these papers use a reform-based approach. Gerard and Nekby (2012) use a Swedish pension reform from 2001 to conduct a triple-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. Selin (2012) also studies a Swedish pension reform which increased incentives to work longer, and finds no evidence of spousal spillovers.

The Norwegian pension reform has been studied by others, for instance Hernæs et al. (2016) who focused on men and women separately, and did not investigate the couple dimension.[3] Essentially, their extensive margin estimates on labor supply are comparable to the first-stage effect in the 2SLS in this paper. They find that the extensive margin response was approximately 17 percentage points at ages 63 and 64, while I find that men and women directly affected by the reform are approximately 20 percentage points more likely to remain employed between ages 63–66. However, their paper does not account for spousal spillovers.

Lastly, one recent contribution on Norwegian data is closely related. Johnsen and Vaage (2017) study spousal spillovers in retirement in Norway using a different identification, 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 the introduction of AFP for public sector workers and centrally negotiated private sector firms. Hence, their reform is a negative work incentive change and it targets a different worker group than the reform from 2011 did. Additionally, their contribution is closer related to the benefit substitution literature; they focus on the effect of a retiring spouse on take-up of disability insurance benefits from the partner.

This paper contributes to the literature in several aspects. The main contribution is

[3]A preliminary draft by Bratsberg and Stancanelli (2017), developed independently of this paper, also studies the Norwegian 2011 reform. However, the draft is currently unavailable online.

that it provides causal estimates on interlinked incentives between spouses, which is rarely causally identifiable even in detailed register data due to lack of individual exogenous variation. Where several previous studies provide non-experimental designs on interlinked incentives, this paper makes use of a quasi-natural experiment setting in a swift and radical reform of the Norwegian pension system to causally identify work-incentives and the incentive spillovers between spouses. The reform created exogenous variation in work incentives, as described briefly above and in detail in Section II. Secondly, the paper shows that leisure complementarity is important for understanding the aggregate effects of policy changes. If we rely on models that disregard complementarity, one will typically underestimate the aggregate effect of a change in labor supply incentives. The aggregation made possible by the IV approach allows for studying the macro implications of policy changes. Lastly, access to administrative data on the entire Norwegian population provides detailed information on the individual level, which is not plagued by attrition and mismeasurement due to self-reporting known to be common problems in survey data (Røed and Raaum, 2003).[4] The details in the administrative data allow for control of potential confounders around retirement ages.

The remainder of this paper is organized as follows. Section II provides background information on the institutional settings. Then, Section III presents the data and descriptive statistics and figures. Section IV explains the empirical strategy, and Section V lays out the results. Finally, Section VI concludes.

II Institutional setting

This section gives an overview of the pension system pre- and post-reform. The pre-reform system essentially covers all workers born in 1948 or earlier, while the post-reform system covers all workers born in 1949 or later, and was implemented January 1, 2011.

[4] Although Kapteyn and Ypma (2007) show that administrative data is not unambiguously an improvement compared to survey data, in particular they argue that administrative data may introduce mismatching due to imperfect linkage information. However, I argue that the data used in this project exhibit minimal risk of such mismatching.

Pre-2011 pension system

The pre-reform system was a pay-as-you-go system consisting of three main pillars. The first pillar was the old-age guarantee pension, granted after forty years of residence and "pension-awarding income".[5] The second pillar was an income-related pension, with a mapping based on the twenty best years of income. The third pillar was a mandatory defined contribution system,[6] where the employers must make a minimum contribution of 2 percent of the earnings to a pension plan.[7] The sum of the three pillars was the individual's *old-age pension*, with a universal access age at 67.

Since its launch on January 1st 1989, approximately two-thirds of the workers (all workers in the public sector, and about half of private sector workers) had access to an early retirement option, known as AFP. To be eligible for AFP, a worker had to be employed and have a pension-awarding income level in a firm covered by an AFP scheme for three of the last five years before retirement. If eligible, the worker could draw full pension benefits from age 62, equivalent to the benefit size that would be paid from the old-age pension system when the worker turns 67, plus an additional flat rate of approximately $2,300 (in 2015 dollars) annually. However, if the worker decided to continue working past the age of 62, he or she faced a *pro rata* earnings test from the first dollar earned. The earnings test was a very high implicit marginal tax rate on continued work past the age of 62 for eligible workers, as benefits were cut one-to-one with earnings. This earnings test was common to private sector and public sector workers.

Post-2011 pension system

The most important change in the 2011 reform, was the *removal of the pro-rata earnings test* in the private sector with AFP, allowing them to freely combine continued work with claiming AFP benefits. This removal only affected this particular subset of workers, which is the variation I use to gain identification. The reform entered into force January

[5]This is defined as having income greater than $1G$ in a year, where G is a base unit set by the government for the purpose of defining this threshold. In 2019, $1G$ is $99,858$ NOK, or approximately $\$10,800$. The unit is indexed annually to match average wage inflation.

[6]It became mandatory in 2006.

[7]It was possible for the employers to choose a defined benefit scheme, however the scheme must deliver at minimum the same benefits as the expected benefits of the 2 percent defined contribution scheme.

1st, 2011. All workers born in 1949 or later were transferred into the new system, while workers born before 1949 remained in the old system.[8] After the reform, workers are facing a requirement linked to the day they turn 62 years. Essentially, workers must be employed in a private sector AFP firm, and must have been employed in either *the same firm covered by private sector AFP* for three of the last five years, or *any firm covered by private sector AFP* for all the five previous years. Therefore, there is no possibility to self-select into treatment, as working outside of a private sector AFP firm at the time of the announcement excludes the worker by default.

For the old-age pension, the universal access age was lowered from 67 years to 62 years. This gave all Norwegian workers access to a pension claim at the same age, whereas previously, workers without access to AFP could only start drawing benefits from age 67 (about half of private sector workers was in this category). However, claiming old-age pension before age 67 is now subject to an actuarial fair recalculation of benefits and life-expectancy adjustment, meaning that drawing old-age pension before the age of 67, as well as increased longevity, leads to a life-long cut in annuities.[9] The changes to the old-age pension system is common to all workers, and hence will affect private sector and public sector workers equally. Hernæs et al. (2019) showed that the actuarial recalculation and flexible withdrawal have had no effect on labor force participation, but a minor effect on high-income earners reducing their earnings by responding on the intensive margin.

The changes to the AFP system in the private sector displayed in Figure 1 imposed large incentive changes. Importantly, the pro-rata earnings test was removed, meaning that benefits are no longer cut if an eligible worker decides to work past the age of 62 while claiming AFP benefits. This is the main driving force of the reform. The effect of these changes is that the timing of retirement and the timing of benefit claiming is

[8]There were some transitional rules, in particular for the 1948 cohort. The workers born in 1948 who had not claimed AFP already were given the option to be transferred to the new system by postponing AFP claiming until after January 1st, 2011. Therefore, the 1948 cohort may be partially treated, and I follow Hernæs et al. (2016) who dropped the entire 1948 cohort.

[9]Additionally, the income-related pension is now an accumulated sum of 18.1 percent of annual income every year, up to a maximum of $7.1G$, rather than based on the 20 best years. However, this change will only partially take effect for cohorts born after 1954, and take full effect for cohorts born in 1963 or later, and is therefore not important in this paper as the affected cohorts are too young to be included here.

decoupled, as one may freely combine continued work with claiming of AFP. The AFP is now a life-long top-up of the old-age pension, and may be claimed alongside old-age pension.[10] However, workers in the public sector did not receive any changes to the AFP, which means that they still face the earnings test, and the AFP benefits can be drawn until 67 years of age with the old benefit levels. I therefore compare two groups. When investigating spillovers onto a public sector worker, say a female worker, I compare employment for the women whose husband works in the private sector with AFP (the treatment group) to those whose husband works in the public sector (the control group).

Table 1
Overview of the pension system, pre-reform and post-reform.

	PRE-2011 SYSTEM	
	Early retirement (62–67)	**Normal retirement (67 →)**
Common to private and public sector workers	AFP: Starting from age 62 and available until age 67. Annuities equal to OAP that would be claimed when retiring at age 67 plus an additional 20,400 NOK ($2,300) annual flat rate. **Earnings tested benefits.**	Old-age pension (OAP): Starting from age 67. Benefits are a mapping based on 20 best years of income, and the system is PAYG, plus a guaranteed pension level and a defined contribution by the employer.
	POST-2011 SYSTEM	
	Early retirement (62–67)	**Normal retirement (62 →)**
Public sector workers	AFP: Equivalent to the pre-reform system	Old-age pension (OAP): Starting from age 62. Can claim partial or full OAP with longevity adjustment and actuarially fair adjustment of pension levels. Optional claiming before 62 is conditional on giving up AFP benefits.
	Early retirement from age 62	
Private sector workers	Starting from age 62, workers can claim AFP benefits with lifelong annuities such that the expected value of the sum of the new AFP annuities is equal to what would be claimed between ages 62–67 in the old system. **No earnings test on benefits.** Can claim partial or full OAP from age 62 alongside AFP, with longevity adjustment and actuarially fair adjustment of pension levels.	

Notes: OAP = old age pension. Reform implementation January 1, 2011. Pre-reform cohorts are cohorts born 1948 or earlier, while post-reform cohorts are those born 1949 or later. The expected-value re-calculation of the new AFP benefit is, highly stylized, such that if B was the total sum paid over the ages 62–67 in the old system, the annuities are now $B/\mathbb{E}$(remaining lifetime).

[10]The new AFP annuities are therefore smaller than for the old AFP, but the *expected value* of the lifetime benefits was unchanged.

III Data and sample restrictions

The data is an administrative register database for all individuals in Norway for the years 2007–2015. Individuals are identified using a unique identification key, and can be linked to their spouse. It contains annual information on labor earnings, wealth, birth cohort, education (in years of non-compulsory education[11]), children and AFP eligibility in private or public sector. The selected cohorts are 1944–1952. Since the 1948 cohort is partially treated by the reform, as explained in Section II, they are removed from the estimations.[12] The removal of this cohort leaves me with four pre-treatment cohorts (1944–1947) and four post-treatment cohorts (1949–1952). Furthermore, I pre-determine employment, marriage and residential status at the time when the spouse (individual i) is age 60, for *both partners*.

Employment is defined as having a wage-income higher than $1G$ (G is approximately \$10,800 in 2019, see footnote 5) in a year, which corresponds to the minimum required income to earn "pension points" in the old system.[13]

Descriptive statistics

Tables 2 and 3 provide descriptive statistics for some key groups in the data. Firstly, it is important to notice in Table 2 that the fraction of men working in the private sector with AFP is much larger than the fraction of women. Furthermore, public sector workers are on average more highly educated than the private sector workers with AFP, while the private sector workers with AFP are on average wealthier than the public sector workers. In terms of average income levels, the two sectors are identical. On average, the private sector workers with AFP retire a little earlier than the public sector workers, even though they share the same minimum legal retirement age of 62 (using the early retirement option). Comparing the private sector workers without AFP shows that this group is

[11]High school is three years, technical college four, bachelor's degree six, master's degree eight and doctoral degree twelve years of non-compulsory schooling.

[12]The same was done in Hernæs et al. (2016) because of this partial treatment that would affect the outcome of the control group.

[13]For comparison, this level is about half of most of the social security benefit *minimum* levels; for instance temporary DI is, measured in basis points, approximately $2.25G$, old-age pension is approximately $2G$ and DI is approximately $2.4G$ at minimum. Therefore, the level is quite moderate, and it is reasonable to assume that individuals with this low income level have in fact exited the labor market.

very different from the other two; firstly, they earn more, they are significantly wealthier, and they retire later on average compared to the public sector workers. Secondly, they are also significantly more likely to exit the labor market through disability insurance (DI), and to be on the temporary DI program. This, in combination with the fact that they do not share the same minimum legal retirement age (they can only retire at 67 before the reform), are the main reasons for why this group should be excluded from the estimation (they do not serve as a proper control group).

Furthermore, inspecting the pre- and post-1949 workers in column 2 and 3 in Table 2, shows that there is little selection into treatment. If there was selection into treatment, we would expect that the private sector AFP group significantly increased in size, as they were given favorable incentives to work and combine take-up of early retirement benefits with continued work. However, we actually see that there is a slight decline in the fraction with private sector AFP (from 21.2 percent to 18.8 percent). As changing workplace just before retirement disqualifies the worker from claiming AFP benefits, selection is in any case trivially ruled out by the regulations.

We also see that there is a significant increase in the labor market participation at age 63 and 64 in the post-reform cohorts, suggesting a strong reform effect. While most of this may be due to the direct treatment effect, comparing the rows for employment at 63 and 64 in Table 3, we see that there is also evidence of a significant spillover effect onto women, as female focal partners greatly increased their labor supply if their husbands worked in the private sector with AFP. There is also a decreased take-up of DI benefits as well as a small decrease in short-term sickness benefits. However, there is a small increase in take-up of temporary DI, which is likely due to the increased labor supply, inducing some workers with poor health to postpone retirement and end up on this particular benefit. Furthermore, both men and women decrease DI take-up, increase temporary DI take-up and decrease short-term sickness leave.

Table 3 splits the sample into the groups that form the treatment and control groups in the empirical estimation. For a clear identification, we would ideally want the observable characteristics to be as equal as possible between the pre- and post-reform cohorts in the

Table 2

Descriptive statistics for cohorts 1944–1952 for the years 2007–2015. 63–66 year-olds married to a 63–66 year-old spouse, pre-determined working when the spouse is aged 60.

		All workers		All men		All women		Sectors		
	All individuals	**Born pre-1949**	**Born 1949 or later**	**Born pre-1949**	**Born 1949 or later**	**Born pre-1949**	**Born 1949 or later**	**Private sector workers**	**Public sector**	**Non-AFP private sector**
Male	49.0	54.4	41.7	67.4	33.4	55.9				
Average age	64.5	64.6	64.2	64.9	64.5	64.4	64.0	64.6	64.4	64.4
Private sector AFP	20.6	21.2	18.8	28.5	28.1	14.2	12.1			
Public sector	41.3	41.9	40.4	29.1	26.4	57.2	50.4			
Years of non-compulsory education	1.4	1.3	1.5	1.5	1.7	1.1	1.3	0.8	2.0	1.0
Number of children	2.3	2.3	2.3	2.3	2.3	2.4	2.4	2.2	2.4	2.3
Income (in G)	3.3	3.4	3.3	4.1	4.4	2.4	2.5	3.1	3.1	3.7
Wealth (in G)	19.6	18.8	20.8	25.8	31.8	10.4	12.5	19.2	12.8	27.7
Employment at age 63	72.9	68.2	79.2	68.0	80.3	68.4	78.3	67.1	73.6	75.2
Employment at age 64	63.8	58.5	70.9	59.0	72.1	57.9	70.1	55.4	63.3	68.9
Disability insurance at age 63	5.9	6.9	4.5	6.2	3.5	7.7	5.3	4.7	4.8	7.7
Disability insurance at age 64	8.1	9.3	6.5	8.0	4.8	10.8	7.7	6.2	7.2	10.1
Temporary DI at 63	1.7	1.3	2.3	0.9	1.6	1.7	2.8	1.1	1.2	2.5
Temporary DI at 64	1.8	1.3	2.6	0.9	1.8	1.7	3.2	1.2	1.3	2.7
Short-term sickness benefits at 63	0.5	0.6	0.4	0.5	0.3	0.8	0.5	0.4	0.3	0.9
Short-term sickness benefits at 64	0.5	0.6	0.5	0.5	0.3	0.7	0.6	0.4	0.3	0.9

Notes: The rows male, private sector AFP, public sector, and all of the employment, disability insurance, temporary DI and short-term sickness benefits are in percent. Average age and years of non-compulsory education are in years. Income and wealth are in G units, where one G-unit is $90,068$ NOK (approximately $10,000 using the average 2015-exchange rate) in 2015 which is the last year included.

Table 3

Treatment versus control group. Descriptive statistics for cohorts 1944–1952 for the years 2007–2015. 63–66 year-old public sector focal partners married to a 63–66 year-old spouse, pre-determined working when the spouse is aged 60.

	Pre-reform						**Post-reform**					
	Men			Women			Men			Women		
	Treatment	Control	p-value	Treatment	Control	p-value	Treatment	Control	p-value	Treatment	Control	p-value
Group size	1,858	14,720		11,324	15,642		1,183	9,771		6,333	7,932	
Years of non-compulsory education	1.8	2.9	0.000	1.1	2.1	0.000	1.9	2.9	0.000	1.2	2.3	0.000
Number of children	2.2	2.4	0.000	2.3	2.4	0.005	2.1	2.4	0.000	2.3	2.4	0.000
Income (in G)	3.1	4.2	0.000	2.1	2.8	0.000	3.8	4.7	0.000	2.5	3.2	0.000
Wealth (in G)	14.5	16.9	0.000	8.4	9.1	0.006	18.3	19.0	0.398	10.1	11.3	0.001
AFP claim (in G)	5.69	5.61	0.293	5.86	5.65	0.244	6.11	6.21	0.495	6.38	6.55	0.211
Employment at age 63	70.8	72.3	0.121	61.4	71.6	0.000	70.3	75.5	0.000	70.5	76.1	0.008
Employment at age 64	57.9	61.9	0.000	49.2	61.9	0.000	60.5	65.3	0.000	56.4	65.7	0.000
Disability insurance at age 63	5.6	5.4	0.716	7.4	4.7	0.015	5.4	4.5	0.096	4.5	3.3	0.224
Disability insurance at age 64	9.1	8.7	0.515	9.2	6.7	0.039	7.8	6.3	0.018	4.4	4.0	0.683
Temporary DI at 63	1.4	1.1	0.291	0.7	0.4	0.412	2.7	2.0	0.079	0.8	0.4	0.346
Temporary DI at 64	1.3	1.4	0.747	0	0.3	0.000	2.7	1.4	0.051	0.8	0.5	0.515
Short-term sickness benefits at 63	0.6	0.4	0.281	0.7	0.2	0.156	0.2	0.2	0.789	0.2	0.05	0.333
Short-term sickness benefits at 64	0.3	0.5	0.259	0.3	0.2	0.534	0.3	0.3	0.844	0	0.2	0.009

Notes: The statistics are for the focal partner. The columns named "Pre-reform" means the spouse is born in the pre-treatment cohorts (1948 or earlier), while "Post-reform" means the spouse is born in the post-treatment cohorts (1949 or later). The column "Treatment" indicates that the spouse works in the private sector, while "Control" indicates the spouse works in the public sector. The rows male, private sector AFP, public sector, and all of the employment, disability insurance, temporary DI and short-term sickness benefits are in percent. Average age and years of non-compulsory education are in years. Income, wealth and AFP claim are in G units, where one G-unit is 90,068 NOK (approximately \$10,000 using the average 2015-exchange rate) in 2015 which is the last year included. AFP claim is the average of the *total potential AFP claim per individual* in the group at the time of retirement. As explained in Section II, AFP can only be claimed between ages 62–66 in the public sector, so the total AFP claim is the sum of the annual benefit multiplied by the number of years of claiming, maximum 5 years. p-value displays the test if treatment and control statistics are significantly different.

different groups, or at least that the treatment and control groups share a common trend in observables. From inspecting educational attainment, children, income and wealth, we see that the groups are fairly stable across the pre- and post-reform cohorts, and at least the observables for the treatment and control groups tend to move in the same direction. However, we do see a clear tendency that observables are different between treatment and control (i.e. they have low p-values). This is, however, not an issue for identification as they are differenced out in the empirical strategy, as the underlying reduced form is difference-in-differences.

Next, we see that employment at age 63 and 64 for male focal partners married to a wife in the private sector with AFP, is almost exactly the same between the pre- and post-reform cohorts, and that the difference between treatment and control is quite small. For female focal partners married to a husband in the private sector with AFP, employment at 63 and 64 significantly increases between the pre- and post-reform cohorts. This suggests that female public sector workers in particular experienced spillover from their husbands.

AFP claim is the average of the *total potential AFP claim per individual*, that is, the entire sum of all annual benefits that the person can receive, given the retirement timing. The average annual benefit for a person is, in basis points, $2.2G$. Say a person with an average annual benefit claims AFP at age 64, then the total AFP claim is $2.2G \cdot (67 - 64) = 6.6G$. The take-up of AFP benefits is increasing from pre- to post-reform cohorts, which largely is due to increased pension accrual.[14]

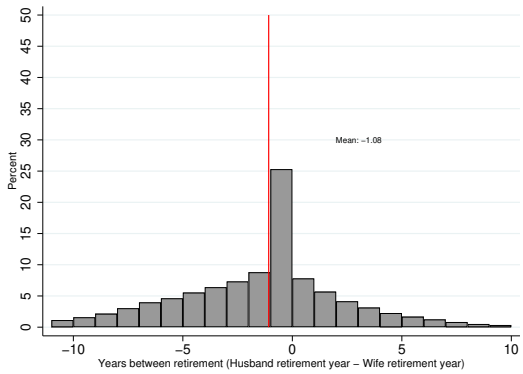
Figure 1 shows the age-difference distribution and the years between retirement within couples in the sample. The graphs are bound on a max age-difference (or year-difference) of 10, but the means are calculated on the entire sample. From panel 1b, we see that the age-difference distribution has quite large dispersion, and a mean of 2 years older husbands. Even though men are older on average, panel 1a shows that more than one quarter of all couples retire within the same calendar year, and more than 45 percent within one year of each other. The average time between retirement is approximately

[14]On average, accrual rates improved by approximately 12 percent as a result of real wage increases between the cohorts. Essentially, accrual is a mapping based on the 20 best years of income, as explained in Section II. For the younger cohorts, i.e. the post-reform cohorts, accrual is thus improved as the average of their best income years are higher due to real wage increases.

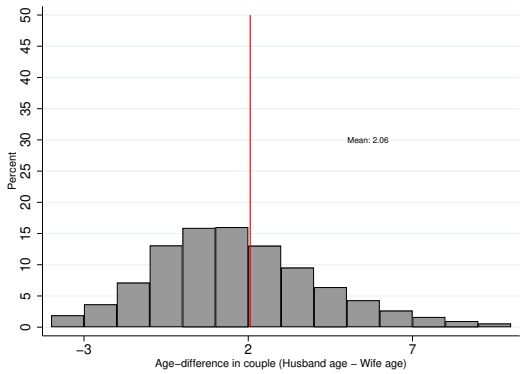
Figure 1

Joint exit and age-difference distribution in couples. In panel (a), left tail are couples where the wife retires last. In panel (b), left tail are couples where the wife is older.

(a) Joint exit



(b) Age-difference distribution



Notes: The vertical lines show the mean in each distribution, calculated on the entire unbound distribution. The graphs are bound to +10 and -10.

1 year (the wives retire after their husbands on average). This suggests that couples coordinate the timing of retirement.

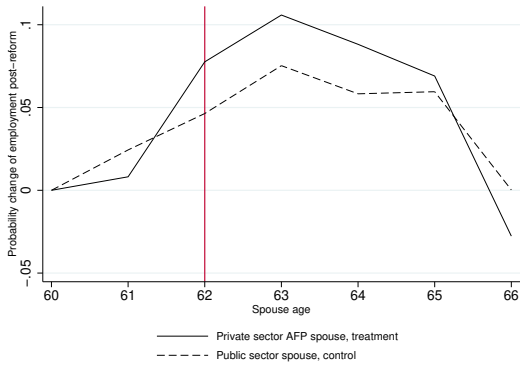
In Figure 2, I plot the difference between the probability of remaining employed at a certain age when being married to a spouse born before and after 1949. This is done separately for spouses in the private sector with AFP (the treatment group, solid line) and spouses in the public sector (the control group, dashed line). In the presence of spousal spillover, we should expect that there is a larger positive effect on the probability of remaining employed if the spouse works in the private sector with AFP, as the treatment group received positive work incentives, as opposed to workers married to a public sector spouse. The figures show that the women married to post-reform cohort husbands have a larger increase in the employment probability when the husband is a private sector worker with AFP, compared to a public sector husband. For men, the evidence is less conclusive.[15] Thus, even if the graphs are only descriptive, they provide some suggestion of spillovers towards women, but less convincing evidence of spillover onto men. If there is spillover onto men, it looks to be stronger at higher ages.

[15]The graphs do not include any correction for time trends or age-composition effects, and are therefore only suggestive.

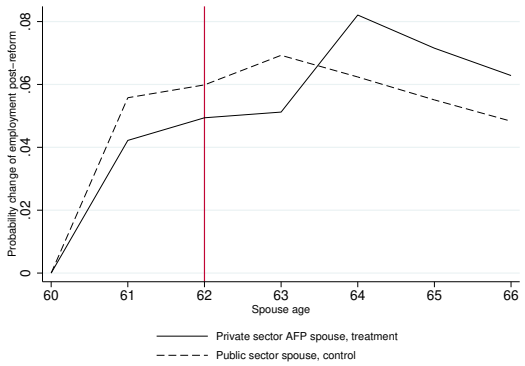
Figure 2

Increased employment post-reform from being married to a private sector spouse with AFP versus public sector spouse.

(a) Female focal partner



(b) Male focal partner



IV Empirical framework

I use an instrumental variable (IV) approach, estimated using two-stage least squares (2SLS). The couples are of two types. First, the focal partner is working in the public sector and the spouse in the private sector with AFP. This is the treatment group, since the focal partner would be affected by the reform through the spouse if the spouse is born in 1949 or later (the treatment period). Second, the focal partner works in the public sector and the spouse works in the public sector. This is the control group, as the focal partner would be unaffected by the reform even if the spouse is born in 1949 or later. Therefore, treatment is essentially defined as the spouse working in the private sector with AFP. The equation system in (1) and (2) displays the first stage and the second stage of the 2SLS, respectively:

$$E_{i,t} = \alpha_0 + \alpha_1 D_i^{private} + \alpha_2 (D_i^{private} \times D_i^{post}) + \alpha_3 X_{ij} + \lambda_t + \varepsilon_{i,t} \tag{1}$$

$$E_{j,t} = \phi_0 + \phi_1 D_i^{private} + \phi_2 E_{i,t} + \phi_3 X_{ij} + \lambda_t + \varepsilon_{j,t} \tag{2}$$

where j indicates the focal partner; the partner that we are interested in studying the spillover effects onto, while i is the spouse of that partner. In the first-stage regression 1, $E_{i,t}$ is an indicator variable for the employment status of spouse i at time t, taking the value one if i is employed at time t. $D_i^{private}$ is a time-invariant dummy variable

taking the value one if i is employed in the private sector and covered by AFP, defined when spouse i is age 60. The variable is time-invariant and does not change when the individual exits the labor market, i.e. when $E_{i,t} = 0$. Similarly, D_i^{post} is a dummy variable taking the value one if i is born in the post-reform cohorts (born in 1949 or later), and is therefore also time-invariant. X_{ij} is includes control variables for the couple, all of which are time-invariant, pre-determined the year when spouse i is aged 60. λ_t is year fixed effects, and includes dummies for each year. $E_{j,t}$ is the indicator for the employment status of the focal partner j, taking the value one if the focal partner retired. Since the focal partner always is a public sector worker, the equation system does not include a sector-variable for spouse j. The second stage regression in Equation 2 then yields the LATE of the spouses' labor market status on the focal partner's employment.

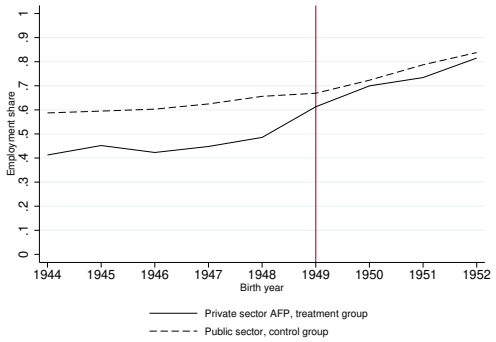
Identifying assumptions

The underlying reduced form difference-in-difference, for which the 2SLS is an extension, implies that the identifying assumption is that public sector focal partners married to a private sector spouse with AFP or public sector spouse share a common trend in their employment among the pre-treatment cohorts (those born before 1949). Next, the IV in the 2SLS relies on two identifying assumptions; instrument relevance and the exclusion restriction. Instrument relevance clearly holds, as own access to AFP clearly affects own employment (this is essentially tested in the first-stage regressions). Second, the exclusion restriction requires that the instrument is conditionally independent of the potential outcome variable, and only affects the potential outcome through the first stage. This assumption cannot be empirically tested, however, I argue here that it holds. Since I do not allow the treatment effect to happen to both partners, the only potential effect on the focal partner when the spouse is treated is that the spouse affects the focal partner indirectly. One might argue that if the spouses' labor supply is interlinked through the budget constraint, then the income effect of the policy change may affect the focal partner even in the absence of leisure complementarity. However, this of course is only relevant if the spouse in fact responds to the treatment, and therefore the effect only comes indirectly through the spouse, not directly onto the focal partner. Firstly, I show

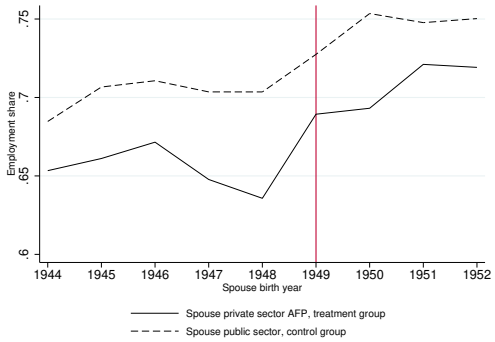
Figure 3

Common trend. Direct effect by own birth cohort. Spillover effect onto 63 year-old focal partners, by spouse's birth cohort. Outcome is employment.

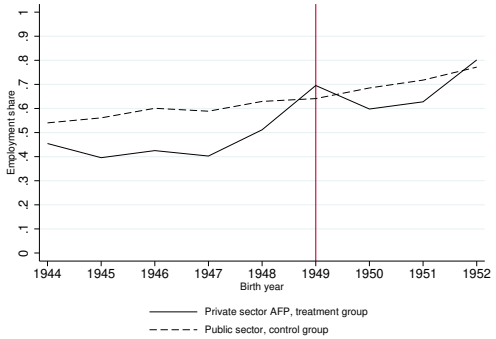
(a) Own-treatment, men



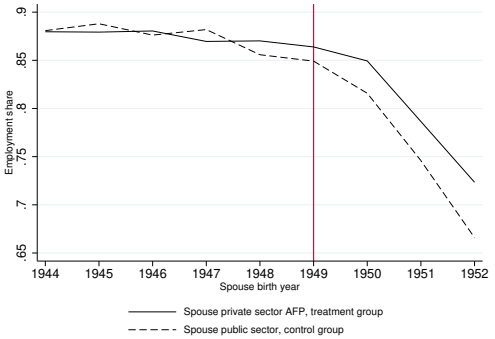
(b) Spouse treated, men aged 63



(c) Own-treatment, women



(d) Spouse treated, women aged 63



Notes: Panels (b) and (d) do not share scaling on the vertical axis, as the reduced form effects of spousal treatment are small. Level differences in panel (b) are quite small compared to panels (a) and (c).

that the pre-trends are parallel visually in Figure 3.

Furthermore, I test the common trend assumption by investigating whether there are slope differences in panels 3a–3d. Essentially, the test of the common trend is:

$$E_{s,t} = \kappa_0 + \kappa_1 D_i^{private} + \kappa_2 (D_i^{private} \times D_{i,t}^{cohort}) + D_{i,t}^{cohort} + \varepsilon_{s,t} \tag{3}$$

where s denotes either the spouse i or the focal partner j, depending on whether we are testing the direct effect ($s = i$) or the spillover effect ($s = j$). The coefficient κ_2 tests if there are slope differences between private sector AFP workers and public sector workers across cohorts, while κ_1 displays average level differences in the pre-treatment cohorts (1948 and earlier). Since the panels are restricted so that the spouse is 63–66 years old, the coefficients show the average employment increase among the individuals either born in cohort t (the direct effect) or married to an individual born in cohort t

(the spillover effect). For the common trend to hold, the coefficients for the cohorts 1948 and earlier must be insignificant, while coefficients in 1949 and later may be significant, which would indicate responses to the reform incentives. Moreover, if κ_2 is significant for the spillover employment (when $s = j$), this is a sign of spousal spillovers. Table 4 shows the results, denoting the relevant indicator for the slope difference between private sector AFP treatment workers and public sector control workers:

$$\mathbb{I}_t = D_i^{private} \times D_{i,t}^{cohort} \tag{4}$$

The direct treatment effect has a parallel trend for both men and women, and as shown in Table 4, there is an immediate treatment effect in the reform-cohorts. We can see that the 1950 and 1951 cohorts do not have slope-coefficients that are different between treatment and control for women, but if we inspect Figure 3 panel (c), we clearly see that there is a smaller level difference between treatment and control for all post-reform cohorts. There is a positive and significant spillover effect onto women from their husbands, but no significant effect onto husbands from their wives, and the pre-trends are parallel.

There are two opposing effects at play which affect partner j when the reform happens. First, as partner i receives increased work incentives, partner i postponing retirement would lead to the focal partner's marginal utility of leisure decreasing (the focal partner would also want to work more, as the partner is now "less at home"). This is the *leisure complementarity effect*.

However, postponing retirement also *increases household total income*, essentially increasing marginal utility of leisure for spouse j. This is the *income effect*. As the two effects are opposing, the sign of the coefficient ϕ_2 in Equation 2 can only become positive if the complementarity effect dominates. If the partner, in response to the reform, chooses to act exactly as he or she would have in absence of the reform, this may lead to short-run decreases in income, but the expected future income stays the same, as explained in Section II. However, if the spouse does not respond to the reform, i.e. does not increase labor supply, the income effect still has a negative relationship with the focal partner's labor supply in the short run, pulling towards a *negative* spillover coefficient. I

Table 4
Test of the common trend assumption, Figure 3 panels (a)-(d). Spouse i age 63–66.

	Direct effect		Spillover effect at 63	
	Panel (a) Direct effect men	**Panel (c) Direct effect women**	**Panel (b) Spillover onto men**	**Panel (d) Spillover onto women**
$\mathbb{I}_{1945}$	0.01	−0.06	−0.01	−0.008
	(0.02)	(0.05)	(0.02)	(0.02)
$\mathbb{I}_{1946}$	−0.01	−0.06	−0.008	0.005
	(0.02)	(0.05)	(0.02)	(0.02)
$\mathbb{I}_{1947}$	−0.02	−0.06	−0.02	−0.01
	(0.02)	(0.05)	(0.02)	(0.02)
$\mathbb{I}_{1948}$	−0.002	−0.04	−0.04	0.02
	(0.02)	(0.05)	(0.02)	(0.02)
$\mathbb{I}_{1949}$	0.10***	0.12***	−0.007	0.02
	(0.02)	(0.05)	(0.02)	(0.02)
$\mathbb{I}_{1950}$	0.14***	0.02	−0.029	0.034*
	(0.02)	(0.05)	(0.02)	(0.02)
$\mathbb{I}_{1951}$	0.11***	0.05	0.005	0.041*
	(0.03)	(0.05)	(0.02)	(0.025)
$\mathbb{I}_{1952}$	0.13***	0.10***	0.0003	0.06*
	(0.03)	(0.05)	(0.02)	(0.03)
$D^{private}$	−0.16***	−0.11***	−0.031	0.001
	(0.02)	(0.04)	(0.02)	(0.01)
Obs	91,186	52,873	15,912	24,916
Clusters	24,921	15,915	15,912	24,916

Notes: Annual data from 2007–2015, using the 1944–1952 cohorts for the *i*-spouse. The 1944 cohort serves as base cohort. Cohort fixed effects included. Standard errors clustered at the individual level.
***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

therefore interpret the results as a *lower bound* estimate of leisure complementarity, as I cannot rule out that there is a negative income effect related to the spouse postponing retirement and thereby increasing household income (or not increasing labor supply, and thereby decreasing household income).

V Results

The main results are presented in Table 5 and 6. The first stage coefficient is interpreted as the direct employment effect on the spouse from working in the private sector with AFP in the cohorts affected by the reform. For instance, the first stage column (1) in Table 5 shows that the husbands in the private sector with AFP born in 1949 or later are approximately 21.6 percentage points more likely to remain employed on average in the

Table 5

Two-stage least squares estimates for employment spillover onto female focal partners from their husbands. Focal partner always public sector female aged 63–66 years, spouse public or private sector worker with AFP aged 63–66.

Female (j-spouse) focal partner 63–66 years old	**First stage (E_i)**		**Second stage (2SLS) (E_j)**		**Reduced form (ITT) (E_j)**	
	(1)	(2)	(1)	(2)	(1)	(2)
	0.216***	0.222***	0.168***	0.163***	0.036***	0.036***
	(0.012)	(0.012)	(0.051)	(0.052)	(0.011)	(0.012)
Constant	[0.730]	[0.784]	[0.529]	[0.571]	[0.651]	[0.699]
Controls	YES	NO	YES	NO	YES	NO
F-stat / Wald χ^2	164.70	96.89	6392.09	3417.43	252.42	249.77
df	24	13	24	13	24	13
Obs	48,188	48,188	48,188	48,188	48,188	48,188
Clusters	20,715	20,715	20,715	20,715	20,715	20,715
Treated obs	9,060	9,060	9,060	9,060	9,060	9,060

Notes: Annual data from 2007–2015, using the 1944–1952 cohorts for the i-spouse and their respective j-partners. The treated observations are the number of couples where the spouse works in the private sector with AFP. Control variables (output in Appendix A) are pre-determined at the time when the i-spouse is age 60, and include demeaned variables of income and income squared (in G-units), wealth and wealth squared (in G-units), children and years of non-compulsory education for each partner. One G-unit is 90,068 NOK (approximately $10,000 using the average 2015-exchange rate) in 2015 which is the last year included. Time fixed effects are included in all specifications. Standard errors clustered at the couple level. df = degrees of freedom.

***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

ages 63–66 after the reform. The reduced form shows the direct effect on the focal partner from being married to a treated spouse (i.e. a spouse working in the private sector with AFP born in 1949 or later). The reduced form column (1) in Table 5 can be interpreted as the intention-to-treat effect on the focal female partners aged 63–66, i.e. as a 3.6 percent employment effect of being married to a treated husband. The second stage coefficient is the 2SLS coefficient, i.e. the local average treatment effect. It is essentially a scaling of the ITT with the first-stage effect. Thus, the 2SLS coefficient in column (1) in Table 5 can be interpreted as an average 16.8 percentage points increase in employment among 63–66 year-old focal female partners, given that their husband postpones retirement by a year.

The first-stage coefficients in Table 6 are also significant and positive, and close to the first-stage coefficients in Table 5 in magnitude. This essentially shows that among the directly affected men and women, the employment effect is roughly the same. However, when we inspect the reduced form (ITT) effect and the second stage effect in Table 6,

Table 6

Two-stage least squares estimates for employment spillover onto male focal partners from their wives. Focal partner always public sector male aged 63–66 years, spouse public or private sector with AFP worker aged 63–66.

Male (j-spouse) focal partner 63–66 years old	**First stage (E_i)**		**Second stage (2SLS) (E_j)**		**Reduced form (ITT) (E_j)**	
	(1)	(2)	(1)	(2)	(1)	(2)
	0.198***	0.204***	0.059	0.079	0.012	0.016
	(0.025)	(0.025)	(0.120)	(0.122)	(0.024)	(0.025)
Constant	[0.679]	[0.730]	[0.652]	[0.678]	[0.693]	[0.736]
Controls	YES	NO	YES	NO	YES	NO
F-stat / Wald χ^2	61.87	41.17	5524.03	1963.87	258.64	145.26
df	24	13	24	13	24	13
Obs	31,686	31,686	31,686	31,686	31,686	31,686
Clusters	13,522	13,522	13,522	13,522	13,522	13,522
Treated obs	1,500	1,500	1,500	1,500	1,500	1,500

Notes: Annual data from 2007–2015, using the 1944–1952 cohorts for the i-spouse and their respective j-partners. The treated observations are the number of couples where the spouse works in the private sector with AFP. Control variables (output in Appendix A) are pre-determined at the time when the i-spouse is age 60, and include demeaned variables of income and income squared (in G-units), wealth and wealth squared (in G-units), children and years of non-compulsory education for each partner. One G-unit is 90,068 NOK (approximately $10,000 using the average 2015-exchange rate) in 2015 which is the last year included. Time fixed effects are included in all specifications. Standard errors clustered at the couple level. df = degrees of freedom.

***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

we see that the spillover effect is insignificant. The coefficients for the reduced form are smaller for men in Table 6 than for women in Table 5, which may be an indication of different spillover effects between the genders, even though the reduced form is insignificant for men. However, the difference in the clustered standard errors in Table 6 and Table 5 are consistent with the treatment group in Table 6 being about 5 times smaller.[16]

Direct effect. The employment effect among 63–66 year-old men that would be predicted by an estimation disregarding spousal spillovers, would be the first-stage effect of approximately 21.6 percentage points. Then, approximately 43 percent of the male workers work in the private sector with AFP, meaning that the aggregate employment effect would be approximately 9.3 percentage points in this age group among men. For women, approximately 11 percent in the age group are working in the private sector with AFP, hence, the aggregate employment effect is approximately 2.2 percentage points in

[16]This will roughly increase the clustered standard errors by $\sqrt{5}$. This has a natural explanation in the fact that fewer women are working in the private sector with AFP, cf. Table 2.

this age group, by using the first-stage effect of 19.8 percentage points in Table 6. In total, this means that workers aged 63–66 increased their labor market participation by approximately 7.9 percentage points as a result of being directly targeted by the reform.

Spillover effect. However, as the male spouses are married to female focal partners who respond strongly to their husbands' labor market decisions, the aggregate employment effect for women must also include the spillover from their spouses. Scaling the first-stage effect in Table 5 by $(1 - 0.168)$ and multiplying by 43 percent, we get that the total employment effect from increasing work incentives for men is approximately 11.1 percentage points, where 9.3 percentage points is attributed to the direct effect on men. The remaining 1.8 percentage points can be attributed to the wives' increasing labor participation. Hence, the total aggregated employment effect on women, including both directly treated women and spillover from husbands for women aged 63–66, is 4 percentage points, where approximately 45 percent of the aggregated employment effect on women aged 63-66 comes from spillover from their husbands.

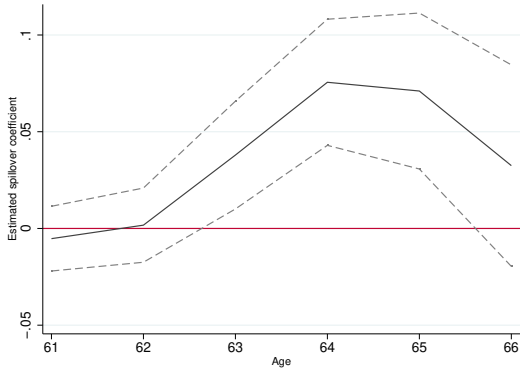
If we ignore the insignificance of the spillover-coefficient for men, the same exercise as above results in a spillover onto men of approximately 0.1 percentage points. Hence, spillover from the wife would account for a negligible fraction of the aggregated employment effect on men, regardless of the imprecise estimation.

Next, I estimate the reduced form (DID) separately for ages 61–66 for female and male focal partners, presented in Figure 4. This essentially allows for studying heterogeneity across ages in terms of spillovers. As panel 4a shows, the spillover is strongest at age 64 for women, and stronger at age 65 than at age 63. For men, panel 4b shows that the spillover is insignificant at all ages, but is closest to being significant at age 64, where the point estimate is also at its largest.

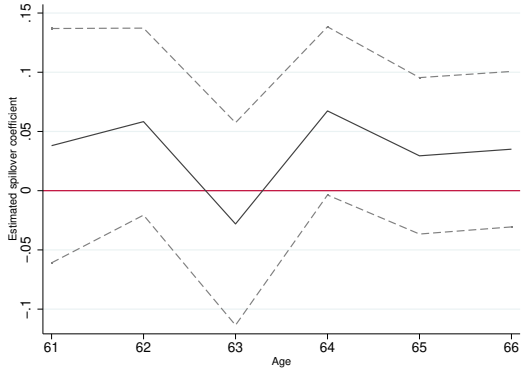
Figure 4

Estimated spillover coefficients (DID), age-by-age, for men and women. Focal partner always public sector worker, spouse public or private sector worker with AFP aged 63–66

(a) Female focal partner



(b) Male focal partner



Notes: Control variables are pre-determined at the time when the i-spouse is age 60, and include demeaned variables of income and income squared (in G-units), wealth and wealth squared (in G-units), children and years of non-compulsory education for each partner. Time fixed effects are included. Standard errors clustered at the couple level. df = degrees of freedom. Confidence bands display the 5 percent significance level.

Heterogeneity in spillovers

One might argue that if there is a clear labor market "breadwinner", this may influence the decision of retiring together, and moreover it may influence who decides on the timing of retirement. I define "breadwinner" as income being at least 25 percent higher than the spouse (when spouse i is 60 years old). Joint retirement may also be more easily available to wealthier households, as they are more likely to afford reducing the combined household income and live off their savings. Furthermore, heterogeneity within the couple with respect to education may be a factor influencing the decision to coordinate the timing of retirement. Table 7 and 8 show different household compositions with respect to income and wealth (when spouse i is 60 years old) and education, respectively.

Clearly, the spillover effect onto women is robust to splitting the sample into household wealth above and below the median, while spillover onto men is still insignificant. However, when the wife is the "breadwinner" (she takes home 25 percent more income than the husband, when the husband is aged 60), the spillover effect is insignificant with a large point estimate. However, if we inspect the observations, there are approximately 11 times fewer treated observations, and approximately 10 times fewer observations overall. This is consistent with the standard error being about $\sqrt{10}$ times higher. For men, the

Table 7

Income heterogeneity, sub-sample analysis. Two-stage least squares estimates for employment spillover onto focal partner from their spouse. Focal partner always public sector individual aged 63–66 years, spouse public or private sector with AFP worker aged 63–66.

Focal partner (j-spouse) 63–66 years old	**Husband breadwinner**		**Wife breadwinner**		**Household wealth above median**		**Household wealth below median**	
Focal partner:	Women	Men	Women	Men	Women	Men	Women	Men
	0.166***	0.146	0.226	0.091	0.168***	0.081	0.154***	0.089
	(0.060)	(0.157)	(0.229)	(0.214)	(0.078)	(0.173)	(0.066)	(0.165)
Constant	[0.522]	[0.620]	[0.473]	[0.819]	[0.522]	[0.596]	[0.870]	[0.993]
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Wald χ^2	4325.17	2555.52	566.86	1474.26	3499.42	5201.63	2966.36	2152.03
df	24	24	24	24	24	24	24	24
Obs	31,239	17,744	3,439	3,307	24,096	15,843	24,092	15,843
Clusters	13,388	7,459	1,514	1,545	6,736	6,925	10,332	6,597
Treated obs	6,736	786	519	236	4,554	744	4,506	756

Notes: Only second-stage in table, first-stage is positive and significant for all specifications. "Breadwinner" defined as taking home 25 percent more than the spouse, in the year when the spouse is age 60. Annual data from 2007–2015, using the 1944–1952 cohorts for the i-spouse and their respective j-partners. The treated observations are the number of couples where the spouse works in the private sector with AFP. Control variables are pre-determined at the time when the i-spouse is age 60, and include demeaned variables of income and income squared (in G-units), wealth and wealth squared (in G-units), children and years of non-compulsory education for each partner. One G-unit is 90,068 NOK (approximately $10,000 using the average 2015-exchange rate) in 2015 which is the last year included. Time fixed effects are included in all specifications. Standard errors clustered at the couple level. df = degrees of freedom.

***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

point estimate for spousal spillover is larger when he is the breadwinner, but the estimate is still imprecise.

In Table 8, we see that when both spouses have a low level of education, the spillover onto women is positive and significant. In this group, we are investigating women in the public sector with low levels of education married to men with low levels of education (either in the private sector with AFP or in the public sector). In the group where both have high levels of education, the spillover onto women has a lower point estimate, and is insignificant (again, consistent with a smaller treatment group). The evidence suggests that women with lower levels of education tend to be influenced by their husband to a larger extent, as also the point estimate is very high when the husband is highly educated while the wife is poorly educated.

Finally, in Table 9, I test whether the within-couple age difference plays a role in spousal spillover. Since Figure 1b shows that men are older than their wives on average,

Table 8

Education heterogeneity, sub-sample analysis. Two-stage least squares estimates for employment spillover onto focal partner from their spouse. Focal partner always public sector individual aged 63–66 years, spouse public or private sector with AFP worker aged 63–66.

Focal partner (j-spouse) 63–66 years old	**Both high education**		**Both low education**		**Husband high wife low education**		**Husband low wife high education**	
Focal partner:	Women	Men	Women	Men	Women	Men	Women	Men
	0.101	−0.061	0.192***	0.037	0.290	0.079	0.109	−0.167
	(0.138)	(0.352)	(0.073)	(0.217)	(0.188)	(0.168)	(0.099)	(0.606)
Constant	[0.464]	[0.654]	[0.481]	[0.578]	[0.467]	[0.710]	[0.570]	[0.672]
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Wald χ^2	3190.27	3027.97	1829.33	521.47	896.19	931.98	842.10	191.62
df	24	24	22	22	23	23	23	23
Obs	18,609	15,770	17,008	7,407	6,910	6,164	5,595	2,317
Clusters	7,909	6,745	7,445	3,195	2,924	2,600	2,411	972
Treated obs	1,642	223	4,981	707	942	513	1,475	55

Notes: Only second-stage in table, first-stage is positive and significant for all specifications. In the last column, the first-stage is significant at a 10 percent level. Annual data from 2007–2015, using the 1944–1952 cohorts for the i-spouse and their respective j-partners. The treated observations are the number of couples where the spouse works in the private sector with AFP. Control variables are pre-determined at the time when the i-spouse is age 60, and include demeaned variables of income and income squared (in G-units), wealth and wealth squared (in G-units), children and years of non-compulsory education for each partner. One G-unit is 90,068 NOK (approximately $10,000 using the average 2015-exchange rate) in 2015 which is the last year included. Time fixed effects are included in all specifications. Standard errors clustered at the couple level. df = degrees of freedom.

***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

testing on a sub-sample of only older is important for understanding whether women only "follow naturally" since they are younger on average.

The results show that the point estimate for spillover onto women is robust to splitting into older wives or older husbands, while the point estimates for men are imprecise with large standard errors.[17]

Implications for public expenditure

One of the main policy goals of the reform was to reduce public expenditure and improve the fiscal adequacy of the pension system. To achieve this goal, incentives towards prolonged working careers were introduced, and as shown they had significant effects on private sector workers. As shown in Section V, public sector women married to private sector men responded by postponing retirement by 16.8 percentage points on average

[17]The first-stage in column 2, where the wife is older and the husband is the focal partner, is significant and positive at all levels.

Table 9

Within-couple age-difference. Two-stage least squares estimates for employment spillover onto focal partner from their spouse. Focal partner always public sector individual age 63–66, spouse public or private sector worker with AFP aged 63–66.

Focal partner (j-spouse) 63–66 years old	**Wife older**		**Husband older**	
Focal partner:	Women	Men	Women	Men
	0.175*	−0.544	0.152*	0.212
	(0.104)	(0.420)	(0.161)	(0.136)
Constant	[0.525]	[1.03]	[0.572]	[0.480]
Controls	YES	YES	YES	YES
Wald χ^2	4545.57	1300.54	2212.86	4174.05
df	24	24	23	23
Obs	25,016	15,582	23,172	16,104
Clusters	8,822	5,430	11,893	8,093
Treated obs	3,643	512	5,417	988

Notes: Only second-stage in table, first-stage is positive and significant for all specifications. Annual data from 2007–2015, using the 1944–1952 cohorts for the i-spouse and their respective j-partners. The treated observations are the number of couples where the spouse works in the private sector with AFP. Control variables are pre-determined at the time when the i-spouse is age 60, and include demeaned variables of income and income squared (in G-units), wealth and wealth squared (in G-units), children and years of non-compulsory education for each partner. One G-unit is 90,068 NOK (approximately $10,000 using the average 2015-exchange rate) in 2015 which is the last year included. Time fixed effects are included in all specifications. Standard errors clustered at the couple level. df = degrees of freedom.

* * *Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

between the ages 63–66. A natural question is therefore how this affects public expenditure. By postponing retirement, public sector workers essentially abstain from taking AFP benefits while working and paying taxes. If they work past age 66, they abstain from claiming any benefits.[18] In the same setup as presented in equations 1 and 2, I use *total AFP take-up* and *taxes paid* as dependent variables. Firstly, I use the sum of all AFP transfers to the individual between observed claiming age and age 66, then, the sum of taxes paid in the same years. I obtain the following estimates, presented in Tables 10 and 11.

Public sector women claim approximately $1.83G$ (approximately $18,100) less AFP benefits and pay approximately $0.39G$ (approximately $3,850) additional taxes, as a result of their husbands extending their working careers by a year. A full year of AFP benefits averages $2.2G$, meaning that this translates to approximately 10 months of reduced

[18]Essentially, for each year the worker extends his or her working career, the total potential AFP benefits are reduced proportionally.

Table 10

Total AFP take-up. Two-stage least squares estimates for spillover onto focal partner from their spouse. Focal partner always public sector individual age 63–66, spouse public or private sector worker with AFP aged 63–66.

Focal partner (j-spouse) 63–66 years old	**Wife's total AFP take-up**		**Husband's total AFP take-up**	
	First-stage (E_i)	Second-stage (AFP_j)	First-stage (E_i)	Second-stage (AFP_j)
	0.216***	−1.828***	0.198***	−1.203
	(0.012)	(0.317)	(0.025)	(0.840)
Constant	[0.730]	[3.174]	[0.679]	[2.819]
Controls	YES	YES	YES	YES
F-stat / Wald χ^2	164.70	16996.27	61.87	17862.92
df	24	24	24	24
Obs	48,188	48,188	31,686	31,686
Clusters	20,715	20,715	13,522	13,522
Treated obs	9,060	9,060	1,500	1,500

Notes: Annual data from 2007–2015, using the 1944–1952 cohorts for the i-spouse and their respective j-partners. Outcome variable AFP_j is measured in G (basis points). One G-unit is 90,068 NOK (approximately $10,000 using the average 2015-exchange rate) in 2015 which is the last year included. Control variables are pre-determined at the time when the i-spouse is age 60, and include demeaned variables of income and income squared (in G-units), wealth and wealth squared (in G-units), children and years of non-compulsory education for each partner. The treated observations are the number of couples where the spouse works in the private sector with AFP. Time fixed effects are included in all specifications. Standard errors clustered at the couple level. df = degrees of freedom.

***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

claiming on average. Meanwhile, income at age 60 on average is approximately $2.7G$ for women, cf. Table 3, hence they pay taxes equivalent to approximately 5 months, given an average tax rate $\tau \approx 0.35$. By computing the product of the reform effect onto men (the first-stage effect) and the reduction in take-up and increased tax payments from women, we can aggregate the *isolated* effect on public expenditure only coming from spillovers onto women. In total, this amounts to approximately $35 million in savings on public expenditure from decreased take-up and $8.4 million from increased tax payments, over the potential years of AFP claiming for the affected cohorts.[19] As we see from Table 11, men also significantly increase their tax payments when their wives are treated by the reform by approximately $11,000 (equivalent to almost 7 months). This indicates an *intensive margin* effect, as the previous results have shown little evidence of an extensive margin effect. The increased taxes paid by the husbands is aggregated at approximately $2.5 million. In total, the savings on public expenditure amount to $46 million, spread across

[19] $0.198 \times (1.828 + 0.386) \times 9,060 \approx 3,900G \approx \$35$ million

Table 11

Taxes paid. Two-stage least squares estimates for spillover onto focal partner from their spouse. Focal partner always public sector individual age 63–66, spouse public or private sector worker with AFP aged 63–66.

Focal partner (j-spouse) 63–66 years old	**Wife's total taxes paid on labor income**		**Husband's total taxes paid paid on labor income**	
	First-stage (E_i)	Second-stage (Tax_j)	First-stage (E_i)	Second-stage (Tax_j)
	0.216***	0.386***	0.198***	0.751**
	(0.012)	(0.072)	(0.025)	(0.320)
Constant	[0.730]	[0.363]	[0.679]	[0.846]
Controls	YES	YES	YES	YES
F-stat / Wald χ^2	164.70	4584.70	61.87	6588.39
df	24	24	24	24
Obs	48,188	48,188	31,686	31,686
Clusters	20,715	20,715	13,522	13,522
Treated obs	9,060	9,060	1,500	1,500

Notes: Annual data from 2007–2015, using the 1944–1952 cohorts for the i-spouse and their respective j-partners. Outcome variable Tax_j is measured in G (basis points). One G-unit is 90,068 NOK (approximately $10,000 using the average 2015-exchange rate) in 2015 which is the last year included. Control variables are pre-determined at the time when the i-spouse is age 60, and include demeaned variables of income and income squared (in G-units), wealth and wealth squared (in G-units), children and years of non-compulsory education for each partner. The treated observations are the number of couples where the spouse works in the private sector with AFP. Time fixed effects are included in all specifications. Standard errors clustered at the couple level. df = degrees of freedom.

***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

all the years the affected cohorts could have claimed these potential benefits and have paid these additional taxes, i.e. from 2012–2015, which means average annual savings are around $11.4 million.

VI Concluding remarks

In this paper, I have shown that spousal spillovers account for substantial parts of aggregate labor market responses to incentive changes. Using a major reform of the Norwegian pension system, I have identified how elderly women in particular increase their labor supply when their husbands increase theirs. Women aged 63–66 are on average approximately 16 percentage points more likely to remain employed if their husband postpones retirement by a year. The effect is strongest on 64 year-old women, and stronger on 65 than on 63 year-olds. For men, the results are inconclusive, but the point estimates are lower than for women, suggesting at least that spillover onto men is likely lower than onto

women. For the 2011 Norwegian reform, spillovers accounted for more than 40 percent of the aggregate increase in labor supply for women at these ages. There is little evidence that combined household wealth or whether either spouse is a "clear breadwinner" affect the estimation results. There is, however, some indication that poorly educated women are more likely to be influenced by their husband's decisions.

From a policy perspective, the aggregate effect on public expenditure is important. By postponing retirement, workers in the public sector abstain from take-up of AFP benefits and pay additional taxes on labor income. I show that women married to a husband affected by the reform on average decrease their take-up of AFP benefits by $18,100 and pay $3,850 in additional tax on labor income, which is equivalent to approximately 10 months of reduced claiming and approximately 5 months of additional tax payment. Additionally, men increase their tax payments by approximately $11,000, equivalent to around 7 months of additional tax payments. By aggregating this response across couples where the husband is affected by the reform, I find that this reduced public expenditure by approximately $46 million over the years 2012–2015.

The macro implications of these findings are first that the utility that wives draw from their own leisure, seems to be non-separable from their husbands' leisure, i.e. that the wife's marginal utility of her own leisure increases when the husband increases his time spent on leisure. I interpret this as leisure complementarity.[20] Since the point estimates show significant and positive coefficients, the spillover effects presented are lower bound estimates of leisure complementarity. This suggests that incentives are interlinked, and thus it is not straightforward to account for the aggregate implications of policy changes affecting individual incentives, meaning that one should be careful to draw conclusions on aggregate effects based on estimates disregarding household compositions and interlinked incentives.

[20] I argue that if the effect was running through the budget constraint it would have the opposite sign; if, for instance, the husband works an additional year because of the reform, his total income would increase, increasing household total wealth. If the income effect dominates, this should *ceteris paribus*, lead the wife to *reduce* her labor supply.

References

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

Bratsberg, B. and Stancanelli, E. (2017), Partial retirement and partners' labor supply: Learning from a Norwegian retirement reform, *The Frisch Center, Oslo* (October 31).

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

Coile, C. (2004), Retirement Incentives and Couples' Retirement Decisions, *The B.E. Journal of Economic Analysis & Policy, De Gruyter 4*(1), 1–30.

Gerard, F. and Nekby, L. (2012), Spousal Joint Retirement: A Reform Based Approach to Identifying Spillover Effects, *The Institute for Labor Market Policy Evaluation 2010: 126.* Working Paper.

Gustman, A. L. and Steinmeier, T. L. (2000), Retirement in Dual-Career Families: A Structural Model, *Journal of Labor Economics 18*(3), 503–545.

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

Hernæs, E., Jia, Z., Piggot, J. and Vigtel, T. C. (2019), Flexible Pensions and Labor Force Withdrawal, *CEPAR* . Working Paper 2019/3.

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

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

Imbens, G. W. and Angrist, J. D. (1994), Identification and estimation of local average treatment effects, *Econometrica 62*(2), 467–475.

Johnsen, J. V. and Vaage, K. (2017), Spouses' retirement and the take-up of disability pension, (University of Bergen). Working Papers in Economics 03/15.

Kapteyn, A. and Ypma, J. Y. (2007), Measurement Error and Misclassification: A Comparison of Survey and Administrative Data, *Journal of Labor Economics 25*, 513–551.

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

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

OECD (2019), Pensions at a Glance 2019: OECD and G20 Indicators, *OECD Publishing, Paris* .

Røed, K. and Raaum, O. (2003), Administrative registers – Unexplored reservoirs of Scientific Knowledge?, *The Economic Journal 113*(488), F258–F281.

Selin, H. (2012), What happens to the husband's retirement decision when the wife's retirement incentives change?, *3772*. CESifo Working Paper note.

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

Stancanelli, E. and van Soest, A. (2012b), Joint leisure before and after retirement: A double regression discontinuity approach. IZA Discussion Papers 6698.

Zweimüller, J., Winter-Ebmer, R. and Falkinger, J. (1996), Retirement of spouses and social security reform, *European Economic Review 40*, 449–472.