Winning the Pension Lottery: Old-age Labor Market Responses to an Increase in Pension Wealth

Evidence from Norway

Yngvar Gotaas Bonde*

Abstract

This thesis investigates the effects of qualifying for Avtalefestet Pensjon (AFP), a supplementary pension scheme in the Norwegian private sector. When turning 62, eligibility requires a minimum tenure at one or more AFP-covered firms. Using Norwegian administrative data, I identify private sector workers who are just below and just above the tenure threshold, employing a fuzzy Regression Discontinuity Design for analysis. The findings reveal strong responses to AFP eligibility and its corresponding increase in individuals' pension wealth. This includes significant reductions in hours worked and labor earnings, earlier withdrawals of public pensions, and earlier retirement. Relatively affluent workers predominantly drive the responses. The calculated elasticity of old-age labor supply with respect to pension wealth is -1.68. Contrary to earlier studies, I find only modest evidence of substitution for other social security. Overall, this thesis establishes a clear link between pension wealth and old-age labor supply and demonstrates substantial indirect fiscal costs of the AFP through this channel.

Keywords: *Pensions; Retirement; Labor Supply; Regression Discontinuity; AFP; Administrative Data*

Spring 2024
Department of Economics
EC9901 Master's Thesis in Economics
Supervisor: Prof. Mårten Palme

Stockholm University

∗. This thesis was completed at The Ragnar Frisch Centre for Economic Research, Oslo, Norway. I thank the center for their hospitality and for providing access to the data from Statistics Norway, which was essential for writing this thesis. Special thanks to Øystein Hernæs for facilitating my stay and providing valuable feedback. I appreciate Herman Kruse and Andreas Myhre at Statistics Norway for generously meeting with me and for their helpful advice. I am also grateful to my supervisor, Prof. Mårten Palme, for his comments and guidance. Any errors are mine alone.

1 Introduction

Rapidly aging populations have made the economic and societal sustainability of pension systems an increasingly pressing issue in most industrialized countries. In 2021, government pension spending in the OECD accounted for 7.7 percent of the countries' aggregated GDP, and pension spending is now the member states' single largest expenditure (OECD 2024). Adding to the direct expenditures, pension systems' design has indirect fiscal implications by influencing labor market behavior through its importance for incentives to work, pension wealth, and norms.

To study pension systems' effects on labor market and retirement behavior, reform-induced changes to benefits and work incentives have been widely used as a source of exogenous variation. An often inherent attribute of pension reforms is that several components of the pension system are changed simultaneously. This makes it difficult to disentangle responses to specific components in isolation, which is useful for understanding the relative impacts of a system's various attributes. As an attempt at such a disentanglement, this thesis investigates the labor market responses to a component of the Norwegian pension system known as *Avtalefestet Pensjon* (hereafter referred to by its acronym AFP) in the private sector. After a comprehensive reform of the Norwegian pension system in 2011, the private sector AFP was changed from an early retirement scheme to a lifelong supplementary pension scheme that is both actuarially fair and decoupled from labor market decisions after age 62. Being eligible for the AFP pension, which half of the Norwegian private sector workers are, represents a large increase in expected old-age non-labor income and pension wealth compared to ineligibility.

This thesis studies the labor market effects of private sector AFP eligibility by applying a Regression Discontinuity Design (RDD). To identify the scheme's causal effects, I exploit one of the qualification rules for private-sector AFP: When turning 62, eligibility requires a minimum number of years employed at one or more AFP firms within a specified age span. Using administrative data, I define AFP tenure relative to the requirement for all Norwegian workers employed in AFP firms when turning 62 between 2011 and 2016. This allows me to compare workers just making the tenure requirement and workers just missing the requirement, due to variation which I argue is exogenous. By following these individuals until the "normal retirement age" of 67, I quantify the causal effect of private sector AFP eligibility and its corresponding increase in pension wealth on old-age labor market behavior. Since changes to AFP status do not change the net return to old-age work, all effects stemming from AFP eligibility can be attributed to the corresponding increase in lifetime expected pension income and thus a pension wealth effect.

The RD analysis is fuzzy, as the assignment variable does not perfectly predict AFP eligibility status. Instead, the estimated probability increase of AFP recipiency, my proxy for AFP eligibility, at the cutoff is 25 percentage points. I therefore distinguish between Intent-To-Treat (ITT) effects and Local Average Treatment Effects (LATE) at the cutoff. The former parameter captures the effect of being assigned to AFP eligibility in my research design, while the latter captures the effect among individuals assigned as AFP eligible and complying with assignment status. I find that being assigned as AFP eligible causes a reduction in total labor market earnings between ages 62 and 67 of approximately $42,000, a 2-hour reduction in average hours of weekly work, and a 6 percent increased probability of retiring before turning 68. When I scale the ITT effects to estimate LATE, I get estimates about four times the magnitude. The estimates are used in a back-of-the-envelope calculation of the old age labor supply elasticity with respect to total pension wealth, yielding an elasticity range of -0.54 using ITT estimates and -1.68 using the LATE estimates, respectively. Although not directly comparable, the calculations indicate labor supply elasticities in the higher (absolute value) range of previously estimated labor supply elasticities with respect to pension wealth and unearned income. I also find evidence that AFP eligibility causes individuals to claim more public old-age pension benefits earlier, despite this pension being actuarially fair and similar in size for individuals just missing and just fulfilling the tenure requirement.

The main implication of my results is that individuals respond strongly to changes in pension wealth, to an extent that should not be neglected when discussing the behavioral and indirect fiscal effects of pension systems. This is especially relevant for the Norwegian context, where public sector AFP will be reformed in 2025 along the lines of today's private sector AFP.

The remainder of the thesis proceeds as follows: Section 2 gives a detailed presentation of the Norwegian pension system and the design of the private sector AFP, including its various requirements. Section 3 provides an overview of earlier literature on labor market responses to changes in pension wealth, previous studies on the Norwegian pension system, and estimates of labor supply responses to changes in other types of unearned income. Section 4 locates the thesis within a conceptual framework and provides predictions for responses to AFP eligibility grounded in economic theory. Section 5 outlines my empirical strategy and presents the data to be used. Section 6 provides the empirical results and evaluates their robustness. Section 7 discusses the results' implications, including a simple welfare analysis, before section 8 concludes.

2 Institutional Setting

Although this thesis investigates the effects of private sector AFP, which is only a subscheme of the Norwegian pension system, it is useful to locate the policy within the broader policy environment. This section provides a broad overview of the Norwegian pension system and the 2011 pension reform, before presenting the history and design of the private sector AFP. The section highlights how the design of the private sector AFP does not alter labor market behavior through changes in work incentives. Instead, it affects labor market behavior through its effect on total pension wealth.

2.1 The Norwegian pension system

The Norwegian pension system can be roughly decomposed into three parts: The public old-age pension system known as Folketrygden (hereafter referred to as FT), employer-based pension schemes (including the private and public sector AFP), and individual pension savings. As a pointer to the components' relative sizes, outstanding pension obligations in 2022 (in present values) were estimated to be 9600 bn. NOK in the FT system, 600 bn. NOK in employer-based pension, and 75 bn. NOK in individual pension savings (NOU 2022:7).

2011 saw the implementation of a comprehensive pension reform, affecting most parts of the Norwegian pension system. Its key elements were the adjustment of retirement ages and benefits to increases in life expectancy, removal of financial disincentives to old-age labor supply, and an updated method for benefit calculation. The reform changed the Norwegian pension system from a Defined Benefit (DB) system towards a Notional Defined Contribution (NDF) system, motivated by a belief in the latter's ability to strengthen work incentives by tightening the link between labor earnings and pension benefit levels.

While the public sector AFP remained largely unchanged under the reform, private sector AFP changed fundamentally. AFP was originally an early retirement scheme covering all public sector workers and about half of private sector workers, being essentially equal in design across the two occupational sectors. The scheme was introduced in the private sector in 1989 after an agreement between Norway's largest employer organization (NHO) and labor union (LO). From 1998 and onwards, AFP-eligible workers could retire from age 62 and claim full pension benefits plus a flat AFP top-up. The AFP benefits stopped at age 67 and were substituted by FT pensions. Thus, the scheme enabled eligible workers to retire before the normal retirement age of 67 with the same yearly pension entitlements as if they had retired at the normal retirement age. The system exhibited strong disincentives for work, however, as AFP recipients faced a *pro-rata* earnings test cutting benefits one-to-one with earnings. In practice, AFP-eligible workers faced a decision of whether to continue working and let go of potential AFP benefits or withdraw from the labor market completely. Gradual transition into retirement was strongly disincentivized.

2.2 Private sector AFP today

The reform changed private sector AFP in several important ways. The earnings test was completely removed, essentially decoupling the labor supply and AFP recipiency decisions and allowing the recipiency of AFP benefits to be freely combined with work. Benefits were no longer restricted to the age span between 62 and 66 years. Instead, the AFP pensions became lifelong top-up annuities that could only be taken out in combination with the FT pensions. The AFP scheme was made actuarially fair and flexible. This allows for AFP to be withdrawn gradually from age 62 and keeps the (net present value) sum of benefits fixed. In practice, the reform changed private sector AFP from an early retirement scheme to a lifelong supplementary pension scheme. AFP's benefit size is calculated as 0.314 percent of the total pension-giving income between age 13 and 61, on average corresponding to roughly a fifth of the FT benefit level. The accrual calculation for AFP stops at age 62, in contrast to the FT accruals which continue until age 70. Two-thirds of the private AFP scheme is funded by the participating firms, with the last third funded by the government. As a result, the scheme's design is decided in collaboration between employee and employer organizations, and the government. Employer contributions are paid into a centrally administered pot similar to a payroll tax. The employer's AFP expenses do not depend directly on AFP recipiency of current or previous employees, but rather on the wages of its current work force.

The most important requirements for qualifying for private sector AFP under the new system are attached to employment history when turning 62 and employment status when claiming the first pension. When turning 62, the individual must have been employed at one or more firms covered by the AFP for a minimum of 7 of the 9 previous years. The tenure requirement was 3/5 years for the 1949 cohort and increased with one year for each subsequent cohort, as shown in Table A2. The requirements applying at the time of withdrawal are also related to tenure, but also to recent earnings and total pension accrual. To be eligible at the time of withdrawal, the workers must have had annual pension-giving earnings above 1 G in both of the two preceding years.[1] Additionally, the calculated level of total pensions at age 67 (both FT and AFP pensions) must exceed a minimum level corresponding to the so-called "guaranteed pension". At the time of the first claim, the worker must have been employed in AFP firms for three consecutive years.[2] Lastly, AFP cannot be combined with the recipiency of disability insurance (DI). DI recipiency after age 62 automatically disqualifies for AFP.

1. Throughot the thesis, *Folketrygdens Grunnbeløp*, referred to as the G, is used to represent monetary values. One G corresponded to 88,370 NOK in 2014, or roughly $14,000 (using the average USD/NOK rate of 0.159 for 2014).

2. It is allowed with up to 52 weeks of certain types of leaves, conditioned on this not being at the end or start of the three years.

In 2022, 163,000 individuals received either FT pensions or AFP pensions, or both, while also participating in the labor market. Unconditional on employment status, 762,000 received FT pensions, and 91,000 received AFP (Statistics Norway 2024).

3 Earlier Literature

Labor market effects of changes to pension wealth have received only modest attention in the literature on labor supply, retirement, and pensions. This is especially prominent compared to the attention devoted to the effect of incentives to work and retirement norms in the pension literature. A possible reason for the asymmetry is the difficulty of obtaining exogenous variation in pension wealth between individuals. Pension wealth is likely correlated with unobservable characteristics such as ability and motivation, thus complicating the identification of causal mechanisms. To overcome endogeneity problems, a common approach in the applied labor supply and retirement literature has been to exploit variation stemming from reform-induced changes. An inherent difficulty with these approaches is to disentangle the effect of reform-induced changes to pension wealth from changes to e.g. returns to work or norms for retirement behavior, as these usually change simultaneously.

Still, reforms have been used to investigate the effect of changes in pension wealth on old-age labor market behavior. A much studied reform was the 1977 Social Security Act amendment in the US. The act reformed the Old Age and Survivor Insurance (OASI), the major public pension scheme in the US social security system. The act abruptly reduced the mean lifetime discounted pension income for the 1917 and later-born cohorts relative to preceding cohorts. Gelber, Iser, and Song (2016) estimated the average reduction in pension wealth to be $6,100 for the 1917 cohort. The sharp difference in social security entitlements between cohorts became known as the "Social Security Notch", and has provided researchers with variation widely used to investigate the US pension system's effect on labor market and retirement behavior. For example, Kruger and Pischke (1992) use the notch to study the role changes in pension wealth had in the downward trend in old-age labor participation among US males, starting in the 1970s. They found that changes in social security wealth explain only a small portion of the reduction in participation, implying a modest pension wealth effect on the retirement decision.

Gelber, Isen, and Song (2016) isolated the effect of changes to the level of future pension benefits from changes in returns to work also caused by the 1977 amendment. Expanding their sample to include both women and men, as well as utilizing administrative data on the whole population, they show that the Notch caused a large increase in old-age worker's earnings. The authors estimate that a $1 increase in OASI benefits caused earnings in the elderly years to decrease by 46 to 61 cents, which they solely attribute to an income effect. Methodologically, they exploit the sharp discontinuity in pension generosity for individuals born just around the Notch date by differences in birth dates around the Notch in an RDD. In a supplementary analysis, the authors find that a $10,000 increase in lifetime discounted pension benefits causes a 0.65 percentage points decrease in the mean yearly participation probability from 1978 to 2012.

Coile and Gruber (2007) also study the effects of the Notch, but on male workers aged 55-69, using survey data between 1992 and 2000. The authors find that 1) individuals do incorporate future income streams in their old-age labor supply decisions, and 2) when isolating the wealth effect from the accrual effect when deciding on the retirement decision in an option value framework, the propensity to retiree increases significantly with pension wealth. They calculate an elasticity of non-participation with respect to pension benefits of 0.16. An important element to consider when using the Notch for estimating labor supply responses is that the information and knowledge of the act's implications may have been limited among the affected cohorts, as argued by Snyder and Evans (2006). This will likely affect the generalizability of studies exploiting the Notch, possibly underestimating the underlying responses.

Costa (1995) uses data from 1900 and studies a historical US pension covering Union Army veterans. The Union Army pension was available to veterans regardless of labor participation and completely decoupled from an individual's earnings history, making it well suited for investigating pension wealth effects. Eligibility was determined by assessing the individual's health status and required army service during the Civil

War. Costa used army service as an instrument in a 2SLS estimation, after thoroughly discussing the instrument's validity. As eligibility for the pension did not alter work incentives, the estimates could be measured directly as the income elasticity. Costa found a non-participation elasticity with respect to pension income of 0.73, indicating a strong response to changes in pension wealth. The result's generalizability, both to different sets of individuals and other periods in time, might be limited.

Non-US estimates of the effect of pension wealth on retirement behavior and labor supply include Laun and Palme (2022). The authors examine the relationship between Social Security Wealth and retirement decisions in Sweden by investigating responses to the 1998 Swedish pension reform. When attempting to isolate the income effect from changes to work incentives also caused by the reform, the authors fail to identify a significant income effect.

In 2014 Germany implemented a pension reform that has been exploited to disentangle the pension wealth effect on labor supply among Geman mothers. The pension wealth of mothers increased on average by 4.4 percent per child born before January 1st, 1992, compared to children born later. Becker et al (2022) used the sharp discontinuity in pension wealth, exogenously given by children's birth dates, to estimate an employment rate elasticity with respect to pension wealth of close to zero before age 55 and of -0.75 close to the retirement age.

Artmann et al. (2023) uses the same reform to quantify the pension wealth effect on a slightly different outcome, namely forward-looking labor supply responses. They found, in contrast to Becker et al, German mothers to be surprisingly forward-looking in their labor supply responses far from retirement. Their estimates give a marginal propensity to earn out of pension wealth of -0.54. The response is completely driven by labor supply adjustments at the intensive margin. The focus on the intensive rather than extensive margin might explain the seemingly contradiction in labor supply responses far from retirement compared to Becker et al.

3.1 Lotteries and labor supply

A setting well suited for investigating income and wealth effects on labor supply can be found in studies of lottery winners. Although not directly analogous, a key attribute of the RDD framework is randomness in individuals' locations around the cutoff. Winning the lottery or qualifying for the AFP is therefore both results of (in the latter case only partly) random processes. Some differences in this comparison should be noted. AFP pensions represent a future income stream and might thus be less salient than a lottery, and AFP workers have not chosen to participate in a lottery and may therefore differ in unobserved characteristics, e.g. risk aversion. Nevertheless, a lottery win represents a shock to wealth exogenous of other labor market characteristics and preferences. Cesarini et al. (2017) use data on Swedish lottery winners from three different lotteries, in which remarkably large portions of Swedes participated. They find that winning the lottery causes significant reductions in the winners' labor supply, both immediately after the win and permanently. The magnitude, however, is quite modest, with pre-tax labor earnings decreasing by about 1 percent of the wealth shock in each of the first 10 years following the win. Interestingly, the authors find that most of the response occurs at the intensive margin and that winners are rather homogenous across ages in their responses. The latter result is relevant in light of this thesis' focus on old-age labor supply, as standard life-cycle models would suggest stronger responses closer to retirement (Blundell, French, and Tetlow 2016). Other lottery studies have generally found larger wealth effects. For instance, Imbens, Rubin, and Sacerdote (2001) estimated around 11 percent of lottery winnings were spent on reducing labor earning in the context of Massachusetts lottery winners, with the effect being larger among older winners.

3.2 The Norwegian pension system

Another arm of relevant literature on the effects of the AFP is constituted by empirical research on the Norwegian pension system. In particular, the effects of the 2011 reform have been thoroughly investigated. Hernæs et al. (2016) studied the labor supply responses to changes in old-age work incentives caused by the 2011 reform. The authors decompose the change in incentives across different worker groups, using differences in how birth cohorts were affected as their identification strategy. The authors find one of the most prominent effects of the 2011 reform to stem from the change in the private sector AFP, and particularly the removal of the earnings test. Despite the absolute value of the private AFP pension, if claimed fully, decreased with the reform, the total private AFP pensions paid out increased as workers no longer faced a *de facto* choice between AFP benefits and the continuation of work. Thus, more eligible workers were observed to actually claim the AFP pension.

When it comes to investigating the effects of eligibility versus ineligibility for private sector AFP, most papers have looked at the pre-reform, or old, AFP. Note here, however, that the pension plan was very different from the post-reform AFP, as discussed in section 2. The studies of the old AFP are still interesting in two aspects: First, the task of identifying the causal effect of AFP was similar with regards to finding exogenous variation in eligibility status. And second, they shed light on the link between information and knowledge about pension wealth and systems and observed labor market behavior. Concerning the first aspect, several methods have been applied in the quest for causal estimates of AFP eligibility on labor market behavior. Bratberg et al (2004) compared workers in private sector firms covered by the AFP and workers in non-AFP firms and argued that these did not differ systematically in unobserved characteristics. The authors found that the economic incentives caused by the scheme did influence retirement decisions. In a conservative estimate, according to the authors, at least 50 percent of AFP retirees would have stayed in the labor force without the scheme. Røed and Haugen (2003) apply a similar identification strategy by comparing workers in AFP and non-AFP firms, while also including public sector employees. They dwell deeper into the plausibility and the level of exogeneity in AFP affiliation and find no significant (observed) differences between workers in affiliated and non-affiliated firms when conditioning on a set of covariates. They also found that the introduction of the old AFP reduced the participation among elderly workers significantly, and this effect became stronger as the program became more settled.

Myhre and Kruse (2023) applied an RDD to investigate the effects of the old AFP scheme in the private sector, with the assignment variable being displacement due to involuntarily job displacements. As eligibility under the old AFP required employment with an AFP firm at age 62, similar to today, they compare workers being displaced just before turning 62 with workers displaced just after turning 62 (when including a notification period). They find no evidence of AFP eligibility impacting the probability of re-employment, but strong evidence of substitution to other social security benefits. Most notably, workers experiencing displacement "just to early" substitute 51 percent of the foregone AFP benefits with disability insurance. The paper thus highlights an important alternative path into retirement for elderly workers, though for a sample entering old age in an initial state of unemployment.

3.3 Structural models

Another approach for estimating labor market responses to changes in pension income is based on the estimation of "deeper" parameters in structural models. The advantage of such models is that they account for complicated interactions between system design and behavior that are difficult to capture in reduced-form models, as well as modeling responses over the whole life cycle. A much-cited example is French (2005) who set up a structural model (focusing on males) for the US Social Security system. The model is calibrated using data from the Panel Study of Income Dynamics. French performs several policy experiments, including a 20 percent reduction in pension benefit levels. The reduction, which corresponds to a reduction in social security wealth of $26, 000 in the model, leads to an increase in labor market earnings after age 62 of $5,500. The response is accompanied by increased labor supply before age 62. The effect is still rather modest, however, especially when compared to the structural estimates of price elasticity of work (i.e. wage changes). This is, however, partly because a reduction in the benefit level in the US context in effect reduced the Social Security Earnings test and thereby the effective marginal tax rate. The complexity of the structural models also allows for the investigation of more complex factors influencing behavioral responses to the pension system. For example, Blundell, French, and Tetlow (2016) emphasize how recent models have devoted increasing attention to liquidity constraints. If present, such constraints are highly relevant for the effect of changes in pension wealth, as individuals usually are unable to borrow against future pensions. Especially when defining housing assets as illiquid, such constraints have indeed been found to reduce the magnitude of the negative effect of pension wealth on labor supply (Blundell, French, and Tetlow 2016).

4 Conceptual Framework

In this section, I sketch a simple theoretical model inspired by standard life-cycle models for optimal labor supply and consumption. The exercise is useful for understanding the mechanisms in play when I later empirically investigate the labor market responses to AFP eligibility.[3]

Consider a worker that maximizes a two-period utility function, where the continued work is an option in period one, and the worker is retired with certainty in period 2. In our setting, it is useful to think of the first period to correspond to the age span between age 62 and 67, and the second period to be the period from the normal retirement age at age 67 until death. If we abstract from discounting, the inter-temporal utility function can be represented as $U = u_1(c_1, l) + u_2(c_2)$, with c_1 and c_2 being consumption in the first and second period, respectively. l is the agent's amount of leisure consumed in period 1. We assume that the two utility functions satisfy standard requirements of concavity and positive first-order derivatives. We further assume that the worker has some liquid private wealth W, faces a fixed wage w, and has the amount P in pension wealth. Under these assumptions, the inter-temporal budget constraint is $c_1 + c_2 = w(1 - l) + W + P$. If we disregard potential credit constraints and initially assume W and P as given, maximization yields standard first-order conditions of marginal utility of consumption being equal across the two periods, and labor supply being chosen such that the marginal utility of leisure over the marginal utility of consumption equalizes the net wage.[4] Importantly, the optimal level of leisure and consumption also relates to the individual's

3. The theoretical model is inspired by the one outlined in Hernæs et al (2016).
4. Specifically, I maximize the Lagrangian $\mathcal{L} = u_1(c_1, l) + u_2(c_2) - \lambda(c_1 + c_2 - (1 - l)w - W - P)$

total wealth level, captured by the first-order conditions of $u'_c = \lambda$ for both periods and $u'_l = \lambda w$ in period 1. The parameter λ captures the increase in the objective inter-temporal utility function to a one-unit increase in the RHS of the individual's budget constraint. It is clear from the first-order conditions that a wealth increase, *ceteris paribus*, increases both the optimal amount of leisure in the first period and consumption in the two periods. This represents the income effect, whose size depends on the explicit functional form of the utility functions.

In our setting, eligibility for AFP can be thought of as a positive and permanent shock to the net present value pension wealth, or $P^{AFP} > P$, without influencing the effective net wage. With only a wealth effect present, we expect the AFP-eligible worker to reduce his labor supply at working age compared to the ineligible worker. The size of the reduction, however, depends on the functional form of the u-functions and the size of the AFP relative to the initial levels of W and P.

The model can be said to be overly simplistic in a couple of ways. A natural expansion involves accounting for a fixed cost of work, likely to affect the choice of l at the extensive margin, and thus be crucial to rationalize the retirement decision. A fixed cost of work might help explain the decision to retire rather than adjusting labor supply at the intensive margin, as it allows for non-convexities in preferences.

More recent theoretical investigations include more factors believed to be important for the relationship between labour supply and pension wealth (see Blundell et al. (2016) for a discussion). An example is health status. A worker's health status is likely to affect work capacity and thereby the wage, which we have previously stated to be exogenous. Health status may also influence the relative valuation of leisure and consumption. Importantly, private health information, and thereby expected mortality information, influences the agent's expected pension wealth by deviating from the number of years-divider used for annual benefit calculations. Formally, annual pension (NPV) benefits are calculated as $\frac{pension\ accrual_i}{n_c}$, where n_c is the individual's birth cohort's expected number of remaining years of life. If the agent believes this number to be smaller (larger) than n_c, P changes accordingly, influencing the agent's lifetime budget constraint. Another element that deserves notice is the possible preference for complementarity in leisure, where the marginal utility of leisure depends on the consumption of leisure by one's family and friends. Recent studies have shown evidence for this complementarity in preferences, also within the Norwegian context (Johnsen, Vaage, and Willén 2022; Kruse 2021).

Lastly, mechanisms that have received increasing attention in the literature recently should be mentioned, namely the role of the saliency of pension benefits. For example, Brinch et al (2017) argue that agents affected by the 2011 Norwegian pension reform were unable or unwilling to take the value of future benefits into account when considering rewards for working. Although such a mechanism can also be present in the context of AFP, for example, due to a lack of information about the scheme and size of future benefits, it is less likely to play a substantial part in the AFP context since the rewards for working, which is usually the hardest to fully understand and respond to, remain unchanged.

The interplay between all three factors presumably affects the optimal leisure and consumption decisions for utility-maximizing individuals, and thus behavioral responses to exogenous changes in pension wealth. As will be shown later, the theoretical considerations can also help rationalize heterogeneity in responses across subsamples of workers.

5 Data and Empirical Strategy

5.1 Data description

The administrative data used in this thesis is provided by Statistics Norway and covers the full Norwegian population. The data contains detailed individual-level records on employment relations, social security take-up, wealth, and labor income, as well as a rich selection of background variables including individuals' month of birth. The individual-level data is matched with employer registers provided by *Fellesordningen for AFP*, the administering body of the AFP scheme in the private sector. The lists contain information on all firms participating in the AFP scheme in the private sector going back to 1996. By combining data on individuals' employment relations and private firms' yearly status as AFP firms, I extract the number of days an individual has been employed at an AFP firm before her 62nd birthday. This constitutes the assignment variable used to determine eligibility for private-sector AFP. The assignment variable is defined as

$$assignment\ var_{i,c} = AFP\ days_{i,c} - requirement_c \tag{1}$$

where $AFP\ days_{i,c}$ is individual i's number of days employed at an AFP firm in the period AFP tenure counts towards the requirement. This is subtracted by the cohort-dependent necessary number of days for AFP eligibility depicted in Table A1. According to *Fellesordningen*, tenure is counted only if the employment at the AFP firm has been the individual's main employment and corresponds to at least 20 percent of a full-time position. When constructing the assignment variable, I first tried to mimic the tenure assessing procedure *Fellesordningen* (2024) report to apply. This included an iterative procedure of keeping and dropping dominating and dominated employment spells, respectively, to be left with only one employment affiliation per person per day. This procedure, however, was vulnerable to miscoding in the form of dropping employment affiliations that seemed to count as AFP tenure. To avoid this problem, a naive approach has been implemented to construct the assignment variable. This is done by simply summarizing the number of days an individual has been reported as an "active" employee at an AFP firm in the tenure-relevant period, thereby disregarding possibly other (main) employment affiliations, working hours, and salary requirements. The consequence is that individuals with several employment relationships at the same time or too few hours worked in the AFP firm, still get these days counted as AFP tenure. As a result, individuals to the left of the cutoff should not have *de facto* more days than coded and therefore be wrongly coded as ineligible with respect to the tenure requirement. Conversely, individuals to the right of the cutoff can have been coded as having longer tenure than they have, thereby being wrongly coded as fulfilling the requirement.[5] Its implications will be investigated when discussing the validity of my

5. Econometrically, the naive approach is likely to induce more non-compliers than crossovers, i.e., more individuals assigned to, but not receiving, treatment than individuals not assigned but receiving

research design.

Another choice made when constructing the variable concerns the lack of exact birth dates, which determines the period for which AFP tenure is relevant. As I only observe an individual's month of birth, I have imposed a birth date of the 15th of the month for all individuals to determine start and stop dates for the tenure period. This is prone to problems when real birth dates differ from the 15th. Unless AFP-employment status is the same at the start and end of months, I will give individuals too few (many) AFP days depending on if they are born after (before) the 15th and employed at an AFP firm at the start (end) month of the tenure period. To account for this, I add 15 days to all individuals' AFP tenure. This follows the naive approach of continuing to be "generous" rather than "strict" when counting the number of AFP days. The implications of these choices are investigated in more detail in section 6.

A remaining concern exists regarding the relatively high proportion of individuals to the left of the cutoff observed to have received AFP between ages 62 and 67, as depicted in Figure 1. The rules are in principle extremely rigid, meaning there should be no individuals not fulfilling the tenure requirement still receiving AFP. The naive approach, where AFP days have been handed out in a liberal rather than conservative fashion, should further reassure this. Hence, the observed share claiming AFP left of the cutoff, which is well above zero, is suspicious. This might be partially explained by tenure being counted for employment spells in AFP firms that take place abroad and for specific, although not the most common, types of work leaves. Unfortunately, I cannot observe these two employment statuses. They therefore constitute sources of measurement errors in the assignment variable.[6] This does not invalidate the identification strategy, however, as long as the presence of measurement errors does not differ systematically across the threshold.

treatment.

6. Other possible explanations for the high share of cross-overs are measurement errors and misreporting in the employment files or the files on AFP status.

5.2 Sample selection

The sample used for analysis consists of individuals born between 1949 and 1954, thus turning 62 between January 2011 and December 2016, and all subject to the "new" private sector AFP rules. The reason for imposing the upper cohort restriction is to facilitate the investigation of outcomes up until the "normal retirement" age of 67. The choice of sample restrictions nevertheless poses a trade-off: Fewer cohorts allow for investigating longer time horizon outcomes at the cost of fewer observations and less precise estimates. A related weakness is that AFP is kept actuarially fair until age 70. Thus, there could be AFP-eligible individuals planning to claim the pension who are not observed as recipients in the analysis due to the upper age-bound of the sample. To check for this, I investigate the number of individuals in the oldest cohorts, therefore observed at older ages, claiming AFP for the first time after turning 67. These workers constitute less than one percent of workers in the sample who have not claimed AFP between 62 and 67. As will be explained, this is only a problem when scaling the effects at the cutoff using the two-stage procedure (which would be upwards biased), as these individuals' behavioral responses to the future AFP pension would be reflected in the outcome variables of interest under the reasonable assumption of individuals reacting to the knowledge of future pension wealth.

Since eligibility for private sector AFP presupposes main employment with an AFP firm at the time of activation, I condition my sample on main employment with an AFP firm at age 62. This further secures against endogeneity between the assignment variable and covariates, as all individuals begin in the same initial employment state when the outcome period starts. Lastly, I restrict the sample to individuals with AFP tenure within 18 months on both sides of the cutoff. The two columns to the right of Table 1 present descriptive statistics of the sample, consisting of 11204 individuals, on predetermined covariates. Public old-age pension entitlements from Folketrygden are calculated based on earnings history at age 60 adhering to the cohort-specific accrual mechanisms for FT pensions, and the same mechanism has been applied to calculate the size of annual AFP benefits. All sums are reported in *Folketrygdens Grunnbeløp*, known as a G, and deflated to the 2014 values, corresponding to approximately $14,000 (using the USD/NOK average exchange rate for 2014.).

5.3 Econometric strategy

To investigate the old-age labor markets effect of qualifying for the AFP pension, I implement a Regression Discontinuity Design (RDD) with Equation (1), AFP tenure relative to the tenure requirement, as the assignment variable. As shown in Figure 1, the assignment variable does not perfectly predict AFP eligibility. Hence, the RDD is fuzzy. I therefore distinguish between intent-to-treat (ITT) effects and Local Average Treatment Effects (LATE). ITT captures the effect of being assigned to treatment on the outcome of interest, while the LATE represents the effect of being assigned to treatment *and* receiving treatment on the outcome of interest.[7]

Intent-to-treat (ITT). The baseline model for investigating the ITT effects is

$$y_{it} = \alpha + f(x_i) + \tau_{ITT} I_{x_i>0} + \delta \mathbf{X}_{it} + \epsilon_{it}, \tag{2}$$

where y_{it} is the outcome of individual i at time t, α is an intercept term, and $\mathbf{X}_{it}$ is a vector of covariates. I denote individual i's realization of the assignment variable as x_i. $f(x_i)$ is an unknown functional form of the assignment variable, and τ_{ITT} represents the reduced form Intent-To-Treat (ITT) effect at the cutoff.[8] This is accomplished by indicator I, which takes the value 1 if $x \geq 0$, and 0 otherwise. τ_{ITT} can thus be interpreted as the effect of being assigned to the treatment condition as opposed to the control condition on the expected outcome y at the cutoff.

For my baseline specification, I follow Cattaneo et al (2019) and estimate two local non-parametric regressions of f, separately for observations to the left and to the right of the cutoff. If not explicitly stated, all bandwidths are chosen based on a data-driven procedure for minimizing the Mean Squared Error (MSE). When fitting the

7. An individual is assigned to the treatment condition if $x_i \geq 0$, where x_i is an individual's tenure relative to the cutoff. Control condition is defined as $x_i < 0$.

8. In the continuity-based RD framework with non-parametric local polynomials, the coefficient is the difference between the intercepts, evaluated at $x = c$, of the fitted lines to the left and right side of the cutoff, respectively.

local polynomials, the procedure aims to balance the trade-off between precision and variance, both increasing with bandwidth size. I allow the bandwidths to vary across specifications while keeping them equal on both sides of the cutoff. This increases the precision of the estimates across specifications since the MSE is minimized for each estimation, at the cost of reduced comparability between estimates. The baseline specification also imposes a triangular kernel weighting of the observations within the bandwidths, in practice attaching more weight to observations located closer to the cutoff. Lastly, the baseline specification implements local regressions of polynomial order one. In the robustness tests in section 7, these choices are modified to explore the result's sensitivity to different specifications. As the estimates do not differ largely across specifications, I adhere to the least complex specification specified above. This is usually recommended in the RD literature (Lee and Lemieux 2010). To graphically represent the discontinuities, I construct quantile-spaced bins and plot fitted regression lines of polynomial order 1 within the MSE-optimal bandwidths. The quantile-spaced bins are constructed to contain roughly the same number of observations within each bin, making each within-bin mean calculated using similar numbers of observations. The bins are non-overlapping, and the method for choosing bins has the advantage of also giving information about the density distribution of the assignment variable.

Local Average Treatment Effect (LATE). The LATE parameter, denoted τ_{LATE}, captures the average effect of AFP eligibility among the observations being assigned to treatment and complying with the treatment status, i.e. individuals who become eligible for AFP because they fulfill the tenure requirement. Identification of the parameter is closely related to the Instrumental Variable (IV) framework and is estimated using a two-step procedure. Here, the assignment variable can be interpreted as an instrumental variable assigning treatment status to individuals, and the observed recipiency of AFP pension between ages 62 and 67 as a binary proxy variable for being treated.[9] I estimate the model

$$T_i = a_0 + \tau_{FS} I_{x_i>0} + f(x_i) + \delta \mathbf{X}_{it} + \epsilon_{it}, \tag{3}$$

9. This is because I do not observe eligibility directly. Its implications will be discussed in the next subsection.

$$y_{it} = \beta_0 + \tau_{LATE} T_i + f(x_i) + \delta \mathbf{X}_{it} + \epsilon_{it}, \tag{4}$$

where x_i and the indicator variable I is defined as in Equation (2). a_0 and β_0 are intercept terms. For both the above equations, $\mathbf{X}_{it}$ is a vector of covariates, ϵ_{it} is an error term, and f is the unknown functional forms of the assignment variable. The LHS variable T_i of Equation (3) is the observed recipiency of treatment. The coefficient τ_{FS} thus represents the discontinuity at the cutoff in receiving treatment. T_i is used as the regressor of interest in equation 4, and takes the value 1 if the individual is assigned to the treatment condition *and* has received treatment, and 0 otherwise. Hence, the coefficient τ_{LATE} represents the average effect on outcome y at the cutoff for the compliers.

For illustration, an alternative LATE estimator can be computed manually as $\tau_{LATE} = \frac{\tau_{ITT}}{\tau_{FS}}$, using the estimates from Equation (2) and Equation (3), respectively. This resembles the two-stage least squares estimator frequently applied in the IV literature (Cattaneo, Idrobo, and Titiunik 2024). However, a simultaneous estimation is necessary to compute valid confidence intervals for inference and to choose MSE optimal bandwidths specific to the LATE. The model specifications are similar to the baseline ITT model outlined above.

In both models, I use bias-corrected RD estimates and standard errors, providing confidence intervals valid for causal inference (Cattaneo, Idrobo, and Titiunik 2019). The choice results in wider confidence intervals than the conventional estimates and is less sensitive to tuning parameters.

5.4 Internal validity and identification

To conduct causal inference in the RDD framework, a necessary condition is the continuity of the conditional expectations of counterfactual outcomes and predetermined covariates in the assignment variable (McCrary 2008). This implies that two main assumptions must hold. First, covariates must be continuous at the cutoff, i.e. individuals should not differ systematically in any observed predetermined covariates just above and just below the cutoff. Second, individuals must not be able to precisely manipulate their location on the assignment variable.

I start by checking for discontinuities in predetermined covariates. If such discontinuities exist, the assumption that individuals' location around the cutoff only influences treatment status, and nothing else, is damaged. I follow Cattaneo et al (2019) and use coverage-error (CER) optimal bandwidths when estimating RDs for the covariates, as I am only interested in investigating the null hypothesis of no discontinuity in the covariates at the cutoff, with point estimates being of no particular interest.[10] The first columns of Table 1 show RD estimates for all 10 predetermined covariates. Plots are shown in Figure 1A and A2.[11]

None of the covariates exhibit discontinuity at the cutoff, with two exceptions: National background and gender. The discontinuities are significant at the 10 percent and 5 percent levels, respectively. Since the number of covariates included is rather large, however, the probability of observing a discontinuity in one or more covariates is high. A reason for concern still exists because both covariates are strongly correlated to the outcome variables. To account for possible spuriousness stemming from these observed discontinuities, I run the estimations of the ITT effects with covariates included. This both helps to restore identification and increase the precision of the estimates (Frölich and Huber 2019). The point estimates are, overall, rather similar to the baseline specifications. In section 6, I also run a separate analysis for females.

The other key assumption for the internal validity of the RDD is the absence of precise manipulation of the running variable. If such manipulation exists, individuals' treatment is not randomly assigned in the cutoff's neighborhood and the RD estimates will be biased. The standard method for testing the assumption, pioneered by McCrary (2008), is to conduct a density test on the distribution around the cutoff. A significant increase in the number of individuals just to the right of the cutoff indicates manipulation. I start by visually examining the distribution of the assignment variable around

10. As MSE-optimal bandwidths by construction are larger than CER-optimal bandwidths, possible size distortions in point estimates also increase.

11. Note here the definition of predetermined. As individuals might behave according to expectations of future AFP status also before turning 62, covariates capturing characteristics chosen these years might not be fully predetermined. The conditioning of the sample of employment in AFP firms at 62 makes such anticipation effects highly unlikely, however, as will be discussed below.

Table 1: Discontinuity checks and descriptive statistics for covariates

Variable	RD estimate (SE)	Bandwidth	Mean (SD)	N
Annual FT pension	0.019 (.055)	101	2.31 (0.48)	11197
Fraction female	0.083 (0.474)*	102	0.27 (0.44)	11204
Fraction married	0.031 (.048)	110	0.68 (0.47)	11204
Fraction industry worker	-0.07 (0.051)	101	0.4 (0.49)	11204
Fraction Norwegian	0.077 (.039)**	90	0.89 (0.31)	11204
Birth year	-0.14 (0.17)	114	1951.6 (1.75)	11204
Fraction, DI claimed 51-61	0.027 (0.037)	112	0.16 (0.37)	11204
Labor earnings 51-61	-1.3 (3.75)	149	61.5 (38.3)	11204
Educational level	-0.1 (0.17)	91	3.74 (1.47)	11122
Net wealth	-3.13 (4.79)	93	9.4 (32.9)	11204

All standard errors are based on NN matching. *** indicates a significance level of 0.01, ** a significance level of 0.05, and * a significance level of 0.1. Industry=1 if the individual works in mining and industry, the petroleum sector, or construction. The educational level consists of 9 levels representing the highest finished education level. ”51-61” means age 51 to age 61. If not otherwise stated, covariates are determined at age 60. The reported RD estimates and standard errors are conventional. P-values are calculated using bias-corrected RD estimates and standard errors. All bandwidths are chosen by a data-driven procedure for obtaining coverage-error (CER) optimal bandwidths. The monetary amounts are measured in 2014 Gs, approximately corresponding to $14,000.

the cutoff. Histograms are presented in Figure A4. The density across the cutoff appears not to be smooth, with a higher number of observations to the right of the cutoff. There are clear indications of ”seasonal” patterns in the data, however. This is not surprising, as it is common for employment contracts to start and end on the first and last dates of a month. Additionally, Norwegian labor market and social security policies may explain the increased density at ”full” years of tenure. For example, sickness benefits can only be claimed for 52 consecutive weeks before ceasing.

To assess whether the discontinuities in the distribution result from manipulation or seasonal effects, I conduct formal density tests at artificial cutoffs in addition to the density test for the real cutoff. If bunching to the right side of the real cutoff results from “gaming” the system, there is no reason to observe similar bunching to the right of other cutoffs irrelevant for AFP eligibility.

The density test I apply is developed by Cattenoe et al (2020) and is related in spirit to the continuity-based framework of RDDs. The test estimates two separate probability density functions, one using observations to the left of the cutoff and one using observations to the right.[12] Both density functions are evaluated at the specified cutoff. The null hypothesis is that density estimates from the two estimations are not different. Conversely, rejecting the null implies a discontinuity in the density at the cutoff, indicating manipulation in the assignment variable. Table A2 shows the difference in the density estimates evaluated at the 10 artificial cutoffs and the real cutoff, and t-statistics for the significance of the difference. I reject the null of continuity in the distribution at the real cutoff, as predicted from the visual inspection. However, the table shows that discontinuities are present, and equal or larger in size, for all the artificial cutoffs. Thus, I conclude that the observed density discontinuity is unlikely to be the result of precise manipulation of the running variable, but instead reflects natural bunching at full months of tenure.

Another factor making manipulation unlikely is the conditioning of the sample on AFP employment at age 62, the last year of the tenure counting period. A worker in the sample can therefore not quit when having met the tenure requirement and still be eligible. The only possible manipulation strategy is to have an unemployment spell at the start or middle of the tenure period, and then be able to get (re)employed at an AFP firm just in time to make the requirement. This would have required impressive foresight and control over one's employment status. Such foresight is especially difficult among the first cohorts, as the eligibility rules were not known before early 2009, giving little time for strategic maneuvering.

A related problem regarding the employed-at-age-62 condition is the possibility of a selection bias in the sample used for estimation. There might be individuals realizing they were employed "too late" for qualifying for AFP, and hence exit the AFP firm before turning 62. This mechanism can affect the composition of individuals assigned to the control condition, thereby threatening identification. A comparison of firm exits

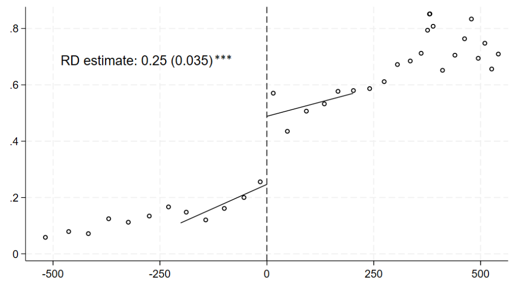
12. As with the RD models, I use triangular kernel weighting. Similarly, the bandwidths used for estimating the density functions are obtained by minimizing the MSEs. Bandwidths are the same on both sides of the cutoff. The density functions uses polynomials of order 2.

for individuals employed just before and just after the date from which continuous employment leads to AFP eligibility could shed light on this possible selection.[13]

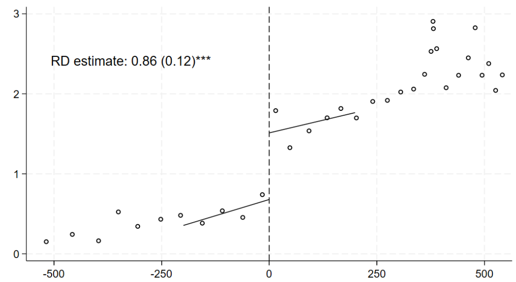
For the identification of the LATE RD estimator, additional assumptions are necessary. The instrument must not affect the outcomes of interest apart from the instrument's effect on AFP status (Cattaneo, Idrobo, and Titiunik 2024). This corresponds to the standard exclusion restriction in the IV framework, and would for example be violated if being just above the cutoff also qualifies for other social security benefits, or is correlated with other factors potentially affecting the outcomes of interest. No obvious reasons for such a correlation exist, but it is impossible to test statistically. For example, the assumption would be violated if some firms rewarded workers with tenure above a certain threshold with higher wages, and this tenure threshold corresponded exactly with the AFP requirement. This is assumed to be highly unlikely. Second, identification requires monotonicity in responses. This would be violated if any individuals would receive treatment when assigned to the control condition but not if counterfactually assigned to the treatment condition, and the converse if assigned the treatment condition. There is no reason to believe this to be the case.

Another possible problem in identifying the LATE arises when the treatment variable exhibits measurement errors. This is likely to be the case in my design because AFP eligibility, and thereby treatment status, is not observed directly. Instead, I use the observed recipiency of AFP as a proxy for eligibility. The proxied treatment might consequently differ from the actual treatment. Ura (2018) examined the implications of identifying the LATE in the presence of measurement error in a binary treatment variable in the IV framework. He shows that the estimated LATE, regardless of the direction of the error, identifies the true LATE effect under an additional assumption to the standard IV assumptions: The instrumental variable must not be correlated to possible measurement errors in the treatment variable. This is violated if deviations between AFP *eligibility* and observed AFP *recipiency* depend on individuals' location of the assignment variable. This could be the case if individuals just above the tenure

13. This could also constitute the basis for an RD analysis on its own, where the start of AFP employment around this cutoff could serve as an instrument (or assignment) variable for AFP eligibility.



(a) Share observed claiming AFP



(b) Total AFP benefits

Figure 1: AFP recipiency ages 62-67

Note: Polynomial fit of order 1 using MSE-optimal bandwidths. Quantile spaced bins. Left-panel RD estimate represents τ_{FS}. X-axis: Number of days of AFP tenure relative to cohort-specific tenure requirement. 3,007 Observations left of the cutoff. 8,197 observations to the right of or at the cutoff. The right panel's Y-axis is in Gs, approximately corresponding to $14,000

cutoff had some reason not to claim AFP despite being eligible, and this reason did not apply similarly to individuals just below the tenure cutoff. Although impossible to rule out, there is no reason to believe such mechanisms exist.[14] Since AFP can be claimed after age 67 and 70 without actuarial punishment, however, there might be individuals in our sample responding to AFP eligibility without being observed receiving AFP, as discussed previously. This leads to the observed take-up of treatment being lower than the actual recipiency of treatment. Hence, the measurement error is likely to be downwards. As a result, the LATE estimator becomes upwards biased (in absolute values). Because of this, the LATE estimator should be interpreted as an upper bound to the true LATE.

6 Results

A necessity for the RDD's ability to identify any causal effect in a fuzzy setting is a significant probability change at the cutoff of receiving treatment. As shown visually in Figure 1 and formally in Table 2, the RD of the jump at the cutoff is .25 for observed

14. If eligible, there are no financial reasons to not claim AFP benefits. The only motivation would be its substitution for DI. However, the share claiming DI pre-age 62 is continuous at the cutoff. Di take up after age 62 is also insignificant when being an outcome in the RD analysis. Hence, there are no indications of AFP benefits being substituted for DI differently for individuals assigned to treatment and control conditions.

private AFP claims, and an increase of .86 Gs in observed AFP benefits claimed. The amount corresponds to an $11,400 increase in total AFP benefits between age 62 and 67.

Two aspects of the relationship between the assignment status and observed AFP recipiency should be noted. As discussed when describing the procedure for constructing the assignment variable, the number of cross-overs is relatively high. Second, it is important to emphasize how compliance is far from absolute to the right of the cutoff. This is partly a product of the naive approach of being "generous" with AFP tenure, but it is also the result of individuals not fulfilling other requirements for AFP eligibility. These reasons are connected to requirements at the time of activation. Ineligibility can stem from an individual's firm no longer being covered by an AFP tariff agreement, unexpected job switches, claiming of disability insurance, or having too small accrued pension-giving income for early retirement. As discussed in the previous section, some eligibles may also have started drawing AFP benefits after age 67.

Table 2 further presents the reduced-form ITT estimates at the cutoff for total labor market earnings, disability insurance recipiency, old age public pensions recipiency from *Folketrygden*, and average weekly hours worked, all between age 62 and 67. In addition, the table's last row shows the share having been observed retired, defined as annual labor income below 1 G, before or during age 67.

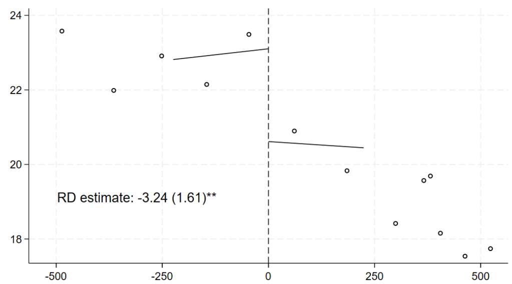
I first focus on ITT effects, for which several statistically significant effects of private sector AFP eligibility can be seen. First, AFP recipients tend to decrease their labor supply between ages 62 and 67 in line with the theoretical predictions from Section 4. The aggregated labor market earnings are reduced with approximately 3 Gs at the cutoff, or $40,000. This is accompanied by a 2-hour reduction in average weekly hours worked, capturing reductions in labor supply at both the extensive and intensive margin. In terms of percentages, the 2-hour reduction in average weekly hours worked represents a 10 percent reduction.[15] The less definitive estimates of pre-68 retirement propensity suggest that much of the reduction is driven by a decrease in hours while

15. For the percentage calculations, the intercept term at the cutoff from the left local polynomial fitted line is used. This captures the counterfactual number of hours worked at the cutoff if treatment-assigned individuals were not assigned to treatment.

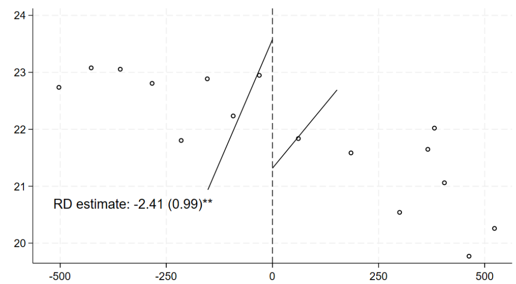
Table 2: RDD estimates. ITT effects and LATE.

Outcome	RD Estimate (SE)	Bandwidth	N
Share claiming private AFP 62-67			
No Covariates	0.25 (0.035)***	201	11204
With Covariates	0.25 (0.034)***	199	11118
Total private AFP claimed 62-67			
No Covariates	0.86 (0.12)***	199	11204
With Covariates	0.89 (0.11)***	184	11118
Total labor earnings 62-67			
No Covariates	-3.24 (1.61)**	224	11204
With Covariates	-2.93 (1.23)**	200	11118
2SLS (IV)	-13.56 (6.9)**	200	11204
Total FT pensions 62-67			
No Covariates	0.89 (0.19)***	149	11204
With Covariates	0.87 (0.16)***	158	11118
2SLS (IV)	3.12 (0.65) ***	164	11204
Total disability insurance 62-67			
No Covariates	-0.39 (0.35)	217	11204
With Covariates	-0.68 (0.28)**	224	11118
2SLS (IV)	-1.34 (1.43)	174	11204
Hours worked 62-67, weekly average			
No Covariates	-2.41 (0.99)**	152	11204
With Covariates	-1.76 (0.87)**	180	11118
2SLS (IV)	-7.35 (3.72)**	195	11204
Share retired pre 68			
No Covariates	0.066 (0.037)*	202	11204
With Covariates	0.06 (0.038)	183	11118
2SLS (IV)	0.28 (0.15)	170	11204

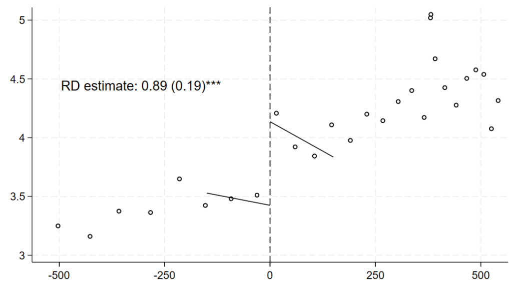
All standard errors are based on NN matching. *** indicates a significance level of 0.01, ** a significance level of 0.05, and * a significance level of 0.1. The reported RD estimates and standard errors are conventional. P-values are calculated using bias-corrected RD estimates and standard errors. All bandwidths are chosen by the data-driven procedure for minimizing the Mean Squared Errors (MSEs). IV estimates are with covariates excluded. Covariates include all variables presented in Table 1. Retired is defined as one year with labor earnings <1G. Effective number of observations (left/right of cutoff) for different bandwidths: 100: 656/824. 150: 880/1137. 200: 1204/1607. 250: 1518/2057. Numbers refer to ages. All monetary variables are measured in 2014 Gs, approximately corresponding to $14,000.



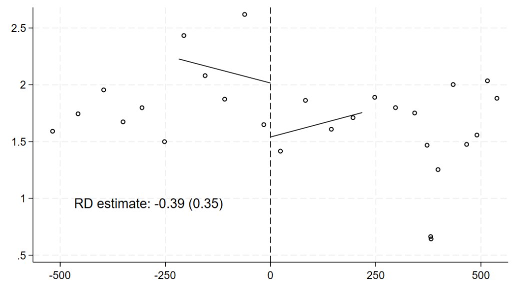
(a) Total labor earnings 62-67



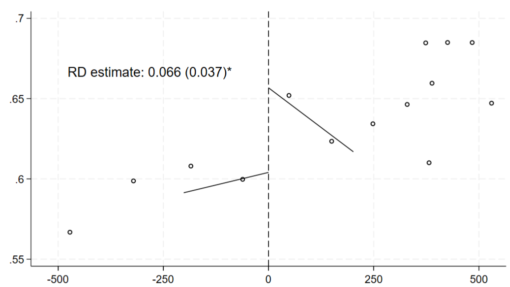
(b) Hours worked 62-67, weekly average



(c) Total FT pensions 62-67



(d) Total disability insurance 62-67



(e) Share retired pre 68

Figure 2: RD Plots of Outcomes

Note: Polynomial fit of order 1 using MSE-optimal bandwidths. Quantile spaced bins. X-axis: Number of days of AFP tenure relative to cohort-specific tenure requirement. 3,007 Observations left of the cutoff. 8,197 observations to the right of or at the cutoff. Numbers refer to ages. All monetary variables are measured in 2014 Gs, approximately corresponding to $14,000.

still in employment and that a gradual substitution of labor earnings with pensions is a more common response than an earlier exit from the labor market. As will be explored further in the next section, the estimates imply large labor supply responses to changes in pension wealth when compared with previous estimates.

Second, AFP recipiency also leads to a large increase in drawings of public old age pensions, amounting to 0.9 G, or approximately $12,000, in cumulative additional pensions claimed between ages 62 and 67. Some of this extra old-age pension benefit at the cutoff is mechanic. As explained in section 4, recipiency of private sector AFP prerequisites simultaneously recipiency of public old age pension from *Folketrygden*. This only requires the minimum amount of FT pension (20 percent of the annual amount), however. Another possible mechanical explanation for the result is the requirement of a minimum level of total pension wealth when first drawing FT benefits. The propensity to exceed this threshold increases with AFP eligibility for all ages before the normal retirement age, thereby relaxing credit constraints on future pension benefits. It is difficult to distinguish between mechanical and "behavioral" effects, but the results suggest a preference for smoothing pension wealth from the later to earlier periods of old age. This is consistent with the theoretical prediction from section 4 of smoothing consumption between periods of continued work and retirement.

Third, and maybe surprisingly, recipiency of disability insurance increases only significantly in the model specification with covariates included. This stands in stark contrast to previous findings about the Norwegian labor market for old-age workers (Kruse and Myhre 2023; Hernæs et al. 2016; Johnsen, Vaage, and Willén 2022), which have shown substitution of pension entitlements for other social security. This has been especially prominent in the context of the old AFP. One would still expect this to be a prevalent mechanism under the new private sector AFP, as claiming disability insurance after age 62 automatically disqualifies for AFP.

Table 2 also depicts the LATE effects from the 2SLS IV estimation. The ITT estimates are roughly multiplied by a factor of four to account for the first step discontinuity in observed APP recipiency of 25 percentage points. The estimates differ slightly from the four-times scaling, because of the selection of new bandwidths for the two-stage estimation procedure.

6.1 Heterogeneity in responses

The results from my main specification show large behavioral effects of AFP eligibility. This subsection digs deeper into potential heterogeneity in responses. Table A4 presents RD estimates for 6 sub-groups of the sample. These are determined based on wealth at age 60, labor earnings before age 62, if the worker is employed in an industry sector at age 60, and gender. The most visible heterogeneity exists between sub-samples determined by wealth and previous labor earnings. For example, while the RD estimate for both labor earnings and hours worked between ages 62 and 67 is insignificant for the sub-samples containing individuals with wealth at the bottom 75 percent of the wealth distribution and the bottom 50 percent of the previous labor earnings distribution, the estimated reduction is large and significant for the relatively well off. This might be counterintuitive at first glance, as AFP eligibility represents a larger percentage increase in pension wealth for the former group compared to the latter. Thus, one would also expect these to be tighter constrained in the absence of AFP and respond stronger as this constraint is relaxed with AFP eligibility. A couple of mechanisms can still help rationalize the result theoretically. In section 4, I discussed how agents may exhibit non-convexities in preferences, for example due to a fixed cost of work. The initially higher wealth levels of the relatively well-off may cause these individuals to be more inclined to adjust their labor supply as they are, at the margin, closer to the decision of withdrawing from the labor market.

Another possible explanation for the heterogeneity concerns individuals' health status, a confounding factor between financial resources and labor supply. It is well known that financial prosperity and health status are strongly correlated. At the same time, health is believed to have a strong impact on old-age labor supply. Most commonly, this relationship has been thought to be negatively correlated, i.e. worsened health status negatively impacts old-age labor supply. Hasselhorn et al (2022) note that several studies have found the opposite correlation. Good health status might instead induce earlier retirement and lower labor supply, motivated by the wish to fulfill life goals while still healthy, especially when combined with financial freedom. This is often accompanied by financial freedom or "affording" to reduce labor supply, which increases with AFP eligibility. This mechanism might be especially prevalent in my research design, as the sample is conditioned on employment at age 62. This reduces the probability of permanently reduced work capacity due to poor health status. It should be mentioned that Cesarini et al (2017) also find a stronger labor supply response among the financially better off in their study on lottery winners. However, the result is not explored in detail by the authors.

While no eye-catching heterogeneity exists based on the occupational sector, the last columns of Table A4 indicate heterogeneity across genders. This is unsurprising, as larger labor supply elasticities for females than men is a common result in the literature (see e.g. Keane (2011) who investigates differences in wage elasticities between men and women).

6.2 Robustness checks

As discussed in Section 5, the result's validity relies on correct model specifications. I therefore check the results' sensitivity to different specification choices. The first check investigates sensitivity to different model specifications. Table A4 compares the baseline ITT estimates to estimates from separate RDD regressions with a local 2nd-order polynomial and a uniform kernel function for weighting of observations. The results exhibit little sensitivity to the choice of the kernel function, but some non-negligible sensitivity to the choice of polynomial order. The second-order polynomial specifications yield larger point estimates, but also higher variance, compared to the baseline model. All estimates remain constant in terms of signs, however. A reassuring finding when comparing the two polynomial choices is that the second-order specification increases the point estimates for both observed AFP recipiency (τ_{FS}) and relevant outcomes (τ_{ITT}). This strengthens the claim that discontinuities in the outcome variable indeed reflect changes in AFP status and not other factors.[16]

16. Or Using the IV framework from section 5, the co-movement in τ_{FS} and τ_{ITT} makes sure the LATE estimate remains similar across polynomial specifications.

Figure A3a shows the sensitivity of τ_{FS} to different bandwidths. Except for very small bandwidths, which create large size distortions of the point estimates as discussed regarding CER versus MSE-optimal bandwidths, the point estimate stays relatively robust to different bandwidth choices. Figure A3b shows the point estimates from so-called donut analyses. Here, observations close to the cutoff are dropped from the estimations in different increments. The point estimate stays robust across the different donut radiuses, indicating a negligible degree of sensitivity to observations in the immediate range of the cutoff. This is reassuring when considering the naive procedure for constructing the running variable, where all individuals were granted 15 days additional AFP tenure to account for possible differences in actual and assumed dates of birth. If this choice had affected the relationship between assignment and observed treatment significantly, the point estimates would have been sensible (most likely increased) when increasing the donut radius.

The last robustness check investigates the τ_{FS} estimate's sensitivity to artificial cutoffs. For this test, separate RD estimations are conducted with the tenure requirement being specified to deviate from the actual requirement in steps of 10 days. The estimations only utilize observations on the control side of the treatment when the artificial cutoff is below zero, and on the treatment side for artificial cutoffs above zero. As can be seen in Figure A5, we do not see significant discontinuities unequal zero for any artificial cutoffs except for the first positive increment. For the smallest steps, however, the effective numbers of observations used in the estimations are low, and the statistical power decreases substantially, as reflected in the standard errors for the artificial cutoff 10 days stricter than the tenure requirement. In conclusion, all the additional robustness checks show that the baseline specification exhibits only a limited degree of sensitivity to the specific model choices and that the method for constructing the assignment variable can be said to be rather robust.

7 Discussion

The previous section highlighted how the AFP scheme strongly impacts old-age labor market behavior, and the pension wealth effect is large compared to previous estimates in the literature. As an exercise to generalize from the results, I perform a back-of-the-envelope calculation of the old age labor supply elasticity with respect to total pension wealth. My parameter of interest is $\varepsilon_{h,P} = \frac{\Delta hours_{oa}}{\Delta P}$, where $\Delta hours_{oa}$ is the average percentage change in weekly hours worked on average between age 62 and 67 caused by AFP eligibility, and ΔP is the average percentage increase in total pension wealth with AFP eligibility. I calculate the denominator based on average calculated FT benefits and (potential) AFP benefits among the individuals within the bandwidths used for the RD estimation.[17] Since both pensions are indexed the same way and with lifelong duration, the percentage change in NPV pension wealth can be computed directly. Intuitively, the elasticity captures the percentage change in average weekly hours worked between the early retirement age of 62 and the normal retirement age of 67 as a response to a one percent increase in the total (NPV) pension wealth. The percentage increase in pension wealth resulting from AFP eligibility is 18 percent and practically the same across all relevant bandwidths. The percentage change in hours is calculated using the intercept term at the cutoff from the left-fitted polynomial line, capturing the counterfactual hours worked at the cutoff.[18]

The elasticity estimates differ depending on whether I use the ITT or LATE estimates for the calculation. Preferably, I would use the LATE to capture the effect of the treatment for the compliers. As the estimated LATE should be interpreted as the upper bound of the true LATE, I calculate elasticity using both ITT and LATE estimates.

The LATE estimates correspond to a 33 percent reduction in average weekly hours worked between ages 62 and 67 at the cutoff, yielding an elasticity of -1.68. When I instead plug in the ITT estimate, the reduction in average weekly hours worked is 10 percent and the elasticity is -0.54. Some considerations should be noted regarding

17. Both calculated for age 60. The average annual FT pension is 2.31 Gs and the average annual potential AFP benefit is 0.42.

18. The left-intercept is 23.81 weekly average hours worked.

the exercise. Pension wealth is not fixed at age 62, since FT accrual continues after age 62. Thus, the elasticity should be interpreted statically, meaning that it captures future responses to an event occurring at age 62, keeping everything else constant. A further complicating factor is that private pension plans and other wealth are excluded, likely underestimating the elasticities as the true denominator is smaller than in my calculation.[19] Keeping these limitations in mind, the estimated elasticity range still represents surprisingly large elasticity estimates (in absolute values), mostly exceeding those earlier reported in both the pension and lottery literature, as presented in section 3.

7.1 Welfare implications

The last point of discussion in this thesis puts the reduced-form estimates to use in a simple welfare analysis. I draw on the policy elasticity framework developed by Hendren (2016) to assess the welfare effect of the AFP scheme. The advantage of Hendren's approach is that it only requires the policy's impact on government revenues, coined by Hendren as the *policy elasticity*. In the framework, the welfare effects of a policy can be represented by the marginal value of public funds (MVPF). The MVPF is a simple benefit/cost ratio, defined as $\frac{WTP}{1+FE}$, where WTP is the willingness to pay for the good or transfer provided by the policy, and FE is the impact of the policy's behavioral response to the government budget per dollar of fiscal expenditures, or the fiscal externality. The denominator hence represents the policy's net fiscal cost. The intuitive interpretation of the statistic is that it measures the amount of welfare that can be delivered to policy beneficiaries per dollar of government spending on the policy.[20]

This parameter is straightforward to compute using my reduced-form estimates.[21] For the examined cohorts, life expectancy at age 62 was 84. This means that the

19. If also including other social security in a broader social security wealth definition, the point estimates for increased social security wealth at the cutoff may lead to a larger denominator and hence move the elasticity estimates towards zero.

20. The MVPF can conversely be interpreted as the shadow price of raising revenue from the beneficiaries of the policy by reducing spending on the policy.

21. In this exercise, I also disregard that 2 thirds of AFP are financed by firms similarly to a payroll tax. This can have fiscal implications on its own, potentially different from other taxes used to fund government policies.

total average (NPV) possible AFP pension wealth, calculated as the average within the bandwidths used for estimation, corresponds to 9.38 Gs.[22] This represents the average mechanical expenditures of the private AFP scheme per recipient. To calculate the behavioral effect of the scheme, I assume the marginal tax rate on labor income in our sample to be 34 percent,[23] and abstract from other taxation possibly affected by responses to AFP eligibility. Using the estimates for the reduction in labor earnings caused by AFP eligibility, a dollar paid out in AFP benefits causes a reduction in collected taxes of 11 cents ($\frac{3.24Gs \times 0.34}{9.38Gs}$) using the ITT estimate, and 49 cents ($\frac{13.56Gs \times 0.34}{9.38Gs}$) using the LATE estimates. The corresponding MVPFs are 0.9 and 0.67, respectively.

The calculated MPVFs are informative when assessing the social desirability of the AFP scheme. As a reference point, a simple non-distortionary transfer has an MVPF of 1. Assuming constant WTPs across policies, both estimated MVPFs of the pension scheme fall below such a neutral benchmark, and raising one dollar in revenues from AFP beneficiaries would be socially desirable in comparison. This assessment rests on a couple of assumptions, however. First, the willingness to pay for private sector AFP among the beneficiaries must not substantially exceed 1. This might not hold for several reasons, e.g. by AFP allowing for pension smoothing or an indirect insurance value, leading to a social marginal value of the scheme higher than one.[24] Second, one of MVPF's advantages is its allowance for comparisons across different policies. To assess the social desirability of the AFP, we would need to know the MVPFs of alternative policies, the willingness to pay among the alternative policies' beneficiaries, and the associated fiscal costs. Hendren and Sprung-Keyser (2020) have shown that MVPFs vary substantially across the ages of beneficiaries with MVPFs being highest for policies directed at health and education for low-income children. This results from

22. The average potential AFP annual benefit (calculated at age 60) is 0.427 G.

23. This is the marginal income tax rate for the average annual labor income of 5.24 G in my sample.

24. The point about insurance value draws upon Baily's (1978) and later Chetty's (2006) influential model of the optimal level of unemployment insurance (UI), where optimal UI increases with (potential) beneficiaries' risk aversion. Kruse and Myhre (2023) used the Baily-Chetty framework to investigate the social desirability of the old private sector AFP and found that the implied risk-aversion must be very large for the scheme to be socially optimal. The exercise suited their research question better than mine, however, as the old AFP closer resembled UI. This was particularly prominent for their sample with elderly *displaced* workers.

the fiscal externalities often being positive for such policies. Hence, we can assume that the AFP scheme, targeted at workers approaching retirement, fares poorly against policies directed towards younger beneficiaries.

The calculation of the MVPF can be extended to also include the estimated reduction in DI benefits. Although insignificant in the baseline specification, including the point estimates for the decrease in DI yields MVPFs of 0.92 and 0.74, using ITT and upper-bund LATE estimates, respectively.[25] This exercise should be interpreted with caution, as I have only used the labor earnings crowd out among the individuals included in my sample. I have abstracted from possible spillover effects to spouses and family members and labor earnings responses after age 67. Both of these effects are likely to increase the actual negative fiscal externality of the scheme, and thus reduce the calculated MVPFs, and increase the propensity for a reduction of the AFP scheme to be welfare improving.[26] As is common when examining the social desirability of policies targeting population groups differently, the weighting of an individual's utilities matters. A preference for redistribution, for example, would affect the desirability of potential policy changes.

8 Conclusion

In this thesis, I have investigated the old-age labor market responses to private sector AFP eligibility, a supplementary pension scheme in Norway. Using administrative data on the full Norwegian population, I exploited the tenure requirement for AFP eligibility in an RDD to obtain estimates of the scheme's causal effect on old-age labor market behavior. As explained in section 2, the pension scheme is actuarially fair and can be freely combined with work. Thus, eligibility for the scheme does not affect returns to work, meaning all labor market responses can be attributed to a pension wealth effect. Section 3 gave an overview of previous literature investigating unearned income

25. The point estimate for the decrease in DI represents a fiscal saving from the government's point of view. This saving is plugged directly into the calculation of the fiscal externality. Using the ITT estimates for illustration, this becomes ($\frac{3.24Gs \times 0.34 - 0.39}{9.38Gs}$).

26. The MVPFs generally measure welfare implications at the margin, and should thus be interpreted as welfare effects caused by small deviations from *status quo*.

and wealth effects on labor supply, while section 4 outlined a conceptual framework providing theoretical predictions of likely labor supply responses to AFP eligibility. An increase in individuals' NPV pension wealth was, *ceteris paribus*, expected to decrease the old-age labor supply. This effect was investigated and quantified in the empirical part of this thesis.

The results showed strong labor market responses to AFP eligibility among old-age workers. I found a significant negative effect on hours worked and labor earnings between age 62 and age 67. The results further indicated an increased propensity to retire before age 68. Another key finding was the increased take-up of public old-age pensions during the earlier parts of old age, which suggests a preference for pension wealth smoothing. Interestingly, the recipiency of Disability Insurance did not decrease significantly. This contrasts with previous literature examining the effects of (the old) AFP, having shown extensive substitution of foregone AFP benefits for other social security. When assessing heterogeneity in responses, I found that much of the labor supply responses were driven by relatively affluent workers. This is surprising and calls for a deeper investigation of the mechanisms which through AFP affect labor market behavior.

The size of the labor supply reduction caused by AFP eligibility depends on whether I estimate ITT or LATE effects. I calculate an old-age labor supply elasticity with respect to pension wealth ranging between -0.54 and -1.68, indicating large wealth effects compared to comparable elasticities previously estimated. The results were put to use in a simple welfare analysis, providing a measure for assessing the social desirability of the AFP scheme. To further develop our understanding of the effects of the AFP scheme, the analysis could be extended to include possible spillover effects, especially regarding spouses and family members.

A few notes of caution should be emphasized when interpreting the results. The two most important assumptions for the internal validity of the RDD are the lack of precise manipulation of the assignment variable and no systematic differences in predetermined covariates at the cutoff. I argue that what looks like manipulation at first glance, instead stems from seasonal patterns in the tenure distribution. Moreover, two covariates are discontinuous at the cutoff. Although possibly the result of chance, I run all ITT estimations with the inclusion of covariates. The estimates remain mostly similar to those of the baseline specifications. Reassuringly, the results appear to be robust across other standard checks for robustness.

An important question concerns the results' generalizability. The RDD identifies, by construction, treatment effects around the cutoff, and identified effects are hence local by nature. Lee and Lemieux (2010) nevertheless note how RD estimates borrow strength from observations away from the cutoff, determined by the weighting procedure and bandwidth selection, and thus exhibit some internal generalizability away from the cutoff. Regarding out-of-sample generalizability, some notes should be made. Different types of unearned income might lead to different types of behavioral effects, e.g. because of social perception of acceptable responses to different types of income. Nor is it clear that the estimated labor supply is constant across different pension wealth levels and program intensity. The results still provide the general insight that labor market responses to pension wealth changes are substantial and that future income streams, which pension wealth in practice is, get incorporated in agents' labor market decisions.

The results presented in this thesis provide several implications for the design of pension systems and future policymaking. In 2025, the public sector AFP will be changed along the lines of the private sector AFP. In this regard, knowledge about the behavioral effects of the private sector AFP is highly relevant. The results also have implications for how to think about pension systems more broadly, both in Norway and elsewhere. A widespread challenge across industrialized economies is the future economic sustainability of pension systems. A commonly shared goal among policymakers is to induce longer careers and later retirement. This thesis has demonstrated the crucial role benefit generosity plays in achieving this. Previous research on pension systems, and pension reforms, has largely focused on the incentives embedded in pension systems. This thesis makes a case for also devoting attention to pension wealth levels.

As a last point, this thesis might provide methodological inspiration for future research. Earlier literature has mostly explored variation stemming from reforms. This complicates the task of investigating the impact of systems' different components in isolation. This thesis exploited variation within rather than between different pension regimes. The methodological approach has the potential to answer new types of questions, especially when for pension system designs exhibiting attributes well-suited for the application of econometric methods of recent times.

References

AFP, Fellesordningen for. 2024. “Vilkår for AFP [Requirements for AFP].” Accessed April 29, 2024. https://www.afp.no/soker/vilkar-for-afp/.

Artmann, Elisabeth, Nicola Fuchs-Schündeln, and Giulia Giupponi. 2023. “Forward-Looking Labor Supply Responses to Changes in Pension Wealth: Evidence from Germany.” IZA Discussion Paper: 16132.

Baily, Martin Neil. 1978. “Some aspects of optimal unemployment insurance.” *Journal of Public Economics* 10 (3): 379–402.

Becker, Sebastian, Hermann Buslei, Johannes Geyer, and Peter Haan. 2022. “The Effect of Pension Wealth on Employment.” IZA Discussion Paper: 15836.

Blundell, R., E. French, and G. Tetlow. 2016. “Chapter 8: Retirement Incentives and Labor Supply.” In *Handbook of the Economics of Population Aging,* 1:457–566. Elsevier.

Bratberg, Espen, Tor Helge Holmås, and Øystein Thøgersen. 2004. “Assessing the effects of an early retirement program.” *Population Economics* 17 (3).

Brinch, Christian N., Erik Hernæs, and Zhiyang Jia. 2017. “Salience and Social Security Benefits.” *Journal of Labor Economics* 35 (1): 265–297.

Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. 2019. *A Practical Introduction to Regression Discontinuity Designs: Foundations.* 1st ed. Cambridge University Press.

———. 2024. *A Practical Introduction to Regression Discontinuity Designs: Extensions.* Cambridge University Press.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2020. “Simple Local Polynomial Density Estimators.” *Journal of the American Statistical Association* 115 (531): 1449–1455.

Cesarini, David, Erik Lindqvist, Matthew J. Notowidigdo, and Robert Östling. 2017. “The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries.” *American Economic Review* 107 (12): 3917–3946.

Chetty, Raj. 2006. “A general formula for the optimal level of social insurance.” *Journal of Public Economics* 90 (10): 1879–1901.

Coile, Courtney, and Jonathan Gruber. 2007. “Future Social Security Entitlements and the Retirement Decision.” *Review of Economics and Statistics* 89 (2): 234–246.

Costa, Dora L. 1995. “Pensions and Retirement: Evidence from Union Army Veterans.” *The Quarterly Journal of Economics* 110 (2): 297–319.

French, Eric. 2005. “The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour.” *The Review of Economic Studies* 72 (2): 395–427.

Frölich, Markus, and Martin Huber. 2019. "Including Covariates in the Regression Discontinuity Design." *Journal of Business & Economic Statistics* 37 (4): 736–748.

Gelber, Alexander M, Adam Isen, and Jae Song. 2016. "The Effect of Pension Income on Elderly Earnings: Evidence from Social Security and Full Population Data." Goldman School of Public Policy Working Paper.

Hasselhorn, Hans Martin, Taina Leinonen, Ute Bültmann, Ingrid Sivesind Mehlum, Jean-Baptist Du Prel, Sibel Kiran, Nicole Majery, Svetlana Solovieva, and Astrid De Wind. 2022. "The differentiated roles of health in the transition from work to retirement – conceptual and methodological challenges and avenues for future research." *Scandinavian Journal of Work, Environment & Health* 48 (4): 312–321.

Hendren, Nathaniel. 2016. "The Policy Elasticity." NBER Working Paper: 19177.

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. "A Unified Welfare Analysis of Government Policies*." *The Quarterly Journal of Economics* 135 (3): 1209–1318.

Hernæs, Erik, Simen Markussen, John Piggott, and Knut Røed. 2016. "Pension reform and labor supply." *Journal of Public Economics* 142:39–55.

Imbens, Guido W., Donald B. Rubin, and Bruce I. Sacerdote. 2001. "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players." *American Economic Review* 91 (4): 778–794.

Johnsen, Julian Vedeler, Kjell Vaage, and Alexander Willén. 2022. "Interactions in Public Policies: Spousal Responses and Program Spillovers of Welfare Reforms." *The Economic Journal* 132 (642): 834–864.

Keane, Michael P. 2011. "Labor Supply and Taxes: A Survey." *Journal of Economic Literature* 49 (4): 961–1075.

Krueger, Alan B, and Jörn-Steffen Pischke. 1992. "The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation." *Journal of Labor Economics* 10 (4): 412–437.

Kruse, Herman. 2021. "Joint Retirement in Couples: Evidence of Complementarity in Leisure." *The Scandinavian Journal of Economics* 123 (3): 995–1024.

Kruse, Herman, and Andreas Myhre. 2023. "Early retirement provision for elderly displaced workers." SSB Discussion Papers: 985.

Laun, Lisa, and Mårten Palme. 2022. "Pension reform, incentives to retire and retirement behavior: empirical evidence from Swedish micro-data." IFAU. Working paper 2022:8.

Lee, David S, and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281–355.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics,* The regression discontinuity design: Theory and applications, 142 (2): 698–714.

NOU 2022:7. 2022. *Et forbedret pensjonssystem [An improved pension system].* Norges offentlige utredninger [Official Norwegian reports]. Arbeids- og inkluderingsdepartementet [Ministry of labor / social inclusion].

OECD. 2024. "Pension spending." Accessed April 16, 2024. https://data.oecd.org/socialexp/pension-spending.htm.

Røed, Knut, and Fredrik Haugen. 2003. "Early Retirement and Economic Incentives: Evidence from a Quasi-natural Experiment." *LABOUR* 17 (2): 203–228.

Snyder, Stephen E., and William N. Evans. 2006. "The Effect of Income on Mortality: Evidence from the Social Security Notch." *The Review of Economics and Statistics* 88 (3): 482–495.

Statistics Norway. 2024. "Tilknytning til arbeid, utdanning og velferdsordninger [Attachment to work, education and social security]." Accessed April 29, 2024. https://www.ssb.no/arbeid-og-lonn/sysselsetting/statistikk/tilknytning-til-arbeid-utdanning-og-velferdsordninger.

Appendix

Table A1: Tenure requirements

Birth cohort	Year turning 62	Tenure requirement
1949-1951	2011-2013	3/5
1952	2014	4/6
1953	2015	5/7
1954	2016	6/8

Source: Fellesordningen for AFP (2024).

Table A2: Continuity tests of density of assignment variable for real and artificial cutoffs

Month relative to real cutoff	| Difference | × 1000	t-stat	Bandwidth
-5	0.75125	-5.010	64
-4	0.56029	-4.655	84
-3	0.70153	-5.152	85
-2	1.67015	-8.888	58
-1	1.19611	-6.516	52
0	**0.95067**	**-4.632**	**49**
1	1.65165	-8.631	48
2	0.88660	-7.189	47
3	2.21682	-10.129	43
4	2.08377	-10.371	45
5	1.71935	-9.022	57

Relative months are constructed using days × 30.4. The second column shows the absolute difference multiplied by 1000 for the density evaluated at the specified cutoffs for separate estimated PDFs using 2nd order and triangular kernels. Formally, difference is $\hat{f}_{-}(\overline{x}) - \hat{f}_{+}(\overline{x})$, where $\overline{x}$ is the specified cutoff. The critical t-value for rejecting $H_0 : \hat{f}_{-}(\overline{x}) = \hat{f}_{+}(\overline{x})$ is 1.96.

Table A3: RD ITT estimates for subsamples. Bandwidths in <>

	Wealth >75 percentile		**Labor Earnings >Median**	
	No	Yes	No	Yes
Share claiming AFP 62-67	.28 (.045)***	.22 (.078)**	.22 (.052)***	.32 (.061)***
	<158>	<163>	<168>	<126>
Total labor earnings 62-67	-1.36 (1.6)	-11.28 (4.64)**	-1.11 (1.11)	-8.52 (3.18)***
	<190>	<144>	<192>	<161>
Hours worked 62-67, weekly avg	-1.56 (1.09)	-3.84 (1.8)**	-1.61 (1.26)	-2.61 (1.22)**
	<169>	<173>	<176>	<207>
Total FT pension 62-67	1.09 (.262)***	.712 (.34)*	.60 (.24)**	.95 (.25)***
	<108>	<168>	<189>	<162>
Share retired pre 68	.043 (.043)	.122 (.074)*	.031 (.052)	.1(.052)*
	<199>	<218>	<177>	<227>
Total DI 62-67	-.567 (.474)	.888 (.54)*	-.31 (.623)	-.22 (.38)
	<175>	<188>	<167>	<199>
N	8403	2801	5601	5603

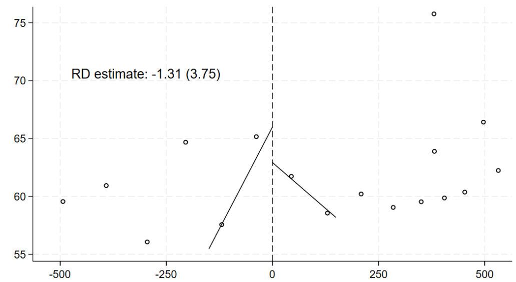
	Work in Industry		**Gender**	
	No	Yes	Male	Female
Share claiming AFP 62-67	.238 (.044)***	.289 (.05749)***	.273 (.039)***	.19 (.07)**
	<194>	<217>	<218>	<214>
Total labor earnings 62-67	-4.155 (2.27)*	-2.9451 (3.139)	-3.03 (2.22)	-2.52 (1.83)
	<144>	<165>	<178>	<154>
Hours worked 62-67, weekly avg	-3.22 (1.24)***	-.618 (1.472)	-1.07 (1.03)	-3.33 (1.77)*
	<153>	<192>	<188>	<154>
Total FT pension 62-67	.718 (.23)***	1.128 (.292)***	.855 (.23)***	1.31 (.30)***
	<165>	<167>	<135>	<205>
Share retired pre 68	.058 (.049)	.097 (.061)	.039 (.045)	.108 (.063)
	<180>	<194>	<187>	<237>
Total DI 62-67	-.21 (.52)	-.737 (.548)	-.426 (.40)	-.786 (.75)
	<163>	<159>	<181>	<246>
N	6772	4432	8173	3031

All standard errors are based on NN matching. *** indicates a significance level of 0.01, ** a significance level of 0.05, and * a significance level of 0.1. Industry=yes if the individual works in mining and industry, the petroleum sector, or construction. The reported RD estimates and standard errors are conventional. P-values are calculated using bias-corrected RD estimates and standard errors. All bandwidths are chosen by the data-driven procedure for minimizing the Mean Squared Errors (MSEs). Numbers refer to ages. The monetary amounts are measured in 2014 Gs, approximately corresponding to $14,000.

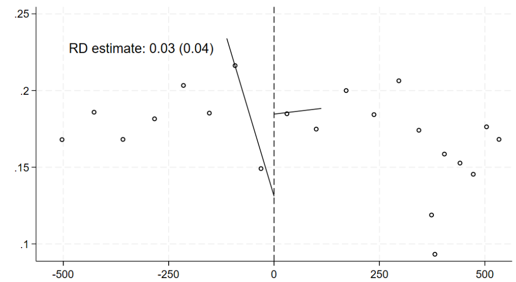
Table A4: Sensitivity to model specifications. ITT-estimates

	RD estimate (SE)	**Bandwidth**
Share claimed AFP 62-67		
Baseline	0.252 (0.035)***	201
Polynomial order 2	0.393 (0.059)***	133
Uniform Kernel	0.261 (0.039)***	135
Total labor earnings 62-67		
Baseline	-3.241 (1.607)**	224
Polynomial order 2	-4.363 (2.291)**	211
Uniform Kernel	-3.358 (1.575)*	211
Hours worked 62-67, weekly average		
Baseline	-2.41 (0.99)**	152
Polynomial order 2	-3.180 (1.274)**	184
Uniform Kernel	-2.029 (0.977)*	127
Total DI 62-67		
Baseline	-0.351 (0.529)	217
Polynomial order 2	-0.517 (0.4)	184
Uniform Kernel	-0.398 (0.398)	132
Share Retired pre 68		
Baseline	0.066 (0.037)*	201
Polynomial order 2	0.089 (0.051)	200
Uniform Kernel	0.068 (0.038)*	162
Total FT pensions 62-67		
Baseline	0.886 (0.194)***	149
Polynomial order 2	1.181 (0.259)***	158
Uniform Kernel	0.762 (0.162)***	139

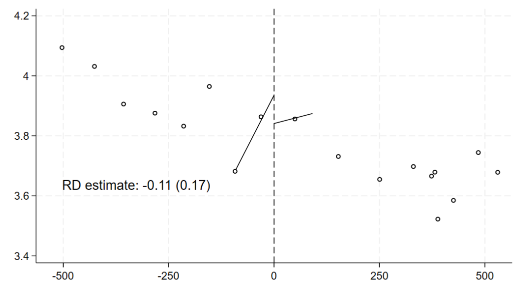
All standard errors are based on NN matching. *** indicates a significance level of 0.01, ** a significance level of 0.05, and * a significance level of 0.1. The reported RD estimates and standard errors are conventional. P-values are calculated using bias-corrected RD estimates and standard errors. All bandwidths are chosen by the data-driven procedure for minimizing the Mean Squared Errors (MSEs). Numbers refer to ages. The monetary amounts are measured in 2014 Gs, approximately corresponding to $14,000.



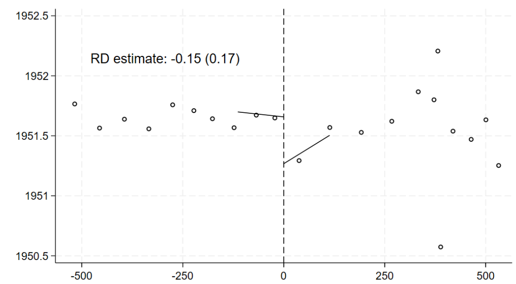
(a) Labor earnings 51-61



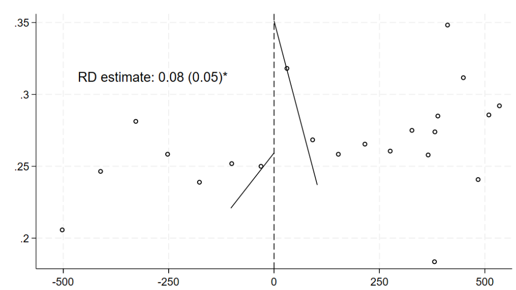
(b) DI claimed 51-61



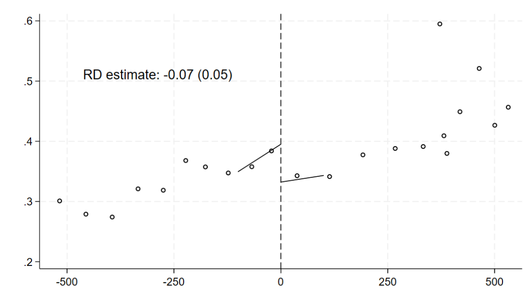
(c) Educational level 60



(d) Year born



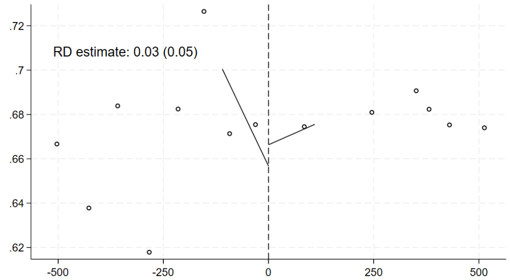
(e) Share female



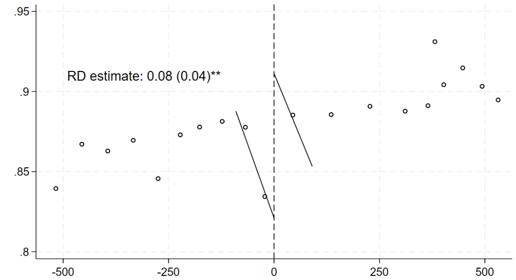
(f) Share industry workers 60

Figure A1: RD Plots for covariates. Part 1

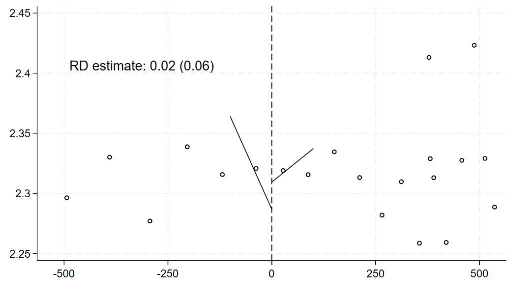
Continues on next page



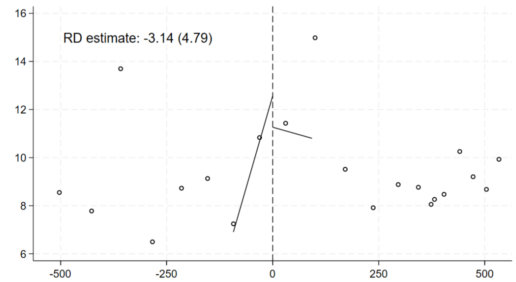
(a) Share married 60



(b) Share Norwegian



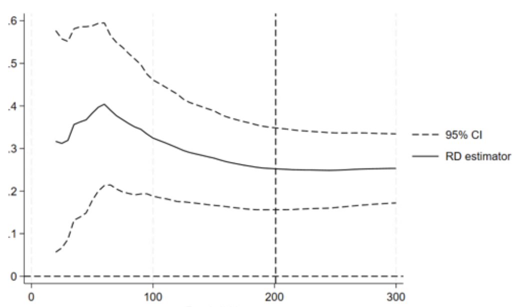
(c) FT benefit level



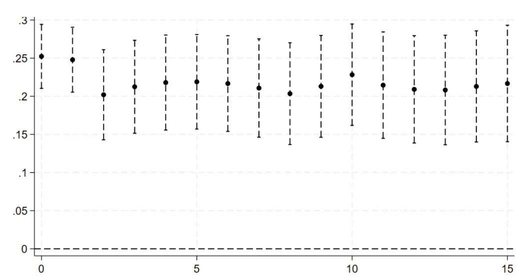
(d) Net wealth 60

Figure A2: RD Plots for covariates. Part 2

Note: Polynomial fit of order 1 using CER-optimal bandwidths. Quantile spaced bins. X-axis: Number of days of AFP tenure relative to cohort-specific tenure requirement. 3,007 Observations left of the cutoff. 8,197 observations to the right of or at the cutoff. All monetary variables are measured in 2014 Gs, approximately corresponding to $14,000.



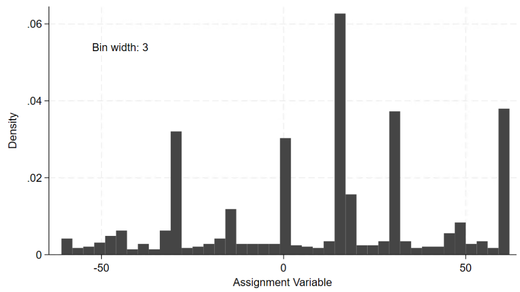
(a) Sensitivity to different bandwidths



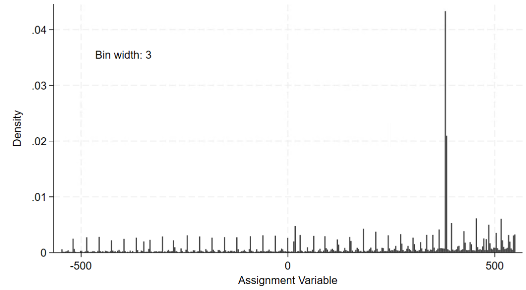
(b) Donut analysis

Figure A3: Robustness checks: RD estimates with AFP recipiency 62-67 as outcome

Note: Left panel shows sensitivity to different bandwidths and 95 percent CI. The dotted line is the MSE optimal bandwidth. The right panel shows RD estimates sensitivity to different donut radiuses. The X-axis depicts values of the assignment variable to both sides of the real cutoff omitted from the sample used for estimation.



(a) Assignment variable <62



(b) Full sample

Figure A4: Density of assignment variable

Note: Histogram showing the relative frequency of observations location on the assignment variable. 3 days within each bin.

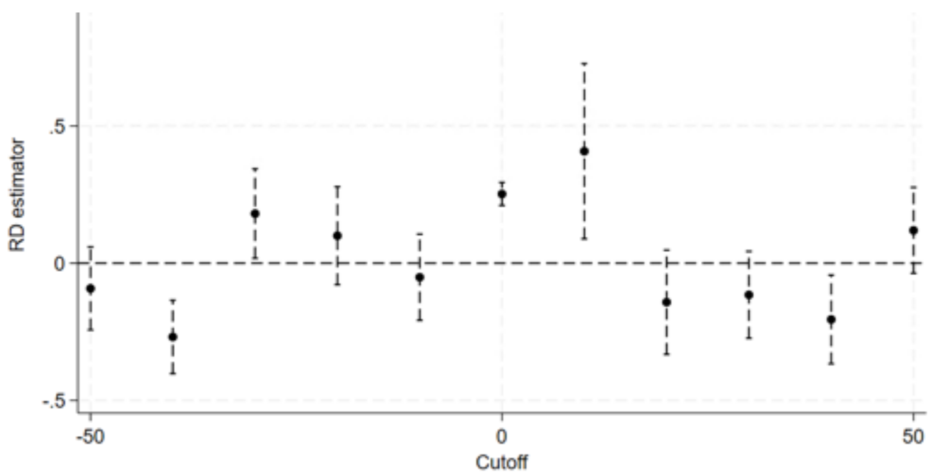


Figure A5: RD estimates for artificial cutoffs. Outcome: AFP recipiency 62-67

Note: RD estimations using only observations to the left side of the real cutoff when the specified cutoff is below zero. Using only observations to the right side of the real cutoff when the specified cutoff is above zero.