ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Andersen, Asbjørn Goul; Markussen, Simen; Røed, Knut

Working Paper

Pension Reform and the Efficiency-Equity Trade-Off: Impacts of Removing an Early Retirement Subsidy

IZA Discussion Papers, No. 12918

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Andersen, Asbjørn Goul; Markussen, Simen; Røed, Knut (2020) : Pension Reform and the Efficiency-Equity Trade-Off: Impacts of Removing an Early Retirement Subsidy, IZA Discussion Papers, No. 12918, Institute of Labor Economics (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/215314

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12918

Pension Reform and the Efficiency-Equity Trade-Off: Impacts of Removing an Early Retirement Subsidy

Asbjørn Goul Andersen
Simen Markussen
Knut Røed

JANUARY 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12918

Pension Reform and the Efficiency-Equity Trade-Off: Impacts of Removing an Early Retirement Subsidy

Asbjørn Goul Andersen
The Ragnar Frisch Centre for Economic Research

Simen Markussen
The Ragnar Frisch Centre for Economic Research

Knut Røed
The Ragnar Frisch Centre for Economic Research and IZA

JANUARY 2020

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Pension Reform and the Efficiency-Equity Trade-Off: Impacts of Removing an Early Retirement Subsidy*

We provide empirical evidence that the removal of work disincentives embedded in retirement earnings tests can increase old-age labor supply considerably, but it does so at the cost of more income inequality. Causal effects are identified based on a reform of the Norwegian early retirement program, which entailed that adjacent birth cohorts were exposed to completely different work incentives from age 62. The reform removed a strict retirement earnings test such that pension wealth was redistributed from early to late retirees. Given the pre-existing employment and earnings patterns, this implied a considerable rise in old-age income inequality. In principle, this could have been offset by changes in the labor supply. We estimate that the reform triggered a 42% increase in hours worked during the ages covered by early retirement options. However, as the labor supply responses were of similar magnitudes across the earnings distribution, they did little to offset the rise in inequality. As measured by the Gini coefficient, inequality in overall old-age income rose by approximately 0.03 (17%).

JEL Classification: H55, D31, J22, J26

Keywords: pension reform, inequality, labor supply

Corresponding author:
Knut Røed
Ragnar Frisch Centre for Economic Research
Gaustadalleen 21
N-0349 Oslo
Norway
E-mail: knut.roed@frisch.uio.no

* The research has received support from the Norwegian Research Council (grant # 270875). Thanks to Erik Hernæs and Axel West Pedersen, and to seminar participants in Bergen, Oslo, and Ås for valuable comments and discussions. Administrative registers made available by Statistics Norway have been essential.

1 Introduction

In recent years, many developed countries have reformed their pension systems to address the rising fiscal costs of population ageing. A key element in many of these reforms has been to encourage senior workers to postpone retirement. One strategy for achieving this is to remove the earnings test on pension income, such that workers above the threshold age for early retirement maintain strong incentives to work. This also removes an important source of economic inefficiency, as the retirement earnings test widens the wedge between employers' wage costs and workers' net pay considerably, discouraging work even when its social value by far exceeds the private value of the forgone leisure. However, the fact that not all workers have equal opportunities for extending their careers, e.g. due to poor health, outdated skills, or arduous work, has raised concerns about the distributional consequences of such policies.

In this paper, we study to what extent removing the retirement earnings test (RET) and introducing actuarial neutrality in the pension system represent a tradeoff between equity and efficiency. We exploit a Norwegian pension reform implemented in 2011, which for a large group of workers transformed an earnings-tested early retirement program into an unconditional life-long pension annuity that could be claimed on actuarially neutral terms by every eligible worker from the age of 62, regardless of own labor earnings. As the reform implies that pension entitlements previously reserved for those who actually left the labor market now are distributed among all workers, the lifetime value of the new unconditional pension is approximately 17% lower than the full (conditional) pension under the old scheme, even though the government "greased" the reform with extra funding. Workers who retire late are generally better off, however. The extent to which workers at different earnings levels are able and willing to respond to such a policy by increasing their labor supply thus has obvious distributional consequences.

Several studies have investigated the labor supply effects of policies relating to a retirement earnings test (RET). In general, the literature separates between two types of RETs, depending on deferral options. When deferral is possible on actuarially neutral terms, the earnings test is in some sense superficial, and, for a rational forward-looking agent, work incentives are largely unaffected. RET reforms of such schemes have been evaluated in both the US (Friedberg, 2000; Song and Manchester, 2007; Haidar and Loughran, 2008; Engelhardt and Ku-

mar, 2009) and in the UK (Disney and Smith, 2002). To the extent that these studies find positive labor supply effects of the removal of RET (e.g., Friedberg, 2000, and Engelhardt and Kumar, 2009), this is likely to reflect risk-aversion, shortsightedness, or simply failure to understand that withheld benefits are not lost, but just paid out later on (Brown et al., 2013; Rabinovich and Perez-Arce, 2019).

When deferral is not an option, the earnings test is definitely real, and the effect on work incentives is obvious: Any postponement of retirement reduces the lifetime pension entitlement. Baker and Benjamin (1999) evaluate a sequential elimination of a real RET in Canada in the 1970s and estimate a 10 percentage points increase in full year work among 65-69 year olds. Brinch et al. (2017) use a difference-in-differences approach to study the effects of a stepwise real RET-removal in Norway during 2008-10 on the earnings of 67-year-old men. They find a sizeable positive earnings effect for workers who are still active at age 66, and show that the bunching of earnings around the old threshold for the earnings test disappears. The pension reform examined in the present paper has also previously been evaluated in this context, disclosing a substantial overall labor supply effect (Brinch et al., 2015; Hernæs et al. 2016).

In summary, the existing empirical evidence suggests that abolishing the (real) earnings test on pension payments is an effective strategy for increasing labor supply among seniors. However, so far the distributional consequences of RET policies have received less attention. One notable exception is Bönke et al. (2018), who investigate the distributional effects of the introduction of an actuarial deferral option in the German early retirement system in 1992, which essentially removed a real RET. Their findings indicate large positive labor supply responses, at the cost of increased inequality. Another exception is Hernæs and Jia (2013), who investigate the distributional effects of a stepwise increase in the earnings threshold for RET in Norway in 2002 (applying at age 67-69) using quantile regression and complementary cumulative distribution functions. They find a positive labor supply effect at the intensive margin, driven by those who were still active at the age of 66 and had earnings around the thresholds. Since these thresholds were quite low, work incentives were primarily improved at low earnings, and, as a result, the reform led to a decrease in old-age earnings inequality.

A priori, it is not clear, how the labor supply responses to the RET removal affects the overall old-age income distribution. On the one hand, effects at the extensive margin should reduce overall inequality, since richer people tend to work regardless of RET, and hence have

less scope for increasing their labor supply. On the other hand, it has been argued that many elderly workers with physically demanding and poorly paid jobs do not really have the option of extending their career much beyond the early retirement age. These worn-out workers will thus become the losers in a regime where annual pensions are tightly attached to the age of actual retirement. Moreover, as pointed out by Etgeton (2018), employees with low education and low pay are generally those who are most exposed to involuntary job loss and therefore have less possibilities to adjust the timing of retirement in accordance with own preferences. Finally, as richer people also have higher hourly wages, they can obtain a given increase in labor earnings through a lower increase in hours worked.

The purpose of the present paper is to examine the distributional consequences of removing the early retirement earnings test on pension income, while keeping the overall pension expenditures roughly fixed. We study the direct distributional effects, given the pattern of employment and earnings, as well as the effects operating through changes in labor supply. The distributional impacts are examined by comparing complete earnings histories up to the early retirement age with expected lifetime earnings and pension income after this age. We also compute standard measures of inequality (Gini coefficients) based on alternative income concepts, such as overall lifetime income and old-age income.

Our empirical analysis builds on complete administrative data, covering the entire Norwegian population, with employer information and individual earnings trajectories from 1967 onwards. The data allow us to single out the group of private sector workers that was exposed to the removal of the earnings test (approximately 23% of the active workforce). Our primary empirical strategy is to compare the last two birth cohorts (1946-47) that were exposed to a real retirement earnings test with the first two cohorts (1949-50) that were exposed to a fully actuarially neutral pension system with no earnings test. The data allow us to compute virtually complete lifetime earnings histories for all these cohorts. We show that while the distribution of prime-age earnings – defined as average annual earnings over the 40-year period from age 21 to 60 – is almost identical for the pre –and post-reform cohorts, their earnings paths after the early retirement age (62 years) diverge considerably. Our analysis confirms the findings in Hernæs et al. (2016) of large average labor supply effects at age 63 and 64, and having access to newer data, we are able to show that these effects remain strong at ages 65-67 also.

We carry out a novel empirical analysis in three parts. First, we explore how the labor supply responses vary across the prime-age earnings distribution. Our main strategy is to divide the sample into deciles based on accumulated labor earnings from age 21 to 60, and estimate the effect of the pension reform separately within each bin.[1] We find that the labor supply responses to strengthened work incentives are surprisingly similar across the distribution of accumulated labor earnings. For all prime-age earnings deciles, except at the very top, employment rates during age 63-65 increased by approximately 20 percentage points, whereas (unconditional) hours worked per week increased by 7-10. During age 66-67, the employment rate increased by 10-15 percentage points and hours worked per week by 3-5. In total, we estimate that the reform caused an increase in hours worked by as much as 42% during the five-year early retirement period. In terms of employment status and hours worked, the weakest response is found among the top-earners, who had relatively high employment rates even prior to the reform and, thus, had less potential for an increase. In terms of absolute earnings, on the other hand, the effects are largest at the top of the prime-age earnings distribution.

Second, we characterize the winners and losers. As the reform essentially shifted pension wealth from early to late retirees, it is no surprise that the clearest winners are those who would have preferred to continue working throughout the early retirement period in both regimes (the "always-workers"). For them, the new pension entitlements can be considered almost as a lump-sum transfer. The clearest losers are those who would have preferred to leave the labor market at the earliest possible occasion in both regimes (the "never-workers"). For them, the only effect of the reform is that their pension becomes smaller. Given the reasonable assumption that nobody decides to leave (remain in) employment as a result of higher (lower) take-home wages, we can identify the definite winners of the Norwegian pension reform as those who continued working as before until the statutory retirement age in the pre-reform period, and the definite losers as those who left the labor market at the lowest early retirement age in the post-reform regime. Defined this way, we find that 15% of the eligible workers can be counted as definite winners, whereas 8% are definite losers. Comparing these two groups, we show that that the "always-working" winners tend to be individuals with higher prime-

[1] To explore the heterogeneity further, we also apply alternative decile-groupings, based on sick-leave history (from age 45 to 60), the social status of the occupation held at age 60, and life expectancy (based on occupation-specific mortality rates).

age earnings, higher education, more prestigious occupations, and much lower sickness absence in the past than the "never-working" losers.

Finally, we examine the distributional consequences of the reform more directly by examining its effect on the distribution of accumulated pension and labor income after the age of 62. In order to do so, we use the pre-reform cohorts to construct a sample that matches the post-reform cohorts on gender, prime-age earnings, and age 60 earnings, and treat the observed old-age outcomes for this adjusted sample as counterfactual observations for the post-reform sample. The resultant trajectories allow us to disentangle the effect of the new entitlement rules – given the pre-existing labor supply behavior – from the consequences of the reform-generated changes in labor supply. Our findings show that while roughly 45% of the workers lost out in terms of lower pension entitlements, the large labor supply responses ensured that the vast majority (93%) came out with higher overall old-age income. The new entitlement rules also led to a considerable increase in old-age income inequality, whereas the labor supply responses were more or less neutral in distributional terms. The resultant increase in income inequality turned out to be considerable. Measured by the Gini coefficient, overall old-age income inequality increased by approximately 17% as a direct result of the reform.

2 Institutional Setting: The Norwegian Pension reform

The Norwegian pension system has three main pillars: (i) a universal public pension (henceforth referred to by the acronym *FTP*), (ii) contractual early retirement schemes (henceforth referred to by the acronym *AFP*), and (iii) occupational pension schemes in the public and private sector. The reform in 2011 entailed a major restructuring of the universal public pension system, introducing a tighter relationship between individual lifetime earnings and pension entitlements, longevity-adjusted annual pensions, and less generous indexation. However, these changes are implemented gradually and, thus, had very limited impact on the cohorts retiring around the time of the reform. Their longer-term distributional impact is evaluated in Nicolajsen and Stølen (2016) and Halvorsen and West Pedersen (2019). In the present paper, we focus on a reform element that had large and immediate consequences for a large group of workers; namely the removal of the retirement earnings test for private sector work-

ers qualifying for early retirement (*AFP*). This reform was implemented in a quasi-experimental fashion, in the sense that adjacent birth cohorts suddenly faced completely different early retirement incentives.

Prior to the reform, the *AFP*-scheme essentially offered a full pension from the age of 62 (starting from the month after the 62nd birthday) until the statutory retirement age of 67. While it was possible to combine pension and labor income, a confiscatory earnings test implied that the effective tax rates on continued work were very high; see Hernæs et al. (2016).[2] There was no deferral option, so postponing retirement would reduce lifetime pension wealth. Moreover, full retirement at age 62 had no consequences for the build-up of future pension entitlements, as they were calculated *as if* the retiree had continued working until age 67. Workers therefore faced substantial disincentives to work after the age of 62. For private sector workers, two elements of the reform greatly changed this; namely: i) the introduction of flexible take-up of *FTP* from age 62 with no earnings test and with actuarially neutral adjustments of the pension; and ii) the restructuring of the *AFP*-scheme into a lifelong annuity, also with no earnings test and with actuarial neutrality.

The new flexibility features implies that the *FTP* can be taken up at different rates (0, 20, 40, 50, 60, 80, or 100%) and at any time between the age of 62 and 75. The pension payments are adjusted correspondingly to ensure that the expected lifetime pension is unaffected by take-up choices. The new *AFP*-scheme offers a lifelong top-up pension that can be taken out only in combination with *FTP* and is subject to the same actuarial adjustment. Perhaps most importantly, pension payments from the new schemes are no longer reduced against income from other sources. Hence, the new system implies a complete decoupling of decisions regarding labor supply and decisions regarding the timing of pension take-out.

In order to qualify for a full pension at the age of 62, the combined *FTP* and *AFP* entitlement must ensure a minimum annual pension level. In this paper, we focus on the majority of workers whose income history is sufficiently stable to satisfy this condition.[3] For this group, the main impact of the reform was a substantial strengthening of the incentives to work after

[2] The pension was reduced in proportion to the income as a share of previous income (defined as the average income in the three best of the last five years).

[3] The reason for excluding individuals who fail to satisfy the condition is that the impact of the reform was more complex for this group. These individuals would either need to postpone take-up (for anything between a month and five years) or take-up *FTP* at some rate below 100 pct. For this group the reform therefore not only affected work incentives but also the earliest access age for the pension.

reaching the age of 62, whereas the earliest age at which they could claim a full pension was unaffected. Work incentives were improved in many ways. First, as we explain in more detail below, the removal of the earnings test implied an increase in the average (net of tax and transfer) take-home wage by as much as 150%. Second, despite additional funding provided by the government, the maximum lifetime value of the *AFP* pension was reduced by approximately 17% for the first affected cohorts, as resources previously reserved for actual retirees became distributed among all eligible workers. Third, while both the *AFP*- and *FTP*-entitlement in the old scheme were calculated *as if* the individual had continued working until age 67, the new scheme was entirely based on actual earnings. This meant that earnings between age 62 and 66 now generally add to the size of the *FTP* entitlement, whereas this was rarely the case before.

The new *AFP*-scheme applied to individuals who had not yet taken up *AFP* by January 2011, implying that the cohort of 1949 was the first to be fully covered by the new scheme. Individuals born in 1948 could choose to enroll in the new scheme by postponing take-up until 2011. This cohort will therefore consist of individuals enrolled in both the old and the new scheme. Individuals born in 1947, 1946, 1945, and 1944, who had still not taken-up *AFP* by January 2011, could also enroll; however, they were offered substantially less generous versions of the scheme (corresponding to 60%, 40%, 20%, and 10% of the full entitlement, respectively). In the following, we shall generally refer to the cohorts born in 1949 or later as the post-reform cohorts, while we refer to the cohorts born in 1947 or before as the pre-reform cohorts.

The restructuring of the private sector *AFP*-scheme was the result of tripartite negotiations between the state and the major associations of employers and employees, starting in 2008. In order to secure an agreement, the government provided extra funding, facilitating an extra "compensation benefit" for all workers born before 1963. Hence, as we show below, the majority (approximately 55%) of the workers came out with higher pensions than under the pre-reform regime. From a fiscal point of view, this turned out to be a good investment, though, as the extra tax revenue generated by the resultant labor supply responses more than compensated for the extra funding; see Hernæs et al. (2016). The outcomes of the *AFP*- negotiations and the main features of the new private sector *AFP* were probably known by most workers from around mid-2009. At this time, it was generally not possible to enroll into or switch between the schemes, since *AFP*-eligibility in both the private- and most of the public sector requires several years of employer- and sector-specific tenure.

3 Data and identification strategy

Our empirical analysis exploits Norwegian administrative data containing detailed information on earnings, employment, occupation, educational attainment, pension entitlements, and demographic characteristics for the entire population. The main analyses will be based on the birth cohorts who reached the age of 62 just before (born 1946-47) and just after (1949-50) the implementation of the reform. To assess pre-reform trends, older cohorts (1943-45) are included in parts of the analysis. We exclude the 1948-cohort from the main part of the analysis because members of this cohort could self-select into either the old or the new *AFP*-scheme.[4] We return to this cohort in Section 5, however, were we use it to identify the workers' own preferences with respect to the choice of early retirement scheme.

Based on the entire earnings history from 1967 and information about the main employer in the years preceding the reform, we identify *AFP*- and *FTP*-entitlements at an individual level. The eligibility requirements for a full pension with *AFP* changed slightly as part of the reform; hence, to avoid selectivity, we sample the analysis population such that it consists of workers who would have qualified by age 62 under both the old and the new rules (see Online Appendix A for a description of eligibility rules before and after the reform). In order to minimize potential endogeneity problems related to anticipation of the reform, our analysis population is conditioned on employment by age 60 rather than by age 61 or 62 (since the incentive to stay on until age 61 or 62 may have been affected by the reform).[5] Descriptive statistics for the pre- and post-reform cohorts are presented in Table 1 and Figure 1. We note that the two groups are similar in terms of demographic composition (gender and fraction of immigrants), educational attainment, work hours, and earnings. The latter is particularly evident when we look at the distribution of prime-age earnings (average annual earnings from age 21 to 60) for the pre- and post-reform cohorts. As can be seen from Figure 1, panels (a) and

[4] In principle, the pre-reform cohorts could also enter into the new *AFP*-scheme by postponing retirement until 2011; however, these cohorts would only be eligible for a substantially less generous version of the scheme. This implies, however, that the 1946 (1947) had better work incentives from age 65 (64) than earlier cohorts, provided that they had not already retired by January 2011. As we show below (Figure 2), this appears to have had little importance in practice – most likely because the typical retirement age with the old *AFP* was age 62. If anything, it might lead to a small underestimation of the true labor supply effects.

[5] Since the negotiations of the reform began in 2008, and the youngest post-reform cohort reached the age of 60 in 2010, we cannot completely rule out behavioral responses to the reform before age 60. As a robustness check, Hernæs et al. (2016) carry out their analyses conditioning on employment at age 58. The fact that this does not noticeable change their results indicates that ex ante selection seems to be a minor concern.

(b), the distribution functions for pre -and post-reform cohorts are hardly distinguishable. The earnings levels observed at age 60 and 61 are somewhat lower for the post-reform cohorts, however, most likely because these cohorts were adversely affected at this age by the economic downturn in 2009-2010 following from the financial crisis.

Table 1: Descriptive statistics

	Pre reform cohorts Born 1946-47	Post reform cohorts Born 1949-50
Number of observations	16,110	15,628
Share of all employed at age 60 (%)	23.5	23.6
Baseline characteristics:		
Women (%)	19.1	21.6
Immigrants (%)	0.7	1.1
Compulsory education only (%)	18.7	17.0
High school (%)	62.4	64.8
College (%)	18.9	18.2
Weekly work hours at age 60	41.3	41.0
Mean earnings (NOK 1,000):		
…at ages 21-60 (annualized)	612.6	614.7
…at age 60	752.9	724.6
…at age 61	720.1	688.7
…at age 62	609.9	625.1
…at age 63	418.5	518.2
…at age 64	323.0	449.8
…at age 65	264.1	376.3
…at age 66	221.9	293.3
…at age 67	167.8	205.5
Sick leave (months with any registered sick-leave per year in last 15 years - annualized)	0.36	0.39
Life expectancy for by occupation by gender at 62	21.5	21.6
Occupation's social status (ISEI)	47.2	47.1

Occupation status at age 60 is based on the International Socio-Economic Index of occupational status (ISEI) suggested by Ganzeboom et al. (1992) and derived from the International Standard Classification of Occupations (ISCO). Life expectancy is also occupation-specific, computed separately for men and women, and based on Borgan and Texmon (2015).

The main outcome variables used in the analyses are employment status, earnings, and weekly work hours in the calendar years at which the individuals reach the age of 63, 64, 65, 66, and 67.[6] Data on earnings come from the public tax records, and individuals with annual earnings exceeding NOK 100,000 (in 2019 value, corresponding to € 10,000 or $ 11,000) are

[6] Given that reliable earnings data are available for whole calendar years only, the outcomes used in this paper are also defined at the calendar year level. We start with the year individuals reach the age of 63 (and thus are 62 years old at the start of the year), since this is the first year where we can observe the full effect of the reform.

classified as employed.[7] This threshold implies that a person is considered employed in a given year if annual earnings exceeded approximately 18% of the average earnings level for a full-time-full-year position. Weekly work hours are calculated using an hourly wage rate imputed from earnings and work hours at age 60.

Figure 1. **The distribution of average annual earnings during age 21-60 and 63-67. Pre –and post-reform cohorts**
Note: All earnings are measured in NOK 1000 and inflated to 2019-value (using the deflator in the Norwegian pension system).

It is clear from Figure 1 that while the distribution of cumulative labor earnings up to age 60 are virtually identical for the pre- and post-reform cohorts (panels (a) and (b)), their earnings after age 62 diverge considerably (panels (c) and (d)). In particular, we note a large drop in the spike at zero earnings and an increase in the probability mass around typical full-time earnings (panel (c)), implying that the old-age cumulative earnings distribution (panel (d)) is significantly shifted to the right for the post-reform cohorts.

[7] Earnings obtained in other years are inflated to 2019 value using the adjustment factor in the Norwegian social insurance system, which corresponds approximately to the annual average wage growth.

Given the striking similarity of the pre- and post-reform cohorts' earnings paths up to age 60, identification of the reform effects will be based on a direct comparison of these cohorts' employment and earnings patterns from age 63 (i.e., from the age at which the reform had a full effect), with controls for observed individual characteristics. The main identifying assumption underlying our empirical strategy is that the two last pre-reform cohorts represent a valid counterfactual for the two first post-reform cohorts. In other words, we assume that if the reform had never been enacted, the labor supply behavior (and outcomes) of the post-reform cohorts would have been largely identical to that of the pre-reform cohorts (after controlling for observable differences between the groups). This translates into three different assumptions, discussed in turn below, namely: (i) no self-selection into or out of the analysis population, (ii) no calendar time effects, and (iii) no spillovers between members of the pre- and post-reform cohorts, implying satisfaction of the so-called *Stable Unit Treatment Value Assumption* (*SUTVA*).

As discussed in section 2, self-selection related to anticipation of the reform cannot be entirely ruled out. While selection *into* the private sector *AFP*-scheme was generally not possible, selection *out of* the scheme and into the public sector scheme may have been an option for some. If post-reform workers, who wish to retire at an early stage, were more likely to shift to the public sector, we might overestimate the true reform effect, because the remaining members of the post-reform group are more prone to continue working. The fact that we condition the sample on employment and *AFP*-affiliation at age 60 leaves little room for such a response, however, since the post-reform cohorts reached this age in 2009 and 2010, respectively, shortly after the content of the reform was known. Hernæs (2017) shows that less than half a percent of private sector workers eligible for the post-reform *AFP* switches to the public sector between age 59 and age 60. Moreover, Hernæs et al. (2016) find that conditioning the sample on employment at age 58 instead, does not alter the estimated labor supply responses noticeably, but does introduce more noise due to a less accurate determination of *AFP*-entitlements. This indicates that endogeneity in the *AFP*-group assignment is unlikely to be driving any of the results.

To assess the validity of the assumption of no calendar time effects, either related to underlying trends or to cyclical fluctuations, we show in Figure 2 how age-specific employment rates and average earnings developed over the last five pre-reform birth cohorts; i.e.,

those born in 1943, 1944, 1945, 1946, and 1947, respectively. For comparison, we also report the same statistics for the first two post-reform (1949 and 1950) cohorts. Focusing on the labor supply at age 63-64, there are no indications of a trend toward increased labor supply among the pre-reform cohorts. It is perhaps possible to see slight trend toward higher employment rates at age 65-66, but that could be related to the fact that the latest pre-reform cohorts were partially treated at this point, provided that they had not already enrolled into the old *AFP*; conf. Section 2. In any case, the main take-away from Figure 2 is that the big shifts coincided with the reform. It is also worth noting that the outcome period used in our analysis was a period of relative macroeconomic stability, particularly during the first four years (2009-2013) where the unemployment rate fluctuated between 3% and 4%. After that, the economy lost some steam, and the unemployment rate peaked around 5% in 2015. If anything, this development should have contributed to lower employment in the post-reform cohorts during the ages covered by early retirement options.

Figure 2. Employment rates and average earnings for five pre-reform (1943-47) and two post-reform (1949-50) cohorts

Note: All earnings are inflated to 2019-value (using the deflator in the Norwegian pension system).

Spillover effects between birth cohorts cannot be entirely ruled out. On the one hand, increased labor supply of the post-reform cohorts at the age of 62 and 63 could harm the employment prospects of pre-reform individuals at the age of 65 and 66, who might be competing for the same kinds of jobs. However, only a small minority of workers are competing for new jobs at this age, whereas the grand majority either remain in their current job (perhaps working

fewer hours) or fully retire. This type of spillover effects should therefore be negligible. Another kind of spillover could arise from the joint retirement decisions of married couples. Kruse (2019) provides empirical evidence from Norway suggesting that spousal spillovers in retirement decisions are asymmetric, such that wives respond to their husbands' choices, but not necessarily vice versa. Given the typical age difference within couples, this implies that the most relevant spillover effect in our data is a situation where a male worker belonging to the pre-reform cohort chooses to retire early due to the poor work incentives, and that this instigates his younger wife, belonging to a post-reform cohort, to retire as well. This implies that the full reform effects will not be revealed until both spouses have entered the post-reform regime. For our analysis, it implies that the ultimate reform effects might be somewhat underestimated.

4 The social gradient in labor supply responses

In order to assess the potential heterogeneity in reform effects, we divide the population into different socioeconomic groups based on information available at age 60, and estimate separate reform effects for each group. Given our focus on the distributional consequences of the reform, we use prime-age earnings as the primary grouping criterion; i.e., we divide the population of workers at age 60 into deciles based on each worker's position in the age 21-60 earnings distribution within own birth cohort. Figure 3, panel (a) presents, for all the four birth cohorts included in our estimation sample, the average age 21-60 earnings levels for each of these deciles, measured in 1,000 NOK (inflated to 2019 value). Average earnings over these 40 years vary from around 325,000 NOK in the lowest decile to more than 1 mill. NOK in the upper decile. Panel (b) then shows, for each decile, the impact of the RET reform on the economic reward (net of tax) associated with continuing another year (at age 63) with the job held at age 60, while panel (c) shows the relative increase in this reward. It is clear that the improvement in work incentives is very large across the earnings distribution, with the average annualized improvement varying between NOK 175,000 and 230,000 measured in absolute terms and between 50 and 200 percent measured in terms of relative improvement. While the absolute increase in the take-home wage was largest at the top of the earnings distribution, the relative increase was largest at the bottom.

Figure 3 Prime-age (21-60) earnings and reform-generated changes in work incentives at age 63. By decile in distribution of prime-age earnings

Note: The reported statistics are based on the total estimation sample, consisting of *AFP*-eligible workers belonging to the 1946, 1947, 1949, and 1950 birth cohorts (N=31,738). All earnings are inflated to 2019-value (using the deflator in the Norwegian pension system). Panel (a) shows average annual earnings over the 40 years from age 21 to age 60 by decile in the same earnings distribution. Panels (b) and (c) show the average absolute and relative reform-generated increase in the take-home wage (after taxes and earnings tests) associated with annual earnings at age 63 equal to the earnings level at age 60. Dotted horizontal lines indicate population averages.

Figure 4 illustrates how a classification of workers based on prime-age earnings correlates with a range of individual characteristics. Panel (a) first shows how the prime-age earnings levels at age 60 vary across the deciles in the accumulated prime-age earnings distribution. A first point to note is that the earnings levels are relatively high at this age for all the deciles in our estimation sample, reflecting that we have conditioned on employment and early retirement eligibility. For the sample as a whole, the observed average earnings level at age 60 of around NOK 650,000 lies around 20% above the average full-time-full-year earnings observed for all workers in Norway. Yet, the earnings differences are substantial, with the top decile earning approximately three times as much as the bottom decile. Panel (b) then illustrates the large gender gap in prime-age earnings within these birth cohorts. While women constitute 20% of the whole sample, they make up as much as 80% of the bottom decile and as little as 1% of the top decile. Panels (c)-(f) show how a range of alternative classification indicators differ across the prime-age earnings deciles; i.e., educational attainment (panel (c)), the social status of the occupation held at age 60 (panel (d)), the expected longevity associated with the occupation held at age 60 (panel (e)), and overall sickness absence during age 45-60 (panel (f)). It is evident that the categorization based on prime-age earnings correlates closely with alternative categorizations based on these characteristics. We return to estimates based on such alternative categorizations after we have presented the main results.

Figure 4: Individual characteristics by decile in the prime-age (21-60) earnings distribution
Note: The reported statistics are based on the total estimation sample, consisting of *AFP*-eligible workers belonging to the 1946, 1947, 1949, and 1950 birth cohorts (N=31,738 except in panels (d) and (e) were missing information on occupational classification reduces the sample to N=31,021 and N=26,211, respectively). The social status of the occupation at age 60 (panel (d)) is based on the International Socio-Economic Index of occupational status (ISEI) suggested by Ganzeboom et al. (1992) and derived from the International Standard Classification of Occupations (ISCO). Computation of life expectancy (panel (e)) is also occupation-specific, and based on Borgan and Texmon (2015). Total sick leave last 15 years (panel (f)) is measured in terms of the number of months with positive sick leave per year.

The estimation of group-specific reform effects is based on a simple ordinary least squares regression of the following type:

$$Y_{it} = \beta x_i + \theta T_i + \epsilon_{it}, \tag{1}$$

where Y_{it} represents the outcome of interest (employment, earnings, weekly work hours) at age t, x_i is a vector of covariates including gender, education (nine fields and eight levels), country of origin for immigrants (five regions), and weekly work hours and earnings at age 60, and T_i is a treatment-dummy equal to 1 for the post-reform cohorts, and 0 for the pre-reform cohorts. The coefficient θ represents the treatment effect. This is exactly the same equation as that used for the whole population in Hernæs et al. (2016), and, for ease of comparison, we also use exactly the same explanatory variables. Note, however, that we use a more restrictive definition of employment, as we require annual earnings to exceed NOK 100,000 (rather

than 10,000). Our definition still allows for relatively minor positions, as NOK 100,000 constitutes less than a fifth of the average earnings level for a full-time position in Norway.

Figures 5-8 present our main results, in terms of estimated effects of the reform on employment status, weekly hours of work, annual labor earnings, and annual labor earnings relative to the earnings level at age 60, respectively. Starting with employment status, the top panels of Figure 5 show the employment rates at age 63, 64, 65, 66, and 67, respectively, within each prime-age (21-60) earnings decile for the pre-reform and post-reform cohorts. We see that the employment rate increases along the distribution of past earnings for both groups. The differences in employment levels between the pre- and post-reform cohorts appear to be roughly constant across the earnings distribution. The bottom panels report the reform effects on employment estimated within each decile with a 95% confidence interval. The effects estimated for the whole sample (indicated by the dashed horizontal line) were roughly 17, 22, 21, 16, and 10 percentage points at age 63, 64, 65, 66, and 67, respectively. The within-decile estimates are generally around the same level across the earnings distribution, with a moderate hump-shape at ages above 63 such that the effects are largest at the upper-medium part of the distribution, but smallest at the very top. This pattern repeats itself also for the hours worked outcome; see Figure 6. At age 63, weekly hours worked increased by approximately 7 throughout the earnings distribution. At higher ages, a more conspicuous hump shape emerges, with largest effects the upper-medium part of the distribution and lower effects at the top.

Although the estimated reform effects on employment and hours worked are roughly the same across the prime-age earnings distribution, measured in absolute terms, it is worth noting that relative to the initial (pre-reform) level of labor supply, the effects are considerably larger at the bottom of the earnings distribution. For example, while the seven added work hours supplied at age 63 by people belonging to the bottom of the prime-age earnings distribution constitutes a 35% increase relative to pre-reform hours, the same number of added hours toward the upper part of the distribution constitute a 25% increase. Considering the reform effects for all years (age 63-67) together, we estimate that weekly hours worked increased by 6.1 on average, or by 42%. For the bottom decile, it increased by 5.4 hours (42.6%). The effect reached its maximum for the 7th decile with 7.3 hours (51%), and its minimum for the very top decile with 4.3 hours (21.9%).

The estimated reform effects on annual earnings are provided in Figure 7. On average, labor earnings increased by 100-150,000 NOK in each year with entitlement to early retirement. For this outcome, there is a marked positive social gradient in the effect pattern, with larger reform effects the higher the position in the prime-age earnings distribution. Note that it is not meaningful to estimate the earnings effects with the conventional log-specification in our case, as the behavioral responses primarily occur at the extensive margin, with earnings typically either equal (or close) to zero or equal (or close) to the age 60 level; see Figure 1, panel (c). A more appropriate alternative may be to define the outcome explicitly in terms of earnings relative to the age 60 level. The results from such a model are presented in Figure 8. The effects are again very similar across the earnings distribution, and conspicuously similar to the employment effects shown in Figure 5. At ages 63-65, the effects on annual earnings constitute approximately 15-20% of the initial (age 60) earnings level for all deciles in the earnings distribution, except for the top decile, where the effects again are significantly smaller than for the other groups.

Figure 5. Observed employment rates for pre- and post-reform cohorts and estimated reform effects by age and decile in the age 21-60 earnings distribution

Note: The top diagrams indicate the employment rate at age 63-67 across the earnings distribution for the pre-reform cohorts (1946-47, gray dots) and post-reform cohorts (1949-50, black dots), respectively. Earnings deciles are based on earnings at age 21-60 and are calculated within cohorts. The lower diagrams report the estimated reform effects (with 95% confidence intervals) for each decile, based on Equation (1), as well as the average effect size across the income distribution (the dashed horizontal line). The population consists of workers affiliated with a private sector *AFP* scheme who were employed at age 60 and qualified for retirement at age 62 both before and after the reform.

Figure 6. Observed weekly hours worked for pre-and post-reform cohorts and estimated reform effects by age and decile in the age 21-60 earnings distribution

Note: The top diagrams indicate average hours worked at age 63-67 across the earnings distribution for the pre-reform cohorts (1946-47, gray dots) and post-reform cohorts (1949-50, black dots), respectively. Earnings deciles are based on earnings at age 21-60 and are calculated within cohorts. The lower diagrams report the estimated reform effects (with 95% confidence intervals) for each decile, based on Equation (1), as well as the average effect size across the income distribution (the dashed horizontal line). The population consists of workers affiliated with a private sector *AFP* scheme who were employed at age 60 and qualified for retirement at age 62 both before and after the reform.

Figure 7. Observed annual earnings for pre- and post-reform cohorts and estimated reform effects by age and decile in the age 21-60 earnings distribution

Note: The top diagrams indicate average earnings at age 63-67 across the age 21-60 earnings distribution for the pre-reform cohorts (1946-47, gray dots) and post-reform cohorts (1949-50, black dots), respectively. The lower diagrams report the estimated reform effects (with 95% confidence intervals) for each decile, based on Equation (1), as well as the average effect size across the income distribution (the dashed horizontal line). The population consists of workers affiliated with a private sector *AFP* scheme who were employed at age 60 and qualified for retirement at age 62 both before and after the reform.

Figure 8. Observed annual earnings relative to earnings at age 60 and estimated reform effects by age and decile in the age 21-60 earnings distribution

Note: The top diagrams indicate average earnings, measured relative to earnings at age 60, at age 63-67 across the age 21-60 earnings distribution for the pre-reform cohorts (1946-47, gray dots) and post-reform cohorts (1949-50, black dots), respectively. The lower diagrams report the estimated reform effects (with 95% confidence intervals) for each decile, based on Equation (1), as well as the average effect size across the income distribution (the dashed horizontal line). The population consists of workers affiliated with a private sector *AFP* scheme who were employed at age 60 and qualified for retirement at age 62 both before and after the reform.

Although the extensive nature of labor supply decisions made during the early retirement age makes it impossible to estimate meaningful labor supply elasticities at the individual level, a natural way to sum up the messages from Figures 5-8 could be to compute such elasticities at the group-level; i.e., divide the decile-specific reform-initiated relative changes in earnings or hours worked by the corresponding relative changes in take-home wages reported in Figure 3, panel (c). If we do this, we obtain elasticity estimates that apparently rise monotonously with prime-age earnings, from 0.2-0.3 for the lowest deciles to 0.3-0.5 for the upper deciles. However, this would arguably give a distorted picture of group-specific labor supply responses. As noted by Hernæs et al. (2016), given that there was a strictly positive labor supply within all groups even before the reform, despite take-home wages close to zero, there are some natural limits to the labor supply elasticities in our context. For example, as the average weekly hours worked at age 63 for the bottom decile were as high as 20 before the reform, it is difficult to imagine anything more than a doubling of the labor supply for this group (in which case absolutely everyone works full time). Since we know from Figure 3 that the take-home wage was more than tripled for this group as result of the reform, this imposes an absolute upper limit on the labor supply elasticity calculated this way of approximately 0.5. By contrast, the top decile would reach fulltime work for everyone with a 50% increase in labor supply; hence, given that their take-home wage also increased by 50% on average (Figure 3), the absolute upper limit on their elasticity calculated this way is approximately 1.0.

Viewed as a whole, we interpret the results in Figures 5-8 as suggestive of relatively homogenous labor supply responses across the different earnings groups, with a possible exception for the very rich. This is somewhat surprising, since we would generally expect to find the most physically demanding jobs and worn-out workers in the lower end of the earnings distribution, presumably with less scope for individual adjustments. One explanation may be that there is quite some overlap in occupational groups between deciles, such that low-wage individuals with long careers may fall into the same category as high-wage individuals with shorter or interrupted careers. This point suggests that it may be of some interest to assess alternative categorizations of socioeconomic groups. Hence, as an alternative to deciles based on accumulated prime-age earnings, we have divided the population into cells based the occupation held by age 60. Figure 9 presents the result from this exercise. To facilitate comparison across the different categorizations, we show the average estimated effects for the age 63-

67 period instead of separate effects for each age. The first column of panels in Figure 9 summarizes the effects already presented in Figures 5-8, by reporting the estimated effects on average annual earnings during the whole early retirement period. The two next columns then present corresponding effects by deciles in distributions based on occupation. In the second column (panels (b), (f), (j), and (n)), the deciles are based on the occupations' socioeconomic status according to the ISEI index (Ganzeboom et al., 1992), whereas in the third column (panels (c), (g), (k), and (o)), they are based on occupation-by-gender-specific life expectancies (Borgan and Texmon, 2015). Finally, the last column in Figure 9 (panels (d), (h), (l), and (p)) presents results by decile in the distribution of accumulated sick-leave days over the past 15 years, sorted from those with most to those with least absence (as approximately 30% of the workers had zero absence, the rightmost data-point comprises more observations than the others). It seems clear that the labor supply responses are similar across the different socioeconomic groups regardless of the specific variable used to construct them. In particular, it is worth noting that labor supply sensitivity is almost unrelated to past sickness absence.

The choice of socioeconomic indicator has a large influence on the gender-composition of the various deciles. This is illustrated in the four lower panels of Figure 9. We already know from Figure 4 (panel (b)) that based on accumulated prime-age earnings, we obtain a distribution heavily dominated by women at the lower end of the distribution and even more dominated by men at the top. Using the occupation-by-gender-specific life-expectancy measure, we get exactly the opposite pattern. This appears to have remarkably little influence on the distribution of estimated effects, however, suggesting that men and women respond similarly to work incentives. This is indeed confirmed by gender-specific estimates, which we report in Online Appendix B.

Figure 9. Estimated effects on average labor market outcomes age 63-67 by deciles based on alternative socioeconomic indicators.

Note: The point estimates (with 95% confidence intervals) indicate the effects on the 5-year average outcomes, measured over the calendar years in which the persons reach the ages of 63-67. The dotted lines indicate the average estimated effects for the total samples. The grey bars in the bottom panels indicate the fraction of females in each bin. See the note to Figure 4 for a description of how we have defined and computed social class, life expectancy, and sick leave.

5 Characterization of winners and losers

The reform created winners and losers. Those who would have been fully employed under both regimes ("*always-workers*") simply got a top-up pension from the new *AFP*-scheme as a bonus, while those who would have retired completely regardless of regime ("*never-workers*") experienced a reduction in lifetime pension entitlements. Individuals who would have retired

later in the new than in the old regime ("*compliers*") could be better or worse off than before. We do not observe the compliers in the data, but if we impose a monotonicity assumption – i.e. assume that the reform had a weakly positive effect on labor supply for everyone – we are able to identify the always-workers in the pre-reform cohorts and the never-workers in the post-reform cohorts. We can think of the always-workers as those who were fully employed throughout the early retirement period in the pre-reform cohorts, despite the strong incentives to retire, and the never-workers as those who retired completely at age 62 in the post-reform cohorts, despite the strong incentives to continue working.

Table 2. Characteristics of definite winners and losers

	Sample		Never-workers (NW) and Always-workers (AW)	
	I	II	III	IV
	Pre	Post	NW	AW
Number of observations	16,110	15,628	1,192	2,483
Share of pre-/post- group (%)	-	-	7.6	15.4
Baseline individual characteristics:				
Women (%)	19.1	21.6	20.6	19.7
Immigrants (%)	0.7	1.1	1.2	0.6
Compulsory education only	18.7	17.0	22.0	13.6
High school	62.4	64.8	67.6	57.9
College	18.9	18.2	10.4	28.5
Weekly work hours at age 60	41.3	41.0	39.6	41.3
Earnings at age 21-60 (annualized and inflated to 1000 2019-NOK)	612.6	614.6	596.5	634.6
Months of sick leave last 15 years (annualized)	0.355	0.391	0.45	0.25
Characteristics of occupation at age 60				
Life expectancy, by gender (years from age 62)	21.5	21.6	21.2	22.0
Social class (ISEI-scale mean)	47.2	47.1	43.9	51.2
Low-ISEI occupations (lower-quartile)	24.2	24.8	27.9	19.5
Medium-ISEI occupations (mid-quartiles)	48.7	49.2	52.8	45.3
High-ISEI occupations (upper-quartile)	24.1	24.4	16.4	33.6

Note: Columns I and II present the distributions of characteristics in the total samples of persons belonging to the pre- and post-reform cohorts, respectively. Column III reports the distribution for persons in the post-reform cohorts who do not work at all after age 62 (the never-workers (NW)) and Column IV reports the corresponding distribution for persons in the pre-reform who continue working as before (at least 80 percent of earnings level at age 60) every year up to (and including) age 66 (Always-workers (AW)). See the note to Figure 4 for a description of how we have defined and computed social class, life expectancy, and sick leave.

Table 2 reports the characteristics of always-workers and never-workers. By comparing the characteristics of these two groups, we can assess the composition of definite winners and losers. A first point to note is that there are more definite winners (15.4%) than there are definite losers (7.6%). Moreover, the group of winners consists of people with better education, higher prime-age earnings, more prestigious occupations, higher life expectancy, and less past sick leave than the group of losers. The differences in prime-age earnings appear to be moderate (6.4% higher in the winner group). For some of the other characteristics, the differences are considerable. For example, the losers have had 80% more sick leave than the winners during the last 15 years. And the occupational status codes suggest that winners to a much larger extent than the losers have high-status occupations. The most heavily overrepresented occupations among the never-workers turn out to be machine –and plant operators, whereas the most overrepresented occupations in the always-worker group are architects, engineers, and managers (not shown in the table).

Figure 10. Fractions of never-workers and always-workers by decile in the prime-age earnings distribution.
Note: The never-workers are persons in the post-reform cohorts who do not work at all after age 62 and the always-workers are persons in the pre-reform cohorts who continue working as before (at least 80 percent of earnings level at age 60) every year up to (and including) age 66. The solid lines are a second order regression lines (OLS) through the ten respective data-points.

Figure 10 shows how the fractions of never-workers (definite losers) and always-workers (definite winners) by decile in the prime-age earnings distribution. With notable exceptions for the bottom and top deciles, the fraction of never-workers appears to decline monotonously

with prime-age earnings rank. The relatively low fraction of never-workers at the extreme bottom is likely to reflect economic necessity, whereas the relatively high fraction at the extreme top reflects abundance and non-necessity. The fraction of always-workers is relatively stable through the bottom half of the prime-age earnings distribution, and then rises steeply with earnings through the upper half, again with the extreme top as an exception.

Another way of assessing the distribution of winners and losers is to look at the behavior of the members of the 1948-cohort who could actually chose between the old and the new *AFP*. As explained in Section 2, the reform was implemented such that the enrolment into the old *AFP* had to be done before January 1, 2011, but not before the month after the 62nd birthday. This implies that people born in November 1948 could almost choose freely between the old and the new scheme, but in order to be part of the old system they would have to retire immediately after reaching age 62. Those who were born earlier in 1948 could also choose between the two schemes, but in order to be part of the new system, and hence avoid the early retirement earnings test, they would have to postpone retirement 2-11 months, depending on birth month. Figure 11 shows the fractions of workers who actually chose the old *AFP* within these two populations, by decile in the prime-age earnings distribution. It is clear that the old *AFP* was more popular, the lower the position in the prime-age earnings distribution. Among workers who could choose between retiring at the earliest possible occasion with the old earnings tested *AFP* or take the new actuarially neutral pension (Figure 11, panel (b)), approximately 30% revealed a preference for the old earnings-tested *AFP*. However, while the fraction preferring the old *AFP* in the bottom decile of the prime-age earnings distribution was approximately 40%, is was just 10% in the top decile. Hence, there is a strong social gradient in the valuation of the reformed scheme.

Figure 11. Fraction choosing the old rather than the new *AFP* in the 1948 birth cohort. By decile in the prime-age earnings distribution
Note: The solid lines are a second order regression lines (OLS) through the ten respective data-points.

6 Consequences for the old-age income distribution

To shed further light on the distributional consequences of the reform, we now examine its overall impacts on old-age income inequality. We do this by matching each member of the post-reform cohort to a similar person in the pre-reform cohort, and then comparing the resultant pre –and post-reform old-age income distributions. More specifically, we employ 1-to-1 nearest neighbor matching (with replacement) consisting of two steps. First, we match exactly on gender and percentile in the prime-age earnings distribution. Among the several potential matches from the first step, we then select the one who is most similar in terms of earnings at the age of 60. We then treat the entire earnings trajectory of the match from age 60 onwards as the counterfactual earnings trajectory. In addition, these counterfactual earnings are used to calculate a counterfactual pension entitlement. The differences in old-age income distributions between the post-reform sample and the matched pre-reform sample can then be interpreted as caused by the reform.[8]

While it is trivial to describe the distribution of earnings/incomes up to age 67 for all the individuals in our dataset – as they are directly observed in the data – we have to make

[8] Given the almost identical distribution of prime-age earnings for the pre –and post-reform cohorts, the matching exercise does not change the comparison between the two cohorts very much. Hence, all the results presented in this section are very similar if we simply compare the pre –and post-reform cohorts directly.

predictions for some income components earned at later stages. For labor earnings, this is done through a second matching procedure; i.e., for each person with positive earnings in the last observation year for that cohort, we find the closest match in a previous cohort based on gender, prime-age earnings and the relevant age-specific earnings level, and fill in the missing data with the observed next-year earnings for this match (i.e., we do not impose any reform effects on transition rates out of employment after age 67); see Online Appendix C for details. At age 75, we assume that everyone leaves the labor force. Expected lifetime is 83 years.

Figure 12. Average annualized pension, labor, and total income during old age, plotted against average annualized prime-age earnings. By decile in the prime-age earnings distribution.
Note: "Old rules – old LS" show results for the pre-reform pension rules and the pre-reform pattern of labor supply. "New rules – Old LS" show results for post-reform rules with unchanged labor supply". "New rules – New LS" show results for post-reform rules with labor supply responses taken into account.

We explore how the reform affected the overall distribution of expected annual old-age income under three alternative assumptions:

i) Pre-reform pension system and pre-reform labor supply

ii) Post-reform pension system and pre-reform labor supply

iii) Post-reform pension system and post-reform labor supply

Figure 12 first provides average old-age earnings plotted against average prime-age earnings for each decile in the prime-age earnings distribution.[9] It is clear that average pension income remained stable or increased slightly across the prime-age earnings decile bins, and it increased more in the upper part of the distribution (panel (a)). Labor earnings increased considerably for all groups, and again they increased more the higher the prime-age earnings

[9] We have assumed that everyone belonging to the post-reform regime start drawing on their pension at the earliest possible occasion (age 62). Since there is no earnings test in this period and the system is actuarially neutral, this choice has negligible impact on the income profiles.

(panel (b)). As a result, the relationship between prime-age earnings and old-age income became steeper (panel (c)).

Figure 13 provides a more complete picture of the old-age total income distribution, in the form of densities (panel (a)) and cumulative distribution functions (panel (b)). Without taking labor supply responses into account, the new entitlement rules shifted probability mass toward the tails of the distribution, and, hence, increased the degree of dispersion. Without labor supply responses, approximately 45% of the workers would have lost and 55% would have gained in terms of pension entitlements. However, the labor supply responses shifted the income distribution considerably to the right, and the vast majority (approximately 93%) of the workers thus came out with higher old-age income than they would have had in the pre-reform pension regime.

Figure 13. The density and cumulative distribution of total old-age income.
See note to Figure 12.

Table 3, panel A, summarizes the estimated distributional impacts in terms of Gini coefficients. Our primary interest lies in how the pension reform affected the old-age income inequality, as reflected in the sum of labor earnings and pension income over the remaining lifetime from age 62. For comparison, we also report inequality metrics for prime-age (age 21-60) earnings (which were not affected by the reform), for labor earnings during the early retirement window (age 63-67), and for total lifetime income. An intuitively appealing property of the Gini coefficient is that, multiplied by two, it gives the expected difference in income between two randomly chosen individuals, relative to the average income. For example, the

Gini coefficient of 0.181 for the prime-age earnings distribution implies that the average difference in prime-age earnings between all possible pairs of individuals constitutes a fraction of $0.181 \times 2 = 0.362$ (36.2%) of the grand average. Focusing on old-age (62-83) income, the Gini coefficient rose from 0.149 to 0.171 as a result of the redistribution of pension income, given the pre-reform labor supply patterns. Adding in the labor supply responses contributes to a further marginal increase in the Gini coefficient to 0.175, so that the overall rise in the Gini coefficient is estimated to 0.026 (17.4%). However, the influence of this rise in old-age income inequality on the overall inequality in lifetime (21-83) earnings is moderate. We estimate that the Gini coefficient characterizing the distribution of total lifetime incomes rose by 0.002 (1.2%).

Table 3. Income inequality (Gini coefficients) before and after the pension reform

	I Pre-reform pension rules and pre-reform labor supply	II Post-reform pension rules and pre-reform labor supply	III Post-reform pension rules and post-reform labor supply	IV Total reform effect (III-I)
A. All				
Labor earnings				
Prime age (21-60)	0.181	0.181	0.181	
Early retirement period (63-67)	0.579	0.579	0.464	-0.115 (-19.9%)
Overall income				
Old age (62-83)	0.149	0.171	0.175	0.026 (17.4%)
Total lifetime (21-83)	0.168	0.170	0.170	0.002 (1.2%)
B. Men				
Labor earnings				
Prime age (21-60)	0.162	0.162	0.162	
Early retirement period (63-67)	0.583	0.583	0.464	-0.119 (-20.4%)
Overall income				
Old age (62-83)	0.149	0.171	0.174	0.025 (16.8%)
Total lifetime (21-83)	0.154	0.156	0.156	0.002 (1.3%)
C. Women				
Labor earnings				
Prime age (21-60)	0.151	0.151	0.151	
Early retirement period (63-67)	0.537	0.537	0.436	-0.101 (-18.8%)
Overall income				
Old age (62-83)	0.114	0.142	0.144	0.030 (26.3%)
Total lifetime (21-83)	0.133	0.136	0.136	0.003 (2.3%)

Panels B and C of Table 3 report corresponding inequality metrics separately for men and women. The reform-initiated rise in within-gender inequality was similar in magnitude as the

rise in overall inequality. Yet, it is clear that the degree of inequality tends to be smaller among women, both in terms of prime-age earnings and old-age income.

7. Conclusion

The usage of real (non-deferrable) retirement earnings tests (RET) in a pension system causes pension entitlements to be disproportionally allocated to people who retire early. As there is a strong social gradient in the preferred timing of retirement, such that workers with good jobs and high earnings typically wish to retire later than workers with bad jobs and low earnings, a retirement earnings test imply a more equal distribution of old-age income. However, when we take into account that the earnings test is likely to affect the labor supply of workers with different occupations and wage rates differently, it is no longer obvious how a RET ultimately affects the old-age income distribution. If low-wage workers respond sufficiently stronger to the work disincentives embedded in the RET than do high-wage workers, it is in principle possible that the reduction in labor earnings caused by RET outweighs the gain associated with higher pensions for those who retire in any case, such that a *removal* of RET actually reduces old-age income inequality.

Exploiting a comprehensive pension reform in Norway, we have examined the effects of RET on labor supply as well as on overall income inequality by comparing adjacent birth cohorts exposed to fundamentally different early retirement systems from age 62 to 67. We find that removal of the real RET which applied for a large segment of the Norwegian workforce until 2011, raised the labor supply over the whole 5-year early retirement period by approximately 6 hours per week, or 42%. Although we identify considerable labor supply responses at all earnings levels, we find that the estimated effect sizes follow a hump-shaped pattern with respect to the prime-age (age 21-60) earnings distribution. The estimated labor supply effects of RET removal vary from 5.4 hours per week (42.6%) for the bottom decile, up to a maximum of 7.3 hours (51%) for the 7th decile and then down again to 4.3 hours (21.9%) for the top decile.

While the redistribution of pension wealth from early to late retirees implied by RET removal did increase inequality considerably, it turns out that the structure of the estimated labor supply responses had little effect on inequality. Adding up the direct effects (given the pre-reform distribution of employment and work hours) and the effects operating through

changes in labor supply, we estimate that old-age income inequality, as measured by the Gini coefficient, rose by approximately 17% as a result of RET removal.

The findings reported in this paper suggest that policy makers face a particularly challenging tradeoff between efficiency and equity in the design of early retirement systems. The large labor supply responses that followed from the RET removal indicate considerable efficiency gains. Before the reform, Norwegian elderly workers could be subjected to real tax rates (incorporating the earnings test) between 80 and 100%. According to the findings in this paper, this instigated workers to leave the marker in large numbers, despite that many of them would have preferred to work with take-home wages somewhat closer to the true value of their labor. The RET essentially drives a huge wedge between the employer's wage costs and the workers net pay, discouraging work even when its social value by far exceeds the private value of the forgone leisure. Thus, the RET appears to be a very expensive way of achieving a more equal income distribution.

To sum up: The removal of the retirement earnings test in the Norwegian early retirement system led to considerable increases in both labor supply (and economic efficiency) and in old-age income inequality. If the rise in income inequality is considered undesirable, a natural question to ask is whether it is possible to design the pension system such that it achieves the preferred redistribution of incomes, but without incentivizing inefficient early retirement and thus imposing large welfare losses on the economy. Within the context of an actuarially fair early retirement system, this can be done by redistributing pension wealth toward workers with low prime-age earnings, i.e., by making the whole pension system more progressive, or by redistributing it toward occupational groups associated with early labor market exit *on average*. Alternatively, given that there is a positive correlation between life expectancy and the prime-age earnings level, it is possible to achieve a more egalitarian distribution of old-age income simply by distributing parts of the pension wealth in the form of time-limited (e.g. 10 or 15 years) rather than lifelong annuities.

References

Baker, M. and Benjamin, D., 1999. How Do Retirement Tests Affect the Labor Supply of Older Men? *Journal of Public Economics* 71, 27–51.

Borgan, J.-K. and Texmon, I., 2015. Levealder og uttak av tidligpensjon i ulike yrker (Longevity and the Timing of Pension Claims in different Occupations); In Norwegian with English summary. Rapporter 2015/39, Statistics Norway

Brinch, C., Vestad, O., and Zweimüller, J., 2015. Excess Early Retirement? Evidence from the Norwegian 2011 Pension Reform. University of Zurich working paper.

Brinch, C., Hernæs, E., and Jia, Z., 2017. Salience and Social Security Benefits. *Journal of Labor Economics* 35 (1), 265-297.

Brown, J., Kapteyn, A., Mitchell, O., and Mattox, T., 2013. Framing the Social Security Earnings Test. Wharton Pension Research Council Working Paper 2013-06

Bönke, T. Kemptner, D., and Lüthen, H., 2018. Effectiveness of Early Retirement Disincentives: Individual Welfare, Distributional and Fiscal Implications. *Labour Economics* 51, 25-37.

Disney, R. and Smith, S., 2002. The Labor Supply Effect of the Abolition of the Earnings Rule for Older Workers in the United Kingdom. *Economic Journal* 112 (478), C136–C152.

Engelhardt, G.V. and Kumar, A., 2009. The Repeal of the Retirement Earnings Test and the Labor Supply of Older Men. *Journal of Pension Economics and Finance* 8 (4), 429-450.

Etgeton, S., 2018. The Effect of Pension Reforms on Old-Age Income Inequality. *Labour Economics* 53, 146-161.

Friedberg, L., 2000. The Labor Supply Effects of the Social Security Earnings Test. *Review of Economics and Statistics* 82, 48–63.

Ganzeboom, H. B., Graaf, P. M. De, and Treiman, D. J., 1992. A Standard International Soco-Economic Index of Occupational Status. *Social Science Research* 21, 1-56.

Haider, S.J. and Loughran, D.S., 2008. The Effect of the Social Security Earnings Test on Male Labor Supply. *Journal of Human Resources* 43, 57–87.

Halvorsen E. and West-Pedersen, A., 2019. Closing the Gender Gap in Pensions: A Microsimulation Analysis of the Norwegian NDC Pension System. *Journal of European Social Policy* 29, No. 1, 130-143.

Hernæs, E. (2017) Pensjonsordninger og mobilitet (pension systems and mobility) Rapport No. 2/2017, Frischsenteret.

Hernæs, E. and Jia, Z., 2013. Earnings Distributions and Labor Supply after a Retirement Earnings Test Reform. *Oxford Bulletin of Economics and Statistics* 75 (3), 410–434.

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

Kruse, H., 2019. Joint Retirement in Couples – Evidence of Complementarity in Leisure. Working paper.

Nicolajsen, S. and Stølen, N. M., 2016. Fordelingsvirkninger av pensjonsreformen. *Søkelys på arbeidslivet* 33 (1-2), 24-44.

Rabinovich, L. and Perez-Arce, F. (2019) Improving Understanding of the Retirement Earnings Test. *Journal of Pension Economics and Finance,* 1-8.

Song, J.G. and Manchester, J., 2007. New Evidence on Earnings and Benefits Claims Following Changes in the Retirement Earnings Test in 2000. *Journal of Public Economics* 91, 669–700

Online appendices

Appendix A: Eligibility in AFP-schemes

Eligibility for the old AFP schemes was determined in part by earnings-history and in part by employment at the time of take-out. The earnings-requirements consisted of three parts that all needed to be satisfied:

- Pensionable income above 1*B* in the take-out year and in the previous year, where *B* is the so-called Basic Amount (Grunnbeløpet) in the Norwegian pension system, currently (2019/2020) equal to approximately NOK 100,000 (≈ € 10,000), and annually adjusted in line with aggregate wage growth.
- Pensionable income earnings above 1*B* in at least 10 years from age 50 (*Last-10-rule*).
- Average earnings above 2*G* in the 10 years with highest earnings after 1967 (*Best-10-rule*).

Pensionable income consists of wage earnings, self-employment earnings, and some temporary social insurance transfers (sick-pay, unemployment insurance, temporary disability insurance). In addition to the earnings-requirements, the individual has to be employed at the time of first take-out. Also, one of these two conditions should be satisfied:

- Employment in the same private sector firm (with an *AFP*-scheme) in the last 3 years.
- Employment in a private sector firm (with an *AFP*-scheme) in the last 5 years.

The eligibility criteria of the new private sector *AFP* are similar to those of the old scheme. As before, they consist of three parts, namely:

- An earnings-requirement (evaluated at the time of take-out)
- An employment requirement (evaluated at the time of take-out)
- An affiliation requirement (evaluated when turning 62)

The earnings-requirement is less strict than that of the old *AFP*-scheme, since the *Last-10* and the *Best-10* rules no longer apply. Thus, the only requirement is that earnings at the time of take-out must exceed 1*B* on an annual basis and that earnings in the preceding year must exceed the *average* of *B* in that year. The second requirement states that, in order to qualify for *AFP*, an individual must be "genuinely" employed in a company affiliated with the *AFP*-

scheme at the time of take-out and must have been so in the previous 3 years. In order to qualify as "genuine" employment, the position should correspond to at least 20% of full-time, and it should represent the primary occupation and source of income. Finally, the affiliation requirement states that the individual must have been covered by the private sector *AFP*-scheme for at least 7 of the previous 9 years when turning 62. This replaces the requirement of affiliation in the previous 5 years applying in the old scheme. In order not to affect the cohorts close to retirement in 2011, this is implemented gradually. For the cohorts analyzed in the present paper, the requirement was 3 out of the last 5 years.

Appendix B: Reform effects by gender

Figure B1. Estimated effects on average labor market outcomes age 63-67 by deciles based on alternative socio-economic indicators. Men.

Note: The point estimates (with 95% confidence intervals) indicate the effects on the 5-year average outcomes, measured over the calendar years in which the persons reach the ages of 63-67. The dotted lines indicate the average estimated effects for the total samples. See the note to Figure 4 for a description of how we have defined and computed social class, life expectancy, and sick leave.

Figure B2. Estimated effects on average labor market outcomes age 63-67 by deciles based on alternative socio-economic indicators. Women.

Note: The point estimates (with 95% confidence intervals) indicate the effects on the 5-year average outcomes, measured over the calendar years in which the persons reach the ages of 63-67. The dotted lines indicate the average estimated effects for the total samples. See the note to Figure 4 for a description of how we have defined and computed social class, life expectancy, and sick leave.

Appendix C: Extrapolation of earnings data series beyond age 67:

Data on earnings is only available until 2017, i.e. until the age of 70-71 for the pre-reform cohorts, and until the age of 67-68 for the post-reform cohorts. In the pre-reform cohorts, the employment rate among 67- and 68-year-olds is around 30 percent and 20 percent, respectively. The corresponding figures for the post-reform cohorts are around 40 percent and 25 percent (cf. figure 2). This implies that even though the pre- and post-reform cohorts essentially face similar rules and work incentives after reaching the statutory retirement age of 67, the gap in employment between the two groups may not immediately disappear. In order to measure the full effect of the reform, we therefore need to extend the earnings trajectories beyond this point. We accomplish this by matching members of younger cohorts with similar individuals from older cohorts, for whom longer earnings trajectories are observed. We then let the younger individuals "inherit" the remaining part of the earnings trajectories from their match. This corresponds to assuming that the rate of decline in labor supply beyond the age of 68 is similar across cohorts.

In practice, we do a 1-to-1 nearest neighbor matching (with replacement) consisting of two steps. First, we match exactly on gender and percentile in the prime-age earnings distribution. Among the several potential matches from the first step, we then select the one who is most similar in terms of earnings at the latest age available for both cohorts. If several matches are equally similar (this will often be the case, when the individual has zero earnings in the latest year available), we match on earnings one year earlier (and if we still have several matches we go back one more year, and so on, potentially until the age of 60 where we have positive earnings for everyone).

The oldest cohort in our data is the 1943-cohort, which means that iterative matching allows us to impute earnings trajectories for all cohorts until the age of 74. Since pension rights can be accumulated until the age of 75, this covers essentially the entire accumulation period. In practice, however, earnings at such late stages, will rarely affect entitlements for the individuals in our sample.