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

Asbjørn Goul Andersen , Simen Markussen , Knut Røed

PII: S0927-5371(21)00085-3
DOI: https://doi.org/10.1016/j.labeco.2021.102050
Reference: LABECO 102050

To appear in: *Labour Economics*

Received date: 9 March 2020
Revised date: 9 June 2021
Accepted date: 11 August 2021

Please cite this article as: Asbjørn Goul Andersen , Simen Markussen , Knut Røed , Pension Reform and the Efficiency-Equity Trade-Off Impacts of Removing an Early Retirement Subsidy, *Labour Economics* (2021), doi: https://doi.org/10.1016/j.labeco.2021.102050

June 9, 2021

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, Oslo, Norway

Simen Markussen

The Ragnar Frisch Centre for Economic Research, Oslo, Norway

Knut Røed

The Ragnar Frisch Centre for Economic Research, Oslo, Norway

Highlights

- We evaluate the distributional impact of removing a retirement earnings test (RET).
- We find large and similar labor supply responses across the earnings distribution.
- However, we show that low-SES individuals are more likely to lose from the reform.
- The RET removal directly increases the Gini-coefficient for old-age income by 21%.
- The labor supply response neither amplifies nor offsets this increase in inequality.

JEL codes: H55, D31, J22, J26

Keywords: Pension reform, inequality, labor supply

*Corresponding author: Asbjørn Goul Andersen, the Ragnar Frisch Centre for Economic Research, Gaustadalleen 21, 0349 Oslo. E-mail: a.g.andersen@frisch.uio.no

Abstract

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. To identify causal effects, we exploit a reform of the Norwegian early retirement program, which entailed that adjacent birth cohorts faced completely different work incentives from the age of 62. The reform removed a strict retirement earnings test such that pension wealth was redistributed from early to late retirees. Given pre-existing employment and earnings patterns, this implied a considerable rise in old-age income inequality. In theory, this direct increase in inequality could be either amplified or offset by changes in labor supply. We estimate that the reform triggered a 42% increase in average hours worked during the period covered by early retirement options; however, as labor supply responses were of similar magnitudes across the earnings distribution, they did little to modify the rise in inequality. As measured by the Gini coefficient, inequality in overall old-age income rose by approximately 0.03 (21%).

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 promoting higher labor supply among the elderly 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.

Many countries still have earnings tests in various forms for individuals who retire before the statutory retirement age, including Germany, France, Belgium, Austria, and USA (see, e.g. OECD (2017, Table 2.A2.1) and Börsch-Supan et al. (2018, Table 1) for recent overviews). In the present paper, we examine to what extent removing a 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. The reform implied that pension entitlements previously reserved for those who actually left the labor market were distributed among all workers. As a result, the *lifetime value* of the new

unconditional early retirement pension was reduced for workers who retired at the earliest possible occasion. For those who continued working into the early retirement period, the pension was increased.

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 Kumar, 2009) and in the UK (Disney and Smith, 2002). To the extent that these studies find positive labor supply effects of the RET removal (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 effect on work incentives is obvious: Any postponement of retirement reduces the lifetime pension entitlement. Baker and Benjamin (1999) evaluate a sequential elimination of such 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. 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 (real) earnings tests 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). 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. There is also a small related literature examining the distributional consequences of raising the early retirement age (Cribb and Emmerson, 2019; Morris, 2019; Geyer et al. 2020).

A priori, it is not clear how the labor supply responses to the RET removal affects the overall old-age labor earnings distribution. On the one hand, effects at the extensive margin should reduce overall income 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.

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 subjected 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 of Hernæs et al. (2016) of large average labor supply effects at age 63 and 64, and we are able to show that these effects remain strong at ages 65-67, and even stretch beyond statutory retirement to age 68 at which point work incentives were unaffected by the reform.

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. We find that the labor supply responses to strengthened work incentives are surprisingly similar across the distribution of prime-age labor earnings. For all earnings deciles, except at the very top, employment rates during age 63-65 increased by approximately 20 percentage points, whereas (un-

conditional) hours worked per week increased by 7-10 hours. During age 66-67, the employment rate increased by 10-15 percentage points and hours worked per week by 3-5 hours. 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. Some of these effects remained even after the end of this period, despite almost unchanged economic incentives at this point. At age 68, we estimate a 4 percentage point increase in employment and a 2 hours increase in work per week. 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, the clearest winners are those who would have preferred to continue working throughout the early retirement period in both regimes (the "always-workers"). For this group, the new pension entitlements can almost be considered an annual 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 these workers, the reform merely reduced the lifetime value of their early retirement pension (by approximately 21%). Assuming 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 pension reform as those who continued working until the statutory retirement age in the pre-reform period. Likewise, we can identify 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 6-7% are definite losers. Comparing these two groups, we show that that the "always-working" winners

tend to be individuals with higher prime-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 age 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 40% 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 21% as a direct result of the reform. However, in contrast to recent studies examining the impact of higher early retirement age (Cribb and Emmerson, 2019; Morris, 2019), we find no indications of severe poverty. This must be interpreted in light of the relatively resourceful group of workers that are included in our analysis. By focusing on private sector workers who are eligible for the early retirement program both before and after the reform, we essentially limit our attention to workers who are still in employment at age 60 and who have had stable careers over many years with relatively high earnings. Although this group of workers is not representative for the full

population of potential retirees in Norway, they may be quite representative for the type of workers that governments might successfully convince to extend their labor market careers.

2 Institutional Setting: The Norwegian Pension reform

The Norwegian pension system has three main pillars: (i) a universal public old-age pension from the National Insurance Scheme (NIS), (ii) contractual early retirement schemes ("*Avtale-festet Pensjon*" 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 workers 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 until the statutory retirement age of 67, when the old-age pension could be claimed. While it was possible to combine the *AFP*-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).[1] There was no deferral option, so postponing retirement would reduce lifetime pension wealth. Moreover, full retirement at age 62 had no consequences for future pension entitlements, which were calculated *as if* the retiree had continued working as before 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 the old-age pension 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 available from age 62 with no earnings test and with actuarial neutrality. The revised system thus implied a complete decoupling of decisions regarding labor supply and decisions regarding the timing of pension claiming.

The new *AFP* applied to individuals who had not yet claimed *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

[1] 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).

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 60%) 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 are 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, we also include older cohorts (1943-45) 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. We return to this cohort in Section 5, however, where 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 the old-age pension and *AFP*-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 restrict our sample to 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, we condition our sample on employment and eligibility by age 60 rather than by age 61 or 62 (because incentives to stay on until age 61 or 62 may have been affected by the reform).[2] 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 gender, 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 (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.

[2] 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.

Table 1: Descriptive statistics

	Pre reform cohorts Born 1946-47	Post reform cohorts Born 1949-50
Number of individuals	16,109	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
Months of sick leave last 15 years (annualized)	0.36	0.39
Earnings (NOK 1,000):		
…at ages 21-60 (annualized)	612.6	614.7
…at age 60	752.9	724.6
…at age 61	720.2	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	221.8	293.3
…at age 67	167.8	205.5
…at age 68	102.8	129.1
Characteristics of occupation at age 60:		
Life expectancy at age 62 (years)	21.5	21.6
Social class (ISEI scale)	47.2	47.1

Note: All earnings are measured in NOK 1,000 and inflated to 2020-value (using the deflator in the Norwegian pension system). Sick leave in the last 15 years (before age 60) is calculated as the average number of months per year with any registered sick leave (only sick-leave periods exceeding 16 days in duration are registered). Life expectancy at age 62 is based the on occupation- and gender-specific estimates of Borgan and Texmon (2015). Social class refers to the status of the occupation held at age 60 according to the International Socio-Economic Index of occupational status (ISEI) suggested by Ganzeboom et al. (1992), which is based on the International Standard Classification of Occupations (ISCO).

The main outcome variables used in the analyses are employment status, gross (pre-tax) earnings, and weekly work hours in the calendar years at which the individuals reach the age of 63, 64, 65, 66, and 67.[3] Data on earnings come from the public tax records, and individuals with annual earnings exceeding NOK 100,000 (in 2020 value, corresponding to €

[3] 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.

10,000 or $ 11,000) are classified as employed.[4] 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 1,000 and inflated to 2020-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

[4] Earnings obtained in other years are inflated to 2020 value using the adjustment factor in the Norwegian social insurance system, which corresponds approximately to the annual average wage growth.

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.

Given the striking similarity of the pre- and post-reform cohorts' earnings paths up to age 60, we base identification of the reform effects on a direct comparison of these cohorts' employment and earnings patterns from age 63 onwards (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, respec-

tively, 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*-eligibility. This indicates that endogeneity in the *AFP*-entitlement is unlikely to be driving any of the results.

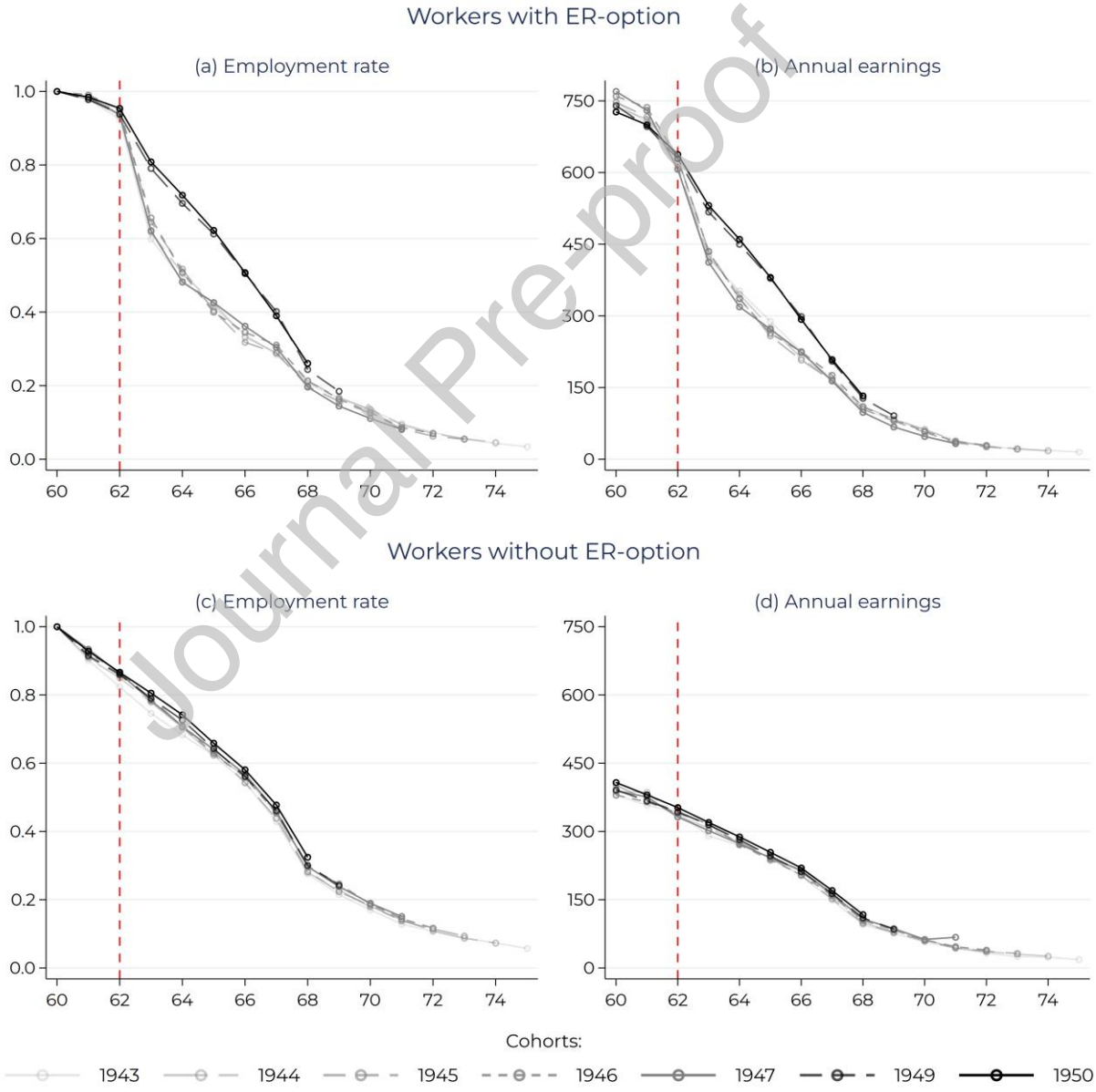

Figure 2. Employment rates and average earnings for five pre-reform (1943-47) and two post-reform (1949-50) cohorts. Conditioned on employment in the private sector at age 60. By AFP eligibility.

Note: All earnings are measured in NOK 1,000 and inflated to 2020-value (using the deflator in the Norwegian pension system). Employment is defined as having annual earnings above NOK 100,000.

To assess the validity of the assumption of no calendar time effects, either related to underlying trends or to cyclical fluctuations, panels (a) and (b) of Figure 2 show age-specific employment rates and average annual earnings for individuals qualifying for the early retirement scheme. The statistics are shown for the last five pre-reform birth cohorts; i.e., those born in 1943, 1944, 1945, 1946, and 1947, and for the first two post-reform cohort (1949 and 1950). 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 a 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. For comparison, panels (c) and (d) of the figure shows the age-specific employment rates and earnings for a group of workers that were not affected by the pension reform. We observe very small changes in employment and earnings for this group.[5]

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) when 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 3 below shows the employment rates for the dif-

[5] Only workers that were neither part of the *AFP* agreement nor eligible for the new early retirement option in the public pension system were unaffected by the pension reform. This group makes up approximately 11% of the workforce and consists of workers with relatively low and/or unstable previous earnings.

ferent age groups of the entire population around the time of the reform. We see that the employment rates are relatively stable for 50-54 year-olds, 55-59-year olds, and 60-62 year old, whereas it jumps discretely at age 63 after the reform.

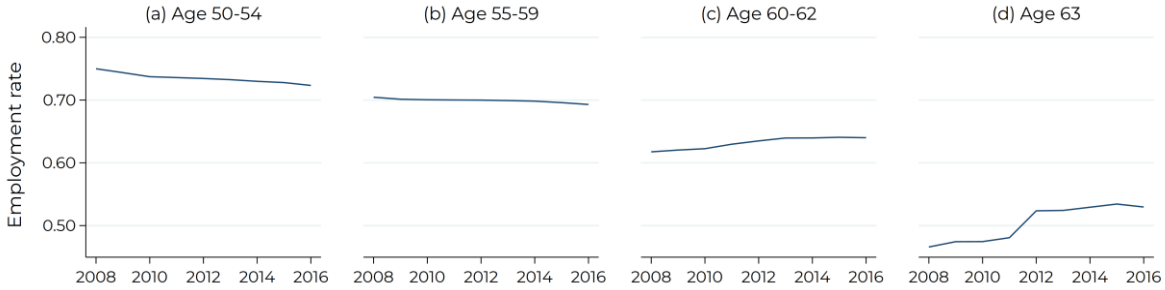

Figure 3. Employment rates for the entire population in Norway. By age and year.
Note: Employment is defined as having annual earnings above NOK 100,000 (measured in 2020-value).

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 will be 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 4, panel (a), presents the average age 21-60 earnings levels for each of these deciles, measured in 1,000 NOK (inflated to 2020-value), for all the four birth cohorts included in our estimation sample. Average earnings over these 40 years vary from around 325,000 NOK in the lowest decile to more than one million 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.

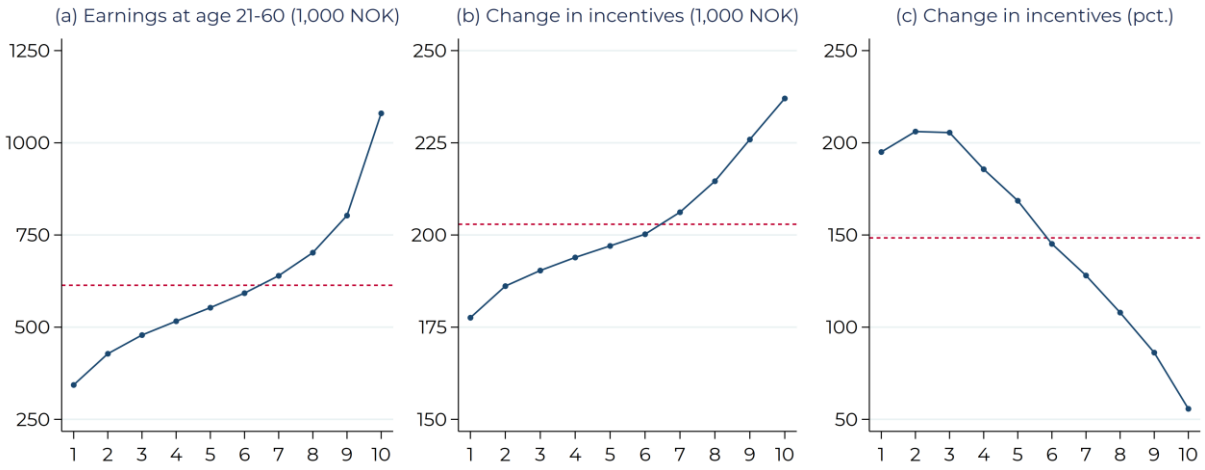

Figure 4: Prime-age (21-60) earnings and changes in work incentives by decile in the prime-age earnings distribution

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). Earnings are measured in NOK 1,000 and inflated to 2020-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 reform-generated increase in the take-home wage at age 63 (in 1,000 NOK and in percent, respectively) along the same deciles as in (a). Changes in incentives are calculated after taxes and earnings tests and assuming that annual earnings are equal to the earnings level at age 60. Dotted horizontal lines indicate population averages.

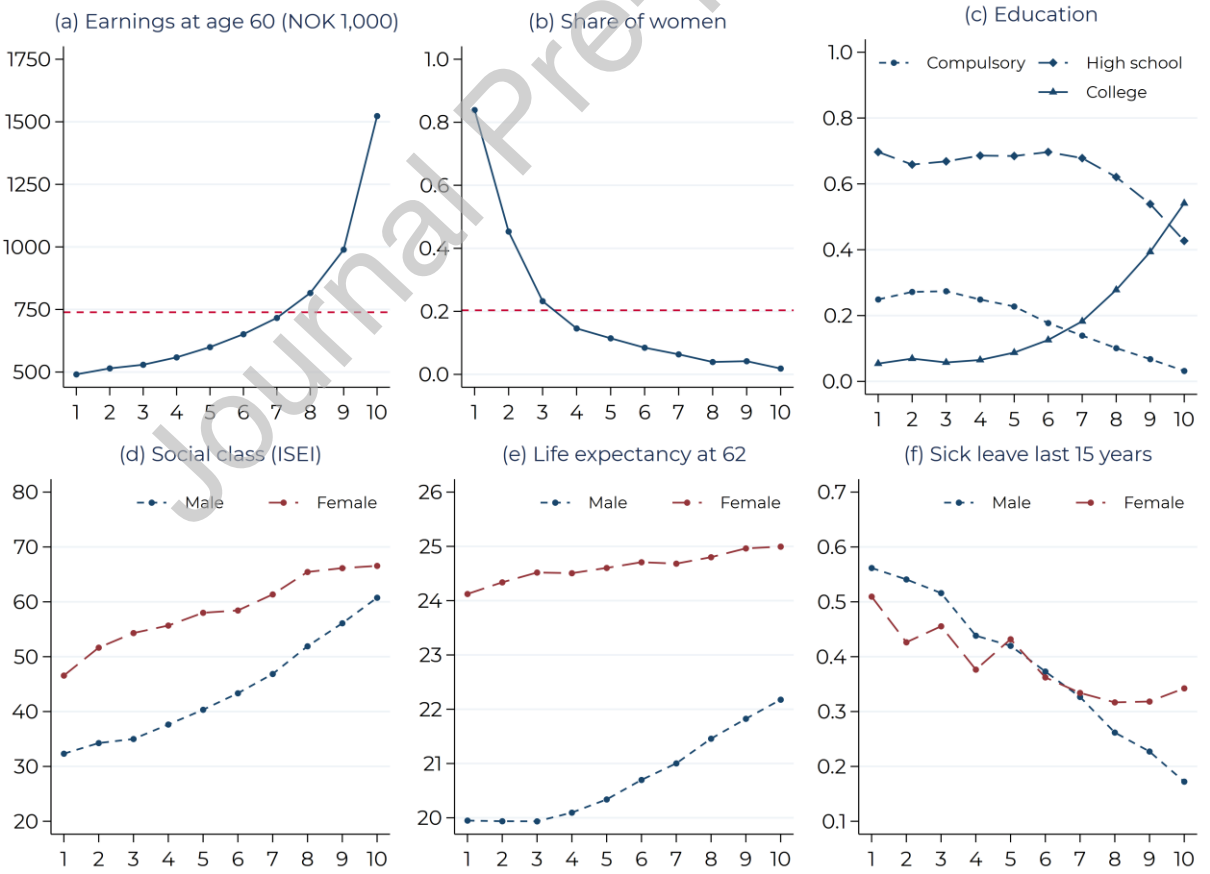

Figure 5: Individual characteristics by decile in the prime-age (21-60) earnings distribution

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) where missing information on occupational classification reduces the sample to N=31,021 and N=26,211, respectively). See the note to Table 1 for a description of how we have defined and computed social class, life expectancy, and sick leave.

Figure 5 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 740,000 lies around 35% 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.

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

$$Y_{idt} = \beta_d x_i + \theta_d T_i + \epsilon_{idt}, \quad (1)$$

where Y_{idt} represents the outcome of interest (employment, earnings, weekly work hours) for a person i belonging to prime-age earnings decile d, measured at age t, and 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. T_i is a treatment-dummy equal to 1 for the post-reform cohorts, and 0 for the pre-reform cohorts, and the coefficient θ represents the treatment effect. This equation is identical to the one 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, given that NOK 100,000 constitutes less than a fifth of the average earnings level for a full-time position in Norway.

Figures 6-9 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 6 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 7. At age 63, weekly hours worked increased by approximately 7 hours throughout the earnings distribution. At higher ages, a more conspicuous hump shape emerges, with the largest effects at 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 individuals in 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 constitutes 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 8. 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 9. The effects are again very similar across the earnings distribution, and conspicuously similar to the employment effects shown in Figure 6. 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.

Figures 6-9 show estimation results for the five-year period that best matches the early retirement period in which work incentives changed (ages 63-67). Given the large labor supply effects identified until the statutory retirement age at age 67, one could hypothesize spillover effects also into higher ages. In Appendix B, we show estimation results also for age 68, indicating positive reform effects on employment (4 percentage point) and work hours (2 hours per week) even at this stage. For one of the post-reform cohorts, we also observe outcomes at age 69, but we do not find any significant effects at this age (not shown).[6]

[6] In Appendix B, we also present results for age 62. As the outcome year corresponding to each age is defined based on the calendar year in which the indicated age is reached, some (but not full) effects could be expected already at this age. The results shown in the appendix indicate a 2-3 hours increase in weekly work, but no impact on employment based on the employment definition used in this paper.

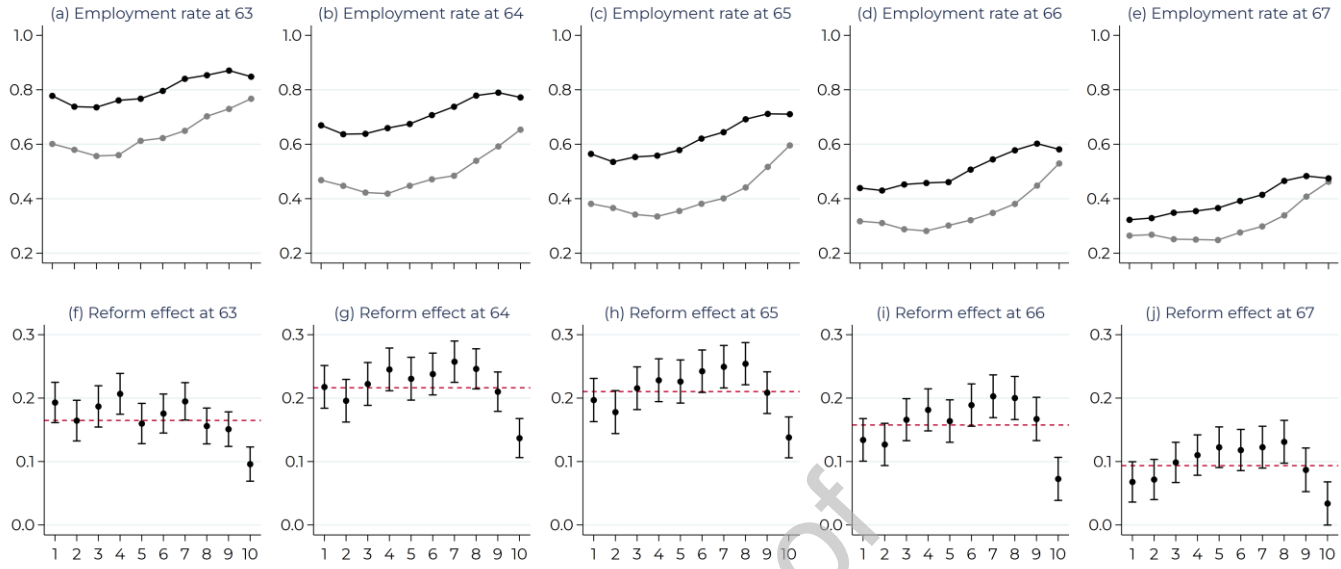

Figure 6. Observed employment rates for pre- and post-reform cohorts and estimated reform effects by age and decile in the age 21-60 earnings distribution

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.

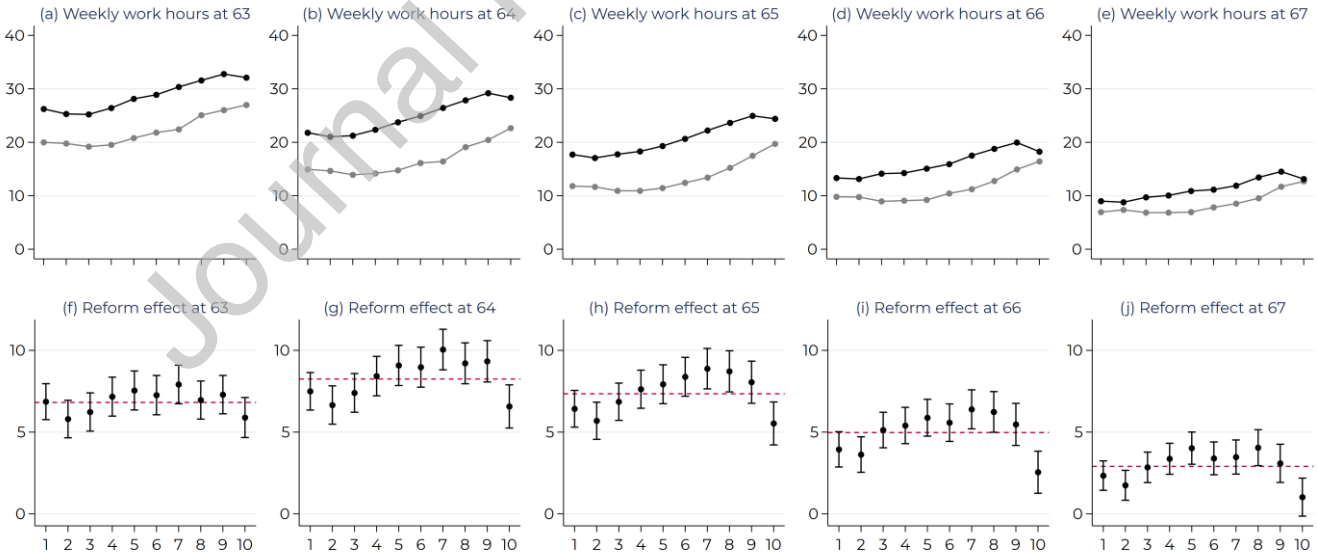

Figure 7. Observed weekly hours worked for pre-and post-reform cohorts and estimated reform effects by age and decile in the age 21-60 earnings distribution

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.

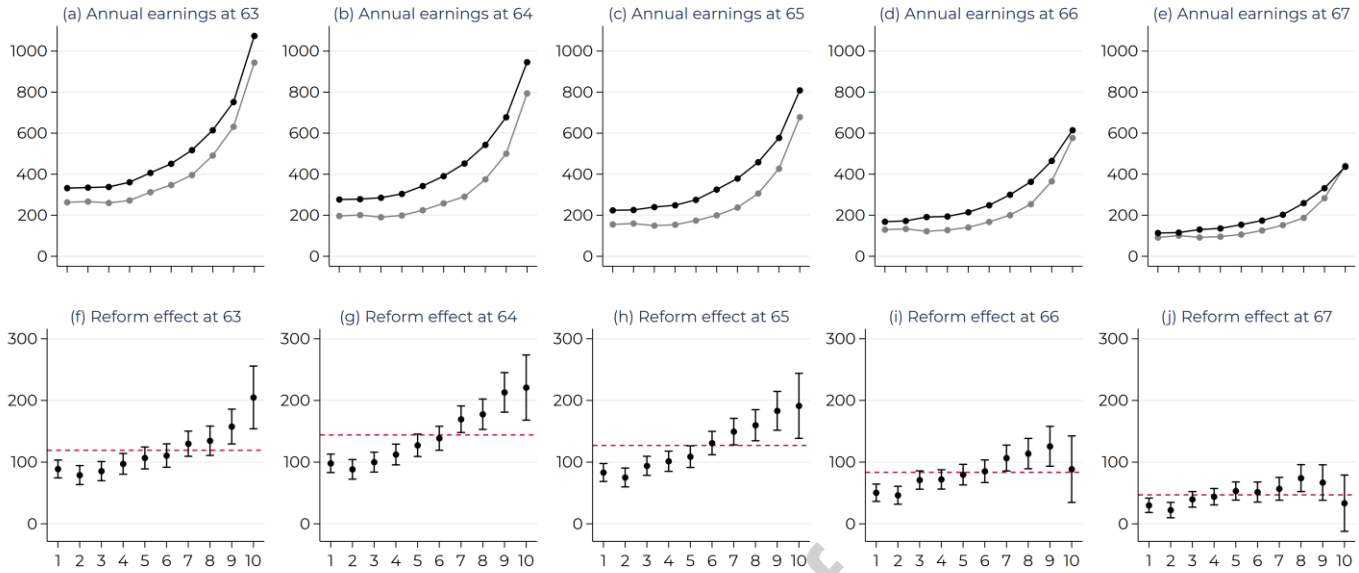

Figure 8. Observed annual earnings for pre- and post-reform cohorts and estimated reform effects by age and decile in the age 21-60 earnings distribution

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 9. 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.

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. However, a natural way to sum up the messages from Figures 6-9 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 4, panel (c). If we do this, we obtain elasticity estimates that apparently rise monotonically 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 4 that the take-home wage was more than tripled for this group due to 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 4), the absolute upper limit on their elasticity calculated this way is approximately 1.0.

Viewed as a whole, we interpret the results in Figures 6-9 as suggestive of relatively homogenous labor supply responses across the different earnings groups, with a possible exception for those with the highest earnings. This is somewhat surprising, since we would generally expect that "worn-out workers", holding the most physically demanding jobs in the lower end of the earnings distribution, should respond less, having less scope for indi-

vidual 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 on the occupation held by age 60. Figure 10 presents the result from this exercise. To facilitate comparison across the different categorizations, we show the average estimated effects for ages 63-67 combined instead of separate effects for each age. The first column of panels in Figure 10 summarizes the effects already presented in Figures 6-9, by reporting the estimated effects on average annual earnings during the full 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 10 (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.

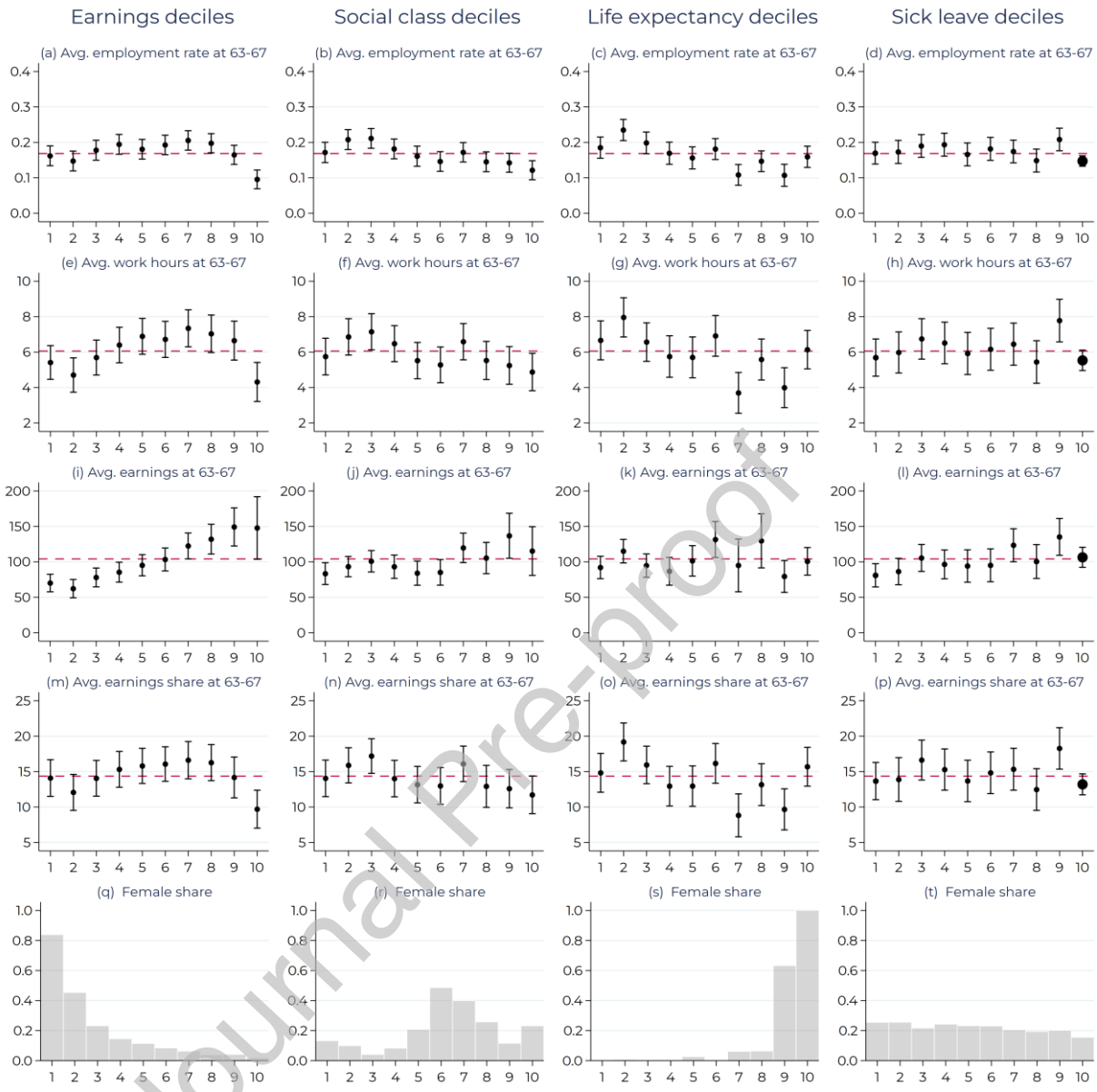

Figure 10. Estimated effects on average labor market outcomes age 63-67 by deciles based on alternative socio-economic indicators.

Note: The point estimates (with 95% confidence intervals) indicate the effects on the 5-year average outcomes, measured over the calendar years in which individuals 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 Table 1 for a description of how we have defined and computed social class, life expectancy, and sick leave. The sick leave scale is reversed such that the lower bins contain workers with more sick leave. The point estimate for the 10th bin is indicated with a larger dot, because this group contains the 30 percent of all workers with no registered absence. The remaining 70 percent are evenly distributed across the bins 1-9. In spite of this, we refer to the bins as sick leave deciles.

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 10. We already know from Figure 5 (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 C.

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 (always-workers) and losers (never-workers)

	Whole sample		Always-workers (Winners) and Never-workers (Losers)	
	I	II	III	IV
	Pre	Post	Winners (Pre)	Losers (Post)
Number of individuals	16,109	15,628	2,415	1,010
Share of pre-/post- group (%)	-	-	15,0	6.5
Baseline individual characteristics:				
Women (%)	19.1	21.6	19.8	19.2
Immigrants (%)	0.7	1.1	0.6	1.5
Compulsory education only (%)	18.7	17.0	13.4	24.5
High school (%)	62.4	64.8	57.7	65.2
College (%)	18.9	18.2	28.9	10.3
Weekly work hours at age 60	41.7	41.3	41.8	39.9
Months of sick leave last 15 years (annualized)	0.355	0.391	0.25	0.65
Earnings at age 21-60 (annualized)	612.7	614.7	637.6	601.3
Characteristics of occupation at age 60				
Life expectancy, by gender (years from age 62)	21.5	21.6	22.0	21.1
Social class (ISEI-scale mean)	47.2	47.1	51.5	43.6
Low-ISEI occupations (lower-quartile, %)	24.2	24.8	18.7	30.3
Medium-ISEI occupations (mid-quartiles, %)	48.7	49.2	45.5	48.8
High-ISEI occupations (upper-quartile, %)	24.1	24.4	34.1	17.9
Lifetime pension entitlements (age 63-83) under different rules (1,000 NOK)				
Pre-reform AFP	1,090	845	0	1,752
Post-reform AFP	1,389	1,394	1,402	1,375
Pre-reform public old-age pension	5,156	5,113	5,319	5,031
Post-reform public old-age pension	5,106	5,077	5,322	4,919
Pre-reform overall annualized pension	291	277	249	315
Post-reform overall annualized pension	302	301	313	293
Change in overall annualized pension (%)	3.8	8.7	25.7	-7.0
Old-age earnings (age 63-83)				
Earnings	1,520	1,994	4,340	0
Earnings annualized	71	93	202	0

Note: Columns I and II present the distributions of characteristics in the total samples of individuals belonging to the pre- and post-reform cohorts, respectively. Column III reports the distribution for individuals in the pre-reform period who continue working as before (at least 80 percent of earnings level at age 60) every year up to (and including) age 66 (i.e. the always-workers or winners). Column IV reports the distribution for individuals in the post-reform cohorts who do not work at all after age 62 (i.e. the never-workers or losers). All monetary amounts are measured in NOK 1,000 and inflated/deflated to 2020-value. See the note to Table 1 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 (15.0%) than there are definite losers (6.5%). 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% higher in the winner group). For some of the other characteristics, the differences are considerable. For example, the losers have had roughly 2.5 times more sick leave than the winners have during the last 15 years. In addition, the occupational status codes suggest that winners are much more likely to have high-status occupations than losers are. 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).

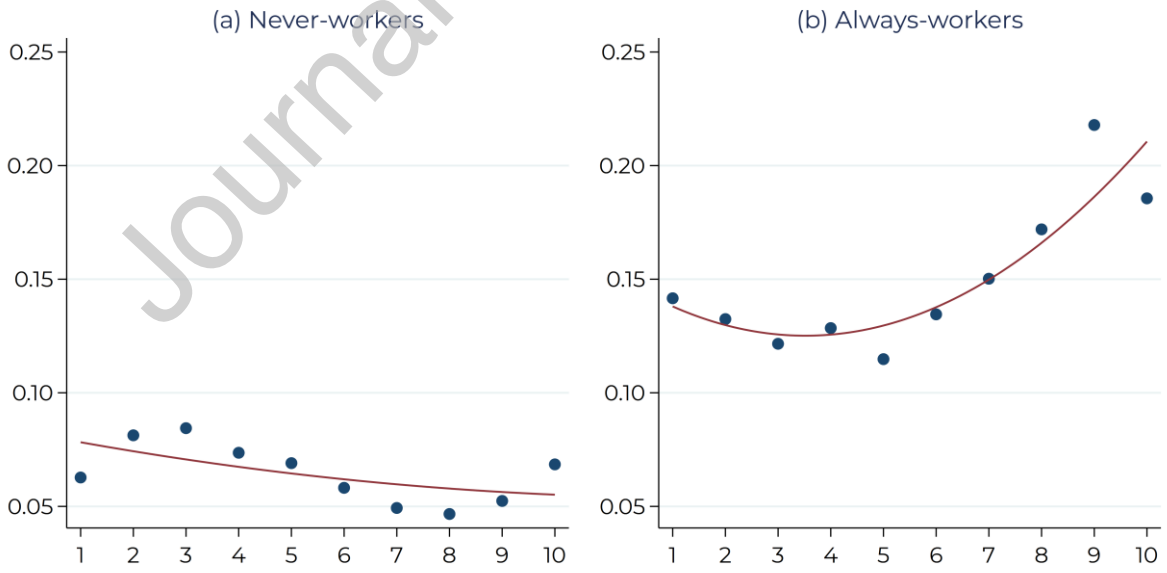

Figure 11. Fractions of never-workers and always-workers by decile in the prime-age earnings distribution.
Note: The never-workers are individuals in the post-reform cohorts who do not work at all after age 62 and the always-workers are individuals in the pre-reform cohorts who continue working as before (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 datapoints.

Figure 11 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 monotonically with prime-age earnings rank. 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 by studying the behavior of the members of the 1948-cohort who could choose between the old and the new *AFP*. As explained in Section 2, the reform was implemented such that the enrollment into the old *AFP* had to be done before January 1, 2011. Given that enrollment was possible from the month after reaching the age of 62, workers born in November 1948 could choose almost freely between the old and the new scheme. In order to be part of the old system, they would have to take up the pension immediately after reaching age 62, whereas postponing take-up by a month (or more) would entail enrollment in the new system. Workers born earlier in 1948 could also choose between the two schemes, but would have to postpone take-up for 2-11 months, depending on month of birth, in order to enroll in the new scheme and avoid the early retirement earnings test. Figure 12 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 among workers in the lower end of the prime-age earnings distribution. Among workers born in November (Figure 12, 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%, it was just 10% in the top decile. Hence, there is a strong social gradient in the valuation of the reformed scheme.

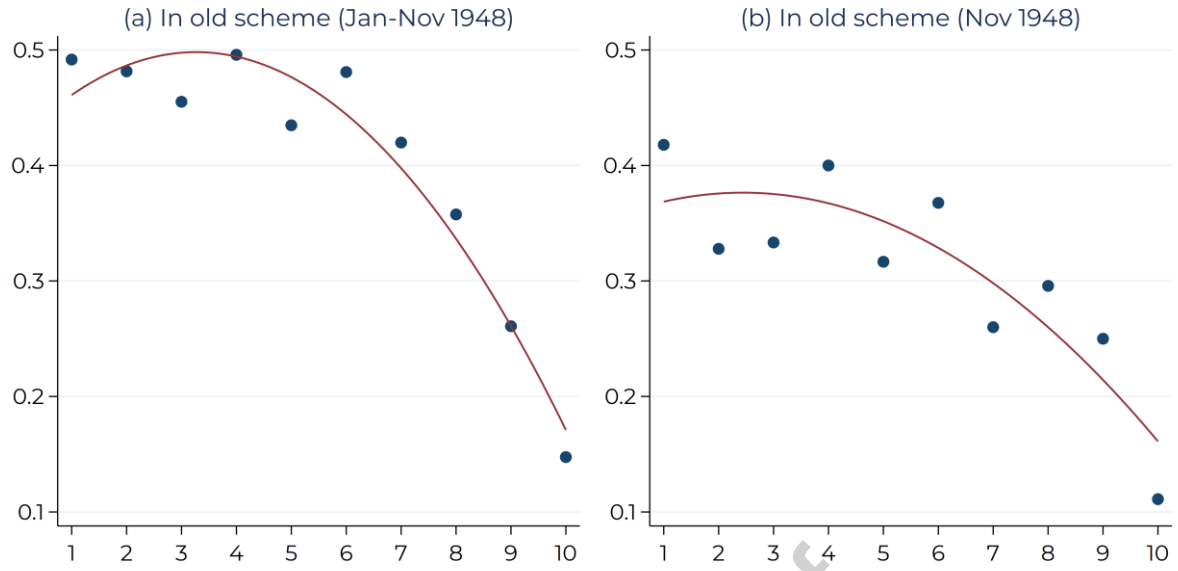

Figure 12. Fraction choosing the old rather than the new *AFP* in the 1948 birth cohort. By decile in the prime-age earnings distribution
Note: The solid lines are a second order regression lines (OLS) through the ten respective data-points.

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 as the counterfactual earnings trajectory. Earnings are observed up to, and including, age 68 for all the individuals in our dataset. Although we would expect some of the labor supply responses during the age 63-68 period to imply slightly higher post-reform employment also from age 69, we choose the more conservative assumption of a zero reform effect after age 68

here. This assumption is implemented by setting labor earnings for everyone to zero from age 69 and onwards. We also set the expected lifetime to 83 years for everyone. Pension entitlements up to this age are then accurately computed.[7]

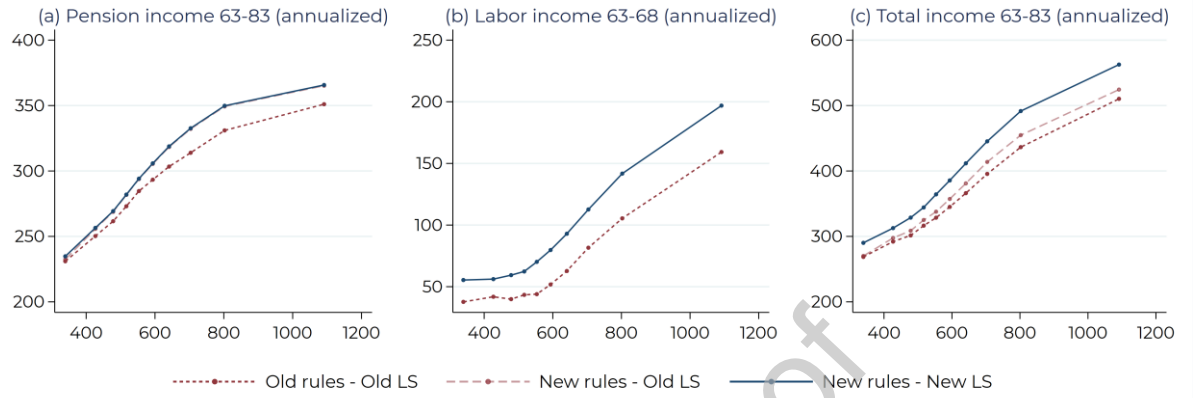

Figure 13. Average annualized pension, labor, and total income during old age, plotted against average annualized prime-age earnings. By decile in the prime-age earnings distribution.
Note: "Old rules – old LS" show results for the pre-reform pension rules and the pre-reform pattern of labor supply. "New rules – Old LS" show results for post-reform rules with unchanged labor supply[8]. "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 by comparing three different situations:

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 13 first provides average old-age earnings plotted against average prime-age earnings for each decile in the prime-age earnings distribution.[8] 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 con-

[7] 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.

[8] 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.

siderably for all groups, and again they increased more the higher the prime-age earnings (panel (b)). As a result, the relationship between prime-age earnings and old-age income became steeper (panel (c)).

Figure 14 provides a more complete picture of the old-age total income distribution, in the form of densities (panel (a)), cumulative distribution functions (panel (b)), and Lorenz curves (panel (c)). Disregarding labor supply responses, 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 40% of the workers would have lost and 60% 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. The degree of inequality increased, as reflected by the (small) shift in the Lorenz curves (panel (c)). The rise in inequality was driven by the redistribution of pension income. The difference between the Lorenz curves with and without labor supply responses taken into account is hardly discernable.

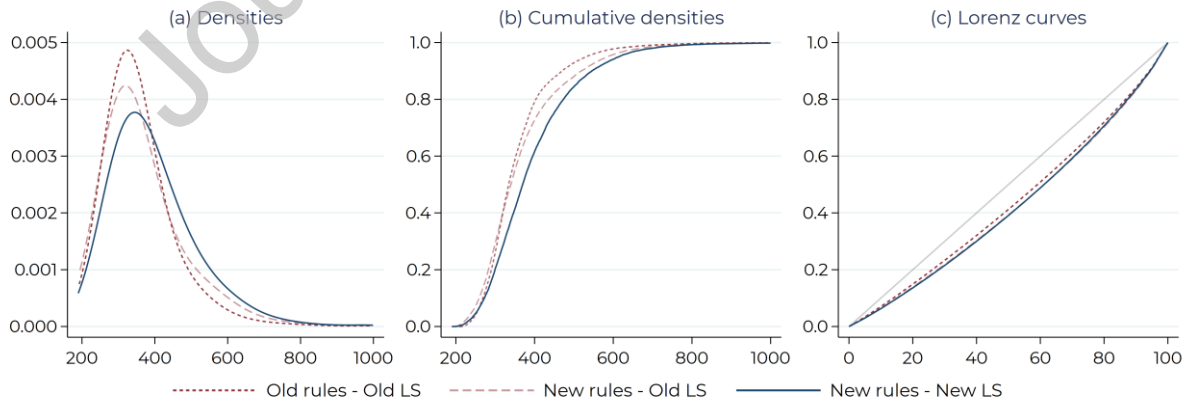

Figure 14. The density and cumulative distribution of total old-age (age 63-83) income.
Note: See note to Figure 13. In the diagrams (a) and (b) the annualized income is capped at 1,000,000 NOK.

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 (age 63-83) income inequality, as reflected in the sum of labor earnings and pension income over the remaining lifetime from age 63.[9] For comparison, we also compute Gini coefficients 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 (age 21-83).

The Gini coefficient of 0.181 for prime-age earnings is relatively low. This reflects that the private sector workers qualifying for early retirement, both before and after the reform, generally had long and stable careers. This Gini coefficient compares to 0.264 for all workers employed at age 60, and 0.353 for all residents at age 60 regardless of their employment status (both based on the same birth cohorts).

By contrast, the Gini coefficient for earnings during the early retirement period was as high as 0.581 for the pre-reform cohorts in our sample, and it fell to 0.464 for the post-reform cohorts when it became more common to continue working. The corresponding numbers for the total population employed at the age of 60 are 0.510 and 0471, for the pre – and post reform cohorts respectively (and 0.548 and 0.507 for all residents of the same cohorts).

The Gini coefficients for old-age income, which mainly consists of pension income (cf. Table 2), are even lower than those of prime-age income. This is because of a minimum pension level and a ceiling on the annual accrual of pension rights. The Gini coefficient, however, rose from 0.129 to 0.157 due to the redistribution of pension income, given the pre-reform labor supply patterns. Adding in the labor supply responses does not change the Gini coeffi-

[9] As the pension system is based on a common life expectancy for everyone, we use the same assumption in our calculations, implying a life expectancy of 83 years.

cient, so that the overall rise in the Gini coefficient is estimated to 0.028 (21.4%). To put this number into perspective, the transfer from the top to the bottom decile of the old-age income distribution in the post-reform sample required to reverse this increase in inequality would amount to approximately 10% of the overall old-age income in the top decile. However, the influence of this rise in old-age income inequality on the overall inequality in lifetime (21-83) income is moderate. We estimate that the Gini coefficient characterizing the distribution of total lifetime incomes rose by 0.001 (0.6%).

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, while the degree of inequality tends to be smaller among women, the rise in in old-age income inequality caused by the reform was larger.

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				
Earnings:				
Prime-age (21-60)	0.181	0.181	0.181	
Early retirement period (63-67)	0.581	0.581	0.464	-0.117 (20.1%)
Overall income:				
Old-age (63-83)	0.129	0.157	0.157	0.028 (21.4%)
Total lifetime (21-83)	0.168	0.169	0.169	0.001 (0.6%)
B. Men				
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 (63-83)	0.128	0.155	0.155	0.027 (21.1%)
Total lifetime (21-83)	0.153	0.155	0.155	0.002 (1.3%)
C. Women				
Earnings:				
Prime-age (21-60)	0.151	0.151	0.151	
Early retirement period (63-67)	0.545	0.545	0.437	-0.108 (19.8%)
Overall income:				
Old-age (63-83)	0.102	0.137	0.137	0.035 (34.3%)
Total lifetime (21-83)	0.132	0.136	0.136	0.004 (3.0%)

Note: Old-age income is the sum of earnings from age 63-68 and pension income from age 63-83. Total lifetime income further includes earnings from age 21-62.

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 implies 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 21% 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 labor market in large numbers, despite the fact 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 old-age 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, as is already the practice in some of Norway's occupational pension schemes.

Acknowledgements: The authors would like to thank Erik Hernæs, Axel West Pedersen, two anonymous referees, and seminar participants in Bergen, Oslo, and Ås for valuable comments and discussions. Administrative registers made available by Statistics Norway have been essential.

Funding information: The research has received support from the Norwegian Research Council (grant # 270875).

Conflict of interests: The authors declare that they have no conflict of interest

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.

Börsch-Supan, A.H., Härtl, K., and Leite, D.N. (2018) Earnings Tests, Non-Actuarial Adjustments and Flexible Retirement. NBER Working Paper 24294.

Cribb, J. and Emmerson, C. (2019) Can't Wait to Get My Pension: The Effect of Raising the Female Early Retirement Age On Income, Poverty and Deprivation. *Journal of Pension Economics & Finance* 18(3), 450-472.

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 Socio-Economic Index of Occupational Status. *Social Science Research* 21, 1-56.

Geyer, J., Haan, P., Hammerschmid, A., and Peters, M., 2020. Labor market and distributional effects of an increase in the retirement age. *Labour Economics 65*, 101817

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.

Morris, T. (2019). The Unequal Burden of Retirement Reform: Evidence from Australia. Working Paper. Available at SSRN 3416304

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

OECD (2017) Pension at a Glance 2017. OECD and G20 Indicators. OECD, Paris.

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.[10]
- Pensionable income earnings above 1*B* in at least 10 years from age 50 (*Last-10-rule*).
- Average earnings above 2*B* 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, and temporary disability insurance). In addition to the earnings-requirements, the individual has to be employed at the time of first take-out. Furthermore, 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)

[10] *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.

- 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: Estimation results for additional ages

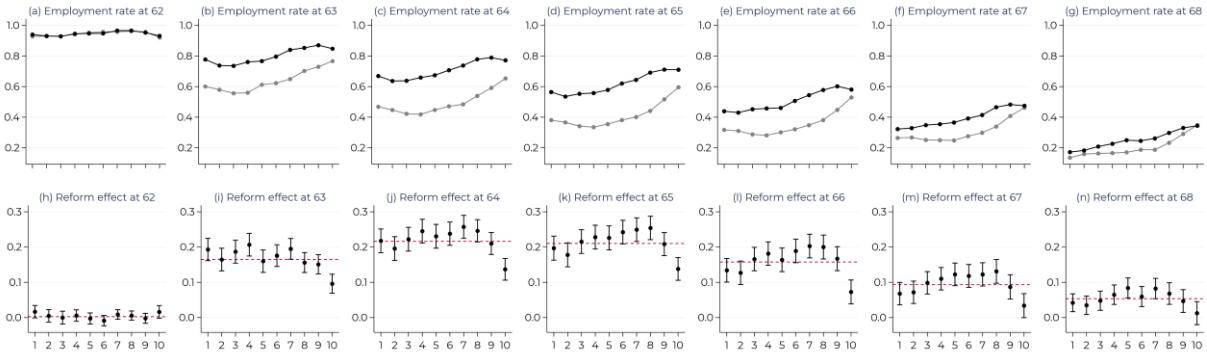

Figure B1. Observed employment rates for pre- and post-reform cohorts and estimated reform effects by age and decile in the age 21-60 earnings distribution

Note: The top diagrams indicate the employment rate at age 62-68 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.

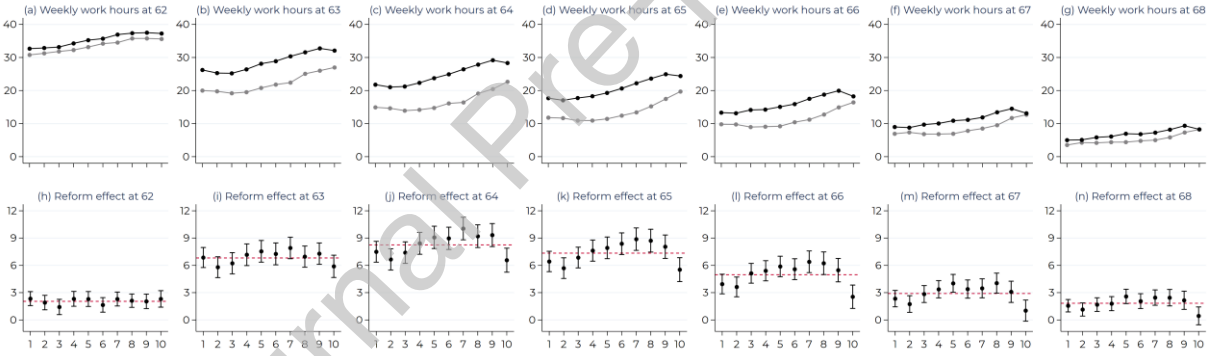

Figure B2. Observed weekly hours worked for pre-and post-reform cohorts and estimated reform effects by age and decile in the age 21-60 earnings distribution

Note: The top diagrams indicate average hours worked at age 62-68 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.

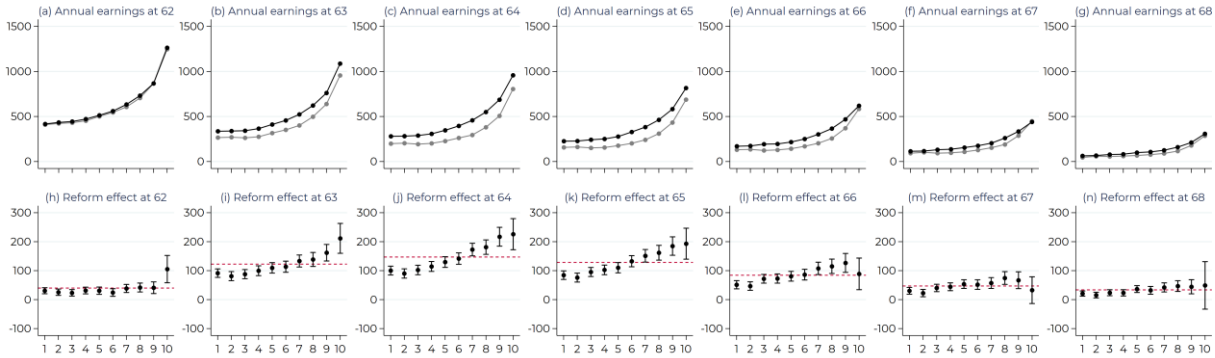

Figure B3. Observed annual earnings for pre- and post-reform cohorts and estimated reform effects by age and decile in the age 21-60 earnings distribution

Note: Earnings are measured 1,000 NOK, inflatesd to 2020-value. The top diagrams indicate average earnings at age 62-68 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.

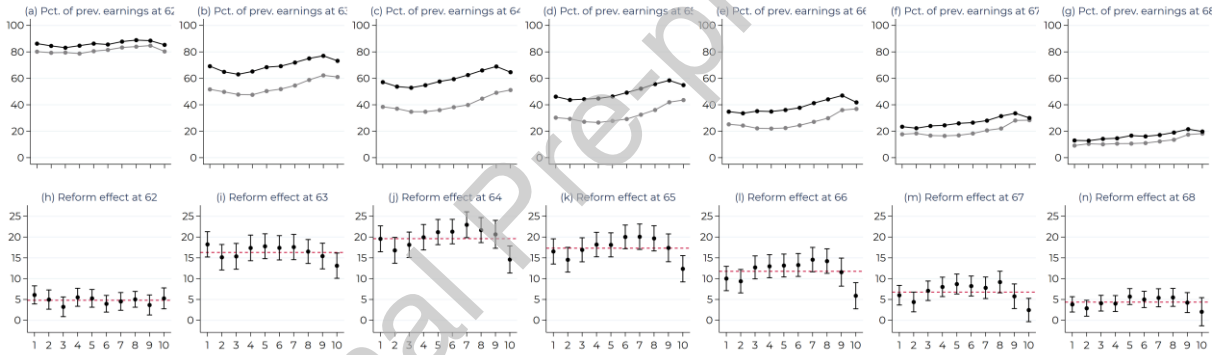

Figure B4. Observed annual earnings relative to earnings at age 60 and estimated reform effects by age and decile in the age 21-60 earnings distribution

Note: The top diagrams indicate average earnings, measured relative to earnings at age 60, at age 62-68 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

Appendix C: Reform effects by gender

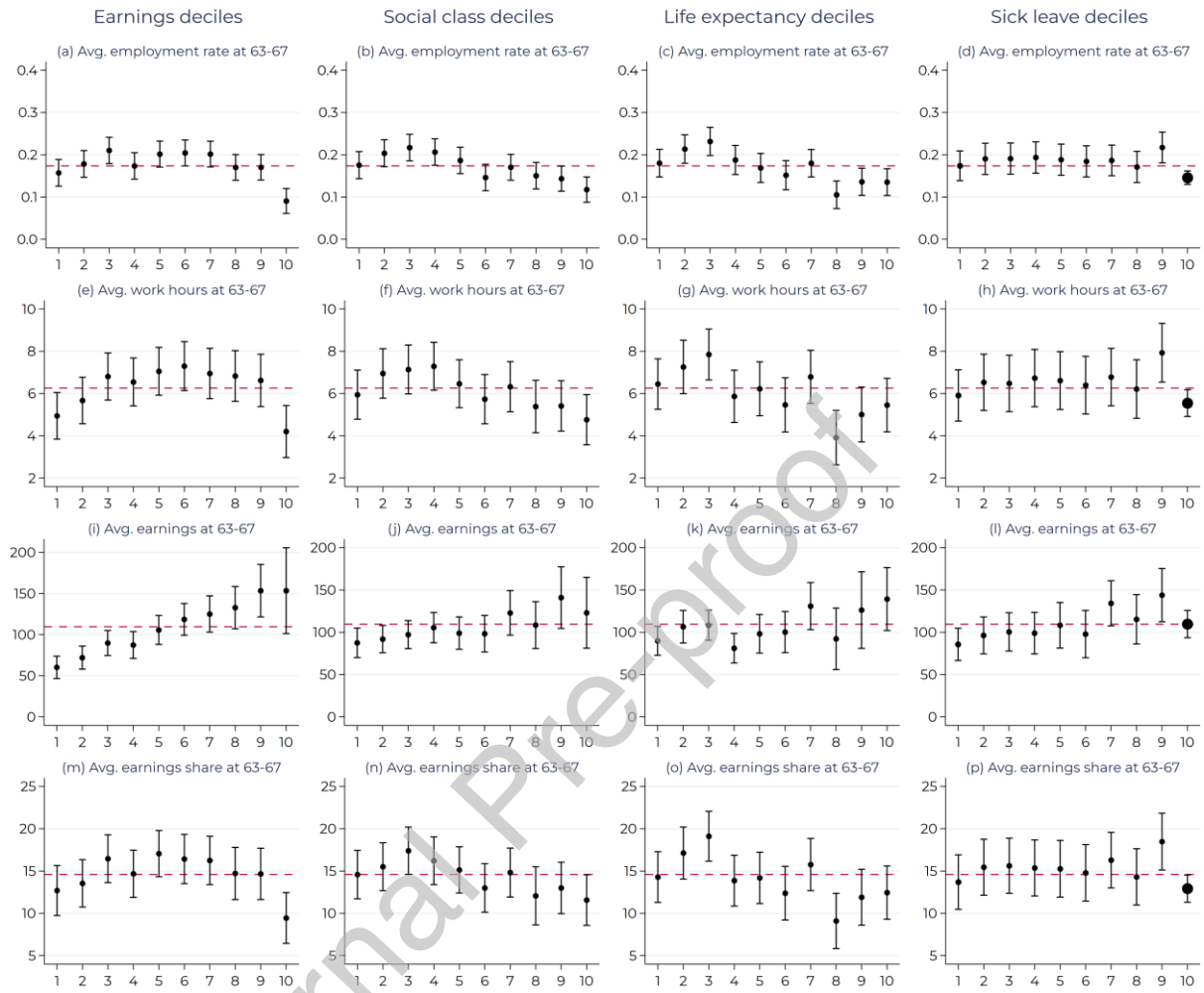

Figure C1. Estimated effects on average labor market outcomes age 63-67 by deciles based on alternative socio-economic indicators. Men.

Note: The point estimates (with 95% confidence intervals) indicate the effects on the 5-year average outcomes, measured over the calendar years in which individuals reach the ages of 63-67. The dotted lines indicate the average estimated effects for the total samples. See the note to Table 1 for a description of how we have defined and computed social class, life expectancy, and sick leave.

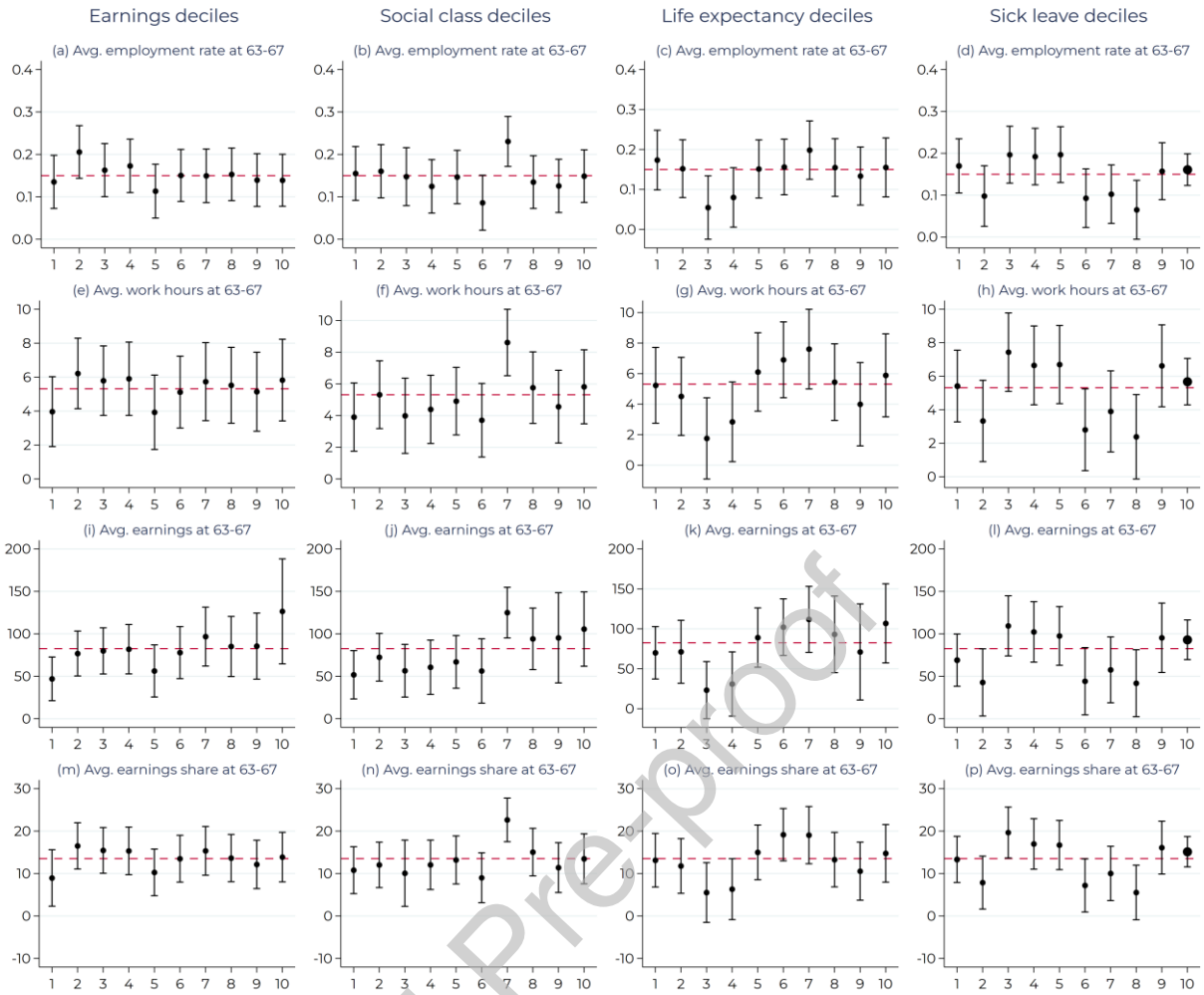

Figure C2. Estimated effects on average labor market outcomes age 63-67 by deciles based on alternative socio-economic indicators. Women.

Note: The point estimates (with 95% confidence intervals) indicate the effects on the 5-year average outcomes, measured over the calendar years in which individuals reach the ages of 63-67. The dotted lines indicate the average estimated effects for the total samples. See the note to Table 1 for a description of how we have defined and computed social class, life expectancy, and sick leave.