Pension reform and labor supply

By

Hernæs, Erik, Simen Markussen, John Piggott, Knut Røed

The definitive publisher-authenticated and formatted version:

Hernæs, Erik, Simen Markussen, John Piggott, Knut Røed, 2016, Pension reform and labor supply, Journal of Public Economics, Vol 142, 39-55., DOI: 10.1016/j.jpubeco.2016.08.009.

is available at:

https://doi.org/10.1016/j.jpubeco.2016.08.009

August 8 2016

Pension Reform and Labor Supply

Erik Hernæs[a], Simen Markussen[b], John Piggott[c], Knut Røed[d]

[a] Ragnar Frisch Centre for Economic Research, Gaustadalleen 21, 0349 Oslo, Norway
erik.hernas@frisch.uio.no (corresponding author)

[b] Ragnar Frisch Centre for Economic Research simen.markussen@frisch.uio.no

[c] ARC Centre of Excellence in Population Ageing Research (CEPAR), UNSW Business School, University of New South Wales, Sydney 2052, Australia j.piggott@unsw.edu.au

[d] Ragnar Frisch Centre for Economic Research knut.roed@frisch.uio.no

Abstract

We exploit a comprehensive restructuring of the early retirement system in Norway in 2011 to examine labor supply responses to increases in work incentives and actuarially neutral reductions in the age of first access to pension benefits. We find that increasing the returns to work is a powerful policy: The removal of an earnings test, implying a doubling of the average net take-home wage, led to an increase in average labor supply by 7 hours per week (30%) at age 63 and by 8 hours (46%) at age 64. The responses primarily came at the extensive margin. In contrast, reducing the access age has almost no effect on labor supply, in our setting with actuarially fair work incentives.

Keywords: early retirement, labor supply, pension reform, program evaluation

Classification: D6, H3, J1, J2

1. Introduction

In response to rapid aging of the population many countries are considering reforms to increase labor supply among workers near to the retirement age by encouraging them to work longer. In this paper we estimate the labor supply responses to a reform package embodying two common elements: changing the earliest age at which workers may access their pension, and increasing work incentives for those already eligible to claim their pension.

Increasing the pension access age implies longer labor force attachment and a reduced life span of pension payouts. More than a dozen countries have either undertaken such a reform, or have announced plans to do so (OECD, 2012). Policy changes of this type will almost certainly result in later retirement (Gruber and Wise, 1999; 2004). In contrast, only a few countries have comprehensively removed pension-related work disincentives for those who have reached the access age. This requires breaking the nexus between retirement age and the access age for benefit payments, leaving workers to decide work patterns and pension drawdown independently. Examples include the US in 1983 and 2000, Canada and Sweden in the 1970s, the UK in 1989, Japan in 1985 and 2002, and Norway in 2011. In an idealized reform of this type, benefits are actuarially adjusted by take-up age.

We use the comprehensive 2011 Norwegian pension reform, which was primarily focused on increasing work incentives, to examine the labor supply responses to alternative reform paths. As we explain below, the reform had widely different implications for different groups of workers, depending on pre-determined factors such as sector of employment and accumulated pension entitlements. Some workers were subject to increases or decreases in access age only, some were subject to large changes in work incentives, and some were more or less unaffected. The reform therefore presents a surprisingly complete quasi experimental set-up for our investigations. No other country has simultaneously implemented such diverse reforms.

We base our analysis on two complementary empirical strategies. Both use comprehensive administrative registers with panel data on employment and earnings for the first birth cohorts potentially affected by the reform and the last unaffected cohorts. First, we compare labor supply patterns before and after reform implementation for groups who were affected in different ways and directions, and use intra-group regression analysis to quantify the labor supply impacts of the reform for the most strongly affected groups. Second, we use the reform-generated changes in work incentives to quantify the relationships between take-home wages and labor supply outcomes, and derive the implied labor supply elasticities. Here, we address the simultaneity problem associated with non-random work incentives by controlling for both hypothetical and actual take-home wages faced under the pre- and post-reform regimes. Intuitively, the *non-causal* associations between labor supply behavior and each of the two incentive variables (calculated on the basis of pre- and post-reform rules, respectively) are likely to be the same before and after the reform, but a causal association will shift between the pre-reform incentives prior to the reform and the post-reform incentives afterwards.

We find that increased work incentives have the potential to raise labor supply considerably. The repeal of the early retirement earnings test in the private sector (leaving the early access age of 62 unchanged) increased average work-hours substantially, with aggregate labor supply elasticities of 0.31 and 0.45 at ages 63 and 64, respectively. Most of the labor supply response occurred at the extensive margin, and the probability of staying on in the labor market with roughly the same work-hours and earnings as at age 60 rose by approximately 17 percentage points at both ages 63 and 64, from initial levels of 41 and 30 per cent, respectively. Although eliminating the earnings test adds a substantial fiscal cost, we show that the labor supply response to the reform under consideration was such that the government budget actually benefited from it, as tax revenues rose more than pension outlays.

We also find that given actuarially fair work incentives, the access age is of minor importance for labor supply behavior. Workers who as a result of the reform faced a lower access age with an actuarially fair early retirement pension (i.e., improved liquidity only) responded by reducing labor supply only slightly at the intensive margin, while maintaining employment status.

Our paper relates to an existing literature which has indicated that earnings tests reduce labor supply both when the tests are “real” in the sense that benefits are not deferred (Baker and Benjamin, 1999; Brinch et al., 2017; Hernæs and Jia, 2013) and when benefits are merely deferred (Friedberg, 2000; Song and Manchester, 2007; Haider and Loughran, 2008; Engelhardt and Kumar, 2009; Disney and Smith, 2002). If labor supply and deferral choices are linked, a labor supply impact of earnings tests with actuarially fair deferral may result from workers perceiving the earnings test as a tax, possibly because the deferral schemes are complicated and poorly understood (Haider and Loughran, 2008), or because actuarial fairness does not apply to persons with high expected mortality (Engelhardt and Kumar, 2009). In the context of Norway specifically, the 2011 reform has also been studied by Brinch et al. (2015), who in line with our findings report a strong labor supply response to the removal of the earnings test and a muted response to the actuarially neutral reduction in access age.

A related literature indicates that reforms which solely reduce access to – or the generosity of – early retirement programs may have the unintended side-effect of increasing the pressure on alternative subsidized escape routes from the labor market, such as disability insurance programs (Duggan et al., 2007; Staubli and Zweimuller, 2011; Bratberg et al., 2004; Vestad, 2013; Røed and Haugen, 2002). Our findings confirm that changes in the access age indeed have spillover effects to disability insurance claims, but that increased work incentives alone do not entail such side-effects.

2. Institutional setting

Before the 2011 reform, the earliest access age for the public pension (hereafter referred to by its acronym FTP) in Norway was 67 years. But all public sector workers and roughly half private sector workers had access to a supplementary early retirement system (hereafter referred to by its acronym AFP), in essence offering a full pension from age 62. Both these pensions were subject to a full earnings test, implying that continued employment after retirement resulted in reduced lifetime pension entitlements. With a full pension, the earnings test became effective from the first dollar earned, such that labor earnings constituting a certain percentage of the pre-retirement earnings level resulted in the same percentage cut in the annual pension.[1] There was no deferral option by delayed take-up, in effect implying very high implicit tax rates on continued work. Hence, the AFP system embodied a strong disincentive to work after the age of 62, particularly for persons with relatively low wages.

The Norwegian 2011 pension reform changed both these systems radically, but the AFP was reformed only in the private sector. The reform implied large and immediate changes in the

[1] To avoid adjustments in cases of "negligible" labor earnings, there was a so-called "tolerance amount" of approximately $2,000 per year that could be earned without adjustment of benefits. All the monetary amounts reported in this paper are inflated to 2013-values using the Norwegian official pension benefit inflator, which in the period covered by this paper roughly corresponds to the wage growth, and then converted to USD ($) with exchange rate of mid-2013, $1=NOK 6.04.

work incentives for many elderly workers. In this paper, we focus on two system parameters of paramount importance for labor supply: i) the earliest access age and ii) the returns to continued work as determined by earnings tests and the degree of actuarial fairness in deferred pension entitlements.[2]

Adjustments to the FTP. The reform reduced the earliest access age to FTP from 67 to 62 years, thus giving all Norwegian workers access to a pension at the same age. Further, this early retirement option is based on an actuarially fair recalculation of annual benefits.[3] Hence, there are no work disincentives at all.

The new system is designed such that the decisions regarding the timing of pension claims and the timing of employment are decoupled; i.e., one is largely free to combine labor and pension income at will, as long as annual pension claims do not exceed the annuitized value of total pension wealth (lump sum withdrawal is thus not possible).[4] A partial pension can be taken in steps of 20, 40, 50, 60, 80 and 100% of the full annual pension. The percentage can be altered annually and a full pension can be taken out at any time.

The actuarial adjustment implies that the annual pension becomes lower with early withdrawal. A precondition for early take-up is that the actuarially adjusted pension entitlement ensures a pension level at age 67 at least as high as the minimum pension, which is effective from that age. A number of workers have such low entitlements that they are prevented from drawing a (full) pension at 62 and thus have to delay claiming, either until age 67 or until their adjusted entitlements provide a pension that at age 67 equals the legislated minimum, which is defined at age 67.

Adjustments to the AFP in the private sector. Concurrently with the FTP reform, the AFP was also radically changed into an actuarially fair system for all private sector workers. The

[2] The reform implied a number of fundamental changes in the Norwegian public pension system which are *not* part of the evaluation in this paper. The most important are i) a transition from a system where pension point accumulation was based on the 20 years with highest earnings to a system where all years count equally much, and ii) the introduction of automatic longevity-adjusted annuities, implying that future increases in longevity will result in lower annual pension entitlements. These reforms will be implemented gradually, however, such that those who were close to retirement age at the time of the reform were completely unaffected by them.

[3] Deferral calculations are based on *average* life-expectancy within birth-cohorts. This implies that individuals with shorter (longer) life-expectancy than the average may find the deferral scheme disadvantageous (advantageous) for them, and thus choose to draw on their pensions as early (late) as possible, regardless of labor supply behavior.

[4] Given the progressivity of the Norwegian tax system, it may still be economically advantageous for some workers to postpone claiming the pension until they have reduced their annual labor earnings.

earnings test was completely removed, and the AFP was redesigned to become a life-long top-up annuity that could be taken only in combination with the FTP. As a result, work incentives increased dramatically for the workers covered by this pension system. Based on the detailed data used in this paper, we have computed that the average hourly net take-home wage (after tax and earnings test deductions) doubled, from $15 to $30 (implicit total tax rates declined from approximately 70 to 40 percent). But the removal of the earnings test also implied that the direct pension costs increased considerably. Our data indicate that although the average total life time benefit in the new AFP is approximately 24% lower than in the old one for persons who claimed the old AFP fully, and thus exited the labor market at age 62, the overall pension costs increased by around 42%, as the system now give valuable pension entitlements to all covered workers. Hence, with unchanged labor supply behavior, the reform of AFP would clearly add to – rather than alleviate – fiscal costs.

<u>AFP in the public sector.</u> In contrast to the private sector, the public sector AFP has not been reformed. It has preserved the pre-reform earnings test, and is still limited to the age range 62-67. Hence, workers in this sector of the economy continue to face strong labor supply disincentives. Moreover, the earnings-tested public sector AFP cannot be combined with early withdrawal of FTP, so that the liquidity option in the new FTP is open only by giving up the AFP option.

<u>Consequences for different worker groups.</u> Table 1 provides an overview of the main consequences of the reform for different worker groups, distinguished by i) their access to AFP and ii) their FTP entitlements at the earliest access age. Some of the groups identified in this table, particularly those who were subjected to changes in work incentives and/or access age (groups 2, 3, and 5) will play an important role in our empirical analysis. Note that group-assignment at the time of the reform was based on predetermined factors. Eligibility to AFP was determined by the employer's membership in the major Norwegian employer associations in combination with the worker's tenure and total work experience.[5] Entitlement to ear-

[5] The most important criteria for eligibility to the old AFP were the following: (i) current employment in a firm belonging to one of the major employer associations in the private sector (private sector AFP) or in the public sector (public sector AFP); (ii) at least 3 years' tenure with the present employer; (iii) at least 10 years of work experience since the age of 50; and (iv) an average of the 10 highest yearly incomes after 1966 exceeding an amount corresponding to approximately one-third of average full-time earnings. In the new private sector AFP, the second (tenure) criterion has been modified to require employment in at least seven of the last nine years in a firm offering private sector AFP.

ly take-up of FTP was determined by the worker's complete history of past earnings (the level of earnings in the "best" 20 years).

Table 1. The Norwegian 2011 pension reform – overview of main consequences for six different worker groups. By AFP affiliation and FTP entitlements at age 62 (percent of workers in parentheses).

	Entitled to full public pension (FTP) at age 62 after the reform	Not entitled to full public pension (FTP) at age 62 after the reform
AFP public sector	**Group 1** (28%) *No changes in either access age or work incentives.* New opportunity to start drawing on a full FTP from age 62 (with actuarial recalculation of benefits), conditional on giving up AFP entitlements.	**Group 4** (12%) *No changes in either access age or work incentives.* Depending on exact pension entitlements, a new opportunity to start drawing on a reduced (full) FTP at some time between age 62 and 66 (with actuarial recalculation of benefits) conditional on giving up AFP entitlements.
AFP private sector	**Group 2** (23%) *No change in the access age, but large increases in work incentives.* Continuation of the opportunity to draw a full AFP/FTP from age 62. Complete removal of the old confiscatory earnings test (actuarial recalculation of benefits).	**Group 5** (3%) *Increases in the access age (reduced liquidity) and large increases in work incentives.* No longer possible to claim a full pension from age 62. Depending on exact pension entitlements, a new opportunity to start drawing on a reduced (full) FTP and AFP at some time between age 62 and 66 (with actuarial recalculation of benefits).
No AFP-entitlement	**Group 3** (23%) *Reductions in the access age (improved liquidity), but no changes in work incentives.* New opportunity to draw a full FTP from age 62. No earnings test adjustment (actuarial recalculation of benefits).	**Group 6** (11%) *No changes in either access age or work incentives.* Depending on exact pension entitlements, a new opportunity to draw on a reduced (full) FTP at some time between age 62 and 66 (with actuarial recalculation of benefits).

Note: The percentage distribution is based on the analysis sample described in the next section.

<u>Announcements, Communication and Anticipation.</u> Since the AFP system has been developed over several years through a tripartite agreement between the major associations of employers and employees and the state, the new reform package was also subject to negotiations between these parties. The negotiations took place in 2008 (the private sector) and 2009 (the public sector). From around May 2009, we can assume that all the main elements of the new early retirement system were known to the workers, around two years before the reform's implementation. This includes the "new" concept of actuarially fair deferral, which was forcefully communicated by policy makers and unions, as well as by the media. By this stage it was typically no longer possible for the workers to switch between the different AFP sys-

tems by changing employer, as access to AFP entitlements in both the private and most of the public sector requires several years of sector-specific tenure.

Transitional rules. Persons in the private sector in the 1948 cohort, who reached the age of 62 in 2010, the last year before the implementation of the reform in 2011, had the possibility of choosing between the "old" earnings tested AFP (which then had to be taken out before January 1, 2011) or waiting until January 2011 to become eligible for the non-earnings-tested "new" AFP: life-long, but with a lower annual amount. Similar, but considerably less valuable, options were offered to those in the four preceding cohorts who had not taken up the old AFP. The 1947 cohort (whose members were 63 at the time of the reform) were offered 60% of the normal new AFP, the 1946 cohort 40%, the 1945 cohort 20%, and the 1944 cohort 10%.

Disability insurance. Throughout the period covered by the analysis in this paper, there has been a disability insurance program providing income replacement to workers below age 67 with at least 50% reduced work capacity due to health problems. The health problems must be assessed by a physician and verified by the social security administration. The replacement ratio in this program is around 66%. In the pension reform process, there has been a concern that restrictions in the access to early retirement may increase the demand for disability insurance.

3. Theoretical background

In order to frame and motivate our empirical analysis, we first provide a brief theoretical discussion of the decision problems facing workers approaching the age of potential retirement. We assume that workers maximize a two-period utility function, where the first period corresponds to the period in which continued work is a realistic option and the second period corresponds to a period of full retirement. Disregarding discounting, the inter-temporal utility function may then be expressed as $U = u_1(c_1, l) + u_2(c_2)$, where c_1 and c_2 are consumption in the first and second period, respectively, l is the amount of leisure, and the two *u*-functions satisfy the standard requirements of concavity and positive first order derivatives. As a starting point, we may assume that agents also face an inter-temporal budget restriction of the form $c_1 + c_2 \leq w(1-l) + W + P$, where w is the wage (which we assume fixed at the individual level), W is liquid private wealth and P is pension wealth.[6] Without credit constraints, and with the two wealth variables (W, P) considered as given, we obtain the standard results that the marginal utility of consumption is equalized across the two periods, and labor supply is determined such that the marginal utility of leisure divided by the marginal utility of consumption equalizes the net wage.

In the Norwegian pre-reform early retirement system, the pension wealth P was not given, but reduced proportionally to actual annual earnings relative to "normal" earnings, which were calculated from earnings in the years prior to the early retirement age.[7] We can write this as $P = P^*[1-(w(1-l)/w(1-l^*))]$, where P^* is the maximum pre-reform pension wealth and $(1-l^*)$ is the labor supplied prior to the early retirement access age. The pre-reform net wage after having reached the early retirement access age thus becomes $w - P^*/(1-l^*)$. For workers with high pension entitlements relative to earnings in recent years, the net wage could be extremely low, and in some cases even negative. The reform of the private sector AFP essentially removed the extra tax implicit in the earnings test and set $P = P^{**}$, with $P^{**} < P^*$, as explained in Section 2. Hence, the reform not only changed the slope of the budget line, it also changed its intercept, affecting the consumption possibilities in the full retirement state.

[6] For simplicity, we disregard general income taxation in this sub-section. These are re-introduced in our empirical specification of work incentives in the next sections.

[7] Normal earnings were defined as the average earnings in the three best out of the five last years prior to the year of early retirement.

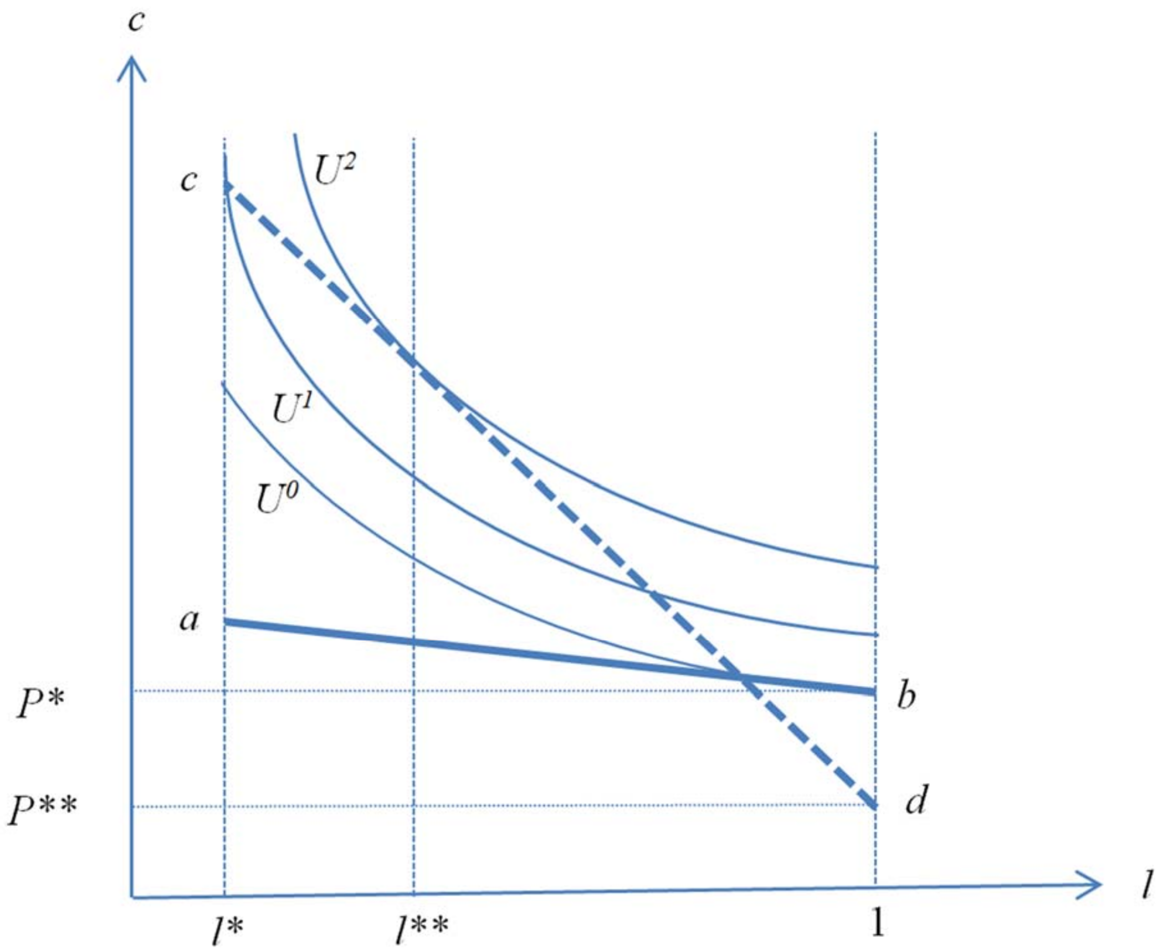

Figure 1. An example of labor supply behavior with pre-reform and post-reform budget lines for workers in group 2.
Note: The pre-reform budget line is marked *ab* and the post-reform budget line is marked *cd.*

Abstracting from the intertemporal nature of the decision problem, Figure 1 shows in a simplified form how the reform changed the situation for the majority of AFP-eligible workers in the private sector, and illustrates with an example what kind of responses we can expect to see.

The line *ab* is the budget line prior to the reform: *cd* is the budget line after the reform. Based on the pre-reform budget line, our exemplar agent would have chosen the corner solution of pulling out of the labor market ($l=1$), and obtained a utility level corresponding to U^0. Based on post-reform incentives, utility is maximized at the level U^2 by choosing the internal optimum $l=l^{**}$.

However, this alternative may not be available, since continuing with reduced hours either requires the existing employer's consent or the ability to find new employment. The latter is typically difficult for mature workers. Hence, the choice set may in many cases be confined to the alternatives of continuing as before or retiring completely, such that the agent must choose between $l=l^*$ (the pre-existing employment contract) and $l=1$ (full retirement). As we have drawn the indifference curves in Figure 1, $l=l^*$ is then the preferred choice, and with the utility level U^1 the agent still obtains higher welfare with post-reform than with pre-reform incentives.

Figure 1 also illustrates that the labor supply responses to the early retirement reform potentially involves both substitution and income effects. As both pension entitlements (P^*, P^{**}), wages (w), and initial employment (l^*) vary across individuals, the slopes and locations of budget lines will also vary and the pre-reform budget line may be almost flat, or even upward sloping. In some cases, the shift from the pre-reform to the post-reform budget line entails an approximately unchanged utility level with a pure substitution effect. But in many cases, the new incentives also involve income effects and either a welfare improvement or a welfare reduction.

For most AFP-eligible workers in the private sector who belonged to the first cohorts affected by the pension reform (group 2 in Table 1), the shifts in the slope and position of the budget line were the *only* reform element of importance; hence, we can use these workers to examine the labor supply response to a pure change in incentives. And with detailed information about the change in net-of tax returns to labor at the individual level, we can relate the labor supply responses quite precisely to the degree of incentive change.

Another potential labor supply distortion comes from liquidity constraints caused by restricted access to pension wealth P. In the context of our simplified model, a reform which raises the access age can be represented by adding to the worker's optimization problem the constraint that period 1 consumption cannot exceed the sum of period 1 labor earnings and private wealth; i.e., $c_1 \leq w(1-l)+W$. With little private wealth (or wealth/earnings from other household members), this restriction may very well be binding, and hence the worker will supply more labor in period 1. An important element of the Norwegian pension reform was to *reduce* the access age to FTP from 67 to 62 years, but with actuarial adjustment. For most private sector workers without AFP (group 3 in Table 1), this was the only reform element of importance. Hence, we can use these workers to examine the labor supply responses to a pure change of access age within an actuarially fair pension system – isolating the influence of liquidity.

As a result of the Norwegian pension reform, there was also a small group of AFP-eligible workers who were exposed to an *increase* in access age (group 5 in Table 1), as their pension entitlements were too small to make them eligible for early claiming in the new system. These faced a higher take-home wage following from the removal of the earnings test. Hence they were exposed to the combination of stronger work incentives and reduced liquidity.

4. Data and descriptive statistics

The analyses in this paper are based on individual data from merged administrative registers. These files are linked by unique encrypted personal identification numbers, and cover the entire population of Norway. The data provide detailed information on individual characteristics and labor market histories. They are not subject to the self-reporting and attrition problems common in survey-based data. Outcomes of interest include employment, earnings, hours worked, and social insurance claims (based on records from the social insurance administration), recorded in the calendar year workers reached age 63 and 64.

Observations on employment and hours worked are derived from administrative records on annual earnings, with the hourly wage rate imputed from earnings and work-hours recorded at age 60, which are available for all persons used in our analysis. Since earnings data are accurately recorded at the annual level only, all outcomes are measured as calendar year averages. We focus on the actual level of labor supply, measured by the average weekly hours of work, as well as on various qualitative outcomes describing labor market behavior relative to behavior at age 60 - two years before any early retirement option becomes available. In particular, we will be interested in the propensity to continue working more or less as before, as opposed to continuing with reduced hours or withdrawing from the labor force altogether.

The main part of our empirical analysis is based on workers who at age 60 were employed and did not receive any form of disability insurance payment. The reason why we condition our analysis on employment at age 60 rather than 61 is that we wish to minimize possible endogeneity problems associated with early (*ex ante*) responses to changes in future work incentives. Since the new early retirement system was formed through the wage agreements in 2008 and 2009 this does not necessarily eliminate the risk of endogeneity completely, as the members of the latest cohort used in our analysis (1950) reached the age of 60 in 2010 and, hence, could respond to the new incentive structures already in this year, e.g., by working more or less than they otherwise would have done. In order to assess robustness with respect to this potential endogeneity problem, we have therefore also done the empirical analysis conditioned on employment at age 58 instead. As this did not change the results to any noticeable extent beyond introducing some extra measurement error in the mapping of persons into the six groups described in Table 1, we have relegated these results to an Appendix.

In the empirical analysis, we will focus on outcomes in the calendar years in which persons reached the ages of 63 and 64, as the reform exposure at age 62 depends on the exact date of birth and since we only have access to annual earnings data. Our analysis will be based on a comparison between the 1946-1947 birth cohorts – who reached 63 in the two years prior to the reform (2009-2010) – and the 1949-1950 cohorts – who reached 63 in the years after the reform year (2012-2013).[8]

From a macroeconomic viewpoint, the outcome period used in our statistical analysis (2009-2013) was relatively stable in Norway, with the aggregate (registered) rate of unemployment varying between 2.5 and 2.9% (2.7% of the labor force in 2009, 2.9% in 2010, back to 2.7% in 2011, down to 2.5% in 2012, and up again to 2.6% in 2013); hence, we can more or less rule out that any significant changes in employment patterns were generated by cyclical fluctuations.

Table 2 provides descriptive statistics for the six groups distinguished in Table 1, including a number of labor market outcomes for the pre-reform and post-reform cohorts. The upper part of the table shows that the six groups are highly different in their composition, e.g., with respect to gender, education, and earnings levels. This potentially makes a direct between-group comparison of labor supply changes difficult to interpret. However, as can be seen in the lower part of the table, and further illustrated in Figure 2, it is only in the two groups for

[8] Because of the special incentives embodied in the transitional rules explained in Section 2, the 1948 cohort is dropped from the analyses. Even workers belonging to the two pre-reform cohorts (1946-1947) were in principle affected by the reform from age 65 or 64, respectively, as they were then allowed to claim a strongly reduced “new” AFP, provided that they had not already taken out the “old” AFP at that point. This could possibly have triggered higher labor supply already at age 63 and thus generated a corresponding reform effect even in our control group. While we will show here that there are no indications of such an effect in the data, it is worth noting that this would make us underestimate the true effects of the reform, and hence that the impact estimates reported for group 2 are on the conservative side.

which work incentives were radically changed by the reform (groups 2 and 5) that we see any major differences in labor supply between the pre- and post-reform cohorts.[9] In these two groups, the post-reform cohorts worked 6-10 more hours per week at ages 63 and 64 than the pre-reform cohorts.

[9] In a previous working paper version of this paper (Hernæs et al., 2015), we show that there were no pre-reform trends in labor supply behavior at age 63 for the 1944-47 birth-year cohorts in any of six groups, with a possible exception for the public sector workers in group 1, where there were indications of a slight pre-reform increase.

Table 2. Descriptive statistics

	Group 1 Public AFP Access to early FTP		Group 2 Private AFP Access to early FTP		Group 3 No AFP Access to early FTP		Group 4 Public AFP No access to early FTP		Group 5 Private AFP No access to early FTP		Group 6 No AFP No access to early FTP	
	Pre-reform	Post-reform	Pre-reform	Post-reform	Pre-reform	Post-reform	Pre-reform	Post-reform	Pre-reform	Post-reform	Pre-reform	Post-reform
Number of observations	18,084	19,305	15,787	15,330	15,363	15,290	8,700	7,582	2,713	2,000	8,040	6,838
Baseline characteristics												
Women (%)	47.3	52.7	17.5	20.2	12.2	14.3	96.8	96.0	91.2	88.4	71.7	69.7
Immigrants (%)	0.9	1.5	1.0	1.5	0.6	1.3	2.6	5.6	5.9	9.8	4.0	6.6
High school (%)	30.8	32.4	62.1	64.7	55.0	54.8	60.4	61.7	60.7	60.3	59.3	58.4
College (%)	63.9	62.5	19.1	18.4	27.7	30.6	18.1	19.5	4.6	5.3	13.4	16.6
Labor earnings age 60 ($1000)	91.9	92.2	104.9	101.5	103.7	105.3	55.9	57.6	57.2	55.1	54.1	55.8
Weekly work hours age 60	39.7	38.9	41.4	41.1	37.9	37.9	32.7	32.9	33.1	33.0	30.1	30.9
Outcomes												
Labor earnings ($1000)												
Age 63	70.3	72.4	58.7	72.8	84.0	84.0	41.0	42.3	31.1	48.3	42.6	44.5
Age 64	61.6	65.4	45.2	62.3	76.1	75.0	35.3	37.2	23.7	35.2	38.1	39.6
Weekly work hours												
Age 63	29.7	30.0	22.1	28.8	30.7	29.9	23.8	24.0	17.8	24.5	24.2	24.3
Age 64	25.7	26.7	16.7	24.3	27.4	26.9	20.4	21.2	13.4	20.5	21.6	21.6
Working as before (%)												
Age 63	61.5	65.7	40.8	57.4	62.2	61.2	58.2	59.9	39.6	61.5	59.6	60.9
Age 64	51.5	55.8	29.5	46.1	54.1	52.2	48.0	50.6	28.2	47.3	51.1	51.7
Working reduced hours (%)												
Age 63	26.1	24.7	35.3	30.2	27.1	28.5	23.8	24.3	32.6	25.7	26.7	25.9
Age 64	25.4	25.7	26.7	30.3	30.9	32.9	21.3	22.9	24.4	30.2	28.1	28.9
Retired without DI (%)												
Age 63	10.4	8.0	20.9	9.5	5.9	7.0	13.6	11.9	21.3	4.5	7.2	7.2
Age 64	19.4	15.7	39.3	19.1	9.3	10.4	23.7	19.6	38.7	9.3	10.7	10.1
Retired with DI (%)												
Age 63	2.0	1.6	3.0	2.9	3.7	3.2	4.4	3.9	6.4	8.4	6.5	6.0
Age 64	3.7	2.7	4.4	4.4	5.7	4.5	7.0	6.9	8.7	13.3	10.0	9.3

Note: For age 64 outcomes, data for the 1950-cohort are not available; hence post-reform outcomes at age 64 are based on the 1949-cohort only. Working “as before” is defined as having work hours at least as high as 80% of work hours recorded at age 60. “Retired without DI” is defined as being non-employed (earn less than $2000) and not receiving any disability insurance benefit. “Retired with DI” is being non-employed and receiving a disability insurance benefit.

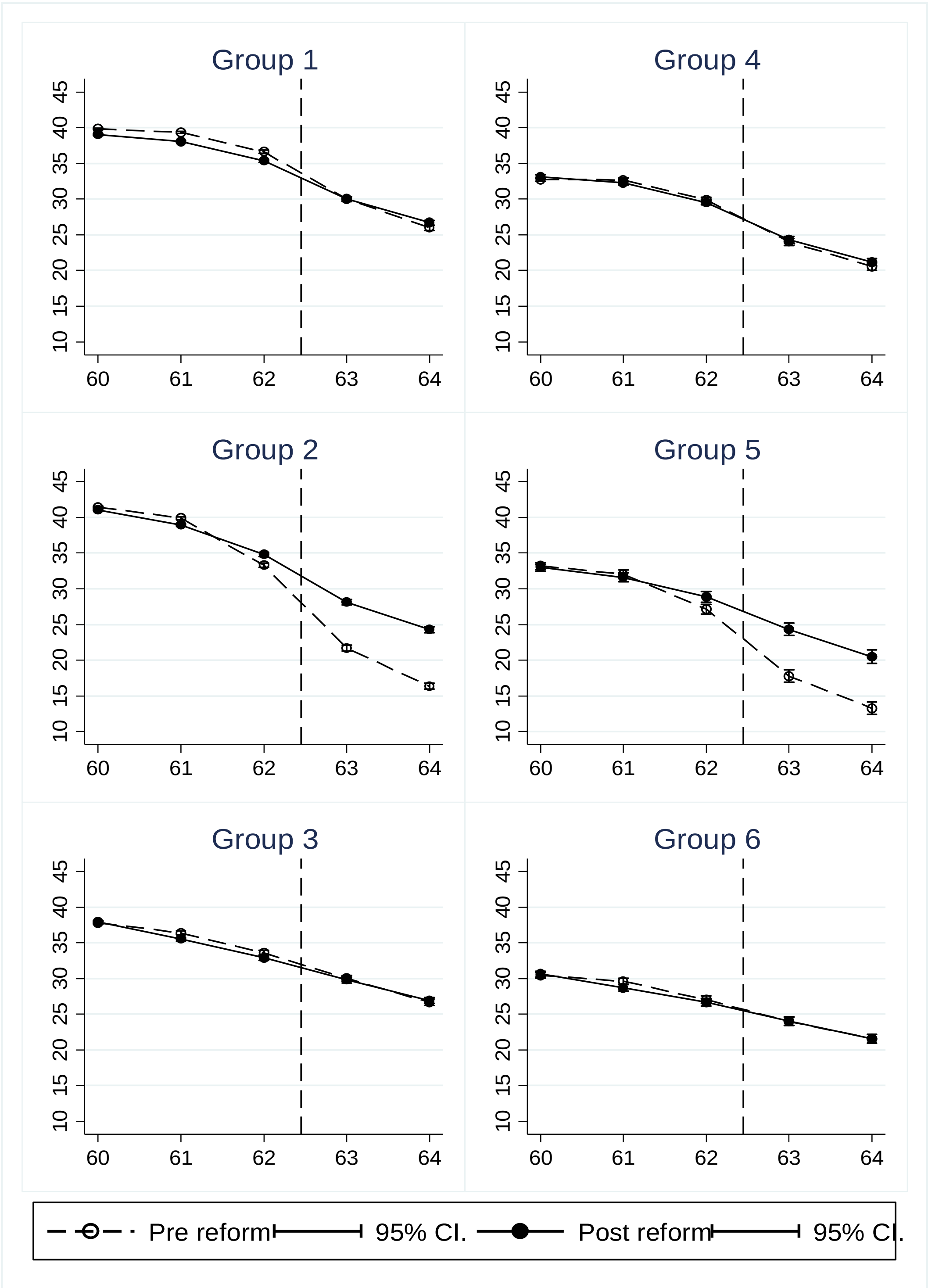

Figure 2. Average weekly work hours at age 60-64 for the last pre-reform cohort (born in 1947) and the first post-reform cohort (born in 1949), conditional on employment at age 60, with 95% confidence intervals.

Note: Since we only have one post-reform cohort that can be followed until age 64 in our data, we use only two cohorts in this graph.

5. Reform effects

Although the nature of the data and policy differentials identified in Table 1 suggests an econometric analysis based on a difference-in-differences approach, there are problems finding good control groups. The natural candidate, public sector workers in group 1 did not experience any change in the return to continued work, but they did get the option of claiming their FTP five years earlier than before. Even though this required that they gave up their AFP, around one fifth took this opportunity. As can be seen from Figure 2, the decline in labor supply at age 63 was also a bit smaller for the post-reform than for the pre-reform cohorts, indicating that the opportunity to combine work and access to the FTP was valued. The public sector workers in group 1 must therefore be viewed as “partly treated” and not well suited for a difference-in-differences approach. Similarly, workers without AFP (group 3) gained the option of claiming the FTP without any conditions other than the actuarial adjustment. In addition, as evidenced in Table 2, group composition varies considerably, particularly with respect to gender, education, and initial earnings and labor supply levels (at age 60). For example, while almost 97% of group 4 is female, women comprise less than 20% of group 2.

Given these problems, we conduct in this section a pure intra-group difference analysis for each of the three treated groups (groups 2, 3, and 5), based on a control variable approach. That is, for each labor supply outcome y_i, we estimate within-group linear regression models of the form

$$y_i = \mathbf{x_i}'\lambda + \theta R_i + \varepsilon_i, \tag{1}$$

where R_i is a dummy variable equal to one for the two cohorts affected by the reform and zero for the others, and θ is the coefficient of interest. The vector of control variables $\mathbf{x_i}$ contain a detailed description of human capital variables and earnings and hours recorded at age 60 (see the note to Table 3 for details), and ε_i is a residual. To shed light on the margins at which the labor supply responses took place, we use a number of different outcome measures in this analysis; i.e., weekly hours of work, annual earnings, employment (with unchanged or reduced hours), and retirement (with or without disability insurance). Figure 3 depicts some of these outcomes by age, for the treated and the non-treated cohorts.

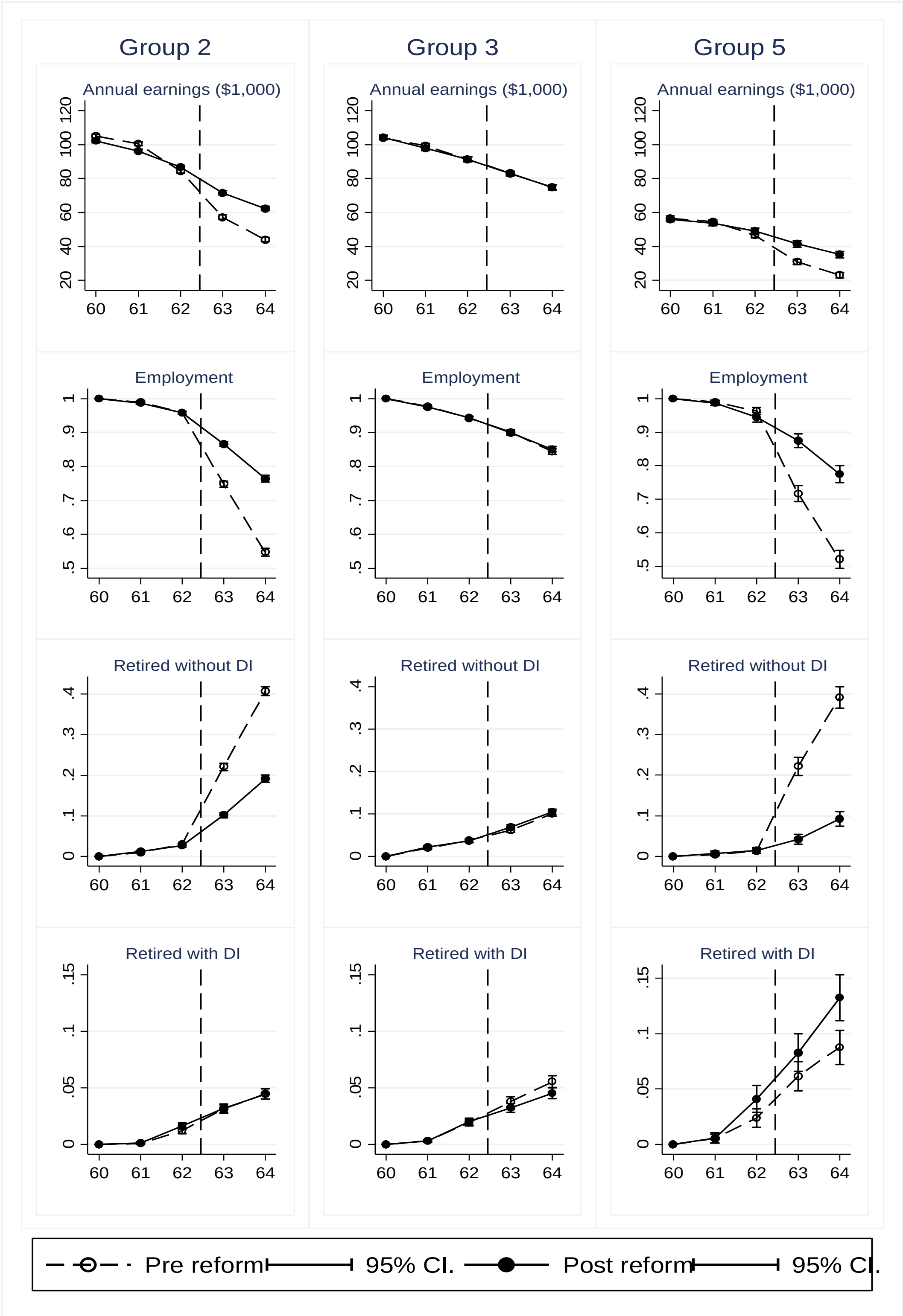

Figure 3. Labor market outcomes at age 60-64 for the last pre-reform cohort (born in 1947) and the first post-reform cohort (born in 1949), conditional on employment at age 60, with 95% confidence intervals.

Note: Since we only have one post-reform cohort that can be followed until age 64 in our data, we use only two cohorts in this graph.

Given that the treated and untreated cohorts described in Figures 2 and 3 appear to have very similar labor supply behavior before they reach the early retirement access age, an alternative strategy could have been to make an intra-group difference-in-differences analysis using the labor supply behavior prior to age 62 as the control observations. However, since we have conditioned on employment at age 60, this would only leave us with a single pure “control year”. In the Appendix, we present results from a model where we have conditioned on employment at age 58 instead, and based on these data we perform both a pure difference analysis (as in this section) and a difference-in-differences analysis with labor supply outcomes at age 59, 60, and 61 used as controls. This is done at the cost of introducing a bit more measurement error in group assignment, however, and also entails the risk that the post-reform cohort altered their labor supply behavior in response to the new incentives already before the age of 62, since they knew about the reform around age 60/61. It is notable, though, that regardless of model specification, the main results remain similar to those provided in the present section. Hence, we consider our results to be highly robust with respect to model specification. The model presented in the Appendix is similar to that used by Brinch et al. (2015) to examine employment and earnings responses to the pension reform at ages 60-64. Their results are also very similar to ours.

Table 3 presents the estimated reform effects for our main sample (employed at age 60) based on Equation (1) for each of the three groups 2, 3 and 5 and for each labor market outcome, measured at age 63 and age 64, respectively. Apart from controls for human capital, previous earnings/hours, gender and country of origin (and the fact that we included additional birth cohorts), these estimates correspond to the vertical differences between the pre- and post-reform cohorts observed at ages 63 and 64 in Figures 2 and 3.

The results are very clear, with significant and strong increases in the labor supply within the two groups subjected to increased work incentives. In the group that was subjected to higher take-home wages *only* (group 2), average labor supply increased by around 7 hours per week (30% of average pre-reform hours) at age 63 and 8 hours per week (46%) at age 64. Earnings increased a bit less in relative terms (by 28 and 41%), suggesting a slightly larger labor supply response at lower initial earnings levels. Most of the response came at the extensive margin, and in particular through a higher probability of continuing with approximately the same earnings levels (and presumably the same job) as recorded at age 60. At both ages 63 and 64, the probability of doing this is estimated to have increased by 17 percentage points. But, while this to some extent substituted for the probability of continuing with reduced hours at age 63, it added to an *increase* in the probability of continuing with reduced hours at age 64. Together, these impacts implied that the employment rate increased by 12 percentage points at age 63 and 20 percentage points at age 64. The higher employment rate substituted virtually one-for-one for regular retirement without disability insurance (DI). We find no effects at all on DI claims.

In the group that in addition to increased work incentives also experienced a removal of their early retirement option (group 5), the labor supply responses in terms of increased numbers of hours were of similar magnitude as in the group subjected to higher take-home wages only. However, since the level of labor supply in this group was much lower prior to the reform the relative hours responses were considerably larger (38% at age 63 and 53% at age 64). Again, most of the effect came at the extensive margin in the form of a higher probability of continuing "as before", with approximately the same hours as those recorded at age 60. For group 5, we also see a considerable and statistically significant spillover to disability insurance. As a result of the lost early retirement opportunity, we estimate that the probability of becoming a DI claimant increased by 1.7 percentage points (27%) at age 63 and 4.4 percentage points (51%) at age 64.

For the group that was subjected to improved liquidity in the form of lower access age to their pension wealth (group 3), the results in Table 3 indicate a slight, but statistically significant, reduction in labor supply (which we did not see in Figures 2 and 3). On average, the reform yielded a 0.9 hours (3%) reduction in weekly labor supply at age 63 and a 0.6 hours (2%) reduction at age 64. The employment rate was reduced by less than one percentage point, however. Thus, the liquidity effect appears to have operated more on the intensive than the extensive margin. Some workers took the opportunity to reduce work hours somewhat, rather than withdrawing from the labor force. We also see evidence indicating that improved liquidity triggered a small drop in disability program participation.

While the impact on hours worked from the increased liquidity was quite small, there was indeed a huge shift in pension claiming, toward the new and lower access age. Around half of those who were given access to the FTP five years earlier started claiming almost immediate-ly (not shown in the table). This may seem at odds with the small reduction in labor supply and indicates that the pension only to a small degree was used to finance reduced labor earnings. Apart from financing work reduction, there could be a number of reasons for the early claiming, such as risk aversion and private mortality information.[10] While this is not the topic of this paper, we note that the claiming spike is in line with other studies. Brinch et al. (2015) report such a spike in their analysis of the 2011 Norwegian reform.

Table 3. Impact of reform on labor market outcomes. Comparing pre- and post-reform cohorts.

	Group 2 Private sector workers eligible for AFP and post-reform FTP (standard error)	Group 3 Not eligible for AFP but for post-reform FTP (standard error)	Group 5 Private sector workers eligible for AFP but not post-reform FTP (standard error)
Weekly hours worked			
Age 63	6.805 *** (0.317)	-0.901*** (0.284)	6.685*** (0.433)
Age 64	7.753 *** (0.322)	-0.591 (0.364)	7.100*** (0.673)
Annual earnings ($1000)			
Age 63	16.617 *** (0.895)	-2.177*** (0.800)	11.001*** (0.689)
Age 64	18.746*** (1.098)	-1.817** (0.924)	12.096*** (1.044)
Probability of employment			
Age 63	0.117 *** (0.007)	-0.008* (0.005)	0.153*** (0.011)
Age 64	0.204 *** (0.008)	-0.002 (0.007)	0.252*** (0.017)
Probability of employment “as before” (>80% of age 60 hours)			
Age 63	0.168*** (0.008)	-0.031*** (0.009)	0.218*** (0.013)
Age 64	0.168*** (0.010)	-0.026*** (0.010)	0.188*** (0.013)
Probability of employment with reduced hours			
Age 63	-0.061*** (0.006)	0.023*** (0.006)	-0.065*** (0.009)
Age 64	0.037*** (0.009)	0.024*** (0.008)	0.064*** (0.012)
Probability of retirement without DI			
Age 63	-0.115 *** (0.006)	0.009*** (0.003)	-0.170*** (0.009)
Age 64	-0.204 *** (0.008)	0.009* (0.005)	-0.295*** (0.016)
Probability of retirement with DI			
Age 63	-0.002 (0.002)	-0.002 (0.002)	0.017*** (0.006)

[10] An additional reason could be to obtain the new AFP, which can be claimed only in combination with the FTP.

Age 64	-0.000 (0.002)	-0.007* (0.004)	0.044*** (0.007)
Number of observations age 63	31,116	21,101	4,711
Number of observations age 64	23,384	15,725	3,737

Note: The table reports estimates of the coefficient θ in Equation (1). The control variable vector $\mathbf{x}_i$ includes gender, education (nine fields and eight levels), country of origin for immigrants (five regions), and labor earnings and weekly hours of labor supply at baseline (age 60). Standard errors are clustered on educational groups (6 levels and up to 10 fields, in total 34 educational groups) and birth cohort, in total up to 136 clusters.
* (**) (***): Significant at the 10 (5) (1) % level.

6. Labor supply elasiticies

The analysis in the preceding section disregards the large variation in reform-initiated incentive changes within the groups of treated workers, and thus fails to relate the labor supply responses to the intensity of incentive increases. In the present section, we therefore focus exclusively on the group exposed to increases in the take-home wage rate (group 2), and seek to quantify the labor responses in relation to the sizes of the incentives changes. For the group as a whole, we have already estimated that that total labor supply (work hours) increased by 30% at age 63 and by 46% at age 64 as a result of the reform. Since, as explained in Section 2, the reform on average entailed a doubling of the take-home wage for workers in this group, the implied aggregate labor supply elasticities are around 0.3 at age 63 and 0.46 at age 64. We now exploit within-group variation in incentive changes to shed further light on the magnitudes of the labor supply responses.

Unfortunately, it is not possible to observe the same individuals' age-63-behavior both before and after the reform; hence we cannot examine the relationship between individual changes in labor supply and individual changes in incentives. But we *can* examine labor supply outcomes for the pre- and post-reform birth cohorts at different fixed positions in the hypothetical/actual incentives-change distribution. Figure 4 illustrates how we can use this idea to provide a graphical illustration of reform effects similar to Figures 2 and 3. Here, we have divided each birth-cohort into quartiles, based on individual positions in the distribution of hypothetical incentives changes relative to age-60-earnings. The figure then shows *by quartile* the differences between average labor market outcomes for the affected 1949/1950-cohorts and for the two unaffected 1946/1947-cohorts.

The pattern is again very clear: The relative increase in labor supply at ages 63 and 64 were larger the larger was the reform-initiated relative change in the take-home wage. For example, work-hours responses at age 63 (age 64) vary from around 25% (38%) for the quartile with the smallest relative incentive increase (a 56% increase in the average take-home wage) to around 42% (68%) for the quartile with the largest increase (a 268% increase in the average take-home wage). It is notable, though that the implied elasticities are largest for the quartiles with the smallest relative incentive increases, a point we discuss in more detail below.

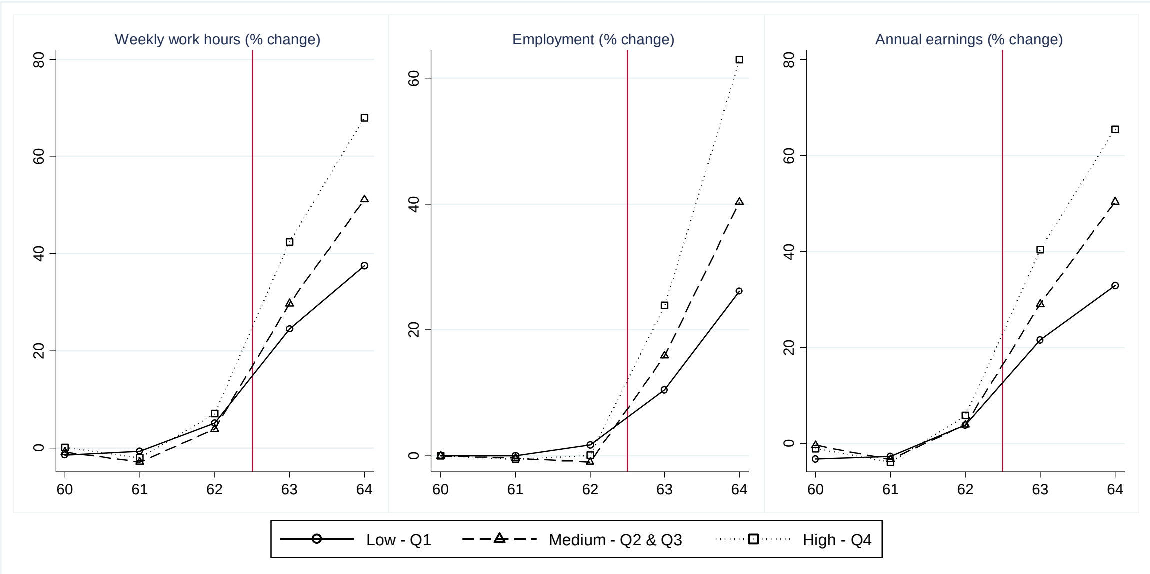

Figure 4. Relative changes in labor supply by quartile in the distribution of relative changes in work incentives.

Note: The graphs show the relative difference in labor supply between the first post-reform cohort (1949) and the last pre-reform cohort (1947) by position in the distribution of relative work incentives. The reform-generated changes in average work incentives was 56% in the first quartile, 116% in the combined second and third quartile, and 268% in the fourth quartile.

We now turn to a more formal statistical analysis intended to exploit the reform-initiated variation in incentives across cohorts in group 2 to quantify the causal relationship between labor supply and work incentives. We set up a regression model similar to (1), but instead of using a reform dummy as the key explanatory variable, we now specify the model directly with the actual work incentive (to be explained in detail below) as the variable of interest. In general, work incentives are not randomly assigned, as persons with different take-home wages are likely to differ systematically along other dimensions as well, such as health status, motivation, job characteristics, and valuation of leisure. Hence, in order to facilitate estimation of the *causal* relationship between work incentives and labor supply, we need to isolate the random assignment-like variation generated by the reform. We do this by estimating regressions where we not only include the actual work incentive as the central explanatory variable, but also add in the corresponding *hypothetical* pre-reform and post-reform work incentives. The basic idea is that while any non-causal correlations between labor supply and pre- and post-reform work incentives should be the same before and after the reform, the causal correlation should shift toward the incentives *actually applying*. Hence, by studying changes in the correlation patterns pre- and post-reform, we can trace out the causal effects.

More specifically, we set up regression models that link labor market outcomes y_i (see Table 4) directly to the economic returns to work for the members of group 2:

$$y_i = \mathbf{x_i}'\lambda + \delta\Delta_i^O + \gamma\Delta_i^N + \varphi\left[(1-R_i)\Delta_i^O + R_i\Delta_i^N\right] + \zeta_i, \quad (2)$$

where (Δ_i^O, Δ_i^N) are individual i's net economic returns to work as they *would have* applied under the old (Δ_i^O) and new (Δ_i^N) pension systems, respectively, and R_i is (still) a dummy variable equal to 1 for workers belonging to the reform cohorts (1949-50) and equal to 0 for the pre-reform cohorts (1947-48). The term $(1-R_i)\Delta_i^O + R_i\Delta_i^N$ thus gives the work incentive actually applying for both the pre- and the post-reform cohort members. The vector of control variables $\mathbf{x_i}$ includes the same variables as before (gender, education, country of origin for immigrants, and labor earnings and weekly hours of labor supply at baseline (age 60)).

The coefficient of interest, φ, represents the causal effect of the work incentive on the outcome variable y_i. We specify the work incentives (Δ_i^O, Δ_i^N) as the net hourly (take-home) wages. Abstracting from the tax system, we would have had $\Delta_i^O = w - P^*/(1-l^*)$ and $\Delta_i^N = w$, where $(1-l^*)$ is labor supplied prior to the early retirement access age; see Section 3. However, when we take the complete tax system into account (including the degree of progressivity and the differential treatment of pension and labor earnings), the net hourly wage rate becomes dependent on the level of labor supply. Hence, in order to use (Δ_i^O, Δ_i^N) as exogenous explanatory variables, we compute the net hourly wage associated with a fixed labor supply level. Since the descriptive evidence in the previous section indicated that the reform primarily changed the probability of continuing working "as before", we have chosen to compute the net hourly wage at the baseline (age 60) level of labor supply; i.e., (Δ_i^O, Δ_i^N) are the net hourly wages at age 63 and 64 derived from continuing working as at age 60. Figure 5 illustrates how these net hourly wages changed as a result of the reform. It is evident that the reform shifted the whole work incentives distribution to the right, and for the vast majority of workers, the hourly take-home wage increased by between $13 and $18. At the individual level, most workers experienced relative increases in the take-home wage exceeding 100%, but the largest percentage increases occurred for persons who prior to the reform had take-home wages close to zero.

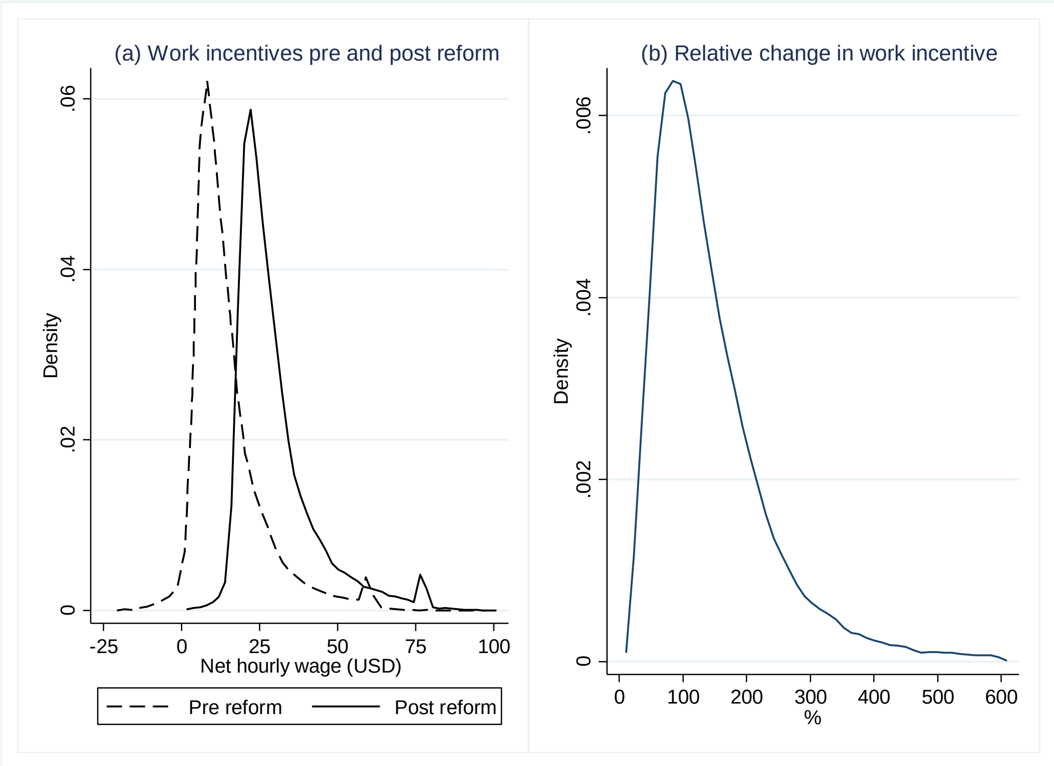

Figure 5. The distributions of predicted net hourly wage rates at percentage changes due to the reform at ages 63 and 64 with pre- and post-reform rules (group 2 only).
Note: Net hourly wage rates are predicted on the basis of the calculated gross wage rates at age 60.

The absorbing nature of retirement also means that there is an option value to continuing to work, which is not reflected in our work incentive variable (Stock and Wise 1990; Gruber and Wise, 1999). This holds both before and after the reform, but after the reform, the return to work is higher. Since this differential extends over more than the first two years, the value of a given option of work will also be higher. Therefore, compared to a situation where workers can freely move in and out of employment, we probably underestimate the increase in the incentive and overestimate the per dollar labor supply responses. Yet, our estimates do appropriately represent the behavioral responses to the repeal of the earnings test.

The main outcome of interest is the weekly number of work-hours supplied, but we also examine discrete outcomes measuring labor supply at age 63/64 relative to labor supply at age 60. With the number of work-hours used as the dependent variable in Equation (1), it may be argued that the relationship between the wage rate and hours worked should be formulated in terms of logarithms, enabling the computation of labor supply elasticities. This is problematic in the present setting, however, both because many workers in the pre-reform cohorts had extremely low, and in some cases even negative, net wage rates, and because persons who exit the labor market work zero hours. We thus estimate the relationship between hours worked and the net wage with a linear model. However, to facilitate comparison with findings in the existing literature, we compute labor supply elasticities at more aggregate levels; i.e., we report the percentage changes in average labor market outcomes relative to the percentage change in average incentives for groups of workers.

Our main results are presented in Table 4. Each dollar increase in the net hourly wage raises weekly labor supply (annual earnings) at age 63 by 0.394 hours ($976) and at age 64 by 0.453 hours ($1107). The labor supply response comes about almost exclusively by inducing more workers to stay on with approximately the same level of employment that they had at age 60. For each dollar increase in the hourly net wage, the probability of staying on "as before" rises by approximately 1 percentage point at both ages 63 and 64. However, while this response to some extent comes about through a reduction in reduced-hours work at age 63 it is accompanied by an increase in reduced-hours work at age 64. Hence, the overall labor supply response is considerably larger at age 64 than at age 63. This reflects the absorbing nature of labor market exits at mature age. Those who leave employment at age 63 due to poor work incentives rarely return to work at age 64 (or later); hence each year's new retirees just add to the total number of exits.

Since we do not follow workers over time, but compare separate cohorts with different sets of incentives, we cannot directly decompose the labor supply responses into extensive and intensive margin effects (Chetty et al., 2011b). We are simply not able to identify the number of hours worked among the workers who would have chosen to be employed with both pre-reform and post-reform work incentives. However, as the main reform effect appears to have been that a much larger group of workers tends to continue working with roughly the same hours as they had at age 60 instead of pulling out of the labor market, it appears that the extensive margin is of major importance.

Table 4. Estimated labor supply effects of the net hourly wage rate

	I Marginal effect measured in hours or percentage points (pp.) of a dollar increase in net wage (standard error)	II Implied aggregate elasticity [95% confidence interval]
Weekly number of work-hours		
Age 63	0.394 hours*** (0.019)	0.305*** [0.263, 0.345]
Age 64	0.453 hours*** (0.019)	0.452*** [0.380, 0.509]
Annual earnings		
Age 63	$976*** (55)	0.250*** [0.215, 0.285]
Age 64	$1107*** (70)	0.368*** [0.314, 0.431]
Probability of working "as before" (>80% of age 60 hours)		
Age 63	0.969 pp. *** (0.051)	0.414*** [0.354, 0.476]
Age 64	0.984 pp. *** (0.054)	0.570*** [0.471, 0.673]
Probability of working with reduced hours		
Age 63	-0.292 pp. *** (0.042)	-0.159*** [-0.214, -0.103]
Age 64	0.206 pp. *** (0.050)	0.129*** [0.031, 0.206]
Probability of being retired without disability insurance		
Age 63	-0.664 pp.*** (0.039)	-0.651*** [-0.727, -0.576]
Age 64	-1.185 pp. *** (0.052)	-0.594*** [-0.670, -0.512]
Probability of being retired with disability insurance		
Age 63	-0.013 pp. (0.012)	-0.118 [-0.349, 0.085]
Age 64	-0.005 pp. (0.016)	-0.051 [-0.310, 0.136]

Note: Column I reports the estimates of the coefficient φ in Equation (1). The control variable vector $\mathbf{x}_i$ includes gender, education (nine fields and eight levels), country of origin for immigrants (five regions), and labor earnings and weekly hours of labor supply at baseline (age 60). Column II reports aggregate elasticities calculated as the predicted reform-initiated percentage change in the average outcome divided by the percentage change in the average net take-home hourly wage. Number of observations: 31,117 for age 63 outcomes and 23,385 for age 64 outcomes. Standard errors are clustered on educational groups (6 levels and up to 10 fields, in total 34 educa-

tional groups) and birth cohort, in total 136 clusters. The estimate and confidence intervals for the aggregate elasticities are based on clustered non-parametric bootstrap with 500 trials.
* (**) (***): Significant at the 10 (5) (1) % level.

For the group as a whole, our estimates imply labor supply elasticities equal to 0.31 at age 63 and 0.45 at age 64; i.e., the same as we obtained based on the reform dummy approach in Section 5. This is at the higher end of typical estimates of the compensated (Hicksian) labor supply elasticities reported in the literature; see, e.g., the recent reviews in Chetty et al. (2011a; 2011b), McClelland and Mok (2012), and Chetty (2012). Given that our estimates also embody income effects that presumably have dampened the substitution effects somewhat (see the previous subsection) this may at first sight appear surprising. However, an important distinguishing feature of our analysis is that we examine the labor supply behavior in a population where everyone is employed to start with (at age 60), and hence can be assumed to have at least one feasible employment option (due to employment protection legislation). Moreover, we study these individuals at a time where non-employment is a realistic alternative, given that they have obtained access to their pension wealth. Finally, the incentive changes we use to identify the labor supply responses are large, compared to margins of variation typically encountered in the literature, implying that inertia caused by frictions of the type discussed in Chetty (2012) is of minor importance.

As we point out above, the elasticity concept is somewhat problematic in our setting, given the large number of zero outcomes and, in particular, the wide range of reform-generated relative incentive changes. To illustrate this, we have divided the group 2 population into 10 deciles based on the size of the percentage change in the net hourly wage rate generated by the reform, and estimated the models separately for each group. Figure 6 reports the resulting estimated marginal effects on hours worked at age 63, as well as the implied group-elasticities (defined as the percentage change in average labor supply divided by the percentage change in average work incentives). On the horizontal axis, we report the average percentage net wage increase generated by the reform in each decile; hence, it can be seen that these incentive increases varied from around 40 to more than 400 percent.

A first point to note from Figure 6 is that the marginal effects are relatively stable across the 10 groups, indicating that our linear model specification probably represents the data fairly well. A second point is that the implied group-elasticities decline monotonically with the size of change in the relative incentives. One reason for this appears to be that there are “natural” limits to the size of the labor supply responses in our setting, given that some persons choose to work almost regardless of work incentives. In our data, the maximum reform response across the 10 deciles appears to be a labor supply increase at age 63 of approximately 30-40%, and responses of this magnitude are achieved in all groups experiencing net wage increases exceeding 100%. Finally, for the more moderate incentive changes observed in the first decile (with an increase in average net wage of 40%), the implied labor supply elasticity is considerably larger than for the population as a whole.

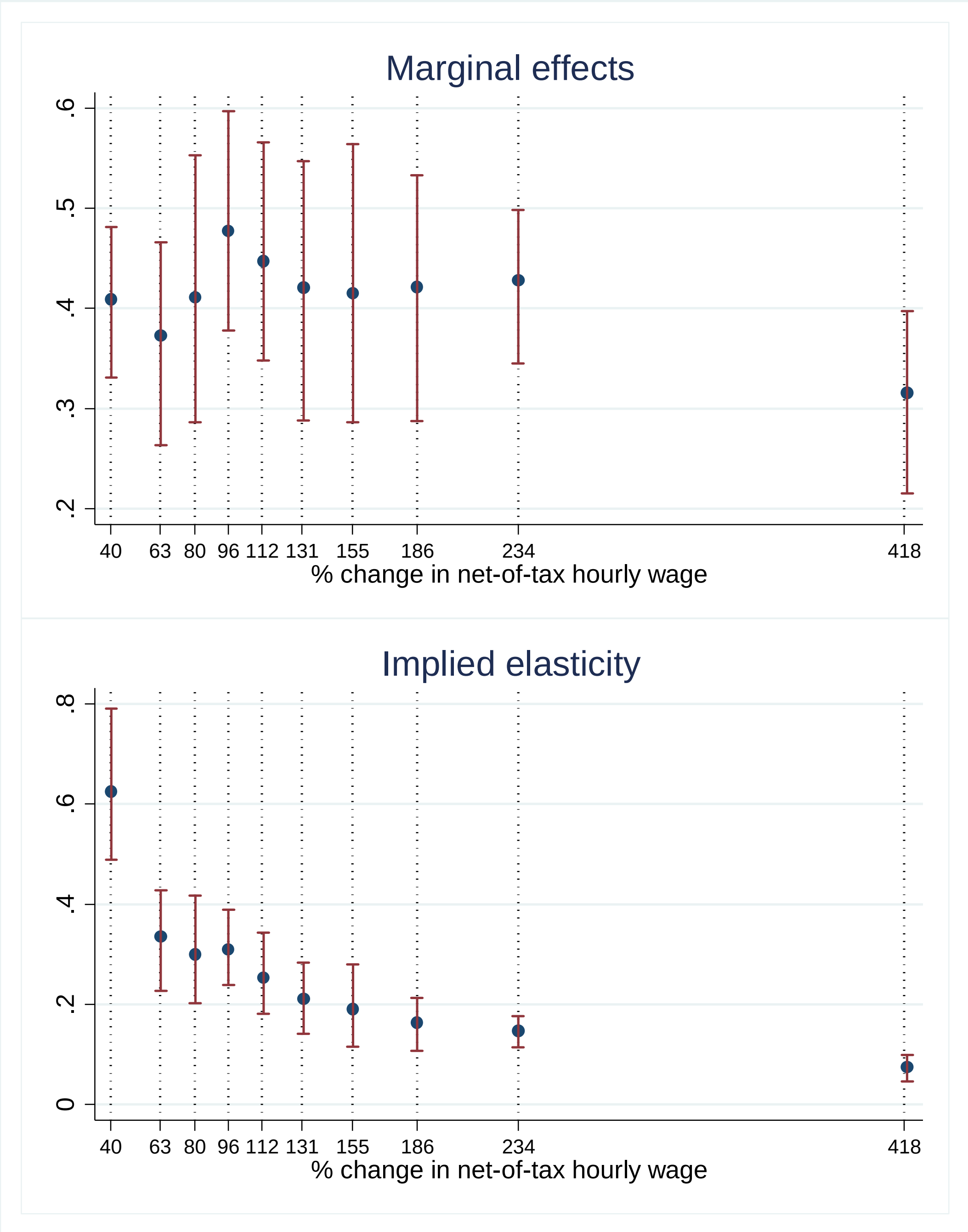

Figure 6. Estimated hours-of-work responses to a dollar increase in the net take-home wage rate, by decile in the distribution of percentage increase in the net wage generated by the reform and with 95% confidence intervals.

Note: The numbers on the horizontal axis indicate the percentage increase in average incentives within each of the 10 deciles, i.e. $100 \times (\overline{\Delta}^{n} - \overline{\Delta}^{o}) / \overline{\Delta}^{o}$. Confidence intervals are based on clustered (by education and birth-year, in total 136 clusters) non-parametric bootstraps with 120 trials.

7. Discussion and concluding remarks

Many countries around the world are trying to encourage greater mature labor force participation, as a response to increasing mature age life expectancy. In this paper we analyze the labor supply response to a comprehensive reform of the Norwegian retirement system introduced in 2011. The reform followed policy debate in which two possible reform options were discussed: To increase the earliest access age (leaving annual pension benefits untouched); or to remove all disincentives to work, allowing workers to claim pension benefits while working, with actuarially fair deferral, rather than increasing the access age.

The Norwegian reform embraced the second of these approaches. Essentially, it combined automatic longevity adjustment with extension of pension entitlements that were previously reserved for those who actually stopped working (or reduced hours considerably) to all covered workers, regardless of labor supply behavior. Among the obvious advantages are that it deals with heterogeneity in the circumstances – health, wealth, family – of older workers. Relatedly, spillovers into disability programs, often a consequence of increasing access age, might be expected to be quite muted, since continuation in the workforce has longer term payoffs, allowing significant wage income to coincide with benefit receipt.

The reform reduced the large differences in retirement income support structures among groups of private sector employees. In this paper, we have analyzed the incentive and liquidity impacts of these changes, with a particular focus on labor supply. Our results suggest a strong labor supply response to the incentive changes, especially at the extensive margin. Many workers, who would have exited the workforce had they confronted the pre-reform incentives, now continue to work more or less as before. This is consistent with the findings of Brinch et al. (2015), who analyse the same reforms with an emphasis on the retirement spike at age 62.

The drive to encourage mature labor force participation is importantly informed by the need for fiscal sustainability in the face of an ageing demographic. Breaking the nexus between pension access and retirement ages looked to be expensive, since the removal of the earnings test meant that pensions would be payable regardless of labor force participation. To contain the overall costs of the system, the maximum (present) value of the early retirement pension was reduced considerably. However, in order to ease the reform process –to make it more acceptable for the workers that were closest to the early retirement age, the first affected cohorts (those born before 1963) were largely sheltered from these reductions through a separate "compensation benefit". Hence, to make the reform a fiscal success, particularly for these first cohorts, a substantial labor supply response is required.

Based on the data and the estimated labor supply responses in this paper we can estimate the changes in public costs as well as the increases in tax payments to assess the overall fiscal consequences of the Norwegian pension reform. Clearly, such a calculation is based on a series of assumptions regarding discount rates (3%), longevity (current life-expectancy tables) and labor supply responses for the ages not considered in this paper (responses at age 62 assumed equal to those at age 63 and responses at ages 65 and 66 assumed equal to those at age 64). All values are in present values net of tax.

Starting with pension costs prior to the reform, we estimate the cost of the early retirement program, for persons who left the labor force at the earliest occasion, at approximately \$137,000 per retiree. But, since take-up was well below 100% the average cost per eligible worker was lower, at only \$73,000. In the new system, pension payments were around \$104,000 per eligible worker. Hence, the direct fiscal impact of the reform was to raise the pension costs by around \$31,000 per worker, or 42%. This increase was primarily caused by the extra compensation benefit given to AFP-eligible workers born before 1963. For subsequent cohorts, who will confront a well-established reform, we estimate that the overall cost of the new AFP is approximately \$78,000, just \$5,000 more than the old one.

However, the reform also raised labor supply, and corresponding tax payments, considerably. Based on the assumptions listed above, we estimate that the reform raised overall income tax by \$32,000 per worker over the five-year period from age 62 to 66.[11] Hence, even though the reform entailed considerable additional pension expenditures for the first affected cohorts, the fiscal contribution was above the self-financing level, even for workers born before 1963. For younger workers, the net fiscal contribution is much larger.

The idea of encouraging workers to compensate for reduced annual pensions by removing all disincentives to work more has something of the flavor of the Feldstein (2005) proposal that the government can make up for a reduced public pension by stimulating occupational pen-

[11] In this calculation we have assumed a 30% average income tax rate on labor earnings plus a 14.1% payroll tax.

sions. The evidence we presented in Section 5 actually indicates that the added labor supply effect obtained by also raising the access age is small. Moreover, whereas higher access age appears to entail a spillover to disability pension programs, no such unintended side effects are identified for those who were exposed to increased work incentives only.

Our findings also suggest that as long as the size of the pension wealth is fixed, there are small – almost negligible – labor supply responses to changes in the access age. Hence, liquidity does not seem to play a major role in the determination of labor supply for mature workers. This result is atypical in relation to the existing literature. However, the studies we have found on changes in the access age all include changes either in work incentives or in the size of pension wealth as well. A study that comes close to our own is by Behaghel and Blau (2012), where impacts of a change in the full retirement age are investigated in a setting where there were no accompanying changes in work incentives. The authors find that this entailed considerable labor supply responses, which they interpret as a result of "reference dependence" with loss aversion. We would argue that such effects are not likely to be relevant in our case, since a majority of workers already had 62 years as the access age, and since the reference point of a full retirement age of 67 was not changed.

The results presented in this paper, as well as other recent studies on preceding pension rules adjustments in Norway (Hernæs and Jia, 2013; Brinch et al., 2015 and Brinch et al., 2017) point to a highly elastic labor supply behavior among elderly workers. Transparent, substantial, and successfully communicated increases in work incentives may constitute a highly effective strategy for increasing mature labor force participation. If these results are supported by studies elsewhere, they will have major implications for pension reform designed to increase mature labor force participation.

Appendix

In this appendix, we present all the estimated reform effects from Section 5 with age 58 used as baseline instead of age 60; i.e., we condition on employment at age 58 and measure the discrete employment outcomes relative to that age. In this exercise, we use both the difference-approach applied in Section 5 and a difference-in-differences (DiD) strategy where we use labor supply at ages 59, 60, and 61 as "control years". The DiD estimates are derived from the following linear regression model:

$$
\begin{aligned}
&y_{ia} = \pi_c + \tau_a + \mathbf{x_i}'\lambda + \theta^{DiD} R_i I(a > 62) + \omega_i, \\
&a = 59, 60, 61, 63 \text{ or } a = 59, 60, 61, 64, \\
&c = 1946, 1947, 1949, 1950.
\end{aligned}
\tag{A1}
$$

Here, π_c is a cohort fixed effect, τ_a is an age fixed effect, $\mathbf{x_i}$ is a vector of individual covariates (the same as in Section 5, only now measured at age 58), R_i is (still) a dummy equal to 1 for the treated cohorts (1949, 1950) and $I(a > 62)$ is a dummy equal to one for age>62 years.

Figure A1 first summarizes the observed labor market outcomes by age. The patterns of labor supply at ages 63 and 64 are similar to those described in Sections 4 and 5 with employment conditioned at age 60. However, what we also see from Figure A1 is that there are hardly any differences at all between the treated and untreated cohorts prior to age 62. This observation forms the basis for the DiD-analysis captured by Equation (A1).

Table A1 presents the estimated reform effects, based on the model used in Section 5 (Equation 1) and based on the DiD-model (Equation A1). An important conclusion coming out of this exercise is that the results based on employment conditioned at age 58 are very similar to those obtained based on employment conditioned at age 60, but in most cases slightly smaller. This is as expected, as a larger fraction of the workers employed at age 58 will have left the labor force before early retirement became a relevant option, and also as the measurement error in group assignment becomes a bit larger.[12] A second conclusion is that the simple difference-methodology (Equation 1) and the DiD-methodology (Equation A1) give very similar results.

[12] The reason for this is that the correct assignment into the six groups described in Table 1 is based on work experience and tenure at age 62. As a basis for our analysis we must predict group assignment (conditional on continued employment) on the basis of observed characteristics in the base-year (at age 58 or age 60, respectively). This probably cause a slight attenuation bias, as a fraction of the workers change their employment status (and/or hours) already before age 62. This attenuation bias is likely to be larger the earlier we make the group assignment.

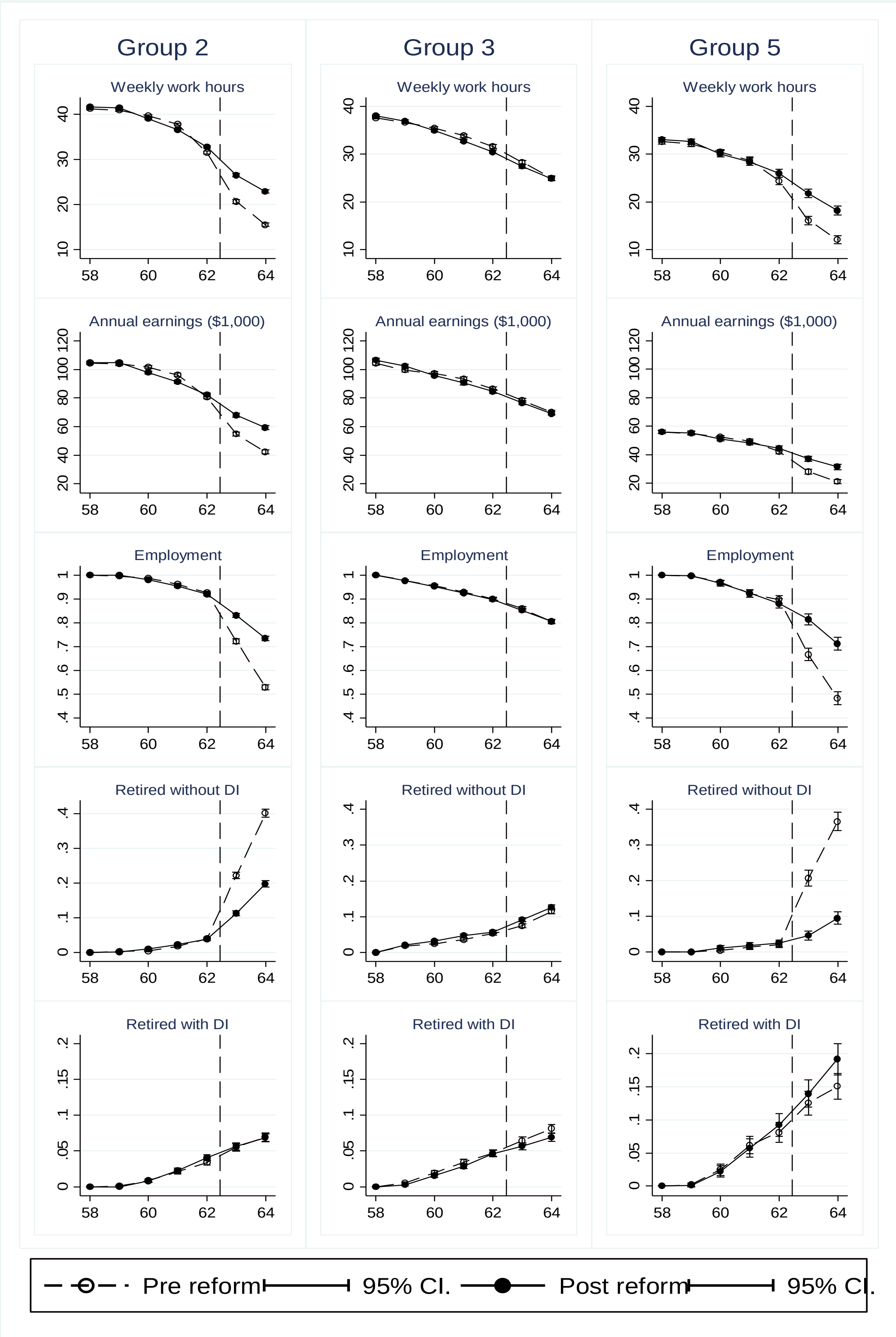

Figure A1. Labor market outcomes at age 58 to 64 for the last pre-reform cohort (born in 1947) and the first post-reform cohort (born in 1949), conditional on employment at age 58. With 95% confidence intervals.

Note: Since we only have one post-reform cohort that can be followed until age 64 in our data, we use only two cohorts in this graph.

Table A1. Impact of reform on labor market outcomes. Comparing pre- and post-reform cohorts based on the population working at age 58 (standard errors in parentheses).

	Group 2 Private sector workers eligible for AFP and post-reform FTP (standard error)		Group 3 Not eligible for AFP but for post-reform FTP (standard error)		Group 5 Private sector workers eligible for AFP but not post-reform FTP (standard error)	
	Diff	DiD	Diff	DiD	Diff	DiD
Weekly hours worked						
Age 63	5.320*** (0.277)	6.472*** (0.721)	-1.614*** (0.284)	-0.615 (0.412)	5.185*** (0.373)	5.958*** (0.619)
Age 64	6.788*** (0.331)	7.296*** (0.863)	-0.940*** (0.314)	-0.317 (0.620)	5.435*** (0.472)	5.993*** (0.561)
Annual earnings ($1000)						
Age 63	12.501*** (0.737)	16.282*** (1.073)	-4.004*** (0.751)	-1.817** (0.795)	8.340*** (0.658)	10.318 (0.804)
Age 64	16.275*** (1.008)	18.282*** (1.487)	-2.725*** (0.886)	-1.255 (1.131)	9.188*** (0.840)	10.699*** (0.851)
Probability of employment						
Age 63	0.102*** (0.006)	0.108*** (0.014)	-0.013*** (0.005)	-0.006 (0.009)	0.132*** (0.011)	0.137*** (0.021)
Age 64	0.193*** (0.008)	0.195*** (0.025)	-0.002 (0.007)	0.003 (0.014)	0.214*** (0.014)	0.221*** (0.022)
Probability of employment “as before”						
Age 63	0.113*** (0.007)	0.152*** (0.018)	-0.047*** (0.008)	-0.030*** (0.010)	0.171*** (0.012)	0.199*** (0.016)
Age 64	0.134*** (0.009)	0.152*** (0.020)	-0.040*** (0.009)	-0.027* (0.015)	0.130*** (0.011)	0.157*** (0.014)
Probability of employment with reduced hours						
Age 63	-0.011* (0.006)	-0.044*** (0.007)	0.034*** (0.006)	0.024*** (0.006)	-0.039*** (0.010)	-0.062*** (0.017)
Age 64	0.059*** (0.008)	0.042*** (0.013)	0.037*** (0.008)	0.030*** (0.010)	0.084*** (0.011)	0.063*** (0.023)
Probability of retirement without DI						
Age 63	-0.102*** (0.005)	-0.106*** (0.008)	0.015*** (0.003)	0.007* (0.004)	-0.149*** (0.009)	-0.154*** (0.009)
Age 64	-0.190*** (0.008)	-0.191*** (0.015)	0.013** (0.005)	0.007 (0.007)	-0.264*** (0.013)	-0.268*** (0.012)
Probability of retirement with DI						
Age 63	0.000 (0.002)	-0.002 (0.007)	-0.002 (0.003)	-0.002 (0.006)	0.017** (0.008)	0.017 (0.015)
Age 64	-0.003 (0.003)	-0.004 (0.011)	-0.011* (0.006)	-0.009 (0.010)	0.050*** (0.009)	0.048** (0.020)
Number of observations age 63	30,422	121,688	23,454	93,816	4,710	18,840
Number of observations age 64	22,665	113,931	17,432	87,794	3,702	17,832

Note: The results in columns marked with “Diff” are estimated on the basis of Equation (1), whereas results marked with “DiD” are estimated on the basis of Equation (A1).
*(**)(***) Significant at the 10(5)(1) % level.

Acknowledgements

Data received from Statistics Norway and from the early retirement administration unit have been essential for the paper. The data are used in compliance with the rules given by the Norwegian Data Inspectorate. We gratefully acknowledge support from the Research Council of Norway (Evaluation of the pension reform) and the ARC Centre of Excellence in Population Ageing Research (CEPAR). We thank Fredrik Haugen, Steinar Holden, Ole Christian Lien, and participants at seminars at ESOP at the University of Oslo, CEPAR, and School of Economics at the University of New South Wales for valuable comments.

References

Baker, M. and Benjamin, D., 1999. How do retirement tests affect the labor supply of older men? *The Journal of Public Economics* 71, 27-51.

Behaghel, L. And Blau, D. M., 2012. Framing social security reform: Behavioral responses to changes in the full retirement age. *American Economic Journal: Economic Policy* 4 (4), 41-67.

Bratberg, E., Holmås, T. H. and Thøgersen, Ø., 2004. Assessing the effects of an early retirement program. *Journal of Population Economics* 17, 387-408.

Brinch, C., Hernæs, E. and Jia, Z., 2017. Salience and social security benefits. *Journal of Labor Economics,* forthcoming.

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

Chan, S. and Stevens, A. H., 2008. What you don't know can't help you: Pension knowledge and retirement decision-making. *The Review of Economics and Statistics* 90 (2), 253 – 266.

Chetty, R., 2012. Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica* 80 (3), 969-1018.

Chetty, R., Guren, A., Manoli, D. S., Weber, A., 2011a. Does indivisible labor explain the difference between micro and macro elasticities? A meta-analysis of extensive margin elasticities. NBER Working Paper Series 16729.

Chetty, R., Guren, A., Manoli, D. S., Weber, A., 2011b. Are micro and macro labor supply elasticities consistent? A review of the evidence on the intensive and extensive margin. *American Economic Review: Papers and Proceedings* 101 (3), 471-475.

Disney, R. and Smith, S., 2002. The labor supply effect of the abolition of the earnings rule for older workers in the United Kingdom. *The Economic Journal* 112 (478), C136-52.

Duggan, M, Singleton, P and Song, J., 2007. Aching to retire? The rise in the full retirement ag and its impact on the social security disability rolls. *Journal of Public Economics* 91, 1327-1350

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,* Cambridge University Press doi:10.1017/S1474747208003892

Feldstein, M., 2005. Structural reform of social security. *Journal of Economic Perspectives*

19 (2), 33-55.

Friedberg, L., 2000. The labor supply effects of the social security earnings test. *The Review of Economics and Statistics* 82, 48-63.

Gruber, J. and Wise, D. A., 1999. *Social security and retirement around the world.* NBER The University of Chicago Press.

Gruber, J. and Wise, D. A., 2004. Social security programs and retirement around the world. Micro-estimation. NBER The University of Chicago Press.

Haider, S. J. and Loughran, D. S., 2008. The effect of the social security earnings test on male labor supply. *The Journal of Human Resources* 43, 57-87.

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., Røed, K., 2015. Pension Reform and Labor Supply: Flexibility vs. Prescription. IZA Discussion Paper No. 8812.

McClelland, R. and Mok, S, 2012. A Review of Recent Research on Labor Supply Elasticities. Working Paper 2012-12, Congressional Budget Office, Washington, D.C.

OECD, 2012. Pensions outlook 2012. OECD Publishing. http://dx.doi.org/10.1787/9789264169401-en

Røed, K. and Haugen, F., 2003. Retirement and economics incentives: Evidence from a quasi-natural experiment. *Labour* 17 (2), 203-228.

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.

Staubli, S. and Zweimuller, J., 2011. Does raising the retirement age increase employment of older workers? Working Paper No 20, University of Zurich

Vestad, O., 2013. Labour supply effects of early retirement provision. Labour Economics 25, 98-109