Labour Economics 82 (2023) 102347

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Ageing and labor productivity

Erik Hernæs[a], Tom Kornstad[b], Simen Markussen[a], Knut Røed[a,*]

[a] The Ragnar Frisch Centre for Economic Research, Oslo, Norway
[b] Statistics Norway, Oslo, Norway

A R T I C L E I N F O

Keywords:
Workforce ageing
Pension reform
Productivity
Labor demand

A B S T R A C T

We exploit a policy-induced shift in the labor supply of elderly (age 63–67) workers in Norway to explore how aging of the workforce within existing firms is likely to affect labor productivity and the demand for younger workers. Our results are imprecise, but indicate that a higher share of age 63–67 workers increases total wage costs and has a small positive effect on labor productivity in the short run. Postponed retirement of existing elderly workers leads to a significant decline in the hiring of younger (below age 30) workers.

1. Introduction

In the present paper, we examine how a policy-induced increase in the number of elderly workers has affected labor productivity and age-specific labor demand within existing firms. Our analysis is based on a quasi-natural experiment in Norway: In 2011 a pension reform radically improved economic incentives for staying employed after the age of 62 for roughly half of the elderly private sector workers. We show that the reform yielded significant exogenous variation in the age structure of employees within and across firms, enabling identification of a crucial and policy-relevant causal relationship between the employees' age composition and labor productivity.

There is obviously no such thing as a universally valid true causal relationship between age and productivity. The influence of age is likely to vary across occupations, industries, and individuals, and also to change over time due to changes in the capital stock, technological innovations, improvements in health conditions, and changes in the relative supply of labor of different ages (Sharpe, 2011; Gordo and Skirbekk, 2013; Acemoglu and Restrepo, 2018). Hence, it is impossible to interpret an effect of age-composition on firm productivity without explicit reference to the source of variation used to identify it. In particular, the variation in age-structure across firms generated by differences in their optimal choices of labor inputs will have other implications for firm productivity than the variation caused by changes in the relative supply of age-specific labor.

Our analysis of the relationship between age and productivity is based on the margin of variation that arguably is the most policy relevant of all, namely the variation that results from public policies designed specifically to increase labor force participation among the elderly. Such policies are discussed and/or have already been implemented in virtually all advanced economies, most often in the form of pension reforms raising the retirement age and/or removing earnings tests. These reforms influence the age composition of the workforce by raising the amount of elderly workers' labor supply in a similar fashion as the aging of populations itself (Coile et al., 2018). Hence, they hold the key to understanding the fundamental relationship between the age composition of a country's labor force and its overall labor productivity.

The policy reform exploited in the present paper drastically improved work incentives for employees aged 63–67 who worked in firms affiliated to a supplementary early retirement program covering roughly 50% of private sector workers in Norway. The program is organized by employer and employee organizations, with part of the cost covered by a public subsidy. Up to 2011, the program provided a pension similar to the public pension over the age span 62–67, but subject to a strict earnings test. In 2011, the public pension became available from 62, conditional on a certain level of accrued pension entitlements, and the supplementary program was transformed into a life-long top-up pension at a correspondingly lower level, both without any earnings test. The reform had a large and immediate influence on the share of 62–63-year-olds that chose to continue in employment (Hernæs et al., 2016; Andersen et al., 2021). Cohorts born after 1948 were subjected to the new system, whereas cohorts born earlier maintained the old system; hence, given that the supplementary early retirement program covered a five year period, the reform's overall influence on the age-specific employment patterns were gradually phased in over five years. The resultant rise in participation rates increased the number of older persons in firms affiliated to the early retirement program, but not in other firms. The reform's influence on the number of mature workers also varied considerably across affiliated firms due to differences in the initial age composition of their workforces. Hence, by exploiting the differences between "treated" and "non-treated" firms as well as the differences between treated firms with different initial age structures within a

* Corresponding author.
E-mail address: knut.roed@frisch.uio.no (K. Røed).

https://doi.org/10.1016/j.labeco.2023.102347
Available online 1 March 2023

difference-in-differences framework, we can identify the causal impacts of exogenous (policy-induced) changes in the number of elderly employees on various firm outcomes, without having to rely on common trends for either firms with and without elderly employees or for firms with and without affiliation to the early retirement program.

We use administrative register data, which cover a description of all limited liability private sector firms' labor inputs, wage costs, profits, sales, and value added over the period from 2002 through 2018, with information on participation in the early retirement program from 2007. For each year 2007–2018 we use the predicted share of (fulltime-equivalent) "treated" workers as an instrument for the firm's actual share of age 63–67 labor input, with the prediction based on the workers' age composition five years before. In the regression analyses, we control for separate time (year) effects by early retirement program affiliation and by the fraction of workers belonging to early retirement age, and identify the causal effect through the interaction of program affiliation, age composition, and time. The identification strategy hinges on the orthogonality of this interaction term and the residual shocks to the outcomes, conditional on the separate time-trends and other controls, including firm-fixed effects. We examine the validity of the identification strategy through event studies (including "placebo" outcome years) and the inclusion of alternative control functions.

Examining labor productivity at the firm level entails some serious measurement problems. Our primary measure of labor productivity at the firm level is based on annual accounts, and defined as the value added (total wage costs plus profits) divided by total labor input (the number of person-years). This can only be a crude proxy for productivity, however, as the timing of realized profits may deviate substantially from the timing of actual value creation. Hence, as a supplement, we also use measures based on reported sales and wage costs. To complicate the analysis, we expect measurement error in the quantities of annual labor input, and also outlier problems related to skewed distributions of key variables, including firm size and profits. These data and measurement challenges are inherent in empirical analyses of firm-level productivity, and inevitably imply that subjective choices must be made regarding, e.g., sample inclusion criteria, outlier treatment, and functional forms. Ideally, such choices should be made strictly prior to the actual data analysis (in order to avoid data-mining). Yet, in practice, it is difficult to fully identify and understand the actual data problems, and thus to choose the appropriate remedies, without having examined the data to some extent. Our solution to this dilemma is to be as transparent as possible regarding the criteria used to adapt the data, and then to assess the consequences for the results of modifying each criterion.

These data challenges also imply that the relationship between age and productivity is estimated with considerable uncertainty. With this caveat in mind, we provide results indicating that overall labor productivity is most likely positively affected by a higher share of age 63–67 workers triggered by postponed retirement due to improved work incentives. The estimated impact on the average wage level is also positive, reflecting that individual wages typically grow with experience and tenure, either due to higher productivity or due to incentive contracts. The estimated effect on total employment is uncertain, but point estimates indicate a negative overall effect. A more detailed analysis of age-specific labor demand patterns shows that the demand for younger (below age 30) workers is significantly reduced. This effect is fully accounted for by reduced hiring of new workers. The employment effect appears to be heterogeneous though, and in industries where we expect the degree of complementarity between young and old workers to be considerable we do not find any displacement effect on younger workers.

Our paper relates to an existing literature examining how particular skills develop over the lifecycle, showing that some skills tend to deteriorate with high age (e.g., physical strength, adaptability, fluid cognitive ability) whereas others are stable or improve (e.g., experience-based knowledge, crystalized cognitive ability), with large variations across individuals (Skirbekk, 2008; Sharpe, 2011). The consequences of aging for labor productivity thus depend on the way the elderly are sorted into and out of jobs, and on how their skills match the skills of other workers. The degree to which workers of different ages are substitutes or complements in the production process plays a key role, as also reflected in our results.

There is a large empirical literature examining the relationship between labor productivity and the age-composition of workers at the firm level. Although some cross-sectional studies find indications of a declining (or hump-shaped) relationship between average age and labor productivity (e.g., Grund and Westergaard, 2008; Lallemand and Rycx, 2009), analyses based on panel data essentially indicate that productivity does not decline with the fraction of elderly employees; see, e.g., Cardoso et al. (2011), van Ours and Stoeldraijer (2011), Göbel and Zwick (2012), and Mahlberg et al. (2013). However, this literature focuses on how the actually chosen age-composition of labor inputs affect firm productivity, given the prevailing labor force participation patterns among the elderly and given the existing age-specific sorting into jobs. It cannot say much about the expected effects of changes in the age structure generated by exogenous changes in the overall supply of mature labor.

A recent paper that does say something about the impacts of exogenous changes in labor supply, and hence is more closely related to our own contribution, is Carta et al. (2021). This paper uses a pension reform in Italy to identify the short-term effects of a policy-induced increase in the supply of workers above 55 years. Not far from our results, they find that the rise in the participation of mature workers did not significantly affect average labor productivity. However, in contrast to us, they estimate a positive impact on the demand for younger workers. Hence, older and younger workers appears to be complements in the firms they study, and they conclude that rising institutional retirement ages can help firms retaining valuable older employees. Interestingly, another paper using exactly the same pension reform to identify causality (Boeri et al., 2022) reaches a conclusion more similar to ours, namely that postponed retirement caused a reduction in the demand for other age groups.

2. Institutional setting: the norwegian pension reform

In 2011, the whole Norwegian pension system was radically reformed (Christensen et al., 2012; Hernæs et al., 2016; Kudrna, 2017; Halvorsen and West-Pedersen, 2019). The main ingredients of the reform was a tightening of the relationship between individual lifetime earnings and pension entitlements, longevity-adjusted annual pensions, and less generous indexation. These changes are phased in gradually, however, and had negligible impacts on the work incentives for cohorts retiring around the time of the reform. In the present paper, we focus on a reform element that had large and immediate consequences for many private sector workers; namely the removal of the retirement earnings test for workers qualifying for a supplementary early retirement pension entitled "AvtaleFestet Pensjon" (hereafter AFP). AFP is a separate pillar of the Norwegian pension system which was established in 1988 through an agreement between the associations of employers and employees in Norway. It applies for workers who are covered by the collective agreements between these associations (all public sector workers and approximately 50% of private sector workers) and comes on top of public and occupational pensions.

The reform of the AFP program was implemented in a quasi-experimental fashion, in the sense that adjacent birth cohorts eligible to AFP suddenly faced completely different early retirement incentives, whereas the pre-existing large difference in incentives between AFP and non-AFP eligible workers was immediately eliminated. Before the reform, workers entitled to AFP could claim an early retirement benefit already from the age of 62, calculated as the old age pension in the public pension system with accrual *as if* the existing job and salary continued up to age 67. There was a strict earnings test, which together with the income tax implied a total tax rate of 75% at average earnings.

The reform repealed the earnings test and introduced an actuarially adjusted pension from age 62, reducing this total tax rate to approximately 40%. From one birth cohort to the next, the average annual take-home pay associated with postponing retirement after the age of 62 and up to age 67 increased by NOK 200,000 (approximately € 20,000) or 150% (Andersen et al., 2021). As a result, labor supply (measured by average gross labor earnings) among 63 and 64 year olds increased by approximately 30% (Hernæs et al., 2016). Before 2011, workers without entitlement to AFP could not claim any pension at all until age 67. For them, the reform entailed no important changes in work incentives, but an opportunity to start drawing on their pension wealth five years before at actuarially neutral terms. Previous empirical evidence has indicated that this new opportunity caused a small decline in labor supply at the intensive margin (Hernæs et al., 2016). A later study focusing exclusively on those without entitlement to AFP over the age range 63–66, found that the decline at the intensive margin was largely offset by increased labor supply at the extensive margin (Hernæs et al., 2021).

Employment protection legislation, as well strong norms regulating employer-employee relationships in Norway, imply that any downwards adjustments of firms' employment normally occur through a combination of voluntary quits and reduced hiring. Hence, if the stronger work incentives following from the AFP reform reduced the quit rate among elderly workers, we expect to see either a corresponding increase in total employment or an offsetting reduction in new hires, the latter typically implying a reduction in the number of young workers.

3. Data and identification strategy

The aim of our empirical analysis is to identify the short-term effects of an externally imposed change in the number of workers in the age 63–67 range on two types of firm outcomes[1]:

i Labor productivity
ii Total employment and employment in age brackets other than age 63–67

The analysis is based on administrative registers containing annual accounts for all limited liability private sector firms in Norway from 2002 through 2018. These data are merged with employer-employee registers with information about all employees, their age, annual earnings, and contracted work hours. Finally, we add in information about each firm's participation in the early retirement program AFP, which is available from 2007.

To identify the causal effects of interest, we exploit the 2011 early retirement reform, which, from a firm perspective, represented an exogenous source of variation in the number of workers above age 62. This forms the basis for an instrumental variables approach, whereby we use the fraction/number of workers expected to be directly affected by the reform as an instrument for the fraction/number of old workers. The idea is that we can then identify and estimate local average treatment effects corresponding to the policy-relevant margin of variation in the age composition, namely the variation generated by a pension reform aimed at encouraging elderly workers to postpone retirement.

Our data will be structured in terms of an initial "base-year" and an "outcome-year", and the statistical analysis is conditioned on firm characteristics in the base-year. Information about a firm's age structure in the base-year will be used to form predictions about the number (or fraction) of age 63–67 workers and of workers eligible for AFP-retirement in the outcome year. To ensure that the firm's employment structure in the base-year is always exogenous with respect to the impacts of the reform, the base-year is the year five years prior to the outcome-year (since none of the current age 63–67 workers had reached early retirement age five years ago). We drop firms with less than 5 full-time-full-year-equivalent (FTE) employees in the base-year to avoid too much noise related to firm exit and unreliable productivity measures. In order to identify AFP-affiliation (and to ensure a meaningful productivity analysis), we also condition on the firm still being active in the outcome year (defined as having at least 1 person-year of labor input).[2] This raises some potential selection problems, which in the empirical analysis will motivate a focus on within-firm changes as the source of identification. Based on these criteria, we have 270,582 firm-year observations in our data, out of which 71,461 (26%) are recorded as affiliated to AFP. In the main part of the analysis, we drop 5265 firm-year observations (1.9%) with recorded labor input inconsistent with recorded total wage costs. More specifically, we require that imputed full-time-full-year gross earnings (including the employers' payroll tax) are between 3 and 30 "Basic Amounts" of the pension system (denoted "G" in Norway, and 1 G is approximately equal to NOK 100,000 (€ 10,000) in 2020).[3] We also drop observations belonging to the top percentile in the distribution of employment growth (2673 firm-year observations with growth larger than 260%), as we suspect that these observations are dominated by unobserved mergers, acquisitions, or other organizational changes unrelated to aging of the workforce.

Fig. 1 shows how the fractions of elderly (age 63–67) workers have developed in firms with and without AFP affiliation during the years before and after the reform, measured in terms of full-time equivalents. For comparison, it also shows "predicted fractions", where the predictions are based on the fraction of age 58–62 workers the firms had five years earlier, assuming a constant ratio (at the pre-reform level) between the previous age 58–62 and the current age 63–67 fractions. We note that without changes in labor supply behavior, we would predict a decline in the fraction of age 63–67 workers in both AFP and non-AFP firms. The main reason for this is a demographic transition caused by the reduced influence of the large cohorts born just after the Second world war (the number of births dropped by 15% between 1946 and 1951). Compared to these predictions, we see that the fraction of elderly workers has increased by approximately 0.7–0.8 percentage point in AFP affiliated firms and dropped by approximately 0.2–0.4 percentage points in non-affiliated firm. While the former has already been shown to be a result of the improved work incentives for workers with access to AFP, the latter is likely to be related to the introduction in 2011 of the opportunity to start drawing on the public pension (at actuarially neutral conditions) already at age 62 instead of at age 67. As it is the differential employment trends for AFP and non-AFP affiliated elderly workers that identify the effects of interest in our analysis, the negative employment trend for elderly workers in non-affiliated firms actually provides a separate contribution to identification.

Table 1, columns I and II, shows some descriptive characteristics for firms with and without AFP-affiliation. It is evident that firms with and without AFP are quite different, particularly in terms of size. This reflects that AFP is a result of an agreement between the major employer and employee associations in Norway, involving the more "organized" parts of the labor market. On average, AFP firms are three times larger than non-AFP firms; hence even though they only make up 27% of the included firm-year observations, they encompass 52% of the person-years.

Although it is possible to control properly for the direct influences of differences in firm size and other characteristics, it appears probable that the associations between age composition and firm outcomes may vary considerably across heterogeneous firms. Hence, we would have

[1] As outcomes and explanatory variables are all defined in terms of calendar years, whereas age is not, we need a calendar-year based definition of age. We define a person as x years old in a calendar year t if the persons reaches the age of x during the course of that year.

[2] Since firms' AFP-affiliation is not observed before 2007, we cannot identify AFP-affiliation for firms that do not survive until this year.

[3] We suspect that average earnings levels outside these ranges have resulted from over –or underreported labor inputs, most likely due to errors in the recorded start or stop dates for employment spells or in the registered number of contracted hours.

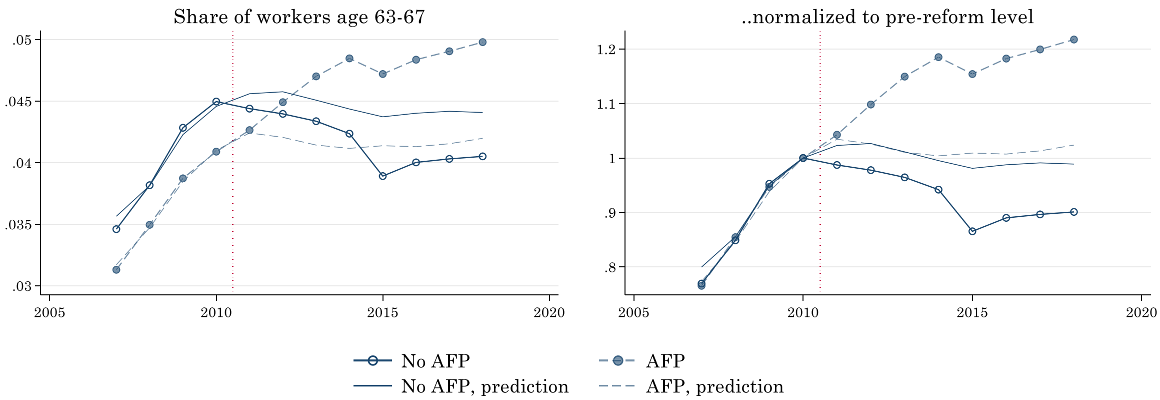

Fig. 1. Share of workers age 63–67 in firms with and without AFP affiliation.

Note: Data include all economically active (at least 1 full-time-full-year-equivalent employee) firm-year observations for firms with at least 5 full-time-full-year-equivalent employees five years before. Observations with unreasonable wage costs per person-year have been dropped (1.9%). Unreasonable wage cost are defined as having average person-year costs below approximately NOK 300,000 or above 3000,000, measured in 2020-value). The number of firm-year observations is 265,317.

Table 1

Descriptive statistics firms.

	Before matching		After matching	
	I Without AFP	II With AFP	III Without AFP	IV With AFP
# Person-years in base-year	16.71	54.37	35.16	36.85
# Person-years in outcome year	18.20	55.29	37.44	38.64
Base-year wage costs per person-year (1000 NOK)	629.45	618.71	627.82	608.72
Base-year profit per person-year (1000 NOK)	210.19	243.00	218.02	229.91
Firm age in outcome year (years)	18.40	22.35	20.69	21.99
Fraction of person-years in age group (outcome year)				
Age<30	0.22	0.20	0.22	0.22
Age 30–62	0.73	0.75	0.72	0.73
Age 63–67	0.04	0.04	0.05	0.04
Age>67	0.01	0.01	0.01	0.01
Selected industries (fraction of firms)				
Construction	0.17	0.19	0.22	0.22
Wholesale	0.13	0.06	0.07	0.07
Retail	0.11	0.09	0.11	0.11
Financial services	0.12	0.05	0.06	0.06
Health care services	0.05	0.03	0.04	0.04
Care repair, petrol stations etc.	0.07	0.07	0.08	0.08
# firm-year observations	192,440	70,204	59,366* (165,681)	59,366

* After matching, there are a total of 165,681 firms non-AFP firms included, with weights such that the number ads up to 59,366.

liked the characteristics of AFP and non-AFP firms to be more similar. To achieve such similarity, we perform a matching exercise, with exact matching on three key base-year characteristics: i) firm size category (with the following 21 categories based on the number of person-years: 5–6,…,18–19, 20–30,…,40–50, 50–100, 100–500, 500–1000, >1000), ii) predicted category of age 63–67 fraction (with the following 11 categories: 0%, 0–10%,…,90–100%), and iii) industry (two digit ISIC; approximately 85 categories in our data). For each AFP-firm-year observation we identify all matches among non-AFP-firms that satisfy these three criteria, and if there is more than one match, we include them all with equal weights adding up to one. If there is no match, we delete the observation. We also delete unmatched non-AFP firms. The result is given in Table 1, columns III and IV. There are 10,838 observations (15.4%) of AFP-firms that do not have a satisfactory match among non-AFP firms. Hence we end up with 59,366 observations of AFP as well as non-AFP firms. In a sensitivity analysis below, we will show that firm size is the most critical matching variable.

4. Event study

To assess the validity of our identification strategy, we perform an event study, encompassing all outcomes of interest, including the endogenous fraction of elderly workers. Let y_{jt} be an outcome measured for a firm j in year t, let S_t be a calendar year dummy variable equal to 1 in year t, and let A_j be a dummy variable equal to 1 for firms affiliated to the AFP early retirement program. Let L^{58-62}_{jt-5} be the number of age 58–62 workers in the firm five years ago, which here represents the number of potential age 63–67 workers in year t, as very few new workers are hired at this age. Finally, let AFP^{58-62}_{jt-5} be the number of age 58–62 workers in the firm five years ago who are eligible for AFP-retirement. In firms affiliated to AFP, we will typically have that $AFP^{58-62}_{jt-5} = L^{58-62}_{jt-5}$, whereas in non-affiliated firms $AFP^{58-62}_{jt-5} = 0$. There will be exceptions from this pattern, however, as some workers have AFP-eligibility determined from main job in another firm. The event study has the following

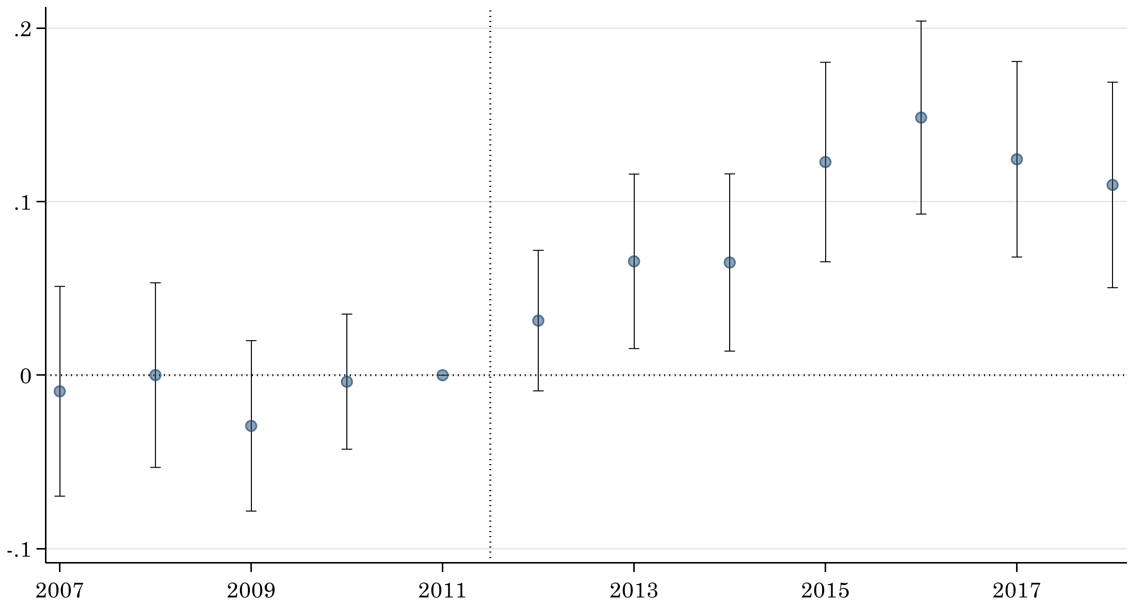

Fig. 2. Event study: The time variation in the effect of the fraction of potential AFP-retirees on the actual fraction of age 63–67 workers (δ_t^{AFP}in Eq. (1), with 2011 as reference)

Note: The reported estimates are based on a single regression where the dependent variable is L_{jt}^{63-67}/L_{jt}, i.e. the number of (fulltime-equivalent) workers aged 63–67 divided by the total number of (fulltime-equivalent) workers in the same (outcome) year. The figure shows point estimates with 95% confidence intervals. Standard errors are clustered at the firm level.

structure:

$$y_{jt} = \underbrace{\gamma_j}_{\substack{\text{Firm fixed}\\ \text{effect}}} + \underbrace{\sigma_t^0 S_t + \sigma_t^{AFP} S_t \times A_j}_{\substack{\text{Separate time effects for}\\ \text{AFP}}} + \underbrace{\beta_t^0 S_t \times \frac{L^{58-62}_{jt-5}}{L_{jt-5}}}_{\substack{\text{Time}\\ -}} + \underbrace{\beta^{AFP} A_j \times \frac{L^{58-62}_{jt-5}}{L_{jt-5}}}_{\substack{\text{Separate}\\ (\text{time}-\text{invariant})}}$$
$$+ \underbrace{\kappa^{AFP} A_j \times \frac{AFP^{58-62}_{jt-5}}{L_{jt-5}}}_{\substack{\text{Separate}\\ (\text{time}-\text{invariant})}} + \underbrace{\delta_t^{AFP} S_t \times \frac{AFP^{58-62}_{jt-5}}{L_{jt-5}}}_{\substack{\text{Time}\\ -}} + \underbrace{\varepsilon_{jt}}_{\text{Residual}}, \qquad (1)$$

where L_{jt-5} is the total number of (full-time-equivalent) workers in firm j five years ago. In order to avoid outlier-problems and excess influence of a small number of very large firms, we scale all the employment variables by the initial (base-year) total employment level in the firm. Motivated by the large heterogeneity in firm characteristics, we include firm-fixed effects (γ_j) in the model. The use of firm fixed effects also mitigates the potential selection problem created by the condition of firm survival until the outcome year.

The coefficient of primary interest in the event study is δ_t^{AFP}, which captures how the influence of the fraction of potential AFP-retirees changes over time. Note that with controls for $L_{jt-5}^{58-62}/L_{jt-5}$ interacted both with time dummy variables and with a (time-invariant) indicator for AFP-firm, and with separate time dummy variables for AFP and non-AFP firms, identification of δ_t^{AFP} relies on a difference-in-differences strategy. It captures the extra effect of the fraction of potential AFP retirees in each year, over and above the effects of the fraction of potential age 63–67 workers (irrespective of AFP entitlement) and of the firm's AFP status.

Fig. 2 first shows the estimated effects of the fraction of potential AFP retirees ($\hat{\delta}_t^{AFP}$ in Eq. (1), $t = 2007, ..., 2018$) on the actual fraction of age-63–67-workers (L_{jt}^{63-67}/L_{jt}), with 2011 as the reference year. This is the last year for which we do not expect any reform influence at age 63 (or higher), as none of the potential AFP retirees had yet been treated by the reform. From 2012 to 2015, the fraction of treated rises year-by-year, and from 2016, all potential AFP retirees had been treated. The estimates reported in Fig. 2 confirm that the impact of having workers in the group of potential AFP-retirees build up gradually after the reform.

These relationships will later form the basis for the first-stage-equation in an instrumental variables analysis.

Fig. 3 shows how the fraction of potential AFP retirees is estimated to have affected the number of (full-time-equivalent) employees in different age groups. These effects are derived from regressions where we use as dependent variable the number of workers in a particular age group in the outcome year divided by the total number of workers in the base-year; hence the reported $\widehat{\delta_t}^{AFP}$ coefficients can be interpreted directly in terms of actual numbers. Panel (c) shows the estimated effects for the directly affected workers.[4] Again, we see that the effects build up gradually from 2012, in accordance with the rising share of reform-treated workers. The resultant increase in the employment of elderly workers appears to have been offset by a reduction in the number of young (below age 30) workers ((panel (a)). Overall, employment appears to have been largely unaffected.

As noted in the introduction, examination of labor productivity entails serious measurement problems. Our preferred indicator of productivity is value added (measured as total wage costs including payroll taxes plus profits) per person-year. However, the timing of realized profits may not correspond to the timing of actual value creation, and many highly valuable firms run deficits for several years before the economic returns materialize in the form of profits. Hence, there is a lot of noise in our value added measure. We therefore also use other proxies for labor productivity, including total wage costs per unit of labor and total sales per unit of labor. The wage level is an appropriate productivity-indicator insofar as labor is paid its marginal product, but problematic in our context if implicit contracts entail higher-than-productivity wages for older workers (Lazear, 1979).

Fig. 4 reports estimates indicating how the fraction of potential AFP retirees for each year have influenced overall wage costs, value added, and total sales per unit of labor, all defined as natural logs of the respective ratios. To circumvent problems related to

[4] Note that whereas Figure 2 shows effects on the *fraction* of workers in the age 63-67 range (i.e., after endogenous responses in the overall number of workers have been accounted for), panel (c) in Figure 3 shows effects on the *number* (normalized by the employment level 5 years ago).

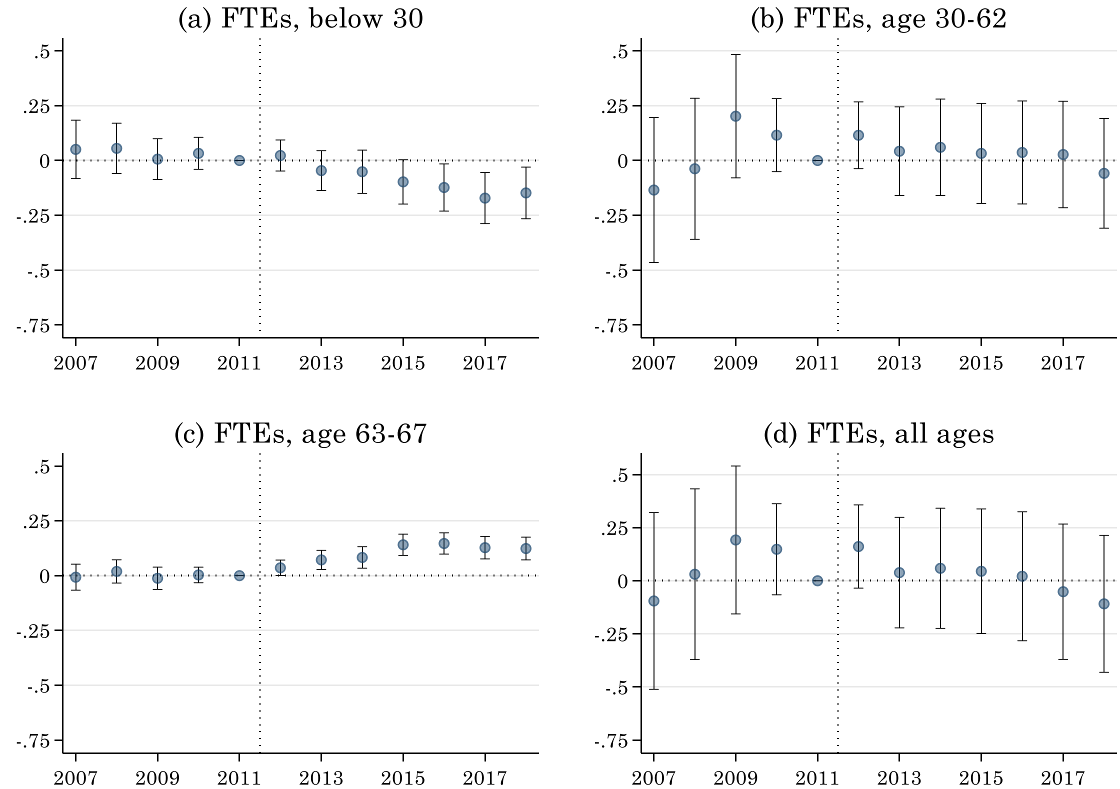

Fig. 3. Event study: The time variation in the effect of the fraction of potential AFP-retirees on the number of workers in outcome year relative to base-year (δ_t^{AFP} in Eq. (1), with 2011 as reference). By age group.
Note: The dependent variable is the number of fulltime-equivalent workers (FTE) in the respective age group divided by the total number of (fulltime-equivalent) workers in the base-year (five years before the outcome year). For example, in panel (c), the dependent variable is L_{jt}^{63-67}/L_{jt-5}. The figure shows point estimates with 95% confidence intervals. Standard errors are clustered at the firm level.

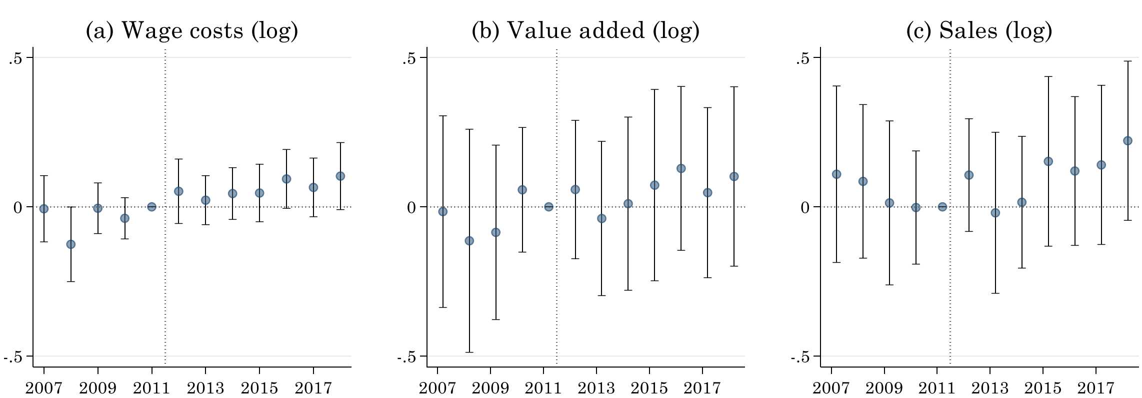

Fig. 4. Event study: The time variation in the effect of the fraction of potential AFP-retirees on the relative growth in wage costs (panel a) and value added (panel b) per person-year (δ_t^{AFP}in Eq. (1), with 2010 as reference).
Note: In the regression reported in panel (a), the dependent variable is the (the log of) firms' total wage costs divided by the total number of (full-time-equivalent) workers, both measured in the outcome year. In panel (b) the dependent variable is (the log of) value added (the sum of total wage costs and total profits) divided by the total number of (full-time-equivalent) workers. In the value-added-analysis, we have shaved the sample by dropping observations with negative value added or with value added exceeding 3 times the wage cost (2.9% of the sample). In panel (c) the dependent variable is (the log of) total sales divided by the total number of (full-time-equivalent) workers. In the sales-analysis, we have shaved the sample by dropping observations with negligible sales (below 1 G) or sales exceeding 20 times the wage cost (4.5% of the sample). The figure shows point estimates with 95% confidence intervals. Standard errors are clustered at the firm level.

computing the log of non-positive numbers and to avoid excessive outlier influence, we have shaved the samples used in the analyses of value added and sales somewhat; see the note to Fig. 4 for details. Although none of the estimates reported in Fig. 4 are statistically significant in isolation (at the conventional 5%-level), it appears to be a pattern that the estimated effects on both wage costs, value added, and sales have become gradually more positive after the reform.

5. Instrumental variables analysis

We now turn to the regression analysis where our aim is to exploit the data more efficiently in order to answer two research questions. The first is how an exogenously imposed change in the number of aged workers affect the demand for workers in other age groups and the firm's total employment. The endogenous regressor in this case is L_{jt}^{63-67}/L_{jt-5}; i.e. the *number* of elderly workers actually employed in the outcome year (normalized by total employment five years ago). The second research question is how the resultant *fraction* of elderly workers in the outcome year affects various indicators of firm productivity. Here, L_{jt}^{63-67}/L_{jt} is the endogenous regressor (i.e., the number of elderly workers as fraction of *current* total employment). The instrument is in both cases the predicted number of reform-treated AFP-eligible workers divided by the number of employees five years ago; i.e. $AFP_{jt}^{TREAT}/L_{jt-5}$, where AFP_{jt}^{TREAT} is defined as follows:

$$\begin{aligned}
AFP_{jt}^{TREAT} &= 0, \text{if } t = 2007, 2008, 2009, 2010, 2011,\\
&= AFP_{jt-5}^{58}, \text{if } t = 2012,\\
&= AFP_{jt-5}^{58} + AFP_{jt-5}^{59}, \text{if } t = 2013,\\
&= AFP_{jt-5}^{58} + AFP_{jt-5}^{59} + AFP_{jt-5}^{60}, \text{if } t = 2014,\\
&= AFP_{jt-5}^{58} + AFP_{jt-5}^{59} + AFP_{jt-5}^{60} + AFP_{jt-5}^{61}, \text{if } t = 2015,\\
&= AFP_{jt-5}^{58} + AFP_{jt-5}^{59} + AFP_{jt-5}^{60} + AFP_{jt-5}^{61} + AFP_{jt-5}^{62},\\
&\text{if } t = 2016, 2017, 2018.
\end{aligned} \tag{2}$$

The regression equations for the employment outcomes have the following structure:

$$\frac{L_{jt}}{L_{jt-5}} = \gamma_j + \delta \underbrace{\frac{L_{jt}^{63-67}}{L_{jt-5}}}_{\text{Endogenous}} + \sigma_t^0 S_t + \sigma_t^{AFP} S_t \times A_j + \beta_t^0 S_t \times \frac{L_{jt-5}^{58-62}}{L_{jt-5}} + \beta^{AFP} A_j \times \frac{L_{jt-5}^{58-62}}{L_{jt-5}} + \kappa^{AFP} A_j \times \frac{AFP_{jt-5}^{58-62}}{L_{jt-5}} + \rho^0 \frac{L_{jt}^{CONTR}}{L_{jt-5}} + \zeta_{jt} \tag{3}$$

where L_{jt}^{CONTR} is defined analogous to AFP_{jt}^{TREAT} with L_{jt-5}^{AGE} substituted for all AFP_{jt-5}^{AGE} terms in Eq. (2), and included to make sure that we control for the age-structure variables with exactly the same functional form as we use for the AFP eligibility variables. In Eq. (3), the outcome is specified as total employment in the outcome year relative to the base-year, but we will also estimate models where the outcome is employment in particular age groups, such that the numerator in the left-hand-side variable is either $L_{jt}^{<30}$ or L_{jt}^{30-62}. In addition, we estimate separate models for the number of person-years associated with newly hired workers (workers that were not employed in the firm five years ago) and the corresponding number associated with employees that were in the firm also in the base-year.

In the productivity and wage cost analysis, the regression equations have the following structure:

$$\log \frac{VA_{jt}}{L_{jt}} = \gamma_j + \delta \underbrace{\frac{L_{jt}^{63-67}}{L_{jt}}}_{\text{Endogenous}} + \sigma_t^0 S_t + \sigma_t^{AFP} S_t \times A_j + \beta_t^0 S_t \times \frac{L_{jt-5}^{58-62}}{L_{jt-5}} + \beta^{AFP} A_j \times \frac{L_{jt-5}^{58-62}}{L_{jt-5}} + \kappa^{AFP} A_j \times \frac{AFP_{jt-5}^{58-62}}{L_{jt-5}} + \rho^0 \frac{L_{jt}^{CONTR}}{L_{jt-5}} + \zeta_{jt} \tag{4}$$

where VA_{jt} denotes value added in the outcome year. Here, we have specified value added as the outcome, but we also estimate Eq. (4) with (log) wage costs and (log) total sales as outcomes. Note that the only difference between the right-hand-sides of Eqs. (3) and (4) is the denominator in the endogenous explanatory variable of interest. The exclusion restriction is that, given the control variables included in Eqs. (3) and (4), the fraction of potential reform-treated AFP-retirees affects the outcome of interest only through its influence on the actual number or fraction of age 63–67 workers. This assumption could be violated if higher employment among the elderly also changed the demand for goods and services differently for different industries. Given the relatively small share of elderly employees (4% of all person-years), we expect such general equilibrium effects to be small, at least in the short run. However, we do provide a validity assessment below by adding into the models alternative (extensive) control variable sets. In particular, to account for possible asymmetric influences on product demand, we include industry-by-year dummy variables.

Table 2
First stage estimates for the endogenous variables in Eqs. (3) and (4) (standard errors in parentheses).

	I Number of elderly (age 63–67) employees L_{jt}^{63-67}/L_{jt-5}	II Fraction of elderly(age 63–67) employees L_{jt}^{63-67}/L_{jt}
$AFP_{jt}^{TREAT}/L_{jt-5}$	0.140***	0.130***
	(0.018)	(0.016)
R-squared	0.738	0.740
F(1,34,271)	58.82	70.16
N	217,960	217,960

Note: Standard errors are clustered at the firm level. */**/*** indicates statistical significance at the 10/5/1 percent level.

Table 3
Second stage estimates (standard errors in parentheses). Employment (Eq. (3)).

	I Number of young (below age 30) employees $L_{jt}^{<30}/L_{jt-5}$	II Number of middle-aged (age 30–62) employees L_{jt}^{30-62}/L_{jt-5}	III Total number of employees L_{jt}/L_{jt-5}
$\widehat{L_{jt}^{63-67}}/L_{jt-5}$	−0.857***	−0.623	−0.760
	(0.321)	(0.575)	(0.779)
Made up by:			
…entry of new workers	−0.789***	−0.412	−1.181*
	(0.304)	(0.485)	(0.677)
…continuation of existing workers	−0.068	−0.212	0.422
	(0.093)	(0.289)	(0.321)
R-squared (total effect)	0.698	0.605	0.526
N	217,960	217,960	217,960

Note: Standard errors are clustered at the firm level. */**/*** indicates statistical significance at the 10/5/1 percent level.

Table 2 first presents results for the first stage analysis. They show that the instrument has a significant and powerful impact on the endogenous regressors. The point estimate of 0.14 implies that the reform raised employment among the treated workers by approximately 14 percentage points. This estimate is somewhat lower than the effects reported by both Hernæs et al. (2016) and Andersen et al. (2021), most likely reflecting that our prediction of "potential AFP retirees" is based on employment as much as five years before and hence include a larger fraction of persons who in practice never become eligible for AFP. This implies that AFP_{jt}^{TREAT} is measured with error, such that our first stage results may be subjected to attenuation bias.

Table 3 shows second stage results for the employment regressions. The estimates indicate that for each additional person-year in the age 63–67 group, the number of person-years in the below-age-30 group is reduced by 0.86; hence our results imply almost full displacement of younger workers. Unsurprisingly, this effect is fully accounted for by reduced hiring of new workers. The impacts on the number of middle-aged workers as well as on total employment are estimated with too much statistical uncertainty for any conclusions to be drawn. What we can say is that we do not find any evidence in support of significant effects on total employment in either direction, but that the hiring of new workers definitely decline. For middle-aged workers it may also be noted that the IV point estimates are negative, in contrast to the small positive point estimates reported in the event study (Fig. 3, panel b). This apparent discrepancy is most likely related to the (somewhat arbitrary) choice of pre-reform reference year in the event study analysis.

Table 4 presents our main results regarding wage costs and labor productivity. Again, the statistical uncertainty is considerable. Yet, there is strong evidence that wage costs rise with the fraction of elderly workers

Table 4
Second stage estimates (standard errors in parentheses). Wage costs, value added, and total sales (Eq. (4)).

	I Wage costs per unit of labor input $\log(WC_{jt}/L_{jt})$	II Value added per unit of labor input $\log(VA_{jt}/L_{jt})$	III Total sales per unit of labor input $\log(TS_{jt}/L_{jt})$
$\widehat{L_{jt}^{63-67}/L_{jt}}$	0.602***	0.783	0.540
	(0.233)	(0.563)	(0.581)
R-squared	0.873	0.669	0.905
N	217,960	211,573	208,097

Note: Standard errors are clustered at the firm level. */**/*** indicates statistical significance at the 10/5/1 percent level. See note to Fig. 4 for a description of sample construction.

and somewhat weaker evidence that labor productivity (measured both by value added and by total sales per unit of labor input) also increases. Based on the results reported in Table 4, we can at least rule out large negative productivity effects resulting from a larger share of age 63–67 workers. Note that the coefficients reported in Table 4 are interpreted as the effect of changing the fraction of elderly workers from 0 to 1, which is of course way outside the central range of variation. As seen in Fig. 1, the reform used to identify causality in our analysis raised the fraction of age 63–67 workers by approximately 1 percentage point. According to the point estimate reported in Table 4, column II, such an increase is expected to raise overall labor productivity by approximately 0.8 percent.

6. Robustness

As noted in previous sections, the analysis in this paper requires adaptations of the raw data in order to deal with challenges related to measurement error, imperfect productivity indicators, large differences between treated and non-treated firms, highly skewed distributions of dependent as well as explanatory variables, and extreme outlier observations. Since these adaptations can be done in many ways, and involve several (subjective) choices of variable value "thresholds", they also involve risks of results-seeking data-mining. Thus, the standard tools of statistical inference may be undermined.

To assess the empirical relevance of such concerns, we now present results for a number of alternative data adaptations. The idea is not to show that "everything is robust" and that we would have reached the same conclusions regardless of data adaptation choices (which we would not), but rather to illuminate which choices that have been important for our findings and which have not. Our hope is that this exercise can form the basis for an informed assessment of the evidence.

The adaptation of the data in this paper involves two major steps. The first is the sample inclusion criteria, which in our case involves decisions regarding a lower and upper threshold on firm's annual wage costs relative to the reported number of hours worked (measured in full-time-full-year-equivalents). The second step is the algorithm that matches non-AFP to AFP firms. In addition, important discretionary choices are made with respect to the inclusion of control variables. In the models reported in Sections 4 and 5, we have added no control variables other than those explicitly included in Eqs. (1), (3), and (4); i.e., apart from the firm-fixed effects, we have only included covariates required for valid difference-in-differences identification. In the present section, we will add to the models extensive sets of covariates in two steps. In the first step, we include covariates describing the firms' base-year situation (firm age, firm size, age composition of employees, value added and sales per employee), and in the second step, we add industry-by-year and firm-size-by-year fixed effects and industry-specific current growth rates (from the last year to the outcome year) in wage costs, value added and employment, each term interacted with the firm's base-year fractions of elderly (age 58–62) workers with and without AFP-entitlement; see note to Fig. 5 for details.[5]

We focus exclusively on the instrumental variables estimates in this section, and for expository reasons, we present the estimates graphically, with confidence "fans" (rather than the more standard confidence intervals) to provide a more comprehensive picture of statistical uncertainty.

Fig. 5 first shows results for the sample used in the previous section, but with alternative control variable sets. For comparison, the leftmost bar in each panel (labelled Firm FE) repeats the baseline estimates already reported in Tables 3 and 4. The next two bars show results obtained when we include base-year firm controls (second bar) and then also add in business-cycle-by-industry controls (third bar). A conclusion from this exercise is that the incorporation of these extra control variables changes almost nothing. Moving on to the three next bars, the same exercise is repeated, only this time without firm-fixed effects. It is clear that not all our findings are robust with respect to this modification of the model. In particular, we now find indications of negative effects on total (as well as middle aged) employment, and the estimated positive effect on value added disappears.

We then examine the impacts of modifying the sample inclusion criteria used to eliminate observations with inconsistent records on wage costs and labor input. Fig. 6 shows results based on alternative thresholds at the bottom (to the left of the vertical stapled line) and the top, with the thresholds marked at the horizontal axis. Although point estimates as well as statistical uncertainty varies somewhat across the different samples, the bottom line here is that our estimation results are robust with respect to the data inclusion criteria.[6]

Finally, we look at the consequences of modifying the matching algorithm. In the baseline analysis, we used three exact (categorized) matching criteria; i.e., firm size, fraction of elderly workers in the base-year, and industry (at two-digit ISIC level). Fig. 7 illustrates the consequences of removing each criterion separately and of dropping the matching exercise completely. A first point to note here is that our results would not have been the same had we not used a matched sample of firms. This is particularly evident for the estimated effects on employment levels. It also appears that the most critical matching variable is firm size.

Viewed as a whole, we conclude that the main results presented in this paper are robust with respect to the inclusion of a wide range of covariates, but that two of the choices made during data adaptation and modeling are important for the estimated employment effects, namely the matching on firm size and the inclusion of firm-fixed effects. We will argue, however, that both these choices are well founded. Firm size is most likely correlated with range of (unobserved) firm characteristics, with implications for subsequent paths of outcomes; hence, large differences in the size distributions of treated and non-treated firms may challenge the identifying assumptions. Firm fixed effects are important to deal with potential selection problems arising from our implicit assumption of firm survival from the base-year to the outcome year (five years later). The estimated effects on wage costs, value added, and sales are less sensitive to data adaptation and modeling choices, although the large statistical uncertainty haunting all our specifications makes it dif-

[5] The inclusion of industry-specific growth rates interacted with the fraction of elderly workers with and without AFP-entitlement is motivated by the possibility that business cycle fluctuations affected AFP and non-AFP workers differently prior to the reform, as AFP sometimes was used by firms as a downsizing tool.

[6] In the baseline model for the value-added effect, we also shaved the sample by dropping observations with negative value added or with value added exceeding 3 times the wage cost (2.9 % of the sample); see note to Figure 4. If we include all observations with value added exceeding 3 times the wage costs, the results change very little (point estimate 0.92, with standard error 0.62). However, if we also include negative values by attributing the log-value of zero, the estimated effect becomes quite different and much more uncertain (point estimate −0.59, with standard error 1.36).

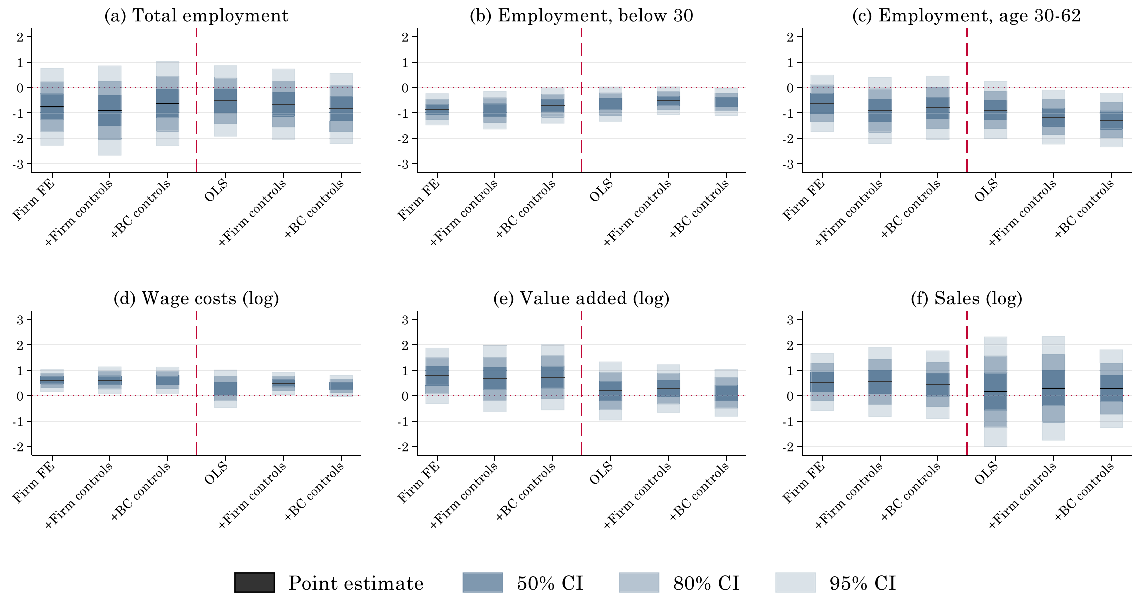

Fig. 5. Instrumental variables estimates with alternative control variable sets (with confidence fans).
Note: The model denoted “Firm FE” is the baseline model used in the previous section. The model denoted “OLS” is the same, only without firm-fixed effects. “Firm controls” include 26 firm age dummy variables (5,6,..,29,>29), age composition (fractions 20–29, 30–62, and 63–67), number of person-years interacted with year dummy variables, log wage costs per employee interacted with year dummy variables, and log value added per employee interacted with year dummy variables. “BC controls” include dummy variables for industry (2-digit ISIC) interacted with year dummy variables, industry-specific growth rates last year for wage costs, value added and employment, all interacted with the fractions of age 58–62-workers with and without AFP-entitlement.

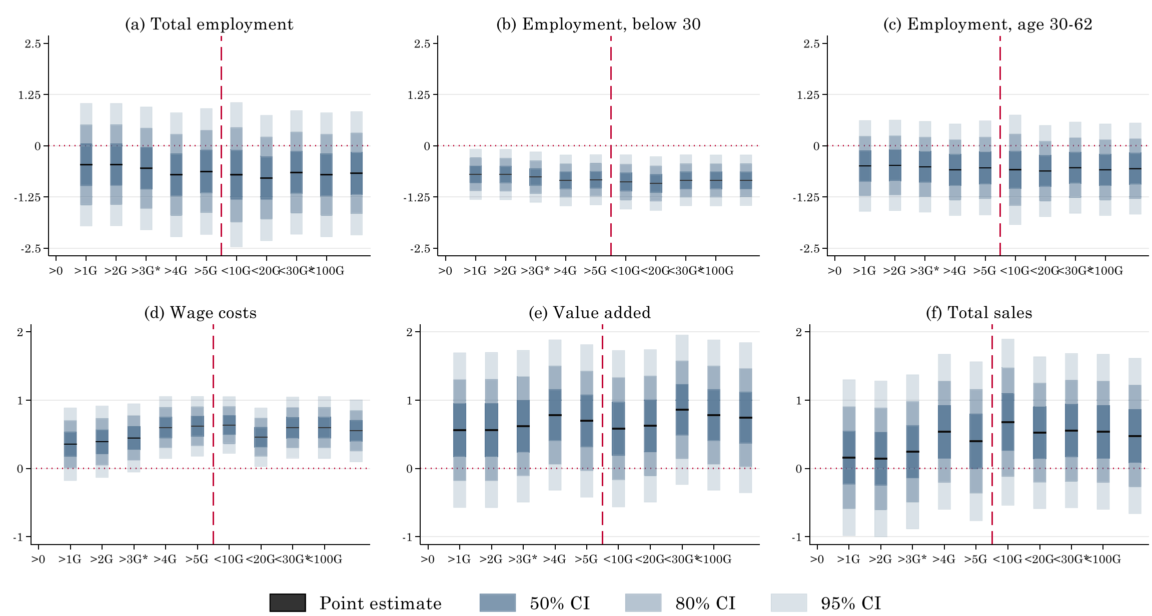

Fig. 6. Instrumental variables estimates with alternative sample selection criteria (with confidence fans)
Note: The estimates to the left of the vertical stapled lines are based on alternative lower data inclusion thresholds on total wage costs per full-time-full-year-equivalent worker. The thresholds are measured in G, which is the wage-growth-adjusted Basic amount used in the social insurance system in Norway (1 G is approximately equal to NOK100,000/€10,000 in 2020). The estimates to the right of the stapled lines are based on alternative upper limits. The sample selection criteria used in the baseline model are marked on the horizontal axis with *. Apart from differences in samples, all estimates are based on the model described in Section 5; i.e. the models denoted “Firm FE” in Fig. 5.

ficult to draw definite conclusions. It is notable, though, that none of the models estimated in this paper indicate negative effects of aging on any of these outcomes.

7. Heterogeneity

As noted in the introduction, we expect the relationship between age composition and firm outcomes to vary across firms with different production technology. In particular, the degree of complementarity between workers of different ages is critical for the way we expect postponed retirement of older workers to affect the demand for younger labor. To shed some light on possible heterogeneity in the influence of aging, we estimate separate effects for firms expected to differ along the dimension of complementarity. To describe the expected degree of complementarity between young and old labor, we use three alternative proxies defined at the industry level. The first is the steepness of

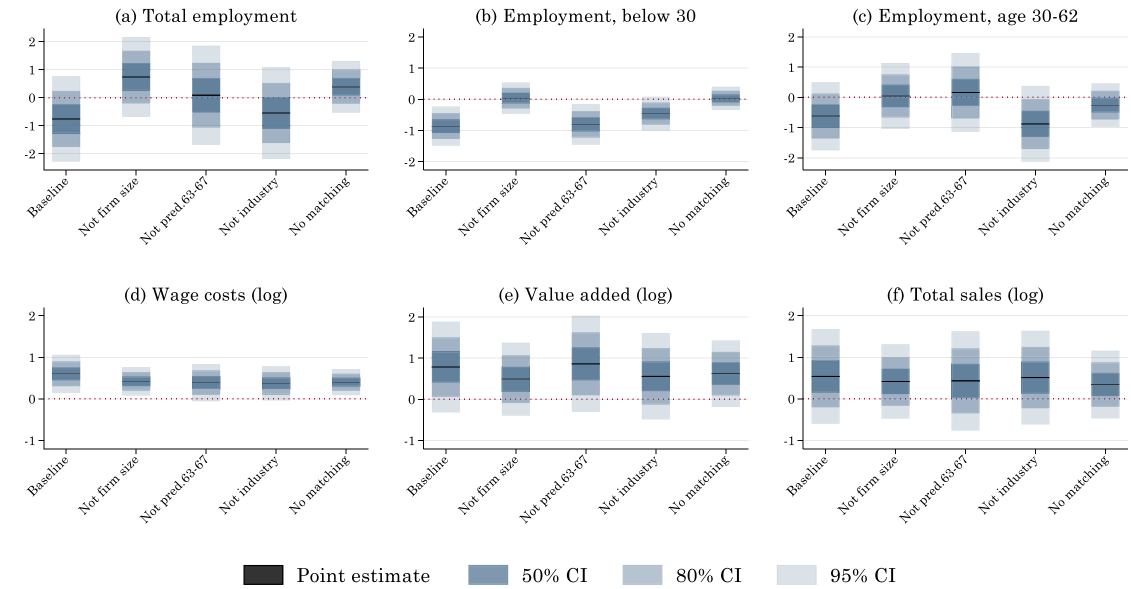

Fig. 7. Instrumental variables estimates with alternative matching algorithms (with confidence fans)
Note: The reported estimates are based on samples constructed from alternative criteria for the matching of non-AFP to AFP firms. In the Baseline model (from Section 5), data are matched based on firm size, the fraction of elderly workers, and industry. Bars 2–4 counted from the left show estimates when each of these matching criteria are dropped (but the other two maintained). The bars to the right show estimates when the data are used directly without any matching at all. Sample sizes (before any trimming) vary across the different panels, from 225,047 in the baseline model to 265,317 in the model without matching; see Table 1.

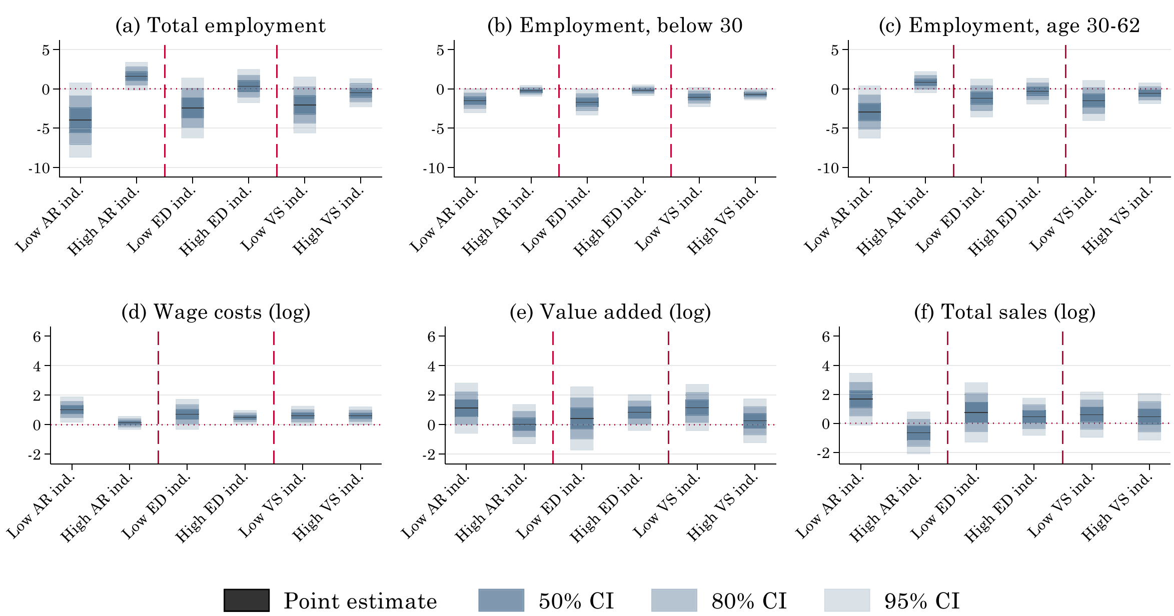

Fig. 8. Instrumental variables estimates for firms belonging to different industries (with confidence fans)
Note: The reported estimates are based on the baseline sample, but grouped into subsamples with separate regressions for each subsample. In the two models to the left (low and high AR ind.), the firms are grouped into industries with large and small age-difference in the wage rates (above and below the median). In the next two models (low and high ED ind.) firms are grouped into industries with high and low average education among its employees (over and above the median). Finally, in the two models to the right (low and high VS ind.), firms are grouped into industries with high and low within-firm variation in the age composition.

the age profile in wages (computed as the ratio of the average wage rate for workers aged 50–65 and workers aged 20–34 at the industry level). The idea is that large wage differences between different age groups indicate that young and old workers are indeed different, suggesting a potential role for complementarity. The second proxy is the average education level at the industry level. A higher education level indicates more sophisticated production technology and perhaps a larger role for experience relative to physical strength. The third proxy is the observed degree of age variation within firms in each industry. More specifically, we decompose the overall age variation among employees in each industry into a within-firm and an across-firm component. We then use the fraction of overall variance accounted for by the within-firm component as a proxy for potential young-old complementarity. Based on each of these three proxies, we divide the sample of firm-year observations into two equally large subsamples – below and above the median.

We realize that our proxies for the degree of complementarity between workers of different ages may be correlated to other industry characteristics that potentially are important for the structure of labor demand. Hence, we see this exercise primarily as an attempt to examine whether the patterns of estimated effects are consistent with a complementarity story, and not as strategy for producing decisive evidence for the importance of complementarity.

The results are presented in Fig. 8. Again, large confidence fans call for cautious interpretation, but there appears to be a pattern that in industries with high potential age complementarity, we do not find that the added elderly workers imply reduced demand for other age groups. Hence, overall employment in these industries appears to increase in response to postponed retirement. In these cases, we also find no effects on average wage costs and no (or very small) effects on productivity. By contrast, in firms characterized by more substitutability, we find negative effects on the demand for other age groups as well as for total labor demand, and indications of positive effects on wage costs and productivity. A plausible interpretation is that older and younger workers perform similar tasks in these firms, but that older and more experienced workers are both more costly and more efficient.[7]

8. Conclusion

In the present paper, we have used a policy-induced shift in the labor supply of elderly (age 63–67) workers to examine how a larger number/share of older workers affects labor productivity and the demand for younger workers. Our results are generally imprecise, but point estimates indicate that increased retention of older workers tend to slightly improve a firm's labor productivity in the short run. Although the individual estimates are subjected to considerable modeling as well as statistical uncertainties, we believe that, viewed as a whole, our findings should alleviate concerns that the aging of the workforce represent a drag on labor productivity. However, concerns that policies leading to postponed retirement of elderly workers may hurt employment opportunities for younger people are to some extent substantiated by our findings, at least in the short run. Higher employment among the elderly is offset by reduced hiring of young workers, ceteris paribus. The estimated impacts on total employment are particularly imprecise; hence we cannot rule out effects in either direction.

The average responses conceal considerable heterogeneity, particularly with respect to the firms' production technology. Based on alternative proxy variables, we have attempted to divide the population of firms into subsamples distinguished by the expected degree of complementarity between young and old workers. In firms expected to have high degree of complementarity, we find positive effects on employment, but negligible effects on average wage costs and productivity. In firms with low degree of complementarity, we find negative effects on employment and positive effects on both wage costs and productivity.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Data availability

The authors do not have permission to share data.

Acknowledgements

We wish to thank two anonymous referees and the Editor Wilbert van der Klaauw for constructive comments and suggestions. Administrative registers made available by Statistics Norway have been essential.

Funding

This research is part of the project "Behavior and sustainability in a new pension system" (PensjonsLAB; internal project number 1378), financed by the Ministry of Labor and Social Affairs (grant number 18/1025).

References

Acemoglu, D., Restrepo, P., 2018. Demographics and Automation. MIT, Department of Economics Working Paper 18-05.

Andersen, A.G., Markussen, S., Røed, K., 2021. Pension reform and the efficiency-equity trade-off: impacts of removing an early retirement subsidy. Labour Econ 72, 102050.

Boeri, T., Garibaldi, P., Moen, E.R., 2022. In medio stat victus. Labor demand effects of an increase in the retirement age. J. Popul. Econ. 35, 519–556. doi:10.1007/s00148-021-00871-0.

Cardoso, A.R., Guimaraes, P., Varejao, J., 2011. Are older workers worthy of their pay? An empirical investigation of age-productivity and age-wage nexuses. Economist (Leiden) 159, 95–111.

Carta, F., D'Amuri, F., von Wachter, T.M., 2021. Workforce aging, pension reforms and firm outcomes. NBER Working Paper 28407.

Chapter 4 in Christensen, A.M., Fredriksen, D., Lien, O.C., Stølen, N.M., 2012. Pension reform in Norway: combining an NDC approach and distributional goals. In: Holzmann, Robert, Palmer, Edward, Robalino, David (Eds.), Nonfinancial Defined Contribution Pension Schemes in a Changing Pension World. Eds. World Bank Publications. Volume 1, Progress, Lessons, and Implementation.

Coile, C., Milligan, K.S. and Wise, D.A. (2018) Social Security programs and retirement around the world: working longer – introduction and summary. NBER Working paper 24584.

Gordo, R., Skirbekk, V., 2013. Skill demand and the comparative advantage of jobs: jobs tasks and earnings from the 1980s to the 2000s in Germany. Labour Econ. 22, 61–69.

Göbel, C., Zwick, T., 2012. Age and productivity: sector differences. Economist (Leiden) 160, 35–57.

Grund, C., Westergaard, Nielsen, 2008. Age structure and the workforce and firm performance. Int. J. Manpow. 29 (5), 410–422.

Halvorsen, E., West-Pedersen, A., 2019. Closing the gender gap in pensions: a microsimulation analysis of the norwegian NDC pension system. J Eur Soc Policy 29 (No. 1), 130–143.

Hernæs, E., Jia, Z., Piggot, J., Vigtel, T.C., 2021. Work Less but Stay Longer. Working Paper 2021/30. CEPAR, University of New South Wales, Sydney.

Hernæs, E., Markussen, S., Piggott, J., Røed, K., 2016. Pension reform and labor supply. J. Public. Econ. 142, 39–55.

Kudrna, G. (2017) The norwegian pension reform: an external perspective. ARC centre of excellence in population ageing research. Working Paper 2017/07.

Lallemand, T., Rycx, F., 2009. Are older workers harmful for firm productivity? Economist (Leiden) 157, 273–292.

Lazear, E.P., 1979. Why is there mandatory retirement? J. Political Econ. 87 (6), 1261–1284. doi:10.1086/260835.

Mahlberg, B., Freund, I., Cuaresma, J.D., Prskawetz, A., 2013. Ageing, productivity and wages in Austria. Labour Econ. 22, 5–15.

Sharpe, A. (2011) Is ageing a drag on productivity growth? A review article on ageing, health and productivity: the economics of increased life expectancy. International Productivity Monitor 21.

Skirbekk, V., 2008. Age and productivity capacity: descriptions, causes and policy options. Ageing Horizons 8, 4–12.

Van Ours, J.C., Stoeldraijer, L., 2011. Age, wage and productivity in dutch manufacturing. Economist (Leiden) 159, 113–137.

[7] Note that even in the dataset with firms from industries with low age-return in the wage rate, the average wage for workers aged 50-65 is 17% higher than for workers aged 20-34. In the dataset with high age-return, the corresponding wage premium is 35%.