Firm downsizing, public policy, and the age structure of employment adjustments*

I. Sebastian Buhai[†] and Hans-Martin von Gaudecker[‡]

December 2012

Abstract

This paper studies the structure of workforce adjustments when firms facing adverse demand conditions are offered public financial incentives for downsizing. In particular, we are interested in how the age composition of employee outflows is shaped by corresponding age-dependent institutional arrangements. Our simple labour demand framework, with stochastic product demand and firing costs heterogenous in workers' early retirement eligibility, has as core prediction that distressed firms will dismiss with predilection those employees eligible to retire early. We test the model's implications on the entire set of mass layoff events in larger Danish private firms over 1980-2001, period covering several reforms to the early pension system. Our empirical conclusion is that firms behave as predicted by our model with regards to their lower-educated workforce, but not towards their higher-educated employees. We suggest that an extension of our firm-level model to narrow within-firm employee categories with potentially asymmetric turnover responses to firm-level demand shocks can rationalize this finding.

Keywords: early retirement, labour demand, employment adjustment, mass layoffs, LEED

JEL codes: H32, H55, J26, J65

*STILL PRELIMINARY. We thank Paul Bingley, Jason Faberman, Eric French, Daniel Hallberg, Ben Heijdra, Marno Verbeek, Till von Wachter, and audiences in seminars and conferences at NASM in Chicago, U. of Wisconsin-Milwaukee, and SOLE in Vancouver, for useful discussions. All remaining mistakes are our own. Financial support under a Netspar research grant is gratefully acknowledged.

[†]Aarhus University; email: sbuhai@gmail.com

[‡]University of Bonn; email: hmgaudecker@uni-bonn.de

1 Introduction

Most economic models of retirement behaviour focus on individual labour supply, ignoring the labour demand side. This contrasts with recent policy debates, where the spotlight is on economically distressed firms shedding their older workers, rather than on workers doing their best to retire early[1]. Supporting evidence for such practice is neither scant, nor controversial. For example, Dorn and Sousa-Poza (2007) highlight large shares of self-reported "involuntary retirement" among individuals from 19 OECD countries, using International Social Survey Programme data. Moreover, explicit and implicit instances of governments incentivizing private-sector employers who face declining demand to push older employees into early retirement are provided by, e.g., Hutchens (1999), Hakola and Uusitalo (2005), Hallberg (2011), and references cited therein[2].

Pure supply side models also have difficulties explaining two related sets of empirical findings. At the micro level, it has been shown that a late-career job loss substantially increases the (early) retirement hazard in subsequent periods (e.g., Chan and Stevens, 1999, 2001; Coile and Levine, 2007), and that job finding rates of the elderly unemployed are low, even in the U.S. (e.g., Maestas and Li, 2006). A similar pattern holds up at the macro level: recessions have been shown to lead to an increased inflow into retirement (e.g., Coile and Levine, 2009; von Wachter, 2007), and into disability (e.g., Autor and Duggan, 2003; Black et al., 2002).

Motivated by the evidence above, this paper attempts to connect—both theoretically and empirically—the labour demand and retirement literatures. We take the firm as observational unit and model the downward workforce adjustments due to negative demand shocks, in the presence of public policies. Our interest is specifically in how the composition of displaced employee outflows reacts to age-related institutional incentives for worker-firm separation. To that aim, we provide a parsimonious labour market demand model where shocks to a firm's product demand map into the firm's employment

[1] A quote from a recent *Economist* article on the Japanese labour market might be a good summary for such worldwide concerns: *"Lifetime employment, although still around is no longer guaranteed. Many companies impose 'early retirement' on workers in their 50's as a way to cut costs without publicly abandoning job security" (*"A new ice age. The perils of a frigid labour market", The Economist, 18 March 2010*)*

[2] Hutchens (1999) mentions the example of Germany's state-subsidized "59er provision", where distressed employers were directly assisted in workforce reduction with early retirement benefits in the form of unemployment insurance or public pensions for older employees. He also points out that—although this type of explicit subsidy does not *de jure* exist—similar financial incentives are implicitly provided also by state and federal governments in USA.

changes, and where the layoff cost per worker depends on the worker's eligibility for publicly financed early retirement benefits. In our framework, we adopt elements of Bentolila and Bertola (1990) and Hutchens (1999). The main testable implication resulting is that distressed employers dismiss relatively more employees eligible to retire, within the age groups targeted by early retirement policies. We test this hypothesis on Danish exhaustive longitudinal linked employer-employee data (LEED), covering the entire populations of workers and firms for more than two decades. The large cross-sectional and time-series dimensions of our data allow us to focus the empirical analysis on massive firm downsizing episodes, providing a means for identification of firm-induced separations (layoffs) as opposed to worker voluntary exits (quits), otherwise observationally indistinguishable in the data. The long time span also enables us to take into account several alterations in the early retirement schemes in Denmark, which were most generous in the beginning of our observation period and became substantially tighter later on, except for a "window of opportunity" during the recession of the early 1990s. We find that massively downsizing Danish firms behave in agreement with the prediction of our model vis-à-vis their lower-educated employees, but not towards their higher-educated co-workers. We discuss how we can rationalize this empirical finding.

The next section overviews related literature; Section 3 presents our theoretical framework; the relevant Danish institutional background is described in Section 4; we summarize in detail the LEED dataset in Section 5; the empirical framework and estimation results are the object of Section 6; in Section 7 conclusions are drawn.

2 Related literature

Briefly mentioned above, we know that early retirement is a frequent consequence of late-career job losses, be it worker self-classified involuntary job separations, e.g. Chan and Stevens (1999, 2001), job losses due to state-specific labour market conditions, e.g. Coile and Levine (2007), or displacements occurring in economy-wide recessions, e.g. von Wachter (2007), Coile and Levine (2009). This type of finding gives rise to interesting welfare considerations. Younger workers tend to fare worse in the long run following a displacement, but early retirement of older employees may be costly for the society as well. A first question arising is normative: where should policy makers set the eligibility for (early) retirement? Here, we investigate a second, positive consideration: what would be the incentives such age-related policy would create for the demand side of the labour market?

Contrary to empirical studies on individual retirement, inter alia Rust and Phelan (1997), Bingley and Lanot (2004), French (2005), or French and Jones (2011), social security programs are rarely treated explicitly in the empirical literature on job loss / business cycles and retirement. One notable exception—and closely related to our paper—is Hakola and Uusitalo (2005), who show, in the context of a unique Finnish pension reform with differential effects for different firms categories, that experience rating of the unemployment-related pension contributions induces a reduction in job exits at employers facing the larger cost increases for early retirement of their employees. Dorn and Sousa-Poza (2007) provide descriptive evidence on the extent of involuntary versus voluntary early retirement, and relate labour market institutions to the proportion of involuntary retirees across OECD countries. More recently, Hallberg (2011) uses an age-dependent variation in collective fee costs across Swedish companies to suggest that early retirement is likely used by firms to lower their labour costs[3].

The latter three empirical studies mentioned relate their findings to implications of Hutchens's (1999) demand-side-driven individual retirement framework[4]. Hutchens develops an implicit contract framework between one worker and one firm, where the firm may use (public or partially private) early retirement benefits as a form of insurance for the worker in case of layoff[5]. Our modelling innovation is to explicitly introduce Hutchens's (1990) publicly subsidized employee dismissal idea in a variant of Bentolila and Bertola's (1990) partial equilibrium model of a firm's employment adjustment responses to its product demand evolution under uncertainty.

Finally, this study is methodologically related to the large literature on job displacements and their consequences. Starting with Jacobson et al (1993), there has been a rapid development in research aiming to quantify the long-term effects of job displacement on various worker outcomes. In particular, the increased availability of large administrative employer-employee datasets has spurred a recent wave of papers, see Carneiro and Portugal (2006), von Wachter et al (2009), Manchester (2009), Sullivan and von Wachter

[3]There are also a couple of related empirical studies that investigate disability program entry in connection with labour demand and/or institutional rules, c.f. Korkeamaki and Kyyra (2010) and references cited by them.

[4]The working paper version of Hakola and Uusitalo (2005) actually contains an early retirement contract model heavily drawn on Hutchens (1999): see Hakola and Uusitalo (2001, Section 3).

[5]Hutchens (1999) builds his model based on arguments from an earlier large literature on (imperfectly experience-rated) unemployment insurance effectively subsidizing worker layoffs, e.g., Feldstein (1976, 1978), Topel (1983, 1984).

(2009), Couch and Placzek (2010) for a subset. One well-established result in this literature is that displaced workers fare worse in terms of earnings or health, even decades after. It is reasonable to assume that there are inputs to worker productivity, which are unobserved by the econometrician but observed by employers, at least after a sufficiently long match. This gives rise to a selection problem, as it is in the firm's best interest to fire those workers whose idiosyncratic productivity is lowest, given observable characteristics. In order to reduce the impact of such issues, this literature has typically focused on mass layoffs, an approach we also adopt.

Next to the mentioned study by Hakola and Uusitalo (2005), two other studies in the literature on worker displacement are particularly related to our paper, for different reasons. The first is Pfann (2006), who uses detailed personnel data and information on bankruptcy/ restructuring for Fokker—an aircraft company responsible for the largest mass layoff in Dutch history—to document how firing costs and outcomes depend on worker-job characteristics such as age, education, tenure, job performance; moreover, he also rationalizes his findings in the context of a dynamic labour demand model. The second is Tatsiramos (2010), who uses information on involuntary versus voluntary job separations using European Community Household Panel data to study differences, by country, nature of separation, and worker demographics, in individual worker exit hazards to distinct labour market states, including early retirement. Part of our contribution is to provide a further input to this type of research, namely to inform on the role of age-related public policy in shaping the pool of displaced workers.

3 Labour adjustment with firing costs heterogenous in employee eligibility for early retirement

3.1 Intuition

In accordance with a large literature on optimal contracts under unemployment insurance, for instance Feldstein (1976, 1978) or Topel (1984), introducing early retirement benefits would generally lead to inefficient outcomes. While we remain silent on the rationale for the presence and the aggregate welfare implications of the early retirement policy, here we explore joint firm-worker behaviour optimization, given public early retirement instruments already in place. We argue that they may be strategically utilized by firms hit by negative economic shocks as implicit subsidy for necessary workforce reductions.

Even in the particular case of Denmark– entailing a "flexicure" labour market with almost inexistent *direct* firing costs and generous unemployment benefits[6]– there are obvious reasons why early retirement of eligible employees could be for a downsizing firm more attractive than indiscriminate layoffs throughout its employees' age distribution. *Indirect* costs may for instance arise as a firm's reputation for poorly justified dismissals would lead to difficulties in recruiting young talent in the future. Hence, in cases where the employer seeks to scale down employment because of declining product demand (or: technological shifts), pushing older workers into publicly financed early retirement schemes is effectively cheaper than pushing them or their younger colleagues into unemployment[7].

Extending Hutchens's (1999) idea of the firm using public early retirement benefits as insurance for laying off its worker to a many-worker environment, translates in the employer being de facto subsidized by the government to force its eligible employees to retire early, when it needs to downsize. We embed this idea in a Bentolila and Bertola (1990)-type framework of firm employment adjustment responses to the (exogenous) stochastic evolution of the firm's product demand, sketching a simple partial equilibrium model of firm-level workforce adjustment with dismissal costs varying in the workers' eligibility for early retirement, in the presence of publicly financed early retirement benefits.

We envisage the following intuitive temporal structure, to be formalized below:

- A profit-maximizing, risk-neutral firm suffers a negative shock to expected profits, which translates in downward adjustment in its labour workforce, the unique production input. The firm determines the required employment adjustment size.

- The firm offers a menu of contracts to each employee, including termination of the employment relationship by either dismissal or early retirement. Except for heterogeneity in eligibility for early retirement, workers are identical[8]. Early retirement is assumed to be less costly than unemployment. [9].

[6] OECD (1999, Table 2.2 or Chart 2.1) shows in detail that Denmark has had very low dismissal protection for full-time employees throughout our entire data observation period, 1980s through late 1990s (similar to other OECD countries like Switzerland, Australia, Canada, UK or US). In Albaek et al (2002, Table 1), out of 53 countries, Denmark is ranked 1st in terms of "flexible hiring and firing", 10th in terms of low "legislative restriction on firing", and 46th in terms of "unemployment insurance meanness".

[7] Using the early retirement channel by distressed firms can be interpreted, for instance, as a way to reduce their partly *psychic* workforce adjustment costs.

[8] In the empirical application we control for observable worker and firm characteristics, such that we compare "otherwise identical" workers inasmuch as possible.

[9] We argued above that indirect costs, e.g., instance reputation-related or psychic, could make straight

- Realized outcomes are materialized through empirically observed employee hazards into different states: continued employment at the same firm, employment at another firm, receiving unemployment insurance benefits, or receiving (early) retirement pension.

For modelling the first stage, we adapt the setup developed by Bentolila and Bertola (1990), compare also Buhai et al (2009) for a different application; the main idea in Hutchens (1999) is used for modelling stage two.

3.2 A variant of Bentolila and Bertola (1990)

Consider a world with risk neutral firms and workers. Workers are infinitely lived and maximize the discounted value of their expected income. Firms maximize the discounted value of their expected profits. Firms face a stochastic constant-elasticity demand curve for their output; using lower case letters for logarithms,

$$n_t = z_t - \eta p_t, \tag{1}$$

where n_t is log demand, z_t is a market index capturing the exogenous stochastic evolution of the product demand, $\eta > 1$ is the price elasticity of demand–implying that the firm has some monopoly power– and p_t is the log price of the firm's product. The diffusion process for z_t is assumed to follow a Brownian motion with drift, such that $\Delta z \sim N(\mu, \sigma^2)$. Labor is the only factor of production. The production function exhibits constant returns to scale. Productivity is normalized to unity, so that output is equal to employment. The firm pays a constant log wage w to each of its workers, and all workers have the same log reservation wage, normalized to 0 [10].

Bentolila and Bertola (1990) assume that employment adjustment of firms is (linearly, asymmetrically) costly, on both the hiring and firing margins[11]. Here, we shall mainly

dismissal effectively more expensive for a firm than pushing of employees in early retirement, even in the Danish context of nearly zero direct firing costs. From the other end, an employee is in practice also (marginally) better off receiving early retirement than unemployment benefits, as described in the institutional background section. In our model we assume for simplicity that all employees value equally their exit options (retirement or unemployment), but conclusions do not change for firm-induced separations as dealt with here.

[10] This simple setup (slightly different from Bentolila and Bertola's) can be generalized—at the cost of complexity in derivation; for instance, we can allow for decreasing returns to scale, stochastic productivity, or wages following an Ito process with known parameters; e.g., in Buhai et al (2009) wages are allowed to depend on the state of demand.

[11] In their paper, workers can also quit with some exogenous probability, at no cost. For simplicity and

focus on the irreversible firing cost per worker, F. In a different application, Buhai et al (2009) focused on the irreversible per-worker hiring cost H, interpreted there as the specific investment I that has to be made at the start of the employment relationship. As argued earlier, we interpret the firing cost F more broadly, as the sum of all direct and indirect firing costs, in the latter including for instance reputation, or psychic costs of firing employees. For the moment think of F as the same for all workers; later on we allow F to vary among workers based on their eligibility for early retirement benefits.

In the benchmark case, where firing and hiring costs are inexistent: $H = F = 0$, the firm can adjust its employment costlessly, at any point in time. The firm is thus maximizing profits Π at any time t:

$$\Pi_t = e^{n_t} \left(e^{p_t} - 1 \right),$$

subject to (1). The FOC for this case is straightforward:

$$\begin{aligned} p_t &= \ln \frac{\eta}{\eta - 1} > 0, \\ n_t &= z_t - \eta \ln \frac{\eta}{\eta - 1} \end{aligned} \tag{2}$$

Parameter $\ln \frac{\eta}{\eta-1}$ is the log of the ratio of price over wage cost, when marginal cost and marginal revenue are equal; it is strictly positive due to the monopoly power of the firm in the market. The firm's log price p is constant over time, while its log labor demand n follows a continuous-time random walk, inherited from the assumption for the log market index z.

The non-trivial case is when the firm cannot costlessly adjust employment: $H+F > 0$. Since H and F are irreversible, hiring new workers or laying off incumbent workers have option values, and employment adjustment will exhibit the typical lumpy behavior. Our starting point for solving the dynamic optimization problem of the firm is similar to that of Bentolila and Bertola (1990). Namely, for one worker to be hired or fired, we take the employment of all his incumbent co-workers as given, and consider when it is optimal for the firm to hire and subsequently fire this new worker. We can therefore consider the decision to hire and fire the N_t-th worker of the firm separately of the hiring and firing of workers before or after her. This essentially becomes then the same exercise as

since this has no impact on our conclusions, we assume here that employee separations are all involuntary; consistently, our model is tested only on mass layoff episodes in the data.

the firm faced with the investment/ abandonment decisions under exogenous uncertainty from Dixit (1989, p. 625-630) or Dixit and Pindyck (1994, p. 218-221).

Denote by $F(p_t|n_t, z_t)$, where (1) applies, the asset value of the firm for its N_t-th worker (i.e. the firm's choices are either keeping or firing the N_t-th worker; this case is analogue to the option value of the "active firm" in Dixit and Pindyck's examples). The standard Bellman equation for $F(p_t)$ reads:

$$\rho F(p_t) = \left[\exp\left(p_t - \ln\frac{\eta}{\eta-1}\right) - \exp(w)\right] + \mu F^{'}(p_t) + \frac{1}{2}\sigma^2 F^{''}(p_t) \quad (3)$$

ρ is the exogenous discount rate[12], satisfying $\rho > \mu + \frac{1}{2}\sigma^2$. The term in square brackets is the marginal revenue product of the N_t-th worker, see (2); the last two terms capture the option value of firing the N_t-th worker.

Denote now by $H(p_t|n_t, z_t)$ the asset value of a firm that looks into hiring the N_t-th worker (analogous to the case of the "idle firm" in Dixit and Pindyck, 1994). Unlike in the case analysed above, there are no current operating revenues or costs of the firm for this N_t-th worker, hence in this case only the option value terms matters. The corresponding Bellman equation reads:

$$\rho H(p_t) = \mu H^{'}(p_t) + \frac{1}{2}\sigma^2 H^{''}(p_t) \quad (4)$$

Given (3) and (4) above, we can write down the value matching and smooth pasting conditions, at the firing and hiring boundaries:

$$\begin{aligned} F(p_F) &= H(p_F) - F, \\ F^{'}(p_F) &= H^{'}(p_F), \\ F(p_H) - H &= H(p_H) \\ F^{'}(p_H) &= H^{'}(p_H). \end{aligned} \quad (5)$$

The first pair refers to the firing decision, the second to the hiring decision. (5) can be expressed, based on the general form solutions for (3) and (4) above, as a system of 4 equations nonlinear in 4 unknowns. Although it is generally hard to compute analytical

[12] Since we assume risk neutrality for firms (hence, discount rate is equal to the risk-free interest rate), we could also solve this dynamic optimization problem by contingent claims analysis.

solutions in this case, it can be proven that an economically meaningful, unique solution exists, see, e.g., the appendix in Buhai et al (2009). Using a similar line of reasoning (with different notation), Bentolila and Bertola (1990, p. 385-390) show that the firm's optimal strategy is to hire workers whenever p_t reaches a constant upper bound $p_H > \ln \frac{\eta}{\eta-1}$ and to fire them whenever p_t reaches a lower bound $p_F < \ln \frac{\eta}{\eta-1}$. The hiring bound p_H exceeds $\ln \frac{\eta}{\eta-1}$ due to the option value of postponing hiring, while the firing threshold p_F is below $\ln \frac{\eta}{\eta-1}$ due to the option value of delaying firing. The implication is that when p_t follows a random walk between p_H and p_F, n_t will be held constant. However, when p_t drifts outside these boundaries, the firm acts on its labour size n_t, which starts drifting, either upwards (if $p = p_H$), i.e. the firm is hiring, or downwards (if $p = p_F$), i.e. the firm is firing. The following proposition summarizes this first implication of our model.

Proposition 1 *The first implication of our model is that, for a firm with demand characterized by (1), log firm size n_t is characterized by a Brownian diffusion process, apart from short-run fluctuations of p_t within the firing-hiring interval [p_H, p_F].*

This prediction is consistent with a variant of Gibrat's law for firm size, which tends to hold for large firms, c.f. Jovanovic (1982).

3.3 Heterogenous dismissal costs à la Hutchens (1999)

The second temporal stage of our model deals with the firm selecting whom to fire. To that end, we seek an implication concerning the variation of the optimal firing cutoff p_F with the firing cost F. We use (4) and (3) to define an auxiliary function, $G(p_t|n_t, z_t) = F(p_t|n_t, z_t) - H(p_t|n_t, z_t)$. Write the corresponding Bellman equation

$$\rho G(p_t) = \exp\left(p_t - \ln\frac{\eta}{\eta-1}\right) - \exp(w) + \mu G'(p_t) + \frac{1}{2}\sigma^2 G''(p_t) \tag{6}$$

The value matching and smooth pasting conditions of (6) that hold at firing threshold p_F, analogous to the first two equations from (5) will then be

$$\begin{aligned} G(p_F) &= -F, \\ G'(p_F) &= 0, \end{aligned} \tag{7}$$

If we now evaluate (6) at the firing boundary $p_t = p_F$, considering (7), we obtain:

$$-\rho F = \exp\left[p_F - \ln\frac{\eta}{\eta - 1}\right] - \exp(w) + \frac{1}{2}\sigma^2 G''(p_F)$$

Changing to level notation for the price cutoff and wage, i.e. $P_F = \exp(p_F)$, $W = \exp(w)$, the above can be written:

$$P_F = -\rho\frac{\eta}{\eta - 1}F + W - \frac{1}{2}\sigma^2 G''(\ln P_F) \tag{8}$$

From (8), noting $\partial G''(.)/\partial F = 0$, it is straightforward to establish:

$$\frac{\partial P_F}{\partial F} = -\rho\frac{\eta}{\eta - 1} < 0 \tag{9}$$

In (9) we show that the firing bound P_F decreases with F, the per worker layoff cost; this result is corroborated by previous studies: Pfann (2006, p. 160) obtains the inequality sign by applying the exact derivation in Dixit (1989, p. 630); a similar conclusion is reached at aggregate level (i.e., an economy-level increase in firing cost implies overall reduction in firing—and hiring) by Bentolila and Bertola (1990, p. 390-391)[13].

Up until now, we have thought of the firing cost F as the same for all (identical) workers of a firm. However, (9) can be interpreted both in the context of varying F among firms (or over time, for the same firms), as well as for varying F among *co-workers within the same firm*. In the latter context, with firing costs varying among (otherwise similar) employees, the more costly employees will be more likely to be retained by the firm, and viceversa[14]. As mentioned earlier, Hutchens (1999) introduces the idea of the firm using existent social security policies, such as early retirement, as effective "subsidy" for necessary layoffs. We incorporate this feature in our model by introducing a source of heterogeneity among workers, namely their being eligible or not for early retirement, and making the firing cost F dependent on that eligibility[15].

[13]This aggregate result has a more general validity; for instance, Hopenhayn and Rogerson (1993) obtain it in a general equilibrium labor demand model.

[14]Keep in mind that wage W is assumed here as the same for each employee. Assuming wage heterogeneity among *otherwise identical* workers, would imply that workers paid more will be fired first; we can see that by taking the partial derivative with respect to wages in (8). Nonetheless, assessing who will be fired with *both wage and layoff cost heterogeneity* needlessly complicates the theoretical analysis. Instead, we account for observed worker wages and other worker demographics in our empirical analyses.

[15]We do not allow the workers' outside option to depend on their eligibility for early retirement, see also footnote 9 earlier. Firstly, though individual early exit preferences might matter, the bulk of empirical

Define the eligibility status for early retirement $R \in \{0, 1|1 = \text{eligible for early retirement}\}$. As argued in earlier sections of this paper, a downsizing firm will find it cheaper to send off employees into publicly financed early retirement rather than in unemployment, even in countries with very low *direct* firing costs and very generous unemployment benefits, such as Denmark; the fact that in the main empirical application we restrict the analysis to mass layoff episodes further strengthens the validity of this assumption[16]. Hence, by assumption:

$$F_{R=0} > F_{R=1} \geq 0 \tag{10}$$

As stated earlier, the firm will determine the required size of its workforce reduction and will select which of its employees to lay off. Denote the firm's number of workers to be displaced by N^{exit}, and the firm's number of workers eligible to early retire by $N_{R=1}$. In light of (9) and (10), we will then have for any worker i in downsizing firm j:

$$\begin{aligned} 1 \quad &\geq \quad \Pr(exit_{ij}|N_j^{exit}, R_i = 1) > \Pr(exit_{ij}|N_j^{exit}, R_i = 0) \geq 0 \\ \text{with} \quad &= \quad \text{instead of } \geq \text{ iff } N_j^{exit} \leq N_{j,R=1} \end{aligned} \tag{11}$$

In words, expression (11) states that the probability the firm of firing employees eligible to retire early is always higher than the probability of firing their co-workers, with the first probability being 1 and the latter 0 whenever the number of displaced workers is lower than or equal to the total number of eligible employees. The following proposition summarizes this second prediction of our model.

Proposition 2 *The main implication of our model is that for a downsizing firm experiencing a demand law described by (1), with worker-heterogenous layoff costs characterized by (10), firing with predilection its employees eligible for early retirement, as formalized in (11), is a profit maximizing strategy.*

studies on early retirement from a labour-supply perspective suggests that voluntary individual early retirement is *not* particularly sensitive to changes in social security benefits, e.g., Hurd (1990) or Krueger and Pischke (1992). Secondly, to the extent that such individual worker incentives materialize through worker quits into early retirement, this is addressed in the empirical application, by focusing only on mass layoff events.

[16] Pfann (2006, Fig. 3) obtains a direct measure of firing cost per worker, using detailed personnel data corroborated with bankruptcy regulations; he shows that while firing costs generally increase by age, they drop dramatically just before the Dutch early retirement age=55. Unsurprisingly, Fokker laid off its employees with ages around the early retirement age with a much higher probability than other workers (e.g, 77% probability for firing of a 54 years old, cf. p. 164). These findings are giving external support for both the validity of our key assumption here, i.e. layoff costs decreasing in employee eligibility for early retirement, as well as for our key model implication. See also the discussion in our final section.

The last stage of the temporal framework presented in the beginning of our theory section is the realization of outcomes following the mass layoff. In the empirical section, we use data on the composition and the labour market destination state of the displaced worker flows to test the empirical correspondent of the above implication.

4 Social security for the elderly in Denmark

Danish labour market institutions famously combine a flexible hiring and firing legislation with a large basis for temporary income support in case of out-of-work, e.g., OECD (1999), Albaek et al (2002). In most other countries – and especially so in the rest of continental Europe– disentangling various incentive effects for (early) exiting from the labour market would be difficult, for instance due to frequently changing, non-uniform legal barriers to firing (especially older) employees. In Denmark the employee dismissal flexibility has been in place for a long time, in particular covering all our data interval. Furthermore, retirement and other social security benefit provisions are typically more straightforward in Denmark than in a host of other countries studied previously, for instance, publicly financed and uniformly applied across firms and workers. The Danish labour market context will thus be of particular help in distinguishing between causes for worker turnover, including for observed early retirement.

We briefly describe the Danish pension system for private sector employees[17], mainly based on Bingley and Lanot (2004, 2007) and Larsen and Pedersen (2008). Public pensions constituted 10% of the Danish GDP in 1994, similar to the OECD average of 9%. Until 2003, all Danes at least 67 years old were covered by the publicly financed old age pension, i.e. "*folkepension*", regardless of previous attachment to the labour market. In 1992, this was worth 40% of average production worker earnings. In the same year, the actual average retirement age was more than five years below, at 61.5 years.

[17] A number of other Danish or international old-age social security provisions do not have implications in our context: i). In the public sector, a so-called "*tjænstemændspension*" = public sector employees pension, is available from age 60, for selected occupations. This does not have a bite here, as we focus on privately owned firms; ii). Within the time window considered, there were no other major social policy reforms *specifically targetting the elderly category* e.g. in terms of disability pensions, social assistance, or unemployment insurance, which could have complicated the early retirement choice; iii). *Private pensions* represent a very low proportion of retirement income over the period considered, and are unlikely to matter empirically for retirement decisions, c.f. Bingley and Lanot (2004). iv). Lastly, health insurance is universal, independent of the employer-employee relationship, and financed through national taxation, hence typical health insurance coverage concerns in the US—which play a role in the incentives for separation, e.g. French and Jones (2011)—are irrelevant for Denmark.

The main policy instrument facilitating early retirement has been the "*efterløn*" (="post-employment wage"). The *efterløn* was introduced in 1979 and required that individuals were at least 60 years old and member of an Unemployment Insurance Fund (UI Fund) for 5 of the last 10 years to be eligible. The latter requirement was substantially tightened in various steps during the 1980s, and once more in 2000; see Table 1 for a complete overview of changes. Participation in the program stabilized to an average of about 100,000 individuals per year, which is a sizable proportion among the elderly out of work: to place this number into financial perspective, 25% of all pension transfers in Denmark were in the form of the *efterløn* in 1994.

The early retirement benefit is equal to the unemployment insurance (UI) benefit, "*arbejdsløshedsforsikring*", namely 90% of former wages, up to a ceiling equalling 60% of average production wages, for 2-and-a-half years, and 80% of the previous amount thereafter, until becoming eligible for the *folkepension.* For eligible out-of-work individuals the *efterløn* is thus more appealing than the UI benefit, the latter being withdrawn after these initial 30 months, when social assistance="*bistandshjælp*", the lowest level of income support (about 24% of average production wages), becomes the alternative. *Efterløn* beneficiaries are also not subject to job search requirements, contrary to UI recipients for whom the requirement is strictly enforced[18]. Hence, *efterløn* is de facto the most attractive early exit route from work[19] for private sector workers in Denmark, and relative to employment it is most appealing for low wage earners[20].

One extension to the *efterløn* was opened up in 1992 with the "*overgangsydelse*", or transitional benefits programme, available to those 55 years and older. In this case benefits were slightly lower and eligibility required a period of unemployment immediately prior to accession to the programme, see again Table 1 above for the details. Two

[18] There were however slight changes over time in the enforcement strength for the older unemployed; e.g., before the late 90s "activation measures" consisting in proof of employment search and/or training, are known to have been more lax for unemployed of >50 years old.

[19] Receiving *efterløn* used to allow the individual to still work up to 200 hours per year (exceeding this amount lead to permanent disqualification from the program). In 1998, the 200 hours limit for working while on efterløn was abolished to be replaced with earnings clawbacks. This change affects the last 3 years of our observation period, but excluding these years does not affect our conclusions. In line with Bingley and Lanot (2004, 2007), we consider the 200 hours level insufficient for either part-time work or partial retirement, Hence, we equate entry into early retirement status with permanent exit from the labour force.

[20] Being equal to the maximum unemployment benefit and hence subject to an income cap, the maximum *efterløn* amount roughly corresponds to the full-time pay of a minimum wage earner; furthermore it is binding for almost all full-time workers, as the lowest observed wages are high and the wage dispersion is relatively low in Denmark. Hence, these benefits are generally independent of previous wages, which makes them more attractive for low-wage, unskilled workers.

Table 1: Overview Danish pension reforms 1979-2003

Year	Scheme	Description of change
1979	Efterløn	Introduction. Replacement rate equals UI benefit for 30 months, thereafter 80% of previous amount. Eligibility: At least 60 years old, member of unemployment insurance fund for at least 5 out of the last 10 years
1980	Efterløn	Eligibility: Requirement of membership in UI fund of 10/15 years.
1985	Efterløn	Eligibility: Requirement of membership in UI fund of 15/20 years.
1990	Efterløn	Eligibility: Requirement of membership in UI fund of 20/25 years.
1992	Overgangsydelse	Introduction. Replacement rate equals 82% of UI benefit. Eligibility: At least 55 years old, membership in UI fund, unemployed for 12/15 months. Note: Upon reaching age 60, programme is replaced by efterløn
1994	Overgangsydelse	Eligibility: At least 50 years old, the other criteria as in 1992.
1996	Overgangsydelse	Closed for entry.
2000	Efterløn	Eligibility: Membership in UI fund 25/30 years.
2003	Folkepension	Eligibility age lowered from 67 to 65.

Efterløn=post-employment wage; Folkepension=state public pension; Overgangsdydelse=transitional benefits program.

years later, the eligibility age was lowered to 50 years, before the programme was shut down for new applicants after two more years. Given its relatively short life span and the prior unemployment window required for eligibility—which makes difficult both the interpretation in the light of our model, as well as assigning with sufficient precision the transitional benefit eligibility in our annual frequency data—we do not make use of the *overgangsydelse* scheme in our empirical analysis.

The reforms in the *efterløn* scheme imply variation in early retirement options over time[21], with changes in the *efterløn* eligibility contributing most to this variation. Our data allow us to calculate individual retirement eligibility for every Danish resident, along with all the changes in their employers' workforces, within the time period under study.

5 Data

5.1 Description and structuring

We use the *Integrated Database for Labor Market Research* (IDA), for the period 1980-2001, produced and maintained by *Statistics Denmark*. IDA tracks every single Danish legal resident between 15 and 74 years old. Each individual's labor market status is recorded at November 30th, every year. At that date we know whether the person is

[21] A second source of variation, in replacement rates, arises in the cross-section; however, since all benefits are subjected to a cap, and as the wage distribution is relatively compressed in Denmark, this becomes of secondary importance, see also the previous footnote.

employed, unemployed, or out of the labour force—including in the latter retirement status subcategories such as early retirement or old-age retirement, disability, etc. The dataset contains individual, plant (establishment), and firm (business unit) identifiers, allowing the linkage between individuals and plants, and respectively between plants and firms. There is no attrition other than due to natural causes or out of country migration in the population of individuals, and similarly, there is no attrition other than due to plant closure within the universe of establishments. There are some typical matched employer-employee data issues regarding the consistency of the firm identifier over time (for instance, when ownership changes in a legal sense because of mergers), but to a large extent Statistics Denmark is able to account for that. There are also some worries of potential misclassification of labour market states, especially relevant before 1996, which we attempt to address as explained in detail in a following subsection.

We have information on the worker's hourly gross earnings, occupation, education, age and gender, and on the establishment's location and industry. Since in our theory framework shocks to the product demand of the firm map in the firm's labour force adjustments, we take the firm—as opposed to the establishment—as our unit of analysis. Industry of a firm is defined as the industry employing the largest share of the firm's workforce. We compute the firm employment size as the number of individuals holding "primary jobs"[22] at any establishments of the firm, and earning a positive wage. The firm tenure of workers hired since 1980 is easily computed by observing the individual working for the same firm in the data, over time; for workers hired between 1964-1980 the tenure can be computed based on a second dataset on individual contribution histories to ATP, the mandatory "Danish labour market supplementary pension fund". Tenure in spells started before year 1964 is left censored (less than 3% of the observations); we discard those observations affected from the individual-level empirical analysis where we control for individual tenure, but not from our aggregate firm-level analysis. We also compute a proxy for firm age as the longest tenure level among all its employees. A worker's potential experience[23] is constructed as age-schooling-6, where schooling is the education

[22] "Primary jobs", as identified by Statistics Denmark, are jobs in which the individuals work more than 50% of their registered working hours; an individual can also hold secondary jobs, which are not included in our data. Primary jobs can be both full-time (more than 30 hours a week) and part-time. However, the firm size computed by our method is very similar to the firm size computed as full-time-equivalent (FTE) at each establishment level–the correlation between our measure and the reported employment size is higher than 0.97 for each of the 21 years of data–hence most jobs in our sample are de facto full-time jobs.

[23] Actual experience is available; however, before 1980 it needs to be constructed based on the ATP pension payments available since 1964, like tenure, see above. Hence, we can select for individual-level

level at the start of the worker's first job. Details for assigning worker eligibility for (early) retirement and a discussion on the classification of labour market states in the data follow in a separate subsection.

One important data caveat is that we do not know the cause of worker-firm separations, i.e., we cannot observationally distinguish between quits and layoffs. This requires strategies that ensure we deal with firm-initiated separations, consistent with our theory model; to that end, we focus on massive yearly downsizing events (equating them with mass layoffs) at the firm level[24]; we give below details of their identification in the data.

We use data on all private sector jobs, discarding those in the farming, fishing and mining industries. As our empirical strategy relies on observing mass layoffs, we focus on the sample of larger employers: we select for analysis those firms with more than 50 employees in any single observation year[25]. Though this accounts for a small fraction of the firms' universe, it does account for a large share of employed individuals in the economy. Hence, our working sample will contain all the firms with more than 50 employees, ever active in the period 1980-2001, and all the individuals ever employed by them over that time period.

5.2 Assigning individual eligibility for early retirement

As already mentioned, the *efterløn* program has been active since 1979, with program eligibility conditional on being between $60-66$ years old (from 67 onwards the folkepension or old-age retirement kicked in), being a current UI Fund member, and having been a UI Fund member for 5 out of the last 10 years. Before 1979 there were no pension rights associated with UI Fund membership. Based on the rules described in Table 1 above, we are able to assign for each individual, in every year, an indicator of whether s/he is eligible for early retirement. Our observation sample starts with the 1980 cross-section; in order to compute individual eligibility in the early years of our sample, we need knowledge

analysis only workers with at most 16 yrs of actual experience in 1980, implying a far larger discarded sample than in the case of left censored tenures. All results are qualitatively robust to using actual experience in the smaller sample.

[24] Mass layoff events are also most likely to be reactions of firms facing large adverse demand shocks, often as consequence of economy-wide recessions—when publicly-funded policies like early retirement are even more relevant as exit options than in periods of milder slack demand.

[25] The reason to select large firms is pragmatic: smaller firms hit by adverse demand shocks are more likely to shut down completely than to reduce their workforce; very small firms might also function more in "family" regimes, often with unusual constraints in their optimization strategies. Our empirical results remain qualitatively the same if the cutoff for selecting firms is set at a minimum of 20, rather than 50, employees per year.

of UI fund membership also for up to 15 years before. Although we do not observe UI fund membership prior to the starting of the sample period, we know the individual employment history for all individuals between 1964-1979, due to the earlier mentioned dataset on mandatory pension contributions (ATP). We follow Bingley and Lanot (2004) and impute UI Fund membership for those missing years using *ATP* data, making the assumption that membership status in 1980 fully reflects membership status during the preceding years. Bingley and Lanot (2004) justify this assumption given that few workers, i.e. only 3% of those over 44, actually change UI Fund membership status.

Having assigned individual eligibility for each individual worker, we compute for each firm the share of early retirement eligibles relative to the *targeted older workforce*, i.e. the workers aged 60-66; we also do this aggregate computation by low- and respectively high-educated worker categories within each firm.

5.3 (Mis)Classification of individual labour market states

Of primary importance for our empirical exercise is the coding precision of the relevant individual labour market states in the data. As earlier stated, IDA directly provides a variable indicating the individual labour market state in the last week of November, each year. The individual can be for instance (self)employed, unemployed, retired—and if retired, he could be receiving *efterløn*, transitional benefits, or old age pension. However, this information has not always been precisely coded, and therefore we cannot rely on it solely to identify the actual labour market state. The following observations address the major potential classification problems:

- the labour market "occupational code" (indicating official retirement, early retirement, employment, unemployment etc.) changed its official definition in 1996; up to that year there was considerably more misclassification among the retirement subcategories. For instance, before 1996 "officially retired", supposed to be the same as "old age retirement" (*folkepension*), bunched anyone else not linked with an employer, or a wage, or reporting a low wage, even when the individual in question was young. While this happened partly because, e.g., disability benefit recipients had been included in the "officially retired" category before a separate category for them was created, there are also instances of included employed individuals, with salaries higher than the median wage, hence misclassification is a concern. The number of "officially retired" persons roughly halfs after 1996, when the definition changes such that it includes only those receiving the *folkepension*

- there exists a residual labour market state in the data entitled "other outside employment"; observations under this code may also contain some of the retirement benefit recipients, as the share of older individuals included in there has been rising before 1996.

Earlier papers working with retirement definitions in Denmark, such as Bingley and Lanot (2004, 2007), use a definition of "retirement state" based on the fulfillment of 3 conditions for a yearly individual observation: zero mandatory pension contributions which are only made in-work; below 20% median labour earnings per year; no attachment to an employer. However, with this type of definition they include the elderly unemployed in the retired category; for our purpose here it is crucial to make the distinction between unemployment and (early) retirement (as destination-state following a separation from the previous employer). Our (re)defining strategy is therefore to use the information on the existent labour market states in the data, and to augment that with the eligibility requirement for (early) retirement in each particular year. Fortunately, misclassification is mainly relevant for the cross-sections 1980 to 1996 and particularly for recipients of the old-age pension, the *folkepension,* category which we do not use per se in the empirical analysis. Miscoding of the "early retirement" benefit recipients is less concerning, both during the pre-1996 and the post-1996 time windows; similarly, the employment labour market state is the most unlikely to have been recorded wrongly.

Finally, directly relevant for the empirical exercises in this paper, in the number of cases where we observe inconsistencies that cannot be corrected based on our (re)definition strategy above, we proceed as follows: i) if "employed, non-eligible for early retirement" individuals are observed to have separated to "(early) retirement", we assume the separation was in fact to "unemployment"; ii) if "employed, eligible for early retirement" individuals are observed to have separated to "any other labour market state than early retirement", we assume the actual state they separated to was "early retirement"[26].

[26] Since our data has annual frequency, workers are observed employed in year t and in the new state in year $t+1$, with the separation from the firm de facto happening in between these observation times. This raises the question of when to assign the eligibility for early retirement, whether before or after the separation event. We deal with this in detail in the empirical subsection on robustness.

5.4 Descriptive statistics

We provide in Table 2 below summary statistics of the variables of interest for the entire population of (former and actual) workers in firms with at least 50 employees in each year of observation. We present these descriptives both for the entire pooled sample, 1980-2001, and for a single cross-sectional year, 2000.

Notice first that the reported means of real wage, tenure, potential experience, education, and age, are very similar in the pooled sample and in the cross-section for 2000, which is comforting in the sense that not many changes in the main individual demographics and job characteristics appear on average to happen throughout the observation period. Caution is needed in interpreting the differences between some of the reported labour market state *shares*: recall that our IDA sample is constructed by first selecting all firms ever active in our observation sample, with >50 employees in any single year of their activity, and subsequently by appending the entire observed time-series record of any individual observed to have *ever worked* in any of those firms. Hence, not very surprisingly, the share of employed people is almost identical in the pooled and in the year-2000 samples. There are some differences in the shares of unemployed, old-age pensioners, and other-out-of-work, which is potentially due also to some misclassification of the latter two categories in the data, as detailed in the earlier subsection on misclassification; the proportion of people on early retirement is however less dissimilar among the pooled and cross-section samples summarized Table 2, which is once more reassuring.

The sub-population of direct interest for our empirical analysis comprises the workers aged $60-66$ years old, employed in the firms with a minimum of 50 employees. Figure **??** depicts the evolution of the aggregate stock of workers in the above-mentioned age range, with three separate plots for: employed overall, employed & eligible-for-early-retirement, and respectively, employed and non-eligible-for-early-retirement individuals, between 1980 and 2001. The stock of employed individuals aged 60-66 eligible for early retirement starts with the high take-up level in the early 80s, decreases [27] till the mid 90s, and stays relatively constant since then, while the stock of non-eligible older workers increases throughout the observation period (except for a small window in the early 90s),

[27]In theory this descending trend might be partly due to our selection of the firm sample, e.g., some older firms may fall out over time of the working sample (firm with min. 50 employees each year), while newer large firms, with younger employees, may be entering that working sample. This is not confirmed to be a practical issue in the data. Moreover, qualitatively these trends look very similar if instead we retain in the working sample all observations on firms that had min. 50 employees *at least once* over 1980-2001.

Table 2: Descriptive Statistics

variable	Pooled 1980-2001	Year 2000
age	39.44 (14.60)	42.33 (15.10)
education years	11.62 (3.17)	12.22 (3.03)
tenure at the firm	6.17 (6.39)	5.77 (6.64)
potential experience	21.82 (15.35)	24.11 (15.77)
firm size	2795.86 (5867.42)	3681.10 (7272.40)
real hourly wage (2000, Kr)	162.20 (60.94)	174.16 (88.10)
observations	16000834	835955
individuals	1877651	835955
firms	3374	1626
firm-worker spells	3075357	588306
percentage employed	73	74
percentage unemployed	10	6
percentage early retired	4	6
percentage old retired	5	7
percentage "other out of work"	8	3

The sample consists of all firms with at least 50 employees in each year of their presence in the data, and of all individuals who ever worked for these firms, observed throughout the entire sample period 1980-2001. The job-related variables (tenure, hourly wage) are averaged over the subpopulation of workers in the time period considered. Individual demographics (age, education, potential experience) are averaged over all individuals present in the sample in the time period considered. (St. dev. in parentheses under the means.)

roughly in line with the official tightening of the eligibility rules for access to this pension scheme, see Table 1. In particular, a consequence of these trends is that the aggregate fraction of eligible to non-eligible 60 to 66 years old employees changes from around 2 : 1 in the early 80s, to 1 : 1 in the early 90s, and reverses to less than 1 : 2 in the early 2000s.

We next quantify the number of major workforce reduction (mass layoff) events within our sample of large firms[28]. We report in Table 3 the number of massive downsizing episodes for the entire pooled sample (with first downsizing events happening between years 1980 and 1981, and last downsizing events between 2000-2001), as well as, se-

[28] Remark that we could have more than one major downsizing event per firm, in different observation years. In practice however, this occurs rarely, and hence we do not/ cannot treat such multiple-downsizing-per-firm events differently than unique downsizing of different firms.

lectively, for annual downsizing episodes occurring every 4-years within our observation period. Since we are not interested in firms completely ceasing operation, we set the maximum downward adjustment at 80% of the firm's employment size in the year preceding the downsizing event[29]; the minimum downward adjustments are 20%, 30%, and respectively 50% of the firm's previous size, with corresponding number of events reported in separate Table columns [30]. We also want to focus solely on downsizing events taking place over a single year—since our data has annual frequency, we can only be sure that all displacements took place somewhere between end November of successive years—and not over a longer time period, thus preventing as much as possible confounding with other potential types of firm re-organization, for instance due to the anticipation of future mass layoffs.

Table 3: Number of major downsizing events in firms with min. 50 employees each year

Time period	min-to-max % annual downsizing		
	20-to-80%	30-to-80%	50-to-80%
Pooled all years	3301	1281	280
2000-2001	187	73	12
1996-1997	132	51	10
1992-1993	195	82	17
1988-1989	160	67	14
1984-1985	118	48	11
1980-1981	233	85	16

Even at first glance, the numbers in Table 3 are suggestive of the recessions in years 2000, 1992 and the early 1980's, as we count a larger number of major downsizing events in those periods. Typically for the empirical literature on mass layoffs, the firm size adjustment considered is at least 30% of the firm size in the previous year, e.g., Jacobson et al (1993), Albaek et al (2002); we report for all our empirical analyses also results obtained using the other two columns of downsizing episodes mentioned in Table 3.

[29] Note that this strategy allows for complete closure of some of the firm's plants, provided the firm does not entirely shut down, overall still keeping at least 20% of its employees from the preceding year.

[30] Employment adjustment is computed here as each firm's *net employment change* from one year to the next. In practice, for this kind of mass downsizing for larger firms, net employment change is essentially identical to the *gross outflow* (very few new workers being hired during the simultaneous mass layoff interval). To eliminate any remaining worry, in the empirical analyses at individual and firm level we expressly use the computed gross outflows of displaced workers.

6 Empirical framework and estimation

6.1 Testable empirical hypotheses

A first implication of our model, see Proposition 1, Section 3, is more general, following from the Brownian nature of the stochastic process assumed for the evolution of the exogenous log product-demand index, which via the model carries over almost perfectly as the process governing the firm employment size evolution over time. This regularity is known as Gibrat's law for firm size and is empirically testable: in the discrete-time setting of our data, it states that log employment size follows roughly (i.e. with eventual small deviations) a random walk over time. We confirm this prediction using various empirical strategies; in order to save space, and since this regularity is of secondary importance in the context of our model, we fully relocate its verification to Appendix A.

The main implication of our model, see Proposition 2 in the theory section, is that the optimizing strategy of downsizing firms is to fire their employees eligible for early retirement with predilection, relative to any of their non-eligible peers, as formally stated captured in (11). This can be phrased as stating that any massive employee outflows (i.e., mass layoffs) will contain either *only eligibles* (if the outflow size is smaller than or equal to the number of eligible employees) or *all the eligibles, besides some non-eligibles* (if the outflow size is higher or equal to the number of eligible employees). Obviously, in the data this implication cannot hold *literally*: on the one hand, there are multiple dimensions of worker heterogeneity that the firm might want to include as constraint in its employment adjustment decision, including foremost the usually unobserved-by-the-econometrician individual ability; on the other hand, there could also simply be instances of imperfectly implementing the optimal employment adjustment strategy or of wrong data coding (see, e.g., the data subsection on misclassification of labour market status). Hence, what we seek to validate empirically is a weaker, though still sharp, variant of the direct model implication from Proposition 2, namely: *downsizing firms are on average more likely to lay off their early-retirement eligible rather than ineligible peers, within the age-category targeted by the early retirement policy.* The peers within the targeted age category by the early retirement policy, ie. individuals aged 60 to 66, are the suitable control group for our empirical tests[31].

[31] This also implies that we will conduct all our analysis on firms which have at least one eligible worker for early retirement (i.e. at least one lower-educated eligible worker, respectively at least one higher-educated eligible worker, for the extension analyses that allow for heterogeneity by low/high education level, see below in the empirical section).

The above empirical formulation of Proposition 2 can be further spelled out in two distinct empirical hypotheses, one at the worker and the other at the aggregate firm level:

i) at the worker level: ceteris paribus, the likelihood of a worker being in the displaced pool of workers ages $60 - 66$ increases with her eligibility to the early retirement scheme;

ii) at the firm level: ceteris paribus, the fraction of displaced workers among employees aged $60 - 66$ is increasing in the fraction of eligible workers to retire early in that age category.

Important, we test hypotheses i) and ii) on the universe of mass layoff events taking place within our time interval, as discussed in the data section, c.f. Table 3. This ensures that we (almost surely) deal with layoffs as opposed to quits, and hence that a worker's separation into early retirement is non-voluntary. Moreover, both empirical analyses are performed over the whole observation period in the data, implicitly accounting for the changes in early retirement eligibility rules over time, which impact accordingly the aggregate number of eligibles for early retirement. The next subsections formalize and discuss in detail the empirical specifications used for checking i) and ii).

6.2 Estimation methodology

6.2.1 Older worker exit hazard conditional on early retirement eligibility

Our first empirical strategy is based on estimating the worker's probability of separation from the firm, conditional on her eligibility status for retirement, cf. point i) in the previous subsection. This estimation is performed for our working sample of firms with at least 50 employees in each year, on the "target" worker subpopulation aged 60 to 66, namely individuals who could be eligible for early retirement benefit, should they fulfill the UI Fund membership criteria. We provide an analysis using separations that are part of the three sets of mass layoff events, over the entire time period 1980-2001, described in Table 3.

Denote by R_{ijt} the eligibility status into early retirement of worker i, working in firm j, at time t, with $R_{ij} = 1$ (0) if the worker is eligible or not for early retirement, c.f. the specific pension rules at the time, see above in Table 1. Let $exit_{ijL} \in \{0, 1\}$ be the observed outcome of exit for worker i from firm j via a mass layoff at time L (=belonging to the set of downsizing episodes pooled over the observation period, described in Table 3). The individual hazard of exiting from the firm, conditional on R_{ijt}, controls X_{ijt}, and on the worker separation being part of a mass layoff of the firm, i.e. at time $t = L$:

$$\Pr[exit_{ijL} = 1 | R_{ijL},\ X_{ijL},\ t = L] = \Lambda(\alpha + \beta R_{ijL} + \gamma X_{ijL}) \tag{12}$$

X_{ijt} includes observed characteristics of the worker and/or the firm—in our reported specification these are the years of education, potential experience, gender, tenure at the firm, hourly wage, full sets of dummy variables for the worker's occupation, the firm's location, and the firm's industry; in addition we include in vector X_{ijt} a full set of time indicators. α is a constant. We use a logit specification[32] for (12), hence $\Lambda(.)$ is the logistic CDF. We compute heteroskedascity-robust standard errors, in addition correcting them for clustering at firm level j and downsizing time L. Estimation of β in equation (12) is consistent if regressor R_{ijL} is not endogenous: below we discuss this assumption in detail. Validation of our hypothesis means verifying that $\widehat{\beta}$ is positive.

The residual term from the estimation of 12 is likely to be correlated to omitted variables in the covariate vector X_{ijL}, characterizing unobserved worker heterogeneity, unobserved firm heterogeneity, or unobserved worker-firm match heterogeneity; a usual problem in simple cross-sectional regression. Since we are only interested in consistently estimating β, the question narrows down to whether early-retirement eligibility R_{ijL} is correlated with any such omitted variables. We know precisely how eligibility for early retirement R_{ijL} is assigned, from the institutional background section above; the target age 60 to 66 is the first determinant. Since we estimate (12) solely for the workers in the targeted age category, 60 to 66, the targeted age-category is kept constant [33]. The other determinant of R_{ijL} is having contributed to a UI Fund for a certain amount of years, depending on the precise rule in effect at the time, cf. Table 1. UI Fund contribution record certainly depends on the time spent in the firm/ labour market, but both tenure and potential experience are included in vector X_{ijL}. It is also difficult to justify correlation between R_{ijL} and worker unobserved characteristics like general ability, or worker-firm unobservables like match-specific quality, which would not work via years of education, potential experience, or tenure. There are only two additional possibilities that require separate discussion.

First, the *worker's (innate) preference for early retirement* would be the most plausible

[32] All our results are qualitatively robust to using instead a probit, or a simple linear probability specification.

[33] We can nevertheless control for individual age level, if we exclude potential experience or years of education; that does have any impact on the qualitative interpretation of our results.

candidate for correlation with R_{ijL}, as an individual who desires to retire early is likely to contribute timely and adequately to a UI Fund [34]; however, given that we estimate (12) only for individual exits that are part of simultaneous mass employee layoffs, the worker's preference for separation is less likely to play a role in the decision to separate. Nevertheless, to alleviate any worries, we perform the following robustness experiment: we discard from the analysis the individuals who become eligible for early retirement exactly in the downsizing year L [35]; we check that our results remain qualitatively unchanged.

Second, it is hard to find reasons for some *firm-specific omitted variables* to be correlated with the individual early retirement eligibility R_{ijL}, unless one is willing to explore the possibility of employee early eligibility status R_{ijL} being endogenous in the firm mass downsizing events, thus implicitly refuting our model's key assumption that mass layoffs are consequences of exogenous shocks to the product demand. This possibility is in fact relevant also for the analogous discussion in the case of the the firm-level analysis in the next subsection, since its implication is that there is a higher proportion of early retirement eligibles among workers aged 60 to 66 in downsizing firms than in firms that do not downsize (with the stronger alternative: on average that proportion of eligibles is higher for firms in downsizing years, than in other years). We tackle this argument in the next subsection, describing our econometric specification at the firm-level.

6.2.2 Older workers exit share conditional on eligible share to retire early

The second strategy to test the main prediction of our model is founded more squarely in the firm's perspective, see point ii) in the above subsection on testable empirical hypotheses. We adapt a simple firm-level turnover analysis approach often used in studies on organisational /technological innovation impacts on (changes in) the structure of a firm's workforce, e.g., Aubert et al (2006). In our context, the "technology" available to the firm is the share of workers eligible for early retirement, cheaper from the perspective of the employment adjustment costs.

For each downsizing event, consistent with the notation used in the theory section c.f. (11), denote by N^{exit}_{60-66} the gross outflow of workers in the 60-66 age category (the age group targeted by early-retirement policies), and by $N_{R=1}$ the number of employees

[34] Remark however that these variables cannot be perfectly correlated, given the (presumably) exogenous changes in the early retirement eligibility rules described in Table 1.

[35] This typically includes the 60 years old among the workers in the age interval 60-66, but also older people in that age range, who just qualified for early retirement given enough years of UI Fund contribution, as per the *efterløn* rule in effect at the time.

eligible for early retirement, before downsizing. We write then the following conditional expectation for $\frac{N^{\text{exit}}_{60-66}}{N_{60-66}}$, the share of displaced workers among the workers aged 60-66, in terms of $\frac{N_{R=1}}{N_{60-66}}$, the share of early-retirement eligible employees aged 60-66, for each firm downsizing episode:

$$\mathrm{E}[\frac{N^{\text{exit}}_{60-66}}{N_{60-66}}|\frac{N_{R=1}}{N_{60-66}},\ X] = G(\alpha + \beta\frac{N_{R=1}}{N_{60-66}} + \gamma X) \tag{13}$$

where we omit the indices for layoff-time L and firm j for ease of exposition; α is a constant, and X includes firm characteristics such as industry, location and firm size [36]; moreover, we add in X a full set of year indicators. The coefficient β is identified from the variation in the share $N_{R=1}/N_{60-66}$ across firms and over time, changes over time capturing those due to the reforms in the early retirement eligibility rules, cf. Table 1. Given that $\frac{N^{\text{exit}}_{60-66}}{N_{60-66}}$ is bounded between 0 and 1, we estimate the general linear model in (13) by a Bernoulli quasi-likelihood method, described in detail in, e.g., Papke and Wooldridge (1996)[37]; the link function $G(.)$ in this case is the logistic CDF, $\Lambda(.)$. We compute heteroskedasticity-robust standard errors to allow for possible model misspecification, in addition correcting them for possible correlation (clustering) by year.

Similar to the individual worker analysis, we estimate (13), for each of the three sets of massive downsizing events, see Table 3. For consistent estimation of β we need $\frac{N_{R=1}}{N_{60-66}}$ to be exogenous; we discuss this assumption below. Parameter β is predicted to be positive in the context of our model.

The concerns for correlation between the residual term of (13) and $\frac{N_{R=1}}{N_{60-66}}$ are to some extent similar to the individual-hazard specification from (12) above; in this case the worry is that a firm-level unobserved component of that residual term would be correlated with $\frac{N_{R=1}}{N_{60-66}}$. The only possibility is the same as the last one discussed above for the individual-level hazard analysis, namely that the mass downsizing episodes are not exogenous, but connected to the proportion $\frac{N_{R=1}}{N_{60-66}}$of early retirement eligible employees in the firm, at the downsizing year L. To eliminate this worry, we verify that $\frac{N_{R=1}}{N_{60-66}}$ does not systematically vary between the downsizing year L versus other years t in firms that downsize, or on average between downsizing firms and the firms that do not downsize.

[36] We use several other specifications, including employee aggregate characteristics like share of females, average tenure, average wage etc.; the qualitative results do not change.

[37] A simple OLS exploratory analysis gives the same qualitative results.

6.3 Estimation results

6.3.1 Individual-level analysis

Before we proceed with the main estimation of (12) on the age category targeted by early retirement policies, we present in Appendix B descriptive relative layoff probabilities for all workers, by four age categories; inter alia, we confirm the nonlinear age-pattern found earlier in the literature, with younger workers and employees in the (early) retirement ranges having a higher layoff probability than full-time working-age workers, even conditional on tenure at the firm, and a host of other observables.

The estimation results of (12) above are displayed for selected covariates in Table 4 below.

Table 4: Individual layoff hazard, conditional on eligibility for early retirement

	Min-max downsizing as % of previous firmsize		
	20%-80%	30%-80%	50%-80%
eligible	.105	.127	.036
	(.060)*	(.078)	(.127)
years of education	-.007	-.014	-.019
	(.015)	(.020)	(.040)
male	-.141	-.137	-.329
	(.069)**	(.093)	(.178)*
potential exp	-.002	-.011	-.021
	(.011)	(.016)	(.032)
hourly wage	.002	.002	.0006
	(.0004)***	(.0004)***	(.0007)
tenure	-.043	-.038	-.006
	(.006)***	(.008)***	(.013)
firm size	.00003	1.00e-05	.00007
	(6.46e-06)***	(8.87e-06)	(.00006)
cons	-.868	.147	.454
	(.655)	(.899)	(1.904)
N. obs.	16107	7872	1931
Log likelihood	-9483.37	-5057.37	-1195.70
Pseudo R^2	0.0394	0.0403	0.1019

The sample contains workers 60 to 66 years old, in firms with at least 50 employees in each year of their observation period, at the time of a mass downsizing episode. These are episodes of the minimum downsizing magnitude described in each column. All regressions control also for a full set of occupation, industry, regional and time indicators. (Std. errors in parentheses under the estimated coefficients are heteroskedasticity-robust and clustered per firm-year.)

Table 4 confirms partly our prediction for the positive impact of the worker eligibility on early retirement on her layoff probability, with a positive and statistically significant coefficient in the first column, which includes all downsizing events of magnitudes 20% to 80% of previous firm size, over 1980-2001; the estimation results obtained on the subsets of downsizing episodes from the second and third column have a positive sign, but are not statistically different from 0 at conventional statistical levels. In addition to the covariates reported in the table, we control for a full set of indicators for firm region and industry, worker occupations, and time dummy variables. We correct the estimated standard errors for possible heteroskedasticity and correlation within firm-year clusters, as explained also above. The result is qualitatively robust to inclusion of higher order polynomials in tenure, experience, firm size, and to exclusion of any subset of worker and firm observables, see more details in the robustness subsection below.

6.3.2 Firm-level analysis

Table 5 shows firm-level estimates of exit fraction among workers aged 60 to 66 on their share of eligibles for early retirement, c.f. (13) from above. Not reported in the table, all regressions control also for a full set of indicators for region, industries, and time. As stated in the empirical framework, we estimate these regressions by quasi-maximum likelihood (QMLE), allowing for robust and clustered-per-year standard errors. The estimated coefficient is positive for the second column, but not statistically significant for the other two. Including a host of other firm-level aggregate worker measures like average tenure/education/wage or proportion of females/managers etc. in the firm etc.—see the subsection on robustness below for more details—does not affect this finding, as the coefficient of interest in the first column remains positive and statistically significant, while the other columns give a result that is not statistically different from null.

In the next subsection, motivated by the ambiguous finding concerning the statistical power of our estimated coefficient in both the individual-level and the firm-level regressions, we extend our approach allowing for firms to differ in their layoff behavior towards workers in higher- vs. lower- educated categories.

6.3.3 Heterogenous layoff behavior by worker education level

Our model assumed all workers to be identical for the firm apart for their eligibility for early retirement. Although we have allowed for a large range of covariates, including educational years, we did not so far consider the situation where the firm has different layoff strategy for different categories of employees. Here, we re-estimate (12) and (13)

Table 5: Firm-level exit fraction on share of eligibles, in 60-66 yrs old

	Min-max downsizing as % of previous firmsize		
	(20%-80%)	(30%-80%)	(50%-80%)
share eligibles	.304	.505	.399
	(.213)	(.255)**	(.529)
fsize	.00003	.00004	.0003
	(.00002)	(.00003)	(.0001)**
cons	-.860	-.362	-.100
	(.179)***	(.223)	(.453)
N. obs.	2188	884	204
Log likelihood	-992.89	-441.57	-101.52

The sample contains firm-year level aggregate variables. The dependent variable is the fraction of 60-to-66 years old workers laid off and the independent variable is the share of eligibles in the 60-66 years old group. All firms have at least 50 employees in each year of their observation period and they are sampled at the time of a mass downsizing episode. These are episodes of the minimum downsizing type described in each column. All regressions control also for a full set of industry, regional and time indicators. (Std. errors in parentheses under the estimated coefficients are heteroskedasticity-robust and clustered per year.)

from above, by higher-educated (i.e., individuals with more than 12 years of education) and respectively lower-educated (i.e., individuals with less than 12 years of education) worker categories.

The following 4 tables summarize the estimated coefficients. Tables 6 and 7 correspond to the higher-educated and respectively lower-educated estimates of the individual worker layoff probability in (12). Tables 8 and 9 are the results for the high-educated and respectively the lower-educated estimations of the firm-level regression from (13).

Both individual and firm-level estimations show a consistent and remarkable difference between the high-educated and low-educated workers in terms of the impact of the eligibility to retire early on their layoff hazard. While the part of the firm's workforce with less years of education confirms the prediction of our model both at individual worker and at firm level, i.e. Tables 7 and 9—with the coefficients of interest becoming larger and more statistically significant[38] than in the case of the whole firm—for the category of the more educated employees the implication of our model does not have a bite at all: the eligibility coefficient estimates become completely insignificant, for all 3 columns, in Table 6 and similarly for the firm-level results in Table 8. These results are robust qualitatively

[38] This is the case only for the first two columns for the firm-level analysis; the estimation from third column has little statistical power given the very low number of observations.

Table 6: High-educated individual layoff hazard, conditional on eligibility for early retirement

	Min-max downsizing as % of previous firmsize		
	(20%-80%)	(30%-80%)	(50%-80%)
high-educated eligible	-.002	.022	-.231
	(.074)	(.096)	(.163)
years of education	.068	.111	.043
	(.049)	(.065)*	(.095)
male	.025	.058	-.140
	(.098)	(.123)	(.270)
potential exp	-.004	.004	-.004
	(.017)	(.024)	(.047)
hourly wage	.001	.001	-.0002
	(.0005)**	(.0004)***	(.0008)
tenure	-.050	-.042	.005
	(.007)***	(.010)***	(.013)
firm size	.00002	9.98e-06	.00002
	(7.04e-06)***	(1.00e-05)	(.0001)
cons	-1.956	-1.974	-.609
	(1.304)	(1.887)	(3.174)
N. obs	7576	3753	976
Log likelihood	-4296.87	-2350.91	-603.95
Pseudo R^2	0.0472	0.0507	0.1072

The sample contains high-educated workers 60 to 66 years old, in firms with at least 50 employees in each year of their observation period, at the time of a mass downsizing episode. These are episodes of the minimum downsizing magnitude described in each column. All regressions control also for a full set of occupation, industry, regional and time indicators. (Std. errors in parentheses under the estimated coefficients are heteroskedasticity-robust and clustered per firm-year.)

Table 7: Low-educated individual layoff hazard, conditional on eligibility for early retirement

	Min-max downsizing as % of previous firmsize		
	(20%-80%)	(30%-80%)	(50%-80%)
low-educated eligible	.261	.280	.378
	(.079)***	(.106)***	(.204)*
years of education	.015	-.013	-.013
	(.041)	(.050)	(.128)
male	-.286	-.315	-.439
	(.093)***	(.141)**	(.247)*
potential exp	-.006	-.028	-.066
	(.015)	(.019)	(.044)
hourly wage	.002	.002	.003
	(.0006)***	(.0007)***	(.001)**
tenure	-.034	-.028	-.010
	(.006)***	(.009)***	(.024)
firm size	.00003	1.00e-05	.00007
	(7.19e-06)***	(8.10e-06)	(.00007)
cons	-1.046	.906	2.233
	(.903)	(1.165)	(2.680)
N. obs	8060	3879	862
Log likelihood	-4849.76	-2523.23	-515.76
Pseudo R^2	0.0398	0.0386	0.1223

The sample contains low-educated workers 60 to 66 years old, in firms with at least 50 employees in each year of their observation period, at the time of a mass downsizing episode. These are episodes of the minimum downsizing magnitude described in each column. All regressions control also for a full set of occupation, industry, regional and time indicators. (Std. errors in parentheses under the estimated coefficients are heteroskedasticity-robust and clustered per firm-year.)

Table 8: Firm-level high-educated exit fraction on share of eligibles, ages 60-66

	Min-max downsizing as % of previous firmsize		
	(20%-80%)	(30%-80%)	(50%-80%)
share high-educ eligible	-.298	.493	-.501
	(.300)	(.306)	(.856)
fsize	.00005	.00004	.0005
	(.00002)**	(.00003)	(.0002)***
cons	-.615	.272	-1.066
	(.190)***	(.282)	(1.137)
N. obs	1493	633	160
Log likelihood	-785.63	-349.71	-71.59

The sample contains firm-year level aggregate variables. The dependent variable is the fraction of 60-to-66 years old high-educated workers laid off and the independent variable is the share of eligibles in the 60-66 years old high-educated group. All firms have at least 50 employees in each year of their observation period and they are sampled at the time of a mass downsizing episode. These are episodes of the minimum downsizing type described in each column. All regressions control also for a full set of industry, regional and time indicators. (Std. errors in parentheses under the estimated coefficients are heteroskedasticity-robust and clustered per year.)

Table 9: Firm-level low-educated exit fraction on share of eligibles, ages 60-66

	Min-max downsizing as % of previous firmsize		
	(20%-80%)	(30%-80%)	(50%-80%)
share low-educ eligible	.379	.546	.479
	(.179)*	(.246)**	(.612)
fsize	.00006	.00005	.0002
	(.00003)**	(.00004)	(.0001)*
cons	-.990	-.439	-.332
	(.172)***	(.215)**	(.537)
N. obs	1621	667	158
Log likelihood	-682.01	-324.47	-76.40

The sample contains firm-year level aggregate variables. The dependent variable is the fraction of 60-to-66 years old low-educated workers laid off and the independent variable is the share of eligibles in the 60-66 years old low-educated group. All firms have at least 50 employees in each year of their observation period and they are sampled at the time of a mass downsizing episode. These are episodes of the minimum downsizing type described in each column. All regressions control also for a full set of industry, regional and time indicators. (Std. errors in parentheses under the estimated coefficients are heteroskedasticity-robust and clustered per year.)

to excluding any of the current control variables or adding others, for instance controlling for several firm-level aggregate employee measures, see also the following section on robustness.

For the individual-level analyses, a concise way of summarizing the key differences between the results in Tables 4, 6, and 7 is to report the respective marginal effects of individual eligibility/ineligibility to early retirement on the probability of being displaced, keeping the sample mean of all other covariates fixed. Table 10 collects and compares these quantities. The upper and middle horizontal panels, corresponding to estimates reported for the entire old-age sample in Tables 4 and respectively for the sample of higher educated workers in Table 6 show that the marginal effects of being eligible, $R = 1$, on the layoff hazard are very similar (upper panel), or even lower (middle panel) than the marginal effect on the layoff hazard if $R = 0$. On the contrary, the marginal effect of $R = 1$ on the layoff hazard are substantially higher than the marginal effect of $R = 0$ for the case of the lower-educated worker category, see bottom panel of Table 10. This is a different, but equivalent, way of stating that our main empirical prediction—see Proposition 2 in the theory section and its empirical formulation in the empirical framework section—is validated for the sample of lower-educated workers.

How do we explain the different results obtained between low- and high-educated worker categories? Our simple demand-side-driven framework assumes quasi-homogenous workers in a firm. However, in reality, firms are composed of potentially very different worker categories (with their existence rationalized, for instance, as due to production complementarities). Examples of such categories could be those imperfectly proxied by the number of years of education below or above 12 years, like here. On the one hand, lower-educated, blue-collar, production workers are typically more homogenous from the perspective of the firm than their higher educated peers, their work tasks being quite well defined within occupation-skill cells; hence, the layoff cost function in their case could be plausibly approximated by a simple function of their early retirement eligibility. As for their higher-educated peers, this category is likely to be much more heterogenous still, with many of them performing multiple, dynamic, job tasks, and thus would need to be further narrowed down to isolate the proper peer comparison subgroups. On the other hand, and more critically, negative shocks to the firm product demand level can have differential effects on different worker categories: for instance, if the demand shocks is related to a technological innovation we know from the vast literature on skill-biased-technological-change that lower-educated, low-skilled workers would be asymmetrically, negatively af-

Table 10: Marginal effects of (in)eligibility to retire early on individual layoff hazard

	Min-max downsizing as % of previous firmsize		
	(20%-80%)	(30%-80%)	(50%-80%)
Individual higher and lower-educated, c.f. Table 4			
(in)eligibility			
R=0	.280	.369	.543
	(.011)	(.016)	(.026)
R=1	.301	.399	.552
	(.008)	(.012)	(.021)
Individual higher-educated only, c.f. Table 6			
(in)eligibility			
R=0	.269	.360	.539
	(.014)	(.021)	(.031)
R=1	.268	.365	.482
	(.010)	(.015)	(.028)
Individual lower-educated only, c.f. Table 7			
(in)eligibility			
R= 0	.276	.366	.529
	(.012)	(.017)	(.041)
R=1	.331	.433	.621
	(.010)	(.016)	(.023)

Marginal effects of (in)eligibility to early retirement are computed keeping all other covariates fixed at their sample mean, based on the logit estimates in Tables 4, 6, and 7. (Std. errors in parentheses are estimated using the Delta method.)

fected, relative to their higher-skilled colleague. Both these types of arguments suggest that an extension of our firm-level model to narrow within-firm employee categories, allowing for potential asymmetric turnover responses to product demand shocks at the firm level, is a good candidate for rationalizing the empirical discrepancies in layoff behavior found towards lower- versus higher-educated co-workers in Denmark.

6.3.4 Robustness

We perform a series of additional empirical exercises meant to test to robustness of our findings above[39].

The first robustness exercises were already mentioned as checks of the exogeneity assumption of the regressors of interest in the individual level and firm level empirical specifications from (12) and respectively (13). For the individual-level analysis, we re-estimate (12) using only workers eligible to early retire at least a year before the mass layoff periods, hence only individuals that could have early retired earlier, but did not. The qualitative interpretations of the results remains the same throughout the analogue of Tables 4, 6, 7 and 10. Furthermore, relevant for both the individual-level and firm-level specifications, we check that the share of eligibles to retire early among workers aged 60 to 66, c.f. (13), is not systematically higher in firms that downsize, at downsizing years, than in other years at the same firm, or on average at firms that do not downsize in our data.

As briefly mentioned in the section concerning assigning individual eligibility for early retirement, this eligibility can be computed either at November 30 of the last year the worker is observed employed in the firm, or the November 30 of the subsequent year, after the mass layoff, when he is observed in the new state, which can be employed at the same firm, or in another labour market state like early retirement, unemployment, working for another firm, etc. The estimates reported in this paper are using the definition of mass layoff a priori eligibility. However, both individual and firm-level results are robust qualitatively and even very similar in terms estimated magnitudes if we assign early retirement eligibility at the observation time subsequent to the mass layoff event. There are drawbacks and advantages for both definitions—however, in general, the more strict definition is the one assuming that the worker had to be eligible for at the observation time preceding the mass layoff—a convincing empirical strategy is to check using both definitions.

[39] All robustness estimates are available on request from the authors.

One concern for both individual and firm-level analyses is that the results are influenced by turnover at very large firms, hence the concern being outliers in the firm size distribution. Indeed in the data we have firms with downsizing episodes included in Table 3, larger at the time of downsizing than, e.g., $30,000$ or $50,000$ employees. To address this concern, we repeat all our empirical analyses by discarding the top 1% of the firm size distribution over our whole period (this eliminates all such outliers): both individual and firm-level results remain qualitatively the same and in several cases they are virtually identical quantitatively as well.

Finally, already mentioned throughout the empirical sections, we perform several robustness checks in terms of the set of control variables in both (12) and (13)[40]. Including higher order polynomials in tenure, experience and firm size in (12) does not alter any of our qualitative results. In the same individual-level specification, including or excluding any of the current covariates, for instance the hourly wage if, e.g., despite these being layoffs, we are concerned about endogeneity of wages in the separation decisions,does not affect either our qualitative results. Including several interaction terms between variables in (12), such as between tenure and hourly wages does not matter for the interpretation of the β coefficient. Now, for the firm-level analysis in (13): including a relevant set of firm-level employee aggregate measures (computed for all employees of the firm at the downsizing year L) such as percentage of female workers, percentage of managers, average tenure, average potential experience, average hourly wage level does not make any difference for the qualitative interpretation of our coefficient of interest, β. Using polynomials or allowing non-linearity of firm size by using firm size category indicators in (13) does not impact either the qualitative findings.

7 Summary and discussion

This paper has proposed to bridge the labour demand and retirement literatures, studying whether existing age-related public policy can give rise to demand-driven incentives in employment adjustment outcomes. Our partial equilibrium framework with stochastic product demand and dismissal costs varying in the workers' eligibility for publicly financed early retirement has as core prediction that distressed firms will lay off with predilection

[40] We have already mentioned that the results are qualitatively robust using logit (reported in the paper), probit, or a simple OLS specification for the individual-level specification in (12); similarly, we have mentioned that results are qualitatively robust for the firm-level specification at (13), using a general linear QMLE Bernoulli model with a logit link function as reported, as well as a simple OLS.

those employees eligible to retire early. We have justified in detail the identification and estimation assumptions of the corresponding empirical hypotheses; we have tested them on the entire set of mass layoff events in larger Danish firms for 1980-2001, identified from an exhaustive register linked-employer-employee data. Using both individual and firm-level analyses, we have unambiguously shown that in Denmark firms push their eligible lower-educated employees into early retirement through mass layoffs; in the case of their higher-educated peers we found no conclusive evidence of this phenomenon. We have further performed a number of critical robustness checks to confirm the stability of our empirical results. We have finally argued that the differential firm layoff behavior by worker education category can be rationalized by extending our firm-level model to a model of narrower within-firm employee groups, potentially allowed to have asymmetric employment adjustment responses to shocks at the firm product demand level.

A couple of earlier empirical studies mentioned above in the literature overview turn out to partially support the key implication of our model, with data of different nature and/or from other countries, and using different empirical methodologies.

Hakola and Uusitalo (2005) provide direct empirical evidence consistent with the key implication of our model, using a Finnish pension reform in year 2000, affecting the employer's unemployment-related contributions for elderly workers; these are effectively early retirement benefits for employees over 60 who were laid off (so-called "unemployment tunnel" to retirement). As a consequence of the reform, the partially-private, experience-rated unemployment pensions became substantially costlier for employers with larger firm size. The authors convincingly show that this change caused the number of dismissals of elderly employees to decrease at large firms, thus checking our model's main prediction on linked employer-employee data from Finland.

Mentioned also in the theory section, Pfann (2006) investigates a massive layoff event at a Dutch firm in demise, by means of a stochastic labour demand model from the same family as ours, providing first indirect empirical support for our model's key assumption of heterogenous firing costs increasing in employee early retirement eligibility. Namely, he shows, constructing direct proxies for the idiosyncratic firing costs, that the age-profile of firing costs sharply drops around early retirement. Corroborating this findings, he then computes that the layoff probability of employees around the official early retirement age is indeed substantially higher than that of other age-category employees, indirectly supporting the main implication of our paper with personnel data of a large company in the Netherlands.

Despite this study having completely different target and interpretation of results, the conclusions of Tatsiramos (2010) indirectly support the mechanism we present in this paper, providing further cross-country empirical evidence in favor of our main prediction. Specifically, he finds that older workers who were displaced in Germany and Spain, are more likely to (early) retire after 60, relative to their age peers who left employment voluntarily [41]. Although these results are given a labour supply interpretation in Tatsiramos (2010), they are implicitly tests of our implication, akin to the individual-level test variant we have performed in the empirical section above. Interestingly, this author also checks that the implication fails to hold in the other two countries studied in his paper, Italy and UK. The contextual difference between the first and the second pair of countries is precisely the presence of specific publicly funded early retirement benefit (and/or other age-targeted unemployment provisions) for displaced elderly workers. As these institutions exist for Germany and Spain, firms can use them as effective subsidy to force eligible elderly workers into early retirement, like in the case of Denmark. Our model's key insight is thus indirectly supported also by a comparative study of individual exit rates on representative worker panel data from four other European countries.

8 References

Abowd, J.M. and D. Card (1989), On the covariance of earnings and hours changes, *Econometrica*, 57(2), 411–445.

Albaek, K., M. van Audenrode and M. Browning (2002), Employment protection and the consequences for displaced workers: a comparison of Belgium and Denmark, in "Losing work, moving on: Work displacement in international perspective", P. Kuhn (ed), Kalamazoo, MI: Upjohn Institute

Aubert, P., E. Caroli and M. Roger (2006), New Technologies, Organisation and Age: Firm-Level Evidence, *The Economic Journal* 116 (509), F73-93.

Autor, D.H. and M.G. Duggan (2003), The rise in the disability rolls and the decline in unemployment, *Quarterly Journal of Economics*, 118 (1), 157

[41] See Tatsiramos (2010, Sections 5.2 and 6) for these conclusions. In the EHCP data he uses some displaced workers who are observed to eventually retire early, might be entering unemployment or inactivity immediately subsequent to the displacement: this is not inconsistent with our approach, since his data is monthly and these intermediary spells are typically short, a vast majority less than one year (which is the frequency of our data).

Bentolila, S. and G. Bertola (1990), Firing costs and labour demands: how bad is eurosclerosis?, *Review of Economic Studies*, 57(3), 381–402.

Bingley, P. and G. Lanot (2004), Employer pay policies, public transfers and the retirement decisions of men and women in Denmark, *European Economic Review*, 48 (1), 181–200.

— and — (2007), Public pension programmes and the retirement of married couples in Denmark, *Journal of Public Economics* 91, 1878-1901

Black, D., K. Daniel, K. and S. Sanders (2002). The impact of economic conditions on participation in disability programs: Evidence from the coal boom and bust. *American Economic Review*, 92 (1), 27–50.

Bond, S., C. Nauges and F. Windmeijer (2005), Unit Roots: Identification and testing in micro panels, Cemmap working paper CWP07/05.

Buhai, I.S, M. Portela, C. Teulings, and A. van Vuuren (2009). Returns to tenure or seniority?, IZA Discussion Paper No. 3302 (updated version).

Carneiro, A. and P. Portugal (2006), Earnings losses of displaced workers: Evidence from a matched employer-employee dataset, IZA Discussion paper No. 2289

Chan, S. and A.H. Stevens, (1999), Employment and retirement following a late-career job loss. *American Economic Review*, 89 (2), 211–216.

— and — (2001), Job loss and employment patterns of older workers. *Journal of Labor Economics*, 19 (2), 484–521.

Coile, C. C. and P.B. Levine (2007), Labor market shocks and retirement: Do government programs matter? *Journal of Public Economics*, 91 (10), 1902–1919.

— and — (2009), The market crash and mass layoffs: How the current economic crisis may affect retirement, NBER Working Paper 15395.

Couch, K.A. and D.W. Placzek (2010), Earnings losses of displaced workers revisited, American Economic Review, 100 (1), p. 572-589

Dixit, A.K. (1989), Entry and exit decisions under uncertainty, *Journal of Political Economy*, 97, 620-638

— and R.S. Pindyck (1994), Investment under uncertainty, Princeton University Press.

Dorn, D. and A. Sousa-Poza (2007), 'Voluntary' and 'involuntary' early retirement: An international analysis. IZA Discussion paper 2714

Feldstein, M. (1976), Temporary layoffs in the theory of unemployment, October, *Journal of Political Economy*, p. 937-958

—- (1978), The effect of unemployment insurance on temporary layoff unemployment, *American Economic Review*, December, p. 834-846

French, E. (2005), The effects of health, wealth, and wages on labour supply and retirement behaviour, *Review of Economic Studies*, 72(2), 395-427

— and J. Jones (2011), The Effects of Health Insurance and Self-Insurance on Retirement Behavior, *Econometrica*, 79 (3), 693–732

Hakola, T. and R. Uusitalo (2001), Let's make a deal. The impact of social security provisions and firm liabilities on early retirement, Finnish Government Institute for Economic Research (VATT) working paper 260

— and — (2005), Not so voluntary retirement decisions? Evidence from a pension reform. *Journal of Public Economics*, 89 (11-12), 2121-2136

Hallberg, D. (2011), Is early retirement encouraged by the employer? Labor-demand effects of age-related collective fees, Uppsala Center for Labour Studies working paper 2011:5

Hopenhayn, H. and R. Rogerson (1993), Job turnover and policy evaluation: A general equilibrium analysis, *Journal of Political Economy*, 101 (5): 915-938

Hutchens, R. (1999), Social security benefits and employer behavior: Evaluating social security early benefits as a form of unemployment insurance, *International Economic Review*, 40 (3), 659-678

Hurd, M.D. (1990), Research on the elderly: Economic status, retirement, and consumption and saving, *Journal of Economic Literature*, 565-637

Korkeamaki, O. and T. Kyyra (2010), Institutional rules, labour demand and retirement through disability programme participation, *Journal of Population Economics*, 2010, 1-30

Krueger, A.B. and J.S. Pischke (1992), The effect of social security on labor supply: A cohort analysis of the notch generation, *Journal of Labor Economics*, 412-437

Jacobson, L.S., R.J. LaLonde and D.G. Sullivan (1993), Earnings losses of displaced workers, *American Economic Review*, 83 (4), 685-709

Jovanovic, B. (1982), Selection and the evolution of industry, *Econometrica*, 50(3), 649–670.

Larsen, M. and P.J. Pedersen (2008), Pathways to early retirement in Denmark, 1984-2000, *International Journal of Manpower*, 29 (5), 384–409.

Maestas, N. and X. Li (2006), Discouraged workers? Job search outcomes of older workers, University of Michigan, Retirement Research Center Working Paper 2006-133.

OECD (1999), Employment protection and labour market performance, in *OECD Employment Outlook*, Ch. 2

Papke, L.E. and J.M. Wooldridge (1996), Econometric methods for fractional response variables with an application to 401(K) plan participation rates, *Journal of Applied Econometrics* 11: 619-632

Pedersen, P.J. and N. Smith (1996), A duration analysis of the decision to retire early. In Wadensjo, E (ed), The Nordic Labour Markets in the 1990's, North Holland, Amsterdam

Pfann, G.A. (2006), Downsizing and heterogenous firing costs, *Review of Economics and Statistics*, 88(1): 158-170

Rust, J. and C. Phelan (1997), How social security and medicare affect retirement behavior in a world of incomplete markets, *Econometrica*, 65 (4), 781-831

Sullivan, D. and T. von Wachter (2009), Job displacement and mortality: An analysis using administrative data, *Quarterly Journal of Economics*, 124 (3), 1265-1306

Sutton, J. (1997), Gibrat's legacy, *Journal of Economic Literature*, 35(1), 40–59.

Tatsiramos, K. (2010), Job displacement and the transitions to re-employment and early retirement for non-employed older workers, *European Economic Review*, 54, p. 517-535

Topel, R. H. (1983), On layoffs and unemployment insurance, *American Economic Review*, October, 541-559

— (1984), Experience rating of unemployment insurance and the incidence of unemployment, *Journal of Law and Economics*, April, p. 61-90

— and M.P. Ward (1992), Job mobility and the careers of young men, *Quarterly Journal of Economics*, 107(2), 145–176.

von Wachter, T. (2007), The effect of economic conditions on the employment of workers nearing retirement age, Boston College, Center for Retirement Research Working Paper 2007-25.

—, J. Song and J. Manchester (2009), Long-term earnings losses due to job separation during the 1982 recession: An analysis using longitudinal administrative data from 1974 to 2004, Columbia University, Dep of Econ Discussion Paper No. 0708-16 (updated)

de Wit, G. (2005), Firm size distributions. An overview of steady-state distributions resulting from firm dynamics models, *International Journal of Industrial Organization*, 23(5–6), 423–450.

A Test of the Gibrat law implication

This appendix section presents the testing methodology and results for the Gibrat law implication for the log firm size, cf. Proposition 1, Section 3. Gibrat's law is known to hold in particular for large firms, c.f. Jovanovic (1982); as in the empirical analysis we work only with firms with more than 50 employees, we expect this law to hold accurately.

There is to date a massive literature on testing variants of Gibrat's law, see for instance reviews by Sutton (1997) or De Wit (2005). As *slight deviations* from Gibrat's law do not invalidate our model's implication, what we seek to show here is that log firm size follows approximately an AR(1) process with transitory shocks, over time. To that aim we utilize two simple tests.

The first approach is laid out in Abowd and Card (1989) and Topel and Ward (1992) for log wages; we adapt this methodology for log firm sizes. First, we estimate

$$\Delta n_{jt} = \delta_0 + \delta_1 Z_{jt} + \varepsilon_{jt}, \tag{14}$$

where Δ is the first difference operator and where Z_{jt} is a vector of controls: age category of the firm, time effects and industry indicators. Second, we construct the autocovariance matrix of the residuals ε_{jt} of this regression. If n_{jt} follows a random walk, ε_{jt} should be uncorrelated across time t. We perform these two steps for various subsamples:

(1) our entire sample of firms with at least 50 employees each year they are observed in the data;

(2) a balanced sub-sample of the sample from point (1), comprising firms that survive for at least 6 continuous years, 1980-1985;

(3) a sub-sample of firms from point (1) with at least 200 employees each year, hence even larger firms; and

(4) the sub-sample of all firms at (1), that are older than 10 years (which means they are either older than 10 years in the beginning of our observation period or that we keep them only after reaching at least 10 years in the data).

These various sub-samples try to uncover the situations for which the Gibrat Law holds better (larger, older, or surviving at least a number of years as a balanced panel). The resulting covariograms for (14), displaying the first 6 lags only, in order to save space, are reported in Table 11. The evidence from Table 11 suggests that Gibrat's law holds closely in all cases. All lagged correlations are really small in magnitude relative to the variance of shocks, reported in the first line, and in the vast majority of cases they are also not statistically different from 0; all our four specifications give very similar results, providing strong support for Gibrat's law.

Our second approach follows Bond et al. (2005). Take a simple dynamic AR(1) panel data model:

$$n_{jt} = \alpha n_{j,t-1} + u_{jt}, \tag{15}$$

where $u_{jt} \equiv (1-)\gamma_j + v_{jt}$ and the initial firm size $n_{j1} = \alpha_0 + \alpha_1 \gamma_j + \varepsilon_j \gamma_j$, with ν_{jt} and γ_j error terms such that E(γ_j) =E$(v_{jt}) = 0$ and E$(v_{jt} v_{js}) = 0$ for $t \neq s$. Under the null of $\alpha = 1$ the OLS estimator of α in (15) is consistent. We refer to this estimator of α as the OLS estimator. Under the alternative $\alpha < 1$, the OLS estimator is biased upwards, more

Table 11: 1st Gibrat's Law Test: Residual Autocovariances

Lag	(1)	(2)	(3)	(4)
0	0.038	0.030	0.025	0.034
	(0.0020)	(0.0027)	(0.0016)	(0.0019)
1	-0.0001	-0.0006	0.00005	-0.00008
	(0.0011)	(0.0016)	(0.0009)	(0.0008)
2	-0.0006	-0.0019	0.0008	-0.0004
	(0.0006)	(0.0014)	(0.0007)	(0.0006)
3	.00004	-0.0009	0.0007	0.0002
	(0.0003)	(0.0009)	(0.0003)	(0.0003)
4	-0.0005	.	-0.0003	-0.0005
	(0.0003)	(.)	(0.0003)	(0.0004)
5	.0001	.	0.0003	0.0003
	(0.0003)	(.)	(0.0004)	(0.0003)
6	.0003	.	-0.0005	0.00003
	(0.0006)	(.)	(0.0005)	(0.0004)
N. obs	29120	5742	10780	19630
N. firms	3373	1177	1514	2012

Specification (1) corresponds to estimates using the entire sample of firms with at least 50 employees each year; (2) to the sub-sample of firms from 1. surviving at least the first 6 years; (3) to the sub-sample of firms from 1. with at least 200 employees in each year of their life spans, and (4) to the sub-sample of firms from 1. at least 10 years old. All regressions control for age of the firm, time and industry effects. (Robust standard errors in parentheses).

so when $\text{Var}(\gamma_j)/\text{Var}(v_{jt})$ is large. In the latter case, one could a transformed statistic, estimating α from:

$$n_{jt} - n_{j1} = \alpha(n_{j,t-1} - n_{j1}) + \varepsilon_{jt} \tag{16}$$

where $\varepsilon_{jt} = v_{jt} - (1-\alpha)\left(n_{j1} - \gamma_j\right)$. The OLS estimator of (16) is consistent again under the null and again upwards biased under the alternative $\alpha < 1$, but this time the bias does not depend on $\text{Var}(\gamma_j)/\text{Var}(v_{jt})$. The results for both these methods, and for all sub-samples described above under points (1) to (4), hence 2x4 columns in total, are shown in Table 12; they are very similar in terms of interpretation to the estimates in Table 11: Gibrat's law holds reasonably well, with α estimated very close to 1, and with the Mean Squared Error (MSE) reasonably small.

As earlier stated, our theory does not require that Gibrat's Law holds perfectly, only that there are no large deviations from it; we have confirmed this with both type of empirical tests above.

Table 12: 2nd Gibrat's Law Test: Unit Root Type Regressions

	(1)		(2)		(3)		(4)	
Coef	OLSa	OLSb	OLSa	OLSb	OLSa	OLSb	OLSa	OLSb
α	.978 (.0019)	.987 (.0048)	.981 (.0023)	.930 (.0226)	.983 (.0019)	1.005 (.0039)	.984 (.0017)	.996 (.0041)
N. obs	29120		5742		10780		19630	
N. firms	3373		1177		1514		2012	
R^2	0.96	0.84	0.97	0.64	0.97	0.93	0.97	0.88
MSE	0.20	0.20	0.21	0.21	0.16	0.16	0.18	0.18

The dependent variable is logfirmsize in OLSa columns and (logfirmsize-initial logfirmsize) in OLSb columns. Columns indexed 1 correspond to estimates using the entire sample of firms with at least 50 employees each year; 2 to the sub-sample of firms from 1. surviving at least the first 6 years; 3 to the sub-sample of firms from 1. with at least 200 employees in each year of their life spans, and 4 to the sub-sample of firms from 1. at least 10 years old. Both regression types (OLSa and OLSb) control for age of the firm, time and industry effects. (Robust standard errors in parentheses).

B Age-profile in individual layoff hazards

As explained in the introduction of the subsection 6.3.1 on individual-level estimation results, this appendix reports results for the general probability of individual job exit conditional on worker age categories and a host of other covariates, on the samples of downsizing events from Table 3, including all workers. The estimation procedure is the same as the one used in (12), i.e. logit, allowing for heteroskedasticity-robust and firm-year clustering standard errors. The baseline age category is younger than 25 years old.

Based on the estimated logit coefficients in Table 13, subsequent Table 14 computes the marginal effect of each of the four age categories, at the sample mean of all the other covariates. Both Tables 13 and 14 show, as expected, that younger and older workers have, ceteris paribus, higher probabilities of separation than workers aged 25 to 60; workers in the age range targeted by early retirement policies (60 to 66) are about equally likely to be laid off as newly hired workers below 25 years old, while the elderly old, 67+ of age, qualified for the official public pension in Denmark, are the most likely to leave the firm among the four age groups considered.

The estimated effects of other covariates[42] such as education, tenure, potential experience, hourly wage, or firm size match also to what was earlier found in the literature: workers who, ceteris paribus, are more educated, have longer job duration, more labour market experience, earn more, or are employed in larger firms, are relatively less likely to separate than their co-workers.

[42]There are also differential effects by occupation, industry, region, and time indicators (all those groups of indicators are jointly statistically significant), stressing again the importance of controlling for those variables: detailed tables are available upon request from the authors.

Table 13: Age-profile in individual layoff probability

	Min-max downsizing as % of previous firmsize		
	(20%-80%)	(30%-80%)	(50%-80%)
25<=age<60	-.120	-.093	-.028
	(.019)***	(.030)***	(.054)
60<=age<67	.051	.022	-.030
	(.034)	(.050)	(.079)
age>=67	.859	.533	.483
	(.102)***	(.107)***	(.238)**
education years	-.017	-.021	-.040
	(.004)***	(.005)***	(.011)***
male	.119	.140	.200
	(.035)***	(.042)***	(.068)***
potential experience	-.015	-.015	-.021
	(.001)***	(.001)***	(.003)***
tenure	-.062	-.054	-.010
	(.006)***	(.008)***	(.009)
hourly wage	.0005	.0006	-.0002
	(.0002)***	(.0003)**	(.0002)
firm size	8.57e-06	-5.00e-07	.00008
	(3.90e-06)**	(5.75e-06)	(.00004)*
cons	-.199	.485	.830
	(.211)	(.275)*	(.282)***
N. obs.	667679	317845	84715
Log likelihood	-420977.12	-210706.73	-52340.13
Pseudo-R^2	0.0411	0.0402	0.0483

The sample contains workers of all ages, in firms with at least 50 employees in each year of their observation period, at the time of a mass downsizing episode. These are episodes of the minimum downsizing magnitude described in each column. The baseline age category is less than 25 years old, age<25. All three specifications include a full set of occupation, industry, regional, and time indicators. (Std. errors in parentheses under the estimated coefficients are heteroskedasticity-robust and clustered per firm-year.)

Table 14: Age-category marginal effects on individual layoff hazard

	Min-max downsizing as % of previous firmsize		
	(20%-80%)	(30%-80%)	(50%-80%)
age category			
age<25	.379	.480	.661
	(.008)	(.012)	(.014)
25<=age<60	.352	.457	.655
	(.006)	(.010)	(.009)
60<=age<67	.391	.486	.654
	(.007)	(.010)	(.017)
age>=67	.591	.611	.760
	(.023)	(.023)	(.042)

Marginal effects for the age categories are computed keeping all other covariates fixed at their sample mean, based on the logit estimates in Table 13 (Std. errors in parentheses are estimated using the Delta method.)