

IZA DP No. 7344

Does Experience Rating Reduce Disability Inflow?

Tomi Kyyrä
Juha Tuomala

April 2013

Does Experience Rating Reduce Disability Inflow?

Tomi Kyyrä

*Government Institute for Economic Research, Helsinki,
University of Helsinki and IZA*

Juha Tuomala

Government Institute for Economic Research, Helsinki

Discussion Paper No. 7344
April 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

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

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

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

ABSTRACT

Does Experience Rating Reduce Disability Inflow?*

This study explores whether the experience rating of employers' disability insurance premiums affects the inflow of older employees to disability benefits in Finland. To identify the causal effect of experience rating, we exploit a pension reform that extended the coverage of the experience-rated premiums. The results show that a new disability benefit claim can cause substantial cost to the former employer through an increased premium. Nonetheless, we find no evidence of the significant effects of experience rating on the disability inflow. The lack of the behavioral effects may be due to the complexity of experience rating calculations and/or limited employer awareness.

JEL Classification: J14, J26, H32

Keywords: experience rating, disability insurance, early retirement

Corresponding author:

Tomi Kyyrä
Government Institute for Economic Research
Valtion taloudellinen tutkimuskeskus
Arkadiankatu 7
PO Box 1279
00101 Helsinki
Finland
E-mail: tomi.kyyra@vatt.fi

* Financial support from the Finnish Centre for Pensions is gratefully acknowledged. We thank Mikko Kautto, Toivo Lehtinen, Juha Rantala and Roope Uusitalo for their helpful comments and Jukka Lampi for constructing the data set.

1 Introduction

Disability benefit programs have attracted considerable interest among economists on both sides of the Atlantic. This is not surprising, given the economic importance of these programs. For example in 2005, 26 million people in the European Union countries were on disability benefits, while in the United States 7.7 million people received benefits through the Social Security Disability Insurance program and 4.4 million through the Supplementary Security Income program (Benítez-Silva et al., 2010). Most economic studies of disability programs have focused on estimating the effects of benefit levels or eligibility criteria on labor supply (e.g. Gruber, 2000; Black et al., 2002; Campolieti, 2004; Autor & Duggan, 2003, 2006; Karström et al., 2008; Kyrrä, 2010; Staubli 2011). Far less effort has been devoted to analyze the role of employers and their incentives, which is the topic of this study.

When a worker applies for a disability benefit, the employer typically has no direct control over the decision that the medical professionals make. Nonetheless, the employer can influence disability outcomes indirectly. For example, the employer may reduce the onset of disabling illnesses at the workplace by investing in workplace health and safety, and by allocating the workload appropriately between employees. When the worker anyway develops a medical condition that reduces his or her working capacity, the employer has the discretion of whether to provide physical aid or retraining or whether to modify the worker's job tasks, which may enable the worker to remain at work. The problem is that the employer's incentives to implement these types of measures can be weak even if their costs to the employer were considerably less than the costs of a new disability benefit claimant to the society.

One policy option that may mitigate this problem is to require individual employers to bear some of the costs of their employees' disability benefit claims through experience-rated disability insurance (DI) premiums. With experience rating, the employer's premium is adjusted to reflect the costs of its workers' disability benefit claims in comparison to other employers. Employers with high disability costs are penalized through a surcharge on top of the base premium, while employers with low disability costs are rewarded by giving a discount on the base premium. If successful, experience rating induces employers to improve their workplace health and safety, and encourages them to implement cost-effective accommodations that enable those employees who have health problems to remain on the job. This should result in fewer individuals claiming disability benefits. However, there is no compelling evidence to what extent, if at all, the experience rating of the DI premiums affects the disability inflow.¹ We address this question by quantifying the effect of the experience rating on the disability inflow of the older workers in Finland.

¹There is an extensive literature on the effects of experience rating in other forms of social security, including workers' compensation for on-the-job injuries (e.g. Ruser, 1991; Bruce & Atkins, 1993; Kralj, 1994; Thomason & Pozzebun, 2002), unemployment insurance (e.g. Topel, 1983; Meyer, 2002, Anderson & Meyer, 1993, 2000) and unemployment-related pensions (Hakola & Uusitalo, 2005). Given that the employer has less control over disability outcomes than layoffs and workplace injuries, these studies do not offer much guidance for understanding the role of experience rating in DI.

This topic is of considerable interest, as reforming disability programs are high on the policy agenda of many governments and experience rating is viewed as a potentially effective policy. For instance, Autor (2011) and Burkhauser & Daly (2011) have recently suggested that the U.S. Social Security DI program should be financed by an experience-rated payroll tax. To motivate this policy proposal, these authors rely on anecdotal evidence from the Netherlands. Over the past two or three decades, the Netherlands has implemented a series of disability program reforms, including the introduction of experience-rated DI premiums in 1998. These reforms have been followed by sharp declines both in the disability inflow and in the share of the Dutch population on disability benefits (for example, see García-Gómez et al., 2011). However, due to the number of simultaneous changes that confound the effects of individual policy measures, it is not clear to what extent adopting the experience rating explains these outcomes.² The analysis of the Finnish DI system provides an alternative - perhaps one that is even better point of reference - for the discussion of the efficiency of experience rating in DI.

In Finland, employers are subject to various degrees of experience rating depending on the size of the firm. To identify the effects of experience rating, we utilize a reform that unified the major pensions Acts in the private sector in 2007. This reform had nothing to do with experience rating, but it accidentally extended the coverage of experience rating to certain new groups of workers and their employers. After this reform, medium-sized and large firms began to pay experience-rated DI premiums for their employees who used to be insured under specific pension Acts. In contrast, the smaller firms in the same industries were not affected but continued to pay flat-rate DI premiums. As a consequence, we can compare disability outcomes under the experience rating and flat-rate schemes in a difference-in-differences type of setting.

Using linked employer-employee data, we construct various measures of disability inflows and compare their changes between the pre-reform and post-reform periods in the firms of different size. This analysis conducted at the firm level does not support the hypothesis that experience rating reduces disability inflow. In the next stage, we compute marginal costs for all the employees who were at risk of being awarded a disability benefit. The marginal cost is defined here as the expected increase in the employer's future DI premium that would result if a disability benefit would have been awarded to the worker. Before the 2007 reform, the marginal cost was zero for all the employees in all firms. In the post-reform period, the marginal costs for the employees of medium-sized and large firms became positive and increased over time due to the gradual adoption of the experi-

²Koning (2009) found that the disability inflow decreased in the Dutch firms that experienced a change in the DI premium when compared to the firms with unchanged premiums. Koning interpreted this as evidence that employers were not completely aware of experience rating and therefore the premium change served as a "wake-up call", which induced preventative measures that reduced the disability events in the subsequent years. Since his data only covered the post-reform years, there were no exogenous changes in the firms' premiums but all the changes were driven by the past change in the disability benefit claims made by the firm's own employees. While the results point to some behavioral responses and information imperfections, the results do not describe the causal effects of experience rating as compared to the counterfactual case of the flat-rate DI premiums, which we analyze in this study.

ence rating system. The resulting exogenous variation in the marginal costs allows us to estimate the causal effects of experience rating. We first demonstrate that the marginal cost can be high, being comparable in size to a worker's annual salary in many cases. The marginal costs are particularly high for relatively young employees in large firms. Then we add the marginal cost as an explanatory variable to various probability models for transitions out of work. Consistently with our findings from the firm-level analysis, the results from these models imply that experience rating has no effect on the transition rate to disability benefits. Thus, in the light of our analysis, the efficiency of experience rating as a disability prevention device seems questionable. This is perhaps an unexpected result, given the size of the costs the employers are liable for. The lack of the behavioral effects may be due to the complexity of the premium calculations, limited employer awareness and/or the transitional provisions associated with the pension reform.

This paper will proceed as follows: In the next section we discuss the disability benefit schemes and DI premiums in Finland. In Section 3 we describe the pension reform that extended the coverage of experience rating. In Section 4 we describe the data and report some descriptive statistics. Section 5 presents the results of our firm-level analysis. In Section 6 we discuss the computations of the marginal costs and report the results for the individual-specific transition rates. The final section concludes.

2 Institutional framework

2.1 Sickness and disability benefits

An employee who is unable to perform his or her job due to illness or injury is entitled to compensation for income loss. In order to receive this compensation, the applicant needs a statement by a doctor certifying that he or she is not capable of work. For the first weeks (typically one to three months), the applicant is fully compensated by the employer, after which he or she can claim a *sickness benefit*. The sickness benefit can be received for a maximum of about one year (300 working days, Saturdays included). Depending on the illness or the injury, the applicant's rehabilitation needs are assessed in a more extensive medical examination during the sickness benefit period. In case of prolonged disability, an individual between the ages of 16 and 62 can qualify for one of four possible disability benefits: (i) a partial disability pension, (ii) a full disability pension, (iii) a partial rehabilitation benefit, or (iv) a full rehabilitation benefit.

When it is unlikely that an applicant will return to work, he or she is awarded a disability pension for an indefinite period of time. Otherwise the applicant is entitled to a rehabilitation benefit (also known as a temporary disability pension), which is granted only for a specific period. The receipt of this benefit also requires that a rehabilitation plan has been drafted. For both benefit types, a full benefit is conditional on a loss in the working capacity of at least 60% and a partial benefit for a loss of at least 40% but below 60%. Disability evaluations are always made by trained professionals. When determining

eligibility, the individual’s age, education, occupation, place of residence and capability to support herself or himself by regular work are all taken into account along with the medical assessment. A disability pension may also be discontinued if the working capacity of the recipient improves, which rarely happens among older recipients. There is no automatic retesting of the disability status, except for new periods of the rehabilitation benefit. The disability benefits can be received until the age of 63 when the entitlement to an old-age pension begins.

2.2 Disability insurance premiums

A major part of disability benefit costs is financed by partially experience-rated premiums (or payroll taxes). The degree of experience rating depends on the firm size, as measured by a firm’s payroll two years earlier. Small firms are not subject to experience rating and they only pay *base premiums*. The base premium is calculated by taking the sum of the age-specific DI taxes over all employees:

$$Q_{kt} = \sum_j \zeta_t(x_{jt})w_{jt}, \quad (1)$$

where k indexes the firm and t indexes the year, ζ_t is the DI tax rate,³ and x_{jt} and w_{jt} are the age and annual salary of employee j , respectively.

Large firms pay *experience-rated premiums*, which are calculated by multiplying the base premium by the *experience multiplier* $m(r_{k(t-2)}, r_{k(t-3)})$. This multiplier takes a value between 0.1 and 5.5, depending on the costs of the disability pension claims made by the firm’s former employees two to three years earlier. These costs are measured by *risk ratios* $r_{k(t-2)}$ and $r_{k(t-3)}$, which are defined below. On the basis of the average of these risk ratios, the firm is allocated to one of 11 possible contribution categories, each of which corresponds to a particular value of m . See the solid line in Figure 1.⁴ The experience-rated premium mQ_{kt} can differ substantially from the base premium Q_{kt} . In principle, a large firm can obtain a 90% discount on the base premium or be obligated to pay a 450% surcharge on top of the base premium.

In contrast to small and large firms, medium-sized firms pay a weighted sum of the base and experienced-rated premiums, and are thus only partially covered by the experience rating. In general, the DI premium is calculated as

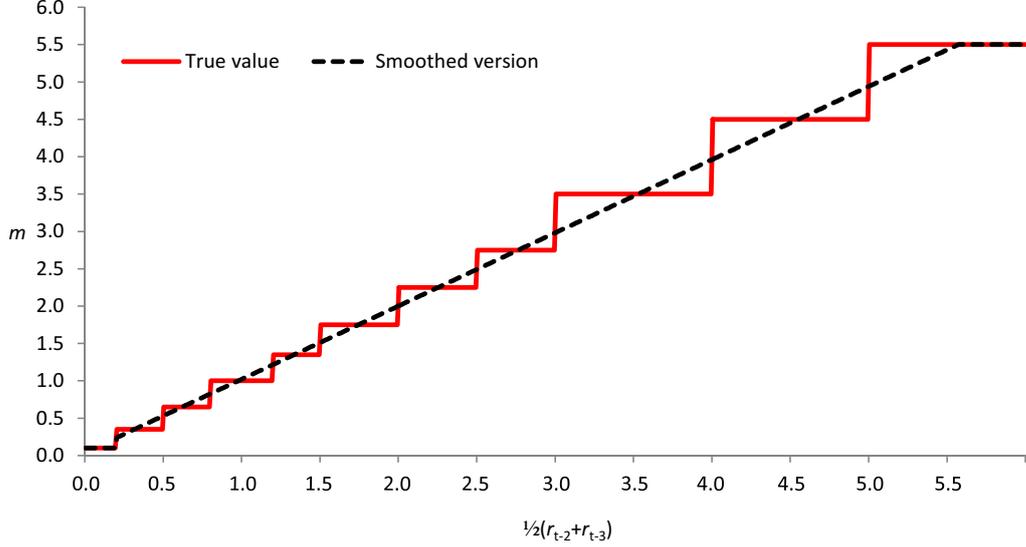
$$C_{kt} = (1 - \alpha(W_{k(t-2)})) Q_{kt} + \alpha(W_{k(t-2)})m(r_{k(t-2)}, r_{k(t-3)}) Q_{kt}, \quad (2)$$

where $W_{k(t-2)}$ is the payroll in year $t - 2$ and α is the *degree of experience rating*; α is 0 for small firms with $W_{k(t-2)} \leq \underline{W}_t$ and 1 for large firms with $W_{k(t-2)} \geq \overline{W}_t$, and between

³The age variation in the DI tax rate reflects the differences in the disability risk and the benefit levels across the age groups.

⁴A smoothed version of the multiplier (the dashed line in Figure 1) is used in our calculations. The reasons for this will be explained later in Section 6.

Figure 1: The experience multiplier in year t , $m(r_{k(t-2)}, r_{k(t-3)})$, as a function of the average of the risk ratios in years $t-2$ and $t-3$, $\frac{1}{2}(r_{k(t-2)} + r_{k(t-3)})$



W_t and \bar{W}_t it increases linearly from 0 to 1 with $W_{k(t-2)}$. The threshold values for small and large firms are updated annually, and they correspond approximately to the firm sizes of 50 and 800 employees with the average salary.

The risk ratio is computed as

$$r_{kt} = \frac{\sum_{j \in D_{kt}} e_{jt}}{R_{kt}}, \quad (3)$$

where e_{jt} is the present value of a (full or partial) disability pension awarded in year t , and D_{kt} denotes the set of individuals who worked for the firm one to two calendar years prior to the year of the *retirement event*.⁵ The retirement event refers to the day when the individual was diagnosed with the disability leading to the disability pension. Due to periods of sickness and rehabilitation benefits, the retirement event is often one or two years before the disability pension is awarded. The present value e_{jt} equals the expected amount of disability pension benefits until the age of 63, i.e. the age when an old-age pension begins. The numerator of the risk ratio is referred to as the *realized disability cost*, as it is a measure of the expected present value of the disability pension claims made by the firm's former employees during the year t . The denominator R_{kt} is the *theoretical disability cost* and it corresponds to the average disability costs in the firms with the same age and wage structure (see the Appendix). This means that if the new claims for disability pension in a firm cause higher than the average costs to the pension system, $r_{kt} > 1$, which tend to push the experience multiplier above one with the delay of two to three years.

⁵If the individual had more than one employer during these two years, the contribution of e_{jt} is divided between the employers according to the share of salaries they have paid during that two-year period.

It is noteworthy that the risk ratio only depends on the disability pension claims, not on the rehabilitation benefit claims. This may induce the employer to encourage those employees with health problems to apply for a rehabilitation benefit rather than for a disability pension. Another important point is that only the first disability pension of each person is taken into account. In particular, if a worker collects a partial disability pension and this is followed by a full disability pension, only the present value of the partial pension has an effect on the firm's risk ratio in the year when that pension was awarded. To minimize the impacts on the risk ratio, the employer may thus encourage those workers who have health problems to apply for a partial disability pension first (i.e. by providing part-time work for a short period of time). These two features of the risk ratio calculations suggest that the effect of the experience rating on the partial disability pension claims and on the rehabilitation benefit claims is ambiguous. To the extent that experience rating leads to general improvements in the firm's health and safety policy, it should reduce the transitions to all types of disability benefits. However, the experience rating may also increase the inflow to partial disability pension as well as the rehabilitation benefits in some cases through the substitution effects.

Experience rating may also affect the relative risk of becoming a disability pension recipient between the employees of different ages within large firms. This is because the effect of a new disability pension claim on the risk ratio is determined by the present value of the pension benefits, which is larger for young claimants who still have many years before being eligible for old-age pensions. Thus, the employer has an incentive to devote extra effort to prevent its young employees from claiming disability pension benefits.

3 TyEL reform and experience rating for former LEL employers

To estimate the effects of the experience rating, we exploit a pension reform that unified the private-sector Pension Acts in 2007. As a by-product of this reform, the experience rating of the DI premiums was extended to cover new groups of workers and their employers. Before the reform, all private-sector employees were covered either by the Employees' Pension Act (TEL), the Temporary Employee's Pensions Act (LEL), or by the Pension Act for Performing Artists and Certain Groups of Employees (TaEL). Whereas a vast majority of these employees were insured under the TEL, the LEL covered dock workers and blue-collar workers in the fields of construction, agriculture and forestry while the TaEL covered artists, journalists and those who worked for households. The employers paid the experience-rated DI premiums for their workers who were insured under the TEL.⁶ But

⁶The experience rating system changed in 2006 when the current system came into effect for the TEL workers. Before this reform, those firms employing more than 50 workers were required to pay a given share of the present value of a new disability benefit claim as a lump sum payment to the pension provider at the time when the disability pension (or rehabilitation benefit) was awarded to their former employee who was insured under the TEL. The medium-sized firms paid only a small share of this present value. However, in addition to the lump sum disability costs, they also paid the flat-rate base premiums on an annual

for those workers insured under the LEL and the TaEL, the employers paid the flat-rate base premiums. That is, the experience rating system was only applied to the employers of the TEL workers.

On January 1, 2007, these three pension Acts were unified into a single Employees Pensions Act (TyEL). While this reform did not affect the eligibility criteria or benefit levels, i.e. the content of DI from the employees' perspective, it did extend the experience-rated DI premiums to also cover those worker groups who used to be insured under the LEL and the TaEL. As a consequence, the former LEL employers and TaEL employers whose payroll exceeded the threshold value of \underline{W}_t became subject to experience rating for the first time, whereas the smaller employers continued to pay only the base premiums. This provides a difference-in-differences type of setting, which we exploit to identify the causal effects of experience rating. In what follows, we focus on the workers insured under the LEL and their employers.

In practice, due to specific rules for the transition period, the transition of the former LEL employers to the experience rating scheme occurred gradually over time. The guiding principle for the transitional provisions was that the present value of the new disability pension claim affects the risk ratio of the former LEL employer only to the extent the underlying employment relationship falls in the TyEL period. Since the present value is assigned to the firms in which the claimant worked one to two calendar years prior to the year of the retirement event, only disability pensions with the retirement event in the year 2008 or later have an effect on the risk ratio of the former employer. The costs of disability pensions that were awarded during the TyEL period but were not assigned to any particular employer are pooled, i.e. collectively covered by all the former LEL employers. To account for this pooling, the risk ratios of the former LEL employers in the transition period were adjusted by adding a calculatory term to the numerator. Namely, the *adjusted* risk ratio for firm k in year $t \geq 2008$ is computed as

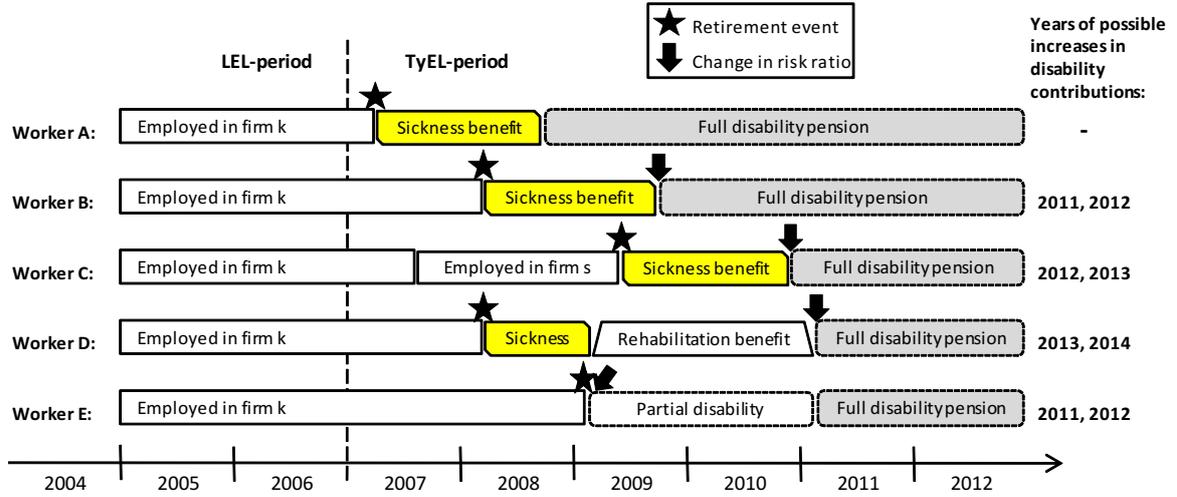
$$r_{kt}^A = \frac{\sum_{j \in D_{kt}} e_{jt} + E_{kt}}{R_{kt}}, \quad (4)$$

where e_{jt} is accounted only to the extent that the underlying employment relationship fell in the TyEL period, and E_{kt} is the calculatory term, which is positive in the years 2008-2010 (see the Appendix for details).

Figure 2 illustrates how the present value of a new disability pension claim e_{jt} is assigned to the former employers in different cases. For example, let us consider worker A who was awarded a disability pension in year 2008 due to an illness diagnosed in 2007 at the beginning of his or her sickness period. The year of the retirement event is 2007, suggesting that the present value would have affected the risk ratios of the firms for which she worked in the years 2006 and 2005 (i.e. firm k), had the TyEL been in force for a longer

basis. In other words, prior to 2006, there were no experience-rated DI premiums but the medium-sized and large firms paid lump sum payments for the disability benefit claims made by their former employees insured under the TEL. Korkeamäki & Kyrrä (2012) discuss this old system and provide evidence for its behavioral effects.

Figure 2: Illustration of the allocation of disability pension costs for the former LEL employers



time. But, as both 2005 and 2006 are LEL years, the present value is not assigned to her employer but is jointly covered by all the former LEL employers. Next, let us consider workers B and D who retired one year later from the same firm but with different sickness and rehabilitation benefit profiles. The year of the retirement event for both workers is 2008. Assuming they earned the same salary in the years 2006 and 2007, only one-half of the present values of their disability pensions (i.e. 2007 share) is accounted for when calculating the risk ratios for firm k . In the case of worker B (D) the risk ratio is affected in the year 2009 (2011), which affects the experience-rated premium for the years 2011 and 2012 (2013 and 2014). A two-year period of rehabilitation benefits for worker D causes a rather long gap of three years between the retirement event and the change in the risk ratio.

In general, only disability pensions with the retirement event in 2009 or later (workers C and E in Figure 2) have a full impact on the former employer's risk ratio. In the case of worker C the present value is split between firms k and s according to the wages these firms paid to the worker in years 2007 and 2008. It is worth noting that the present value of worker E 's disability pension is based on the assumption that a partial disability pension would have lasted from 2009 until the time when the worker reaches the age of 63, so that only a (possibly very small) portion of the actual disability costs is assigned to firm k when computing its risk ratio for the year 2009, whereas a transition to a full disability pension in 2011 has no effect at all.

It is obvious that the calculatory term dominates the numerator of the adjusted risk ratio in 2008. This is because $e_{j08} > 0$ only if the retirement event was in the same year when the disability pension was awarded, which is applicable to only 17% of the cases. Furthermore, the relative weight of the calculatory term declines gradually to zero by 2010.

Table 1: Sample statistics for firm data

	Year					
	2005	2006	2007	2008	2009	2010
# of firms	573	597	627	663	663	663
# of firms with $\alpha > 0$	80	93	98	105	117	128
Mean firm size	38.2	43.2	53.6	56.5	46.4	49.9
Median firm size	20	21	22	22	21	22
Mean α for firms with $\alpha > 0$	0.180	0.199	0.207	0.202	0.197	0.191

Notes: Firm size is the number of all employees, regardless of the pension Act under which they are insured.

4 Data and descriptive statistics

The data were compiled by merging various administrative registers of the Finnish Centre for Pensions, which co-ordinates the entire pension system in Finland. The data include comprehensive records on employment periods and the wages for all the Finns who had some work history, as well as the detailed pension information for all retirees. Each employment relationship can also be matched to the firm records on the industry, payroll and the number of employees. However, apart from age, the data do not contain background information for individuals, nor is there information on the receipt of sickness benefits. But we do observe the retirement events, that is, the days when a diagnosis was made for the illness or disability that eventually led to a rehabilitation benefit or disability pension. This is important because the disability pension costs are assigned to the employers on the basis of the year of the retirement event.

Our analysis covers the period 2005-2010, but in order to compute the risk ratios and marginal costs, information is also needed from earlier years. First, we selected all workers insured under the LEL in 2005 and 2006. From the years 2007 to 2010, we included workers who would presumably have been insured under the LEL in the absence of the TyEL reform. Next, we traced the employers of these workers. The smallest employers were excluded by requiring that the firm employed at least 10 LEL workers each year during the period 2005-2010. For each firm in a given year we also need the payroll from two years earlier (to determine the value of α) as well as the information on the wages and ages for its LEL employees over the past four-year period (to compute the risk ratios). Due to some missing records, the firms included in the analysis do not necessarily appear in the data each year.

The resulting panel includes 663 firms, of which 573 are continuously observed over the period 2005-2010. On average, these firms are rather small. The mean size of the workforce varies over time, ranging from 38 to 57 workers (Table 1). The median firm size, however, is much smaller, being around 21 workers every year. About one-sixth of the firms have positive α , and thereby transferred gradually from the flat-rate scheme to the experience-rated scheme during the years 2007-2010. As these firms are by definition relatively large, a much larger share of the workers than the firms in the sample became

Table 2: Sample statistics for worker data

	Year at risk				
	2006	2007	2008	2009	2010
# of workers	8479	10016	10403	11451	11156
Mean age	52.2	52.4	52.5	52.4	52.5
Mean job tenure	9.6	8.9	9.2	8.9	9.4
Industry:					
Manufacturing	14.3	11.0	10.6	8.0	7.7
Building construction	40.3	48.1	48.1	46.8	46.3
Civil engineering	5.1	5.4	5.4	5.2	5.8
Specialized construction	18.0	16.8	17.1	18.1	19.3
Transport and storage	15.5	13.1	13.2	12.4	11.4
Other	6.8	5.6	5.7	9.5	9.5
Fraction of workers with $\alpha > 0$	0.591	0.631	0.627	0.629	0.629
Mean α for those with $\alpha > 0$	0.546	0.630	0.614	0.596	0.569
# of transitions to:					
partial rehabilitation benefit	2	0	1	1	0
full rehabilitation benefit	94	81	93	98	105
partial disability pension	10	16	18	13	17
full disability pension	58	69	84	90	77
# of firms in year $t - 1$	561	586	613	655	657
Mean firm size in year $t - 1$	39.6	45.9	58.6	46.7	50.2
Median firm size in year $t - 1$	21	22	22	21	22

Notes: Firm characteristics are for the firm for which the individual worked for the two years prior to the present year. Firm size is the number of all employees, regardless of the pension Act under which they are insured.

exposed to experience rating after the TyEL reform (over 60%, see Table 2).

In the individual-level analysis we focus on an older subgroup of LEL workers who fulfill certain conditions. To be included in the risk set in year $t \in \{2006, 2007, \dots, 2010\}$, we require that the individual (*i*) was between the ages of 45 and 61, (*ii*) had not received any pension benefits before the year t , and (*iii*) had been working during the years $t - 1$ and $t - 2$ (or in the TyEL period would have been working) under the LEL scheme for the same employer that was included in the firm panel. The younger workers are excluded from the analysis due to their very small risk of disability. The tenure restriction is required in order to detect employer whose risk ratio will be affected if a disability pension is awarded to the worker. It is noteworthy that the individual at risk in year t does not necessarily work during that year (for example, due to a sick leave or layoff).

The worker data include 18,197 individuals in 661 firms. The average age of the worker at risk is slightly higher than 52 years each year (see Table 2). On average, the workers had worked at their current firms for 9 to 10 years. Approximately two-thirds of these

Table 3: Risk set and transitions out of work by age

Age	Number of observations				Transition rates		
	Risk set	Rehab benefit	Disab Pension	Other exit	Rehab benefit	Disab Pension	Other exit
45	3478	13	2	253	0.0037	0.0006	0.0727
46	3496	25	1	270	0.0072	0.0003	0.0772
47	3378	18	3	237	0.0053	0.0009	0.0702
48	3354	32	2	224	0.0095	0.0006	0.0668
49	3313	17	3	205	0.0051	0.0009	0.0619
50	3273	29	3	194	0.0089	0.0009	0.0593
51	3324	42	11	205	0.0126	0.0033	0.0617
52	3203	31	15	196	0.0097	0.0047	0.0612
53	3184	24	14	189	0.0075	0.0044	0.0594
54	3158	33	22	169	0.0104	0.0070	0.0535
55	3163	37	20	180	0.0117	0.0063	0.0569
56	3083	39	26	172	0.0127	0.0084	0.0558
57	3020	40	54	215	0.0132	0.0179	0.0712
58	2915	35	59	215	0.0120	0.0202	0.0738
59	2623	40	67	264	0.0152	0.0255	0.1006
60	2186	15	79	182	0.0069	0.0361	0.0833
61	1794	5	71	166	0.0028	0.0396	0.0925
All	51,945	475	452	3536	0.0091	0.0087	0.0681

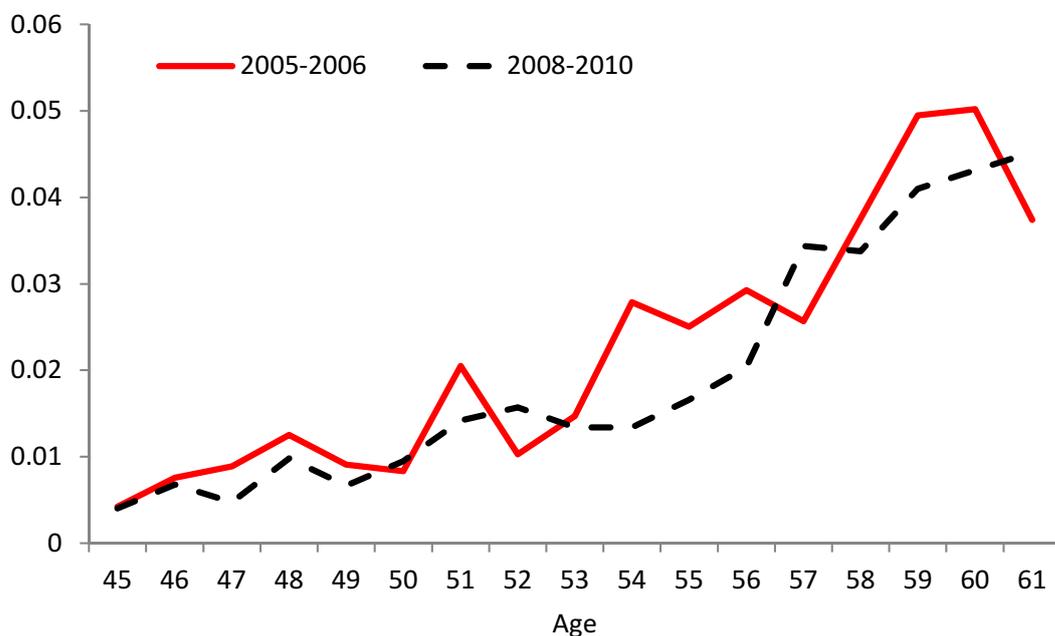
Notes: The risk set consists of all individuals who worked in the same firm for the past two calendar years. Other exits include transitions to all other destinations than disability benefits. Transition rates are obtained by dividing the number of transitions to the state into question by the number of individuals in the risk set.

workers were employed in one of three construction industries, and thus the data represent only a narrow sector of the economy. Over 60% of the individuals were employed in a firm with $\alpha > 0$ and consequently, they were affected by the experience rating system during the TyEL period. Furthermore, the average degree of experience rating within this group varies between 0.57 and 0.63 during the TyEL years.

By the end of the observation period, 927 workers had left the labor market to collect disability benefits. Most of these workers were awarded a full disability pension. The numbers of the recipients of partial benefits is relatively small: 74 workers were awarded a partial disability pension and only 4 were awarded a partial rehabilitation benefit. This makes a distinct analysis of the transitions to partial benefits infeasible, and for this reason we do not usually make a distinction between those receiving partial and those receiving full benefits.

Table 3 shows the size of the risk set and the number of transitions to the different exit destinations as well as the corresponding transition rates by age. Not surprisingly, the number of people at risk declines sharply with age. Only a small part of that decline can be explained by transitions to disability benefits. While the average transition rates

Figure 3: The transition rate to disability benefits (rehabilitation benefit or disability pension) as a function of age in the pre-reform and post-reform periods



to rehabilitation benefits and disability pension are of the same level, being about 0.009, the age pattern is rather different. Workers who are 56 years and under are much more likely to be granted a rehabilitation benefit than a disability pension, whereas the opposite occurs for the older age group, which is more likely to be granted a disability pension, not a rehabilitation benefit. The likelihood of becoming a disability pension recipient is particularly pronounced at ages 60 and 61. These observations are not surprising, given that the cost-benefit analysis of rehabilitation measures favors the younger workers who have a longer potential working career.

The other exit refers to the case where the individual left the firm without becoming a recipient of disability benefits. This outcome is a kind of residual state, which includes layoffs, employer changes and all states outside the labor force other than being on disability benefits. On average, 7% of workers leave their firm each year without claiming disability benefits. This rate varies less with age than does the transition rates to the disability benefits.

Figure 3 plots the overall transition rate to disability benefits as a function of age for the years 2005-2006 (pre-reform period) and 2008-2010 (post-reform period). The year 2007 is excluded, which was the first TyEL year but that particular year was when the experience rating did not yet have an effect. In both periods, the disability risk is first very low but increases with age, reaching the level of about 0.04 by age 60. Compared to the pre-reform period, the disability risk in the post-reform period is similar until the age of 53 years, but is slightly lower at older ages (except at ages 57 and 61). As a result, the average disability risk declined from 0.021 to 0.018 between the pre-reform and post-reform

periods. This decline is not necessarily related to the adoption of experience rating because the overall disability risk has been declining since 2003.

5 Risk ratios and disability inflow rates

The risk ratios measure the relative costs of the disability pension claims accounting for differences in the age structure of the workforce between firms. The risk ratios also determine the experience-rated premiums during the TyEL period, and hence should be of direct economic interest for the firms that were affected by the reform. The TyEL reform in 2007 may have encouraged large firms to adopt measures to reduce their risk ratios in order to gain from a lower DI premium in the future. Since the reform did not affect the incentives of the small firms that kept on paying flat-rate base premiums, the risk ratios of the large firms should have declined in comparison to those of the unaffected small firms. Moreover, as the financial gain from a lower risk ratio is positively related to the degree to which the DI premium is experience-rated in the TyEL period (as measured by α), the relative drop in the risk ratio among the larger firms is expected to be proportional to the firm's α .⁷

These hypotheses are tested by applying the following two-period model:

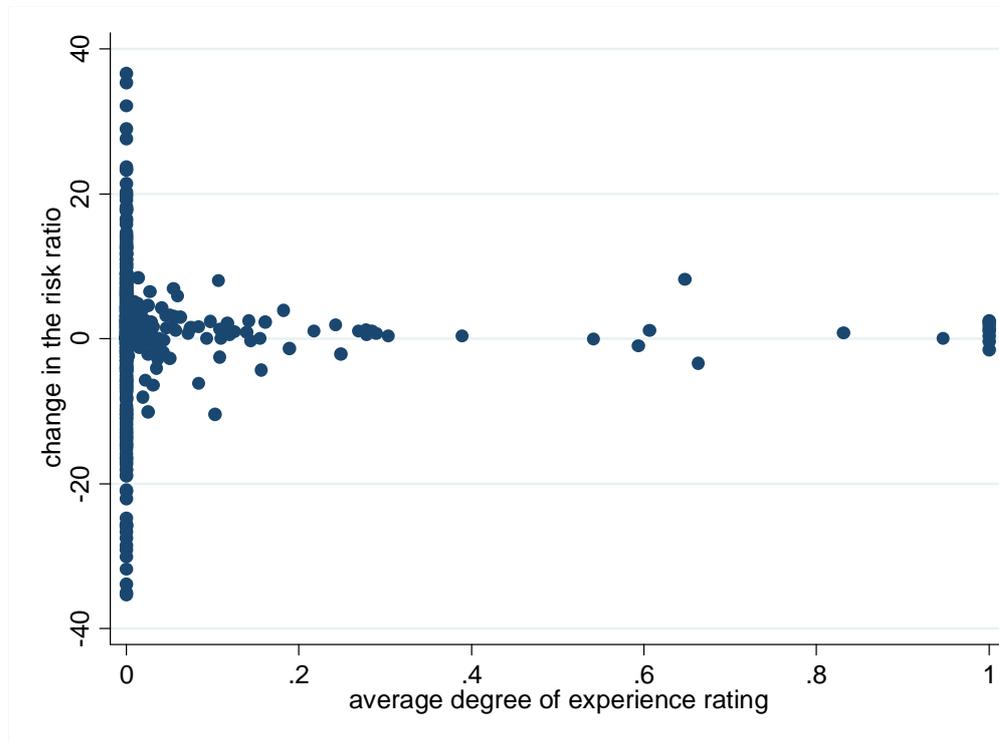
$$\bar{r}_{kt} = \psi + \lambda \cdot t + \eta \cdot \bar{W}_{kt} + \boldsymbol{\mu} \cdot \mathbf{X}_{kt} + \theta \cdot (\bar{\alpha}_{kt} \cdot t) + \varepsilon_{kt}, \quad (5)$$

where k indexes the firm and t indexes the time period, being 0 for the LEL-period 2005-2006 and 1 for the TyEL-period 2008-2010 (excluding the first TyEL year of 2007 when the experience rating had no effect yet). The outcome variable \bar{r}_{kt} is the average *unadjusted* risk ratio during a period t ,⁸ which is a measure of the relative disability pension costs that were caused to the pension system by the firm's former employees. On the right-hand side, \mathbf{X}_{kt} is the vector of industry dummies, \bar{W}_{kt} is the average payroll, $\bar{\alpha}_{kt}$ is the average degree of experience rating during period t , and ε_{kt} is the error term. Because $\bar{\alpha}_{kt}$ is a function of the payroll, and because the firm size (for which \bar{W}_{kt} is a proxy) may also have a direct effect on the disability outcomes, it is important that the possible payroll effect is controlled for. With \bar{W}_{kt} held constant, the effect of the experience rating is captured by θ . Since $\bar{\alpha}_{kt}$ does not occur during the period 0, the effects of \bar{W}_{kt} and $\bar{\alpha}_{kt}$ are easily sorted out. The change in the average risk ratio from period 0 to period 1 that is unrelated to experience rating is captured by λ , which is identified from the data on the small firms for which

⁷Comparing the changes in the risk ratios may seem odd because the risk ratio is a relative measure. However, the reference level of disability pension costs, i.e. the theoretical disability cost in the risk ratio formula, describes the average disability pension cost across all similar firms in the private sector. Because the former LEL employers are a relatively small group, this reference level is mainly determined by disability outcomes in other firms.

⁸The unadjusted risk ratio refers to the risk ratio that has been computed as if the TyEL would have been in force for a long time. That is, we do not include the calculatory term in the risk ratio but instead also assign the realized disability costs that are associated with the employment relationships that occurred during the LEL period to the employer. Due to some outliers for a few small firms, the year-specific risk ratios were top coded at the 99th percentile.

Figure 4: The change in the average unadjusted risk ratio from 2005-2006 to 2008-2010, $\bar{r}_{k1} - \bar{r}_{k0}$, (y-axis) versus the average degree of experience rating, $\bar{\alpha}_{k1}$, (x-axis).



$\bar{\alpha}_{kt} = 0$ in both periods. Under the plausible assumption that $\bar{\alpha}_{kt}$ is uncorrelated with the error term, the Ordinary Least Squares (OLS) estimate of θ has a causal interpretation.

Before turning to the regression results, it is illustrative to consider the scatter plot in Figure 4 that depicts $\bar{\alpha}_{k1}$ against $\bar{r}_{k1} - \bar{r}_{k0}$.⁹ The change in the average risk ratio is strikingly large for several small firms with $\bar{\alpha}_{k1} = 0$. This is because the small firms generally have a much wider range of the risk ratios than the large firms do. It is difficult to detect any relationship between the change in the risk ratio and the degree of experience rating in the graph. In particular, we do not see a negative association between these variables that would support our hypothesis. The OLS estimates of θ in Table 4 confirm this observation.

Model 1 corresponds to the baseline specification in (5). This is followed by a model with a more flexible specification of the firm-size effect (the 3rd order polynomial for the payroll). The effect of experience rating in both models is statistically insignificant (with a “wrong” sign). In the last two models we relax the restriction that the change in the risk ratio among the large firms is linearly related to the degree of experience rating. The coefficient on the dummy variable for the experience-rated firms in Model 3 suggests that the risk ratios changed identically in both those firms that became subject to experience rating and in those firms that were not affected by the TyEL reform. In Model 4, a non-monotone effect for the degree of the experience rating is allowed for by using dummy

⁹For expositional purposes, the graph does not include those four small firms with $\bar{\alpha}_{k1} = 0$ for which the absolute change in the average risk ratio is greater than 50.

Table 4: The effect of the degree of experience rating on the average unadjusted risk ratio

	Model 1		Model 2		Model 3		Model 4	
	Coeff	<i>t</i>	Coeff	<i>t</i>	Coeff	<i>t</i>	Coeff	<i>t</i>
$\bar{\alpha}_{kt} \cdot t$	0.3407	0.58	0.7474	0.96				
$1(\bar{\alpha}_{kt} \cdot t > 0)$					-0.0005	-0.00		
$1(\bar{\alpha}_{kt} \cdot t \in (0, .15])$							-0.0928	-0.28
$1(\bar{\alpha}_{kt} \cdot t \in (.15, .5])$							0.2237	0.33
$1(\bar{\alpha}_{kt} \cdot t > .5)$							0.6156	0.90
Payroll control:	Linear		Polynomial		Polynomial		Polynomial	

Notes: Models were estimated by Ordinary Least Squares. The data contains 1146 observations on 573 firms. $1(A)$ is 1 if A is true and 0 otherwise. All models include industry dummies. Model 1 also includes the payroll, whereas Models 2 to 4 include the 3rd order polynomial for payroll. The t statistics are based on the robust standard errors clustered at the firm level.

variables for the three distinct intervals of $\bar{\alpha}_{k1}$ (these intervals were chosen on the basis of the number of the available observations). The coefficients on these dummy variables do not exhibit a clear pattern, nor do they differ from zero at the conventional risk levels.

None of the models that were considered provides support for the hypothesis that the realized disability costs in the TyEL period would have declined in the firms that were exposed to experience rating as compared to the non-affected firms.¹⁰ However, the changes in the risk ratios may not reveal the whole story. This is because the risk ratio reflects only the disability pension costs, but the experience rating may also affect the rehabilitation benefit claims. For these reasons, to complete our analysis, we also examine the relationship between the degree of experience rating and the inflow rates of older workers to different disability benefits.

The dependent variable here is the share of 45 to 61 years old workers who were awarded a given type of disability benefit. We include only those firms that employed at least five workers in the relevant age category each year, even though the results are not sensitive to this restriction. The outcome variable is bounded between zero and one, and its distribution has a mass point at zero. To address with these data features, we follow Papke & Wooldridge (1996) and specify a fractional logit model of the form

$$E(\bar{y}_{kt} | \bar{W}_{kt}, \mathbf{X}_{kt}, \bar{\alpha}_{kt}, t) = \Lambda(\psi + \lambda \cdot t + \eta \cdot \bar{W}_{kt} + \boldsymbol{\mu} \cdot \mathbf{X}_{kt} + \theta \cdot (\bar{\alpha}_{kt} \cdot t)), \quad (6)$$

where \bar{y}_{kt} is the average annual inflow rate in period t (the years 2005-2006 or 2008-2010) and $\Lambda(z) \equiv \exp(z) / [1 + \exp(z)]$ is the logistic function. The annual inflow rate equals the fraction of the older employees who were awarded a certain type of disability benefit (disability pension or rehabilitation benefit). To be included in the firm's risk set in a given year, the worker had to be employed by the firm for the past two calendar years (but not necessarily during the current year). This definition allows a one-year sickness benefit

¹⁰Estimating the models by the Fixed Effect method leads to the same conclusion.

Table 5: The average partial effect of the degree of experience rating on the average inflow rates to rehabilitation benefits and disability pension

	Model 1		Model 2		Model 3		Model 4	
	APE	z	APE	z	APE	z	APE	z
A. Rehab benefit inflow								
$\bar{\alpha}_{kt} \cdot t$	0.0050	1.29	0.0033	0.74				
$1(\bar{\alpha}_{kt} \cdot t > 0)$					0.0016	0.94		
$1(\bar{\alpha}_{kt} \cdot t \in (0, .15])$							0.0015	0.82
$1(\bar{\alpha}_{kt} \cdot t \in (.15, .5])$							0.0013	0.44
$1(\bar{\alpha}_{kt} \cdot t > .5)$							0.0036	0.77
B. Disab pension inflow								
$\bar{\alpha}_{kt} \cdot t$	-0.0013	-0.51	-0.0039	-1.19				
$1(\bar{\alpha}_{kt} \cdot t > 0)$					-0.0010	0.61		
$1(\bar{\alpha}_{kt} \cdot t \in (0, .15])$							-0.0014	0.84
$1(\bar{\alpha}_{kt} \cdot t \in (.15, .5])$							0.0035	1.18
$1(\bar{\alpha}_{kt} \cdot t > .5)$							-0.0065	1.59
Payroll control:	Linear		Polynomial		Polynomial		Polynomial	

Notes: The models were estimated by maximizing the Bernoulli quasi-likelihood function. The sample includes 664 observations on 332 firms employing at least 5 workers aged 45-61 in each year between 2005 and 2010. $1(A)$ is 1 if A is true and 0 otherwise. All models include industry dummies and the average age of the employees between age 45 and 61. Model 1 also includes the payroll, whereas Models 2 to 4 include the 3rd order polynomial for payroll. Reported estimates are average partial effects (APEs) or average marginal effects. In Model 1 the APE is the partial derivative of the expected inflow rate, averaged across firms subject to experience rating in period 1. In Models 2, 3 and 4 the APE is the difference in the expected inflow rate compared to the case of no experience rating, averaged across firms subject to experience rating in period 1. The z statistics are based on the robust standard errors clustered at the firm level.

period between the worker's exit from work and his or her receipt of a disability benefit. Here \mathbf{X}_{kt} does not only include the industry dummies, but also the average age of the firm's relevant workforce (i.e. employees aged 45 to 61 with at least two years of tenure).

We estimate the model by maximizing the Bernoulli quasi-likelihood function. The resulting estimator is consistent regardless of the conditional distribution of \bar{y}_{kt} . That is, we only specify the conditional expectation but leave the conditional distribution of the outcome variable unspecified. Apart from the logit transformation $\Lambda(\cdot)$ and one additional control variable, the modeling setting and the considered specifications are similar to the linear risk ratio models that were discussed above.

The results are shown in Table 5. We only report the average partial effect (APE) of the experience rating variable, which is defined as the effect on the expected value of \bar{y}_{kt} averaged across the firms subject to experience rating during period 1. That is, we consider the average effect of the experience rating on the disability inflow during the TyEL period among the firms that became exposed to experience rating. When interpreting the

estimates, one should note that the average inflow rates are rather low. For example, during the TyEL period, these are approximately 0.008 and 0.010 for rehabilitation and disability pension benefits, respectively.

Panel A shows the effects on the rehabilitation benefit claims. These effects appear to be positive in all the model specifications, but none of them is statistically significant. The effects on the disability pension inflow (Panel B) are negative except for the firms with $0.15 < \bar{\alpha}_{k1} \leq 0.5$ in the last specification, but they do not differ from zero at the conventional risk levels. These results are not sensitive with respect to our decision to combine partial benefits and full benefits. The distinct models for the inflows into partial and full benefits also did not produce significant effects (not reported here). We thus conclude that the disability inflow rates of the older employees in the TyEL period did not change differently in the firms that became subject to experience rating and in those smaller firms that were not affected by the reform.

6 Marginal costs of disability pension claims

In this section, we examine the economic incentives at the employer-employee level. We define the *marginal cost* of a new disability pension claim as the expected increase in the employer's future DI premium. This can be computed for each worker who is at the risk of becoming disabled. The marginal cost measures the cost of a new disability pension claimant to his or her former employer, and how this cost varies across workers and firms. The distribution of these costs is of obvious interest. In addition, the marginal costs provide an alternative way to examine the potential effects of experience rating. Using the marginal cost as a regressor in various probability models, we test whether the disability cost risk explains job tenure and the transitions from work to disability benefits.

6.1 Computing marginal costs

Let us assume that a disability pension is awarded in year $t \geq 2008$ to worker j who worked in firm k . The marginal cost of this event to the former employer is

$$c_{jkt} = E_t \left[\delta^2 (C_{k(t+2)}^j - C_{k(t+2)}) + \delta^3 (C_{k(t+3)}^j - C_{k(t+3)}) \right], \quad (7)$$

where δ is the annual discount factor, C_{ks}^j is the DI premium in year s given that the pension was awarded to worker j in year t , and C_{ks} is the counterfactual premium, had the worker continued to work without receiving disability benefits until the end of year $t + 3$. At the beginning of year t , the employer does not know the future values of the components of the DI premium formula, which explains the expectation operator in (7). Let us further assume that the employer knows all the historical values, and uses the latest realized values of the payroll and base premium, $W_{k(t-1)}$ and $Q_{k(t-1)}$, to predict their future values. Now, by substituting (2) into (7) and replacing all the future values with

their expected values, we find that

$$c_{jkt} = \alpha(W_{k(t-1)})Q_{k(t-1)} \left(\delta^2 \left[m \left(\hat{r}_{kt}^A + \Delta \hat{r}_{kt}^{Aj}, r_{k(t-1)}^A \right) - m \left(\hat{r}_{kt}^A, r_{k(t-1)}^A \right) \right] + \delta^3 \left[m \left(\hat{r}_{k(t+1)}^A, \hat{r}_{kt}^A + \Delta \hat{r}_{kt}^{Aj} \right) - m \left(\hat{r}_{k(t+1)}^A, \hat{r}_{kt}^A \right) \right] \right), \quad (8)$$

where \hat{r}_{kt}^A and $\hat{r}_{k(t+1)}^A$ are the expected counterfactual risk ratios, had the worker continued to work until the end of year $t + 3$, and $\Delta \hat{r}_{kt}^{Aj}$ is the expected change in the risk ratio that would result from worker j 's disability pension claim in year t .¹¹ In what follows, we set $\delta = 1/1.03$ but compute all the other parameters in (8) from the data. Here we discuss the most important aspects of these computation. The details are provided in the Appendix.

It is evident that the key parameter is the expected change in the risk ratio, $\Delta \hat{r}_{kt}^{Aj}$. A disability pension awarded in year t may or may not increase the numerator of the employer's risk ratio in that year, depending on the year of the retirement event (the year when the disabling illness was diagnosed). If the year of the retirement event is 2009 or later, the expected increase in the numerator equals \hat{e}_{jt} , the expected present value of disability pension benefits until the age of 63. If the year of the retirement event is 2008, only a part of \hat{e}_{jt} (we assume one-half) is accounted for when determining the employer's risk ratio. And if the year of the retirement event is 2007 or earlier, the disability event has no effect on the employer's risk ratio. We estimate \hat{e}_{jt} using data on the disability pension benefits of the workers who retired at different ages. Note that the year of the retirement event is a latent variable as it is only observed if the disability pension was actually awarded in year t . However, we need to compute c_{jkt} for all the individuals. To address this issue, we assume that with a certain probability the impact of a new disability pension on the risk ratio is only partial (corresponding the case when the year of the retirement event is 2008) or non-existent (the case when the retirement event is before the year 2008). We estimate these probabilities using the observed differences in the years of the retirement event and the pension receipt among those who were awarded a disability pension. Let $\hat{\phi}_t$ denote the probability that \hat{e}_{jt} is assigned to the former employer (taking into account the case when only one-half of \hat{e}_{jt} is assigned to the employer) if a disability pension is awarded to worker j in year t . Now, we can write

$$\Delta \hat{r}_{kt}^{Aj} = \frac{\hat{\phi}_t \hat{e}_{jt}}{\hat{R}_{kt}}, \quad (9)$$

where \hat{R}_{kt} is the expected theoretical disability cost and $\hat{\phi}_t \hat{e}_{jt}$ is the expected realized cost of a new disability pension. The expected change in the risk ratio decreases with age (as \hat{e}_{jt} decreases with age) and firm size (as \hat{R}_{kt} increases with firm size), and increases with

¹¹We assume that the disability pension awarded to worker j in year t does not affect $\hat{r}_{k(t+1)}^A$. In principle, it may have an effect through a change in the age and wage structure of the workforce affecting the theoretical disability cost and calculatory term (see the formula in the Appendix). Nonetheless, such an effect is negligible for large firms, while it is of no consequence for the smaller firms that are not subject to a notable degree of experience rating.

time (as $\hat{\phi}_t$ increases with t).

It is not obvious what we should assume about the expected counterfactual risk ratios, \hat{r}_{kt}^A and $\hat{r}_{k(t+1)}^A$. The latest realized value, $r_{k(t-1)}^A$, would be a problematic choice. It ignores the gradually diminishing role of the calculatory term in the risk ratio formula, and it would be a particularly poor predictor for relatively small firms because disability events among their employees are rare and their risk ratios are volatile. The risk ratio in the small firm soars at a high level in the year when the disability pension is awarded, but usually declines in the next year. Instead of using the latest realized value, we use the collective risk ratio of all LEL employers (the average risk ratio computed by the pension provider), which we then adjust for the transitional provisions.¹² The resulting estimates of \hat{r}_{kt}^A and $\hat{r}_{k(t+1)}^A$ are independent of the firm's own disability experience, which enhances the identification of the causal effects of the marginal costs in the econometric analysis.

We have computed two versions of the marginal costs using different versions of the experience multiplier m . The first is based on the step-function version of the experience multiplier (the solid line in Figure 1), while the second is based on the smoothed function (the dashed line). In the former case, the marginal cost is often zero for the employees of large firms, as the impact of a new pension claim on the risk ratio, $\Delta\hat{r}_{kt}^{Aj}$, is typically too low to raise m . This feature is problematic and unrealistic for two reasons. First, a disability pension for one worker increases the probability that the next disability event in the same year will increase m . The rational employer should take this into account. Second, the employer does not know the relevant risk ratios.¹³ The employer probably knows the current value of m and perhaps its determinants $r_{k(t-2)}$ and $r_{k(t-3)}$, but does not know the risk ratios that determine the value of m in years $t+2$ and $t+3$. Therefore the employer cannot accurately assess whether or not a new pension claim made by a given worker this year will increase the future m . Due to this uncertainty about the number of new disability pension claims within the year and about the future values of the risk ratios, the employers should worry each possible disability case and this should also be reflected in our incentive measure.

Additional problem that arises with the step-function version of m is that the marginal costs become very sensitive with respect to the assumed values of \hat{r}_{kt}^A and $\hat{r}_{k(t+1)}^A$. To illustrate this, let us consider a large firm, so that $\Delta\hat{r}_{kt}^{Aj}$ is very small for all its employees. Now, suppose that the average of \hat{r}_{kt}^A and $\hat{r}_{k(t+1)}^A$ (or that of \hat{r}_{kt}^A and $r_{k(t-1)}^A$) is only marginally below the threshold value after which m rises to the next level. In this case, a new disability pension claim, despite its modest effect on the risk ratio, invariably raises m , resulting in high marginal costs for all employees of the firm (as the effect of the increase in m is proportional to the payroll). However, with slightly different values assumed for \hat{r}_{kt}^A and $\hat{r}_{k(t+1)}^A$, this situation may not happen in which case the marginal cost is zero for

¹²Another alternative was considered where the counterfactual risk ratios are based on the firm's own three-year disability history. However, the results were not altered when these counterfactual risk ratios were used.

¹³The employer may ask the past values of its risk ratios from the pension provider. However, the marginal cost depends on the future risk ratios, which are not yet defined at the beginning of year t .

all these employees. Because there are no obvious values for these risk ratios, the marginal costs based on the step-function version of m are essentially arbitrary.

For these reasons we prefer the smoothed version of m to the step-function version. The use of the smoothed version provides a less extreme alternative by placing positive values for many otherwise zero marginal costs and by cutting extremely high values of marginal costs. A major advantage of this approach is that the marginal costs are not sensitive with respect to the assumed values of \hat{r}_{kt}^A and $\hat{r}_{k(t+1)}^A$. In fact, the only role of these risk ratios is to restrict the value of m to the interval $[0.1, 5.5]$, which boundaries are rarely binding for large firms. It follows that the marginal cost is determined primarily by a worker's age, firm size and time period. By using the smoothed version of m , we mimic the situation in which the employer knows its current m and the likely increase in the risk ratio, $\Delta\hat{r}_{kt}^{Aj}$, but is uncertain of the counterfactual risk ratios and hence the distance to the threshold value for the next level of m . In the subsequent analysis, we report only the results for the marginal costs that are based on the smoothed version of m .

6.2 Descriptive evidence

Figure 5 shows the cumulative distribution function of marginal costs for the employees of experience-rated firms during the last three years of the observation period (before 2008 the marginal costs are zero by construction). In 2008, there are no considerably high values, and the mass of the marginal costs are close to zero. This was expected since a disability pension awarded in 2008 affects the employer's risk ratio only if the retirement event occurred in the same year, which probability is about 0.17. Even in that case, the effect is relatively small because one-half of the present value of the disability cost is ignored and pooled among all the employers. In the next two years, both the average level and the dispersion of the marginal costs increase, and in many cases the marginal costs are of considerable size.

Figure 6 depicts the average marginal cost by age and α category in 2010 when the transitional provisions no longer played a role anymore but when the employers' risk ratios were solely determined by their own disability expenditures. Not surprisingly, the average costs are rather low in relatively small firms with $0 < \alpha \leq 0.25$. The average level is almost threefold for the next size group with $0.25 < \alpha \leq 0.5$, but it is considerably higher for the two largest groups. Furthermore, the average marginal costs of workers around age 50 in the firms with $\alpha > 0.5$ are comparable in size to the annual salary. In all cases, the average marginal cost begins to decline near the age of 54.¹⁴

Overall, there is considerable variation in the marginal costs across individuals and firms at least during the years 2009 and 2010. Often these marginal costs are also rather high, which is why the employers should respond to experience rating. Yet no evidence of such behavioral effects were found in the analysis of the risk ratios and the inflow rates in Section 5. In the following subsection, we estimate the effect of the marginal

¹⁴This is due to the age patterns of the benefit level and multiplier γ , which jointly determine the present value of a new pension claim (see the Appendix).

Figure 5: The cumulative distribution function of the marginal costs for the employees of firms with $\alpha > 0$ in the years 2008, 2009 and 2010

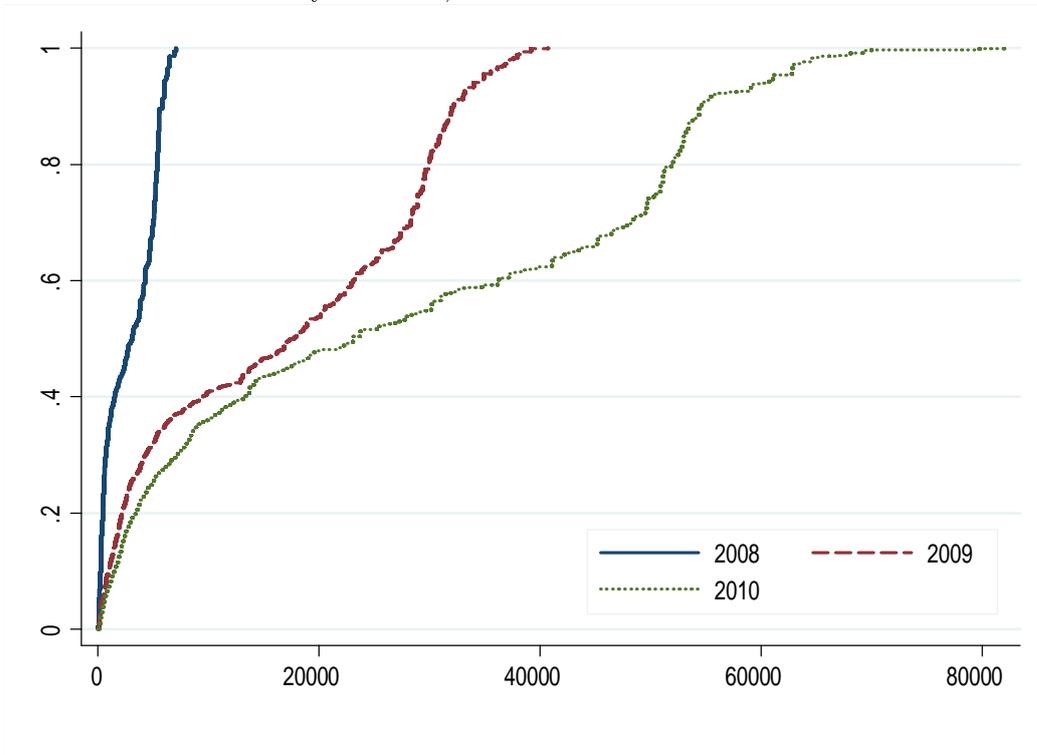
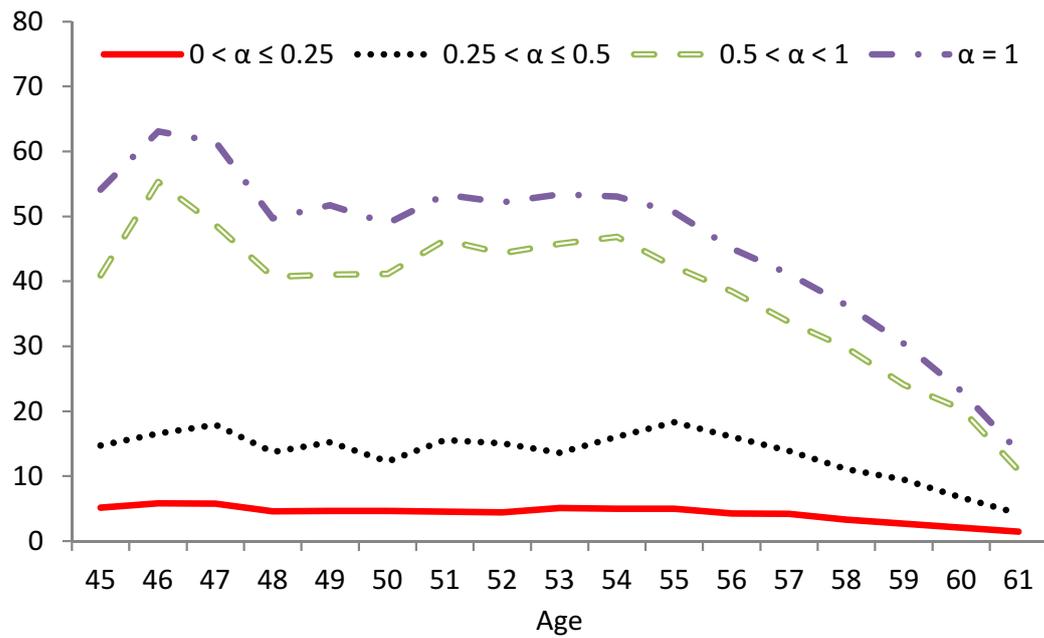


Figure 6: Average marginal cost (€1000) as a function of age by α group in the year 2010



cost on the transition rates out of work to different exit destinations. The larger number of observations and within-firm variation in the cost measure allow us to estimate more accurately the effects of the experience rating than is possible in the firm-level analysis. With the individual-level data, we can also examine whether the effect of experience rating varies with age and job tenure.

6.3 Effects of marginal costs

Consider a worker who has worked in a given firm for the past two calendar years. In the current year, the worker may still be employed in the same firm, may have stopped working due to a disability, in which case he or she receives either rehabilitation or disability pension benefits,¹⁵ or this worker may have left the firm without becoming a disability benefit recipient. To test whether experience rating affects these outcomes, we apply a standard logistic regression. Specifically, we estimate a number of the binary logit models of the form

$$\ln \left(\frac{p_{jkt}^s}{1 - p_{jkt}^s} \right) = \lambda_t + \boldsymbol{\mu} \cdot \mathbf{X}_{jkt} + \theta \cdot c_{jkt}, \quad (10)$$

where p_{jkt}^s is the probability that worker j occupies state s in year t , given that he or she worked in firm k in years $t - 1$ and $t - 2$. The vector of control variables \mathbf{X}_{jkt} includes the worker's age x_{jt} , lagged payroll $W_{k(t-1)}$, job tenure and the industry dummies. The variable of primary interest is the marginal cost, c_{jkt} , which we defined above. In some models, we also add interactions of c_{jkt} with the age and job tenure.

The marginal cost varies along with a number of dimensions. To make this explicit we write

$$c_{jkt} = g_t(x_{jt}, W_{k(t-1)}, Q_{k(t-1)}, Q_{k(t-2)}, Q_{k(t-3)}; \boldsymbol{\Omega}_t),$$

where $\boldsymbol{\Omega}_t$ denotes the set of the pension system parameters and \hat{e}_{jt} . The marginal cost is a function of worker's age (through the expected present value of a disability pension \hat{e}_{jt}), the lagged payroll (through α) and the lagged base premiums (through the expected calculatory term and theoretical costs), all of which vary over time. In addition, the underlying technical parameters $\boldsymbol{\Omega}_t$ change annually. However, the most significant point is that $g_t(\cdot)$ changes over time, which indicates that the mapping of the inputs of $g_t(\cdot)$ into the marginal cost changes over time. This occurs due to the gradual implementation of experience rating, which is captured by the increasing probability of employer's liability for the costs of a new disability pension (i.e. $\hat{\phi}_t$ and $\hat{\phi}_{t+1}$) and by the diminishing role of the calculatory term in the risk ratio formula. It should be emphasized that, apart from the worker's age and the lagged payroll (if interpreted as a proxy for firm size), the inputs of $g_t(\cdot)$ cannot have a direct effect on the outcome probability.

Because c_{jkt} only depends on the worker's age, firm characteristics, system parameters

¹⁵Due to the small numbers of transitions to partial benefits, we do not make a distinction between full and partial rehabilitation benefits, nor do we differentiate between full and partial disability pension benefits. Our results are not sensitive with respect to this.

and time, it is independent of unobserved individual characteristics such as health and preferences for work. We should not be concerned about unobserved firm characteristics either. As the collective risk ratio of the LEL employers was used for the expected counterfactual risk ratios, c_{jkt} is independent of firm-specific shocks that affect the firm's past and current disability inflow. Conditional on the age and lagged payroll, the variation in c_{jkt} stems from the gradual implementation of experience rating, i.e. changes in $g_t(\cdot)$, and to some extent from changes in the lagged base premiums and from the changes in the technical parameters over time. Since these sources of variation are exogenous from the perspectives of both the employer and employee, we can estimate the causal effect of c_{jkt} from the model (10).

Using the data pooled over the years 2006-2010, we have estimated several model variants. For the sake of brevity, we focus on the effects of the marginal cost, which are reported in Table 6. In addition to the coefficients on c_{jkt} and its possible interactions, we report the average partial effect (APE), which is defined as the average effect of c_{jkt} on p_{jkt}^s (accounting for the possible interactions with age and/or job tenure) across the employees of firms with $\alpha > 0$ in 2010. This gives the average response after the transition period on workers whose employers were subject to experience rating. The reported z statistics are based on the robust standard errors accounting for heteroscedasticity and cross-section dependence within firms.¹⁶

The results in Panel A are from the baseline specification that is outlined in (10). These estimates indicate no effect on the probability of being awarded a disability benefit because the APE on the receipt of both types of disability benefits is zero up to the third decimal place. On the other hand, the effects on the probability of continuing to work in the current firm and on that of other exits are statistically significant at the 5% and 10% levels, respectively. These imply that an increase of €10,000 in the marginal costs among all the employees whose employers were subject to experience rating in 2010 would raise the average exit rate out of work for reasons other than disability by 0.005. This means a less than 10% increase in the exit probability for the average worker (see Table 3). This would also reduce the average probability of remaining employed in the current firm for one additional year by 0.006. This is a very small effect in relative terms because the average probability of remaining employed in the current firm is as high as 0.93.

In Panels B, C and D, we allow the effect of the marginal cost to change with age and/or job tenure. Adding these interactions does not change the average effect on the disability benefit claims as the APE remains robustly zero for both types of benefits. However, the age interaction on the probability of claiming a rehabilitation benefit is significant at the 10% level (Panels B and D), indicating that experience rating may have some positive effect

¹⁶For the corresponding linear probability models, we computed one-way clustered (i.e. clustering on firm *or* individual) and two-way clustered (i.e. clustering on both firm *and* individual) covariance-variance matrices using the method by Cameron et al. (2011). Comparison of these estimates suggested that ignoring the cross-sectional dependence in the unobservables would lead to clearly downward-biased standard errors. By contrast, we found no evidence of significant time-dependence in the unobservables for the same individual. Thus it suffices to cluster the standard errors at the firm level.

Table 6: The effects of marginal costs in logit models

	Outcome:							
	Stay employed		Rehab benefit		Disab pension		Other exit	
	Estimate	z-value	Estimate	z-value	Estimate	z-value	Estimate	z-value
A. Baseline model								
MC	-0.0629**	-2.11	0.0295	0.77	0.0012	0.02	0.0629*	1.85
APE	-0.0062**	-2.02	0.0003	0.77	0.0000	0.02	0.0053*	1.79
B. Age interaction								
MC	-0.0407	-1.07	0.0771	1.52	0.0031	0.05	0.0426	0.90
MC×(Age - 55)	0.0043	1.19	0.0132*	1.68	0.0112	0.61	-0.0037	-1.07
APE	-0.0049	-1.45	0.0006	1.25	0.0003	0.36	0.0044	1.25
C. Tenure interaction								
MC	-0.0490*	-1.76	0.0275	0.72	0.0017	0.03	0.0458	1.30
MC×(Tenure - 8)	0.0089**	1.97	0.0041	1.19	-0.0010	-0.19	-0.0089*	-1.90
APE	-0.0056**	-2.19	0.0003	0.84	0.0000	0.00	0.0049*	1.86
D. Age and tenure interactions								
MC	-0.0330	-0.90	0.0746	1.45	0.0036	0.06	0.0316	0.67
MC×(Age - 55)	0.0032	1.03	0.0131*	1.65	0.0112	0.61	-0.0027	-0.87
MC×(Tenure - 8)	0.0086**	2.00	0.0040	1.14	-0.0009	-0.18	-0.0087*	-1.92
APE	-0.0047	-1.59	0.0006	1.31	0.0003	0.36	0.0043	1.36
E. Level dummies								
1 (0 < MC ≤ 1)	-0.0333	-0.43	-0.0792	-0.52	0.0790	0.48	0.0485	0.54
1 (1 < MC ≤ 2)	-0.0966	-0.80	0.3594	1.41	0.3875	1.31	0.0199	0.15
1 (2 < MC ≤ 3)	-0.1703	-1.30	-0.1778	-0.61	0.1173	0.41	0.2281	1.45
1 (3 < MC ≤ 4)	-0.3017**	-2.11	0.3434	1.11	-0.0095	-0.02	0.3033**	1.98
1 (4 < MC ≤ 5.5)	-0.2224	-1.02	0.0593	0.20	-0.2139	-0.63	0.2556	1.03
1 (MC > 5.5)	-0.3894**	-2.24	0.3617	1.03	-0.0133	-0.01	0.3343*	1.81
APEs:								
1 (0 < MC ≤ 1)	-0.0030	-0.42	-0.0007	-0.52	0.0006	0.48	0.0037	0.54
1 (1 < MC ≤ 2)	-0.0089	-0.79	0.0038	1.30	0.0036	1.22	0.0015	0.14
1 (2 < MC ≤ 3)	-0.0161	-1.28	-0.0015	-0.63	0.0010	0.40	0.0186	1.42
1 (3 < MC ≤ 4)	-0.0298**	-1.96	0.0036	1.00	-0.0001	0.00	0.0254*	1.89
1 (4 < MC ≤ 5.5)	-0.0214	-0.98	0.0005	0.20	-0.0015	-0.67	0.0211	0.99
1 (MC > 5.5)	-0.0396**	-2.12	0.0039	0.93	-0.0001	0.00	0.0283*	1.77

Notes: Marginal cost (MC) was computed using the smoothed experience multiplier, measured in 2010 euros and divided by 10,000. All models includes age dummies, industry dummies, and the 3rd order polynomials for lagged payroll and job tenure. $1(A)$ is 1 if A is true and 0 otherwise. All models were estimated by Maximum Likelihood. Unless otherwise stated, the estimates are the coefficients of marginal cost and its interactions with age/tenure. In Panels A to D, the average partial effect (APE), or average marginal effect, is the partial derivative of the outcome probability accounting for possible interactions, averaged across the distribution of the covariates for workers employed in experience-rated firms in 2010. In Panel E it is the similarly averaged discrete change in the probability compared to the case of zero marginal cost. The z -values are based on the robust standard errors clustered at the firm level. Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

for the oldest workers. In Panels C and D, the tenure interaction on the probability of other exits and on that of remaining at work is statistically significant at least at the 10% level. Thus, for a recently hired worker, the experience rating may increase the probability of other exits and, consequently, decrease the probability of continuing employment in the current firm. As an example, for a worker with two years of tenure, an increase of €10,000 in the marginal cost increases the probability of other exits by 0.016 (z -value 3.15) or 0.014 (z -value 2.59) according to the models in Panel C and D, respectively. It is possible that the experience rating induces the employer to get rid of recently hired workers who develop health problems. This effect may not exist for more tenured workers who are better protected due to the seniority rules.

The models considered so far imply a constant effect on the odds ratios throughout the range of marginal cost. To relax this assumption, we replaced the continuous marginal cost variable by a set of dummy variables in the models reported in Panel E. The reference category is no experience rating, corresponding to the case $c_{jkt} = 0$. As in other models, there is no effect on the likelihood of becoming a disability benefit recipient. However, the experience rating affects other outcomes, but only through the relatively large values of marginal costs. The probability of other exits and that of remaining employed in the current firm do not differ from the reference level until the marginal cost exceeds €30,000. After this point, the probability of other exits (remaining employed) increases (decreases). These high values of marginal costs increase the probability of other exits by 0.02 to 0.03, and decrease the likelihood of staying employed in the current firm by 0.03 to 0.04 on average. The small difference between these effects appears to be due to the increased probability of the rehabilitation benefit claim, but the estimated increase is very imprecise and therefore needs to be interpreted with great caution. Note that the effects of the largest marginal costs on the probability of other exits are rather large also in relative terms, albeit they are significant only at the 10% level.

Due to a small number of partial benefit claims, the transitions were combined into a single outcome comprising partial and full benefits of a given type. This may conceal the true effect of the marginal cost if experience rating reduces the inflow to full benefits but increases the inflow to partial benefits. This is a relevant concern at least in the case of disability pension claims. If one of the two disability pensions is to be claimed, the employer has an incentive to induce the employee to claim a partial pension instead of the full pension because the partial pension has a smaller impact on the risk ratio of the firm. For this reason, we also estimated distinct models for full and partial benefits. Without reporting the detailed results here, we emphasize that these models do not produce statistically significant effects, and thereby the results in Table 6 are not sensitive with respect to our decision to combine partial and full benefits.

As a robustness check, we also estimated similar linear probability models. Table 7 presents these results for the baseline specification (Panel A) and the model with the marginal cost dummies (Panel B). The OLS estimates can be compared to the APEs from the logit models in Panels A and E in Table 6, albeit one should keep in mind that the

Table 7: The effects of marginal costs in linear probability models

	Outcome:							
	Stay employed		Rehab benefit		Disab pension		Other exit	
	Coeff	<i>t</i> -value	Coeff	<i>t</i> -value	Coeff	<i>t</i> -value	Coeff	<i>t</i> -value
A. Baseline model								
MC	-0.0058	-1.56	0.0003	0.73	0.0001	0.33	0.0054	1.45
B. Level dummies								
1 ($0 < MC \leq 1$)	-0.0030	-0.45	-0.0007	-0.52	0.0006	0.42	0.0030	0.47
1 ($1 < MC \leq 2$)	-0.0115	-1.07	0.0037	1.26	0.0050	1.36	0.0029	0.30
1 ($2 < MC \leq 3$)	-0.0137	-1.07	-0.0016	-0.65	0.0007	0.31	0.0146	1.12
1 ($3 < MC \leq 4$)	-0.0304*	-1.65	0.0035	1.01	0.0000	0.01	0.0269	1.54
1 ($4 < MC \leq 5.5$)	-0.0175	-0.88	0.0005	0.18	0.0001	0.07	0.0168	0.87
1 ($MC > 5.5$)	-0.0414	-1.55	0.0030	0.91	0.0017	0.73	0.0367	1.45
C. Baseline model with firm fixed effects								
MC	-0.0025	-1.17	0.0006**	2.01	-0.0001	-0.15	0.0019	0.88
D. Level dummies with firm fixed effect								
1 ($0 < MC \leq 1$)	-0.0011	-0.15	0.0006	0.30	-0.0010	-0.53	0.0015	0.22
1 ($1 < MC \leq 2$)	-0.0079	-0.64	0.0053*	1.66	0.0039	0.87	-0.0012	-0.12
1 ($2 < MC \leq 3$)	-0.0073	-0.89	-0.0001	-0.04	-0.0005	-0.16	0.0079	0.96
1 ($3 < MC \leq 4$)	-0.0166	-1.46	0.0064*	1.80	-0.0007	-0.24	0.0109	1.20
1 ($4 < MC \leq 5.5$)	-0.0048	-0.27	0.0023	0.90	-0.0014	-0.62	0.0039	0.24
1 ($MC > 5.5$)	-0.0113	-0.71	0.0078**	2.34	-0.0005	-0.17	0.0040	0.26

Notes: Marginal cost (MC) was computed using the smoothed experience multiplier, measured in 2010 euros and divided by 10,000. All models includes age dummies, industry dummies, and the 3rd order polynomials for lagged payroll and job tenure. $1(A)$ is 1 if A is true and 0 otherwise. The reference category for the level dummies is $MC = 0$. The models in Panels A and B were estimated by Ordinary Least Squares, and those in Panels C and D by the Fixed Effects method. The t statistics are based on the robust standard errors clustered at the firm level. Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

APEs describe the average effects for a given subgroup (the employees of experience-rated firms in 2010). By and large, the OLS coefficients are similar to the corresponding APEs in Table 6, having always the same sign and being of the same magnitude. These OLS estimates, however, are less precise: Out of the 28 OLS coefficients in Panels A and B, only 1 is (marginally) statistically significant at the 10% level.

In Panels C and D, we report the results for the linear probability models when estimated by the Fixed Effects (FE) method. The FE results do not imply any effect on the risk of becoming a disability pension recipient, which is in line with the OLS and logit estimates. But, contrary to the OLS and logit estimates, some of the FE estimates on the receipt of the rehabilitation benefit are statistically significant. In Panel C, the implied effect is rather small: An increase of €10,000 in the marginal cost would raise the probability of being awarded a rehabilitation benefit by 0.006, corresponding to a relative increase

of about 7% in the average rehabilitation claim rate of 0.0091 (see Table 3). There are greater effects for the three marginal cost categories in Panel D, but only one of them is statistically significant at the 5% level and the sizes of the effects do not exhibit any clear pattern. The FE estimates on the probabilities of other exits and remaining employed in the current firm are smaller in absolute value than the corresponding OLS estimates and APEs from the logit models. Moreover, none of these estimates are statistically significant at the conventional risk levels.

To conclude, our results suggest that experience rating has no impact on the likelihood of being awarded a disability pension. This finding is highly robust, as the same conclusion can be drawn from various specifications of the linear probability and logit models. While this result is counter evidence for the ultimate goal of experience rating, it is in accordance with our analysis of the risk ratios and inflow rates in Section 5.

Our other findings are less robust. In the logit analysis we found that the experience rating may have a negative effect on a worker's job tenure. This is because the probability of a worker leaving his or her current firm without claiming disability benefits increases with the marginal cost. This effect appears to be pronounced for recently hired employees, and it is driven by the relatively high values of the marginal costs. However, the result is only evident in the logit analysis as the corresponding effects lose their statistical significance in the linear probability models.

Among the oldest workers, the experience rating may slightly increase the probability of being awarded a rehabilitation benefit. This conclusion should be interpreted with caution, however, as the underlying age interaction is estimated rather imprecisely in the logit and the FE models (significant at the 10% level) while the OLS estimate is not significant at all (these linear probability model results were not reported). In addition, the FE models without the age interactions also indicate some positive effects, but these effects do not appear in the corresponding logit and OLS specifications. Despite these weak effects in some model specifications, the estimated effects are predominantly zero. Taken together with our fractional logit results that indicate no effect on the rehabilitation benefit inflow, this leads us to conclude that experience rating has a negligible effect on the transitions to rehabilitation benefits.

7 Concluding remarks

This study contributes to understanding better the efficiency of experience rating as a disability prevention device. Our calculations of the marginal costs demonstrate that a new disability pension claimant can cause substantial cost to the former employer through an increase in the DI premium. Given the size of the potential costs, the experience rating should promote preventive health and safety practices, and encourage employers to keep their employees with work limitations at work.

However, our results suggest that experience rating is not succeeding in reducing disability benefit claims. We did not find declines in the post-reform risk ratios or in the

disability inflow rates for the large firms that became exposed to experience rating. We also did not find evidence that the marginal cost would have reduced the individual-specific transition rates to disability pension benefits. The only possible effects of the marginal cost (yet not very robust) were an increase in the likelihood of leaving the current employer for reasons other than disability and an associated decrease in job tenure. Thus, experience rating does not have desired effects on disability benefit claims and employment among older workers.

From the description of the institutional framework and our marginal cost calculations it is evident that the design of the Finnish experience rating scheme is rather complex. A long delay, of possibly several years, between a medical diagnosis of disability and a possible increase in the experience-rated premium may hinder employers from recognizing the causes for premium changes. The transitional provisions of the TyEL reform further complicated the assignment of the disability pension costs to the former LEL employers. It might be that many employers were unable to respond to economic incentives because they were unaware of the details of the new experience rating system. If so, a desired effect of experience rating may become apparent only after a longer period of time when the employers had better understood how the complex DI system works and have gradually experienced the impact of their own claim history on the DI premiums.

It should be emphasized that we considered a rather special group of workers and employers. The former LEL sector covered workers in short-term employment. If the effect of the experience rating varies by sector, our pessimistic results may not apply to other sectors. Due to the short-term employment contracts, the former LEL employers may have both weaker incentives and weaker opportunities to influence their employees' health and well-being. The short-term contracts may also offer employers a better opportunity to avoid cost liabilities by laying off the workers with a high disability risk. These factors may have diluted the effect of the experience rating among the former LEL employers.

For the reasons discussed above, it may be a slightly premature to conclude that the experience rating in DI has no effect whatsoever. And even without the desired behavioral effects, the experience rating system still provides a means of allocating the overall costs of the disability benefits more equitably among individual employers.

References

- [1] Anderson, P., M., and B. D. Meyer (1993). "Unemployment insurance in the United States: Layoff incentives and cross subsidies", *Journal of Labor Economics* 11, S70-S95.
- [2] Anderson, P., M., and B. D. Meyer (2000). "The effects of the unemployment insurance payroll tax on wages, employment, claims and denials", *Journal of Public Economics* 78, 81-106.

- [3] Autor, H., D., and M. G. Duggan (2003). "The rise in the disability rolls and the decline in unemployment", *Quarterly Journal of Economics* 118, 157-205.
- [4] Autor, H., D., and M. G. Duggan (2006). "The growth in the social security disability rolls: A fiscal crisis unfolding", *Journal of Economic Perspectives* 20, 71-96.
- [5] Autor, H., D. (2011). "The unsustainable rise of the disability rolls in the United States: Causes, consequences, and policy options", NBER Working Paper 17697.
- [6] Benítez-Silva, H., R. Disney, and S. Jiménez-Martín (2010). "Disability, capacity for work and the business cycle: an international perspective", *Economic Policy* July 2010, pp. 483-536.
- [7] Black, D., K. Daniel, and S. Sanders (2002). "The impact of economic conditions on participation in disability programs: Evidence from the coal boom and bust", *The American Economic Review* 92, 27-50.
- [8] Bruce, C., J., and F. J. Atkins (1993). "Efficiency effects of premium-setting regimes under workers' compensations: Canada and the United States", *Journal of Labor Economics* 11, S38-S69.
- [9] Burkhauser, R., and M. Daly (2011). *The declining work and welfare of people with disabilities: What went wrong and a strategy for change*, Washington, D.C.: American Enterprise Institute
- [10] Cameron, A. C., J. B. Gelback, and D. L. Miller (2011). "Robust inference with multiway clustering", *Journal of Business & Economic Statistics* 29, 238-248.
- [11] Campolieti, M. (2004). "Disability insurance benefits and labor supply: Some additional evidence", *Journal of Labor Economics* 22, 863-888.
- [12] García-Gómez, P., H-M. von Gaudecker, and M. Lindeboom (2011). "Health, disability and work: patterns for the working age population", *International Tax and Public Finance* 18, 146-165.
- [13] Gruber, J. (2000). "Disability insurance benefits and labor supply", *The Journal of Political Economy* 108, 1162-1183.
- [14] Hakola, T., and R. Uusitalo (2005). "Not so voluntary retirement decisions? Evidence from a pension reform", *Journal of Public Economics* 89, 2121-2136.
- [15] Karlström, P., M. Palme, and I. Svensson (2008). "The employment effect of stricter rules for eligibility for DI: Evidence from a natural experiment in Sweden", *Journal of Public Economics* 92, 2071-2082.

- [16] Koning, P. (2009). “Experience rating and the inflow into disability insurance”, *De Economist* 159, 315-335.
- [17] Korkeamäki, O., and T. Kyrrä (2012). “Institutional rules, labor demand and retirement through disability programme participation”, *Journal of Population Economics* 25, 439-468.
- [18] Kyrrä, T. (2010). “Early retirement policy in the presence of competing exit pathways: Evidence from policy reforms in Finland”, VATT Working Papers 17.
- [19] Kralj, B. (1994). “Employer responses to workers’ compensation insurance experience rating”, *Industrial Relations* 49, 41-61.
- [20] Krueger, A. B., and J. F. Burton, Jr. (1990). “The employers’ costs of workers’ compensation insurance: Magnitudes, determinants, and public policy”, *The Review of Economics and Statistics* 72, 228-240.
- [21] Meyer, B. D. (2002). “Unemployment and workers’ compensation programmes: Rationale, design, labor supply and income support”, *Fiscal Studies* 23, 1-49.
- [22] 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.
- [23] Ruser, J. W. (1991). “Workers’ compensation and occupational injuries and illness”, *Journal of Labor Economics* 9, 325-350.
- [24] Staubli, S. (2011). “The impact of stricter criteria for disability insurance on labor force participation”, *Journal of Public Economics* 95, 1223-1235.
- [25] Thomason, T., and S. Pozzebon (2002). “Determinants of firm workplace health and safety claims management practices”, *Industrial and Labor Relations Review* 55, 286-307.
- [26] Topel, R. H. (1983). “On layoffs and unemployment insurance”, *The American Economic Review* 73, 541-559.

Appendix: Technical details

Here we discuss the computation of various terms needed for the risk ratios and marginal costs. Some underlying parameter values are shown in Table 8.

Table 8: Some parameters

t	β_t^0	β_t^1	β_t^2	r_t^{LEL}	\underline{W}_t	\overline{W}_t	$\hat{\phi}_t$
2004	0.14	0.63	0.12	1.85	1 500 000	24 000 000	
2005	0.16	0.41	0.33	2.07	1 542 000	24 672 000	
2006	0.56	0.19	0.21	1.38	1 639 146	26 226 336	
2007	0.04	0.64	0.22	1.46	1 803 061	28 848 970	
2008	0.04	0.14	0.49	2.11	2 026 640	32 426 242	0.0863
2009	0.05	0.14	0.26	2.22	2 415 755	38 652 080	0.4329
2010	0.05	0.13	0.24				0.7673
2011							0.8766

Terms for risk ratios

We need to compute the risk ratios for the years 2005-2010. The theoretical disability cost, the denominator of the risk ratio, is calculated as a weighted sum of the base premiums:

$$R_{kt} = \begin{cases} \beta_t^0 Q_{kt} + \beta_t^1 Q_{k(t-1)} + \beta_t^2 Q_{k(t-2)}, & t < 2006 \\ \beta_t^0 Q_{k(t-1)} + \beta_t^1 Q_{k(t-2)} + \beta_t^2 Q_{k(t-3)}, & t \geq 2007, \end{cases} \quad (11)$$

where β_t 's are parameters computed by the pension provider.

For the adjusted risk ratios we also need the calculatory term which is given by

$$\begin{aligned} E_{k08} &= \frac{1}{2} (r_{08}^{LEL} + r_{07}^{LEL}) \cdot (\beta_{08}^1 Q_{k06} + \beta_{08}^2 Q_{k05}), \\ E_{k09} &= \frac{1}{2} (r_{09}^{LEL} + r_{08}^{LEL}) \cdot \beta_{09}^2 Q_{k06}, \\ E_{kt} &= 0, \quad t \geq 2010 \end{aligned} \quad (12)$$

where r_t^{LEL} is the collective risk ratio of all the (former) LEL employers in year t , which is computed by the pension provider.

Expected counterfactual risk ratio in year t

To compute \hat{r}_{kt}^A we need expected values for $\sum_{j \in D_{kt}} e_{jt}$, R_{kt} , and E_{kt} for the years 2008-2010; see (4). We compute the expected theoretical disability cost as

$$\hat{R}_{kt} = \beta_{t-1}^0 Q_{k(t-1)} + \beta_{t-1}^1 Q_{k(t-2)} + \beta_{t-1}^2 Q_{k(t-3)}, \quad (13)$$

which differs from (11) only in that the unknown β_t 's are replaced by their lagged values, which are assumed to be known at the beginning of year t . The expected calculatory term

is obtained by replacing the unknown parameters r_t^{LEL} and β_t 's in (12) with their latest realized values from year $t - 1$, so that

$$\begin{aligned}\hat{E}_{k08} &= r_{07}^{LEL} \cdot (\beta_{07}^1 Q_{k06} + \beta_{07}^2 Q_{k05}), \\ \hat{E}_{k09} &= r_{08}^{LEL} \cdot \beta_{08}^2 Q_{k06}, \\ \hat{E}_{k10} &= 0.\end{aligned}\tag{14}$$

To derive the expected realized disability cost, we first assume that the expected *unadjusted* risk ratio for year t equals r_{t-1}^{LEL} , that is, the risk ratio without the transitional provisions (i.e. without the calculatory term but assigning all realized disability costs, including those associated with the employment relationships that fell in the LEL period, to the former employer) equals the last year's collective risk ratio. In other words, the relative costs of the firm's disability pension claims equals the last year's sectoral average. Under this assumption $r_{t-1}^{LEL} \hat{R}_{kt}$ equals the expected realized disability cost, had the TyEL been in force for a long time. In reality, the employer's accountability for these costs only covers the TyEL years of the underlying employment relationships. To take this into account, we multiply these costs with the probability that the costs of disability pensions awarded in year t would be assigned to the former employer, $\hat{\phi}_t$ (this parameter is defined below). It follows that the expected realized disability cost under the transitional provisions, i.e. the expected value of $\sum_{j \in D_{kt}} e_{jt}$ in the adjusted risk ratio, equals $\hat{\phi}_t r_{t-1}^{LEL} \hat{R}_{kt}$. Now, we obtain

$$\hat{r}_{kt}^A = \frac{\hat{\phi}_t r_{t-1}^{LEL} \hat{R}_{kt} + \hat{E}_{kt}}{\hat{R}_{kt}} = \hat{\phi}_t r_{t-1}^{LEL} + \frac{\hat{E}_{kt}}{\hat{R}_{kt}}.\tag{15}$$

Expected counterfactual risk ratio in year $t + 1$

In an analogous way to \hat{r}_{kt}^A we compute

$$\hat{r}_{k(t+1)}^A = \hat{\phi}_{t+1} r_{t-1}^{LEL} + \frac{\hat{E}_{k(t+1)}}{\hat{R}_{k(t+1)}},\tag{16}$$

where the expected unadjusted risk ratio for year $t + 1$ is assumed to be equal to r_{t-1}^{LEL} . To obtain $\hat{R}_{k(t+1)}$ and $\hat{E}_{k(t+1)}$, we simply replace the parameters values in (4) and (12) that are not known at the beginning of year t (these also include Q_{kt}) by their latest realized values from year $t - 1$. In doing so, we get

$$\hat{R}_{k(t+1)} = (\beta_{t-1}^0 + \beta_{t-1}^1) Q_{k(t-1)} + \beta_{t-1}^2 Q_{k(t-2)}\tag{17}$$

and

$$\begin{aligned}
\hat{E}_{k09} &= r_{07}^{LEL} \cdot \beta_{07}^2 Q_{k06}, \\
\hat{E}_{k10} &= 0, \\
\hat{E}_{k11} &= 0.
\end{aligned} \tag{18}$$

The present value of a new disability pension claim

The present value is estimated as $\hat{e}_{jt} = \gamma(x_{jt})\hat{b}_{jt}$, where γ is a decreasing function of the age at which the disability pension is awarded, x_{jt} , and \hat{b}_{jt} is the estimated annual disability pension benefit. The multiplier γ is used by all pension providers. It depends on the average duration of disability pension receipt of persons who were awarded a disability pension at a given age. It also accounts for the (average) probability that the recipient returns to work and the (average) survival probability until the old-age pension. For the age group included in our analysis, γ takes values from 10.6 at age 45 to 1.4 at age 61.

To obtain \hat{b}_{jt} we regress (earnings-related) disability pension benefits on age and linear time trend, using the data on employees who were insured under the LEL and were awarded a disability pension between the ages of 45 and 61. We use only the data from the LEL period because for the pension benefits awarded in the TyEL period it is difficult to distinguish the benefits accrued from employment spells covered by the LEL scheme from other pension benefits. Since we do not make a distinction between a partial and full disability pension, the predicted pension benefit equals the expected value over the two benefit types, suggesting that the likelihood of being awarded either a partial or a full disability pension is implicitly accounted for.

Employer's accountability for new pension claims

We estimate $\hat{\phi}_t$ using average differences in the years of the retirement event and disability pension receipt among all new pension recipients who worked under the LEL scheme. Let φ_s be the fraction of disability pensions with difference s between the year of the retirement event and the year when the disability pension was awarded. Using data on all disability pensions awarded in the period 2007-2010 for the former LEL employees, these values are: $\varphi_0 = 0.1725$, $\varphi_1 = 0.5208$, $\varphi_2 = 0.1479$ and $\varphi_3 = 0.0707$. We define

$$\begin{aligned}
\hat{\phi}_{2008} &= 0.5\varphi_0, \\
\hat{\phi}_{2009} &= \varphi_0 + 0.5\varphi_1, \\
\hat{\phi}_{2010} &= \varphi_0 + \varphi_1 + 0.5\varphi_2, \\
\hat{\phi}_{2011} &= \varphi_0 + \varphi_1 + \varphi_2 + 0.5\varphi_3.
\end{aligned} \tag{19}$$

Note that the underlying assumption is that one-half of the costs of a disability pension awarded in 2008 is assigned to the former employer, that is, one-half of the present value of such a pension is accounted for when determining the employer's risk ratio.

As an example, let us consider the allocation of the costs of disability pensions that were awarded in 2010 for individuals who worked for a given employer. The employer is fully liable for the costs of those pensions which retirement event was either in 2010 or 2009 (expected shares of φ_0 and φ_1 , respectively), and partially liable for those with the retirement event in 2008 (expected share of $0.5\varphi_2$). The rest of the costs (expected share of $1 - \varphi_0 - \varphi_1 - 0.5\varphi_2$) are pooled as the underlying employment relationships took place entirely in the LEL period.