Labour Economics 41 (2016) 304–317

Contents lists available at ScienceDirect

Labour Economics

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

Assessing the effects of disability insurance experience rating. The case of The Netherlands

Nynke de Groot[a,*], Pierre Koning[b,c,d,e]

[a] VU Amsterdam, Department of Economics, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
[b] Leiden University, Leiden, The Netherlands
[c] VU Amsterdam, The Netherlands
[d] IZA, Bonn, Germany
[e] Tinbergen Institute, Amsterdam, The Netherlands

HIGHLIGHTS

- We study the effects of disability insurance (DI) experience rating in The Netherlands.
- Our analysis exploits the removal of experience rating for small firms in 2003–2004.
- The removal of experience rating caused an increase of DI inflow of about 7%.
- DI outflow decreased by 12% as a result of the reform.
- These effects seemed to be confined to the first year of DI benefit receipt.

ARTICLE INFO

Article history:
Received 30 August 2015
Received in revised form 28 April 2016
Accepted 10 May 2016
Available online 13 May 2016

JEL classification:
H22
I12
C23

Keywords:
Disability insurance
Experience rating
Differences-in-differences

ABSTRACT

Experience-rated disability insurance (DI) premiums are often advocated as a means to stimulate firms to reduce DI inflow and increase DI outflow. To assess the size of these intended effects of experience rating, this study provides an empirical analysis of the effects of DI experience rating in The Netherlands. We use a difference-in-difference approach with administrative matched firm- and worker data that exploits the removal of experience rating for small firms in 2003 and 2004. According to our results, removing experience rating caused an increase of DI inflow of about 7% for small firms, while DI outflow decreased by 12% as a result of the reform. We argue that these effects were largely confined to the first year of DI benefit receipt.

1. Introduction

According to the literature, one of the most important conditions for preventing work disability is that workers should receive timely interventions and work adaptations (OECD, 2010). In this respect, firms can play a key role by facilitating the return to work from sickness (Autor and Duggan, 2010). Setting disability insurance (DI) premiums that are experience rated may therefore be effectively increase the awareness of firms concerning DI benefit costs, which eventually will reduce the number of DI beneficiaries. Even so, the literature on the effects of experience rating is limited (Tompa et al., 2012).

In this context, The Netherlands provides an interesting setting in which to study the effects of experience rating. After DI enrollment peaked at 12% of the labor force in the mid-nineties, the Dutch government implemented several reforms to reduce the number of DI beneficiaries. One of these measures was the introduction of firm experience rating in 1998. While most countries that provide Workers' Compensation use experience rating to finance disability benefits, The Netherlands and Finland are the only countries with experience rating for public DI benefits.

* Corresponding author. Tel.: +31 20 444 6120.
E-mail addresses: nynke.de.groot@vu.nl (N. de Groot), p.w.c.koning@vu.nl (P. Koning).

http://dx.doi.org/10.1016/j.labeco.2016.05.007
0927-5371/

In The Netherlands, the DI premium for both firms and governmental agencies is based on the DI costs of the (former) workers of the particular firm or agency. In the period investigated here, annual firm disability risks were defined as the disability costs of DI benefit recipients who entered into the program over a time frame of five preceding years, divided by the average wage sum over the same time frame. Next, the DI risk was translated into the DI premium that was paid by firms over their current wage sum. This premium was capped by both a maximum and a minimum premium. Over the years, the maximum DI premium peaked in 2004 at about 9% of the wage sum for firms classified as large. For the remaining group of small firms, DI maximum premium rates were proportionally lower.

To study the effects of experience rating, this paper exploits the removal of experience rating for the group of small firms that took place in 2003. This removal of experience rating allows us to use a difference-in-difference (DiD) design, with large firms as a control group for which the experience-rating incentive did not change. We study whether the removal of experience rating increased the DI inflow and decreased DI outflow rates, using 2001 and 2002 as pretreatment years and 2003 and 2004 as successive years in which the reform was enacted and may have affected DI inflow and DI outflow. In the empirical analysis, we use matched administrative data from Statistics Netherlands on firms and (former) workers between 1999 and 2011. We enrich these data with DI spells as well as other demographic and labor market characteristics. This resulted in a data set with over 250,000 unique firms and almost ten million workers who are eligible for DI benefits.

The choice of these particular years for the analysis had to do with two important reforms that took place in 2005 and 2006. These reforms probably affected small and large firms in different ways, which led us to limit the time period we use for our DiD design. In particular, the reform in 2005 extended the sickness period that precedes DI benefit receipt — and for which firms are financially responsible — from one to two years. In 2006, a substantial reform of the DI system introduced the distinction between two types of DI benefits: one for workers who were permanently and fully disabled, and one for those who are partially and/or temporarily disabled. Experience rating did not apply to the new scheme for permanently and fully disabled individuals, thus restricting the experience-rating incentive to new partially and/or temporarily disabled individuals. Overall, both reforms substantially reduced the inflow into DI and the coverage of experience rating.

Although our preferred model focuses on the pre-2005 period, we also present DiD analyses that exploit the re-introduction in 2008 of experience rating for small firms; this yields estimates of the effect of experience rating on DI inflow and DI outflow. Moreover, we re-estimate the pre-2005 analysis on a sample of individuals that excludes workers who would not have been entitled to DI benefits if they had applied for DI benefits after the reforms in 2005 and 2006. This provides us with more insight into the specific ways the reforms may have altered the potential impact of experience rating.

Generally, our findings are in line with economic predictions. In the time period under investigation, experience rating reduced inflow into DI and increased outflow from DI. These results are robust with respect to sensitivity analyses on the setup of our data and the specification of common trends. As to DI outflow, we find effects to be confined to partially disabled workers only. There is no evidence of experience-rating effects in the post-2005 period. We argue that this decrease in the impact can largely be attributed to the extension of the sick period before DI benefits commence — from one year to two years.

This paper adds to a literature on experience rating that is still limited. For The Netherlands, Koning (2009) studies the unanticipated effects of experience rating of firms who experienced an increase in their DI premium. Van Sonsbeek and Gradus (2013) estimate the effect of experience rating in The Netherlands, using aggregated sector data. Both studies find that experience rating reduced the inflow into DI by about 15%. Korkeamäki and Kyyrä (2012) exploit a pension reform in Finland to study the effect of experience rating. They find significant effects of experience rating for older workers on both the inflow into sick leave and the transition from sick leave into disability retirement.[1]

Experience rating is more widespread in private Workers Compensation (WC) schemes than in DI schemes that are provided as a public scheme. Most studies on WC focus on such outcome measures as fatality- and injury rates. The picture that emerges from these studies is that experience rating reduces disability claim costs (see Hyatt and Thomason (1998) or Ruser and Butler (2009) for survey studies).[2] At the same time, evidence points to certain unintended effects of experience rating, such as increased claims control and increased pressure not to report injuries (Ison, 1986; Lippel, 1999; Strunin and Boden, 2004).

This paper proceeds as follows. In the next section we describe the Dutch DI system and in Section 3 we discuss the method of experience rating. Section 4 presents our data. We discuss the empirical implementation in Section 5 and present the results from the estimations in Section 6. Section 7 concludes.

2. Institutional setting

Until recently, the Dutch DI system could be characterized as one of the most generous schemes of all OECD countries (OECD, 2010). Although several reforms have been introduced to make the scheme less susceptible to moral hazard problems, the Dutch DI scheme still differs from most DI schemes in other countries in some important aspects. The level of the benefits is based on the difference between the pre-disability (covered) earnings and the residual earnings capacity, where the residual earnings capacity is the income the individual could earn conditional on his or her disability. This means that disability is measured as a percentage, rather than as an all-or-nothing condition. Moreover, The Netherlands is one of the few countries where the DI program covers all workers against all income losses that result from both occupational and non-occupational injuries (LaDou, 2011). DI claims are assessed by the public benefit administration called UWV (Uitvoeringsinstituut Werknemersverzekeringen, roughly translated as Employee Insurance Agency).

Since the introduction in 1967 of the generous DI scheme known as the WAO, the Dutch DI stock has steadily increased and the DI inflow has remained persistently high (Fig. 1). The generosity of the system made it susceptible to moral hazard problems; for both firms and workers the scheme functioned as an attractive alternative pathway into unemployment (Koning and van Vuuren, 2007, 2010). Starting from 1996, the Dutch government implemented various reforms to increase the incentive of both employers and workers to decrease DI enrollment (Fig. 2).

First, the sickness benefit program was privatized in 1996, making employers fully financially responsible for the first year of sickness benefits of their workers. Employer incentives were further enhanced by the system of DI experience rating that started in 1998.[3] Since then, the DI premium for Dutch firms has been based on the

[1] There is a related literature that studies the effect of experience rating in the context of sickness benefits; see e.g. Fevang et al. (2014) and Böheim and Leoni (2011).

[2] For the US, we refer to Ruser (1985, 1991), Seabury et al. (2012) and Bruce and Atkins (1993) as studies on experience rating. In addition, Campolieti et al. (2006) presents evidence for Canada and Lengagne (2014) for France.

[3] The incentives of sickness benefits and DI experience rating both applied to all employers, including governmental agencies. For ease of exposition, in the remainder of the paper we refer to all employers — also including governmental agencies — as 'firms'.

Fig. 1. Dutch stock and inflow of workers in disability insurance as a percentage of the insured population (1967–2012).

actual DI benefit costs of their (former) workers. The calculation of the DI premiums will be explained in the next section. The ability of firms to deter DI claims was (and still is) limited, as claims follow automatically once the sickness period ends.

In 2002, another reform increased the responsibility of firms by means of a more stringent system of gatekeeping; see De Jong et al. (2011) for a detailed description of the gatekeeper protocol. Firms have thus become responsible for the work resumption of sick workers, with the obligation to draft a rehabilitation plan together with the sick worker. In 2005, the sickness period for which firms are responsible, was further extended from one to two years. This measure, which effectively increased employer incentives to prevent sickness, also implies that (as of 2005) individuals entered disability benefits after two years of sick leave instead of after one year. This caused a substantial drop in DI inflow in 2005 (see Fig. 1).

Finally, the most recent reform in 2006 entailed the start of two different types of DI benefits: the IVA (Income scheme for Fully Disabled) benefit for individuals who are fully and permanently disabled and the WGA (Act for Partially Disabled workers) benefit for those with partial, or temporarily full, disability.

Fig. 1 shows that there are strong reasons to believe that the DI reforms have been largely successful in curbing DI inflow since the start of this century. Koning and Lindeboom (2015) argue that the key to this success has been the intensified role of firms in preventing long-term sickness absence and subsequent disability, with a strong emphasis on early interventions. Substantial economic incentives increased the urgency among firms to increase their efforts to orevent sickness and accidents and to help reintegrate disabled workers, while the Gatekeeper protocol facilitated employer awareness and guided firms in their new role. That said, the extent to which the experience rating system has contributed to this process remains unclear.

3. Experience rating in The Netherlands

In this section we explain the calculation of the experience-rated DI premium of Dutch firms. We first discuss the general method of calculating DI premiums in 1998 and then present an overview of changes in the calculation of the premiums over the years. To shed some light on the consequences of these changes, we also assess yearly variation in the size of DI experience-rated premiums, which is measured as a percentage of the annual wage costs of a firm.

3.1. Setting of experience rating

The experience-rated DI premium of Dutch firms is based on the individual disability risk of a firm. The disability risk is defined as

$$d_{it} = \frac{\sum_{s=0}^{T} S_{t-2,t-2-s}}{\sum_{s=0}^{T} W_{t-2-s}/(T+1)}, \tag{1}$$

where $S_{t,\tau}$ are the disability costs of firm i in year t for recipients that entered into the program at time τ ($t \geq \tau$). As the equation shows, disability costs are divided by the insured wage costs W_t at time t, so as to obtain the disability risk d_t. Both the DI benefit costs and the wage sum are registered with a delay of two years and are summed over several successive cohorts of workers. In 1998, the time window for the disability risk was five years, so $T = 4$. Particularly for starting firms, the information that is needed to calculate the disability risk is incomplete. The disability cost percentage is then calculated over the

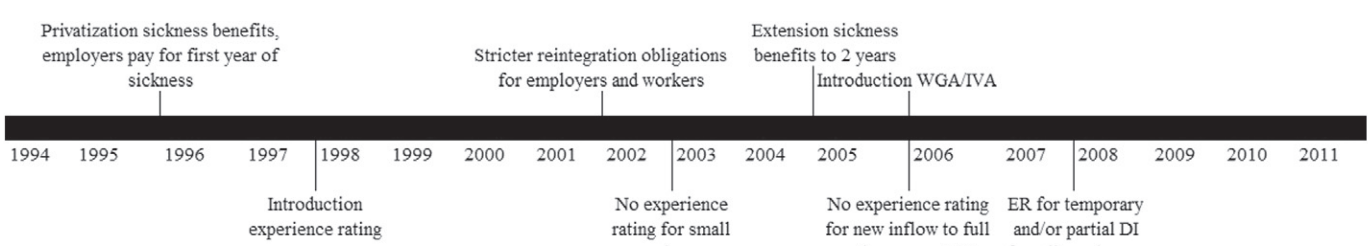


Fig. 2. Recent changes in disability insurance employer incentives in The Netherlands (1994–2011).

longest available time window, and subsequently rescaled to a time window of five years. Although this way of rescaling (artificially) increases the spread of DI risks, the effective impact in actual premiums that are paid is limited; in almost all cases rescaling applies to small firms that either have no disability costs or would have paid maximum premiums also in the absence of rescaling.

Next, the firm DI premium p_{it} that follows the individual disability risk is capped by the minimum premium $p_{\min}$ and the maximum premium $p_{\max}$:

$$p_{it} = \min\ (p_{\min} + d_{it},\ p_{\max})\,. \tag{2}$$

This means that every firm pays at least a uniform minimum premium. Moreover, the premium cap implies that the experience-rating system is ‘incomplete’ to some extent: higher disability costs result in proportionate increases in the DI premium up to the maximum premium, but over-users do not pay the additional costs they impose on the system. Next to DI benefit costs that originate from firm start-ups and firm bankruptcies, the costs of over-users are financed by the minimum premiums.

In the time period under investigation, the values of the minimum and maximum premiums vary with respect to firm size, the argument being that small firms are more susceptible to exogenous variation in their DI cost percentage. Initially, small firms were defined as having total wage costs that are smaller than the average wage costs per worker in The Netherlands, multiplied by 15 (workers). Maximum premiums are set equal to four times the average premium for large firms and to three times the average premium for small firms. Then, using an iterative algorithm, the minimum premiums are set at the level that balances the total disability costs with the collected premiums. As DI cost percentages of small firms are more likely to be bounded by the maximum, the minimum premium is higher for small firms.

For ease of exposition, Eq. (2) abstracts from any differences in DI benefits that stem from the two-year delay in the experience-rating system. That is, if the current average DI risk exceeds (is smaller than) the DI risk at $t-2$, the premiums are increased (decreased) proportionally. In the years before 2005, the DI risks were downscaled by at most 17%, but after 2005 upscaling of around 30% was applied.

As a final remark, the introduction of experience rating was combined with the possibility for firms to opt out of the public system to private insurance providers. Between 2001 and 2004, at most 3.8% of the firms opted out of the public system (Deelen, 2005). Also, Hassink et al. (2015), who investigate the years 2007–2011 wherein the share of privately insured firms equaled about 30%, show that opting out had no effect on DI inflow rates. We thus do not expect opting out to substantially change the incentive of DI experience rating.

3.2. Experience rating over the years

Over the years, the calculation method of DI experience rating has not changed fundamentally. This does not mean, however, that the effective impact of experience rating on individual DI premiums has remained constant over time. In 2003, experience rating was abolished for firms that were classified as ‘small’, and was replaced by a system of sectoral premium rates. In 2004, the coverage of experience rating across firms was further reduced, as the group of ‘small’ firms was extended from 15 to 25 times the average wage costs in The Netherlands. Firms with wage costs between 15 to 25 times the average wage were thus still experience rated in 2003. In 2008, however, experience rating was re-introduced for smaller firms. The scheme now covers the DI benefit costs of the old WAO scheme and the new WGA scheme for temporary and/or partial disability. As the total costs of these two new benefit schemes together are gradually decreasing over time, the total sum of DI costs that are experience rated decreases over time as well.

Due to the above-mentioned changes, we observe substantial variation in the potential range of the experience-rated premiums across years (see Fig. 3). With additional DI benefit cohorts that were annually added to the individual disability risk, the spread of experience-rated premiums increased in the first years of DI experience rating between 1998 and 2003. However, lower experience-rated DI costs caused by the extension of sick leave benefits in 2005 and the new DI scheme in 2006 have effectively reduced the spread of DI premiums to levels that have been fairly constant since 2007.

To shed more light on the importance of the minimum and maximum DI premiums, Fig. 4 presents the distribution of the premiums for all firms, using administrative data from UWV. Clearly, the vast majority of small firms — without disabled workers that were assigned to them — pay the minimum premium. In the years 1999–2002, around 5% of the small firms paid the maximum premium; in 2008–2011 this percentage decreased to around 3%. While most small firms pay either the minimum or the maximum premium, the majority of the firms that are classified as ‘large’ pay a premium somewhere between the minimum and maximum premium.

4. Data

In our analysis, we use various administrative data sets from Statistics Netherlands that contain information on DI benefits and employment spells that are observed between 1999 and 2011. Data sets from Statistics Netherlands can be linked with unique firm and worker identifiers. As to firms, we also observe the administrative information from UWV that is needed to calculate their DI risks, including their status as ‘small’ or ‘large’.

Unfortunately, until 2009, firms in the UWV data do not have similar identifiers as those of Statistics Netherlands. This means that the classification of firms into ‘small’ or ‘large’ can only be derived from the information of wage sum costs in the data of Statistics Netherlands. In this context, care should be taken in two respects.

First, the exact calculation of wage costs in the data of Statistics Netherlands may differ from calculations from UWV due, for instance, to differences in the reference date and the inclusion or exclusion of additional income such as leased cars or compensation for travel costs. This in turn implies the presence of measurement errors in the data from Statistics Netherlands, causing some employers to be wrongly classified as small or large. To shed more light on the potential impact of measurement errors, we can, however, merge the firm data for 2009–2011. We then find that about 0.5% of the small firms have been wrongly classified as large; the percentage of large firms that have been wrongly classified as small then decreases from 6.4% in 2009 to 4.6% in 2011. In light of these small fractions, we do not expect a large estimation bias. If anything, we would underestimate the potential effects of the removal of experience rating for small firms because some of the classified small firms are actually experience rated and vice versa.

Second, firms in the data from Statistics Netherlands may consist of different plants, each paying distinct experience-rated premiums. An example is a large chain of supermarkets in The Netherlands. Statistics Netherlands merges these supermarkets to one large firm, while UWV regards them as separate entities with different risk premiums. To solve this matter, we restrict our analysis to firms with single plants.[4] This results in a loss of around 20% of the firms and

[4] For example, in 2009 91% of the firms in the UWV data correspond to exactly one firm in the data of Statistics Netherlands, 7% to two firms, 2% to three or more firms. As a robustness test, we present model outcomes that also employ data from firms with multiple plants, assuming that plants all have similar experience-rating incentives.

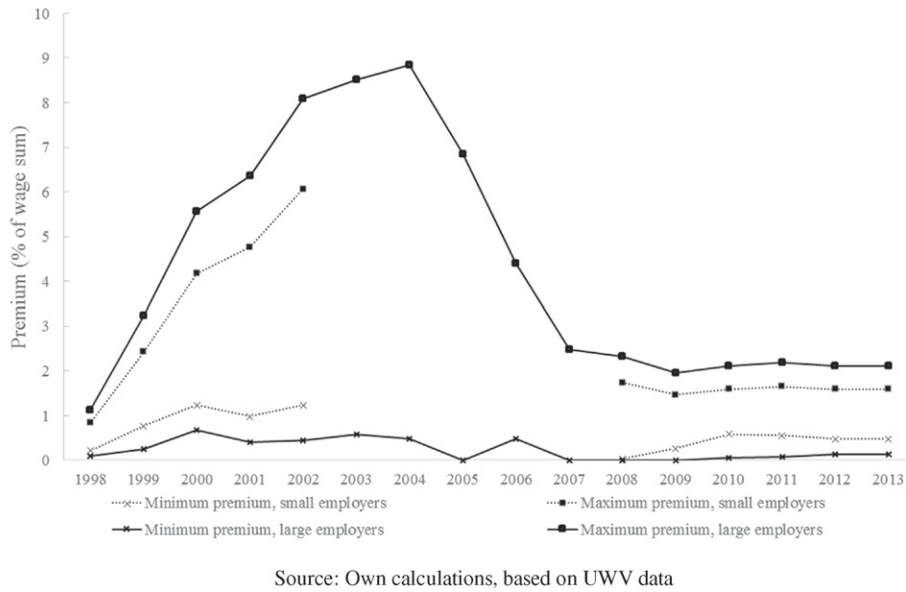


Fig. 3. Range of experience-rated DI premiums, measured as a percentage of wage costs and stratified with respect to firm size (1998–2013). Firm size is based on the total wage costs of the firm.

30% of the workers in our sample. These are predominantly larger firms.

Table 1 summarizes the main characteristics of the combined data sets from Statistics Netherlands. We only present the statistics for the selected sample of firms with a single plant. Recall that the data also include governmental agencies, as DI experience rating also applies to these employers. Also, note that the statistics on DI recipients represent only the benefits of individuals who were assigned to a firm because of experience rating. As a result, we observe a decrease in the percentage of individuals with DI benefits, especially since the extension of the sick leave benefits in 2005 and the introduction of the new WGA and IVA schemes in 2006 (see Koning and Lindeboom, 2015; Van Sonsbeek and Gradus, 2013). Accordingly, the average premium has decreased substantially after these reforms.

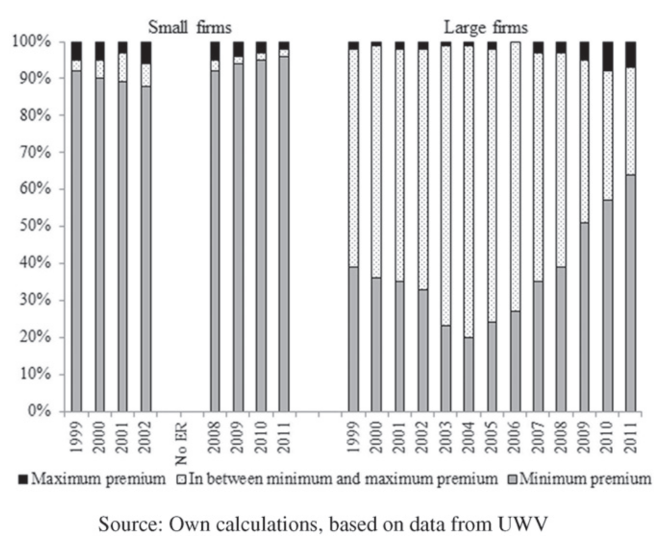


Fig. 4. Distribution of experience-rated DI premiums of firms: minimum premiums, maximum premiums, and premiums in-between minimum and maximum (1999–2011).

Finally, we observe a decrease in the number of firms in our sample after 2005. This is due to a change in the source of employment contracts in 2006 in the data of Statistics Netherlands.

5. Empirical implementation

5.1. General estimation strategy

Obviously, the experience-rating system in The Netherlands aims at an increase in preventative and reintegration activities. In line with this, one would expect a decrease in the inflow into DI and an increase of the outflow out of DI of those disabled workers that were assigned to firms.[5] We now test whether experience rating had these intended effects on DI, using a difference-indifference approach that exploits the removal of experience rating for small firms in 2003.[6]

Recall from Section 2 that several DI reforms took place after the introduction of experience rating in 1998. These reforms may have altered the effectiveness of DI experience rating. Specifically, in 2005 the sickness benefits period was extended to two years, and in 2006 the new DI scheme with two distinct schemes was enacted. It is likely that the reform in 2005 led to a lower DI inflow rate, with DI recipients having more severe impairments compared to the period when the assessment of claims was performed after one year of sickness benefit receipt, and the eligibility standards were less stringent. In addition, the introduction of a graduated DI system may have triggered complex behavioral responses among individuals — see e.g. Autor and Duggan (2007) and Marie and Vall Castello (2012).

Since the reforms in 2005 and 2006 changed the size and composition of the DI inflow substantially and may have affected small and large firms in different ways, the primary focus of our analysis will

[5] Experience rating could also have unintended effects, such as substitution to unemployment insurance(UI) benefits, changes in hiring policies or an increase of firm exits. These effects are, however, beyond the scope of the current paper.

[6] Although there are two distinct experience-rating systems for small and large firms, the use of regression discontinuity designs to estimate the impact of experience rating is not straightforward in the current context. In particular, firms in a close interval around the threshold can switch from being classified as small to large, or the reverse.

Table 1
Descriptive statistics of the Statistics Netherlands data for all firms with one plant, for the years 2001 to 2011 (only odd years are shown).

	2001	2003	2005	2007	2009	2011
Number of firms	252,400	216,254	203,503	122,542	157,129	151,689
Number of workers (×1,000)	6803	5908	5582	3214	4108	3534
Average firm size (workers)	27.0	27.3	27.4	26.2	26.1	23.3
% of large firms	8.4	9.4	9.4	8.2	8.9	5.6
% pays the minimum premium	94.4	86.5	83.6	87.7	90.9	93.7
% pays the maximum premium	2.4	4.9	7.7	8.7	6.7	4.8
Average premium (% of wage sum)	1.73	2.30	1.87	0.79	0.76	0.87
Average risk percentage	0.6	2.2	2.8	2.3	2.1	1.9
Sector (%)						
- Trade	23.1	23.0	23.2	26.7	25.2	22.9
- Industrial	13.7	14.4	14.5	15.8	14.1	10.7
- Business	10.9	10.8	11.5	11.7	12.7	10.7
- Health	11.0	11.3	11.1	13.1	11.4	11.6
- Food	9.1	9.1	8.8	10.0	9.3	9.5
Worker characteristics						
Average age	36.8	37.8	38.5	38.3	38.9	39.8
Male (%)	53.1	52.4	51.6	51.2	50.3	48.1
Immigrant (%)	16.7	16.5	16.4	16.8	17.9	18.4
Permanent contract (%)	–	–	–	72.0	68.9	69.5
Pre-disability earnings (€)	19,955	21,513	22,253	23,284	26,023	27,475
Characteristics of DI recipients[a]						
Number of DI recipients	195,973	220,445	187,095	80,762	81,338	69,174
DI, % of workers	3.6	4.5	4.0	2.9	2.3	2.3
- % WAO	100	100	100	84.6	60.8	41.3
- % WGA	–	–	–	12.3	30.4	43.7
- % IVA	–	–	–	3.1	8.8	15.0
- % fully disabled	48.8	50.2	49.0	52.0	55.9	59.1
Inflow into disability	65,861	40,828	14,267	11,043	11,381	9559
Inflow, % of workers	1.2	0.8	0.7	0.4	0.3	0.3
Outflow from disability	22,417	22,345	22,886	5691	4913	4021
Outflow, % of workers	0.4	0.4	0.5	0.2	0.1	0.1
Average annual DI benefits (€)	6714	9150	10,567	12,328	13,469	14,321

[a] DI statistics only include the DI spells of individuals that could be linked to a firm. If an individual has not been employed for the last five years, the DI spell is not included either. This explains why the number of worker observations is considerably smaller than the total DI inflow.

be on the time period from 1999 to 2004.[7] In these years, our treatment group consists of small firms for which experience rating was removed in 2003–2004. As an additional analysis, we also present model outcomes for the period between 2006 and 2011. With experience rating being re-introduced for small firms in 2008, this means that the treatment group in this period consists of small firms that were not experience rated in the years 2006 and 2007.

5.2. Identification issues

The research design for both the inflow model and the outflow model essentially relies on three identifying assumptions. First, the DiD setup assumes that the outcome measures of the treatment group and the control group share a common time trend. Second, firms should not anticipate the wage costs threshold that determines the experience-rating incentive. Finally, there should be no firms that switch over time between the treatment group en the control group.

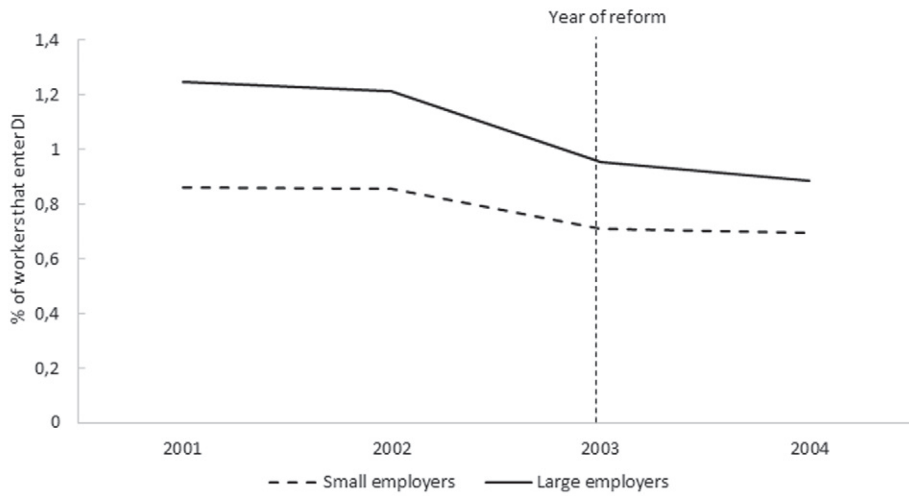
To start with, the common trends assumption implies that sick or disabled individuals who were employed at a small firm respond similarly to calendar time effects as their counterparts employed at large firms. As an eyeball test on this assumption, Fig. 5 explores the evolution of DI inflow and DI outflow as pre-treatment trends. The upper panel portrays the inflow into DI as a percentage of the total numbers of workers for small and large firms in the years 2001–2004. Before the reform, we observe similar trends in inflow between 2001 and 2002.[8]

Similarly, the lower panel of Fig. 5 shows the survival curves of those receiving DI by year of inflow into DI and size of the firm. We do not observe a difference in the survival curves of individuals from small and large firms between 2001 and 2002. The survival rates of individuals who worked at small firms are similar to their counterparts from large firms until the end of the first year of DI. After the first year, the survival probability of individuals who worked at a large firm drops below that of individuals who worked at a small firm. Nevertheless, more formal robustness tests are needed on time trends in DI inflow and outflow. Toward that end, we will formulate a placebo test and use samples of the treatment and control groups with more similar employer sizes.

Our second assumption is that firms do not anticipate the wage costs threshold that determines the size of the experience-rating incentive. Anticipation effects would occur if firms keep the wage costs just below the threshold to avoid experience rating, or the reverse. We argue that such effects are unlikely to exist, since the threshold, which is set in the year before the actual year of experience rating, applies to the wage costs of the two years beforehand. Moreover, the removal of experience rating for small firms in 2003 was announced in July 2002. Large firms were thus not able to decrease their wage costs to escape from experience rating. This is

[7] To clarify this point, consider the 2005 extension of the sick leave period. According to Kok et al. (2013), small firms responded to this change by increasing private insurance, whereas larger firms did not. This renders it likely that the decrease of DI inflow due to the extension of sick pay was higher for large firms than for small firms.

[8] We repeated the explanatory analysis for the small sample of firms that could be matched to UWV data, as we then observe the years 1999 and 2000. Again, we observe similar trends in inflow for small and large firms in the pre-treatment period. This figure is available from the authors upon request.



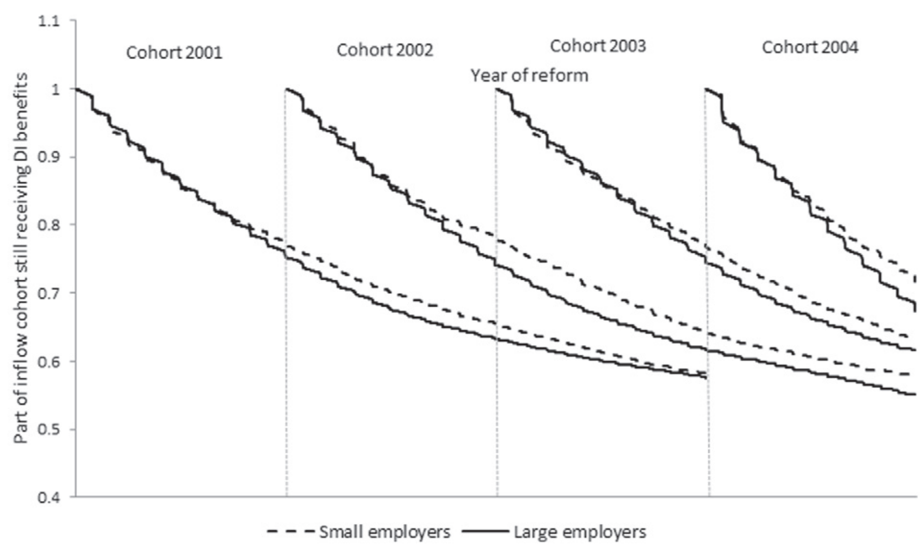


Fig. 5. Inflow into DI and survival curves of DI receipt by year of DI inflow, stratified by size of the firm based on wage costs.

confirmed by Fig. 6, which displays the distribution of firms with total wage costs around the threshold of experience rating. In particular, there is no evidence that the wage costs of firms concentrate just below the threshold value. We also tested this formally with the discontinuity test suggested by McCrary (2008). The null hypothesis of a continuous wage sum around the threshold could not be rejected for any year between 2001 and 2011, except for 2007.[9]

Third, our estimation strategy assumes that firms are classified as small or large over a longer stretch of time. In practice, however, firms may switch from small to large in the next year, or the reverse. In this respect, recall that the thresholds for experience rating are set with a time delay. Consequently, the ex-ante incentive effect of experience rating will be almost equal for firms with wage costs that are just below and just above the threshold. As there are many firms close to the threshold that switch between experience-rating statuses, one therefore may expect the effect estimates of experience rating to be biased toward zero. This effect applies particularly to firms with wage costs that are close to the threshold, as firms just below the experience-rating threshold are likely to be subject to experience rating in the following year (and vice versa).

To assess the size of a potential attenuation bias close to the threshold, Table 2 shows the percentage of firms that switched from one classification to another classification in the following year. The first two rows show the percentage of small and large firms classified as the opposite size in the following year. For small firms, this percentage is relatively small, at most 1%. We do observe a more substantial percentage of large firms that drop in the next year below the experience-rating threshold, with 7.0% of large firms at the most. When calculating the number of switches per firm, we find that the vast majority of firms never switches classification. Only 3.5% of the firms change from small to large or the other way around, and most of those firms only switch once (2.3%). We therefore expect that the bias of switching of firms is relatively small. If small firms take into account that they might be subject to experience rating the next year (or the reverse), this would cause a small underestimation of the effect of experience rating.

Table 2 also shows that yearly switches between firm statuses are much more prominent if we zoom into wage sums that differ less than €100,000 from the threshold value. About 20% of the small firms close to the threshold are classified as a large firm in the following year, whereas the opposite holds for about 27% of the large firms.

[9] The McCrary test yielded a p-value of the null hypothesis of continuity in the density around the experience-rating threshold that was equal to 0.02 for the year 2007. For all other years, the p-value was well above 0.10.

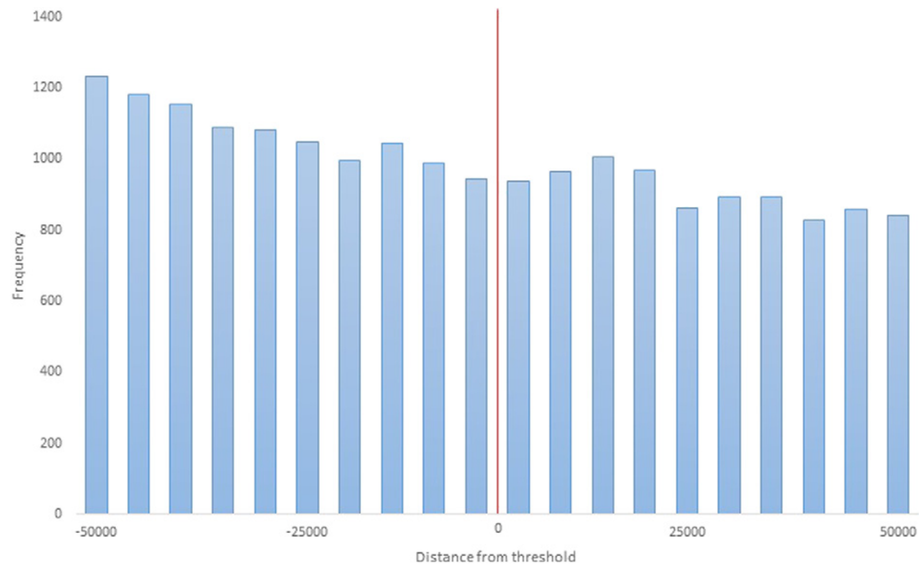


Fig. 6. Wage cost distribution of employers, stratified with intervals of €5000 around the experience-rating threshold, aggregated over 2003–2007.

Table 2
Percentage of firms that switch from small to large (or the reverse), based on the experience-rating threshold of the wage costs (2002–2011)[a].

Actual size	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
All firms										
Small to large	0.7	0.7	1.0	0.6	0.3	0.3	0.5	0.6	0.2	0.2
Large to small	4.8	4.3	7.0	5.6	4.8	5.0	5.6	1.6	5.6	2.9
Wage sum close to threshold[b]										
Small to large	15.2	21.4	28.6	22.1	17.8	17.4	24.2	22.9	18.5	21.7
Large to small	25.5	16.2	36.9	27.2	25.0	29.6	26.8	15.5	38.6	26.0
All	19.9	20.8	37.2	24.6	21.7	24.2	27.2	22.3	30.5	23.9

[a] The wage costs are measured with a delay of two years. Before 2004, the experience-rating threshold was equal to 15 times the average wage; after 2004 it was equal to 25 times the average wage.
[b] Only firms with a wage sum that differs less than €100,000 from the threshold.

This suggests that a Regression Discontinuity design will probably underestimate the effect of DI experience rating.

5.3. DI inflow model

Thus far, we have discussed the assumptions that are necessary for our difference-in-difference design. We present next the empirical specification that is used to implement this design, using DI inflow and DI outflow as outcome variables of interest.

As the experience-rating incentive is directed to individual firms, we aggregate the individual data on DI inflow at the level of individual firms. An alternative would be to estimate an individual duration model for the time until inflow into DI. The main disadvantage of this approach is that we do not observe employment before 1999. We thus would have to estimate the model on a stock sample, which could lead to biased estimates.[10]

We define the inflow y_{jt}^{inflow} as the fraction of workers who worked for firm j in the year of risk ($t-1$ before 2005, $t-2$ after 2005), entering DI in year t. With the dependent variable that is expressed as a fraction of the workers per firm, we propose the fractional probit estimator described in Papke and Wooldridge (2008) that incorporates the longitudinal nature of the data. This essentially implies that the effect of the removal of experience rating is identified from 'within-firm' variation. We estimate the model using the pooled Bernoulli quasi maximum likelihood estimator, as described in Papke and Wooldridge (2008). This estimator assumes a conditional mean of the following form:

$$E\left(y_{jt}^{inflow} \middle| S_{jt}^{s}, D_{jt}, X_{jt}, \rho_j\right) = \Phi\left(\alpha + \kappa^s S_{jt}^s + \bar{\kappa}^s \overline{S_j^s} + \delta D_{jt} + \bar{\delta}\overline{D_j} + \beta X_{jt} + \bar{\beta}\overline{X_j} + \mu_t + \rho_j\right) \quad (3)$$

where Φ is the standard normal cumulative distribution function and ρ_j is a firm effect that is assumed to follow a normal distribution, conditional on the regressors S_{jt}^s, D_{jt}, X_{jt} and μ_t.[11] α is a constant and the variable D is our treatment dummy: this variable is equal to 0 if the firm is classified as large in all years, as well as for firms that are classified as small in the years from 1999 to 2002 (before the removal of experience rating). Note that in the additional analyses for the period after 2005, the treatment variable is set to 0 from 2008 to

[10] Although one may argue that biases due to stock sampling apply to both large and small firms, we cannot rule out that these biases are different. In particular, job turnover is likely to be larger for small firms. Even so, we ran a logit specification for DI inflow with individual data. We briefly discuss these results in the robustness checks in Section 6.

[11] See Papke and Wooldridge (2008) for a derivation of this conditional mean.

2011 (after the re-introduction of experience rating). Consequently, D_{jt} is set equal to one if the firm is classified as small between 2003 and 2007 and was not subject to experience rating.

Vector X_{jt} contains both firm characteristics (dummies for sector, average wage) and characteristics of the workers of the firm (average age, percentage of men, percentage of immigrants). Recall from Section 3 that in 2004 the threshold value of wage sums for small versus large firms was increased from 15 to 25 times the average wage per worker. In our analysis we therefore define 'medium-sized firms' as those that have a wage sum between 15 times and 25 times the average wage. For both small firms with a wage sum that is smaller than 15 times the average wage and medium-sized firms, we estimate control dummies S^1 and S^2. The time trend μ_t is specified using dummy variables for every year. This vector controls for calendar time variation in inflow probabilities and is identified by the control group of large firms. $\overline{S_j^s}$, $\overline{D_j}$ and $\overline{X_j}$ are the time-averages of S_{jt}^s, D_{jt} and X_{jt} for firm j.

In our regression we cluster the standard errors at the level of the firm and obtain them using 500 bootstrap replications. Unfortunately, there is no validated method existing at present to estimate the fractional probit model on an unbalanced sample. We therefore estimate the model on a balanced sample of firms.[12]

5.4. DI outflow model

To estimate the effect of experience rating on DI outflow, we use data on the level of the individual workers instead of firms. We thus avoid losing individual information on DI durations that would occur if we aggregate the outflow to the level of firms. We model the duration of DI benefits on a flow sample of individuals entering DI by using a hazard rate model with a Cox proportional hazard specification that can be estimated with standard Maximum Likelihood techniques:

$$y_{ij\tau,t}^{outflow} = \lambda(t)\exp\left(\kappa^s S_{jt}^s + \delta^{1st} D_{jt}^{1st} + \delta^{2nd} D_{jt}^{2nd} + \beta X_{ijt} + \mu_\tau\right) \tag{4}$$

where $y_{ij\tau,t}^{outflow}$ denotes the outflow hazard on day t for an individual i who entered DI at calendar time τ and worked for firm j before entering DI. $\lambda(t)$ represents the duration dependence in outflow from DI benefits. Again we include two firm-size dummies S_{jt}^s to control for the size of the firm (based on the total wage costs), as well as dummies for the year of inflow μ_τ. X_{ijt} includes both firm characteristics (i.e., sector and average wage of the firm) as well as worker characteristics (i.e., gender, immigrant, wage category, regional and household status). We allow the potential effect of experience rating to vary with respect to the DI duration, allowing for distinct treatment effects in the first $\left(D_{jt}^{1st}\right)$ and second year of DI benefit receipt $\left(D_{jt}^{2nd}\right)$.

6. Estimation results

6.1. Baseline specification

Table 3 shows the main estimation results for the fractional probit model for DI inflow, which is measured as a percentage of the workers at the firm (see columns two and three, respectively). A table with all coefficient estimates can be found in the Appendix to this paper.

Our key finding is that the removal of experience rating increased DI inflow in the period prior to 2005. The implied average partial effect of experience rating for small firms in this period is equal to an increase of 0.00051 in the annual DI inflow rate. As the average annual DI inflow rate for small firms equaled 0.0074 before the removal of experience rating, this implies a relative increase of 7%. This effect corresponds to about half of the size of the effect found by Koning (2009) and Van Sonsbeek and Gradus (2013). One explanation for this difference may be that the effects of experience rating are smaller for the treatment group of small firms than for the control group of large firms. Like Koning (2009), one may also argue that firms typically responded to unanticipated increases in premiums, rather than being fully informed and this able to anticipate the incentives.[13]

Table 3
Fractional probit estimations (quasi-MLE) for the fraction of workers per firm that is awarded with DI benefits (2001–2004) and Cox proportional hazard estimates (no hazard ratios) of outflow from DI for individuals who entered DI between 2001 and 2004.

	Inflow		Outflow	
Removal of ER	0.027**	(0.009)	–	–
Removal of ER, first year after inflow	–	–	−0.154**	(0.022)
Removal of ER, second year after inflow	–	–	−0.039	(0.024)
Small firm	0.041	(0.040)	−0.037**	(0.014)
Middle-sized firm	0.040	(0.024)	0.029	(0.019)
Year effects	Yes		Yes	
Worker characteristics	No		Yes	
Firm characteristics	Yes		No	
Sector dummies	Yes		Yes	
Regional dummies	No		Yes	
Observations	183,665		119,631	

Standard errors in parentheses, for inflow estimations obtained using bootstrap with 500 replications. *Significant at a level of 10%. **Significant at a level of 5%.

In a broader perspective, our results are comparable to those obtained by Campolieti et al. (2006) and Hyatt and Thomason (1998) for Workers' Compensation in Canada. Moreover, the coefficient estimates of the control variables are in line with expectations (see Appendix A: firms with older workers, a lower average wage and operating in the sectors construction and transport have a higher inflow into DI).

As to the estimation of effects on DI outflow, recall that we use data on individuals who entered the DI scheme between 2001 and 2004 and who can be assigned to a particular firm, and then estimate the DI duration using a Cox proportional hazard specification. The resulting coefficient estimates and standard errors are given in columns four and five of Table 3. Loosely speaking, the coefficient values that are presented in the fourth column can be interpreted as a percentage increase or decrease in the exit rate out of DI. Again, a full table that includes all estimated coefficients can be found in the Appendix to this paper.

In line with expectations, the coefficient values of the removal of experience rating on DI outflow are negative. This implies that the removal of DI experience rating decreases the probability of an exit from DI, and thus increases the DI duration. Still, we find a significant impact only for the first year of DI benefit receipt. Our impact estimates correspond with a drop in the DI exit probability of 3.0 percentage points after one year (from 24.7% to 21.7%) and of 4.7 percentage points after two years (from 34.1% to 28.4%). These results correspond roughly to those of Van Sonsbeek and Gradus (2013), who find a positive, borderline significant effect of experience rating on DI outflow.

Our estimates indicate that individuals who worked for small firms are less likely to exit DI. Arguably, small firms may have fewer

[12] We did estimate the fractional probit model on the unbalanced panel following the method proposed in Wooldridge (2010). The main conclusions do not change when using these estimation results.

[13] The study of Korkeamäki and Kyyrä (2012) supports this hypothesis. They estimate the effect of a lump-sum payment by employers at the moment of DI entry. This effect is markedly larger than the effect of conventional experience-rating systems.

Table 4
Coefficient estimates of the effect of the removal of experience rating on DI inflow and DI outflow: Heterogeneity.

	DI inflow		DI outflow			
			First year		Second year	
Baseline specification	0.027**	(0.009)	−0.154**	(0.022)	−0.039	(0.024)
By degree of DI						
DI ≤35 %	−0.075	(0.077)	−0.270**	(0.056)	0.023	(0.056)
DI 35–80 %	0.012	(0.040)	−0.297**	(0.069)	0.035	(0.069)
DI >80%	0.034	(0.053)	−0.048	(0.040)	−0.002	(0.041)
By level of DI						
Below the median	−0.031	(0.027)	−0.191**	(0.036)	0.028	(0.036)
Above the median	0.140	(0.148)	−0.103**	(0.052)	0.058	(0.053)

Every cell represents a separate analysis. Estimations include the same control variables as in the main analysis. Standard errors in parentheses, for inflow estimations obtained using bootstrap with 500 replications. *Significant at a level of 10%. **Significant at a level of 5%.

possibilities to arrange work adaptations or to offer job opportunities elsewhere. Conditional on work resumption, the probability of employment at the previous employer is about 50%. Finally, the remaining control variables of the DI outflow model are again in line with expectations: older individuals, women, immigrants, individuals with a low previous wage, single parents and individuals without children are less likely to exit DI.

The estimated effects of the removal of experience rating for small firms on both DI inflow and outflow can be translated into an effect on the total DI stock. In particular, the estimates imply that the total DI stock in 2004 was 0.4% larger because of the removal of experience rating for small firms. About two-thirds of this effect can be attributed to the effect on DI inflow, and one-third to the DI outflow effect. Assuming that the effects of experience rating on DI inflow and outflow are similar for large firms, the DI stock in 2004 would have been 1.7% larger if the removal of experience rating had been implemented for all firms.

Equipped with the individual information of employed workers and DI recipients, we are able to stratify the effect of experience rating with respect to various worker characteristics. Table 4 shows the coefficient estimates of the removal of experience rating for individuals with different degrees of disability and for different levels of DI benefits. The estimation results of the DI inflow model show no significant differences in effects between worker groups, which is probably due to the fact that (share) variables are calculated per firm. As to DI outflow, we find the experience-rating effect to be confined to partially disabled workers only. This suggests that the effects of experience rating are strongest for individuals with some job possibilities. Also, DI outflow effects are larger for workers with low pre-disability wages.[14]

6.2. Robustness analyses

In this subsection, we assess in greater detail our estimation strategy for both DI inflow and DI outflow effects. The results of the corresponding robustness analyses are presented in Table 5.

First, we focus on the selection of firms that is used in our analyses. So far, we have restricted our sample to firms with one plant only, so as to exclude firms for which it was impossible to ascertain whether or not they were experience rated. As a robustness check on the DI inflow and DI outflow models, we therefore expanded our sample with firms that have multiple plants. We thus aggregated the wage costs for firms with multiple plants. We next assumed that the total wage costs determine whether or not the plants of these firms are experience rated. The first lines of Table 5 show that adding firms with multiple plants to our data in this way does not substantially change our estimation results for either model.

Second, our estimation strategy relies on the assumption that small firms (i.e. those without experience rating in 2003 and 2004) share a common trend with large firms. Although our graphical analyses in the previous section did not reveal substantial differences in the trends between small and large firms, we can also perform formal analyses by adapting our sample of firms and adjusting model specifications. One simple test on the common trends assumption is to exclude firms with wage costs extending far beyond the experience-rating threshold. We do so by including only firms with more than five and less than 250 workers. We thus relax the common trends assumption, since firms in the treatment and control group become more comparable. Table 5 shows that this causes coefficient estimates to decrease somewhat, while the coefficient estimates for the DI outflow model do not change significantly.

Adding another robustness check on the common trends assumption, we also performed a placebo test on the experience-rating incentive. By pretending that the removal of experience rating for small firms occurred in 2001 instead of 2003, we thus created a placebo dummy that is equal to one if the firm is small in the years 2001 or 2002.[15] We substituted the treatment variable by the placebo variable and re-estimated our model for the years 1999–2002. For both outcome measures, Table 5 shows that this yields insignificant estimates for the placebo variables.

Third, one may argue that the impact estimate of experience rating on DI outflow can be considered as a lower bound. Higher DI inflow rates for the treatment group of smaller firms may have affected the composition of DI recipients, with the additional inflow consisting of individuals with better job prospects and, consequently, higher DI exit probabilities. We test for the potential importance of these compositional effects by concentrating on a stock sample of individuals who entered DI before 2003, which is the year the reform took place. As the fourth panel of Table 5 shows, this yields substantially stronger impact estimates of experience rating on DI outflow. From this we conclude that compositional effects do attenuate the impact of experience rating on DI outflow levels.

[14] Note that the coefficient estimates of the removal of experience rating do not differ across gender, age or sector that corresponds to the last job before the start of a DI spell.

[15] Since we need information on the years before 2001, we use data from UWV to measure the size of the firm for all outcome measures. The downside to this data set is that we can only account for the firms that still existed in 2009. For this reason we do not use this data set in the main analyses.

Table 5
Coefficient estimates of the effect of the removal of experience rating on DI inflow and DI outflow: Robustness tests.

	DI inflow		DI outflow			
			First year		Second year	
Baseline specification	0.027**	(0.009)	−0.154**	(0.022)	−0.039	(0.024)
Selection of firms						
All firms (multiple plants)	0.028**	(0.008)	−0.140**	(0.017)	−0.059**	(0.021)
Test common trend, firm selection						
Without very small firms[a]	0.020**	(0.007)	−0.166**	(0.031)	0.037	(0.031)
Without very large firms[b]	0.026**	(0.026)	−0.136**	(0.032)	0.033	(0.033)
Without very small and large firms	0.014**	(0.007)	−0.152**	(0.034)	0.049	(0.035)
Test common trend, placebo test[c]						
Placebo variable	−0.011	(0.049)	−0.033	(0.061)	0.112	(0.076)
Selection of inflow						
Stock sample before 2003	–	–	−0.342**	(0.047)	−0.060*	(0.033)
Separate effects for first and second half of the year						
First half	–	–	−0.104**	(0.027)	−0.096**	(0.031)
Second half	–	–	−0.219**	(0.030)	0.037	(0.034)
Individual data						
Logit (coefficient)	0.1530**	(0.0137)	–		–	
Logit, without small and large firms	0.0993**	(0.0175)	–		–	

Every cell represents a separate analysis. Estimations include the same control variables as in the main analysis. Standard errors in parentheses, for inflow estimations obtained using bootstrap with 500 replications.
* Significant at a level of 10%.
** Significant at a level of 5%.
a Fewer than five workers.
b More than 250 workers.
c Based on data UWV, 1999–2002.

Fourth, we also investigated the pattern of DI outflow effects by adopting a more refined specification of incentive effects, using intervals of six months instead of one year of DI benefit receipt. We then find significant and similar effects on outflow for the first one and a half years after DI inflow. Experience-rating effects become insignificant in the second half year of the second year, suggesting that, over time, the impact is hump-shaped.

Finally, we re-estimated the DI inflow model with individual instead of firm data, while using a logit specification. When interpreting these findings, one should bear in mind that we do not control for the employment duration of workers. The lower part of Table 5 shows the coefficient estimate of the removal of experience rating that follows from this strategy. In particular, we then find that the removal of experience rating increased DI inflow by roughly 15%. This is more than two times larger than the fractional probit estimate. One explanation may involve the oversampling of individuals from (very) large firms, which may violate the common trends assumption. We therefore repeated the estimation omitting individuals from very small firms (fewer than five workers) and large firms (with more than 250 workers). As a result, the estimated effect significantly reduces in size and no longer differs significantly from the estimate based on firm-level data.

6.3. Additional analyses

6.3.1. The effect of premium caps

So far we have assumed that the effect of experience rating does not depend on the level of the experience-rated DI premium, but applies to all firms in the control group equally. However, we explained earlier that premia are capped at minimum and maximum rates, causing experience-rating incentives along the premium distribution to differ at the margin. In particular, firms with premiums that are capped at the maximum premium have no incentive to curb new DI inflow.

To estimate the importance of adverse effects of the maximum premium, we calculated the experience-rated DI premium rates for firms in our sample.[16] This sample does not include the treatment group of small firms that were not experience rated in 2003 and 2004; for this group, we estimate a separate dummy. If firms are aware they are paying the maximum premium, one would expect experience-rated firms paying the maximum premium to have higher DI inflow rates and lower DI outflow rates than those firms paying premiums below the maximum.

Clearly, the effect of paying the maximum premium on DI inflow and DI outflow is subject to endogeneity bias. Firms with few prevention and reintegration activities have higher DI risks and higher corresponding DI premiums — and thus exhibit a higher likelihood of paying the maximum premium. To avoid this endogeneity problem, we estimate model specifications for DI inflow and DI outflow conditioning the initial DI risk of a firm. More specifically, we extend our model by including a (third order) polynomial of DI risks. The impact of the maximum premium can thus be identified as a Regression Discontinuity effect at a certain level of the DI risk.

Table 6 shows the estimation results that follow from this estimation approach for both the DI inflow model and the DI outflow model. For the DI inflow model we find a strong discontinuity effect for experience-rated firms with maximum premiums. This impact is substantial when compared to other estimates. However, account should be taken of the fact that only a minority of firms pays the maximum premium, which implies also that local treatment effects will

[16] Because we do not observe exactly the same information as UWV had when they calculated the premiums, the constructed DI risk and DI premium may be subject to measurement error.

Table 6
Coefficient estimates of the effect of the removal of experience rating on DI inflow and DI outflow with interaction terms of premium caps.

	DI inflow		DI outflow			
			First year		Second year	
Baseline specification	0.027**	(0.009)	−0.154**	(0.022)	−0.039	(0.024)
Estimation with interaction terms and risk premium						
Reference: pays premium below max	–	–	–	–	–	–
Pays the maximum premium	0.111**	(0.023)	−0.128**	(0.025)		
Removal of ER	0.030**	(0.005)	−0.166**	(0.022)	−0.051**	(0.024)
Risk percentage	0.081**	(0.039)	−0.054	(0.034)		
Risk percentage^2	−0.002	(0.005)	0.0004*	(0.0002)		
Risk percentage^3	0.0001	(0.0001)	−0.00001*	(0.000003)		

Estimations include the same control variables as in the main analysis. Standard errors in parentheses, for inflow estimations obtained using bootstrap with 500 replications.
* Significant at a level of 10%.
** Significant at a level of 5%.

Table 7
Coefficient estimates (average partial effect for DI inflow) of the effect of the removal of experience rating on DI inflow and DI outflow: Before and after 2005 and for different selections of DI spells before 2005.

	DI inflow		DI outflow			
			First year		Second year	
Before 2005	0.0005**	(0.0002)	−0.154**	(0.022)	−0.039	(0.024)
After 2005	0.0001	(0.0001)	0.068	(0.079)	0.053	(0.137)
Before 2005, different samples						
Exclusion DI spells, ≤35%	0.0005**	(0.0001)	−0.106**	(0.034)	0.016	(0.034)
Expansion sick leave period, >35%	0.0003**	(0.0001)	−0.047	(0.034)	0.084**	(0.040)

Every cell represents a different estimation. Estimations include the same control variables as in the main analysis. *Significant at a level of 10%. **Significant at a level of 5%.

apply only to a specific group of firms as well. In line with our earlier results, we also find DI inflow rates to be higher for the group of firms that is not experience rated. As to DI outflow, Table 6 also shows disincentive effects of the maximum premium. These effects are comparable in size to the effect of the removal of experience rating.

6.3.2. Experience-rating effects after 2005

We argued earlier that the reforms taking place after 2004 have changed the size as well as the composition of the cohort of (new) DI recipients in ways that may well have been different for the treatment and control group of firms. This is the reason why we restricted our analysis from 2001 to 2004. Still, we are able to perform a similar DiD analysis for the period between 2006 and 2011, which includes the re-introduction of experience rating for small firms in 2008. In this context, the treatment is defined as the absence of experience rating in 2006 and 2007. As the common trends assumption may well be more restrictive in the period after 2005, estimation results should be taken with caution (see Section 5.1).

Table 7 presents the coefficient estimate of the removal of experience rating that follows from this research design for 2006–2011, compared to the coefficient estimate that was obtained for the period before 2005. For both the DI inflow and DI outflow models, we find the effects of the removal of experience rating to be insignificant for the period after 2005. This suggests that firms became unresponsive to the experience-rating incentive. When interpreting this finding, recall that the DI scheme and the incentive of DI experience rating differ between the periods before and after 2005 at least in three ways. First, in the new DI scheme that started in 2006, experience rating no longer applies to individuals with a disability degree of less than 35% — as these are excluded from DI benefits in the new scheme. It is likely that this change increased the share of workers in DI with bad job prospects. Second, in 2005, the period of continued wage payments during sickness was extended from one to two years. This reform may have decreased the (additional) effect of experience rating as well, as re-employment probabilities usually decrease over time. Third, both the range of the experience-rating premiums as well as the level of the maximum premiums decreased substantially after 2005 (see Fig. 3), causing the effective impact of the experience-rated premium on the employers wage costs to decrease accordingly.

With this in mind, the pertinent question is how changes in the size and composition of the DI inflow since 2005 have affected the impact of experience rating. To shed light on this question, we re-estimated our benchmark model for the pre-2005 period for the sample of workers that would still be entitled to DI benefits in the post-2005 period. Stated differently, we exclude from our sample those workers who would no longer have been entitled to DI benefits in the post-2005 period.

When following this strategy, we obtain coefficient estimates for the DI inflow and DI outflow models that are presented in the lower panel of Table 7. According to the table, the exclusion of workers with disability degrees below 35% does not significantly affect our model estimates for the DI inflow and DI outflow models. When excluding workers with DI spells that are shorter than one year, however, the effect estimates for the pre-2005 period become significantly smaller. The average partial effect on DI inflow drops from 0.0005 to 0.0003, whereas the effect on DI outflow in the first year becomes insignificant. This suggests that the lower impact of DI experience in the post-2005 period is partially due to the extension of the sickness period that precedes DI.[17]

7. Conclusion

This paper studies the effect of firm experience rating on DI inflow and DI outflow in The Netherlands, using matched firm- and worker data. We exploit the removal of experience rating for small firms in 2003, which allows us to use a difference-in-difference

[17] At the same time, there are reasons to believe that the impact of the extension may be underestimated. In particular, it is likely that financial incentives due to wage continuation in the sickness period are perceived by employers as more direct than the delayed impact of experience rating.

design. Our focus is on the period until 2005, as there were other reforms in 2005 in 2006 that may well have affected small and large firms in different ways. In particular, the 2005 reform extended the sickness benefit period that precedes DI claims from one to two years, and the 2006 reform split the disability scheme into separate schemes for permanently and fully disabled individuals and for temporarily and/or partially disabled individuals.

Our main finding is that the removal of experience rating in 2003 increased the DI inflow for small firms by about 7%, whereas DI outflow of individuals from small firms decreased by about 12%. We estimate that the DI stock in 2004 was 0.4% larger because of the reform. As to DI inflow, our results are about half the size of the effects on inflow found by Koning (2009) and Van Sonsbeek and Gradus (2013). Moreover, there is strong evidence that the decrease in DI outflow for the treatment group of small firms is confined to partially disabled workers and workers with relatively high DIbenefits. Interestingly, we also find evidence that the cap that was used for experience-rated premiums had substantial disincentive effects. That is, firms paying the maximum premium had higher DI inflow rates and lower DI exit rates, suggesting that they responded to the absence of prevention and reintegration incentives (at the margin).

We also broadened our perspective by assessing the specific context that may or may not have contributed to the effectiveness of experience rating. We thus estimated our model for the period after 2005, exploiting the re-introduction of experience rating for small firms in 2008. We then found no evidence of experience-rating effects, on either DI inflow or DI outflow. To investigate the potential role of post-2005 reforms in explaining these outcomes, we re-estimated our benchmark model for the pre-2005 period omitting the workers that would no longer have been entitled to DI benefits in the post-2005 period. Based on this analysis, we argue that particularly the extension of the sickness benefit period from one to two years has lowered the potential impact of experience rating on both DI inflow and DI outflow.

Acknowledgments

We would like to thank two anonymous referees, Anne Gielen, Philip de Jong, Maarten Lindeboom, Jan-Maarten van Sonsbeek and participants of the SOLE/EALE 2015 Conference, the 11th IZA Conference on Labor Market Policy Evaluation, the VU University lunch seminar and the VU-UVA Labour Economics seminar for useful comments and suggestions. UWV and Statistics Netherlands are gratefully acknowledged for giving access to the data. This study was funded by Instituut Gak.

Appendix A. Full estimation results of baseline specifications

Table A1
Fractional probit estimations for the fraction of workers per firm that receives with DI benefits (2001–2004) and Cox proportional hazard estimates of outflow from DI, for individuals who entered DI between 2001 and 2004.

	DI inflow		DI outflow	
Effects experience rating				
Removal of ER	0.027**	(0.009)	–	–
Removal of ER, first year after inflow	–	–	−0.154**	(0.022)
Removal of ER, second year after inflow	–	–	−0.039	(0.024)
Firm characteristics				
Small firm	0.041	(0.040)	−0.037**	(0.014)
Middle-sized firm	0.040	(0.024)	0.029	(0.019)
Average age	0.007**	(0.001)	–	–
Percentage of men	−0.031	(0.047)	–	–
Percentage of immigrants	0.063	(0.056)	–	–
Percentage of single households	0.054	(0.040)	–	–
Percentage of single parents	0.031	(0.048)	–	–
Percentage of parents	0.089**	(0.019)	–	–
Annual wage below €7500	0.372**	(0.047)	–	–
Annual wage of€7500–15,000	0.333**	(0.044)	–	–
Annual wage of€15,000–25,000	0.255**	(0.042)	–	–
Annual wage of€25,000–40,000	0.164**	(0.040)	–	–
Sector				
- Agriculture	0.089**	(0.019)	−0.029	(0.031)
- Industry	0.180**	(0.014)	−0.104**	(0.032)
- Government	0.131**	(0.013)	−0.025	(0.033)
- Construction	0.375**	(0.015)	−0.183**	(0.038)
- Trade	0.130**	(0.013)	0.013	(0.032)
- Food	0.033**	(0.017)	−0.019	(0.035)
- Transport	0.222**	(0.019)	0.133**	(0.035)
- Financial	0.255**	(0.061)	0.253**	(0.057)
- Business	0.116**	(0.015)	−0.055*	(0.033)
- Education	0.095**	(0.017)	−0.065*	(0.034)
- Health care	0.110**	(0.015)	−0.008	(0.031)
Worker characteristics				
Age, 25–35	–	–	−0.086**	(0.024)
Age, 35–45	–	–	−0.291**	(0.024)
Age, 45–55	–	–	−0.592**	(0.024)
Age, 55–65	–	–	−0.771**	(0.025)
Male	–	–	0.005	(0.010)
Single household	–	–	0.026	(0.033)

Table A1 (*continued*)

	DI inflow		DI outflow	
Couple	–	–	−0.029	(0.032)
Single parent	–	–	0.050	(0.035)
Has children	–	–	0.152**	(0.010)
Wage, €10,000–20,000	–	–	0.052**	(0.011)
Wage, €20,000–30,000	–	–	0.114**	(0.012)
Wage, €30,000–40,000	–	–	0.226**	(0.016)
Wage, €40,000–50,000	–	–	0.249**	(0.025)
Wage, >€50,000	–	–	0.189**	(0.022)
Year effects	Yes		Yes	
Regional dummies	Yes		Yes	
Observations	183,665		119,631	
Log pseudolikelihood	−30,352		−689,144	

References

Autor, D.H., Duggan, M.G., 2007. Distinguishing income from substitution effects in disability insurance. Am. Econ. Rev. Am. Econ. Assoc. 97 (2), 119–124.

Autor, D.H., Duggan, M.G., 2010. Supporting Work: A Proposal for Modernizing the U.S. Disability. The Center for American Progress and The Hamilton Project.

Bruce, C.J., Atkins, F.J., 1993. Efficiency effects of premium-setting regimes under workers' compensation: Canada and the United States. J. Labor Econ. 11 (1), S38–S69. Part 2: U.S. and Canadian income.

Böheim, R., Leoni, T., 2011. Firms Moral Hazard in Sickness Absences. Economics Working Papers 2011–2013. Department of Economics, Johannes Kepler University Linz, Austria.

Campolieti, M., Hyatt, D., Thomason, T., 2006. Experience rating, work injuries and benefit costs: some new evidence. Relat. Indust. Ind. Relat. 61 (1), 118–145.

De Jong, Ph., Lindeboom, M., van der Klaauw, B., 2011. Screening disability insurance. J. Eur. Econ. Assoc. 9 (1), 106–129.

Deelen, A.P., 2005. Adverse selection in disability insurance: empirical evidence for Dutch firms. CPB Discussion Paper No. 46.

Fevang, E., Markussen, S., Røed, K., 2014. The sick pay trap. J. Labor Econ. 32 (2), 305–336.

Hassink, W.H.J., Koning, P.W.C., Zwinkels, W., 2015. Employers opting out of disability insurance: selection or incentive effects. IZA Discussion Paper Series No. 9181.

Hyatt, D.E., Thomason, T., 1998. Evidence on the Efficacy of Experience Rating in British Columbia A Report to The Royal Commission on Workers′ Compensation in BC. Unpublished manuscript.

Ison, T.G., 1986. The significance of experience rating. Osgoode Hall Law J. 24 (4), 723–742.

Kok, L., Heyma, A., Lammers, M., 2013. Verlaag kosten loondoorbetaling voor kleine bedrijven. TPE Digit. 7 (3), 4–17.

Koning, P.W.C., 2009. Experience rating and the inflow into disability insurance. De Economist 157 (3), 315–335.

Koning, P.W.C., Lindeboom, M., 2015. The rise and fall of disability insurance enrollment in the Netherlands. J. Econ. Perspect. 29 (2).

Koning, P.W.C., van Vuuren, D., 2007. Hidden unemployment in disability insurance. Labour 21 (4), 611–636.

Koning, P.W.C., van Vuuren, D., 2010. Disability insurance and unemployment insurance as substitute pathways. Appl. Econ. 42 (5), 575–588.

Korkeamäki, O.I., Kyyrä, T., 2012. Institutional rules, labour demand and retirement through disability programme participation. J. Popul. Econ. 25 (2), 439–468.

LaDou, J., 2011. The European influence on workers' compensation reform in the United States. Environmental Health 10(103).

Lengagne, P., 2014. Workers compensation insurance: incentive effects of experience rating on work-related health and safety. Irdes Working Paper 64.

Lippel, K., 1999. Therapeutic and anti-therapeutic consequences of workers' compensation. Int. J. Law Psychiatry 22 (5–6), 521–546.

Marie, O., Vall Castello, J., 2012. Measuring the (income) effect of disability insurance generosity on labour market participation. J. Public Econ. 96, 198–210.

McCrary, J., 2008. Manipulation of the running variable in the regression discontinuity design: a density test. J. Econ. Lit. 142 (2), 698–714.

OECD, 2010. Sickness, Disability and Work: Breaking the Barriers. A Synthesis of Findings Across OECD Countries. OECD Publishing, Paris.

Papke, L.E., Wooldridge, J.M., 2008. Panel data methods for fractional response variables with an application to test pass rates. J. Econ. 145, 121–133.

Ruser, J.W., 1985. Workers' compensation insurance, experience-rating, and occupational injuries. RAND J. Econ. 16 (4), 487–503.

Ruser, J.W., 1991. Workers' compensation and occupational injuries and illnesses. J. Labor Econ. 9 (4), 325–350.

Ruser, J., Butler, R., 2009. The economics of occupational safety and health. Found. Trends Microecon. 5 (5), 301–354.

Seabury, S.A., McLaren, C.F., Reville, R., Neuhauser, F., Mendeloff, J., 2012. Workers' compensation experience rating and return to work. Policy Pract. Health Saf. 10 (1), 97–115.

Strunin, L., Boden, L.I., 2004. The workers' compensation system: worker friend or foe? Am. J. Ind. Med. 45 (4), 338–345.

Tompa, E., Cullen, K., McLeod, C., 2012. Update on a systematic literature review on the effectiveness of experience rating. Policy Pract. Health Saf. 2, 47–65.

Van Sonsbeek, J.M., Gradus, R.H.J.M., 2013. Estimating the effects of recent disability reforms in the Netherlands. Oxf. Econ. Pap. 832–855.

Wooldridge, J.M., 2010. Correlated random effects models with unbalanced panels. Manuscript.