Employer Responsibility in Disability Insurance: Evidence from the Netherlands*

Daniel Prinz
Harvard University
(Job Market Paper)

Bastian Ravesteijn
Erasmus University Rotterdam

November 1, 2020
Click here for most recent version

Abstract

We develop a framework to study optimal disability insurance when employers exhibit moral hazard and show that the optimal system takes into account employer-side moral hazard and selective hiring. We illustrate these insights using a reform in the Netherlands that extended experience rating to temporary workers. Using this reform, we document a 24% decrease in disability inflow. We also find an increase in worker selection, accounting for 14% of the overall decrease. Using our model, we evaluate the normative implications of the experience rating policy. We conclude that, given reasonable assumptions, the policy improved welfare and additional employer responsibility would further add to social welfare.

*Prinz: dprinz@g.harvard.edu. Ravesteijn: ravesteijn@ese.eur.nl. We thank Anikó Bíró, Sam Burn, David Cutler, Ed Glaeser, Nathan Hendren, Yunan Ji, Judit Krekó, Tim Layton, Attila Lindner, Nicole Maestas, Tom McGuire, Sam Moy, Ljubica Ristovska, Mark Shepard, Liyang Sun, and seminar participants at the Harvard Proseminar on Inequality and Social Policy, the Harvard Medical School Department of Health Care Policy Health Economics Seminar, the Harvard PhD Seminar in Health Policy, the Harvard Health Economics Tea, the Harvard Kennedy School Applied Microeconomics Seminar, the Harvard Center for Population and Development Studies, the Harvard PhD Seminar in Labor and Public Economics, the NBER Aging and Health Trainee Seminar, the Boston University Health Economics Reading Group, the Hungarian Academy of Sciences Institute of Economics Summer Workshop, and the University of Groningen for useful comments. We gratefully acknowledge financial support from the Social Security Administration and the Disability Research Center at the National Bureau of Economic Research through a Pre-Doctoral Fellowship in Disability Policy Research; the Harvard Kennedy School Malcolm Wiener Center for Social Policy through the Multidisciplinary Program in Inequality & Social Policy and the James M. and Cathleen D. Stone Ph.D. Scholarship in Inequality and Wealth Concentration; and the Horowitz Foundation for Social Policy through its grant program and the Robert K. Merton Award. The findings and conclusions expressed are solely those of the authors and do not represent the opinions or policy of the Social Security Administration, any agency of the Federal Government, the National Bureau of Economic Research, or Statistics Netherlands.
1 Introduction

The dramatic rise of disability insurance (DI) participation over the last several decades has caused concern among policymakers in the United States and other developed nations. In the U.S., 8.5 million former workers received Social Security Disability Insurance (SSDI) benefits in 2018, up from 2.8 million in 1988 and 4.7 million in 1998. Over the same time period, spending on benefits more than tripled in real terms (Social Security Administration, 2019). These increases in DI participation occurred while the physical demands of occupations decreased (Johnson, 2004; Pew Research Center, 2016) and medical care improved. Given these developments, it is important to understand what policies could increase the productive capacity of the labor market. One likely important, but poorly understood factor is the behavior of employers.

Employers can take a variety of actions that impact the work capacity and productivity of their workers, as well as their attachment to the labor force. Employment conditions may impact the development of medical conditions that can qualify a worker for DI benefits. Employers can also impact the trajectory of a worker after the development of such conditions, by retaining workers at a decreased level of productivity, and/or by offering accommodations. Employers can exert effort to influence workers’ DI claims and that absent appropriate incentives, there is underprovision of such effort, a form of moral hazard. This employer-side moral hazard is separate from worker-

---

1 Accommodations are defined as any change in the work environment or in the way things are customarily done that enables an individual with a disability to enjoy equal employment opportunities (Code of Federal Regulations, 1991; Equal Employment Opportunity Commission, 2002). Depending on the type of condition, such accommodations may include accessible worksites and workstations, modified schedules, job restructuring, retraining, and reassignment, among other policies.
side moral hazard, the idea that DI distorts labor supply decisions. Employer-side moral hazard becomes an important concern if employer effort to prevent DI entry is imperfectly observable by the insurer and is not fully contractible between employers and workers. This moral hazard means that the DI benefit receipt of each employer’s former workers imposes an externality on all other beneficiaries in the form of higher DI premiums. For example, in the U.S., if the Social Security Administration’s Disability Insurance Trust Fund is to stay solvent, payroll taxes need to cover benefit outlays, and in principle more spending on benefits leads to higher payroll taxes, lower net wages and profits, and dead-weight loss from taxation. The introduction of incentives in the form of experience rating should make employers internalize these costs and exert greater effort to retain workers who would otherwise receive DI benefits. This in turn translates into higher overall productivity, lower payroll taxes required to keep the insurance program solvent, and consequently higher net wages, profits, and lower deadweight loss from taxation.

An important concern about experience rating is that it increases the expected cost of employing workers who are more likely to experience health-related productivity losses. If employers can observe information about candidates that is predictive of a future health deterioration, then experience rating strengthens their incentives to select workers along these dimensions in the hiring process, resulting in inefficient under-employment of some workers and reducing aggregate productivity.2

This paper develops a framework to understand the trade-off between these two effects, the intended retention effect and the unintended selection effect. We show that absent employer responsibility for DI costs, employers exhibit moral hazard: there is underprovision of effort that could decrease workers’ DI entry in a socially efficient way. This employer-side moral hazard generates a welfare loss because all employers and workers bear the costs of lost productivity and higher DI spending and higher premiums (in the U.S., in the form of

---

2 Another potential concern is that such a policy imposes greater risk on employers, who face a financial penalty when any of their workers claim DI, even in cases where an exogenous health shock reduces work capacity to zero. This may be problematic especially among small employers.
payroll taxes) generate dead-weight loss. When employers become (partially) responsible for their former workers’ DI costs, they internalize (part of) this cost, but they may also engage in stronger selection of workers. Workers who are more likely to need DI benefits become more expensive to employ, so unless they are able to fully pass through these higher costs to wages, employers are less likely to employ them. If employers behave in a risk-neutral way, then the optimal employer cost sharing depends on the parameters governing retention and selection responses.

We illustrate these mechanisms empirically using a 2012-2013 reform in the Netherlands, which extended experience rating for both short-term and long-term DI to fixed-term and temporary workers. Experience-rated premiums were capped at the industry level at relatively low rates, with the exception of one industry: the temporary work agency industry which, due to relatively short contract durations and high DI inflow, had been an important contributor to DI rolls. We compare changes in DI benefit receipt by temporary work agency workers to changes in DI benefit receipt by permanent workers (for whom DI was already experience-rated) in a difference-in-differences framework. We find that DI benefit receipt declined by 24% among temporary work agency workers due to the policy. We also find large decreases in short-term DI use. Marginal workers remained employed at higher rates rather than dropping out of the labor force or entering other programs, such as unemployment insurance.

Our results on short-term DI suggest that the combination of experience rating short-term and long-term DI jointly reduced DI benefit receipt. We view this as an important lesson for policy: employer-side incentives may be most effective when they are not only strong but immediate. When short-term DI became experience-rated firms had the incentive to facilitate workers’ return to work in the short-term which they may be better equipped to do than to implement policies that change outcomes that only manifest in the longer term (e.g., DI benefit receipt two years later). Qualitative information based on discussions with temporary work agencies suggest that they invested in services that helped workers return
to work earlier and remain employed.

Examining heterogeneity in outcomes, we find that the decline was larger for mental health conditions and among women and middle-aged workers. Relative declines were larger among higher-earning and healthier temporary work agency workers. This suggests that the reform may have succeeded in increasing the employment of workers with relatively high remaining working capacity, a key goal of policies that incentivize the retention of workers who could potentially receive DI.

Importantly, we also find that while DI benefit receipt was reduced, there is also evidence that the reform caused employers to modify hiring practices to avoid workers at high risk of entering DI. To generate this result, we rely on workers’ prior healthcare spending, which we find is highly predictive of benefit receipt. The likelihood that a temporary work agency worker was in the top quartile of the distribution of predicted DI benefit receipt based on the previous year’s healthcare spending decreased by 4.2%. The likelihood that they were in the top decile, decreased by 3%. Using healthcare spending to predict DI benefit receipt, we find that changes in worker composition can explain 14% of the overall decline in DI benefit receipt due to the policy, suggesting that both the intended retention effect and the unintended selection effect of experience rating may be quantitatively important, implying an important trade-off for policy.

Viewing these empirical estimates through the lens of our model, the introduction of experience rating improved welfare. Our model suggests that at the optimal level of experience rating the retention and selection effects are balanced. Since we find that the retention effect is substantially larger than the selection effect, our model suggests that the degree of experience rating introduced improved welfare relative to the previously existing flat-rate system, reducing lost productivity, and a higher degree of experience rating would represent further improvement. Under strong assumptions we can extrapolate our estimates to conclude that full experience rating could be optimal in this setting. Whether these results would hold in a setting with workers more attached to their employers is an open question: on the one hand,
due to high turnover, selection may be easier in the temporary worker setting we study, but on the other hand, longer-term employment relationships mean that employing less healthy workers is costlier.

Our work contributes to a small literature on the consequences of employer-side incentives in disability insurance (Acemoglu and Angrist, 2001; Koning, 2009; Korkeamäki and Kyyrä, 2012; Kyyrä and Tuomala, 2013; De Groot and Koning, 2016; Kyyrä and Paukkeri, 2018; Hawkins and Simola, 2020).\textsuperscript{3} We also build on the literature that has discussed the potential advantages and drawbacks of such policies, and has argued that employer-side incentives should be introduced (Autor and Duggan, 2010; Autor, 2011; Burkhauser and Daly, 2011; Liebman, 2015). The only existing studies that we are aware of on the consequences of employer cost sharing in DI that rely on policy variation are Koning (2009), De Groot and Koning (2016) studying the Netherlands, and Hawkins and Simola (2020) studying Finland, which study reforms in the Netherlands and Finland. All three studies find that experience rating reduces the DI flows of the impacted employers and Hawkins and Simola (2020) also find an increase in worker selection. Other studies from Finland (Korkeamäki and Kyyrä, 2012; Kyyrä and Tuomala, 2013; Kyyrä and Paukkeri, 2018) find mixed evidence on the impact of experience rating on benefit receipt. We advance this literature in several ways. First, we develop a novel framework to analyze the incentives created by employer cost sharing, and to derive optimal policies. Second, we exploit a recent and large reform in the Netherlands and use administrative data on earnings and disability receipt, as well as on health, healthcare utilization, and the qualifying diagnosis of each enrollee. Third, our focus on the population of temporary workers is also novel when studying social insurance policies in general and DI policies in particular. This population is becoming increasingly important

\textsuperscript{3}A somewhat larger literature, summarized by Tompa, Cullen and McLeod (2012), examines employer-side incentives in workers’ compensation. Unemployment insurance (UI) is also partially experience rated in the U.S.. Anderson and Meyer (1993) find that under imperfect experience rating, there is persistent variation across industries and across employers within the same industry in UI claims. Anderson and Meyer (2000) show that when experience rating is strengthened, UI claims are reduced. In recent work, Johnston (2020) finds that unemployment insurance tax increases reduce employer hiring and employment, but have no effect on layoffs or worker earnings.
(Collins et al., 2019; Katz and Krueger, 2019a,b), and there has been little work on how to design social insurance policies for these types of contingent workers. Our results suggest that employer incentives may be especially important for these types of workers as employers seem highly responsive to incentives and these employees are more weakly attached to the labor force.

We complement the recent literature that studies the aggregate economic sources of disability applications (e.g., Autor and Duggan, 2003; Maestas, Mullen and Strand, 2015, 2018) and the employment consequences of disability receipt (e.g., Maestas, Mullen and Strand, 2013; French and Song, 2014; Autor, Maestas, Mullen and Strand, 2015; Gelber, Moore and Strand, 2017). This literature focuses on individual workers, largely taking the role of employers as given. We do the opposite, focusing on the employer sources of disability receipt behavior, while holding worker-side incentives fixed.

Our theoretical framework contributes to the literature on the optimal design of social insurance policies. We complement models of optimal social insurance with worker-side moral hazard in UI (Baily, 1978; Chetty, 2006) and DI (Diamond and Sheshinski, 1995) and extend the literature on employer responsibility developed in the unemployment insurance context (Topel, 1984; Burdett and Wright, 1989; Fath and Fuest, 2005; Blanchard and Tirole, 2008) by incorporating worker heterogeneity and by applying these models to the DI context.

Our work is also connected to recent research on alternative work arrangements (Collins et al., 2019; Katz and Krueger, 2019a,b). Contingent, flexible work often comes with lower-quality benefits and less employer responsibility for worker outcomes such as disability or health. Workers participating in these arrangements are also different. For example, they are less educated and have weaker labor market attachment. In recent work examining how workers in the contingent workforce and with alternative work arrangements fare in disability programs, Broten, Dworsky and Powell (2018) document larger disability and employment impacts of workplace injuries for temporary and contract workers than for

---

4 For an overview of this literature, see Chetty and Finkelstein (2013).
5 For recent structural work, see Kim and Rhee (2018) and Aizawa, Kim and Rhee (2020).
permanent workers. Rutledge, Zulkarnain and King (2018) find that in general, likely due to differences in eligibility, they are less likely to apply for and be awarded disability benefits. In this paper, we examine the consequences of increasing the responsibility that employers take for their contingent workers. Our findings suggests that policymakers can change the behavior of employers towards their contingent workforce if incentives are sufficiently strong and direct. Many characteristics of these workers, in particular their weaker labor force attachment, may make it difficult to offer them social insurance benefits like DI. Experience rating may make such benefits feasible by requiring employers to cover the costs of benefits.

Finally, we build on our own work on the employment and industry sources of disability insurance claims (Maestas, Prinz and Ravesteijn, 2018; Lurie, Maestas, Miller and Prinz, 2019). In this work, we have documented variation across industries in disability insurance benefit receipt and studied the consequences of health shocks by industry, as well as the relationship between wage premiums and disability insurance benefit receipt. This paper advances our understanding of the role of employers in disability insurance benefit receipt by developing quasi-experimental evidence based on a policy reform and a framework in which estimates can be interpreted.

The remainder of this paper proceeds as follows. In Section 2 we develop a framework that allows us to analyze retention and selection responses by employers to employer cost sharing in DI. We proceed by describing the institutional background of disability insurance programs in the Netherlands and the 2012 reform in Section 3. We then describe our data in Section 4 and methods in Section 5. We present our results in Section 6 and welfare analysis in Section 7. Section 8 concludes.
2 Theoretical Framework

2.1 Motivation

Most disability insurance systems, including the U.S. Social Security Disability Insurance (SSDI) program, are currently financed through a flat-rate payroll tax. This means that firms pay a fixed percentage of their wage bill as disability insurance premium. Several proposals to reform this system include elements that would incentivize employers to decrease the number of their workers who receive benefits, including proposals to experience rate DI premiums (Autor and Duggan, 2010; Autor, 2011; Burkhauser and Daly, 2011; Liebman, 2015).

The economic idea behind experience rating is that employers exhibit moral hazard because absent appropriate incentives, there is an underprovision of effort by employers to reduce the rate at which their former workers claim DI benefits. This employer-side moral hazard is separate from worker-side moral hazard, the idea that DI uptake by workers is responsive to the generosity of benefits. Employer-side moral hazard becomes an important concern if employer effort to prevent DI uptake is imperfectly observable by the insurer and is not (fully) contractible between employers and workers. (If it were contractible between employers and workers, then the insurer could incorporate incentives for employer effort in the worker-side replacement rate.) Employer-side moral hazard leads to more workers receiving DI payments, which drives up overall DI costs and imposes an externality on all employers and workers in the economy. Experience rating can make employers internalize these costs and retain more workers who would otherwise receive DI payments. This is the intended retention effect of employer cost sharing policies.

There are two concerns that have been raised about such proposals. First, they could expose employers, especially small employers, to excess risk. Second, they could lead to employers changing the composition of their workforce. If employers can observe information about potential workers that is predictive of their future DI uptake and hence their expected
cost to the employer, they could choose to hire or retain workers who are less likely to receive DI benefits. This becomes a problem if the insurer does not observe the predictive information. If the insurer could also observe the same information, they could use it to risk adjust experience-rated premiums. This is the unintended selection effect of employer cost sharing policies.

In this section we develop a framework in which these different effects of employer cost sharing, and in particular employers’ retention and selection responses, can be studied in order to understand the welfare consequences of experience-rated DI premiums.

2.2 Setup

The economy consists of measure 1 of risk-averse workers, characterized by their health-related productivity type $\theta$ and (at least) measure 1 of risk-neutral firms. An employer first decides whether to hire a worker, observing the productivity distribution $F_\theta(\lambda)$ of that worker. Next, once a worker is hired, productivity is revealed and the employer decides whether to retain the worker, in which case the worker produces $\lambda$.

In order to simplify the exposition, in the model, we abstract from the time dimension of the revelation of productivity $\lambda$. This is somewhat different from what plays out in most employment relationships, where disabling conditions are revealed over time and workers remain productive for some time after being hired. While this is intuitive, it does not add much to the model since worker output between the time of hiring and the realization of a disabling condition can be summarized by $\lambda$. If one worker remains productive for longer before a health-related productivity decline than another worker with the same decline, we can just say that the first worker has a higher realized $\lambda$.

In case the worker is retained, she earns wage $w$, and in case she is not retained, she is paid a disability benefit of $\rho w$ where $\rho$ is the replacement rate. In a competitive market, employers will make zero profit, but have some opportunity cost or fixed cost $M$ of setting up their business.
Workers derive utility $u(c)$ from their consumption $c$. Thus if they are employed, their utility is $u(w)$ and if they receive disability benefits, their utility is $u(\rho w)$. If there were no DI benefits, workers who are not employed would have utility $u(b)$, which means that $b$ is the wage equivalent of non-employment.

The benefits $\rho w$ of a worker who enters DI are financed from employer cost sharing $\eta \rho w$ plus a flat-rate payroll tax. Experience rating $0 \leq \eta \leq 1$, which can be viewed as a special case of cost sharing in the wider insurance literature, is the parameter that defines the extent to which an employer’s DI premiums depends on the DI uptake of the employer’s former workers. Even when paying an actuarially fair premium, the employer does not take into account the impact of his own actions on the non-experience-rated part of the premium.

### 2.3 Worker Heterogeneity

Workers are heterogeneous in two ways: in terms of their \textit{ex ante} productivity distribution summarized by their type $\theta$, and their realized productivity level $\lambda$. A worker of type $\theta$ has an \textit{ex ante} health-related productivity distribution $F_{\theta}(\lambda)$ (density $f_{\theta}(\lambda)$) on $[0, 1]$. The insurer does not observe $\theta$ but employers and workers do. In practice, the types $\theta$ correspond to worker characteristics that are informative about future health-related productivity but that employers and workers are not allowed to condition wages on in their contracts. For example, past and current health status, age, gender, smoking status, lifestyle could predict what future health-related productivity draws a worker is likely to experience but employers may be legally barred from using them to offer different contracts. \textit{Ex ante} health types $\theta$ are distributed according to $G(\theta)$ (density $g(\theta)$) on $[0, 1]$ (for notational simplicity). Employers have a common ranking of types: all employers prefer a worker of health type $\theta'$ to a worker of health type $\theta$ whenever $\theta' > \theta$. This is a general formulation that incorporates several special cases. For example, if employers are risk neutral, then they rank \textit{ex ante} health types based on expected health-related productivity: $\theta \equiv E_{\theta}[\lambda]$. If employers behave as risk averse, then they rank \textit{ex ante} health types based on expected health-related productivity
and the risk premium. The key is that $\theta$ summarizes the *ex ante* ranking of workers for all employers.

Second, workers are *ex post* heterogeneous in their realized health-related productivity, denoted by $\lambda$. This is the health-related productivity that a worker ultimately draws from her *ex ante* distribution $F_\theta(\lambda)$. The insurer does not observe $\lambda$ but employers and workers do. In practice, the realized health-related productivity $\lambda$ corresponds to health events that workers experience and which change their productivity. For example, a worker might experience a health deterioration that changes their ability to perform certain physical tasks at work or develop a mental health condition that makes it more difficult to continue working effectively. The general health-related productivity term $\lambda$ reflects net productivity, after any costs of retraining and/or accommodation have been deducted.

The distinction between *ex ante* health type $\theta$ and *ex post* realized productivity $\lambda$ is that $\theta$ is what the employer observes when they decide to hire or not hire a worker, while $\lambda$ is additionally observed when they decide whether to retain or not retain a worker. We argue that choices about which *ex ante* worker health types $\theta$ are hired correspond to (in)efficiency in selection, while choices about which workers to keep based on *ex post* realized health-related productivity $\lambda$ correspond to (in)efficiency in retention. Having these two dimensions of heterogeneity is useful to capture both selection and moral hazard. If there was no *ex ante* heterogeneity in expected costs (i.e., heterogeneity in $\theta$) or this heterogeneity was not observable to firms, there would be no selection, only moral hazard, as in the benchmark case discussed in the next subsection. If there was no heterogeneity in realized productivity (i.e., heterogeneity in $\lambda$), because type $\theta$ was perfectly deterministic, then there would be no moral hazard, only selection.

2.4 Benchmark: No *Ex Ante* Worker Heterogeneity

We start with a benchmark case with no *ex ante* heterogeneity across workers. (A case with *ex ante* heterogeneity that is not observed by employers would be equivalent.) All workers’
realized health-related productivity is drawn from the same distribution: \( F_\theta(\lambda) = F(\lambda) \) for all \( \theta \). That is, we assume that all workers look the same at the point when employers decide whether to hire them and the only policy-relevant decision that employers make is whether to retain a worker after \( \lambda \) is realized.

Employers make their retention decisions based on the observed health-related productivity \( \lambda \) of a worker. Employers retain a worker if her productivity net of wage is larger than the employer’s share of the benefit costs:

\[
\frac{\lambda - (1 + \tau)w}{-\eta \rho w}.
\]

Recall that realized health-related productivity \( \lambda \) is a general term that can capture a variety of underlying differences in productivity and costs. For example, workers may get a health shock such that they are able to produce output \( l(h) \) and impose an accommodation cost \( a(h) \) on the employer in order to keep working. In this case \( \lambda = l(h) - a(h) \). The condition in Equation (1) implies a health-related productivity threshold \( \bar{\lambda} = (1 + \tau - \eta \rho)w \) above which workers will be retained once their health-related productivity is realized. What this condition misses is the full cost of benefits: \( \rho w = \eta \rho w + (1 - \eta) \rho w \). When they make their retention decision, firms only care about the cost sharing component \( \eta \rho w \), even though through the flat-rate payroll tax \( \tau \), the full cost of benefits will impact all firms in the economy. This is the source of the potential inefficiency in this system.

Employers pay wage \( w \) independent of realized health-related productivity \( \lambda \). If workers have no private information on their future realized health-related productivity \( \lambda \) and no ability to engage in \textit{ex post} wage bargaining after their health-related productivity is revealed, the optimal contract between risk-averse workers and risk neutral employers would set the wage \( w \) independent of the realized health-related productivity of the worker.

From the social planner’s point of view, a worker should be retained if her health-related productivity is higher than the value of non-employment \( b \). The threshold \( \bar{\lambda} \) is in general
higher than the socially efficient threshold $b$. The lost productivity from the inefficiently high retention threshold for health-related productivity $\bar{\lambda}$ is

$$L(\eta) = \int_{b}^{1} \lambda dF(\lambda) - \int_{\lambda}^{1} \lambda dF(\lambda) = \int_{b}^{\bar{\lambda}} \lambda dF(\lambda).$$  \hspace{1cm} (2)

This lost productivity, illustrated in Figure 1, is the driver of welfare loss from moral hazard. It lowers overall payments to workers in the form of wages and benefits.

The wage $w$ must satisfy the overall resource constraint in the economy:

$$\frac{F(\bar{\lambda})}{\text{Disabled}} \rho w + \frac{(1 - F(\bar{\lambda}))}{\text{Working}} w = \int_{\lambda}^{1} \lambda dF(\lambda) - M$$ \hspace{1cm} (3)

or

$$w = \frac{\int_{\lambda}^{1} \lambda dF(\lambda) - M}{\rho F(\bar{\lambda}) + (1 - F(\bar{\lambda}))}$$ \hspace{1cm} (4)

where we allow the term $M$ to capture employers’ opportunity cost and fixed cost of production. We focus on lost productivity as a driver of welfare loss because there is a mapping from total productivity to wages and benefits which drive worker utility. Focusing on productivity has the advantage that we can directly examine the implications of retention decisions based on realized productivity $\lambda$ (and in the next subsection based on the $ex \ ante$ productivity distribution $F_\theta(\lambda)$).

The welfare loss from lost productivity is minimized by setting $\eta^* = 1$ and making employers fully responsible for benefit costs. To see this, consider Equation (1): with $\eta = 1$, the right hand side of the equation is the full cost of benefits: a worker is retained if her net of wage productivity is higher than the full social cost of her disability benefits. With any $\eta < 1$, it is always optimal to increase $\eta$ towards 1, moving $\bar{\lambda} = (1 + \tau - \eta \rho)w$ towards $b$ and decreasing the lost productivity $L(\eta)$. This is also illustrated in Figure 1. Two important assumptions are needed for this result: the risk neutrality of employers and no $ex \ ante$ heterogeneity of workers. We now relax the latter assumption.
2.5 *Ex Ante* Heterogeneous Workers and Selection

We now consider the case where workers are *ex ante* heterogeneous in their health type $\theta$ and employers are able to choose whether they hire a worker of health type $\theta$.

Notice that the socially optimal employment threshold $b$ and the threshold used by employers $\bar{\lambda}$ do not depend on worker health type $\theta$. That is, whether a worker remains employed (or should remain employed) only depends on her realized productivity, not on her *ex ante* health type. The key change from the baseline case is that employers may not hire some worker health types because their expected value to the employer net of their wage and benefits (for a risk neutral employer, the expected profit) is not sufficiently high.

We illustrate this in Figure 2. In this example, there are two worker health types $\theta_1$ and $\theta_2$ with their *ex ante* productivity distributions in blue and red, respectively. Their realized health-related productivity is drawn from distributions $F_{\theta_1}(\lambda)$ (density $f_{\theta_1}(\lambda)$) and $F_{\theta_2}(\lambda)$ (density $f_{\theta_2}(\lambda)$), respectively. Employers set their retention threshold at productivity level $\bar{\lambda}$ for each worker type, higher than the socially optimal level $b$. For each type $\theta$, the expected total lost productivity is the integral of the density of the productivity distribution between the socially optimal employment threshold $b$ and the threshold set by employers $\bar{\lambda}$:

$$L(\eta) = \int_b^1 \lambda dF_\theta(\lambda) - \int_1^{\bar{\lambda}} \lambda dF_\theta(\lambda) = \int_b^{\bar{\lambda}} \lambda dF_\theta(\lambda).$$

The figure shows that the socially optimal retention threshold $b$ and the retention threshold adopted by firms $\bar{\lambda}$ are the same across types. Consequently, the interval from $b$ to $\lambda$ over which one needs to integrate to calculate the lost productivity from inefficient retention is the same for both types. What differs is the density of each distribution over this interval: type $\theta_1$ workers are *ex ante* more likely than type $\theta_2$ workers to fall into this interval where it would socially efficient to retain them, but optimizing firms will not do so.

As discussed above, if workers have no private information on their future realized health-related productivity $\lambda$ and no ability to engage in *ex post* wage bargaining after their health-related productivity is revealed, the optimal contract between risk-averse workers and risk neutral employers would set the wage $w$ independent of the realized health-related produc-
tivity of the worker.

Not allowing wages to depend on observed health types is a stronger and more consequential assumption. If employers were able to fully discriminate based on observed health types and offer different wages \( w(\theta) \) to different health types, they could potentially pass through the differential change in the cost of employment of different health types to the workers. We provide a more formal discussion below in Section 2.6, but the key is that increasing employer cost sharing \( \eta \) only has distributional consequences. Specifically, lower \( \theta \) workers see a wage increase and higher \( \theta \) workers see a wage decrease, but all types are hired by firms. If the social planner does not care about these distributional consequences it remains efficient to set \( \eta^* = 1 \).

In the case which is more often discussed in policy circles when experience rating considered, wages exhibit rigidity: some health characteristics that are informative about the future productivity of workers (summarized in \( \theta \)) are observable to employers, but employers are unable to condition wages on them. For example, there might be legal restrictions on using certain worker characteristics for wage setting. In this case, the incidence of increased employer cost sharing \( \eta \) will be on the extensive margin: some worker health types \( \theta \) are not going to be hired by employers and their productivity will be lost. This of course has distributional consequences too: instead of getting paid lower wages, ex ante higher-risk workers have a lower likelihood of being employed.

In the case of a risk neutral employer, a health type \( \theta \) worker is hired if

\[
\int_\lambda^1 \lambda dF_\theta(\lambda) - \left[F_\theta(\bar{\lambda}) \eta \rho w + (1 - F_\theta(\bar{\lambda})) w\right] > M. \tag{5}
\]

This implies a threshold \( \bar{\theta}(\eta) \) below which workers will not be hired. This threshold \( \bar{\theta}(\eta) \)
satisfies:

$$\int_{\tilde{\lambda}}^{1} \lambda dF_{\theta}(\lambda) - \left[ F_{\theta}(\tilde{\lambda}) \eta \rho w + \left( 1 - F_{\theta}(\tilde{\lambda}) \right) w \right] = M. \quad (6)$$

This is in contrast with the socially optimal hiring standard under which all workers should be hired (but not all workers should be retained after their productivity is realized). Given this observation, we can write the total lost productivity in the economy as

$$L(\eta) = \int_{0}^{\tilde{\theta}(\eta)} \int_{b}^{1} \lambda f_{\theta}(\lambda)g(\theta)d\lambda d\theta + \int_{\tilde{\theta}(\eta)}^{1} \int_{b}^{\tilde{\lambda}(\eta)} \lambda f_{\theta}(\lambda)g(\theta)d\lambda d\theta. \quad (7)$$

This is illustrated in Figure 3. Suppose that the two worker health types $\theta_1$ and $\theta_2$—with their \textit{ex ante} productivity distributions in blue and red, respectively—are on the margin of being hired. If the health type $\theta_1$ worker is not hired, her productivity $L_S$ is lost because of selection. $L_S$ is the integral of the density of the productivity distribution above the threshold set by employers $\tilde{\lambda}$: $L_S(\eta) = \int_{\tilde{\lambda}}^{1} \lambda dF_{\theta_1}(\lambda)$. If the health type $\theta_2$ worker is hired, her productivity $L_R$ is lost because the retention threshold $\tilde{\lambda}$ is higher than the socially optimal level $b$. $L_R$ is the integral of the density of the productivity distribution between the socially optimal employment threshold $b$ and the threshold set by employers $\tilde{\lambda}$: $L_R(\eta) = \int_{b}^{1} \lambda dF_{\theta_2}(\lambda) - \int_{\tilde{\lambda}}^{1} \lambda dF_{\theta_2}(\lambda) = \int_{b}^{\tilde{\lambda}} \lambda dF_{\theta_2}(\lambda)$. Notice that we only attribute lost productivity over the retention threshold $\tilde{\lambda}$ to selection, since the productivity between $b$ and $\tilde{\lambda}$ would be lost to inefficient retention even if the worker were hired.

The optimal cost sharing $\eta^*$ now needs to balance the increase in lost productivity through increased selection (increasing $\eta$ increases $\tilde{\theta}$, the threshold below which a worker will not be hired) and the decrease in lost productivity through increased retention (increasing $\eta$ lowers $\tilde{\lambda}$, the threshold below which a worker will not be employed after her productivity is realized). This trade-off is captured in the first order condition:
\[ \frac{d}{d\eta} L(\eta^*) = \]
\[ \frac{d}{d\eta} \bar{\theta}(\eta^*) \cdot g(\bar{\theta}(\eta^*)) \cdot \left( E_{\bar{\theta}(\eta^*)} \left[ \lambda \mid b \leq \lambda \right] - E_{\bar{\theta}(\eta^*)} \left[ \lambda \mid b \leq \lambda \leq \bar{\lambda}(\eta^*) \right] \right) \]
\[ + \frac{d}{d\eta} \bar{\lambda}(\eta^*) \cdot \int_{\theta(\eta^*)}^{1} f_{\lambda}(\bar{\lambda}(\eta^*)) g(\theta) d\theta \cdot \bar{\lambda}(\eta^*) \]
\[ = 0. \]

This condition can be interpreted to say that the optimal cost sharing level trades off the cost of selection and the benefit of retention.\textsuperscript{6} When \( \eta \) is set optimally, a marginal change in \( \eta \) leads to selection and retention responses that cancel each other out. When \( \eta \) is set below optimum, it can be increased, because the increase generates selection costs that are lower than the retention benefits. When is set \( \eta \) is set above the optimum, too many workers are selected out of employment relative to the benefit generated by workers retained.

The first part of the first order condition summarizes the cost of selection when increasing \( \eta \). This consists of the term describing the set of workers affected by the additional selection, \( \frac{d}{d\eta} \bar{\theta}(\eta^*) \cdot g(\bar{\theta}(\eta^*)) \) and the expected loss of productivity among these workers, \( E_{\bar{\theta}(\eta^*)} \left[ \lambda \mid b \leq \lambda \right] - E_{\bar{\theta}(\eta^*)} \left[ \lambda \mid b \leq \lambda \leq \bar{\lambda}(\eta^*) \right] \). Notice that the latter term takes into account that were these selected-out workers hired, their retention would still be inefficiently low as discussed in the previous subsection. This “cost of selection” term is increasing in \( \eta \): as employers bear a higher share of benefit costs, workers of higher \textit{ex ante} productivity become too costly to employ.

The second part of the first order condition summarizes the benefit of retention when increasing \( \eta \). This consists of the term describing the set of workers affected by the additional

\textsuperscript{6}See Appendix Section C.1 for the derivation of the first order condition.
retention, \( \frac{d}{d\eta} \bar{\lambda}(\eta^*) \cdot \int_{\bar{\theta}(\eta^*)}^{1} f_\theta(\bar{\lambda}(\eta^*)) g(\theta) d\theta \) and the lost productivity among these workers \( \bar{\lambda}(\eta^*) \). This “benefit of retention” term is decreasing with \( \eta \): as employers bear a higher share of benefit costs, they are retaining more workers of lower realized productivity.

As before, the wage must satisfy the budget constraint

\[
F(\bar{\theta}) \rho w + \int_{\bar{\theta}}^{1} F_\theta(\bar{\lambda}) \rho w dG(\theta) + \int_{\bar{\theta}}^{1} (1 - F_\theta(\bar{\lambda})) w dG(\theta) = \int_{\bar{\theta}}^{1} \int_{\lambda}^{1} \lambda dF_\theta(\lambda) dG(\theta) - M \tag{9}
\]

or

\[
w = \frac{\int_{\bar{\theta}}^{1} \int_{\lambda}^{1} \lambda dF_\theta(\lambda) dG(\theta) - M}{\rho \left( F(\bar{\theta}) + \int_{\bar{\theta}}^{1} F_\theta(\bar{\lambda}) dG(\theta) \right) + \int_{\bar{\theta}}^{1} (1 - F_\theta(\bar{\lambda})) dG(\theta)}. \tag{10}
\]

This means that the previous finding that welfare is maximized by setting \( \eta^* = 1 \) and making employers fully responsible for benefit costs no longer necessarily holds. It is welfare improving to increase \( \eta \) only if the productivity lost among the workers who are selected out of the market because they are not productive enough in expectation is less than the extra productivity realized by workers who are retained by firms after their productivity is realized.

### 2.6 Discussion

The results above relied on a number of assumptions. Here we briefly discuss these assumptions and the generalizability of the results.

**Ex Post Wage Bargaining** We assumed throughout that there is no *ex post* wage bargaining. This means that wages are not dependent on realized health-related productivity \( \lambda \). This is the optimal contract between a risk-averse worker and a risk-neutral employer. We can allow for *ex post* wage bargaining by allowing a worker with realized health-related...
productivity $\lambda$ to capture a portion $\gamma$ of the surplus she brings to the employer:

$$w(\lambda) = \bar{w} + \gamma(\lambda - \bar{\lambda}).$$

If workers who realize higher-than-average productivity $\lambda$ are able to bargain their wage $w(\lambda)$ up, to satisfy the resource constraint, the threshold $\bar{\lambda}$ has to be higher than in the benchmark case above. If ex post bargaining allows both lower and higher wages for lower-than-average and higher-than-average productivity workers, respectively, then $\bar{\lambda}$ will decrease, moving towards the optimal level $b$. In the extreme case where $\gamma = 1$ and all workers are paid exactly their realized productivity, firms retain all workers, paying lower wages to workers with decreased health-related productivity. This is of course unlikely to happen in practice (and is illegal in most contexts) and is also not optimal with risk-averse workers and risk-neutral firms.

**Risk-Averse Employers** We assumed that employers are risk-neutral. It is possible that employers act as if they were risk averse, perhaps because they are small and not fully diversified (Greenwald and Stiglitz, 1990, 1993). If they maximize a concave and separable function of productivity and cost, then the hiring threshold must solve

$$\int_{\bar{\lambda}}^{1} v(\lambda) dF_\theta(\lambda) - [F_\theta(\bar{\lambda}) c_1 (\eta \rho w) + (1 - F_\theta(\bar{\lambda})) c_2 (w)] = M. \quad (12)$$

This can also be understood as implying a different ranking of types where risk-averse employers don’t only take into account expected profits as above but the riskiness of profits. Then under this new ranking, a new, $\bar{\theta}$ can be found.

**Health Type-Specific Wages** We assumed that $\theta$ summarizes health-related characteristics of workers that are observable to employers but employers are unable to condition wages on $\theta$, possibly because of legal constraints. Consider a case where employers are allowed
and able to offer wages that at least partially depend on the productivity of different worker health types:

$$\hat{w}(\theta) = \alpha w(\theta) + (1 - \alpha) \bar{w} \tag{13}$$

where $\hat{w}(\theta)$ is the wage received by a worker of health type $\theta$, $w(\theta)$ is the wage she would receive if wages were allowed to be fully type-specific (i.e. workers would receive their expected productivity net of cost) and $\bar{w}$ is the average wage set based on average expected productivity.

In the section on selection above, we considered the case where $\alpha = 0$ and firms do not incorporate any information from $\theta$ in wage setting. The other extreme case is where $\alpha = 1$ and firms fully incorporate workers observed health type $\theta$ into wages. In this case the wage offered to a health type $\theta$ worker fully reflects the worker’s expected productivity and cost to the employer:

$$w(\theta) = \frac{\int_\lambda^1 \lambda dF_\theta(\lambda) - M}{\rho F_\theta(\lambda) + (1 - F_\theta(\lambda))}. \tag{14}$$

This implies that when $\eta$ changes, the differential change in the cost of hiring workers of different type $\theta$ is fully passed through to workers. This means that firms would not engage in selection behavior, since they could just set the wage for each type to make employing that type profitable. Therefore increasing employer cost sharing $\eta$ only has distributional consequences (lower $\theta$ workers see a wage increase, higher $\theta$ workers see a wage decrease) and if the social planner does not care about distributional consequences it remains efficient to set $\eta^* = 1$.

3 Background

3.1 Disability Insurance in the Netherlands

Disability insurance in the Netherlands is coordinated by the Dutch social security administration, the UWV. The 2004 Work and Income According to Work Capacity Act (WIA)
offers mandatory income insurance against health-related productivity loss to all working individuals, except those who are self-employed, and those who become ill while on unemployment insurance. Employers pay premiums that cover the costs of these benefits. Since 1998, premiums for permanent workers have been experience-rated (i.e., based on a firm’s own costs), while until the 2012-2013 reform we study they were not experience-rated for fixed-term and temporary workers. Instead, firms paid premiums, based on the average benefit receipt in their sector.

Sick workers can enter DI after a two-year (104-week) period. During this period, employers pay mandatory sick pay for sick permanent workers at a replacement rate of 70% of prior earnings. Fixed-term workers, whose employment contract ends before the two-year period is over, receive sick pay through a social insurance scheme called Ziektewet (Sickness Act). Temporary workers, the focus of this study, receive all their sick pay from Ziektewet. This scheme became experience-rated at the same, during the 2012-2013 reform. For the remainder of the paper we will use Ziektewet sickness payments interchangeably with short-term DI, as in effect these payments go to workers who are unable to work in the short term, but can then plausibly return to work. Figure 4 shows the timeline of sickness and DI benefits by worker type. Appendix B summarizes further details about the DI system and prior reforms.

3.2 The 2012 Experience Rating Reform

The high rate of DI benefit receipt among workers not included in the 1998 experience rating policy for permanent workers was a concern for policymakers and a reform passed in 2012 included these workers in experience rating as well. Prior to the introduction of experience rating, employers paid flat-rate DI premiums for the disability and sickness costs of their flex workers. These premiums were set at the industry level, and reflected the average disability risk in each of the 69 industries (e.g., agriculture, construction, transportation, temporary work agencies). After the introduction of experience rating, employers paid a premium that
was based on their own former workers’ sickness and disability costs relative to the industry average, up to an industry-specific maximum.

The premium formula used by the Dutch social security administration (Uitvoering Werknemersverzekeringen, 2014) can be summarized as follows. For each employer $j$, a employer-specific premium $\text{Premium}_j$ can be calculated as the combination of the following terms:

$$\text{Premium}_j = \text{Balance Correction} + \text{Correction Factor} \times (\text{Risk}_j - \text{Average Risk}). \quad (15)$$

In this equation, the Balance Correction ($\text{Rekenpercentage}$) and the Correction Factor ($\text{Correctiefactor Werkgeversrisico}$) terms ensure the overall balance of the social security fund. An employer’s Risk $\text{Risk}_j$ is calculated as the ratio of benefit costs of the former workers of the employer and wages:

$$\text{Risk}_j = \frac{\text{Benefit Costs}_j}{\text{Wage Bill}_j}. \quad (16)$$

These amounts are calculated over a certain lookback period. Initially (premium adjustment in 2014), the lookback period for DI benefits was benefits paid out in 2012 with first entitlement commencing in 2012 and the lookback period for the average wage bill was 2008-2012. Then the lookback period was widened for benefits to 2012-2013 in the 2015 premium adjustment, and subsequently widened up to 10 years. For Ziektewet (short-term DI) benefits, the lookback period in 2014 was 2012 and is subsequently always year $t - 2$.

### 3.3 Degree of Experience Rating

Premiums were not fully experience rated for all employers even after the reform for several reasons. First, small employers (employers with a wage bill below 10 times the average wage) continued to pay the average rate of their sector, while medium employers (employers

---

7The system introduced for flex workers during the 2012 reform follows closely the system that was in place for permanent workers, discussed in De Groot and Koning (2016).
with a wage bill between 10 times the average wage and 100 times the average wage) paid a
linear combination of the average rate in their own sector and their own cost (based on their
exact position on the scale between 10 and 100 times the average wage), and large employers
paid their own cost. We focus on large temporary work agencies that paid their own cost.
Second, the final premium was not allowed to below a minimum rate or above a maximum
rate determined at the level of the sector. The maximum rates were set by far the highest
for the temporary work agencies we study, at 3.28% for DI benefits and 7.77% for Ziektewet
(short-term DI) benefits. In all other sectors, the maximum was as low as 0.68% for DI
benefits and 1.24% for Ziektewet (short-term DI) benefits. Third, since the the lookback
period used to calculate firm-level risk is limited, firms are not eventually responsible for all
benefit costs of their former employees.

A key object in our model presented in Section 2 is the degree of experience rating \( \eta \). To
calculate the empirical \( \eta \) introduced by the reform for large temporary work agencies,
we make some simplifying assumptions. We assume that the sectoral maximum rates of
3.28% for DI benefits and 7.77% for Ziektewet (short-term DI) benefits are not binding.
This assumption would be violated if some firms have so many workers receive DI already,
that the they face a marginal cost of zero for additional workers. UWV aggregate statistics
suggest that this assumption holds, in particular because the maximum rates are set at
1.75-times and 4-times higher than the average rate for DI and Ziektewet (short-term DI),
respectively. We also assume that the lookback period is 10 years. The lookback period
starts from 1 year and subsequently widens to 10 years (the lookback period for permanent
workers), so a worker who enters DI in 2014 will eventually remain in the set of former
workers whose benefits a firm covers through premiums for 10 years, though initially new
entries change a firms calculated risk much more than when the “steady state” with a 10-year
lookback period is reached.

The key to estimating the degree of experience rating \( \eta \) is calculating the number of
years for which a worker actually contributes to overall DI costs (as opposed to the firm’s
experience rated premiums). If we assume that workers who enter DI are not going to be reemployed and remain on DI (which is the case for the vast majority of workers), then we can take the difference between the retirement age of 65 (when workers transition to pensions) and the average age of DI entry among temporary work agency workers (43.9) to calculate that the average beneficiary will spend 21.1 years on benefits and firms will be responsible for 10 years. Under these simplifying assumptions, we find that $\eta = 10/21.1 = 0.47$, or that temporary work agencies became responsible for 47% of their workers’ disability benefit costs.

### 3.4 Temporary Work Agency Workers in the Netherlands

Most of the Dutch labor market is characterized by a high degree of employment protection. However, substantial variation in employment protection exists, ranging from permanent workers who can only be fired with a substantial severance payment, to temporary work agency workers who have no employment protection during the first year and a half of working for a temporary work agency. Between 2003 and 2019, the share of permanent contracts has been declining from 77% to 65%, in favor of fixed-term contracts and alternative work arrangements. Temporary agencies make up a large part of employment in the Netherlands. The number of individual workers from whom temporary work was their primary income sources at a particular point in time rose from 185,000 in 2003 to 266,000 in 2019. Counting all workers who worked for a temporary work agency for at least one day during 2018, the total number of temporary work agency workers in 2017 was 857,000, out of about 8 million workers (de Wit, Vermeulen, Leest and van Druten, 2018).

A temp contract, offered by temporary work agencies, is the only employment contract in the Netherlands that is similar to at-will employment in the U.S. But a temp contract can only be offered during the first 78 weeks of employment. Continuation of the employment relationship between 78 weeks and four years of employment requires a fixed-term contract, which provides employment security similar to fixed-term workers, which means that the
temporary work agency pays the salary until the end of the fixed-term contract duration. The worker is obliged to accept so-called appropriate work during this period. Throughout this second phase, at most six contracts can be had. Continuation of employment beyond four years requires a permanent contract with the highest degree of employment protection. The proportion of workers in the at-will phase relative to all temporary work agency workers fell from 87% in 2007 to 78% in 2017 (de Wit, Vermeulen, Leest and van Druten, 2018).

An important limitation of our analysis is that we do not observe the “user companies” at which temporary workers perform their jobs. Instead, we observe the temporary work agency as their employer and the temporary work agency industry as their industry. The industry composition of the “user companies” has been stable between 2007 and 2017, with a quarter of workers performing jobs in the manufacturing and roughly ten percent of temporary workers in each of the following industries: construction; wholesale and retail trade; transportation and storage; administrative and support service activities; public administration and defense (de Wit, Vermeulen, Leest and van Druten, 2018).

4 Data and Sample

4.1 Administrative Datasets

We use universal register data from the Netherlands provided to us through a contract with Statistics Netherlands (CBS). We have access to several distinct datasets for the years 2009-2018, all of which cover everyone who lives in the Netherlands. We summarize the key datasets and data elements used here. Appendix D provides more information on our data.

**Personal Information** We use the Municipal Personal Records Database to establish a baseline sample of residents of the Netherlands. We derive a number of key characteristics of each person from this data, including month and year of birth, gender, municipality, and (parental) national origin.
**Employment Information**  We use the Jobs and Wages Administrative Database to document employment characteristics. From this database, we are able to link workers to employers, and observe monthly earnings, working hours, industry, and the type of work arrangement “contract type” which determines job security.

**Disability Insurance**  We use the Disability Benefit Database to identify individuals who take up disability benefits. From this database, we are able to derive information on the period of benefit eligibility, the disability program used, diagnosis, the disability percentage class, and benefit amount. We do not have access to information on disability applications.

**Short-Term Disability Insurance**  We use two social security datasets to find individuals who take up Ziektewet (short-term DI) benefits. This is only applicable to fixed-term and temporary workers, since sickness benefits for permanent workers are paid by employers. From these databases, we are able to derive information of the start and end dates of short-term disability insurance benefit claims.

**Healthcare Spending**  We use the Health Insurance Act Cost database to calculate annual curative healthcare spending for each individual. From this database, we are able to calculate curative healthcare spending overall and by category (inpatient hospital, mental health, prescription drugs, etc.). Long-term care costs, which are used to fund home care and homes for the elderly and disabled, are not included in this database.

**Education**  We use the Highest Education Level Tabulation to categorize individuals by their level of education. From this database, we are able to derive information on detailed educational attainment, which we aggregate into broad categories. A limitation of this data is that degree information is only complete for birth cohorts starting in 1983.
4.2 Sample Construction

We start by building a dataset that contains annual employment records for 2009 to 2018. Employment information is available at the worker-employer-month level. To build annual employment records, we take the following steps. We first calculate total annual wages and hours for each person-employer-contract type combination. Then for individuals employed in multiple jobs in a year, we keep the employment relationship with the highest earnings. Our baseline sample is for the period 2009-2016; we incorporate data up to 2018 in order to establish a two-year look-ahead period necessary for tracking disability insurance benefit receipt and employment outcomes.

We then match these employment records for the 2009-2016 period with records of short-term disability insurance benefit (Ziektewet) receipt and records of disability insurance benefit receipt for the 2009-2018 period. For a given person-year, we define short-term disability benefit receipt as a record starting that year. We define disability insurance benefit receipt as a disability insurance record starting two years later, to allow for the two-year waiting period. In addition, we merge in data that contains the personal information, healthcare spending measures, and educational attainment variables described in Section 4.1.

Table 1 shows summary statistics for permanent workers and temporary work agency workers. Overall, permanent workers are in a better labor market position. They have higher hourly wages (22.3 vs 13.0 euros on average). They are more attached to the labor market overall (96% stay employed the following year vs 87% among temporary work agency workers) and to their own employers in particular (84% stay employed at the same firm vs 39% among temporary work agency workers). Most temporary work agency workers are young, with 54% under 30 (15% among permanent workers). They are also less educated: 64% (vs 48% among permanent workers) have vocational qualifications, but only 27% (vs 45% among permanent workers) completed a tertiary degree. In our main event study, we control for these differences across the two groups of workers, as well as firm fixed effects in some specifications. We also carry out heterogeneity analyses where we estimate our event
study *within* important categories of workers (e.g., among low-wage workers).

## 5 Empirical Framework

To study the impact of experience rating, we leverage the 2012-2013 introduction of experience rating among temporary workers. As discussed in Section 3.2, the reform introduced by far the strongest financial incentives to large (wage bill above 100 times the average wage) firms in the temporary work agency sector. We compare the outcomes of workers of large temporary work agencies to permanent workers, for whom there was no policy change, in a difference-in-differences framework.

In the Netherlands, workers start receiving DI payments two years after the initial onset of sickness (see Section 3.1 for details). Therefore we examine DI benefit receipt outcomes of workers in an index year $t$ two years later in year $t + 2$. We report employment outcomes over both 1-year and 2-year horizons. We also examine what happens to temporary agency workers in the intervening two-year period, in particular, whether they receive sickness benefits through the government’s Ziektewet (short-term DI) scheme which became experience rated at the same time.

### 5.1 Regression Framework

**Disability Insurance Benefit Receipt and Employment**  
Our main estimating equation takes the form

$$y_{ijt} = \beta_0 + \sum_{t=2009}^{2016} \beta_{1t} T_{it} + X_{it} \beta + \rho_{rt} + \gamma_j + \epsilon_{ijt}$$  \hspace{1cm} (17)

where $i$ indexes individuals, $j$ indexes employers, and $t$ indexes time, $y_{ijt}$ is an outcome (e.g., DI benefit receipt, employment), $T_{it} \in \{0, 1\}$ is an indicator for whether worker $i$ is a temporary worker instead of a permanent worker in period $t$, $X_{it}$ is vector of controls (age, gender, education, healthcare spending percentile), $\rho_{rt}$ are region-by-year fixed effects, $\gamma_j$
are employer fixed effects. The coefficients of interest are $\beta_{1t}$, which describe the differential change between temporary and permanent workers for each year.

We also estimate a pooled version of this regression, where we replace the year-by-year comparison with a comparison of the pre-period and the post-period:

$$y_{ijt} = \beta_0 + \beta_1 T_{it} \times 1[t > 2012] + X_{it} \beta + \rho_{rt} + \gamma_j + \varepsilon_{ijt}$$ (18)

where $i$ indexes individuals, $j$ indexes employers, and $t$ indexes time, $y_{ijt}$ is an outcome (e.g., DI benefit receipt, employment), $T_{it} \in \{0, 1\}$ is an indicator for whether worker $i$ is a temporary worker, $1[t > 2012]$ is an indicator for our post-period, $X_{it}$ is vector of controls (age, gender, education, healthcare spending percentile), $\rho_{rt}$ are region-by-year fixed effects, $\gamma_j$ are employer fixed effects. The coefficients of interest is $\beta_1$, which describes the differential change between temporary and permanent workers in the post-period.

**Short-Term Disability Insurance Benefit Receipt** As discussed above, the Ziektewet (short-term DI) scheme which provides sickness benefits to temporary work agency workers (as well as to workers with fixed-term contracts that end during their sickness) became experience rated at the same as disability insurance. An important challenge for empirically evaluating the effect of experience rating on short-term DI benefit receipt is that permanent workers are not available as a control group because their sickness payments are provided directly by employers and we do not reliably observe this in the data. To provide suggestive evidence on short-term DI benefit receipt, we examine the evolution of short-term DI benefit receipt among temporary work agency workers, estimating the equation

$$y_{ijt} = \beta_0 + \sum_{t=2009}^{2016} \beta_{1t} + X_{it} \beta + \rho_r + \gamma_j + \varepsilon_{ijt}.$$ (19)

where $i$ indexes individuals, $j$ indexes employers, and $t$ indexes time, $y_{ijt}$ is an outcome (e.g., any receipt, days of receipt), $X_{it}$ is vector of controls (age, gender, education, healthcare spending percentile), $\rho_r$ are region-by-year fixed effects, $\gamma_j$ are employer fixed effects. The coefficients of interest is $\beta_1$, which describes the differential change between temporary and permanent workers in the post-period.
spending percentile), \( \rho_{rt} \) are region fixed effects, \( \gamma_j \) are employer fixed effects. The coefficients of interest are \( \beta_{1t} \), which describe the evolution of short-term DI benefit receipt, adjusting for observables. We also estimate a pooled version of this regression, where we replace the year-by-year coefficients with a comparison of the pre-period and the post-period:

\[
y_{ijt} = \beta_0 + \beta_1 1[t > 2012] + \mathbf{X}_it \beta_2 + \rho_r + \gamma_j + \epsilon_{ijt}
\]

(20)

where \( i \) indexes individuals, \( j \) indexes employers, and \( t \) indexes time, \( y_{ijt} \) is an outcome (e.g., any receipt, days of receipt), \( \mathbf{X}_it \) is vector of controls (age, gender, education, healthcare spending percentile), \( \rho_{rt} \) are region fixed effects, \( \gamma_j \) are employer fixed effects. The coefficient of interest is \( \beta_1 \), which describes the difference in short-term DI benefit receipt between the pre-period and the post-period, adjusting for observables.

### 5.2 Identification

In order for our estimates to represent the causal effect of the introduction of experience rating on DI benefit receipt and employment among temporary agency workers, it must be the case that there was no other change that affected these outcomes among temporary agency workers (the treatment group) but not among permanent workers (the control group).

The first challenge that must be considered is shifts in the composition of workers in the control and treatment firms. This is especially important in our context because our theoretical model predicts that treatment firms would react by adjusting the composition of their workers, selecting workers who are less likely to take up DI benefits. We address the issue of compositional shifts in three ways. First, we provide estimates including firm fixed effects. These estimates reflect within-firm changes in the outcomes, ruling out the possibility of changes in the composition of firms (e.g., through strategic bankruptcy). Second, we provide estimates that control for various observable worker characteristics (age, gender, education, healthcare spending percentile). To the extent that any compositional shifts
are picked up by these observable characteristics, these controls eliminate concerns about changing worker composition. Third, in Section 5.3, we discuss how we address selection directly using compositional shifts to predict changes in disability insurance benefit receipt and estimate what share of the overall change could be explained by changing composition.

A second challenge we need to consider is the role of other economic factors and trends that may affect the two types of workers differently. In our event studies, there do not appear to be strong differential trends in DI entry in the 4 years preceding the reform and we are also not aware of other reforms that would have affected DI benefit receipt differentially in the two groups. We control in some specifications for worker characteristics (age, education, health) that could reflect the impact of economic trends on worker composition, as well as firm fixed effects that take into account the impact of economic trends on the composition of firms. We also control in some specifications for region-year fixed effects, to take into account potential region-specific, time-varying shocks.

5.3 Estimation of Worker Selection

To understand changes in worker composition and worker selection, we focus on the previous year’s healthcare spending, a natural proxy for health status. We show that healthcare spending is highly predictive of DI benefit receipt in Figure A5. We then show how predicted disability insurance benefit receipt based on the distribution of prior healthcare spending changes among temporary work agency workers and permanent workers over time.

To quantify how much employers may have changed the composition of their workers to lower their expected disability costs, we use data on worker’s health status to predict their likelihood of DI benefit receipt. We then examine how much predicted DI benefit receipt changed among temporary workers employed by large temporary work agencies relative to permanent workers employed in other sectors of the economy. To predict DI benefit receipt,
we estimate the regression

\[ d_{ijt} = \alpha + \sum_{p=1}^{100} \gamma_p \delta_p + \varepsilon_{ijt} \]  

where \( i \) indexes individuals, \( j \) indexes employers, and \( t \) indexes time, and \( \delta_p \) are indicators for spending in the \( p \)th percentile of the healthcare spending distribution in the previous year. This prediction method allows us to account for the highly non-linear relationship between health status and DI benefit receipt. Using the predicted DI measure \( \hat{d}_{ijt} \) from, we estimate the difference-in-differences Equations (17) and (18). The coefficients estimating these equations indicate how much predicted DI benefit receipt changed, or in other words, how much DI would have declined absent a retention effect, purely because of shifting worker composition (selection effect).

An important caveat about our selection model is that by definition it can only pick up selection on predicted DI benefit receipt that is predicted by the variables (prior healthcare spending) included in the model. To the extent that there is selection on variables orthogonal to prior healthcare spending, we would be underestimating selection. To the extent that there are compositional changes that go in the opposite direction (i.e., increasing predicted DI benefit receipt among temporary work agency workers), we would be overestimating selection.

6 Results

6.1 Main Results

The model presented in Section 2 suggests that the introduction of experience rating would have two effects. The intended retention effect would decrease the productivity threshold above which firms choose to retain workers, corresponding to a decrease in \( \bar{\lambda} \) in our model. Empirically, we would expect the number of workers receiving DI to decrease and the number of workers remaining employed to increase. The unintended selection effect would increase
the expected productivity threshold above which firms hire workers, corresponding to an increase in \( \bar{\theta} \) in our model. Empirically, we would expect the composition of workers employed to shift towards healthier individuals who are \textit{ex ante} less likely to receive disability benefits. To examine the first mechanism, the intended retention effect, we start by reporting results on DI and Ziektewet (short-term DI) benefit receipt and employment outcomes. We then describe heterogeneity along various dimensions to understand who were the marginal workers affected by the policy, and examine the robustness of our estimates.

**Disability Insurance Benefit Receipt** Figure 5 shows difference-in-differences event-study estimates for DI benefit receipt. These results are from estimating Equation (17), controlling for age (in 1-year bins), gender, healthcare spending (in percentile bins), and region-by-year fixed effects, as well as firm fixed effects. For each person working in year \( t \), we measure DI benefit receipt starting in year \( t + 2 \) to account for the two-year sickness period that precedes DI benefit receipt. Standard errors are clustered at the firm level.

The figure suggests that there is no clear differential trend in DI benefit receipt between temporary work agency workers and permanent before 2012. Then in 2013, DI benefit receipt among temporary work agency workers dropped deferentially by 0.127 percentage points, or 25%. In the following years, the difference in DI benefit receipt appears roughly stable, suggesting that the two groups of workers followed similar trends. (Appendix Figure A1 shows raw trends for permanent and temporary workers.) To quantify the overall impact of the policy, comparing the pre-period (2009-2012) and the post-period (2013-2016), Table 2 shows pooled difference-in-differences estimates based on Equation (18) for DI benefit receipt. Column (1) suggests that pooling over the entire post-period, the decrease in DI benefit receipt among temporary workers relative to permanent workers is 0.120 percentage points (24%).

To examine the robustness of this result, we estimate several alternative specifications, showing the results in Table 2. Observing in Figure 5 and Appendix Figure A1 that there
was a large drop in DI benefit receipt by temporary work agency workers relative to permanent workers between 2009 and 2010, we show results excluding the year 2009 in column (2). If we exclude 2009 from the estimation, the drop is still large 0.102 percentage points (21%). In an alternative specification we drop firm fixed effects. These results, displayed in columns (5) and (6) are fairly similar, albeit the effect appears somewhat smaller. In another alternative specification, we re-estimate Equation (18) without controls for age, gender, healthcare spending, and region-by-year fixed effects (controlling instead for year). Columns (7), and (8) show these results, which are quite close to our main specification.

To put our results in context, we can compare our results to previous estimates of the impact of experience rating on DI benefit receipt in other work. We are aware of three existing quasi-experiment estimates: Hawkins and Simola (2020) use the introduction of experience rating for larger firms in Finland in 1996, Koning (2009) uses the introduction of experience rating for permanent workers in the Netherlands in 1998, and De Groot and Koning (2016) use the removal of experience rating for permanent workers in small firms in the Netherlands in 2003. All of these studies focus on permanent workers. Appendix Figure A4 shows the main estimates from these papers for DI benefit receipt (in blue) and our main estimate (in red). It suggests that our estimates are higher than existing estimates in the literature: Koning (2009) find a 15% reduction in disability insurance benefit receipt, De Groot and Koning (2016) find a 7% reduction, and Hawkins and Simola (2020) find a 13% reduction. Among other institutional differences, our finding of a larger reduction could be explained by larger and more immediate incentives or by our focus on temporary work agency workers.

Short-Term Disability Insurance Benefit Receipt  Ziektewet (short-term DI) benefits which provide income replacement to temporary workers before they are eligible for DI became experience-rated at the same time as DI benefits. An important limitation to our analysis of the receipt of these benefits is that permanent workers, who serve as a control
in our analyses of DI benefit receipt and employment are not insured under the Ziektewet scheme. Instead, they receive benefits from their employers when they are sick but not yet receiving DI, which we cannot reliably observe. Therefore they are not available as a control group in this analysis. We provide suggestive evidence on the evolution of short-term DI benefit receipt but caution that we do not have a natural control group.

Figure 6 shows how short-term DI benefit receipt evolved over this period among temporary workers. Panel (a) shows the extensive margin of benefit receipt, displaying the percent of temporary workers who spend any time on short-term DI benefits in each year. It suggests that there is an overall downward trend in the share of temporary work agency workers who receive benefits. As the experience rating policy is introduced, there is an 1.8 percentage point (9.2%) decrease in any benefit receipt. (Panel (a) of Appendix Figure A2 shows raw trends.) Column (1) of Panel (a) of Table 3 shows results pooling over the pre-period and post-period years. It suggests that the frequency of any Ziektwet receipt was 3.8 percentage points (19.6%) lower in the post-period than in the pre-period. We examine the robustness of this result under several alternative specifications, including versions excluding 2009, dropping controls, and dropping firm fixed effects. These results are reported in columns (2)-(8) of Panel (a) of Table 3. The results appear fairly similar under these alternative specifications.

Panel (b) of Figure 6 shows the average length of time spent on benefits (including the zeros) changes over time. Here it is apparent that a particularly large drop occurs in 2013 as experience rating is introduced. The average length of benefits decreases by 7.2 days (34%). (Panel (b) of Appendix Figure A2 shows raw trends.) Column (1) of Panel (b) of Table 3 shows results pooling over the pre-period and post-period years. It suggests that the average number of days spent on short-term DI benefits was 7.7 days (36.3%) lower in the post-period than in the pre-period. Columns (2)-(8) of this table report results excluding 2009, dropping controls, and dropping firm fixed effects. These results are again quite close to the main specification.
Panel (c) of Figure 6 isolates the intensive margin, showing the average number of days, conditional on any benefit receipt. This figure suggests that conditional on any benefit receipt, the length of time spent on short-term DI benefits decreases by 30 days (27.6%) in 2013, though it rebounds in the following years. (Panel (c) of Appendix Figure A2 shows raw trends.) Column (1) of Panel (b) of Table 3 shows results pooling over the pre-period and post-period years. It suggests that conditional on any short-term DI benefit receipt, the number of days spent on short-term DI benefits was 27.7 days (25.5%) lower in the post-period than in the pre-period. This result is somewhat sensitive to the set of controls and firm fixed effects included as displayed in columns (2)-(8) of the same table.

Finally, to better understand the full distribution of the length of time spent on benefits, Panel (d) of Appendix Figure A2 shows the distribution of benefit days, conditional on any benefits in 2012 and 2013. This figure shows that a large share of the decrease in average short-term DI benefit days between 2012 and 2013 was attributable to spells that last under 4 months. The probability that a short-term DI spell lasts at least one month decreases by 8%, the probability that it lasts at least two months decreases by 5%, the probability that it lasts at least three months decreases by 3%, and the probability that it lasts at least four months decreases by 3%. These shifts in the distribution of short-term DI spells suggest that the experience rating policy has lead to the quicker return to work of less sick workers, while having less impact on either being sick at all or on the sickest workers with longer short-term DI spells. This suggests that the policy may have succeeded in targeting workers with the most remaining working capacity.

Taken together, these results suggest that firms may take steps to encourage workers to return to work more quickly following the onset of an illness and therefore to prevent longer term-reliance on disability insurance benefits. This is consistent with conversations we have had with temporary work agencies. Our data makes it hard to directly observe what firms are doing, but there is some qualitative information available from public documents (e.g., Randstad, 2014) and our discussions with large temporary work agencies. For example,
at Randstad (the largest agency), they regularly audit whether workers who claim to be sick are actually sick. Once sickness is determined, their dedicated health and absenteeism department assesses workers’ working capacity and recommends actions for recovery. If illness persists, they check in again and discuss what tasks the worker would still be capable of performing and develop a re-integration plan together with the worker. This plan describes how the worker can get back to work and how the firm can help, e.g. by offering training or hiring a separate outside firm specialized in re-integration. The employee is obliged to draft, evaluate and modify the re-integration plan and is expected to do as was agreed. Once a worker returns to work, they follow up with her about the progress of re-integration. If the worker cannot perform their original tasks, they evaluate what work would be appropriate given the changes in her working capacity.

We think that these results, though suggestive in the absence of a natural control group, may be particularly important for policy when considering how to structure benefits and employer-side incentives. An important insight is that experience rating short-term DI benefits represents an early, immediate intervention that incentivizes firms to work to return workers to employment early on. This type of policy may be more effective than policies that only incentivize reduction in longer-term DI claims because firms may be less able to implement changes that only affect longer-term outcomes.

**Employment** We also examine the impact of experience rating on employment outcomes more broadly. Figure 7 shows estimates for whether a worker employed in a given year is employed in the following year. These results are again from estimating Equation (17), controlling for age (in 1-year bins), gender, healthcare spending (in percentile bins), and region-by-year fixed effects. The figure suggests that the employment prospects of temporary work agency workers relative to permanent workers are relatively stable if not declining prior to 2012, followed by a marked improvement starting in 2013. After the introduction of experience rating, there is a 1.7 percentage point (2%) increase in the probability of
remaining employed. (Appendix Figure A3 shows raw trends.) Over time, the differential increase among temporary workers in employment increases, Table 4 based on estimating Equation (18), suggests that pooling over the entire pre-period (2009-2012) and post-period (2013-2016), the increase in being employed two years later is 2.5 percentage points (2.9%).

To examine the robustness of this result, we estimate several alternative specifications, showing the results in Table 4. The increase in employment remains very similar when dropping firm fixed effects (columns (5) and (6)) or dropping controls (columns (7) and (8)) controls. We also report versions of each specification dropping 2009 (in columns (2), (4), (6), and (8)) and find similar effects. In an additional robustness check displayed in Appendix Table A1 we estimate employment effects over a two-year horizon, finding somewhat larger impacts.

One important point that can be observed comparing our results for DI benefit receipt and for employment is that in absolute terms the increase in employment (2.5 percentage points) is substantially larger than the decline in DI benefit receipt (0.127 percentage points). This finding lines up with the only existing study we are aware of that examines the impact of experience rating both on DI benefit receipt and employment (Hawkins and Simola, 2020), which finds an employment increase of 1.6 percentage points and a decline in DI benefit receipt of 0.079 percentage points. One possible mechanism could be that firms retain in the short term a wider set of at-risk workers than the subset of these workers who would have counterfactually received DI because firms cannot identify which workers among those at risk would end up on DI. Below in Section 6.2 we examine heterogeneity in employment outcomes by predicted DI benefit receipt.

The only other paper that, in addition to estimating impacts on DI benefit receipt, also reports estimates for employment is Hawkins and Simola (2020). In the overall sample of workers they find a decrease in employment, presumably due to pre-trends. Focusing on the sample of high-risk workers (top quartile of disability risk), they find a 1.9% increase in employment over a two-year time horizon, quite close to our 2.5% overall estimate and 3.2%
estimate for the top quartile (reported in Section 6.2). They also document an increase in employment that is substantially larger in absolute terms than the decrease in DI benefit receipt.

These results suggest that the intended retention effect (the increase in $\lambda$ in our model) of the policy was fairly large. On average, making firms responsible for 47% percent of DI costs lead to a 24% decline in DI benefit receipt and a 2.5% increase in employment. To the extent that this is a result of firms’ internalization of the costs of DI benefit receipt and a reduction in firm-side moral hazard, decreased DI benefit receipt and increased employment can improve efficiency since over the longer term they can result in lower tax rates, corresponding to less dead-weight loss and higher wages. Our model suggests that policymakers need to trade off this effect against the selection effect of the policy. We examine this selection effect below in Section 6.3 and present a welfare calculation comparing the retention and selection effects in Section 7.

### 6.2 Heterogeneity

To fully characterize the impact of the introduction of experience rating, it is useful to examine who the marginal workers moved by the policy were. This can help us understand whether the policy was well-targeted. We examine heterogeneous effects along several dimensions, including diagnosis type, gender, age, wage, and health status (predicted disability insurance benefit receipt based on prior healthcare spending). We report heterogeneity estimates for each of the three outcomes discussed above (disability insurance benefit receipt, short-term disability insurance benefit receipt, and employment).

**Disability Insurance Benefit Receipt** We start by examining heterogeneity by diagnosis type, focusing on the two most common broad diagnosis categories, musculoskeletal and mental health diagnoses. Columns (2) and (3) of Table 5 present pooled difference-in-differences estimates. These results suggest somewhat larger (though not statistically
significantly different) effects for mental health diagnoses. We find that disability insurance benefit receipt based on mental health diagnoses declined by 0.05 percentage points (22%), while disability insurance benefit receipt based on musculoskeletal diagnoses declined by 0.025 percentage points (20%). To our knowledge, our paper is the first to document heterogeneity in the effects of experience rating by type of diagnosis. One interpretation of our results is that mental health conditions may be slightly more amenable to accommodation and workers with these types of conditions represent a slightly higher return on investment by firms in efforts to retain them.\footnote{Prior work on the worker side, examining the labor supply effects of DI benefit receipt using variation in the allowance rates of examiners has found that individuals with mental impairments are the most likely to be on the margin of program entry (Maestas, Mullen and Strand, 2013). On the other hand, Kostøl and Mogstad (2014) find that return-to-work incentives have smaller effects among DI recipients suffering from mental illness.}

Next, we examine heterogeneity by gender in columns (4) and (5) of Table 5. We find slightly larger effects among women for whom disability insurance benefit receipt declines by 0.14 percentage points (24.9%), while among men the decline is 0.11 percentage points (23.2%). Breaking down our estimates by age, we re-estimate our main event study specification separately for four age groups (18-30, 31-40, 41-50, and 51-65). There is no effect in the youngest age group. In the 31-40 age group, there is a 0.20 percentage point (27.4%) decrease relative to the control group when experience rating is introduced. The decrease is 0.24 percentage points (24%) in the 41-50 age group. Finally, it’s 0.26 percentage points (22%) in the oldest age group. These results are consistent with the overall low incidence of disability among younger workers and with the view that when workers in the youngest group do receive DI, they are likely to be quite sick and therefore less amenable to accommodation. On the other hand, middle-aged workers may be the most amenable to accommodation efforts and have more working years left relative to the oldest age group.\footnote{Maestas, Mullen and Strand (2013) find that younger individuals are more likely to be on the margin of program entry and Kostøl and Mogstad (2014) find no effect of return-to-work incentives among older workers.}

We report results by quartile of wage in columns (11)-(14) of Table 5. We divide workers into four quartiles based on the distribution of wages among temporary workers to account for
the fact that these workers fall into the bottom of the overall wage distributions. We compare temporary workers who earn in each quartile of the temporary worker wage distribution to permanent workers with similar earnings. Effects are larger for higher quartiles: 10.6% (not statistically significant) in quartile 1, 12.9% in quartile 2, 24.2% in quartile 3, and 23.9% in quartile 4. This suggests that higher-earning workers may be easier or more worth to accommodate, likely because they are more productive, have more human capital, and are more attached to the labor market.

Finally, we also examine heterogeneity by health status and predicted (based on prior healthcare spending) DI benefit receipt. For this analysis, we drop 2009 to allow the use of one year of prior healthcare spending, since we only have healthcare spending from 2009. Columns (15)-(18) of Table 5 suggest that the relative drop is similar in each quartile: 27.1% in quartile 1, 22.5% in quartile 2, 21.6% in quartile 3, and 22.5% in quartile 4. Because the overall baseline is much higher in the top quartile than the rest of the distribution, the overall effect is driven by the effect in the top quartile.

Overall, we find that the impact of the introduction of experience rating on disability insurance is heterogeneous. We find that our results are somewhat larger for mental health diagnoses and among women, middle-aged workers, workers who are relatively high-earning within the distribution of the treated group. These results are suggestive that, in line with policymakers’ intentions, experience rating has the largest impact on workers who are the most likely to have substantial remaining working capacity.

**Short-Term Disability Insurance Benefit Receipt** We examine heterogeneity in impact of experience rating on short-term DI (Ziektewet) benefit receipt in Table 6. Panel (a) reports results for the extensive margin (any benefit receipt). Columns (2) and (3) suggest that the effect is somewhat larger for men (21.3%) than for women (17.6%). Columns (4)-(7) break down the effect by age group, showing changes decreasing with age: 21.3% among workers age 18-30, 19.4% among workers age 31-40, 17.6% among workers age 41-50, and
14.4% among workers age 51-65. Columns (9)-12) report results by quartile of hourly wage. We find a positive relationship between the size of the decline and wages: the decline is 10.3% in the bottom quartile, 16.9% in quartile 2, 22.3% in quartile 3, and 29.6% in the top quartile. Columns (13)-(16) report results by predicted DI benefit receipt based on healthcare spending. The relationship is negative, suggesting that the decrease is largest among workers who are least likely to \textit{ex ante} to receive benefits. We find a 23.1% decline in the bottom quartile of predicted DI benefit receipt, 20.8% in quartile 2, 18.1% in quartile 3, and 14.3% in the top quartile.

Panels (b) and (c) examine two further outcomes, the number of days spent on short-term DI benefits and the number of days spent on short-term DI benefits, conditional on any benefit receipt. Men (40.1%, 28.0% conditional on any) show slightly larger declines than women (32.5%, 22.7% conditional on any). The change is slightly larger in the middle age group: 37.8% (24.3% conditional on any) in the 18-30 age group, 35.3% (25.1% conditional on any) in the 31-40 age group, 36.1% (27.5% conditional on any) in the 41-50 age group, and 29.3% (23.3%) in the 51-65 age group. Intensive margin estimates are slightly larger in the middle wage quartiles: 33.7% (24.3% conditional on any) in the bottom quartile, 38.4% (26.7% conditional on any) in quartile 2, 38.5% (28.7% conditional on any) in quartile 3, and 32.9% (23.0% conditional on any) in the top quartile. Finally, we again find a negative relationship between predicted likelihood of DI benefit receipt and the decrease in days spent on short-term DI benefits: 41.9% (30.7% conditional on any) in the bottom quartile, 39.9% (27.1%) in quartile 2, 36.4% (25.3% conditional on any), and 29.7% (21.1% conditional on any) in the top quartile.

**Employment** We examine heterogeneity in the change in employment in Table 7. Columns (2) and (3) suggest that the employment increase is slightly larger among men (2.7 percentage points or 3.1%) than among women (2.3 percentage points or 2.6%). The change is again larger among older workers, as displayed in Columns (4)-(7): 2.1 percentage points (2.3%)
in the 18-30 age group, 3.5 percentage points (4.2%) in the 31-40 age group, 3.8 percentage points (4.6%) in the 41-50 age group, and 3.8 percentage points (4.7%) in the 51-65 age group.

Breaking down the effect by quartile of the wage distribution of temporary work agency workers in Columns (9)-(12), we find roughly similar increases in employment in each quartile: 3.3 percentage points (4.0%) in the bottom quartile, 2.5 percentage points (2.9%) in quartile 2, 2.8 percentage points (3.2%) in quartile 3, and 2.3 percentage points (2.6%) in the top quartile. Columns (13)-(16) show that the effect size is also similar across quartiles of the distribution of predicted DI benefit receipt: 2.4 percentage points (2.6%) in the bottom quartile, 1.6 percentage points (1.8%) in quartile 2, 1.9 percentage points (2.2%) in quartile 3, and 2.7 percentage points (3.2%) in the top quartile.

6.3 Worker Selection Estimates

We now turn to examining the second mechanism discussed in our model presented in Section 2, the selection of workers by firms. When experience rating policies have been proposed to incentivize firms to reduce the number of workers taking up DI benefits, a potential concern has been that firms would start to select workers who are less likely to enter DI in the future, potentially even hurting the employment chances of these workers relative to the status quo. Our model formalizes this intuition, suggesting that the optimal policy would trade-off these selection incentives against the retention incentives of experience rating. It predicts that when experience rating is strengthened, the composition of workers would shift towards workers who are \textit{ex ante} less likely to receive disability benefits, corresponding to an increase in the selection threshold parameter $\bar{\theta}$.

We start by estimating a predictive relationship between health status (proxied by prior-year healthcare spending) and disability insurance benefit receipt, based on Equation (21). The underlying idea is that if employers select workers who are predictably more likely to receive disability insurance and who are under the experience rating system more costly
to employ in expectation, then we will observe a decline not only in disability insurance benefit receipt but also in predicted disability insurance benefit receipt. Appendix Figure A5 shows the relationship between log healthcare spending in the prior year and the likelihood of disability insurance benefit receipt, separately for temporary work agency workers and permanent workers. The relationship is non-linear, with most disability insurance benefit receipt concentrated at the top end of the distribution. For example, among temporary work agency workers, 1.67% receive disability insurance at the 99th percentile of the log healthcare spending distribution, 1.18% at the 90th percentile, 0.75% at the 75th percentile, and 0.48% at the median.

In Figure 8, we report difference-in-differences event study estimates for predicted disability insurance benefit receipt (blue triangles), putting it on the same scale as actual disability insurance benefit receipt (red circles). To allow us to use the same scale and account for the compositional shift in prior healthcare spending and consequently predicted disability insurance benefit receipt that happens in 2012 (rather than in 2013 when the overall decline in disability insurance benefit receipt happens), on this figure we omit the year 2010 rather than 2012. Because healthcare spending data is only available from 2009, the first year for which we can use prior-year healthcare spending to predict disability insurance benefit receipt is 2010. We find that predicted disability insurance benefit receipt drops by 0.01 percentage points (3%) in 2012, while actual disability benefit insurance receipt first drops in 2013, by 0.1 percentage points (25%).

Table 8 reports estimates pooling the pre-period and post-period. It suggests that predicted disability insurance benefit receipt is just under 0.01 percentage points (3%) lower in the post-period relative to the pre-period. At the same time, actual disability insurance benefit receipt by temporary agency workers declined by just under 0.1 percentage points (20.6%). Columns (3)-(4) show the same results dropping firm fixed effects. To further characterize the shift in the composition of workers by predicted disability insurance benefit receipt, Appendix Figure A6 shows the distribution of predicted disability insurance benefit
receipt among temporary agency workers in 2011 and 2012. The likelihood that a temporary work agency worker was in the top quartile of the distribution of predicted DI benefit receipt based on the previous year’s healthcare spending decreased by 4.2%. The likelihood that they were in the top decile, decreased by 3%.

These results suggest that the effect of the introduction of experience rating on selection (as measured by changes in predicted disability insurance benefit receipt based on prior-year healthcare spending) are relatively small. Using our preferred estimates, selection can account for 14% of the overall decline in disability insurance benefit receipt. The only other paper that considers potential selection responses to experience rating disability insurance is Hawkins and Simola (2020). They find that at most 7% of the overall drop in disability insurance benefit receipt could come from worker selection.

Our results on selection suggest that shifts in the composition of workers employed by temporary work agencies can account for a relatively small share of the overall decrease in disability insurance benefit receipt. However, this share is still substantial and policymakers should take it into account when designing policy. First, selection responses may take longer to manifest as firms are able to adjust their hiring practices more slowly in some contexts. Second, a social planner may place particular weight on the types of workers who are selected out of the labor market as a consequence of the introduction of experience rating.

7 Welfare Estimation

7.1 Welfare Impact of the Reform

In Section 2 we concluded that the optimal degree of experience rating $\eta$ depends on two key quantities: the cost of selection and the benefit of retention. The intuition is that as $\eta$ increases, more workers, and in particular \textit{ex ante} healthier workers are “selected out”, i.e., not hired by firms, but at the same time more workers, and in particular less healthy workers are retained by firms. The optimal $\eta$ is reached when the marginal impact of changing $\eta$
is zero, because the two impacts are balanced. This first order condition is summarized in Equation (8) and repeated here:

\[
\frac{d}{d\eta} L(\eta^*) = \\
\frac{d}{d\eta} \bar{\theta}(\eta^*) \cdot \frac{g(\bar{\theta}(\eta^*))}{\text{Density Affected}} \cdot \left( E_{\bar{\theta}(\eta^*)} [\lambda | b \leq \lambda] - E_{\bar{\theta}(\eta^*)} [\lambda | b \leq \bar{\lambda}(\eta^*)] \right) \\
\frac{\text{Expected Lost Productivity at } \bar{\theta}}{\text{Cost of Selection}} \\
+ \frac{d}{d\eta} \bar{\lambda}(\eta^*) \cdot \int_{\bar{\theta}(\eta^*)}^{1} f_{\theta}(\bar{\lambda}(\eta^*))g(\theta)d\theta \cdot \bar{\lambda}(\eta^*) \\
\frac{\text{Density Affected}}{\text{Benefit of Retention}} \cdot \bar{\lambda}(\eta^*) \\
= 0.
\]

Rearranging the first order condition as a ratio of the cost of selection and the benefit of retention, we have that at the optimum,

\[
\frac{d}{d\eta} \bar{\theta}(\eta^*) \cdot \frac{g(\bar{\theta}(\eta^*))}{\text{Density Affected}} \cdot \left( E_{\bar{\theta}(\eta^*)} [\lambda | b \leq \lambda] - E_{\bar{\theta}(\eta^*)} [\lambda | b \leq \bar{\lambda}(\eta^*)] \right) \\
\frac{\text{Expected Lost Productivity at } \bar{\theta}}{\text{Cost of Selection}} \\
= -1.
\]

That is, at the optimum the two quantities must equal each other and we can rely on the ratios of their components, as we do below.

Based on the logic of the model, and with some additional assumptions, we can ask whether the experience rating policy as introduced increased welfare. Further, we can also use the elasticities estimated from the introduction of the policy to calculate the optimal \(\eta\). We caution that these calculations involve a number of assumptions, but nevertheless we provide this exercise to show how estimates from various studies could inform policy.
We start by assuming that the estimated 3% decline in predicted DI benefit receipt is proportional to the change in selection threshold and the workers affected:

\[
\frac{d}{d\eta} \bar{\theta}(\eta^*) \cdot \underbrace{g(\bar{\theta}(\eta^*))}_{\text{Density Affected}}.
\]

We also assume that the estimated 17% additional decline in actual DI benefit receipt (on top of the predicted decline) is proportional to the change in the retention threshold and the workers affected:

\[
\frac{d}{d\eta} \bar{\lambda}(\eta^*) \cdot \int_{\bar{\theta}(\eta^*)}^{1} f_\theta(\bar{\lambda}(\eta^*)) g(\theta) d\theta.
\]

The remaining parameter is the ratio of the expected lost productivity of the marginal worker “selected out” \(E_{\bar{\theta}(\eta^*)} [\lambda \mid b \leq \lambda] - E_{\bar{\theta}(\eta^*)} [\lambda \mid b \leq \lambda \leq \bar{\lambda}(\eta^*)]\) and the productivity of the marginal worker retained \(\bar{\lambda}(\eta^*)\). One way to estimate these quantities is taking the ratio of the wages (a proxy for productivity) of those affected by decreased DI benefit receipt and of those affected by decreased employment in temporary work agencies.

Table 5 suggests that the largest decline in disability insurance benefit receipt in relative terms happens in the bottom quartile of predicted disability insurance benefit receipt, though because overall most disability insurance benefit receipt happens in the top quartile, the largest absolute effect is observed there. The largest negative shift in the probability that a worker is employed by a temporary work agency happens in the third quartile, with a smaller drop in the top quartile. Importantly, the average hourly wage of temporary agency workers is almost the same in every quartile: 13.2 euros in the bottom quartile, 12.9 euros in quartile 2, 13.0 euros in quartile 3, and 13.3 euros in the top quartile. This would suggest that the ratio of the productivity of the workers positively impacted by the retention change and the workers negatively impacted by the selection change is 1.

An alternative estimation strategy may be to rely directly on the wage distribution, and
estimate changes in retention in selection across parts of this distribution. Table 5 suggests that the decline in disability insurance benefit receipt is concentrated in the top two quartiles of the distribution of wages among temporary agency workers. If we weight the average hourly wage in each quartile (8.77 euros, 11.1 euros, 13.1 euros, and 18.4 euros) by the overall decline in each quartile, the weighted average hourly wage among those affected by decreased DI benefit receipt is 14.95 euros. At the same time, the likelihood of employment falls in the bottom quartile of the temporary work agency wage distribution and stays the same or increased in the other quartiles. This would suggest that the average wage of those selected out is 8.77 euros.

Therefore the ratio of the above components is

\[
\frac{\frac{d}{d\eta} \bar{\theta}(\eta^*) \cdot g(\bar{\theta}(\eta^*))}{\text{Expected Lost Productivity at } \bar{\theta}} \cdot \frac{\frac{d}{d\eta} \bar{\lambda}(\eta^*) \cdot \int_{\bar{\theta}(\eta^*)}^{1} f_\theta(\lambda(\eta^*)) g(\theta) d\theta}{\text{Productivity}} = \frac{0.03}{0.17} \cdot \frac{13}{13} = 0.176
\]

using the predicted disability insurance benefit receipt method to calculate the productivity ratio and

\[
\frac{\frac{d}{d\eta} \bar{\theta}(\eta^*) \cdot g(\bar{\theta}(\eta^*))}{\text{Expected Lost Productivity at } \bar{\theta}} \cdot \frac{\frac{d}{d\eta} \bar{\lambda}(\eta^*) \cdot \int_{\bar{\theta}(\eta^*)}^{1} f_\theta(\lambda(\eta^*)) g(\theta) d\theta}{\text{Productivity}} = \frac{0.03}{0.17} \cdot \frac{8.77}{14.95} = 0.104
\]

using the direct wage method to calculate the productivity ratio. Since this ratio is under 1, our model suggests that the policy introduced increased welfare but that the degree of experience rating was smaller than under the optimal policy.
7.2 Optimal Experience Rating

To calculate the optimal degree of experience rating $\eta^*$ from our empirical estimates, we need to make a further assumption. The policy introduced moved $\eta$ from zero to 0.47 (as calculated in Section 3.3). If we assume that there is a constant ratio between the change in $\eta$ and the selection-to-retention impact calculated above, we can extrapolate from the observed impact to conclude that the selection-to-retention ratio would remain under 1 even with $\eta = 1$, suggesting that full experience rating could be optimal, when abstracting from the potential excess risk faced by employers.

This calculation relies on strong assumptions. In particular, we assume that the selection and retention responses would scale linearly with changes in the degree of experience rating. This may not be the case in practice, as firms may not be able to easily accommodate workers as they go further down in the realized productivity distribution, i.e., as they consider the retention of workers with more and more serious health problems. A particular challenge that our model has abstracted from is that workers preferences for work are likely different when they have more serious health conditions. We would reasonably expect workers in worse health to have higher disutility from work, suggesting that as we go lower in the realized productivity distribution, workers would be more likely choose to apply for disability insurance (and would have a higher chance of being approved) rather than remaining employed. To fully incorporate both worker preferences and firm incentives, one would need a more complex model that could be estimated structurally. Recent work by Aizawa, Kim and Rhee (2020) makes progress in this direction.

Our model can also be applied to findings from other papers. Hawkins and Simola (2020) report a 13% decrease in disability insurance benefit receipt, an a 1.9% increase in employment, following the introduction of experience rating that moved the rate from 0% to 13% on average. They argue that at most 7% of the effect could be attributed to changes in selection, with 93% attributable to increased retention. Assuming a small productivity wedge between workers affected by the retention and selection responses in their context,
the Finnish reform was also welfare improving and close to full experience rating could also be optimal in their context.

8 Conclusion

Observing the increasing enrollment and costs of disability insurance programs, several proposals have suggested that firms bear more responsibility for the DI benefit costs of their former workers in order to incentivize them to decrease the flow of workers to DI. This paper presents a model of DI that accounts for both the intended consequence of this reform (reduced flow of workers to DI through increased retention) and its unintended consequence (selection against workers who are more likely to receive benefits in the future). The model allows us to describe the trade-off between these two mechanisms and derive a formula for the optimal experience rating policy.

Using the introduction of an experience rating policy among temporary workers in the Netherlands, we then estimated the size of these consequences empirically. We found that when temporary work agencies had to pay a sizable penalty in the form of increased premiums, the share of their workers who took up benefits decreased substantially. We documented a 24% decrease in DI benefit receipt, as well as a substantial fall in short-term DI use and an increase in employment. The policy also affected the workers who likely have the highest remaining working capacity, suggesting that it was well-targeted. Anecdotal evidence from conversations with temporary work agencies also suggests, that the reform may have succeeded in improving the employment prospects of potentially disabled workers in this sector.

The reform also lead to a compositional shift in their worker pool, consistent with selection, though selection could only account for a small share, 14% of the overall reduction in benefit receipt. This suggests that the selection effect was relatively small but meaningful, implying that policymakers do need to consider the potential selection effects when designing
programs. Risk adjustment or employment subsidies may be the policy tools to use to ensure that no workers are hurt when experience rating is introduced, though we cannot directly speak to the potential effects of these tools.

Connecting these estimates back to our model, we argued that given that the retention effect was substantially larger than the selection effect, the reform improved welfare. Moreover, extrapolating from our results, we found that in this context full experience rating may be optimal. Further work is needed to understand whether these results generalize to other contexts: employment protections in the Netherlands are very different from employment protections in the U.S. with potentially important implications for the design of employer incentives in disability insurance. The temporary work agency workers examined in this paper generally have weak employment protections, close to the U.S. system of at-will employment. But they also have weaker labor market attachment than the average U.S. worker. Further, while it appears that employer experience rating does well in the temporary worker context, whether selection responses would be weaker or stronger in the general labor market is not ex ante clear. Generally high turnover in the temporary work agency context means that it may be easier to select workers, but in contexts where longer-term employment contracts are used, the incentive to select workers is stronger.

Our work has important implications for labor market and social insurance policy in a world where contingent labor is becoming more common. First, our model and evidence suggest that a system of employer responsibility in the disability insurance system may be a good way to increase the efficiency, equity, and fiscal sustainability of this crucial social insurance program. Second, the evidence presented here suggests that it is possible to provide social insurance benefits to weakly attached contingent workers. In fact, employer responsibility may make it possible to efficiently provide such benefits, since employers need to internalize the social costs of contingent work arrangements.
References


Randstad. 2014. “Verzuimreglement (Absenteeism Regulations).” [https://www.randstad.nl/binaries/content/assets/randstadnl/werknemers/documenten/verzuimreglement-health-at-work](https://www.randstad.nl/binaries/content/assets/randstadnl/werknemers/documenten/verzuimreglement-health-at-work).


Figure 1: Benchmark Case With No *Ex Ante* Worker Heterogeneity

Notes: Figure illustrates the distribution of realized health-related productivity and the total productivity lost due to employer moral hazard in the benchmark case with no *ex ante* worker heterogeneity. All workers’ realized health-related productivity $\lambda$ is drawn from the distribution $F(\lambda)$ (density $f(\lambda)$). Employers set their retention threshold at productivity level $\bar{\lambda} = (1 + \tau - \eta \rho)w$, higher than the socially optimal level $b$. The expected total lost productivity is the integral of the density of the productivity distribution between the socially optimal employment threshold $b$ and the threshold set by employers $\bar{\lambda}$: $L(\eta) = \int_b^1 \lambda dF(\lambda) - \int_0^1 \lambda dF(\lambda) = \int_b^{\bar{\lambda}} \lambda dF(\lambda)$. For more details, see Section 2.4.
Figure 2: Worker Heterogeneity and Selection

Notes: Figure illustrates the distribution of realized health-related productivity and the total productivity lost due to employer moral hazard in a case where workers are *ex ante* heterogeneous. In this example, there are two worker health types $\theta_1$ and $\theta_2$ with their *ex ante* productivity distributions in blue and red, respectively. Their realized health-related productivity is drawn from distributions $F_{\theta_1}(\lambda)$ (density $f_{\theta_1}(\lambda)$) and $F_{\theta_2}(\lambda)$ (density $f_{\theta_2}(\lambda)$), respectively. Employers set their retention threshold at productivity level $\bar{\lambda}$ for each worker type, higher than the socially optimal level $b$. For each type $\theta$, the expected total lost productivity is the integral of the density of the productivity distribution between the socially optimal employment threshold $b$ and the threshold set by employers $\bar{\lambda}$: $L(\eta) = \int_b^1 \lambda dF_\theta(\lambda) - \int_{\bar{\lambda}}^1 \lambda dF_\theta(\lambda) = \int_b^{\bar{\lambda}} \lambda dF_\theta(\lambda)$. For more details, see Section 2.5.
Figure 3: Worker Heterogeneity and Selection

Notes: Figure illustrates worker heterogeneity and the potential for selection in a case where workers are heterogeneous. In this example, there are two worker health types $\theta_1$ and $\theta_2$ with their \textit{ex ante} productivity distributions in blue and red, respectively. Their realized health-related productivity is drawn from distributions $F_{\theta_1}(\lambda)$ (density $f_{\theta_1}(\lambda)$) and $F_{\theta_2}(\lambda)$ (density $f_{\theta_2}(\lambda)$), respectively. Employers set their retention threshold at productivity level $\bar{\lambda}$ for each worker type, higher than the socially optimal level $b$. If the health type $\theta_1$ worker is not hired, her productivity $L_S$ is lost because of selection. $L_S$ is the integral of the density of the productivity distribution above the threshold set by employers $\bar{\lambda}$: $L_S(\eta) = \int_{\bar{\lambda}}^{\eta} \lambda dF_{\theta_1}(\lambda)$.

If the health type $\theta_2$ worker is hired, her productivity $L_R$ is lost because the retention threshold $\bar{\lambda}$ is higher than the socially optimal level $b$. $L_R$ is the integral of the density of the productivity distribution between the socially optimal employment threshold $b$ and the threshold set by employers $\bar{\lambda}$: $L_R(\eta) = \int_b^{\bar{\lambda}} \lambda dF_{\theta_2}(\lambda) - \int_{\bar{\lambda}}^{b} \lambda dF_{\theta_2}(\lambda) = \int_{\bar{\lambda}}^{\bar{\lambda}} \lambda dF_{\theta_2}(\lambda)$. For more details, see Section 2.5.
Notes: Figure shows how individuals who become disabled while employed, or while on unemployment insurance, are eligible for DI benefits. DI benefits typically start two years after the worker calls in sick. Employers replace at least 70% of prior earnings for the duration of the employment contract, for a maximum of two years. Fixed-term workers receive sick pay through the Ziektewet (Sickness Insurance Act) scheme (in effect short-term DI) if their fixed-term employment contract ends within two years after the onset of disability. Temporary workers immediately receive sick pay through the Ziektewet (short-term DI) program when they call in sick because their employment contracts immediately end. Prior to the 2012-2013 reform, experience rating existed for DI premiums for permanent workers only. The 2012-2013 reform introduced experience rating for short-term DI and DI premiums for fixed-term and temporary workers.
Figure 5: Disability Insurance Benefit Receipt

Notes: Figure shows difference-in-differences estimates of the impact of the introduction of experience rating on disability insurance benefit receipt by temporary agency workers, relative to permanent workers. For each person working in year $t$, we measure disability insurance benefit receipt starting year $t + 2$ to account for the two-year sickness period that precedes disability insurance benefit receipt. The coefficients are from estimating Equation (17), controlling for age (in 1-year bins), gender, healthcare spending (in percentile bins), and region-by-year fixed effects, as well as firm fixed effects. Standard errors are clustered at the firm level. For more details, see Section 5.1. ($N = 42,516,521$ worker-years)
Figure 6: Short-Term Disability Insurance Benefit Receipt

(a) Any Short-Term DI Benefit Receipt

(b) Average Total Short-Term DI Days

(c) Average Total Short-Term DI Days, Conditional on Any

Notes: Figure shows estimates of the impact of the introduction of experience rating on short-term DI (Ziektewet) receipt by temporary agency workers. For each person working in year $t$, we measure short-term DI benefit receipt starting in the same year. The coefficients are from estimating Equation (19), controlling for age (in 1-year bins), gender, healthcare spending (in percentile bins), and region fixed effects, as well as firm fixed effects. Standard errors are clustered at the firm level. Panel (a) shows the share with any short-term DI benefit receipt start in year $t$. Panel (b) shows the average number of short-term DI benefit receipt days for sickness spells starting in year $t$. Panel (c) shows the average number of short-term DI benefit receipt days for sickness spells starting in year $t$, conditional on any. For more details, see Section 5.1. ($N = 2,020,368$ worker-years in Panel (a). $N = 2,020,368$ worker-years in Panel (b). $N = 360,551$ worker-years in Panel (c).)
Notes: Figure shows difference-in-differences estimates of the impact of the introduction of experience rating on the employment of temporary agency workers, relative to permanent workers. For each person working in year $t$, we measure employment in year $t + 1$. The coefficients are from estimating Equation (17), controlling for age (in 1-year bins), gender, healthcare spending (in percentile bins), and region-by-year fixed effects, as well as firm fixed effects. Standard errors are clustered at the firm level. For more details, see Section 5.1. ($N = 42,516,521$ worker-years)
Notes: Figure shows difference-in-differences estimates of the impact of the introduction of experience rating on disability insurance benefit receipt by temporary agency workers. For each person working in year \( t \), we measure disability insurance benefit receipt starting year \( t + 2 \) to account for the two-year sickness period that precedes disability insurance benefit receipt. The blue triangles show “predicted” disability insurance benefit receipt, based on the prediction described in Section 5.3. The red circles show actual disability insurance benefit receipt. These coefficients are from estimating Equation 17. Standard errors are clustered at the employer level. For more details, see Section 5.1. (N=36,998,766 worker-years)
Table 1: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Permanent Workers</th>
<th>Temporary Work Agency Workers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Hourly Wage (euros)</td>
<td>22.3</td>
<td>13.0</td>
</tr>
<tr>
<td>Female (%)</td>
<td>46</td>
<td>54</td>
</tr>
<tr>
<td>Age 18-30 (%)</td>
<td>15.2</td>
<td>54.3</td>
</tr>
<tr>
<td>Age 31-40 (%)</td>
<td>23.3</td>
<td>18.7</td>
</tr>
<tr>
<td>Age 41-50 (%)</td>
<td>29.2</td>
<td>15.5</td>
</tr>
<tr>
<td>Age 51-65 (%)</td>
<td>32.4</td>
<td>11.4</td>
</tr>
<tr>
<td>Mean Healthcare Spending (euros)</td>
<td>1482</td>
<td>1224</td>
</tr>
<tr>
<td>Vocational Education (%)</td>
<td>48.1</td>
<td>63.7</td>
</tr>
<tr>
<td>Secondary Education (%)</td>
<td>6.6</td>
<td>9.0</td>
</tr>
<tr>
<td>Tertiary Education (%)</td>
<td>45.3</td>
<td>27.3</td>
</tr>
<tr>
<td>Disability Insurance Benefit Receipt (%)</td>
<td>0.36</td>
<td>0.47</td>
</tr>
<tr>
<td>Employed Next Year (%)</td>
<td>96.3</td>
<td>87.3</td>
</tr>
<tr>
<td>At Same Firm Next Year (%)</td>
<td>83.9</td>
<td>38.9</td>
</tr>
<tr>
<td>Worker-years</td>
<td>40,496,153</td>
<td>2,020,368</td>
</tr>
</tbody>
</table>

Notes: Table shows summary statistics for permanent workers and temporary work agency workers. Data used covers the 2009-2016 period. For more details, see Section 4.2.
Table 2: Disability Insurance Benefit Receipt

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post</td>
<td>0.0297</td>
<td></td>
<td>0.0645*</td>
<td>0.0307***</td>
<td>0.0344***</td>
<td>-0.0244</td>
<td>0.0292</td>
<td>0.0675***</td>
</tr>
<tr>
<td></td>
<td>(0.0425)</td>
<td></td>
<td>(0.0366)</td>
<td>(0.0034)</td>
<td>(0.0035)</td>
<td>(0.0396)</td>
<td>(0.0351)</td>
<td>(0.0024)</td>
</tr>
<tr>
<td>Treatment</td>
<td>0.02819***</td>
<td>0.2571***</td>
<td>0.1629***</td>
<td>0.1468***</td>
<td>0.3459***</td>
<td>0.3246***</td>
<td>0.01664***</td>
<td>0.1411***</td>
</tr>
<tr>
<td></td>
<td>(0.0345)</td>
<td></td>
<td>(0.0338)</td>
<td>(0.0184)</td>
<td>(0.0179)</td>
<td>(0.0182)</td>
<td>(0.0171)</td>
<td>(0.0394)</td>
</tr>
<tr>
<td>Post × Treatment</td>
<td>-0.1203***</td>
<td>-0.1022***</td>
<td>-0.0835***</td>
<td>-0.0624***</td>
<td>-0.098***</td>
<td>-0.0766***</td>
<td>-0.1182***</td>
<td>-0.0996***</td>
</tr>
<tr>
<td></td>
<td>(0.0113)</td>
<td></td>
<td>(0.0099)</td>
<td>(0.0169)</td>
<td>(0.0163)</td>
<td>(0.0154)</td>
<td>(0.0143)</td>
<td>(0.0106)</td>
</tr>
<tr>
<td>Pre-2013 Mean</td>
<td>0.508</td>
<td>0.483</td>
<td>0.508</td>
<td>0.483</td>
<td>0.508</td>
<td>0.483</td>
<td>0.508</td>
<td>0.483</td>
</tr>
<tr>
<td>Observations</td>
<td>41,323,389</td>
<td>36,056,213</td>
<td>42,516,521</td>
<td>36,998,766</td>
<td>41,323,389</td>
<td>36,056,213</td>
<td>42,516,521</td>
<td>36,998,766</td>
</tr>
<tr>
<td>Controls</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm Fixed Effects</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exclude 2009</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Table shows difference-in-differences estimates of the impact of the introduction of experience rating on disability insurance benefit receipt by temporary agency workers, relative to permanent workers. For each person working in year $t$, we measure disability insurance benefit receipt starting year $t+2$ to account for the two-year sickness period that precedes disability insurance benefit receipt. These coefficients are from estimating Equation (18). In columns (1), (2), (5), and (6), we control for age (in 1-year bins), gender, healthcare spending (in percentile bins), and region-by-year fixed effects. In columns (1), (2), (7), and (8), we control for firm fixed effects. In columns (2), (4), (6), and (8) we exclude data for 2009. Standard errors are clustered at the firm level. For more details, see Section 5.1.
Table 3: Short-Term Disability Insurance Benefit Receipt

(a) Percent with Any Short-Term Disability Insurance Benefit Receipt

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-2013 Mean</td>
<td>19.5</td>
<td>18.7</td>
<td>19.5</td>
<td>18.7</td>
<td>19.5</td>
<td>18.7</td>
<td>19.5</td>
<td>18.7</td>
</tr>
<tr>
<td>Observations</td>
<td>1,913,287</td>
<td>1,662,019</td>
<td>2,020,368</td>
<td>1,746,064</td>
<td>1,913,287</td>
<td>1,662,019</td>
<td>2,020,368</td>
<td>1,746,064</td>
</tr>
<tr>
<td>Controls</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm Fixed Effects</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exclude 2009</td>
<td></td>
<td></td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Total Short-Term Disability Insurance Days

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>1,913,287</td>
<td>1,662,019</td>
<td>2,020,368</td>
<td>1,746,064</td>
<td>1,913,287</td>
<td>1,662,019</td>
<td>2,020,368</td>
<td>1,746,064</td>
</tr>
<tr>
<td>Controls</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm Fixed Effects</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exclude 2009</td>
<td></td>
<td></td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(c) Total Short-Term Disability Insurance Days, Conditional on Any

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre Mean</td>
<td>(4.800)</td>
<td>(5.030)</td>
<td>(4.997)</td>
<td>(5.298)</td>
<td>(4.063)</td>
<td>(4.505)</td>
<td>(5.847)</td>
<td>(5.962)</td>
</tr>
<tr>
<td>Observations</td>
<td>349,857</td>
<td>292,879</td>
<td>360,551</td>
<td>300,675</td>
<td>349,857</td>
<td>292,879</td>
<td>360,551</td>
<td>300,675</td>
</tr>
<tr>
<td>Controls</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm Fixed Effects</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exclude 2009</td>
<td></td>
<td></td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Table shows estimates of the impact of the introduction of experience rating on short-term DI (Ziektever) receipt by temporary agency workers. For each person working in year $t$, we measure short-term DI benefit receipt starting in the same year. These coefficients are from estimating Equation (20). In columns (1), (2), (5), and (6), we control for age (in 1-year bins), gender, healthcare spending (in percentile bins), and region fixed effects. In columns (1), (2), (7), and (8), we control for firm fixed effects. In columns (2), (4), (6), and (8) we exclude data for 2009. Standard errors are clustered at the firm level. For more details, see Section 5.1.
Table 4: Employment

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post</td>
<td>-0.0233</td>
<td>-0.2717</td>
<td>-0.0003</td>
<td>0.0597</td>
<td>0.9830***</td>
<td>0.7550***</td>
<td>-0.7666***</td>
<td>-0.7522***</td>
</tr>
<tr>
<td></td>
<td>(0.1903)</td>
<td>(0.2060)</td>
<td>(0.0411)</td>
<td>(0.0407)</td>
<td>(0.2107)</td>
<td>(0.1651)</td>
<td>(0.0249)</td>
<td>(0.0248)</td>
</tr>
<tr>
<td></td>
<td>(0.8585)</td>
<td>(0.7821)</td>
<td>(0.3605)</td>
<td>(0.3906)</td>
<td>(0.3340)</td>
<td>(0.3537)</td>
<td>(0.9360)</td>
<td>(0.8538)</td>
</tr>
<tr>
<td>Post $\times$ Treatment</td>
<td>2.5264***</td>
<td>2.5712***</td>
<td>2.7485***</td>
<td>2.6697***</td>
<td>2.2655***</td>
<td>2.2425***</td>
<td>3.0867***</td>
<td>3.064***</td>
</tr>
<tr>
<td></td>
<td>(0.1515)</td>
<td>(0.1787)</td>
<td>(0.3142)</td>
<td>(0.2674)</td>
<td>(0.2767)</td>
<td>(0.1639)</td>
<td>(0.1967)</td>
<td></td>
</tr>
<tr>
<td>Pre-2013 Mean</td>
<td>86</td>
<td>86.1</td>
<td>86</td>
<td>86.1</td>
<td>86</td>
<td>86.1</td>
<td>86</td>
<td>86.1</td>
</tr>
<tr>
<td>Observations</td>
<td>41,323,389</td>
<td>36,056,213</td>
<td>42,516,521</td>
<td>36,998,766</td>
<td>41,323,389</td>
<td>36,056,213</td>
<td>42,516,521</td>
<td>36,998,766</td>
</tr>
</tbody>
</table>

Notes: Table shows difference-in-differences estimates of the impact of the introduction of experience rating on employment by temporary agency workers, relative to permanent workers. For each person working in year $t$, we measure employment in year $t+1$. These coefficients are from estimating Equation (18). In columns (1), (2), (5), and (6), we control for age (in 1-year bins), gender, healthcare spending (in percentile bins), and region-by-year fixed effects. In columns (1), (2), (7), and (8), we control for firm fixed effects. In columns (2), (4), (6), and (8) we exclude data for 2009. Standard errors are clustered at the firm level. For more details, see Section 5.1.
Table 5: Disability Insurance Benefit Receipt: Heterogeneity

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>Musculo-skeletal</td>
<td>Mental Health</td>
<td>Male</td>
<td>Female</td>
<td>Age 18-30</td>
<td>Age 31-40</td>
<td>Age 41-50</td>
<td>Age 51-65</td>
</tr>
<tr>
<td>Post</td>
<td>0.0297</td>
<td>0.0039</td>
<td>0.0504***</td>
<td>0.0119</td>
<td>0.0714</td>
<td>-0.0419</td>
<td>0.0557</td>
<td>-0.1358</td>
</tr>
<tr>
<td>Treatment</td>
<td>0.2819***</td>
<td>0.0853***</td>
<td>0.1203***</td>
<td>0.2573***</td>
<td>0.2953***</td>
<td>0.0985***</td>
<td>0.4455***</td>
<td>0.4795***</td>
</tr>
<tr>
<td>Post × Treatment</td>
<td>-0.1203***</td>
<td>-0.0251***</td>
<td>-0.0451***</td>
<td>-0.1100***</td>
<td>-0.1392***</td>
<td>-0.0110</td>
<td>-0.2002***</td>
<td>-0.2359***</td>
</tr>
<tr>
<td>Pre-2013 Mean</td>
<td>0.508</td>
<td>0.123</td>
<td>0.206</td>
<td>0.473</td>
<td>0.559</td>
<td>0.212</td>
<td>0.731</td>
<td>0.969</td>
</tr>
<tr>
<td>Observations</td>
<td>41,323,389</td>
<td>41,323,389</td>
<td>41,323,389</td>
<td>22,241,289</td>
<td>19,082,100</td>
<td>7,063,749</td>
<td>9,489,728</td>
<td>11,781,079</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>Wage Q1</td>
<td>Wage Q2</td>
<td>Wage Q3</td>
<td>Wage Q4</td>
<td>Predicted DI Q1</td>
<td>Predicted DI Q2</td>
<td>Predicted DI Q3</td>
<td>Predicted DI Q4</td>
</tr>
<tr>
<td>Post</td>
<td>0.0297</td>
<td>-0.0908</td>
<td>0.0206</td>
<td>0.2726*</td>
<td>0.0046</td>
<td>0.0621</td>
<td>0.0399</td>
<td>-0.0142</td>
</tr>
<tr>
<td>Treatment</td>
<td>0.2819***</td>
<td>-0.054</td>
<td>-0.009</td>
<td>0.1867***</td>
<td>0.5588***</td>
<td>0.1401***</td>
<td>0.2412***</td>
<td>0.2940***</td>
</tr>
<tr>
<td>Post × Treatment</td>
<td>-0.1203</td>
<td>-0.0241</td>
<td>-0.0581***</td>
<td>-0.1412***</td>
<td>-0.1880***</td>
<td>-0.0693***</td>
<td>-0.0798***</td>
<td>-0.1263***</td>
</tr>
<tr>
<td>Pre-2013 Mean</td>
<td>0.508</td>
<td>0.123</td>
<td>0.206</td>
<td>0.473</td>
<td>0.559</td>
<td>0.212</td>
<td>0.731</td>
<td>0.969</td>
</tr>
<tr>
<td>Observations</td>
<td>41,323,389</td>
<td>2,123,230</td>
<td>2,869,229</td>
<td>4,631,303</td>
<td>31,699,627</td>
<td>9,073,541</td>
<td>8,935,310</td>
<td>8,841,373</td>
</tr>
</tbody>
</table>

Notes: Table shows heterogeneity in our difference-in-differences estimates of the impact of the introduction of experience rating on disability insurance benefit receipt by temporary agency workers, relative to permanent workers. For each person working in year \( t \), we measure disability insurance benefit receipt starting year \( t + 2 \) to account for the two-year sickness period that precedes disability insurance benefit receipt. These coefficients are from estimating Equation (18). Columns (1) and (10) replicate our overall estimate from Table 2. Columns (2) and (3) break down disability insurance benefit receipt by diagnosis category. Columns (4) and (5) provide estimates for male and female subsamples, respectively. Columns (6), (7), (8), and (9) provide estimates for age 18-30, age 31-40, age 41-50, and age 51-65 subsamples, respectively. Columns (11)-(14) provide estimates by wage quartile (defined over the distribution of wages among temporary work agency workers). Columns (15)-(18) provide estimates by quartile of the distribution of predicted DI benefit receipt. In all columns, we control for age (in 1-year bins), gender, healthcare spending (in percentile bins), and region-by-year fixed effects, as well as for firm fixed effects. Standard errors are clustered at the firm level. For more details, see Section 5.1.
Table 6: Short-Term Disability Insurance Benefit Receipt: Heterogeneity

(a) Percent with Any Short-Term Disability Insurance Benefit Receipt

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>Male</td>
<td>Female</td>
<td>Age</td>
<td>Age</td>
<td>Age</td>
<td>Age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18-30</td>
<td>31-40</td>
<td>41-50</td>
<td>51-65</td>
</tr>
<tr>
<td></td>
<td>(0.3439)</td>
<td>(0.3685)</td>
<td>(0.4413)</td>
<td>(0.408)</td>
<td>(0.4143)</td>
<td>(0.2816)</td>
<td>(0.2954)</td>
</tr>
<tr>
<td>Pre-2013 Mean</td>
<td>19.5</td>
<td>18.5</td>
<td>21.0</td>
<td>17.2</td>
<td>24.9</td>
<td>22.7</td>
<td>18.3</td>
</tr>
<tr>
<td>Observations</td>
<td>1,913,287</td>
<td>1,168,804</td>
<td>744,483</td>
<td>1,055,882</td>
<td>353,359</td>
<td>291,982</td>
<td>212,064</td>
</tr>
<tr>
<td>Controls</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Firm Fixed Effects</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

(b) Total Short-Term Disability Insurance Days

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>Male</td>
<td>Female</td>
<td>Age</td>
<td>Age</td>
<td>Age</td>
<td>Age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18-30</td>
<td>31-40</td>
<td>41-50</td>
<td>51-65</td>
</tr>
<tr>
<td></td>
<td>(1.1781)</td>
<td>(0.9529)</td>
<td>(1.5764)</td>
<td>(1.0069)</td>
<td>(1.6348)</td>
<td>(1.5532)</td>
<td>(1.0338)</td>
</tr>
<tr>
<td>Pre-2013 Mean</td>
<td>21.15</td>
<td>18.83</td>
<td>24.56</td>
<td>14.76</td>
<td>29.84</td>
<td>31.25</td>
<td>28.1</td>
</tr>
<tr>
<td>Observations</td>
<td>1,913,287</td>
<td>1,168,804</td>
<td>744,483</td>
<td>1,055,882</td>
<td>353,359</td>
<td>291,982</td>
<td>212,064</td>
</tr>
<tr>
<td>Controls</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Firm Fixed Effects</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>
Table 6: Short-Term Disability Insurance Benefit Receipt: Heterogeneity

(c) Total Short-Term Disability Insurance Days, Conditional on Any

<table>
<thead>
<tr>
<th></th>
<th>(1) Overall</th>
<th>(2) Male</th>
<th>(3) Female</th>
<th>(4) Age 18-30</th>
<th>(5) Age 31-40</th>
<th>(6) Age 41-50</th>
<th>(7) Age 51-65</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-2013 Mean</td>
<td>108.5</td>
<td>101.8</td>
<td>117.2</td>
<td>85.77</td>
<td>119.7</td>
<td>137.9</td>
<td>153.8</td>
</tr>
<tr>
<td>Observations</td>
<td>349,857</td>
<td>203,487</td>
<td>146,370</td>
<td>171,151</td>
<td>81,576</td>
<td>60,796</td>
<td>36,334</td>
</tr>
<tr>
<td>Controls</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Firm Fixed Effects</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-2013 Mean</td>
<td>108.5</td>
<td>97.42</td>
<td>105.6</td>
<td>106.9</td>
<td>119.6</td>
<td>82.56</td>
<td>93.29</td>
<td>110.6</td>
<td>139.3</td>
</tr>
<tr>
<td>Observations</td>
<td>349,857</td>
<td>59,089</td>
<td>89,324</td>
<td>103,372</td>
<td>98,072</td>
<td>53,397</td>
<td>70,358</td>
<td>74,258</td>
<td>80,715</td>
</tr>
<tr>
<td>Controls</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Firm Fixed Effects</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

Notes: Table shows difference-in-differences estimates of the impact of the introduction of experience rating on employment by temporary agency workers, relative to permanent workers. For each person working in year t, we measure employment in year t+1. These coefficients are from estimating Equation (18). These coefficients are from estimating Equation (18). Columns (1) and (8) replicate our main overall estimate from Table 4. Columns (2) and (3) provide estimates for male and female subsamples, respectively. Columns (4), (5), (6), and (7) provide estimates for age 18-30, age 31-40, age 41-50, and age 51-65 subsamples, respectively. Columns (9)-(12) provide estimates by wage quartile (defined over the distribution of wages among temporary work agency workers). Columns (13)-(16) provide estimates by quartile of the distribution of predicted DI benefit receipt. In all columns, we control for age (in 1-year bins), gender, healthcare spending (in percentile bins), and region-by-year fixed effects, as well as for firm fixed effects. Standard errors are clustered at the firm level. For more details, see Section 5.1.
Table 7: Employment: Heterogeneity

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>Male</td>
<td>Female</td>
<td>Age 18-30</td>
<td>Age 31-40</td>
<td>Age 41-50</td>
<td>Age 51-65</td>
</tr>
<tr>
<td>Post</td>
<td>-0.0233</td>
<td>0.2487</td>
<td>-0.5186*</td>
<td>-0.3879</td>
<td>-1.3757***</td>
<td>-0.9889***</td>
<td>1.7182***</td>
</tr>
<tr>
<td>(1.903)</td>
<td>(0.1983)</td>
<td>(0.3006)</td>
<td>(0.3906)</td>
<td>(0.2259)</td>
<td>(0.1917)</td>
<td>(0.4079)</td>
<td></td>
</tr>
<tr>
<td>(0.8585)</td>
<td>(0.7592)</td>
<td>(0.9773)</td>
<td>(0.7704)</td>
<td>(0.6664)</td>
<td>(0.9527)</td>
<td>(1.1812)</td>
<td></td>
</tr>
<tr>
<td>Post × Treatment</td>
<td>2.5264***</td>
<td>2.7077***</td>
<td>2.2325***</td>
<td>2.0561***</td>
<td>3.5479***</td>
<td>3.8426***</td>
<td>3.7762***</td>
</tr>
<tr>
<td>(0.1515)</td>
<td>(0.1418)</td>
<td>(0.2331)</td>
<td>(0.1761)</td>
<td>(0.1879)</td>
<td>(0.1732)</td>
<td>(0.4376)</td>
<td></td>
</tr>
<tr>
<td>Pre-2013 Mean</td>
<td>86.0</td>
<td>86.0</td>
<td>86.1</td>
<td>88.0</td>
<td>84.8</td>
<td>84.2</td>
<td>79.5</td>
</tr>
<tr>
<td>Observations</td>
<td>41,323,389</td>
<td>22,241,289</td>
<td>19,082,100</td>
<td>7,063,749</td>
<td>9,489,728</td>
<td>11,781,079</td>
<td>12,988,833</td>
</tr>
<tr>
<td>Controls</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Firm Fixed Effects</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
<th>(13)</th>
<th>(14)</th>
<th>(15)</th>
<th>(16)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>Wage Q1</td>
<td>Wage Q2</td>
<td>Wage Q3</td>
<td>Wage Q4</td>
<td>Predicted DI Q1</td>
<td>Predicted DI Q2</td>
<td>Predicted DI Q3</td>
<td>Predicted DI Q4</td>
</tr>
<tr>
<td>Post</td>
<td>-0.0233</td>
<td>-0.4469</td>
<td>-0.778</td>
<td>-1.0057***</td>
<td>0.2212</td>
<td>-0.4525**</td>
<td>-0.5605**</td>
<td>0.4168</td>
<td>0.009</td>
</tr>
<tr>
<td>(0.1903)</td>
<td>(1.2131)</td>
<td>(0.7135)</td>
<td>(0.9411)</td>
<td>(0.1871)</td>
<td>(0.2161)</td>
<td>(0.2405)</td>
<td>(0.2858)</td>
<td>(0.3738)</td>
<td></td>
</tr>
<tr>
<td>(0.8585)</td>
<td>(0.9026)</td>
<td>(1.6078)</td>
<td>(1.1024)</td>
<td>(0.69825)</td>
<td>(0.6707)</td>
<td>(0.7614)</td>
<td>(0.7715)</td>
<td>(0.9203)</td>
<td></td>
</tr>
<tr>
<td>Post × Treatment</td>
<td>2.5264***</td>
<td>3.3352***</td>
<td>2.4907***</td>
<td>2.8234***</td>
<td>2.2733***</td>
<td>2.3399***</td>
<td>1.5961***</td>
<td>1.9149***</td>
<td>2.7113***</td>
</tr>
<tr>
<td>(0.1515)</td>
<td>(0.3981)</td>
<td>(2.2500)</td>
<td>(2.0203)</td>
<td>(0.2094)</td>
<td>(0.2323)</td>
<td>(0.1441)</td>
<td>(0.1990)</td>
<td>(0.2564)</td>
<td></td>
</tr>
<tr>
<td>Pre-2013 Mean</td>
<td>86.0</td>
<td>83.5</td>
<td>86.6</td>
<td>88.0</td>
<td>86.1</td>
<td>88.4</td>
<td>87.9</td>
<td>87.2</td>
<td>83.9</td>
</tr>
<tr>
<td>Observations</td>
<td>41,323,389</td>
<td>2,123,230</td>
<td>2,869,229</td>
<td>4,631,303</td>
<td>31,699,627</td>
<td>9,073,541</td>
<td>8,935,310</td>
<td>8,841,373</td>
<td>8,664,911</td>
</tr>
<tr>
<td>Controls</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Firm Fixed Effects</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

Notes: Table shows difference-in-differences estimates of the impact of the introduction of experience rating on employment by temporary agency workers, relative to permanent workers. For each person working in year $t$, we measure employment in year $t+1$. These coefficients are from estimating Equation (18). These coefficients are from estimating Equation (18). Columns (1) and (8) replicate our main overall estimate from Table 4. Columns (2) and (3) provide estimates for male and female subsamples, respectively. Columns (4), (5), (6), and (7) provide estimates for age 18-30, age 31-40, age 41-50, and age 51-65 subsamples, respectively. Columns (9)-(12) provide estimates by wage quartile (defined over the distribution of wages among temporary work agency workers). Columns (13)-(16) provide estimates by quartile of the distribution of predicted DI benefit receipt. In all columns, we control for age (in 1-year bins), gender, healthcare spending (in percentile bins), and region-by-year fixed effects, as well as for firm fixed effects. Standard errors are clustered at the firm level. For more details, see Section 5.1.
Table 8: Disability Insurance Benefit Receipt: Actual and Predicted

<table>
<thead>
<tr>
<th></th>
<th>(1) Actual</th>
<th>(2) Actual</th>
<th>(3) Predicted</th>
<th>(4) Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post</td>
<td>0.0601***</td>
<td>0.0344***</td>
<td>0.0061***</td>
<td>0.0045***</td>
</tr>
<tr>
<td></td>
<td>(0.0025)</td>
<td>(0.0035)</td>
<td>(0.0002)</td>
<td>(0.0004)</td>
</tr>
<tr>
<td>Treatment</td>
<td>0.1411***</td>
<td>0.1468***</td>
<td>-0.0164***</td>
<td>-0.0245***</td>
</tr>
<tr>
<td></td>
<td>(0.0402)</td>
<td>(0.0179)</td>
<td>(0.0035)</td>
<td>0.0018</td>
</tr>
<tr>
<td>Post × Treatment</td>
<td>-0.0996***</td>
<td>-0.0624***</td>
<td>-0.0098***</td>
<td>-0.0083***</td>
</tr>
<tr>
<td></td>
<td>(0.0100)</td>
<td>(0.0163)</td>
<td>(0.0010)</td>
<td>(0.0013)</td>
</tr>
<tr>
<td>Pre-2013 Mean</td>
<td>0.483</td>
<td>0.483</td>
<td>0.33</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Notes: Table shows difference-in-differences estimates of the impact of the introduction of experience rating on actual and predicted disability insurance benefit receipt by temporary agency workers, relative to permanent workers. For each person working in year $t$, we measure disability insurance benefit receipt starting year $t + 2$ to account for the two-year sickness period that precedes disability insurance benefit receipt. These coefficients are from estimating Equation (18). In columns (1) and (2), the outcome is actual disability insurance benefit receipt. In columns (3) and (4), the outcome is predicted disability insurance benefit receipt, estimated based on Equation (21). In columns (1) and (3), we control for firm fixed effects. For more details, see Section 5.1 and Section 5.3.
A Additional Figures and Tables

Appendix Figure A1: Disability Insurance Benefit Receipt

Notes: Figure shows raw trends in disability insurance benefit receipt for permanent workers (blue triangles) and temporary work agency workers (red circles). For each person working in year $t$, we measure disability insurance benefit receipt starting year $t+2$ to account for the two-year sickness period that precedes disability insurance benefit receipt. For more details, see Section 5.1. ($N = 42,516,521$ worker-years)
Appendix Figure A2: Short-Term Disability Insurance Benefit Receipt

(a) Any Short-Term DI Benefit Receipt

(b) Average Total Short-Term DI Days

(c) Average Total Short-Term DI Days, Conditional on Any

(d) Distribution of Short-Term DI Days in 2012 and 2013, Conditional on Any

Notes: Figure shows changes in short-term DI (Ziektewet) benefit receipt over time for temporary work agency workers. Temp agency workers receive benefits through the Ziektewet program during their first two years of sickness, before they are eligible for disability insurance benefits. Panel (a) shows the share with any short-term DI benefit receipt start in year $t$. Panel (b) shows the average number of short-term DI receipt days for sickness spells starting in year $t$. Panel (c) shows the average number of short-term DI benefit receipt days for sickness spells starting in year $t$, conditional on any. Panel (d) shows the distribution in 30-day bins of the average length of sickness spells starting in 2012 (in blue) and in 2013 (in red), conditional on having any sickness spells. For more details, see Section 6. ($N = 2,020,368$ worker-years in Panel (a). $N = 2,020,368$ worker-years in Panel (b). $N = 360,551$ worker-years in Panel (c). $N = 360,551$ worker-years in Panel (d).)
Appendix Figure A3: Employment

Notes: Figure shows raw trends in employment for permanent workers (blue triangles) and temporary work agency workers (red circles). For each person working in year $t$, we measure employment in year $t+1$. For more details, see Section 5.1. ($N = 42,516,521$ worker-years)
Notes: Figure shows the main estimates of the impact of experience rating on DI benefit receipt in prior studies (in blue) and our paper (in red). Koning (2009) studied the introduction of experience rating for permanent workers in the Netherlands in 1998, De Groot and Koning (2016) use the removal of experience rating for permanent workers in small firms in the Netherlands in 2003, and Hawkins and Simola (2020) use the introduction of experience rating for larger firms in Finland in 1996. All of these studies focus on permanent workers. For more details, see Section 6.1.
Appendix Figure A5: Prediction Fit: Log Healthcare Spending vs. Disability Insurance Benefit Receipt

Notes: Figure shows the relationship between log total healthcare spending and DI benefit receipt on a binned scatter plot. Each red dot represents one percentile of the distribution of log total healthcare spending. The grey dashed line represents a quadratic fit. For more details, see Section 5.3.
Appendix Figure A6: Changes in Worker Composition

Notes: Figure shows the change in worker composition by predicted likelihood of disability insurance benefit receipt among temporary agency workers for 2011 (in blue) and 2012 (in red). For more details, see Section 5.3.
**Appendix Table A1: Employment (Two-Year Horizon)**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post</td>
<td>0.4573**</td>
<td>0.6860***</td>
<td>0.2746***</td>
<td>0.5043***</td>
<td>1.6855***</td>
<td>2.0689***</td>
<td>-0.9086***</td>
<td>-0.7600***</td>
</tr>
<tr>
<td></td>
<td>(0.2328)</td>
<td>(0.2554)</td>
<td>(0.0679)</td>
<td>(0.0679)</td>
<td>(0.2551)</td>
<td>(0.2471)</td>
<td>(0.0402)</td>
<td>(0.0400)</td>
</tr>
<tr>
<td></td>
<td>(1.3728)</td>
<td>(1.2549)</td>
<td>(0.5563)</td>
<td>(0.5873)</td>
<td>(0.4891)</td>
<td>(0.5158)</td>
<td>(1.4698)</td>
<td>(1.33336)</td>
</tr>
<tr>
<td>Post x Treatment</td>
<td>4.6839***</td>
<td>5.125***</td>
<td>4.6547***</td>
<td>5.0506***</td>
<td>3.7572***</td>
<td>4.2377***</td>
<td>5.7752***</td>
<td>6.0996***</td>
</tr>
<tr>
<td></td>
<td>(0.2347)</td>
<td>(0.2451)</td>
<td>(0.3510)</td>
<td>(0.3712)</td>
<td>(0.3048)</td>
<td>(0.3182)</td>
<td>(0.2590)</td>
<td>(0.2651)</td>
</tr>
<tr>
<td>Pre-2013 Mean</td>
<td>79.5</td>
<td>78.9</td>
<td>79.5</td>
<td>78.9</td>
<td>79.5</td>
<td>78.9</td>
<td>79.5</td>
<td>78.9</td>
</tr>
<tr>
<td>Observations</td>
<td>41,323,389</td>
<td>36,056,213</td>
<td>42,516,521</td>
<td>36,998,766</td>
<td>41,323,389</td>
<td>36,056,213</td>
<td>42,516,521</td>
<td>36,998,766</td>
</tr>
<tr>
<td>Controls</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm Fixed Effects</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exclude 2009</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Table shows difference-in-differences estimates of the impact of the introduction of experience rating on employment by temporary agency workers, relative to permanent workers. For each person working in year $t$, we measure employment in year $t + 2$. These coefficients are from estimating Equation (18). In columns (1), (2), (5), and (6), we control for age (in 1-year bins), gender, healthcare spending (in percentile bins), and region-by-year fixed effects. In columns (1), (2), (7), and (8), we control for firm fixed effects. In columns (2), (4), (6), and (8) we exclude data for 2009. Standard errors are clustered at the firm level. For more details, see Section 5.1.
B  Institutions

B.1  Disability Insurance Process

Disability insurance eligibility starts after 104 weeks of sickness, after a two-step determination procedure by UWV examiners. First, a physician examiner fills out a standardized 65-item Functional Capacities List (FCL), based on a one-hour interview and the applicant’s medical file. The functional capacities on the FCL relate to limitations in personal functioning (e.g. memory, attention span); social functioning (e.g. sensory capacities, emotional regulation); physical environment (e.g. temperature, noise); dynamic movements (e.g. bending, lifting); static positions (e.g. standing, kneeling); and work hours (per day/week).

A second examiner, called a labor expert, records information on the applicants educational background, and compares the applicants profile with a database consisting of 5,000 actual jobs in the Dutch economy, grouped into 300 occupational titles. This jobs database is actively maintained by a team of 25 professionals, who ensure that the database contains only current jobs, which exist in all five regions of the Netherlands. These job profiles are classified into approximately 300 occupational titles, based on standardized task descriptions. The much larger number of job profiles stems from the fact that multiple job profiles belonging to the same occupational title are entered in the database as long as they each refer to a job in a different region, or with a different work arrangement, or if the number of work hours varies.

Each job profile consists of a standardized questionnaire of occupational demands. The majority of these occupational demands are directly linked to specific FML items describing the applicant’s functional capacities. After an automated preselection algorithm drops job profiles that are automatically rejected because its demands cannot be met according to the applicant’s functional capacity or educational profile, they are ranked by their listed earnings. Starting from the highest-paying remaining job profile in the database, the labor expert examiner determines whether each job profile is appropriate for the applicant. The
three highest-earning appropriate jobs then determine the applicants estimated earnings capacity.

The disability percentage, calculated as the ratio between the estimated earnings capacity and actual prior earnings multiplied by 100, estimates the relative health-related earnings loss. This percentage determines eligibility for (partial) disability benefits, which replace a proportion of prior earnings. In the Netherlands, people with partial disabilities can receive partial benefits. There is no DI payment eligibility under 35 percent disability, payments are progressively more generous between 35 percent and 79 percent disability, and full payments start at 80 percent disability. Importantly, this means that in the context we are studying, certain individuals who receive DI payments are also working.

B.2 Prior Reforms

A key difference between recent Dutch and American disability policy is that after observing that the disability rolls were increasing, policymakers in the Netherlands implemented several new policies. As Koning and Lindeboom (2015) explain, Dutch reforms proceeded in three waves and included both employer-side and worker-side changes. In 1996, employers were made responsible for paying sickness benefits, which then covered the first year of sickness before disability benefits could be claimed. In 1998, experience rating was introduced for permanent workers, which meant that employers would pay different premiums based on their past contribution to the DI rolls. The 2002 introduction of the so called Gatekeeper protocol meant that employers were required to work together with workers on drafting a return-to-work plan and on making sure workers resumed work to the extent possible. In 2004, the sickness benefit period, during which employers are responsible for paying at least 70% of workers’ previous wages was expanded from 1 year to 2 years. Starting in the same year, a large fraction of DI recipients were reexamined, further decreasing the DI rolls. In 2006, another wave of DI reform introduced the distinction between full/permanent disability and partial/temporary disability and tightened eligibility criteria, increasing the minimum
disability threshold for payments from 15% to 35%. At the same time, wage subsidies for partially disabled workers were also introduced. The WIA is the Dutch equivalent of SSDI in the U.S. It includes two schemes, IVA (full, permanent disability) and WGA (partial disability). The WGA scheme applies if the disease duration is more than 2 years, but there is prognosis of (partial) recovery, and the IVA scheme applies if the disease duration exceeds 2 years and recovery is very unlikely. In contrast, in the U.S., Social Security Disability Insurance (SSDI) payments are awarded to all disabled individuals with sufficient work history if their disability has lasted or is expected to last for at least one year or to result in death. In the U.S., in addition to SSDI, the Supplemental Security Income (SSI) program provides benefits to disabled individuals without sufficient work history. This is similar to the WAJONG program in the Netherlands.
C Derivations

C.1 Cost Sharing: Worker Heterogeneity and Selection

We use $\bar{\theta}(\eta)$ and $\bar{\lambda}(\eta)$ to emphasize they are functions of $\eta$. The total lost productivity is

$$L(\eta) = \int_{0}^{\bar{\theta}(\eta)} \lambda f_\theta(\lambda) g(\theta) d\lambda d\theta + \int_{\bar{\theta}(\eta)}^{1} \lambda f_\theta(\lambda) g(\theta) d\lambda d\theta$$

We use the short-hand $h(\theta) = \int_{0}^{\bar{\theta}(\eta)} \lambda f_\theta(\lambda) d\lambda = E_\theta [\lambda \mid b \leq \lambda]$, and $\ell(\theta, \eta) = \int_{0}^{\bar{\lambda}(\eta)} \lambda f_\theta(\lambda) d\lambda = E_\theta [\lambda \mid b \leq \lambda \leq \bar{\lambda}(\eta)]$. Then the total lost productivity can be written as

$$L(\eta) = \int_{0}^{\bar{\theta}(\eta)} h(\theta) g(\theta) d\theta + \int_{\bar{\theta}(\eta)}^{1} \ell(\theta, \eta) g(\theta) d\theta$$

Applying Leibniz integral rule

$$\frac{d}{d\eta} L(\eta) = h(\bar{\theta}(\eta)) g(\bar{\theta}(\eta)) \cdot \frac{d}{d\eta} \bar{\theta}(\eta) - \ell(\bar{\theta}(\eta), \eta) g(\bar{\theta}(\eta)) \cdot \frac{d}{d\eta} \bar{\theta}(\eta)$$

$$+ \int_{\bar{\theta}(\eta)}^{1} \frac{d}{d\eta} \ell(\theta, \eta) g(\theta) d\theta$$

where we apply Leibniz rule again to the last integrand

$$\frac{d}{d\eta} \ell(\theta, \eta) = \frac{d}{d\eta} \int_{b}^{\bar{\lambda}(\eta)} \lambda f_\theta(\lambda) d\lambda$$

$$= \bar{\lambda}(\eta) f_\theta(\bar{\lambda}(\eta)) \cdot \frac{d}{d\eta} \bar{\lambda}(\eta)$$
Plugging terms back in and use the expectation notation we have

\[
\frac{d}{d\eta} L(\eta) = E_{\bar{\theta}(\eta)} [\lambda \mid b \leq \lambda] g(\bar{\theta}(\eta)) \cdot \frac{d}{d\eta} \bar{\theta}(\eta) - E_{\bar{\theta}(\eta)} [\lambda \mid b \leq \lambda \leq \bar{\lambda}(\eta)] g(\bar{\theta}(\eta)) \cdot \frac{d}{d\eta} \bar{\theta}(\eta) \\
+ \bar{\lambda}(\eta) \cdot \frac{d}{d\eta} \bar{\lambda}(\eta) \int_{\bar{\theta}(\eta)}^{1} f_{\theta}(\bar{\lambda}(\eta)) g(\theta) d\theta \\
= g(\bar{\theta}(\eta)) \cdot \frac{d}{d\eta} \bar{\theta}(\eta) \cdot (E_{\bar{\theta}(\eta)} [\lambda \mid b \leq \lambda] - E_{\bar{\theta}(\eta)} [\lambda \mid b \leq \lambda \leq \bar{\lambda}(\eta)]) \\
+ \bar{\lambda}(\eta) \cdot \frac{d}{d\eta} \bar{\lambda}(\eta) \cdot \int_{\bar{\theta}(\eta)}^{1} f_{\theta}(\bar{\lambda}(\eta)) g(\theta) d\theta
\]
D Data Description

D.1 Personal Information

Personal information, including birth date, gender, ethnic origin, and number of parents born abroad is built from the Municipal Personal Records Database (Gemeentelijke Basisadministratie Persoonsgegevens, or GBA). Appendix Table D2 lists the variables used. In addition, we use the crosswalk provided by Statistics Netherlands called LANDAKTUELREF to convert birth countries into ethnic origins (Native Dutch, Moroccan, Turkish, Surinamese, Netherlands Antillean and Aruban, Other Non-Western, Other Western, Unknown). We use the database VSLGWB that converts addresses into municipalities and an additional crosswalk provided Statistics Netherlands to convert municipalities into regions.

Appendix Table D2: Municipal Personal Records Database (Gemeentelijke Basisadministratie Persoonsgegevens, or GBA)

<table>
<thead>
<tr>
<th>Data Element</th>
<th>Dataset</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person ID</td>
<td>GBAPERSSONTAB 2016</td>
<td>rinpersoons, rinpersoon</td>
</tr>
<tr>
<td>Birth Year</td>
<td>GBAPERSSONTAB 2016</td>
<td>gbageboortejaar</td>
</tr>
<tr>
<td>Gender</td>
<td>GBAPERSSONTAB 2016</td>
<td>gbagecles</td>
</tr>
<tr>
<td>Birth Country</td>
<td>GBAPERSSONTAB 2016</td>
<td>gbageboorteland</td>
</tr>
<tr>
<td>Number of Parents Born Abroad</td>
<td>GBAPERSSONTAB 2016</td>
<td>gbaaantaloudersbuitenland</td>
</tr>
<tr>
<td>Person ID</td>
<td>GBAADRESOBJECTBU.S. 2016</td>
<td>rinpersoons, rinpersoon</td>
</tr>
<tr>
<td>Address Begin Date</td>
<td>GBAADRESOBJECTBU.S. 2016</td>
<td>gbadatumaanvangadreshouding</td>
</tr>
<tr>
<td>Address End Date</td>
<td>GBAADRESOBJECTBU.S. 2016</td>
<td>gbadatumeindeadreshouding</td>
</tr>
</tbody>
</table>

D.2 Employment Information

Employment information, including firm identifiers, contract types, wages, and hours, is built from the Jobs and Wages Administrative Database (Banen en lonen op basis van de Polisadministratie, or SPOLIS). Appendix Table D3 lists the variables used.
Appendix Table D3: Jobs and Wages Administrative Database (Banen en lonen op basis van de Polisadministratie, or SPOLIS)

<table>
<thead>
<tr>
<th>Data Element</th>
<th>Dataset</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person ID</td>
<td>SPOLIS</td>
<td>rinpersoons, rinpersoon</td>
</tr>
<tr>
<td>Firm ID</td>
<td>SPOLIS</td>
<td>sbeid</td>
</tr>
<tr>
<td>Sector ID</td>
<td>SPOLIS</td>
<td>ssect</td>
</tr>
<tr>
<td>Contract Type</td>
<td>SPOLIS</td>
<td>scontractsoort</td>
</tr>
<tr>
<td>Labor Relation</td>
<td>SPOLIS</td>
<td>sarbeidsrelatie</td>
</tr>
<tr>
<td>Job Type</td>
<td>SPOLIS</td>
<td>ssoortbaan</td>
</tr>
<tr>
<td>Wage</td>
<td>SPOLIS</td>
<td>slnlbph</td>
</tr>
<tr>
<td>Hours</td>
<td>SPOLIS</td>
<td>sbasisuren</td>
</tr>
</tbody>
</table>

D.3 Disability Insurance Benefit Receipt

Information on disability insurance benefit receipt, including benefit receipt dates, disability program used, diagnosis, degree of disability, and benefit amount is built from the Disability Benefit Tabulation (Aototuitkeringtab). We do not have information on disability applications. Appendix Table D4 lists the variables used.

Appendix Table D4: Disability Benefit Tabulation (Aototuitkeringtab)

<table>
<thead>
<tr>
<th>Data Element</th>
<th>Dataset</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person ID</td>
<td>Aototuitkeringtab</td>
<td>rinpersoons, rinpersoon</td>
</tr>
<tr>
<td>Program</td>
<td>Aototuitkeringtab</td>
<td>AOTotWetcode</td>
</tr>
<tr>
<td>Begin Date</td>
<td>Aototuitkeringtab</td>
<td>AOTotDatumAanvang</td>
</tr>
<tr>
<td>Degree of Disability</td>
<td>Aototuitkeringtab</td>
<td>AOTotUitkeringspercentage</td>
</tr>
<tr>
<td>Diagnosis Code</td>
<td>Aototuitkeringtab</td>
<td>AOTotDiagnose</td>
</tr>
</tbody>
</table>

D.4 Short-Term Disability Insurance Benefit Receipt

Information on short-term disability insurance (Ziektwet) benefit receipt, including start and ending dates is built from two social security datasets (Ovuitkpersoonbus and Zwpersoonmndbedragbus). Appendix Table D5 lists the variables used.
Appendix Table D5: Sickness Benefit Uptake (Ovuitkpersoonbus and ZwPersoonMndbedragbus)

<table>
<thead>
<tr>
<th>Data Element</th>
<th>Dataset</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person ID</td>
<td>Ovuitkpersoonbus</td>
<td>rinpersoons, rinpersoon</td>
</tr>
<tr>
<td>Begin Date</td>
<td>Ovuitkpersoonbus</td>
<td>begingeldigheid</td>
</tr>
<tr>
<td>End Date</td>
<td>Ovuitkpersoonbus</td>
<td>eindgeldigheid</td>
</tr>
<tr>
<td>Person ID</td>
<td>ZwPersoonMndbedragbus</td>
<td>RINPERSOONS, RINPERSOON</td>
</tr>
<tr>
<td>Begin Date</td>
<td>ZwPersoonMndbedragbus</td>
<td>AANVMNDBEDRAGZW</td>
</tr>
<tr>
<td>End Date</td>
<td>ZwPersoonMndbedragbus</td>
<td>EINDMNDBEDRAGZW</td>
</tr>
</tbody>
</table>

D.5 Healthcare Spending

Information on healthcare spending We use the Healthcare Cost Tabulation (Zvwzorgkosten) to calculate annual healthcare spending for each individual. The data contain annual spending by category (inpatient hospital, mental health, prescription drugs, etc.) and we add up these categories to calculate overall annual spending. Appendix Table D6 lists the variables used.
Appendix Table D6: Healthcare Cost Tabulation (Zvwzorgkostentab)

<table>
<thead>
<tr>
<th>Data Element</th>
<th>Dataset</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person ID</td>
<td>Zvwzorgkostentab</td>
<td>rinpersoons, rinpersoon</td>
</tr>
<tr>
<td>Primary Care Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKHUISARTS</td>
</tr>
<tr>
<td>Prescription Drug Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKFARMACIE</td>
</tr>
<tr>
<td>Oral Care Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKMONDZORG</td>
</tr>
<tr>
<td>Inpatient Care Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKZIEKENHUIS</td>
</tr>
<tr>
<td>Paramedical Care Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKPARAMEDISCH</td>
</tr>
<tr>
<td>Medical Aids Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKHULPMIDDEL</td>
</tr>
<tr>
<td>Medical Transportation Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKZIEKENVERVOER</td>
</tr>
<tr>
<td>Childbirth Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKGEBOORTEZORG</td>
</tr>
<tr>
<td>Care Costs Incurred Abroad</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKBUITENLAND</td>
</tr>
<tr>
<td>Other Basic Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKOVERIG</td>
</tr>
<tr>
<td>Primary Mental Health Care Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKEERSTELIJNPSYCHO</td>
</tr>
<tr>
<td>Secondary Mental Health Care Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKGGZ</td>
</tr>
<tr>
<td>Other Basic Mental Health Care Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKGENBASGGZ</td>
</tr>
<tr>
<td>Other Specialist Mental Health Care Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKSPECGGZ</td>
</tr>
<tr>
<td>Geriatric Rehabilitation Care Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKGERIATRISCH</td>
</tr>
<tr>
<td>Home Nursing Costs</td>
<td>Zvwzorgkostentab</td>
<td>ZVWKWYKVERPLEGING</td>
</tr>
</tbody>
</table>

D.6 Education

We use the Highest Education Level Tabulation (Hoogsteopltab) to categorize individuals by their level of education. The data contain detailed educational attainment in 18 different categories and we aggregate these into three broad categories. Appendix Table D7 lists the variables used.

Appendix Table D7: Highest Education Level Tabulation (Hoogsteopltab)

<table>
<thead>
<tr>
<th>Data Element</th>
<th>Dataset</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person ID</td>
<td>Hoogsteopltab</td>
<td>rinpersoons, rinpersoon</td>
</tr>
<tr>
<td>Education Level</td>
<td>Hoogsteopltab</td>
<td>oplnivsoi2016agg4hbetnirwo</td>
</tr>
</tbody>
</table>