DISCUSSION PAPER SERIES

IZA

IZA DP No. 9692

The Effectiveness of Medical and Vocational Interventions for Reducing Sick Leave of Self-Employed Workers

Stijn Baert
Bas van der Klaauw
Gijsbert van Lomwel

January 2016

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Effectiveness of Medical and Vocational Interventions for Reducing Sick Leave of Self-Employed Workers

Stijn Baert
Ghent University, University of Antwerp, Université catholique de Louvain and IZA

Bas van der Klaauw
VU University Amsterdam, Tinbergen Institute, CEPR and IZA

Gijsbert van Lomwel
UWV

Discussion Paper No. 9692
January 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

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

IZA Discussion Paper No. 9692
January 2016

ABSTRACT

The Effectiveness of Medical and Vocational Interventions for Reducing Sick Leave of Self-Employed Workers*

We investigate whether interventions by (i) medical doctors and (ii) occupational specialists are effective in reducing sick leave durations among self-employed workers. To this end, we exploit unique administrative data comprising all sick leave claims by self-employed workers insured with the major Dutch private insurer between January 2009 and March 2014. We estimate a multivariate duration model dealing with non-random selection into the two intervention types by controlling for observable and unobservable claimant characteristics. We find adverse treatment effects for both interventions, which are heterogeneous by the physical toughness of the claimants' occupation.

JEL Classification: C41, I13, J22, R31

Keywords: sickness absenteeism, self-employment, medical interventions, dynamic treatment effects

Corresponding author:

Stijn Baert
Ghent University
Sint-Pietersplein 6
9000 Gent
Belgium
E-mail: Stijn.Baert@UGent.be

* The authors acknowledge Achmea for providing the data. In addition, Stijn Baert is grateful to the department of Economics of the VU University Amsterdam for giving him the opportunity to work on this study as a visiting researcher at the department. Lastly, the authors acknowledge Hans Bloemen, Bart Cockx, Jochen Mierau, Matteo Picchio and the seminar participants at Aarhus University, Ghent University and Université Catholique de Louvain for their insightful comments and suggestions, which have helped to improve this study considerably.

1 Introduction

Over the past decades, many studies have shown the presence of moral hazard in public sickness insurance systems. For example, Barmby et al. (1991), Johansson and Palme (2002), Henrekson and Persson (2004), Johansson and Palme (2005), Puhani and Sonderhof (2010), Ziebarth (2010), Markussen et al. (2011), Ziebarth (2013), De Paola et al. (2014), Fevang et al. (2014) and Ziebarth (2014) indicate that more generous sick leave benefits increase the incidence and/or the duration of sickness absenteeism. In addition, some studies have investigated the effectiveness of medical practitioners in reducing this sick leave. Carlsen and Nyborg (2009) find, based on focus group interviews in Norway, that general practitioners fail as gatekeepers. They relate this empirical finding to the fact that (i) general practitioners are unable to distinguish shirkers from truly sick, and that (ii) patients, truly sick or not, prefer – and, therefore, engage – physicians who give priority to healing over gatekeeping. However, Hartman et al. (2013) find that medical certification is an important instrument for managing sickness absenteeism in Sweden. In the same direction, Markussen (2010) shows that the introduction of stricter regulations for physicians' sick leave certification in Norway resulted both in lower sick leave entry rates and in lower recovery rates.

This evidence with respect to both moral hazard in sickness insurance and the effectiveness of medical practitioners to reduce moral hazard may, however, not be generalised to self-employed workers. In many OECD countries, self-employed workers are not covered by the public sickness insurance system so that they have to buy insurance on the private market. The fact that self-employed workers decide themselves on the level of private sickness insurance may lead to different insurance, intervention and recovery dynamics. Furthermore, there are several reasons why self-employed workers have more interest in reducing their absence durations, even beyond what is optimal from a health point of view. First, financial incentives to avoid sickness absenteeism are often larger for self-employed workers. A long period of absence may lead to lost investments and irrecoverable loss of market share because finding an adequate substitute might be difficult (Hyytinen and Ruuskanen, 2007). Second, although they experience, on average, more stress than employees, self-employed workers are found to be more satisfied and involved with their jobs (Blanchflower and Oswald, 1998; Parasuraman and Simmers, 2001; Parker, 2004; Hyytinen and Ruuskanen, 2007). Finally, following Parker (2004) and Lechmann and Schnabel (2014) self-employed workers are characterised by a higher need for achievement, love of independence, risk taking propensity and optimism level. These personal characteristics seem in favour of short absence durations.

In this study, we are the first to investigate whether medical interventions are effective in reducing sick leave durations among self-employed workers. Therefore, we exploit unique administrative data from the major Dutch private insurance company.

Our analysis is based on all (i.e. more than 15,000) sickness benefit applications of self-employed workers suffering from a physical condition between January 2009 and March 2014. As the insurance company uses both “medical track” and “labour track” interventions, we are able to compare the relative effectiveness of medical doctors (offering medical support) versus occupational specialists (offering ergonomic advice and coaching) in reducing sick leave durations. In the data, claims and interventions are recorded with a daily precision. Moreover, the company’s database provides detailed claimant information.

In the spirit of Abbring and Van den Berg (2003), we exploit the time variation in the start of the intervention tracks to capture causal effects on the timing of recovery. More precisely, we develop a multivariate duration model that deals with the non-random and dynamic selection into the tracks by controlling for observable and unobservable intervention determinants. This model allows to identify heterogeneous treatment effects with respect to the moment of intervention and claim(ant) characteristics.

From a policy perspective, an answer to our research question is crucial to insurers who are responsible for paying sickness benefits and decide about engaging doctors and occupational specialists in order to minimise these payments. Furthermore, self-employed workers may financially suffer from sickness absenteeism so they might be interested in the effectiveness of medical intervention, ergonomic advice and coaching themselves. Lastly, the results of this study are relevant to public policymakers who are interested in a stronger role for doctors and occupational specialists as gatekeepers of the welfare state.[1]

The outline of this paper is as follows. Section 2 describes the institutional background concerning the private sickness insurance system in the Netherlands and the medical interventions. In Section 3 we present our data and provide a descriptive analysis. Section 4 introduces the econometric model and Section 5 contains our estimation results. The final section concludes.

2 Institutional Setting

In the Netherlands, as in many OECD countries, self-employed workers are exempted from the public sickness and disability insurance that is provided to employees. Therefore, they have to buy this insurance on the private market. Slightly over 10% of the

[1] Recent reforms in the Dutch sickness and disability insurance system focussed on empowering employers. In particular, more financial incentives for employers were introduced and employers were given more responsibilities in stimulating a fast return to work. This is often argued to be an important determinant in the recent reduction of long-term sickness absenteeism and the inflow into disability insurance in the Netherlands (Koning, 2004; De Jong et al., 2011). Implementing this for self-employed workers is problematic, simply because one cannot separate between the employer and the worker.

Dutch labour force consist of self-employed workers, which is comparable to other Western European countries. Spierdijk et al. (2009) show that long-term sickness prevalence among Dutch self-employed workers is about 6%.[2]

Sickness insurance plans for self-employed workers may differ between insurance companies and often insurance companies offer various plans. We analyse data from the major Dutch private insurance company. When buying sickness insurance from this company, a self-employed worker has to decide on a number of modalities. Two modalities are particularly relevant for our study. The first is the *deferment period*, which is the time period between falling sick and the start of benefit payment. The second relevant modality is the *insured income*, which is at most 80% of the income of the self-employed worker. In Subsection 3.2 we present statistics for both modalities.

Since 2003, the insurance company employs active case management in order to enhance recovery rates. The program starts with an intake interview conducted by a caseworker. During this interview, an initial medical diagnosis is determined. In addition, information is gathered about the type of business of the self-employed worker and her/his existing health limitations. Next, within the first weeks after intake of the claim, the caseworker discusses all gathered information with a medical doctor and an occupational specialist engaged by the private insurer. Together they decide about the most appropriate intervention. The potential interventions are classified into two tracks, which are used independently.

The first track is the *medical track*, in which physicians are engaged to speed up recovery. This track takes off with the claimant visiting a medical doctor who thereafter provides a second opinion concerning the degree of disability. Based on her/his advise, eventually supplemented by information collected from the claimant's general practitioner, further medical interventions are carried out by medical doctors. The second track is the *labour track*, where an occupational specialist is assigned to the claimant. The occupational specialist provides ergonomic advice to the claimant and coaches him/her back to work.

3 Data

3.1 Sample of Analysis

Our data are provided by the major private insurance company in the Netherlands and contain all sickness spells between January 20, 2009 and March 31, 2014 of self-employed workers insured against income loss due to absenteeism. The exact number of days of

[2]For more background information on relevant Dutch labour market institutions, we refer to Spierdijk et al. (2009), De Jong (2012) and Gautier and van der Klaauw (2012).

sickness absenteeism is recorded either until recovery or until March 31, 2014. For each claim, the start of medical and labour track interventions are recorded with a daily precision. In total, the data include 19,488 claims.

For 19,138 claims, an initial medical diagnosis is determined during the intake interview. From the claims with an initial medical diagnosis, we retain the 15,616 claims with a physical condition. Excluded are 1,668 claims for maternity leave, which all have a fixed duration of 113 days, and 1,854 claims with a psychological condition. The latter claims have very different recovery and intervention dynamics than physical claims.[3] The subsample of psychological claims (alone) is too small for an empirical analysis. Next, we drop 16 claims with negative duration times, 26 claims with missing explanatory variables and 26 claims where an intervention start after recovery.

Our data suffer from the problem that short sickness spells may not be reported. If the self-employed worker knows that she/he will recover before the end of the deferment period, there is no direct incentive to report the sickness to the insurance company (even though the company requests to report all spells). Therefore, we set the start of our duration model for recovery to ten days after reporting sick and drop the 136 spells with a sickness duration less than ten days. Short spells are often due to the flu or less serious injuries, and the insurance company never intervenes within the first few days of sickness. Furthermore, we censor durations after 548 days (one and a half year). This avoids that we have to model outliers and less than 1% of the medical track interventions and 5% of the labour track interventions start after 548 days. Sensitivity checks (see Subsection 5.3) show that both choices do not substantially affect our estimation results.

3.2 Descriptive Analysis

We observe for each claim three durations: the *duration until recovery*, the *duration until entering the medical track* and the *duration until entering the labour track*. There are 6252 individuals who enter the medical track and 2888 individuals entering the labour track. In 2498 cases both tracks start during the period of sickness absenteeism. Figure 1 and Figure 2 show the Kaplan-Meier estimates for the survival functions with respect to entering the medical and labour track. The median duration until entering the programs is 134 days for the medical track and 388 days for the labour track.

Figure 3 reports Kaplan-Meier estimates for the survival function until recovery (before right censoring) by intervention. The median sick leave duration is 57 days for claimants who do not participate in any track, it is 184 and 238 days for those who are treated exclusively by the medical track and the labour track respectively, and more

[3] Recovery rates are lower for claimants with a psychological condition during the first seven months of sick leave and higher afterwards. In addition, claimants with a psychological condition have a higher probability of entering the labour track and also the medical track during the first three months of sick leave.

Figure 1: Kaplan-Meier estimates for entering the medical track.

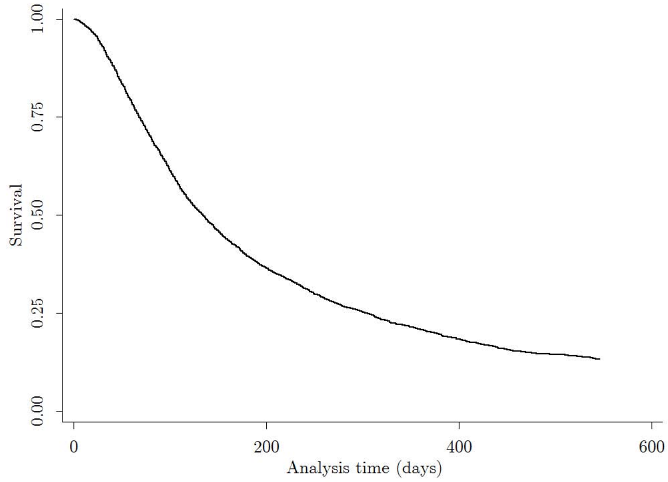


Figure 2: Kaplan-Meier estimates for entering the labour track.

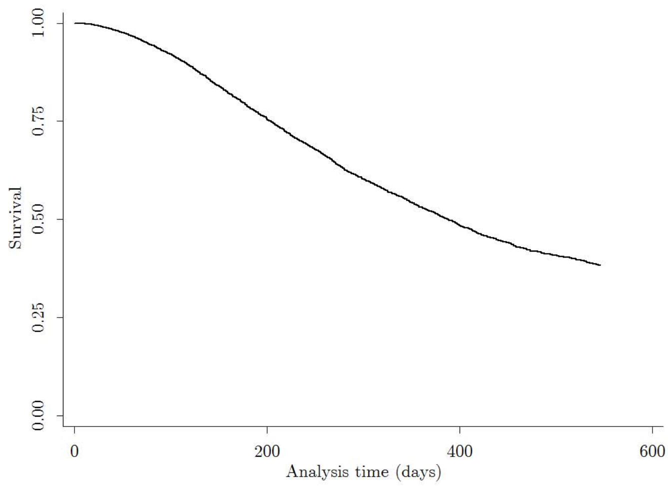
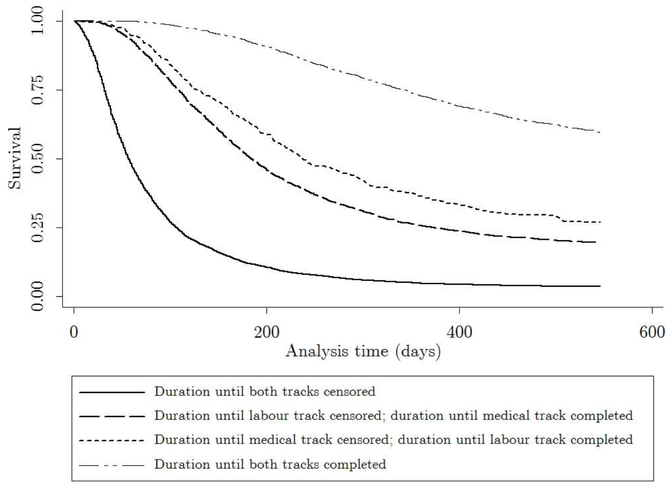


Figure 3: Kaplan-Meier estimates for recovery.



than 548 days for those who are treated by both tracks. These differences should not be given a causal interpretation. The composition of individuals varies between the four groups, and there is the *dynamic selection problem.* As intervention tracks do not start immediately when a claimant enters sick leave, participation can only be observed for claimants who are absent sufficiently long.

Our data contain an extensive set of observed claim(ant) characteristics. In Table 1 we present summary statistics for these variables. We report these statistics both for the total sample and for the four subsamples by undergone intervention. The majority of the individuals are between 36 and 55 years old (at the start of the claim). The subsample of control claimants contains both more individuals who are younger than 36 and individuals who are older than 55, while the treated claimants are overrepresented within the middle age categories. Women have a relatively higher probability of entering the labour track intervention. There is no large compositional difference in the region from which the subsamples of individuals come.

Concerning the occupational type and its toughness, Table 1 shows that the medical track is used relatively more for agriculturalists, small and medium entrepreneurs and – more general – for tough occupations while the labour track is used more among liberal and (rather) light occupations. Of particular interest are the deferment period and the insured income, both captured by four indicator variables. The shorter the deferment period and the higher the insured income are, the more generous is the sick leave compensation. However, we do not find evidence for systematic higher intervention rates for claimants with more generous compensations.

Table 1: Summary Statistics.

Subsample:	All	C	M	L	ML
Age					
< 36	0.187 (0.390)	0.207 (0.405)	0.176 (0.380)	0.115 (0.320)	0.143 (0.350)
36 – 45	0.332 (0.471)	0.315 (0.465)	0.344 (0.475)	0.367 (0.483)	0.370 (0.483)
46 – 55	0.357 (0.479)	0.343 (0.475)	0.357 (0.479)	0.415 (0.493)	0.398 (0.489)
> 55	0.124 (0.329)	0.134 (0.341)	0.123 (0.329)	0.103 (0.304)	0.090 (0.286)
Gender					
Female	0.129 (0.335)	0.136 (0.342)	0.111 (0.314)	0.174 (0.380)	0.127 (0.332)
Region					
North	0.408 (0.492)	0.402 (0.490)	0.428 (0.495)	0.415 (0.493)	0.400 (0.490)
South	0.429 (0.495)	0.427 (0.495)	0.423 (0.494)	0.405 (0.492)	0.450 (0.498)
Center	0.162 (0.369)	0.171 (0.376)	0.148 (0.356)	0.179 (0.384)	0.150 (0.357)
Occupation type					
Agricultural	0.391 (0.488)	0.383 (0.486)	0.425 (0.494)	0.354 (0.479)	0.376 (0.484)
SME	0.456 (0.498)	0.440 (0.497)	0.473 (0.500)	0.410 (0.492)	0.488 (0.500)
Liberal profession	0.153 (0.360)	0.177 (0.381)	0.101 (0.302)	0.236 (0.425)	0.136 (0.343)
Toughness of occupation					
(Rather) light	0.175 (0.380)	0.183 (0.387)	0.127 (0.332)	0.282 (0.451)	0.203 (0.402)
Rather tough	0.168 (0.374)	0.187 (0.390)	0.123 (0.328)	0.218 (0.413)	0.163 (0.369)
Tough	0.657 (0.475)	0.630 (0.483)	0.751 (0.433)	0.500 (0.501)	0.635 (0.482)
Insured income					
< €100M	0.156 (0.363)	0.158 (0.365)	0.174 (0.379)	0.087 (0.282)	0.130 (0.336)
€100M – €500M	0.366 (0.482)	0.341 (0.474)	0.378 (0.485)	0.441 (0.497)	0.422 (0.494)
€500M – €1000M	0.277 (0.447)	0.274 (0.446)	0.278 (0.448)	0.277 (0.448)	0.285 (0.451)
> €1000M	0.202 (0.401)	0.227 (0.419)	0.170 (0.375)	0.195 (0.397)	0.163 (0.370)
Deferment period					
< 14 days	0.351 (0.477)	0.349 (0.477)	0.396 (0.489)	0.244 (0.430)	0.305 (0.461)
14 days – 3 months	0.483 (0.500)	0.494 (0.500)	0.468 (0.499)	0.515 (0.500)	0.463 (0.499)
3 months – 1 year	0.107 (0.309)	0.099 (0.299)	0.088 (0.283)	0.138 (0.346)	0.157 (0.364)
> 1 year	0.059 (0.236)	0.058 (0.233)	0.049 (0.216)	0.103 (0.304)	0.075 (0.263)
Observations	15412	8870	3754	390	2498

Means and standard deviations in parentheses.

C: duration until both tracks censored.

M: duration until labour track censored; duration until medical track completed.

L: duration until medical track censored; duration until labour track completed.

ML: duration until both tracks completed.

4 Econometric Model

The goal of our econometric analysis is to estimate the causal effects of entering the medical track and/or the labour track on recovery from sickness absenteeism. To this end, we jointly model the process of recovery and the entry processes into both tracks. Our model builds on the timing-of-events framework of Abbring and Van den Berg (2003). This framework is ideal for studying interventions in a dynamic setting because it deals, under certain identifying assumptions, with both selective participation and dynamic selection.

4.1 Econometric Framework

Consider a self-employed worker who first reports sick at (calender) date τ_0. Our model is a continuous-time duration model in which t describes the elapsed sickness duration and t_m and t_l the durations until entering the medical and labour track, respectively. Let θ_r denote the rate at which self-employed workers recover from sickness. This recovery rate can depend on the elapsed sickness duration t, observed characteristics x, calendar time $\tau_0 + t$, unobserved characteristics v and variables indicating whether the medical track $I(t_m < t)$ and labour track $I(t_l < t)$ have been started (with $I(\cdot)$ the indicator function).

We denote the unobserved term v in the recovery rate by v_r. This term is assumed to be independent of x and τ_0. Since the variables in x are mainly used as control variables and we will not causally interpret their covariate effect, this is not a strong assumption. Conditional on x, τ_0, v_r, t_m and t_l, the rate of recovery after t periods of sickness absenteeism follows a mixed proportional hazard specification as described in Van den Berg (2001):

$$\begin{aligned} \ln \theta_r(t|x, \tau_0, v_r, t_m, t_l) &= \lambda_r(t) + \psi_r(\tau_0 + t) + x'\beta_r + \delta_m(t|t_m, x)I(t_m < t) \\ &\quad + \delta_l(t|t_l, x)I(t_l < t) + v_r. \end{aligned} \tag{1}$$

In this specification $\psi_r(\tau_0+t)$ is a genuine calendar-time effect modelled by dummies for each quarter. These calendar-time effects control both for seasonal effects in recovery and for the macroeconomic context. In addition, the function $\lambda_r(t)$ represents the duration dependence. The functions $\delta_m(t|t_m, x)$ and $\delta_l(t|t_l, x)$ are the key parameters of interest as they describe the causal effects of participation in the medical track and the labour track, respectively. We return to the parameterisation of the functions at the right-hand side of equation (1) in the next two subsections.

The timing of entering the medical and labour track is most likely not exogenously determined. Therefore, we jointly model the timing of entering these tracks as mixed proportional hazard specifications:

$$\begin{aligned} \ln \theta_m(t|x, \tau_0, v_m) &= \lambda_m(t) + \psi_m(\tau_0 + t) + x'\beta_m + v_m; \\ \ln \theta_l(t|x, \tau_0, v_l) &= \lambda_l(t) + \psi_l(\tau_0 + t) + x'\beta_l + v_l. \end{aligned} \tag{2}$$

Both hazard rates describe the rate of entering the tracks given that the sick self-employed worker has not yet entered this track. The hazard rates depend on the same set of observed characteristics x as those determining the recovery rate.

Now consider the joint distribution of t_r, t_m and t_l. Conditional on τ_0, x, v_r, v_m and v_l, the only possible relation between t_r and (t_m, t_l) goes via the direct effects of participating in the medical track and the labour track (on the recovery rate). In case of independence between v_r and (v_m, v_l), we have a standard duration model for $t_r|x, \tau_0, t_m, t_l$ with $I(t_m < t)$ and $I(t_l < t)$ time-varying regressors which are orthogonal to the unobserved heterogeneity v_r. However, if v_r and (v_m, v_l) are not independent, inference on $t_r|x, \tau_0, t_m, t_l$ should be based on $(t_r, t_m, t_l)|x, \tau_0$.

It is straightforward to derive the likelihood contributions. The use of a flow sample of self-employed workers starting a period of sickness absenteeism implies that we do not have any initial conditions problems. The right-censoring in the data is independent, and is, therefore, solved in a straightforward manner. In particular, let c_r equal one if a self-employed worker is observed to recover from sickness absenteeism, and c_m and c_l equal one if the worker entered, respectively, the medical track and the labour track during the period of sickness absenteeism. If $i = 1, \ldots, n$ denote the observations, then the loglikelihood function equals:

$$
\begin{aligned}
\log \ell \;=\; \sum_{i=1}^{n} \log \Bigg\{ \int_{v_r} \int_{v_m} \int_{v_l} & \theta_r(t_{r,i}|x_i, \tau_{0,i}, v_r, t_{m,i}, t_{m,i})^{c_{r,i}} \\
& \exp\left(- \int_0^{t_{r,i}} \theta_r(t_{r,i}|x_i, \tau_{0,i}, v_r, t_{m,i}, t_{m,i}) dz \right) \\
\theta_m(t_{m,i}|x_i, \tau_{0,i}, v_m)^{c_{m,i}} & \exp\left(- \int_0^{t_{m,i}} \theta_m(z|x_i, \tau_{0,i}, v_m) dz \right) \\
\theta_l(t_{l,i}|x_i, \tau_{0,i}, v_l)^{c_{l,i}} & \exp\left(- \int_0^{t_{l,i}} \theta_l(z|x_i, \tau_{0,i}, v_l) dz \right) dG(v_r, v_m, v_l) \Bigg\}; \qquad (3)
\end{aligned}
$$

where $G(v_r, v_m, v_l)$ is the joint distribution of the unobserved characteristics (v_r, v_m, v_l). If a worker is still sick at the end of the observation period ($c_r = 0$), then t_r equals the duration until right-censoring. Furthermore, if during the period of sickness absenteeism no medical track intervention has been observed ($c_m = 0$), then t_m is set equal to the observed period of sickness absenteeism t_r (which is the moment of censoring the duration until a medical intervention). Analogously, if no labour track intervention has been observed ($c_l = 0$), then t_l is set equal to t_r.

4.2 Identification of the Treatment Effects

The main parameters of interest are the effects of participating in the medical track and the labour track on recovery from sickness absenteeism. There are two complications in their empirical evaluation. First, there may be selection on (un)observable claim(ant) characteristics when assigning sick workers to both tracks. Second, since participation in the tracks does not start at the beginning of sick leave but during the spell, those with a long sickness spell are more likely to enter the tracks. The second complication is solved by the dynamic structure of the model, which explicitly accounts for the length of sickness spells. The first complication deals with the essential identification problem in dynamic settings.

Abbring and Van den Berg (2003) provide an extensive discussion on the identification of dynamic treat effects in duration models. The key assumption for assigning a causal interpretation to the effects $\delta_m(t|t_m, x)$ and $\delta_l(t|t_l, x)$ of participation in the medical and labour track is that the moment of starting the tracks is not anticipated. No anticipation implies that conditional on both observed and unobserved characteristics, the recovery rate at each moment of time does not depend on the exact timing of track participation in the future. This does not imply that participating in the tracks is exogenous. Based on both observed and unobserved characteristics sick workers may have different intervention rates, and these intervention rates may change during the spell of sickness absenteeism. The timing-of-events framework explicitly allows for selection on unobservables.

In our institutional setting, the insurer aims at minimising the waiting times before entering treatment. Not only because the insurer wants to act quickly, but also because she/he wants to avoid uncertainty for self-employed workers. In practice, this implies that once the caseworker decides that a given track is useful for the worker, the worker enters this track as soon as possible. The waiting times between the caseworker announcing entering a track and the start of the track are, therefore, more in terms of days than in terms of weeks. This implies that the anticipation period is short and likely unimportant.

If the assumption of no anticipation is satisfied, no exclusions restrictions are necessary to identify the causal intervention effects.[4] However, Abbring and Van den Berg (2003) show that the mixed proportional hazard rate specifications are required. We are concerned that the proportionality assumption (i.e. observed and unobserved determinants affect the transition rates to recovery, the medical track and the labour track proportionally) may not be satisfied across individuals with different deferment periods and that this may be a source of bias. One might argue that those with long deferment periods will not start a claim in case of light diseases. As a result, those with a long deferment period may have longer sickness durations ceteris paribus, and are, therefore, a negatively selected subsample of the population of self-employed workers with a high deferment period. This may cause non-proportionality with respect to the unobserved determinants of recovery: one might expect v_r to be lower for those with a long deferment period. For that reason, in Subsection 5.3, we present a sensitivity analysis in which we estimate our model separately for two subsamples stratified by deferment

[4] Throughout the remainder of this article, "treatment effect" and "intervention effect" are used interchangeable.

period.

In case the mixed proportional hazard assumption is satisfied, the causal effects of participating in the intervention tracks can depend on the elapsed duration of the sickness spell t, the moment of entering the tracks t_m and t_l, and observed characteristics x.[5] As a benchmark specification, we choose homogenous and constant effects: $\delta_m(t|t_m, x) = \delta_{m,0}$ and $\delta_l(t|t_l, x) = \delta_{l,0}$. We refer to this model as the *constant effects model.*

In a second specification, we allow the effects of both tracks to depend on the elapsed sickness duration at the start of participation in the track. Thereby, we are able to distinguish between the impact of early, middle late and late interventions. Early interventions start in the first six weeks of the claim, middle late interventions in week 7 until week 13 and late interventions (i.e. the reference category) after 13 weeks. We refer to this specification as the *duration varying effects model*:

$$\begin{aligned} \delta_m(t|t_m, x) =& \; \delta_{m,0} + \delta_{m,t_m \leq 42} I(t_m \leq 42) + \delta_{m,43 \leq t_m \leq 91} I(43 \leq t_m \leq 91); \\ \delta_l(t|t_l, x) =& \; \delta_{l,0} + \delta_{l,t_l \leq 42} I(t_l \leq 42) + \delta_{l,43 \leq t_l \leq 91} I(43 \leq t_l \leq 91). \end{aligned} \tag{4}$$

Finally, in Subsection 5.2 we also present some analyses where we allow for (other types of) heterogeneous treatment effects. The *heterogenous effects model* is specified, in its application for heterogeneity by gender, as follows:

$$\begin{aligned} \delta_m(t|t_m, x) =& \; \delta_{m,0} + \delta_{m,\text{female}} I(\text{female}); \\ \delta_l(t|t_l, x) =& \; \delta_{l,0} + \delta_{l,\text{female}} I(\text{female}). \end{aligned} \tag{5}$$

4.3 Parameterisation and Estimation

As mentioned in Subsection 4.1, $\psi_r(\tau_0 + t)$ is modelled using dummies for each quarter. In addition, $\delta_m(t|t_m, x)$ and $\delta_l(t|t_l, x)$ are parameterised as outlined in the previous subsection. Finally, we need to parameterise the baseline hazards $\lambda_r(t)$, $\lambda_m(t)$ and $\lambda_l(t)$ and the joint distribution of the unobserved heterogeneity terms $G(v_r, v_m, v_l)$. For both functions, we adopt the most flexible specifications used to date.

For the baseline hazard function, we use a piecewise constant specification:

$$\begin{cases} \lambda_r(t) & = \alpha_s^r; \\ \lambda_m(t) & = \alpha_s^m; \text{ for } t \in [t_{s-1}, t_s), \\ \lambda_l(t) & = \alpha_s^l; \end{cases} \tag{6}$$

where s is an indicator of the duration interval. In particular, we fix the number of

[5]Richardson and van den Berg (2013) show that causal effects are even allowed to depend on unobserved characteristics v.

intervals to 8 with as cut-off points $t_0 = 0$, $t_1 = 10$, $t_2 = 20$, $t_3 = 40$, $t_4 = 70$, $t_5 = 100$, $t_6 = 140$, $t_7 = 190$ and $t_8 = +\infty$ (days).[6]

Concerning the joint distribution of the unobserved heterogeneity terms $G(v_r, v_m, v_l)$, we follow Heckman and Singer (1984) and assume this distribution to be discrete with unrestricted mass-point locations. In particular, we assume that (v_r, v_m, v_l) is randomly drawn from a discrete distribution with a finite and a priori unknown number K of points – in fact vectors of three points – of support.[7] The probabilities associated to these points of support are specified as logistic transforms:

$$p_k = \frac{\exp(q_k)}{\sum_{j=1}^{K} \exp(q_j)}; \quad \text{for } k = 1, \ldots, K. \tag{7}$$

We normalise q_0 to be equal to zero.

The log-likelihood described in equation (3) is maximised according to the procedure described in Gaure et al. (2007). In particular, we add points of support until the likelihood function does not show any improvement. Subsequently, we select the number of mass-points K that minimises the Akaike Information Criterion (AIC).

5 Results

In this section we first present and discuss the estimation results for our benchmark model in which we estimate homogenous and constant effects of participating in the medical and labour track. Next, we look into heterogeneity in the treatment effects by the timing of the interventions and by claim(ant) characteristics. In a third subsection, we discuss robustness tests for our main results. We end this section with a discussion of our findings. In the main text, we present the estimated treatment effects. Detailed estimation results and model selection statistics can be found in Appendix A.

5.1 Constant Effects Model

The AIC indicates that $K = 7$ is the optimal number of unobserved heterogeneity types.[8] Table 2 presents for the constant effects model the estimates for the key parameters of interest.

[6] We also estimate models with 4 and 13 intervals (see Subsection 5.3). Concerning the actual choice of time intervals, in the spirit of Craiu and Lee (2005), cut-off points were added one by one at the location where they produced the largest increase of the likelihood value after splitting.

[7] As a normality restriction, we do not include an intercept in x.

[8] Table 7 in Appendix A shows that results are very robust with respect to changing the number of unobserved heterogeneity types.

Table 2: Estimated Intervention Effects in the Constant Effects Model.

Medical track		
$\delta_{m,0}$	-0.364^{***}	(0.055)
Labour track		
$\delta_{l,0}$	-0.363^{***}	(0.097)
Duration dependence	yes	
Calendar time effects	yes	
Observed heterogeneity	yes	
Unobserved heterogeneity	yes	
N	15412	
Parameters	158	
Loglikelihood	-133056.758	

Standard errors in parentheses.

*** indicates significance at 1% level.

Detailed estimation results are in Table 6 in Appendix A.

Model selection statistics are in Table 7 in Appendix A.

The estimated treatment effects are highly significantly negative for both interventions, with a comparable magnitude. Recovery rates drop by about $(1 - \exp(-0.36) =)$ 30% when starting one of the tracks ceteris paribus. From the moment both tracks are started, the recovery rate drops by about 52%. So, the homogeneous effects model does not show any benefits of offering the interventions on recovery rates of self-employed workers. In Subsection 5.4 we discuss the interpretation of these results and we provide an explanation for the negative finding.

Before inspecting heterogeneity within and robustness of the estimated intervention effects, we briefly highlight some secondary results based on the other estimated parameters (see Table 6 in Appendix A). The intervention tracks are used more for expensive claims: Low deferment periods and high insured incomes predict earlier entry in the intervention tracks. A lower deferment period and a higher insured income also result in higher recovery rates. On the one hand, this finding is in line with our estimated intervention effects, as both point in the direction of no moral hazard. On the other hand, this finding supports the idea that due to underreporting of short sickness spells the individuals with a high deferment period observed in our data are a negatively selected subsample of the population of sick self-employed workers with high deferment periods (see Subsection 4.2). We come back to this issue in Subsection 5.3. Finally, the recovery rate is higher for younger claimants and claimants with tough occupations.

The calendar time effects show that the use of the medical track decreased over our observation period. This coincides with information from the private insurance provider about its policy. Next, although non-parametric Kaplan-Meier estimates (see Figures 1 to 3) indicate negative duration dependence of the modelled hazard rates, after controlling for observable and unobservable claim(ant) characteristics and quarter dummies, we observe positive duration dependence in all hazard rates. The longer the sickness duration, the more likely it is that a self-employed worker will recover (or get treated), which makes sense. The increase in recovery probability is most substantial during the first 30 days, and is likely related to reporting behaviour. The first sickness day is the day that the self-employed worker consults a physician, but the worker only has the obligation to inform the insurer somewhere during the deferment period. So if the self-employed expects to recover quickly, she/he can wait with reporting to the insurer. Short sickness spells are, therefore, especially for those with a substantial deferment period, likely not always reported in our data. On the other hand, the number of individuals entering the tracks is low early in the sickness spell. The caseworker may start the intervention tracks when recovery takes longer than expected.

Concerning the unobserved heterogeneity distribution, we observe that there are groups of individuals who never enter one of the two or even both tracks. In particular, there is a group which recovers quickly and never enters any track. In addition, we observe that those individuals with unobserved characteristics associated to the lowest recovery rates (the fourth heterogeneity type) also never enter the labour track. There is thus strong selectivity is the assignment of tracks.

Table 3: Estimated Intervention Effects in the Duration Varying Effects Model.

Medical track		
$\delta_{m,0}$	$-0.424{*}{*}{*}$	(0.051)
$\delta_{m,t_m \leq 42}$	0.046	(0.064)
$\delta_{m,43 \leq t_m \leq 91}$	0.026	(0.049)
Labour track		
$\delta_{l,0}$	$-0.330{*}{*}{*}$	(0.092)
$\delta_{l,t_l \leq 42}$	-0.038	(0.145)
$\delta_{l,43 \leq t_l \leq 91}$	$-0.162{*}$	(0.089)
Duration dependence	yes	
Calendar time effects	yes	
Observed heterogeneity	yes	
Unobserved heterogeneity	yes	
N	15412	
Parameters	162	
Loglikelihood	-133055.133	

Standard errors in parentheses.

***(*) indicates significance at 10%(1%) level.

Model selection statistics are in Table 8 in Appendix A.

5.2 Heterogeneous Treatment Effects

Table 3 presents the intervention effects for the duration varying effects model.[9] The adverse effects of both interventions are present for early, middle late and late interventions. In particular, for the medical track there is no significant heterogeneity in the intervention effect by its timing. There is, however, some weak evidence for a more adverse effect of the labour track when this track starts more than 13 weeks after the start of the sickness spell. The labour track intervention decreases recovery rates by about $(1 - \exp(-0.33) =)$ 28% if the intervention is started early in the spell of sickness absenteeism (within six weeks) and by about $(1 - \exp(-0.33 + -0.16) =)$ 39% if this intervention is started late (later than 13 weeks after the start of the spell).

Table 4: Estimated Intervention Effects in the Heterogeneous Effects Model.

	(1)	(2)	(3)	(4)
Medical track				
$\delta_{m,0}$	-0.415*** (0.048)	-0.338*** (0.061)	-0.414*** (0.052)	-0.323*** (0.052)
$\delta_{m,\,female\,gender}$	0.099 (0.076)			
$\delta_{m,\,tough\,occupation}$		-0.080 (0.055)		
$\delta_{m,\,ins.\ inc.\ >\ €500K}$			0.014 (0.051)	
$\delta_{m,\,def.\ per.\ <\ 14\ days}$				-0.216*** (0.050)
Labour track				
$\delta_{l,0}$	-0.348*** (0.091)	-0.244** (0.100)	-0.376*** (0.102)	-0.384*** (0.091)
$\delta_{l,\,female\,gender}$	-0.071 (0.105)			
$\delta_{l,\,tough\,occupation}$		-0.219*** (0.074)		
$\delta_{l,\,ins.\ inc.\ >\ €500K}$			0.042 (0.073)	
$\delta_{l,\,def.\ per.\ <\ 14\ days}$				-0.101 (0.074)
Duration dependence	yes	yes	yes	yes
Calendar time effects	yes	yes	yes	yes
Observed heterogeneity	yes	yes	yes	yes
Unobserved heterogeneity	yes	yes	yes	yes
N	15412	15412	15412	15412
Parameters	160	160	160	160
Loglikelihood	-133056.328	-133048.453	-133056.898	-133042.855

Standard errors in parentheses.

***(**) indicates significance at 10%(5%) level.

Model selection statistics are in Table 9, Table 10, Table 11 and Table 12 in Appendix A.

Next, we explore other dimensions of heterogeneity in the intervention effects. Table 4 presents the intervention effects for the related heterogeneous effects models. The adverse effects of both interventions are present for all subsamples by gender (column (1)), toughness of the occupation (column (2)), insured income (column (3)) and deferment period (column (4)). We find only evidence for two aspects of heterogeneity in the

[9]The estimates for the other parameters of this model, which are very comparable to those outlined in Table 6, are available on request. In addition, the model selection statistics for the duration varying effects model can be found in Table 8 in Appendix A.

intervention effects. First, the labour track is more adverse for claimants with tough occupations, which makes sense. Tough occupations may result in frequent (and minor) physical claims for which interventions are not effective. Second, the medical track is more adverse in case of smaller deferment periods. This finding is complementary to the idea that the observed individuals with a long deferment period are a negatively selected subsample of the population of self-employed workers with a long deferment period. For these individuals, medical treatment might be less adverse. We discuss this further in the next subsection.

5.3 Robustness Checks

In this subsection, we present additional analyses to test the robustness of our main results. In Subsection 4.2, we mentioned that the mixed proportional hazard assumption may fail across claimants with different deferment periods. Therefore, we re-estimate our benchmark model separately for two subsamples defined according to their deferment period. Table 5 (and Table 13 in Appendix A) indicate that the estimates of the intervention effects are not different for those with a short deferment period (shorter than 14 days) and those with a more substantial deferment period.[10] These estimates do not confirm the idea of a more adverse effect of the medical track for claimants with a short deferment period.

Table 5: Estimated Intervention Effects in the Constant Effects Model, Subsamples by Deferment Period.

	Deferment period < 14 days	Deferment period ≥ 14 days
Medical track		
$\delta_{m,0}$	-0.490*** (0.090)	-0.435** (0.182)
Labour track		
$\delta_{l,0}$	-0.266*** (0.069)	-0.367*** (0.106)
Duration dependence	yes	yes
Calendar time effects	yes	yes
Observed heterogeneity	yes	yes
Unobserved heterogeneity	yes	yes
N	5408	10004
Parameters	149	155
Loglikelihood	-45854.300	-87031.331

Standard errors in parentheses.

***(**) indicates significance at 10%(5%) level.

Model selection statistics are in Table 13 in Appendix A.

[10]Other subsamples by deferment period turned out to be too small to obtain robust estimates for our econometric model.

In further robustness checks, announced in Subsection 3.1, we re-estimated the constant effects model (i) without setting the starting time of the modelled durations to ten days after their start in the source data and (ii) without censoring the duration times after 548 days. In addition, as mentioned in Subsection 4.3 and to anticipate the critique that selection on unobservables might not be well identified as a consequence of overfitting, we estimated our benchmark model for four intervals in the baseline hazard function instead of eight. We also tested the robustness of our results after increasing the number of intervals to 13. However, these operations influenced the findings only negligibly.

5.4 Discussion of the Main Results

The estimation results show robust evidence for an adverse effects of interventions by medical doctors and occupational specialists with respect to reducing sick leave durations of self-employed workers. A first potential explanation for this finding is that moral hazard in sickness insurance is probably low among self-employed workers. Given their personal and inherent job characteristics self-employed workers have a strong interest in keeping their absence durations as short as possible, even shorter than optimal from a health point of view. Our secondary results with respect to the effect of benefit generosity (captured by insured income and deferment period) on recovery rates seem to confirm this hypothesis.

A second possible explanation is that medical doctors may not be effective in reducing moral hazard. This explanation is in agreement with Carlsen and Nyborg (2009), who show that due to information asymmetries, medical doctors might be unable to distinguish shirkers from certain groups of truly sick.

Furthermore, a principal-agent problem may exists between the insurance company and the medical doctors and occupational specialists. The insurance company does not provide any (financial) incentives to speed up recovery and the engaged medical doctors and occupational specialists may apply all measures for the benefit of the sick avoiding health risks on the patient's behalf. It is not unlikely that they advise even longer sick leave periods than necessary (Hartman et al., 2013). This might a fortiori be the case for the medical doctors as, notwithstanding their gatekeeping role, these physicians work under the Hippocratic Oath.

This focus on the well-being of the claimants brings us to a next possible explanation. While self-employed workers may aim to return to the work floor as soon as possible (taking into account only the short-term perspective), the medical doctors and occupational specialists may also take into account the long-term perspective (avoiding relapses). Thereby, participation in one of the tracks can slow self-employed workers down in their ambition to return to work and convince them about a more realistic trajectory. Unfortunately, we are not able to take this long-term perspective into account based on our data.

Finally, the other way round, it may be that the engaged medical doctors and occupational specialists simply maximise their profits by keeping the patients home for a longer period (yielding more paid visits).

6 Conclusion

In this study, we investigate the effectiveness of medical doctors (offering medical support) and occupational specialists (offering ergonomic advice and coaching) in reducing sick leave durations among self-employed workers. While the effectiveness of medical practitioners in reducing sick leave occurrence and sick leave duration has been studied by several researchers for (publicly insured) employees, this has not yet been investigated for (privately insured) self-employed workers. We exploit unique administrative data from the major Dutch private insurance company. From these data we use all sickness benefit applications with a physical condition by self-employed workers between January 2009 and March 2014. We estimate a multivariate duration model dealing with the non-random and dynamic selection into the intervention tracks engaging medical doctors ("medical track") and occupational specialists ("labour track") by controlling on observable and unobservable intervention determinants.

We find adverse treatment effects for both the medical and labour track interventions, which are robust against various sensitivity checks. Moreover, these treatment effects are very similar in magnitude for both interventions. After starting a track, recovery rates drop by about 30%. The negative effects of both interventions are present for early, middle late and late interventions. In addition, they are present for all tested subsamples by gender, toughness of the occupation, insured income and deferment period. Interestingly, the labour track intervention is more adverse for claimants with tough occupations.

Finally, we provide several potential explanations for our main finding. First, we hypothesise that moral hazard in sickness insurance is low among self-employed workers given their personal and job characteristics. The finding that claimants with more generous sickness benefits (due to a lower deferment period and/or a higher insured income) do not recover faster is consistent with this hypothesis. In addition, the engaged medical doctors and occupational specialists may be ineffective in reducing sick leave durations as they (i) might be unable to distinguish shirkers from truly sick, (ii) might give priority to healing over gatekeeping or (iii) might just maximise their number of visits. Lastly, medical doctors and occupational specialists may be more focussed on long-term health than the self-employed workers, who may be more interested in restarting work as soon as possible. To what extent the adverse short-term effect of their interventions are compensated by potentially beneficial long-term effects with respect to the productivity of the self-employed workers, seems a fruitful direction for future research.

References

Abbring, J.H. and G.J. van den Berg (2003), The Nonparametric Identification of Treatment Effects in Duration Models, *Econometrica*, 71, 1491–1517.

Barmby, T.A., C. Orme and J.G. Treble (1991), Worker Absenteeism: An Analysis Using Microdata, *Economic Journal*, 101, 214–229.

Blanchflower, D. and A. Oswald (1998), What Makes an Entrepreneur?, *Journal of Labor Economics*, 16, 26–60.

Carlsen, B. and K. Nyborg (2009), The Gate is Open: Primary Care Physicians as Social Security Gatekeepers, *University of Oslo Department of Economics Memorandum Papers*, 13/2009.

Craiu, R.V. and C.M. Lee (2005), Selection for the Competing-Risks Model With and Without Masking, *Technometrics*, 47, 457–467.

De Jong, P. (2012), Recent changes in Dutch disability policy, *APE The Hague Working Papers*, unnumbered.

De Jong, P., M. Lindeboom and B. van der Klaauw (2011), Screening Disability Insurance Applications, *Journal of the European Economic Association*, 9, 106–129.

De Paola, M., V. Scoppa and V. Pupo (2014), Absenteeism in the Italian Public Sector: The Effects of Changes in Sick Leave Policy, *Journal of Labor Economics*, 32, 337–360.

Fevang, E., S. Markussen, and Røed (2014), The sick pay trap, *Journal of Labor Economics*, 32, 305–336.

Gautier, P.A. and B. van der Klaauw (2012), Labor market policy and participation over the life cycle, *Netspar Panel Papers*, 26.

Gaure, S., K. Røed and T. Zhang (2007), Time and causality: A Monte Carlo assessment of the timing-of-events approach, *Journal of Econometrics*, 141, 1159–1195.

Hartman, L., P. Hesselius and P. Johansson (2013), Effects of eligibility screening in the sickness insurance: Evidence from a field experiment, *Labour Economics*, 20, 48–56.

Heckman, J.J. and B. Singer (1984), A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data, *Econometrica*, 52, 271–320.

Henrekson, M. and M. Persson (2004), The Effects on Sick Leave of Changes in the Sickness Insurance System, *Journal of Labor Economics*, 22, 87–113.

Hyytinen, A. and O.-P. Ruuskanen (2007), Time Use of the Self-Employed, *Kyklos*, 60, 105–122.

Johansson, P. and M. Palme (2005), Assessing the effect of public policy on worker absenteeism, *Journal of Human Resources*, 37, 381–409.

Johansson, P. and M. Palme (2005), Moral hazard and sickness insurance, *Journal of Public Economics*, 89, 1879–1890.

Koning, P. (2004), Estimating the impact of experiencing rating on the inflow into disability insurance in the Netherlands, *CPB Discussion Papers*, 37.

Lechman, D.S.J. and C. Schnabel (2014), Absence from Work of the Self-Employed: A Comparison with Paid Employees, *Kyklos*, 67, 368–390.

Markussen, S. (2010), How physicians can reduce sick leave – evidence from a natural experiment, *University of Oslo Department of Economics Memorandum Papers*, 19/2009.

Markussen, S., K. Røed, O.J. Røgeberg and S. Gaure (2011), The anatomy of absenteeism, *Journal of Health Economics*, 30, 277–292.

Parker, S. (2004), *The Economics of Self-Employment and Entrepreneurship*, Cambridge University Press, Cambridge.

Parasuraman, S. and C.A. Simmers (2001), Type of employment, work-family conflict and well-being: a comparative study, *Journal of Occupational Behaviour*, 22, 551–568.

Puhani, P.A. and K. Sonderhof (2010), The effects of a sick pay reform on absence and on health-related outcomes, *Journal of Health Economics*, 29, 285–302.

Richardson, K. and G.J. van den Berg (2013), Duration dependence versus unobserved heterogeneity in treatment effects: Swedish labor market training and the transition rate to employment, *Journal of Applied Econometrics*, 28, 325–351.

Spierdijk, L., G. van Lomwel and W. Peppelman (2009), The determinants of sick leave durations of Dutch self-employed, *Journal of Health Economics*, 28, 1185–1196.

Van den Berg, G.J. (2001), Duration models: Specification, identification, and multiple durations, in: Heckman J.J. and E. Leamer (eds) *Handbook of Econometrics*, 5, North Holland, Amsterdam.

Ziebarth, N.R. (2013), Long-term absenteeism and moral hazard. Evidence from a natural experiment, *Labour Economics*, 24, 277–292.

Ziebarth, N.R. and M. Karlsson (2010), A natural experiment on sick pay cuts, sickness absence, and labor costs, *Journal of Public Economics*, 94, 1108–1122.

Ziebarth, N.R. and M. Karlsson (2014), The Effects of Expanding the Generosity of the Statutory Sickness Insurance System, *Journal of Applied Econometrics*, 29, 208–230.

Appendix A: Additional Tables

Table 6: Estimation Results for the Constant Effects Model.

	θ_r		θ_m		θ_l	
Medical track						
$\delta_{m,0}$	−0.364∗∗∗	(0.055)				
Labour track						
$\delta_{l,0}$	−0.363∗∗∗	(0.097)				
Age						
< 36	0.235∗∗∗	(0.040)	−0.037	(0.055)	−0.365∗∗∗	(0.094)
36 – 45 (ref.)						
46 – 55	−0.017	(0.035)	−0.020	(0.044)	−0.025	(0.072)
> 55	0.001	(0.056)	−0.169∗∗	(0.071)	−0.282∗∗	(0.120)
Gender						
Female	−0.051	(0.038)	−0.158∗∗∗	(0.050)	−0.070	(0.080)
Region						
North	−0.029	(0.027)	−0.077∗∗	(0.034)	−0.248∗∗∗	(0.057)
South (ref.)						
Center	−0.037	(0.038)	−0.327∗∗∗	(0.050)	−0.309∗∗∗	(0.083)
Occupation type						
Agricultural	0.004	(0.028)	−0.109∗∗∗	(0.035)	−0.135∗∗	(0.059)
SME (ref.)						
Liberal profession	0.231∗∗∗	(0.048)	−0.439∗∗∗	(0.064)	−0.254∗∗	(0.098)
Toughness of occupation						
(Rather) light	−0.154∗∗∗	(0.042)	−0.087	(0.054)	0.485∗∗∗	(0.087)
Rather tough	−0.075∗∗	(0.036)	−0.109∗∗	(0.047)	0.226∗∗∗	(0.075)
Tough (ref.)						
Insured income						
< €100M	−0.396∗∗∗	(0.044)	−0.315∗∗∗	(0.054)	−0.659∗∗∗	(0.095)
€100M – €500M (ref.)						
€500M – €1000M	0.147∗∗∗	(0.036)	−0.014	(0.045)	−0.013	(0.074)
> €1000M	0.284∗∗∗	(0.047)	−0.074	(0.063)	0.142	(0.101)
Deferment period						
< 14 days	0.310∗∗∗	(0.030)	0.257∗∗∗	(0.037)	−0.032	(0.065)
14 days – 3 months (ref.)						
3 months – 1 year	−0.505∗∗∗	(0.046)	−0.371∗∗∗	(0.054)	0.113	(0.085)
> 1 year	−0.947∗∗∗	(0.063)	−1.205∗∗∗	(0.075)	−0.465∗∗∗	(0.113)
Quarter						
01/2009 – 03/2009 (ref.)						
04/2009 – 06/2009	−0.090	(0.089)	−0.094	(0.095)	0.123	(0.288)
07/2009 – 09/2009	−0.118	(0.089)	−0.075	(0.097)	0.164	(0.288)
10/2009 – 12/2009	−0.213∗∗	(0.089)	−0.206∗∗	(0.100)	0.046	(0.287)
01/2010 – 03/2010	−0.106	(0.087)	−0.310∗∗∗	(0.100)	0.099	(0.286)
04/2010 – 06/2010	−0.131	(0.089)	−0.247∗∗	(0.099)	−0.007	(0.287)
07/2010 – 09/2010	−0.191∗∗	(0.090)	−0.134	(0.100)	−0.045	(0.287)
10/2010 – 12/2010	−0.361∗∗∗	(0.090)	−0.594∗∗∗	(0.101)	−0.262	(0.289)
01/2011 – 03/2011	−0.197∗∗	(0.088)	−0.732∗∗∗	(0.101)	−0.209	(0.287)
04/2011 – 06/2011	−0.202∗∗	(0.089)	−0.691∗∗∗	(0.102)	−0.103	(0.286)
07/2011 – 09/2011	−0.508∗∗∗	(0.093)	−1.508∗∗∗	(0.116)	−0.083	(0.286)
10/2011 – 12/2011	−0.506∗∗∗	(0.092)	−1.672∗∗∗	(0.113)	0.062	(0.284)
01/2012 – 03/2012	−0.303∗∗∗	(0.090)	−1.769∗∗∗	(0.113)	−0.216	(0.288)

04/2012 – 06/2012	−0.246*** (0.091)	−1.762*** (0.113)	−0.253 (0.289)
07/2012 – 09/2012	−0.284*** (0.092)	−2.125*** (0.122)	−0.240 (0.289)
10/2012 – 12/2012	−0.405*** (0.093)	−1.811*** (0.114)	−0.239 (0.290)
01/2013 – 03/2013	−0.389*** (0.093)	−1.824*** (0.114)	−0.234 (0.290)
04/2013 – 06/2013	−0.235** (0.093)	−1.924*** (0.117)	−0.223 (0.292)
07/2013 – 09/2013	−0.356*** (0.096)	−1.821*** (0.118)	−0.188 (0.292)
10/2013 – 12/2013	−0.387*** (0.096)	−1.555*** (0.113)	0.107 (0.288)
01/2014 – 03/2014	−0.160* (0.096)	−1.591*** (0.118)	0.266 (0.290)
Duration dependence			
t=[1,10] (ref.)			
t=[11,20]	0.660*** (0.066)	0.685*** (0.093)	1.044*** (0.292)
t=[21,30]	1.168*** (0.066)	1.218*** (0.089)	1.662*** (0.276)
t=[31,50]	1.417*** (0.072)	1.722*** (0.084)	1.945*** (0.264)
t=[51,110]	1.472*** (0.088)	2.158*** (0.088)	2.561*** (0.265)
t=[111,190]	1.436*** (0.104)	2.517*** (0.098)	3.330*** (0.279)
t=[191,330]	1.402*** (0.116)	2.665*** (0.115)	3.746*** (0.288)
t > 330	1.047*** (0.131)	2.803*** (0.145)	3.966*** (0.306)
Unobserved heterogeneity distribution[11]			
v_1	−5.761*** (0.186)	−6.014*** (0.157)	−10.079*** (0.505)
v_2	−7.453*** (0.214)	−7.621*** (0.188)	−10.704*** (0.448)
v_3	−6.559*** (0.195)	−4.047*** (0.167)	−6.677*** (0.397)
v_4	−10.347* (5.591)	−5.074*** (0.211)	$-\infty$
v_5	−6.905*** (0.228)	−5.593*** (0.138)	−8.783*** (0.415)
v_6	−6.405*** (0.301)	$-\infty$	−5.203*** (0.469)
v_7	−4.652*** (0.422)	$-\infty$	$-\infty$
q_2		−2.280*** (0.202)	
q_3		−2.348*** (0.185)	
q_4		−3.439*** (0.608)	
q_5		−0.615** (0.288)	
q_6		−4.685*** (0.376)	
q_7		−1.926** (0.952)	
N		15412	
Parameters		158	
Loglikelihood		-133056.758	
AIC		266429.516	

Standard errors in parentheses.

***(**)((*)) indicates significance at the 1%(5%)((10%)) level.

[11] Some heterogeneity parameters were estimated as a large negative number causing a 0 probability with respect to being selected into the medical track. This is numerically problematic. When we faced this problem, in the spirit of Gaure et al. (2007), we marked the offending parameter as “negative infinity” and kept it out of further estimation.

Table 7: Model Selection: Constant Effects Model.

K	N	Parameters	Loglikelihood	AIC	$\delta_{m,0}$	$\delta_{l,0}$
1	15412	134	-133420.113	267108.227	-0.452*** (0.026)	-0.454*** (0.036)
2	15412	138	-133235.125	266746.250	-0.611*** (0.033)	-0.595*** (0.042)
3	15412	142	-133133.179	266550.359	-0.585*** (0.033)	-0.509*** (0.047)
4	15412	146	-133124.921	266541.841	-0.464*** (0.051)	-0.486*** (0.054)
5	15412	150	-133096.373	266492.746	-0.375*** (0.056)	-0.295*** (0.099)
6	15412	154	-133076.500	266461.001	-0.379*** (0.055)	-0.398*** (0.096)
7	15412	158	-133056.758	266429.516	-0.364*** (0.055)	-0.363*** (0.097)
8	15412	162	-133055.262	266434.525	-0.361*** (0.055)	-0.356*** (0.097)

AIC: Akaike Information Criterion.

*** indicates significance at the 1% level.

Table 8: Model Selection: Duration Varying Effects Model.

K	N	Parameters	Loglikelihood	AIC	$\delta_{m,0}$	$\delta_{m,t_m\leq 42}$	$\delta_{m,43\leq t_m\leq 91}$	$\delta_{l,0}$	$\delta_{l,t_l\leq 42}$	$\delta_{l,43\leq t_l\leq 91}$
1	15412	138	-133415.960	267107.919	-0.437*** (0.036)	-0.036 (0.046)	-0.017 (0.039)	-0.395*** (0.043)	-0.075 (0.111)	-0.197*** (0.075)
2	15412	142	-133231.029	266746.057	-0.575*** (0.041)	-0.074 (0.047)	-0.049 (0.040)	-0.563*** (0.050)	-0.040 (0.111)	-0.167** (0.075)
3	15412	146	-133130.214	266552.428	-0.578*** (0.040)	-0.038 (0.051)	-0.022 (0.040)	-0.483*** (0.051)	-0.053 (0.113)	-0.179** (0.078)
4	15412	150	-133122.752	266545.503	-0.468*** (0.054)	-0.010 (0.052)	-0.011 (0.040)	-0.468*** (0.057)	0.015 (0.114)	-0.154* (0.080)
5	15412	154	-133093.533	266495.066	-0.386*** (0.057)	0.061 (0.069)	0.021 (0.053)	-0.271*** (0.099)	-0.033 (0.139)	-0.192** (0.094)
6	15412	158	-133073.495	266462.990	-0.397*** (0.056)	0.108 (0.070)	0.049 (0.052)	-0.367*** (0.097)	0.033 (0.133)	-0.144 (0.092)
7	15412	162	-133055.133	266434.265	-0.424*** (0.051)	0.046 (0.064)	0.026 (0.049)	-0.330*** (0.092)	-0.038 (0.145)	-0.162* (0.089)
8	15412	166	-133051.980	266435.961	-0.380*** (0.057)	0.105 (0.069)	0.050 (0.051)	-0.324*** (0.098)	-0.035 (0.150)	-0.169* (0.094)

AIC: Akaike Information Criterion.

***(**)((*)) indicates significance at the 1%(5%)((10%)) level.

Table 9: Model Selection: Gender Varying Effects Model.

K	N	Parameters	Loglikelihood	AIC	$\delta_{m,0}$	$\delta_{m,female\,gender}$	$\delta_{l,0}$	$\delta_{l,female\,gender}$
1	15412	136	-133419.538	267111.076	-0.460*** (0.027)	0.069 (0.065)	-0.449*** (0.038)	-0.039 (0.097)
2	15412	140	-133234.539	266749.077	-0.619*** (0.034)	0.070 (0.066)	-0.590*** (0.044)	-0.050 (0.097)
3	15412	144	-133132.626	266553.252	-0.592*** (0.034)	0.067 (0.065)	-0.502*** (0.049)	-0.051 (0.065)
4	15412	148	-133124.042	266544.083	-0.458*** (0.052)	0.058 (0.067)	-0.469*** (0.056)	-0.038 (0.098)
5	15412	152	-133095.905	266495.810	-0.382*** (0.057)	0.072 (0.077)	-0.295*** (0.101)	-0.064 (0.110)
6	15412	156	-133075.559	266463.117	-0.390*** (0.056)	0.101 (0.078)	-0.389*** (0.098)	-0.085 (0.107)
7	15412	160	-133056.328	266432.656	-0.415*** (0.048)	0.099 (0.076)	-0.348*** (0.091)	-0.071 (0.105)
8	15412	164	-133054.397	266436.795	-0.373*** (0.056)	0.099 (0.078)	-0.345*** (0.098)	-0.076 (0.106)

AIC: Akaike Information Criterion.

*** indicates significance at the 1% level.

Table 10: Model Selection: Toughness of Occupation Varying Effects Model.

K	N	Parameters	Loglikelihood	AIC	$\delta_{m,0}$	$\delta_{m,tough\,occupation}$	$\delta_{l,0}$	$\delta_{l,tough\,occupation}$
1	15412	136	-133410.575	267093.150	-0.361*** (0.043)	-0.127*** (0.048)	-0.359*** (0.056)	-0.158** (0.069)
2	15412	140	-133226.034	266732.069	-0.515*** (0.048)	-0.129*** (0.048)	-0.504*** (0.061)	-0.149** (0.069)
3	15412	144	-133124.363	266536.725	-0.492*** (0.048)	-0.128*** (0.048)	-0.425*** (0.063)	-0.144** (0.069)
4	15412	148	-133115.726	266527.452	-0.376*** (0.061)	-0.118** (0.049)	-0.383*** (0.070)	-0.173** (0.069)
5	15412	152	-133089.775	266483.551	-0.327*** (0.067)	-0.071 (0.057)	-0.132 (0.109)	-0.206*** (0.078)
6	15412	156	-133069.069	266450.139	-0.323*** (0.068)	-0.077 (0.057)	-0.278** (0.107)	-0.209*** (0.077)
7	15412	160	-133048.453	266416.906	-0.338*** (0.061)	-0.080 (0.055)	-0.244** (0.100)	-0.219*** (0.074)
8	15412	164	-133047.125	266422.251	-0.304*** (0.067)	-0.076 (0.057)	-0.238** (0.105)	-0.221*** (0.075)

AIC: Akaike Information Criterion.

***(**) indicates significance at the 1%(5%) level.

Table 11: Model Selection: Insured Income Varying Effects Model.

K	N	Parameters	Loglikelihood	AIC	$\delta_{m,0}$	$\delta_{m,\,insured\ income\ >\ €500K}$	$\delta_{l,0}$	$\delta_{l,\,insured\ income\ >\ €500K}$
1	15412	136	-133419.377	267110.754	-0.429*** (0.033)	-0.045 (0.042)	-0.483*** (0.050)	0.056 (0.066)
2	15412	140	-133234.861	266749.722	-0.601*** (0.040)	-0.017 (0.043)	-0.618*** (0.055)	0.046 (0.067)
3	15412	144	-133132.886	266553.773	-0.577*** (0.039)	-0.016 (0.042)	-0.535*** (0.059)	0.049 (0.066)
4	15412	148	-133124.324	266544.648	-0.437*** (0.057)	-0.044 (0.045)	-0.507*** (0.068)	0.045 (0.068)
5	15412	152	-133095.332	266494.664	-0.379*** (0.063)	0.009 (0.053)	-0.363*** (0.110)	0.098 (0.078)
6	15412	156	-133075.738	266463.477	-0.379*** (0.062)	-0.001 (0.052)	-0.443*** (0.107)	0.087 (0.077)
7	15412	160	-133056.898	266433.796	-0.414*** (0.052)	0.014 (0.051)	-0.376*** (0.102)	0.042 (0.073)
8	15412	164	-133054.754	266437.507	-0.360*** (0.062)	-0.005 (0.052)	-0.393*** (0.108)	0.075 (0.075)

AIC: Akaike Information Criterion.

*** indicates significance at the 1% level.

Table 12: Model Selection: Deferment Period Varying Effects Model.

K	N	Parameters	Loglikelihood	AIC	$\delta_{m,0}$	$\delta_{m,\,deferment\ period\ <\ 14\ days}$	$\delta_{l,0}$	$\delta_{l,\,deferment\ period\ <\ 14\ days}$
1	15412	136	-133401.571	267075.142	-0.356*** (0.031)	-0.243*** (0.044)	-0.449*** (0.043)	-0.024 (0.070)
2	15412	140	-133214.673	266709.346	-0.509*** (0.037)	-0.260*** (0.044)	-0.587*** (0.049)	-0.018 (0.070)
3	15412	144	-133112.224	266512.447	-0.484*** (0.037)	-0.264*** (0.044)	-0.513*** (0.053)	-0.013 (0.070)
4	15412	148	-133100.236	266496.472	-0.337*** (0.053)	-0.231*** (0.050)	-0.213** (0.090)	-0.068 (0.076)
5	15412	152	-133083.543	266471.086	-0.308*** (0.059)	-0.233*** (0.051)	-0.244** (0.101)	-0.055 (0.078)
6	15412	156	-133063.459	266438.917	-0.297*** (0.059)	-0.227*** (0.052)	-0.413*** (0.099)	-0.061 (0.077)
7	15412	160	-133042.855	266405.710	-0.323*** (0.052)	-0.216*** (0.050)	-0.384*** (0.091)	-0.101 (0.074)
8	15412	164	-133041.455	266410.910	-0.279*** (0.058)	-0.222*** (0.051)	-0.379*** (0.097)	-0.092 (0.075)

AIC: Akaike Information Criterion.

***(**) indicates significance at the 1%(5%) level.

Table 13: Model Selection: Constant Effects Model, Subsamples by Deferment Period.

K	N	Parameters	Loglikelihood	AIC	$\delta_{m,0}$	$\delta_{l,0}$
Subsample with deferment period < 14 days						
1	5408	125	-45947.692	92145.384	-0.542*** (0.043)	-0.395*** (0.062)
2	5408	129	-45904.490	92066.980	-0.552*** (0.051)	-0.417*** (0.087)
3	5408	133	-45875.688	92017.377	-0.669*** (0.055)	-0.506*** (0.083)
4	5408	137	-45870.735	92015.470	-0.525*** (0.087)	-0.609*** (0.089)
5	5408	141	-45870.040	92022.079	-0.564*** (0.091)	-0.280 (0.174)
6	5408	145	-45865.787	92021.574	-0.520*** (0.092)	-0.484*** (0.173)
7	5408	149	-45854.300	92006.600	-0.490*** (0.090)	-0.435** (0.182)
8	5408	153	-45851.200	92008.401	-0.473*** (0.090)	-0.440** (0.184)
Subsample with deferment period ≥ 14 days						
1	10004	131	-87269.343	174800.687	-0.389*** (0.033)	-0.489*** (0.044)
2	10004	135	-87142.162	174554.324	-0.551*** (0.043)	-0.747*** (0.062)
3	10004	139	-87080.897	174439.795	-0.510*** (0.042)	-0.526*** (0.061)
4	10004	143	-87066.575	174419.150	-0.328*** (0.057)	-0.154* (0.089)
5	10004	147	-87056.395	174406.791	-0.311*** (0.067)	-0.160 (0.098)
6	10004	151	-87042.181	174386.361	-0.279*** (0.069)	-0.387*** (0.107)
7	10004	155	-87031.331	174372.662	-0.266*** (0.069)	-0.367*** (0.106)
8	10004	159	-87030.629	174379.259	-0.263*** (0.069)	-0.367*** (0.106)

AIC: Akaike Information Criterion.

***(**)((*)) indicates significance at the 1%(5%)((10%)) level.