De Economist (2011) 159:361–380
DOI 10.1007/s10645-011-9169-2

Return to Work After Long Term Sickness

The Role of Employer Based Interventions

T. P. Everhardt · Ph. R. de Jong

Published online: 29 May 2011

Abstract In this paper the impact of vocational interventions on the duration until reemployment of long term sick workers is studied, taking account of self-perceived health and other personal, health and vocational characteristics. We use a longitudinal dataset containing three waves of interviews with employees who were still on the sickness rolls 9 months after calling in sick. We compare a semi-parametric Cox model with a Weibull model to allow for unobserved heterogeneity. The results from these models are similar. They point to the importance of the set of accommodation and rehabilitation activities that employers and occupational health agencies employ and of a timely start of such activities. Their effectiveness shows that the series of reforms in sickness and disability schemes that were introduced from 2002 onwards have been quite successful.

Keywords Return to Work · Sickness · Interventions · Duration analysis · Cox · Weibull · Unobserved heterogeneity · Interval censoring

JEL Classification C14 · C33 · C34 · C35 · C41 · J64

1 Introduction

This paper investigates the effect of activities that employers and external providers undertake to help long-term sick employees to resume their work. Dutch employers

T. P. Everhardt
Aarts de Jong Wilms Goudriaan Public Economics (APE), Amsterdam, The Netherlands
e-mail: t.everhardt@ape.nl

Ph. R. de Jong (✉)
University of Amsterdam, Aarts de Jong Wilms Goudriaan Public Economics (APE), Amsterdam, The Netherlands
e-mail: ph.dejong@ape.nl

cannot dissolve a labor contract until two years after the first day of sickness and are subject to fierce, financial and administrative, incentives to reduce the risk of long-term sickness and disability benefit dependency. Dutch firms, therefore, have a strong interest in using effective return-to-work (rtw) activities, much stronger than firms in other countries do. A unique cohort study enables us to assess the extent to which employers are successful in helping their long-term (9 months) sick employees return to work. The purpose of this survey was to track how many employees resume their old or other work, what their personal, medical and vocational characteristics are and whether they were helped to return to work. We describe the setup of the survey and the data in Sect. 4. Section 2 provides an overview of the recent changes in the Dutch sickness and disability benefit schemes. Section 3 scans the literature: a relatively small number of studies by economists dwell on the efficacy of vocational interventions for sick workers. Such studies are more the realm of occupational medicine and social science. In Sect. 5 we specify two alternative duration models that take account of common limitations of the data such as unobserved heterogeneity and incomplete observations. Section 6 presents the results for both models to test the robustness of the outcomes. Section 7 concludes.

2 The Dutch Sickness and Disability Policy Reform

According to calculations by the OECD the entry rate into the Dutch disability system was 1.2 % of the working age (20–64) population in 2,000. This was the one to highest rate among 25 countries. In 2008 the Dutch rate had dropped to 0.38 and was below OECD average (OECD 2010). This remarkable drop was induced by a series of policy changes. These reforms culminated in the establishment of a new disability insurance scheme (WIA) in 2004 that replaced the WAO scheme that had been in place since 1967. At their heart were changes that increased the incentives of both employees and their employers to invest more time and effort in accommodation and rehabilitation following the onset of a disability (Burkhauser et al. 2008).

Foremost among the reforms that caused this drop was the extension from 1 to 2 years of the mandate that firms (including small employers) bear full responsibility for sick pay. Workers are considered for eligibility for the longer-term disability benefit system only after two years of sickness benefits. During these first 2 years, employers must allow workers receiving sickness benefits to continue with the firm. Employers can only dismiss employees who refuse to cooperate in a reasonable work-resumption plan. Firms have a set of prescribed rehabilitation and accommodation activities that they (via a contracted private occupational health agency) must provide to try either to retain disabled employees or to find alternative employment for them during those 2 years.

The so-called Gatekeeper Protocol structures these activities. After a maximum of 6 weeks of sickness an occupational physician has to make a Problem Analysis—an assessment of medical cause, functional limitations and prognosis regarding work resumption. On the basis of this assessment the employer and sick employee together draft a return-to-work plan (rtw-plan) in which they specify an aim (resumption of current/other job under current/accommodated conditions) and the steps needed to

reach that aim. They appoint a case-manager, and fix dates at which the plan should be evaluated, and modified if necessary. The rtw-plan should be ready in the eighth week of sickness. It is binding for both parties, and one party may summon the other when considered negligent.

Disability insurance benefit claims are only admissible if they are accompanied by a report containing an assessment as to why the plan has not resulted in work resumption. The claim is not processed by the National Social Insurance Institute (NSII), the public organization responsible for adjudicating disability benefit claims, if the report is delayed or incomplete or if it is clear that the rehabilitation efforts were insufficient. Depending on the seriousness of the negligence, the NSII can return the reintegration report and give the employer the opportunity to complete it, or the NSII can start a sanction procedure against the employer. In 2007, nearly 11 % of disability insurance claims were returned to employers and the employer continued to be responsible for employing the worker until the employee had returned to work, or sufficient rtw activities were tried the benefit claim was considered admissible (Kenniscentrum UWV 2010).

Determination of disability insurance benefits was also dramatically reformed. The reforms split the previous all-encompassing disability benefits scheme into two separate programs. The first provides benefits to those judged to have an unrecoverable loss of earnings capacity of at least 80 %. These individuals are eligible for full and permanent disability benefits replacing 75 % of gross earnings (with a cap on covered earnings of €49,300 per year in 2011). The second provides benefits to those judged to have a loss of earnings capacity between 35 and 80 %, and those that are fully disabled when examined but with the prospect to recover part or all of their capacity. These individuals are eligible for partial benefits or temporary full benefits. Partial and temporary full beneficiaries can receive up to 70 % of gross earnings, but the percentage varies depending on actual work status and significant incentives have been built in to the program to encourage beneficiaries to work to their estimated earnings capacity.

Employers pay to fund the publicly run partial and temporary full disability program, but they can opt out of it by enrolling their workers with a private insurer instead. Either way, employers pay experience-rated premiums that cover the first 10 years of partial disability benefit receipt. After 10 years, benefits are covered by the uniform pay-as-you-go rates that also cover the fully and permanently disabled and the stock of current beneficiaries under the old system. This unprecedented set of incentives increases the potential gains for firms of providing effective accommodation and rehabilitation.

3 Literature Review

Compared to the literature on unemployment duration the body of econometric literature on the determinants of work resumption after prolonged sickness is limited but growing.[1] The set of incentives to which Dutch firms are subject have

[1] There is vast amount of studies on the effects of medical interventions and vocational rehabilitation on the probability of return to work in the realm of occupational medicine. See, for instance, Franche et al.

inspired a number of studies on financial and other measures taken to reduce the use of sickness and disability benefits. Koning (2004) studies the effects of experience rating in financing disability benefits using a difference-in difference approach. The impact of experience rating proves to be substantial, amounting to a 15% reduction of the inflow into the Disability Insurance (DI) scheme. Employers seem to have been triggered to increase preventive activities, once they experienced increases in DI premium rates (ex post incentives). De Jong et al. (2011) also use difference-in difference techniques to investigate the effects of stricter screening of DI benefit claims. It is based on an experiment where some district offices of the NSII screened disability claims much stricter than elsewhere. Strictness appears to mainly reduce the number of claims, rather than the number of awards. Also here preventive effects turn out to be more important than the punitive effect of increased strictness.

Neither of these studies, however, identifies the nature of those preventive measures or their effect on the rtw hazard. Two more Dutch cohort studies investigate the role of interventions that are legally mandated under the Gatekeeper Protocol. Using logistic regression Corten (2009) finds that the frequency of contacts with an occupational doctor influences the probability of return to work. He uses a sample of 889 construction workers who had been sick for at least 3 months, and who were surveyed 2 years later on their work status and which return-to-work activities they had undertaken. Oomens et al. (2010) use a sample of 1,091 employees who had been sick for at least 3 months. They were surveyed at 19 and at 56 weeks after their first day of sickness. They also apply logistic regression to analyze the probability of being fully back at work at 56 weeks. Taking each of the first two steps of the Gatekeeper Protocol in time (making a problem analysis within the first 6 weeks of sickness and setting up a rtw plan before lapse of the eighth week) doubles the probability of having resumed ones work 9 months later. Despite the fact that the Gatekeeper Protocol mandates these steps to be taken before the sixth and eighth week of sickness about 30% report that the problem analysis was lacking and 46% state that there is no rtw plan.

Frölich et al. (2004) study the effects of participation in rehabilitation programs for long-term sick in Sweden. They focus on different sorts of interventions, like passive, workplace, educational, medical and social ones. Workplace interventions are found to be superior to the other ones. However, none of the rehabilitation measures in the program outperform non-participation when it comes to the rtw process. Their research is based on 6,287 employees who were sick for at least 60 consecutive days. By non-parametric multidimensional propensity score matching they sort their sample into homogeneous groups, and then use multinomial probit analysis to estimate the effects of interventions on rtw. The analysis is multinomial because they define rtw conditional on the type of intervention that was used.

Two final papers use duration models that are comparable to our approach. Spierdijk et al. (2009) analyze the determinants of sick leave durations among self-employed in the Netherlands who take out private sickness and disability insurance. They find that insurer-based case management improves the recovery rate of self-employed with a sick spell up to 1 year but it does not positively affect the recovery rate of those trapped

(2005) and Krause et al. (1998). Another source is the Workers' Compensation literature. An example is Butler et al. (1995).

in longer-term sickness. The sample used consists of 1,248 sickness claims starting in a period between 2000 and 2005. The data stems from an insurance company. They employ a mixed proportional hazard model with a gamma frailty distribution to take account of unobserved heterogeneity. To check the robustness of their estimates they also use accelerated failure time and proportional hazard Weibull models with gamma and heckman-singer frailty. In its econometric approach their study is similar to ours.

Høgelund et al. (2010) investigate the effectiveness of a Danish national program that allows sick-listed workers to gradually return to work. They find a significant and positive effect of this program on the probability of returning to pre-sickness working hours, using a simultaneous discrete mixed proportional hazard model of program participation and of returning to regular working hours, thus taking account of selectivity in program participation. Their analysis is based on a sample of 934 workers under age 60.

4 Description of the Data and Definition of the Explanatory Variables

The data used in this paper stems from a longitudinal study among a group of 3,736 persons who reported sick around January 1, 2007 and had not (fully) returned to work 9 months later. Three waves of interviews, at 10, 18 and 27 months after reporting sick were held giving participants a choice between web or paper surveys. Table 1 shows the sample sizes across the three waves and also documents the cumulative number of work resumptions. Notice that we do not distinguish between partial or full work resumption: among those who are found to have partly resumed in one wave 63% have fully resumed their pre-sickness work in the next wave. Moreover, we consider (partial or full) work resumption to be an absorbing state: if someone goes back to work he or she is assumed to not report sick again within the observation period of 27 months. This ignores the fact that 14% call in fully sick again after having resumed their work.

The full sample counts 3,919 employees. Of those 26% apply for disability benefit, 1 year after being sampled, and 17% are awarded a DI benefit. Among those who apply 43% work partially (as compared to their pre-sickness weekly hours) and among the new DI recipients 32% work. These numbers show that rtw and DI benefit recipiency are not mutually exclusive states.

Table 1 demonstrates that at the first interview, ten months after reporting sick, 71% had already fully or partially resumed. At 9 months after reporting sick all sample persons were still fully or partially sick listed but this percentage shows that most

Table 1 Number of sample persons by work status and survey participation

Respondents in wave	Partial or full work resumption	No work resumption	Attrition
1	2,782 (71%)	1,137 (29%)	0 (0%)
1 & 2	1,900 (48%)	391 (10%)	1,628 (42%)
1 & 2 & 3	1,360 (63%)	183 (4%)	748 (33%)

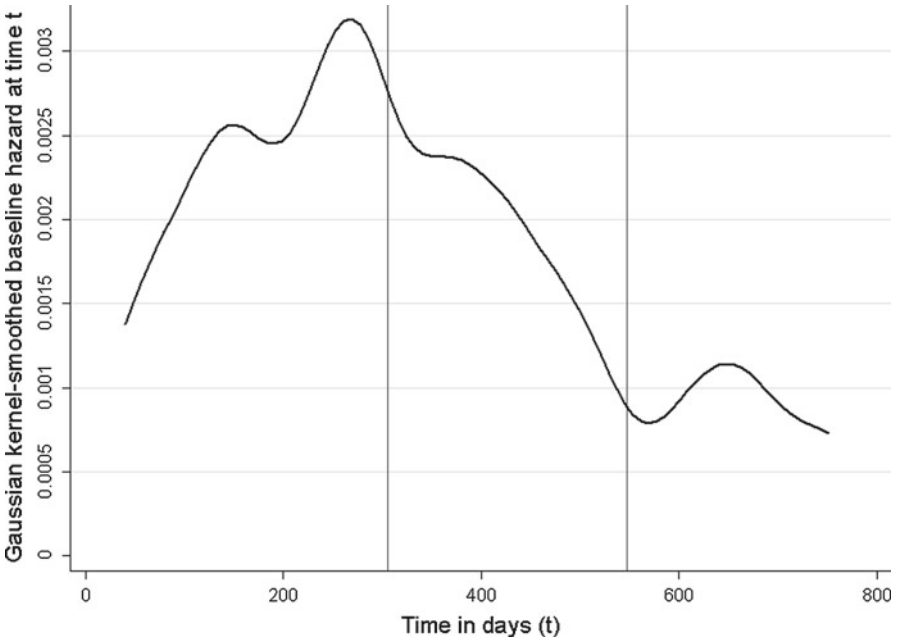

Fig. 1 Smoothed baseline hazard function with vertical lines at 10 and 18 months after the beginning of the research

of the 9 months sick workers were on their way back to work. Figure 1 shows how steeply the work resumption hazard increases over the first nine months.

In each wave questions were asked about current health and employment status, reemployment attempts, expectations with regard to recovery and work resumption and rtw activities undertaken with or without help of their employer or occupational health service (OHS). These variables may change across waves and are therefore time-varying. Background characteristics, like type of health complaints, work conditions and duration of complaints before reporting sick, education, age, gender and marital status were recorded in the first interview. Most of the questions are binary response variables. Questions with more than two answer categories are also transformed into binary variables defining dummy variables for each category.

Tables 2, 3, 4 and 5 present the definition and features of the variables used in the analyses. The number of sample persons who participated in at least the first wave and for whom we have a full record is 3,736. In this sample we have 4,562 observations indicating whether or not a transition from being fully sick to partial or full work resumption took place. There are two reasons why the number of observations may seem small relative to the size of the sample. First, as mentioned we assume partial or full work resumption to be an absorbing state. This means that we ignore transitions back into full sickness. Hence, for 71 % (see Table 1) we only have one observation. Additionally, attrition reduces the number of observations.

Rtw activities, or vocational interventions, encompass workplace adaptation, accommodation of working conditions and/or vocational training. Such activities can

Table 2 Description of personal characteristics used as explanatory variables. The reference category for education and income is the modal (most frequent) category

Personal characteristics				
Variable	Values	Meaning	Time-varying	Mean (σ)
Age at t_0	18	–	No	47.24
	19	–		(9.78)
	20	–		
	etc.	–		
Female	0	No	No	0.45
	1	Yes		
Married at t_0	0	No	No	0.72
	1	Yes		
Higher education	0	No	No	0.57
	1	University, higher professional education, higher general and pre-university education and secondary vocational education.		
Intermediate education	0	No	No	0.35
	1	Secondary general education, primary vocational education		
Lower education	0	No	No	0.07
	1	Primary education		
Non-western immigrant	0	No	No	0.10
	1	Yes		
High net income	0	No	No	0.09
per month	1	In excess of €2200		
Intermediate net income	0	No	No	0.38
per month	1	In excess of €1400 and below €2201		
Low net income	0	No	No	0.54
per month	1	Below €1401		
Number of individuals				3,736
Number of observations				4,562

be offered by the employer or by an OHS contracted by the firm. If a worker cannot be reemployed by his or her current employer a private reintegration agency or the public employment service (PES) may be enlisted to help him or her find a new job. This may be the case if the current employer is unable to offer work commensurate with the functional limitations of the disabled worker, or if the work relations have been harmed by a conflict. In our sample no more than 10% leave their pre-sickness employer and are reemployed by a new firm.[2] Evidently, all variables are self-reports.

[2] Detailed information in Dutch on this study can be found at www.ape.nl under 'WIA'

Table 3 Description of health indicators used as explanatory variables. The reference category for perceived health is the modal (most frequent) category

Health indicators				
Variable	Values	Meaning	Time-varying	Mean
Longer existing health complaints	0	No	No	0.65
	1	Yes		
Other health complaints	0	No	No	0.46
	1	Yes (other than mental or musculo-skeletal)		
Perceived health is good	0	No	Yes	0.30
	1	Yes		
Perceived health is fair	0	No	Yes	0.54
	1	Yes		
Perceived health is bad	0	No	Yes	0.16
	1	Yes		
Number of individuals				3,736
Number of observations				4,562

Table 4 Description of intervention(s) used as explanatory variables

Vocational interventions				
Variable	Values	Meaning	Time-varying	Mean
Interventions by the employer	0	No	Yes	0.77
	1	Yes		
Interventions by an OHS	0	No	Yes	0.71
	1	Yes		
Interventions by other agencies	0	No	Yes	0.29
	1	Yes		
Rtw-plan	0	No	No	0.63
	1	Yes		
Number of individuals				3,736
Number of observations				4,562

As Table 2 shows the typical sample person is male, married, has a higher level of schooling and is not an immigrant. The majority has a personal monthly income below €1401, net of taxes, which is low compared to an average net income of €1614 in 2010. As regards health, 65% have health complaints that already existed before reporting sick, and 46% have 'other' health complaints than mental or musculo-skeletal ailments, as reason for reporting sick (Table 3). In the remainder of the article we take a person with modal health, income and education as the reference employee, with all other variables are set at zero. Self-perceived health and being offered an interven-

Table 5 Description of work characteristics used as explanatory variables

Work characteristics				
Variable	Values	Meaning	Time-varying	Mean
Career prospects	0	No	No	0.34
	1	Yes		
Complaints caused by	0	No	No	0.13
a conflict at work	1	Yes		
Complaints caused by	0	No	No	0.30
time-pressure	1	Yes		
Number of individuals				3,736
Number of observations				4,562

tion are time-varying characteristics. Across the three waves 70% perceive their health as fair or worse. Table 4 shows that 77% receive interventions from their employer and 71% from an Occupational Health Service (OHS). Only 29% receive interventions from public agencies, like the NSII or the PES, or from other private firms, like re-integration companies. These last organizations get involved if the current employer is unable to offer commensurate work. 63% report that a rtw-plan has been drafted after 8 weeks sickness. Table 5 contains a number of work characteristics that we want to take account of. A previous analysis shows that employers offer vocational interventions selectively: workers with good career prospects and those who report sick as a result of stress complaints due to time-pressure have a higher probability of being offered rtw activities, whereas workers who report sick due to a conflict are less frequently selected for vocational interventions (De Jong et al. 2010). To control for selectivity we include these variables in the set of covariates.

Transitions can be observed until a sample person drops out of the survey or until the observation period has lapsed. 17% drop out before the last wave while no work resumption is observed. Furthermore, at the end of the observation period, 27 months after the first day of sickness, 5 % have not resumed their work. All such observations are right censored.

5 The Model

Throughout this paper we consider transitions from being fully sick to partial or full work resumption as an absorbing state, thus ignoring observations of 'fall-backs': those that report fully sick again after having resumed. Our focus is on rtw, and not on its duration. Taking account of consecutive spells of sickness and work would require competing risk models. Given the relative scarcity of fall-backs in our sample we would like to argue that such an extension would not critically affect our results.

5.1 The Cox Model

As a first approach we posit a Cox proportional hazard model, which means that the hazard rate conditional on covariates ($\lambda(\mathbf{x}, \beta)$) is proportional to a baseline hazard (λ_0). This model is called semi-parametric because the baseline hazard is non-parametric and the proportional scaling of this baseline hazard, conditional on the covariates x, is parametrized. Cox (1972) shows that when the main interest lies in estimating the effects of the covariates the baseline hazard does not need to be specified. This makes the Cox model considerably more flexible than parametric models. To get the conditional hazard rate the baseline hazard is multiplied by a scale factor depending on the covariates ($\phi(\mathbf{x}, \beta)$).

T denotes the duration of being fully sick-listed. Across the sample the ordered discrete rtw times are given by $t_1 < t_2 < \cdots < t_j < \cdots < t_k$. Here k is the number of discrete rtw times; $k \leq N$, with N being the sample size. The probability of interest for the Cox model, proposed by Breslow and Peto, discussed in Cox and Oakes (1984), is given by:

$$\Pr[T_j = t_j | R(t_j)] \approx \frac{\prod_{m \in D(t_j)} \phi(\mathbf{x_m}, \boldsymbol{\beta})}{\left[\sum_{l \in R(t_j)} \phi(\mathbf{x_l}, \boldsymbol{\beta})\right]^{d_j}}, \tag{1}$$

where

$$\begin{aligned} \phi(\mathbf{x}, \beta) &= \text{scale factor depending on covariates } \mathbf{x} \\ d_j &= \text{number of sick spells completed at time } t_j, \\ D(t_j) &= l : t_l = t_j = \text{set of sick spells completed at time } t_j \\ R(t_j) &= l : t_l \geq t_j = \text{set of sick spells at risk at time } t_j, \end{aligned}$$

To ensure that the scale factor, $\phi(\mathbf{x}, \beta)$, is positive we choose the exponential form. Given that in this specification the baseline hazard drops out of the equation, its functional form does not need to be specified.

Following Cox, estimation of the parameters is done by maximizing the log partial likelihood:

$$P\mathcal{L}(\beta) = \sum_{j=1}^{k} \left[\sum_{m \in D(t_j)} (\mathbf{x_m}(\mathbf{t_j})' \boldsymbol{\beta}) - d_j \cdot \ln \sum_{l \in R(t_j)} \left(e^{\mathbf{x_l}(\mathbf{t_j})' \boldsymbol{\beta}} \right) \right], \tag{2}$$

Each sample person has three observations at most, one for every wave. An observation of an individual contributes to the likelihood in Eq. (2) when the individual is still observed as being fully sick at time t_j, or when the sick spell is completed (rtw is observed) at time t_j. As mentioned before, observations in which no work resumption is reported are right-censored. Sample persons who report that they resumed their job, but for whom we do not know when, are treated differently. For those observations we arbitrarily put the moment of work resumption in the middle of the interval between two waves.

5.2 Parametric Alternatives

The flexibility of the semi-parametric Cox model, as specified in Eqs. (1 and 2), comes at a price: it does not properly treat transitions of individuals for whom the exact date of reemployment is unknown (i.e., interval censoring) and it ignores unobserved heterogeneity. As an alternative we choose the Weibull model and take account of interval-censoring and unobserved heterogeneity to see to what extent the results obtained with the Cox model are robust against such alternative specifications.

Contrary to the Cox model, the Weibull model specifies the scale factor as well as the functional form of the baseline hazard as shown by the hazard conditional on covariates x:

$$\lambda(t|\mathbf{x}, \boldsymbol{\beta}) = \alpha t^{\alpha-1} \exp(\mathbf{x}'\boldsymbol{\beta}). \tag{3}$$

The Weibull model has a monotonic increasing hazard rate when the baseline parameter $\alpha > 1$, a monotonic decreasing hazard rate when $\alpha < 1$ and a constant hazard rate when $\alpha = 1$. The baseline hazard is thus $\lambda_0(t) = \alpha t^{\alpha-1}$ and the scale factor is again assumed to be exponential. The likelihood for a Weibull model taking account of interval censored observations is then as follows:

$$L = \prod_i f(t|\mathbf{x}, \boldsymbol{\beta})^{c_{r,i}} (S(t_{-1}|\mathbf{x}, \boldsymbol{\beta}) - S(t|\mathbf{x}, \boldsymbol{\beta}))^{c_{int,i}} S(t|\mathbf{x}, \boldsymbol{\beta})^{(1-c_{r,i})(1-c_{int,i})}, \tag{4}$$

where $c_{int,i}$ is an indicator whether the observation in interval censored or not and where $c_{r,i}$ is equal to one when the exact resumption date is observed, and S is the usual survival (in sickness) function:

$$S(t|\mathbf{x}, \boldsymbol{\beta}) = 1 - F(t|\mathbf{x}, \boldsymbol{\beta}). \tag{5}$$

The contribution to the likelihood of an observation in which the employee resumes work, where the exact date of resumption is known, is the Weibull density evaluated at the time instance of work resumption (the first part of the likelihood in Eq. (4)). The contribution to the likelihood of interval censored observations is the probability of return to work between the beginning and the end of the observation (the second part of the likelihood in Eq. (4)). Finally, the survival function evaluated at the end of the observation (the third part of the likelihood in Eq. (4)) represents observations in which no work resumption took place.

Next, we take account of unobserved heterogeneity by adding a multiplicative term, θ_i, to the individual hazard function. Since the hazard rate should be positive θ_i has to be positive. The unobserved heterogeneity term is assumed to follow a certain distribution that we approximate using a discrete mass-point approach. Such a nonparametric approach for dealing with unobserved heterogeneity is also used and discussed in Heckman and Singer (1984), Baker and Melino (1999) and Gaure et al. (2005).

We use non-parametric maximum likelihood estimation to simultaneously estimate the regression parameters and the mass points (together with their probabilities) of the

heterogeneity distribution and apply an algorithm developed by Gaure et al. (2005) to find mass points. According to Røed and Raaum (2006) and Gaure et al. (2005), the number of mass points increases with the sample size. The size of our sample is smaller than any considered in the papers cited. The expected number of mass points to be found is therefore small. In articles by Carrasco and Garcia Pérez (2008) and Van den Berg (2001) it is argued that it is often hard to find more than 2 or 3 mass points.

6 Results

6.1 Results for the Cox Model

Table 7 presents the results of the semi-parametric Cox model, using the partial likelihood method (see Eq. (2)). The second column in Table 7 shows the factors by which the baseline hazard has to be multiplied to obtain marginal effects. Perceiving oneself as being in good health induces a rtw probability which is 1.31 times the rtw probability of the reference employee (with fair health). The baseline hazard applies to a person with the reference characteristics and puts all other variables at 0, except age for which we take the mean (47). Controlling for selectivity in the provision of rtw activities and for health, interventions distinguished by provider have a relatively strong impact on the rtw probability. Moreover, the presence of a rtw plan, which is prescribed by the Gatekeeper Protocol but not enforced, has its intended work enhancing effect. As mentioned before, selectivity is captured by the last three variables, two of which are significant. Time-pressure, or work stress, as a reason for long-term sickness increases the rtw hazard. It indicates that stress complaints are easier to cope with than other causes of prolonged sickness. Apart from the obvious effects of self-perceived health, age, education and income are statistically significant and have their expected signs. Calling in sick with other complaints than musculo-skeletal or mental ones decreases the probability of work resumption.

The coefficient of interventions by other agencies has a negative sign, indicating that individuals receiving interventions from other agencies have a rtw chance that is even lower than that of those without an intervention. This is a counterintuitive result, which may be caused by the fact that only a few employees get interventions from other agencies without getting interventions from the employer or OHS too, see Table 6. The effect of interventions by other agencies is therefore hard to extract from the data.

Table 6 Numbers of observations on interventions according to provider

	Interventions by other agencies	
Interventions by the employer or OHS	Yes	No
Yes	1,205	2,926
No	76	508

Table 7 Estimates of the Cox duration model

Variable	Coefficient [a]	Factor
Age at t_0	−0.006**	0.99
	(−2.34)	
Female	−0.109**	0.90
	(−2.17)	
Married at t_0	0.003	1.00
	(0.06)	
Higher education [b]	Reference	
Intermediate education	−0.006	0.99
	(−0.11)	
Lower education	−0.183*	0.83
	(−1.72)	
Non-western immigrant	−0.060	0.94
	(−0.75)	
High net income per month (>2200€)	0.208**	1.23
	(2.32)	
Intermediate net income per month (>1400, <2201€)	0.236***	1.27
	(4.53)	
Low net income per month (<1401€)[b]	Reference	
Other health complaints	−0, 148***	0.86
	(−3.18)	
Longer existing health complaints	0.023	1.02
	(0.49)	
Good health	0.294***	1.34
	(6.27)	
Fair health [b]	Reference	
Poor health	−0.692***	0.50
	(−8.17)	
Intervention(s) by an OHS	0.290***	1.34
	(4.55)	
Intervention(s) by the employer	0.525***	1.69
	(8.37)	
Intervention(s) by other agencies	−0.223***	0.80
	(−4.47)	
Rtw-plan	0.207***	1.23
	(4.22)	

Table 7 continued

Variable	Coefficient[a]	Factor
Career prospects	−0.011	0.99
	(−0.22)	
Complaints caused by conflict with the employer	−0.153**	0.86
	(−2.21)	
Complaints caused by time-pressure	0.229***	1.26
	(4.52)	
Number of individuals		3,736
Number of observations		4,562
Partial log likelihood		−15, 464.7

[a] t-values in parentheses: $* p < 0.10$ ($t \geq 1.64$), $** p < 0.05$ ($t \geq 1.96$) and $p < 0.01$ ($t \geq 2.58$)
[b] The baseline hazard for the regular employees is based on these covariates at value 1 and all other covariates at value 0

6.1.1 Graphical Inspection

Using the estimated coefficients the trajectory of the baseline hazard over the observation period can be derived (see for instance Cameron and Trivedi (2005)). The estimated baseline hazard is plotted in Fig. 1. It is a smoothed curve fitted through the derived point-estimates using a gaussian kernel.

The two vertical lines in the graphs mark the points at which the first and the second wave (10 and 18 months after calling in sick) took place. The hazard rate first rises sharply and then declines just as fast, with a small increase (followed by a decrease once more) just before the 2 year mark. Although the survey population is selected at 9 months sick leave, a major share had already partly resumed before being sampled. This pattern reflects that graded, or therapeutic modes of work resumption have become a standard form of vocational rehabilitation, especially in the case of stress-related and pain complaints. It also confirms the previously observed behavioral phenomenon that the rtw hazard increases at the end of an entitlement period (Lindeboom and Theeuwes 1993)—in this case at the end of the two-year waiting period before possible disability benefit enrollment.

Next rtw distribution functions for different combinations of characteristics are calculated. Figures 2, 3, 4 show the resulting distribution functions according to health status and by provider. We draw these trajectories for the reference person (male, 47 years old, higher education, etc.) by varying his health and type of provider.

Figure 2 depicts the distribution functions by health status category. The vertical lines mark the one and two year after the first day of sickness points. The probability of work resumption before the end of the first year is twice as large for a person in a good health as for one in poor health. The rtw probability for those in fair or good health increases steeply in the first year and only moderately in the second year. This pattern underscores that an early start with graded work resumption is part of the success of the new Dutch sickness and disability benefit system.

Figure 3 demonstrates that the probability of work resumption in the first year is highest if the employer provides guidance towards work. Interventions by occupa-

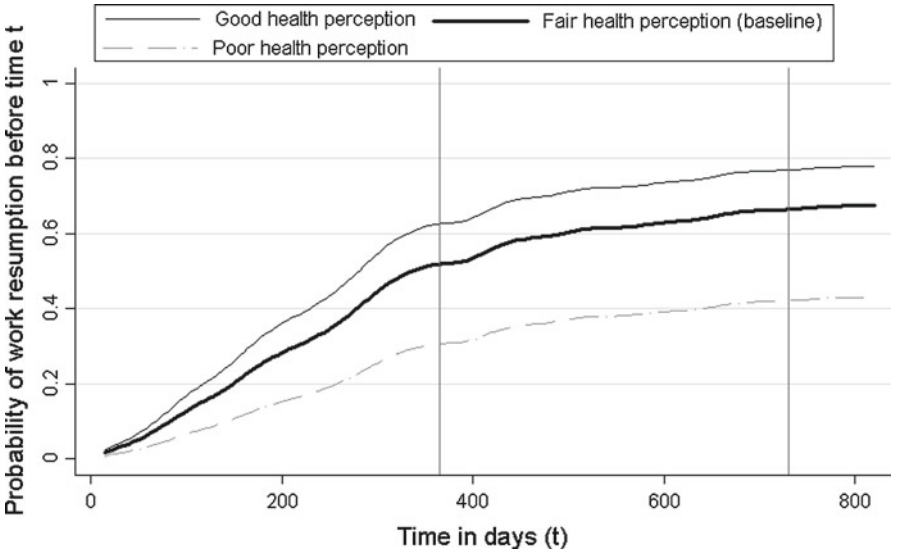

Fig. 2 Work resumption distribution for regular employees with different health categories

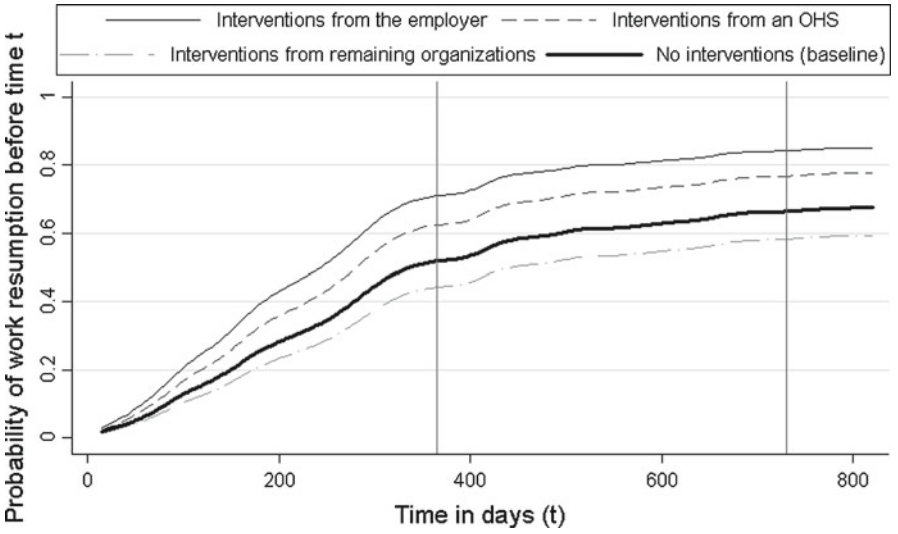

Fig. 3 Work resumption distribution for regular employees with or without received interventions

tional health service agencies, contracted by the employer, are also effective but not as strongly as provision by the employer. The effect of interventions by other providers, including public agencies, might be distorted, as it is based on few observations, as explained earlier. As stated before, this adverse effect should be interpreted carefully.

Figure 4 shows the combined effects of health and vocational interventions by both the employer and the OHS. Interestingly, interventions are found to compensate for bad health because the rtw pattern for those in good health without interventions is

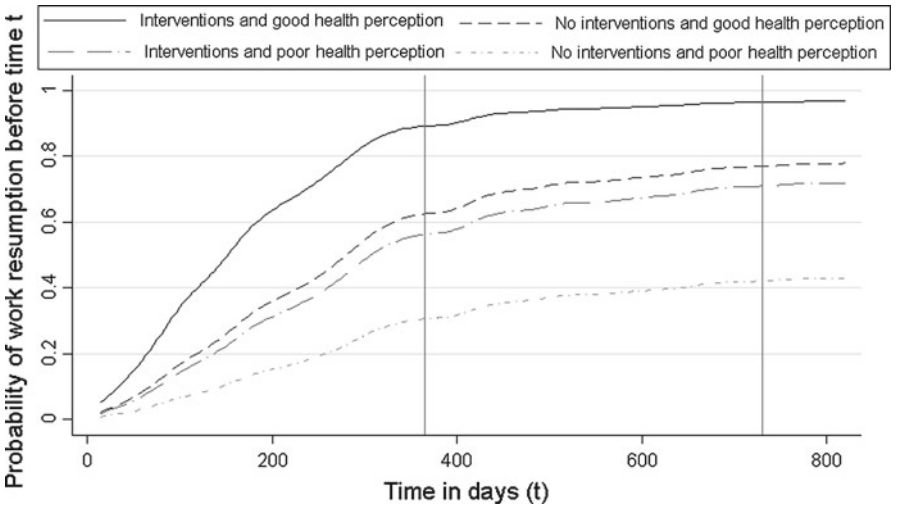

Fig. 4 Work resumption distribution for regular employees with different combinations of the various health categories and received private interventions

only slightly better than that for those in bad health who receive interventions from both their employer and the OHS.

6.2 Results for Weibull Models

In Table 8, the column denoted as 'Weibull' contains the results for the parametric duration model (see Eq. (4) for the maximized log likelihood). Apart from the shape parameter and interval censoring the estimates closely resemble those of the Cox model in terms of size, sign and significance. Notice furthermore that the shape parameter is larger than 1. This implies an increasing hazard function which differs from the inversed bathtub shape found for the Cox model (see Fig. 1). When estimating the parametric models without taking into account interval censoring, the value of the shape parameter is much higher (results not presented in this paper, but available upon request from the authors). Ignoring interval censoring in a parametric model appears to overestimate the duration dependence parameter.

The column 'Weibull +UH' gives the results when accounting for unobserved heterogeneity. We were able to find two mass points θ_1 and θ_2. The probability of the second mass point (1 minus the probability of the first mass point) was estimated simultaneously, and turns out to be small but significant. But the mass-point itself, θ_2 is insignificant. Notice that taking account of unobserved heterogeneity increases the effects of almost all covariates a bit. Furthermore, in this specification the constant term is indeterminate, because whatever value one chooses for the constant the values of the mass-points will adapt accordingly.

Adding unobserved heterogeneity with only one significant mass point does not improve the likelihood significantly. This may be due to the fact that we use a sample of employees who were sick for at least 9 months. A sample that would include those

that fully resume at an earlier stage would yield a dataset that undoubtedly contains more (unobserved) heterogeneity.

7 Conclusions and Recommendations for Further Research

The spectacular drop in DI enrollments can be in part attributed to the fact that all Dutch employers are legally mandated to pay sickness benefit during the two-year waiting period before entitlement for disability benefit may start. Apart from bearing the financial risk of sickness absenteeism they are also obliged to follow the so-called Gatekeeper Protocol prescribing the steps that the firm and the sick employee have to take to prevent prolonged sickness and DI enrollment.[3].

In this paper we investigated the role of employer based vocational interventions that the Gatekeeper Protocol describes. Two previous Dutch studies using longitudinal data proved that taking the prescribed steps increases the rtw probability but they don't fully exploit the longitudinal features of their data. We use a semi-parametric Cox model, as well as a parametric Weibull model to control for interval censoring and unobserved heterogeneity, to analyze the effects of vocational interventions by employers and other providers, controlling for self-perceived health and selectivity in the provision of interventions. By using such models we can take account of the timing of interventions and changes in health status over the 27 months observation period. Over that period a cohort of 3,736 employees who reported sick around January 1, 2007 and had not (fully) returned to work 9 months later, took part in a longitudinal survey that consisted of three waves of written interviews, at 10, 18 and 27 months after the first day of sickness.

Controlling for self-perceived health, age, gender, education, personal net income, duration and type of complaints, career prospects, and time pressure and conflicts at work as causes for reporting sick, we find strong impacts of vocational interventions by employers themselves and by OHS agencies contracted by employers. Despite their long sick spells 71% of the sample population have already partly resumed their work when the first wave is held (10 months after reporting sick). This, and other results of our analyses, confirm that an important part of the success of the new regime is an early start with graded work resumption as a standard form of vocational rehabilitation, especially in the case of health complaints that are difficult to assess objectively. Such complaints used to be a major source of long-term absenteeism and entry into the disability benefit scheme. The incentives of the new regime, however, spurred the demand for vocational interventions that prevent prolonged sickness, and OHS agencies and occupational physicians learned to meet that demand in increasingly effective ways.

Despite the strong effects reported here, several limitations remain. First, we did not distinguish between partial and full rtw. Within our observation period 56% of those that resumed partially end up working at their pre-sickness hours. On the other hand, part of those that partially resume enter the DI scheme. Distinguishing between partial and full work resumption may change the effects reported here, and may be informative

[3] Van Sonsbeek and Gradus (2010) show that 22 points of the 70% drop in DI enrollments between 2000 and 2009 can be attributed to the Gatekeeper Protocol.

Table 8 Estimates of Weibull duration models

	Weibull		Weibull + UH	
	Coefficient[a]	Factor	Coefficient[a]	Factor
Female	−0.109**	0.90	−0.109**	0.90
	(−2.11)		(−1.99)	
Married	0.007	1.01	0.031	1.03
	(0.13)		(0.58)	
Higher education [b]	Reference		Reference	
Intermediate education	0.009	1.01	0.027	1.03
	(0.18)		(0.48)	
Lower education	−0.191*	0.83	−0.204*	0.82
	(−1.78)		(−1.86)	
Non-western immigrant	−0.026	0.97	−0.054	0.95
	(−0.32)		(−0.65)	
High net income per month (>2200€)	0.206**	1.23	0.281***	1.32
	(2.23)		(2.73)	
Intermediate net income per month (>1400, <2201€)	0.237***	1.27	0.242***	1.27
	(4.40)		(4.21)	
Low net income per month (<1401€)[b]	Reference		Reference	
Other health complaints	−0.174***	0.84	−0.199***	0.82
	(−3.65)		(−3.90)	
Age at t_0	−0.005**	0.99	−0.006**	0.99
	(−2.12)		(−2.08)	
Longer existing health complaints	−0.027	0.97	0.001	1.00
	(−0.57)		(0.02)	
Good health	0.284***	1.33	0.284***	1.33
	(5.85)		(5.56)	
Fair health [b]	Reference		Reference	
Poor health	−0.747***	0.47	−0.757***	0.47
	(−8.84)		(−8.77)	
Intervention(s) by an OHS	0.391***	1.48	0.405***	1.50
	(6.08)		(6.09)	
Intervention(s) by the employer	0.639***	1.89	0.629***	1.88
	(10.31)		(9.80)	
Intervention(s) by other agencies	−0.277***	0.76	−0.263***	0.77
	(−5.41)		(−4.88)	
Rtw-plan	0.220***	1.25	0.246***	1.28
	(4.42)		(4.61)	

Table 8 continued

	Weibull		Weibull + UH	
	Coefficient [a]	Factor	Coefficient [a]	Factor
Carreer prospects	−0.018	0.98	−0.022	0.98
	(−0.37)		(−0.424)	
Complaints caused by conflict with the employer	−0.186***	0.83	−0.208***	0.81
	(−2.624)		(−2.744)	
Complaints caused by time-pressure	00.206***	1.23	0.222***	1.25
	(3.94)		(3.92)	
Constant	−7.026		−7.026	
	(−36.39)		–	
Shape-parameter	1.092***		1.144***	
	(42.08)		(35.46)	
θ_1			0.745***	
			(4.49)	
θ_2			0.005	
			(0.23)	
$P(\theta_2)$			0.023**	
			(2.55)	
Number of individuals	3,736		3,736	
Number of observations	4,562		4,562	
Log likelihood	−8, 220.9		−8, 217.1	

[a] t-values in parentheses: * $p < 0.10$ ($t \geq 1.64$),** $p < 0.05$ ($t \geq 1.96$) and $p < 0.01$ ($t \geq 2.58$)
[b] The baseline hazard for the regular employees is based on these covariates at value 1 and all other covariates at value 0

on the timing and eventual success of graded rtw. Second, we distinguished interventions by provider but not by type of activity as, for instance, Frölich et al. (2004) do. Adding such time-varying factors may be a useful elaboration because they may tell us something about the efficacy and optimal timing of distinct activities. Third, for lack of a more objective measure, we used self-reported health as an indicator of true health. To the extent that the difference between true and self-reported health is determined by unobserved taste factors our results are biased. We may further explore this matter by using doctors' assessments in administrative records. And, finally, we have assumed that rtw is an absorbing state although our data tell us that one in seven of those who resume their work fall back into full sickness later on. In future work we will try to remove this limitation.

Acknowledgments The authors are grateful for the financial support from Stichting Instituut gak which made the research, on which this paper is based, possible.

References

Baker, M., & Melino, A. (1999, June). Duration dependence and nonparametric heterogeneity: A monte carlo study. Working Papers melino-99-01, University of Toronto, Department of Economics.

Burkhauser, R. V., Daly, M. C. & de Jong, Ph. R. (2008). Curing the dutch disease: Lessons for united states disability policy. Working paper 188, University of Michigan, Michigan Retirement Research Center.

Butler, R. J., Johnson, W. G., & Baldwin, M. L. (1995). Managing work disability: Why first return to works is not a measure of success. *Industrial and Labor Relations Review, 48*(3), 452–469.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics*. Cambridge: Cambridge University Press.

Carrasco, R., & García Pérez J. I. (2008, December). Unemployment duration among immigrants and natives: Unobserved heterogeneity in a multi-spell duration model. Working Papers 08.13, Universidad Pablo de Olavide, Department of Economics.

Corten, I. W. (2009). Reïntegratie van langdurig zieke werknemers in de bouw. Economisch instituut voor de Bouwnijverheid.

Cox, D. R., & Oakes, D. (1984). *Analysis of survival data. Monographs on statistics and applied probability*. London: Chapman and Hall.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistics Society, Series B, 34*, 187–220.

De Jong, P. R., Lindeboom, M., & Vander Klaauw, B. (2011). Screening disability insurance applications. *Journal of the European Economic Association, 9*(1), 106–129.

De Jong, P. R., Veerman, T., Van der Burg, C., & Schrijvershof C. (2010). De Weg naar de WIA. APE/AStri.

Franche, R. L., Cullen, K., Clarke, J., Irvin, E., Sinclair, S., & Frank, J. (2005). Workplace-based return-to-work interventions: A systematic review of the quantitative literature. *Journal of Occupational Rehabilitation, 15*(4), 607–631.

Frölich, M., Heshmati, A., & Lechner, M. (2004). A microeconometric evaluation of rehabilitation of long-term sickness in Sweden. *Journal of Applied Econometrics, 19*(3), 375–396.

Gaure, S., Røed, K., & Zhang, T. (2005, August). Time and causality: A monte carlo assessment of the timing-of-events approach. Memorandum 19/2005, Oslo University, Department of Economics.

Heckman, J., & Singer, B. (1984). The identifiability of the proportional hazard model. *Review of Economic Studies, 51*(2), 41–231.

Høgelund, J., Holm, A., & McIntosh, J. (2010). Does graded return-to-work improve sick-listed workers' chance of returning to regular working hours?. *Journal of Health Economics, 29*(1), 158–169.

Kenniscentrum U. W. V. (2010). UWV Kwartaal Verkenning 2010-III.

Koning, P. (2004). Estimating the impact of experience rating on the inflow into disability insurance in the Netherlands. CPB Discussion Paper 37, Centraal Planbureau.

Krause, N., Dasinger, L. K., & Neuhauser, F. (1998). Modified work and return to work: A review of the literature. *Journal of Occupational Rehabilitation, 8*(2), 113–139.

Lindeboom, M., & Theeuwes, J. (1993). Search, benefits and entitlement. *Economica, 60*(239), 327–346.

OECD (2010). Sickness, disability and work: Breaking the barriers. A synthesis of findings across OECD countries. Organisation for Economic Co-operation and Development, Paris.

Oomens, S., Koppes, L., van den Bossche, S., & Houtman, I. (2010). Langdurig zieke werknemers en hun werkhervattingen. In G. Jehoel-Gijdbers (Ed.), *Beperkt aan het werk*. Netherland: Sociaal Cultureel Planbureau.

Røed, K., & Raaum, O. (2006). Do labour market programmes speed up the return to work? *Oxford Bulletin of Economics and Statistics, 68*(5), 541–568.

Spierdijk, L., Van Lomwel, G., & Peppelman, W. (2009). The determinants of sick leave durations of Dutch self-employed. *Journal of Health Economics, 28*(6), 1185–1196.

Van Sonsbeek, J. M., & Gradus, R. H. J. M. (2010). De beleidsmaatregelen voor de WAO gekwantificeerd. *Economisch Statistische Berichten, 95*(4594), 586–588.

Van den Berg, G. J. (2001, March). Duration models: Specification, identification and multiple durations. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of Econometrics*, (Vol. 5, Chap. 55, pp. 3381–3460). Elsevier.