Does VDAB Training Influences Re-Employment? A Duration Analysis

Rembert De Blander

Steven Groenez

2016 nr. 03

WSE Report

Steunpunt Werk en Sociale Economie

Naamsestraat 61/3551 - 3000 Leuven

T:+32 (0)16 32 32 39

steunpuntwse@kuleuven.be

www.steunpuntwse.be

Does VDAB Training Influences Re-Employment? A Duration Analysis

Rembert De Blander
Steven Groenez
HIVA, KU Leuven

Een onderzoek in opdracht van de Vlaamse minister van Werk, Economie, Innovatie en Sport in het kader van het Vlaams Programma Strategisch Arbeidsmarktonderzoek.

De Blander, R., Groenez, S. (2016). Does VDAB Training Influences Re-Employment? A Duration Analysis (WSE Report 2016 nr. 03). Leuven: Steunpunt Werk en Sociale Economie / HIVA, KU Leuven.

ISBN: 9789088731297

Naamsestraat 61 /3551 – B-3000 Leuven

T: +32(0)16 32 32 39

steunpuntwse@kuleuven.be

www.steunpuntwse.be

Does VDAB Training Influences Re-Employment? A Duration Analysis

Rembert De Blander*and Steven Groenez
HIVA - KU Leuven

June 30, 2016

Abstract

We estimate the effect of training on the duration of the unemployment spell for people who became unemployed. To that effect we use the timing of events approach, jointly modeling the waiting time till training and unemployment duration with a common frailty. Our results point to a significant locking in effect but also to a positive effect of training lasting for 2 years.

1 Introduction

The question whether some active labor market programs (ALMP), such as job search assistance, subsidized employment[1] or labor market training, has an effect on the labor market outcomes of the participants, seems pertinent to all parties concerned: the unemployed, unemployment agencies and governments. Their main goal is to integrate unemployed or otherwise disadvantaged individuals into the workforce, using one of the following strategies Heckman et al. (1999, p.2043): increasing human capital, improving work habits, helping job search or some combination of them. Consequently, the substantial body of literature on ALMPs mainly considers labor market outcomes (Heckman et al., 1999, p.2044), such as earnings and wages, employment rates, unemployment exit rates, claims of unemployment benefits, ... In general following conclusions seem to hold: (i) wage subsidies (WS) and job search assistance (JSA) programs can be effective in increasing participants' employment probability; (ii) training programs, be it classroom or on the job, are ineffective in the short run, but seem to exhibit some positive medium term effects; (iii) direct employment programs in the public sector are frequently not effective; (iv) all impact estimates are very sensitive to the method used (Heckman et al., 1999, p.2066); (v) all impact estimates exhibit heterogeneity both geographically and among demographic and skill groups (Heckman et al., 1999, p.2054). For meta-analyses (and, hence overviews of the literature) see Heckman et al. (1999) and more recently Card et al. (2010). Studies covering specific geographic areas include Greenberg et al. (2003) for the US and Kluve (2010) for the EU.

Existing research on the effectiveness of training programs as ALMP instrument comes in in two flavors. On the one hand there are the so-called experimentalists: proponents of (quasi-)randomized controlled trials (RCT) (for example Ham and Lalonde, 1996; Eberwein et al., 1997), while on the other hand we have the advocates of the (structural) econometric method (EM), which jointly models the outcome together with an explicit choice mechanism, thus alleviating or even completely bypassing the need for RCTs in order to evaluate the impact of social policies such as training programs. For a

*Corresponding author. Address: KU Leuven, HIVA, Parkstraat 47, B-3000 Leuven, Belgium. Tel: +32 16 323318. Email: rembert@deblander.eu (De Blander); steven.groenez@kuleuven.be (Groenez).

[1]Subsidized employment not only consists of wage subsidies, but also of direct (temporary) job creation in the public sector, although the latter seems recently less fashionable.

comprehensive overview of issues involved and in-depth comparisons between assumptions underlying both methods, see Heckman et al. (1999); Heckman and Navarro-Lozano (2004), and references therein.

In contrast with some third world, mainly world bank sponsored, evaluation studies[2], European evaluation studies of training programs are almost exclusively non-experimental. Moreover most studies use the *timing of events* evaluation design (Abbring and van den Berg, 2003). By jointly modeling the duration of the unemployment spell and the time elapsed before a training is started, the effect of the training on the unemployment duration is purged of confounding unobserved common factors, such as motivation. The intuition is that a training can only affect the probability of finding a job ex-post, while confounders have an effect both before and after the training. At least if we rule out any anticipation effects! Examples of this approach include Richardson and van den Berg (2013) for Sweden, Hujer et al. (2006) using data for East Germany and Lalive et al. (2008) who find no effect for the Swiss unemployed. More recent approaches not only consider exit probabilities out of unemployment, but characteristics of subsequent jobs as well (Crépon et al., 2012; Osikominu, 2013).

In the Belgian context, the only readily available quantitative results on the effectiveness of training as an ALMP come from Cockx (2002), who estimates the change in exit probabilities out of unemployment due to a training organized by the Walloon public employment service (FOREM). For Flanders, sparse studies include overviews of methodological issues (Bollens2007, 2014), and a survey where participants rate the training they followed (De Rick and De Cuyper, 2014).

In this paper, we try to fill this perceived gap for Flanders by assessing the effect of training programs on the subsequent employment probability, using administrative data, which are described in section 2. The method employed is presented in section 3 and results are described in section 4, while section 5 concludes.

2 Data

We use data from the public employment service (*PES*) in Flanders[3] consisting of individual records[4], whereby anyone entering unemployment is recorded at the end of the month of entrance. Moreover, we have information on the labor market position of the individual at the end of all subsequent calendar months[5]. From these data, we selected those individuals for which an unemployment spell involving unemployment benefits, started in 2007[6]. In addition were removed: the impaired[7], those younger than 25 and older than 65[8], disabled persons and individuals living outside Flanders. For each subject, the data runs from the first entrance in unemployment in 2007 up to the end of September 2010.

Quite obviously, duration is defined as the amount of time passed until the current state ends in favor of some other state of interest (i.e. employment), while alternative exits[9] as well as the end of the observation window, censor the current state. Similarly, we define a waiting time between the onset of unemployment and the start of some kind of training.

We obtained a second data-set containing data on 13067 training sessions from the *PES*, corresponding with all individuals we selected in the previous step. From these training session records, we remove the modules "person-centered training" (persoonsgerichte vorming), "job search training" (sollicitatietraining) and an "introductory module" (trajectbepaling) as well as all internet-based types of training, on the grounds that either their contents are very basic or that the training effort can not be verified. This operation removes 7622 (training) observations. Next, only entries are retained

[2]Examples include Hirshleifer et al. (2015), Attanasio et al. (2011), Card et al. (2011) and Cho et al. (2013).

[3]Its official acronym is VDAB, or "Vlaamse Dienst voor ArbeidschBemiddeling" in full.

[4]The primary source of these data is the (federal) unemployment register.

[5]Our data are thus left-truncated and independently interval censored.

[6]Thus removing voluntary registered job seekers and school-leavers.

[7]Someone is labeled impaired ("arbeidsgehandicapt") when (s)he is less able to obtain or perform a job due to an illness (p.ex. burn-out) or a handicap.

[8]With ages recorded at the onset of the unemployment spell.

[9]Alternative exits include: inactivity, i.e. voluntary unemployment, sickness, retirement, ...

that have a meaningful status: finished, interrupted or ongoing.[10] We further distinguish between successfully finished training spells on the one hand and prematurely interrupted or failed training spells on the other hand. Next, we group all training entries according to a 79-fold classification of remaining training types[11], after which we remove duplicates based on the timing and the result of the training. Next, we group adjacent training spells, which we define as being no further apart than three months, into one training block, which we assign the status and result of the final spell within the block. After cleaning up the dataset on training, we end up with the situation given in Table 1.

#trainings/spell	#spells	#trainings
1	3602	3602
2	60	120
3	1	3
Total	3663	3725

Table 1: Number of training episodes per unemployment spell

If an unemployment spell contains more than one training, the onset of the second training censors the remaining unemployment spell. Likewise, if an individual's records contain more than one unemployment spell, we use only the initial one, again, for convenience reasons. The finally concocted database thus consists of 108647 individuals starting an initial unemployment spell somewhere during 2007, of which 2389 received at least one training during their first unemployment spell.

Table 2 presents descriptive statistics of our finally resulting dataset. Note that the provinces Vlaams Brabant and West Vlaanderen are slightly underrepresented, which might reflect higher provincial unemployment rates. Slightly more than half are women, 80% have Dutch as native language, while 90% have the Belgian nationality. 80% possess a drivers license. Previous unemployment history is synthesized in the number of months the individual was unemployed in the last two years prior to the start of the current unemployment spell. On average, the people in our sample have been unemployed 25% of the time. 40% in the sample have only obstained a degree in primary education, 36% secundary education, while 23% have succesfully followed tertiary education, roughly 25% of which university. Educational track refers to the objective of the subject chosen in secondary education. The vocational track prepares for a direct transition to a profession, whereas the general and the technical tracks prepare for a transition to tertiary education. During 2.4% of the unemployment spells a training was followed which lasted on average 11 months. Roughly ²/₃ finished the training successfully.

3 Modeling

Participation in a training is possibly selective, meaning that the observed and unobserved characteristics of both groups may be different, a different outcome for both groups may thus not only be a consequence of the treatment, but could also be due to these other differences. Over and above this classical selection problem, we have to take into account a dynamic selection problem, which is driven by the mechanism that the more employable workers will on average leave unemployment sooner, and therefore will have a smaller probability to be treated.

To overcome these difficulties, we control for differences between the treated group and the control group based on both observed and unobserved individual characteristics. Selection on observables is taken into account by conditioning the hazard rates on explanatory variables. Selection on unobserved characteristics is taken into account by making use of a "frailty", which models an unobserved factor influencing both the labor market fitness and the willingness to follow a training (Abbring and van den Berg 2003, 2004). This approach is also called the Timing of Events approach. It exploits the fact that

[10]We thus remove 311 entries with a status that has only an administrative meaning, such as "in process", "ready to start", "not started", "annulled", . . .

[11]There exists a much finer, i.e. 577-fold classification of training types which proved not very helpful.

Name	Mean
Vlaams Brabant	0.1460928
West Vlaanderen	0.1647355
Oost Vlaanderen	0.2414247
Limburg	0.1466731
Antwerpen	0.3010739
Sex	0.5030764
Language Dutch	0.7987436
Belgian nationality	0.8919018
Drivers license	0.7955935
Months unemployed (past 2 years)	6.082078
Primary education	0.4073668
Secondary education	0.3649694
Tertiary education: college	0.1671948
Tertiary education: university	0.060469
Educational track: general	0.0914081
Educational track: technical	0.1662921
Educational track: vocational	0.2777338
Age (standardized)	2.64e-17
Age squared	0.9999908
#trainings/spell	0.0236257
N	108568

Name	Mean
Duration of training	11.40104
Passed	0.660026
N	2309

Table 2: Descriptive statistics

unobserved heterogeneity affects the transition to regular employment throughout the unemployment spell, whereas the treatment may only influence this transition from its onset. This difference in timing allows identification of both the treatment effect and the selection effect without imposing 'exclusion restrictions' on the observed explanatory variables. In what follows, we specify the econometric model and discuss the identification of the treatment effect.

3.1 Likelihood

Consider a state of affairs where people find themselves out of a job. During their unemployment spell they will be offered different kinds of training, which they can turn down or accept. The main question of interest in this paper is whether (different types of) training influence(s) the hazard of finding a job (differently).

At the moment of being sacked, two processes start up: a job search process, J, ending with finding a job, and a waiting process, W, ending with starting some training, O_k of type $k = 1, \ldots, K$. The first sub-spell of an unemployment spell, ends by starting a training or by finding a job. Once the training started, one either completes the training, fails or finds a job. After the training finished, the spell ends by finding a job or by being censored.

Define the hazard rate $\theta_P(t \mid \mathcal{S}_P)$ of process $P = J, W,\ k = 1, \ldots, K$ as

$$\lambda_P(t \mid \mathcal{S}_P) = \lim_{dt \to 0} \frac{\Pr[t \leq T < t + dt \mid T \geq t; \mathcal{S}_P]}{dt},$$

with survival function

$$S_P(t \mid \mathcal{S}_P) = \exp\left\{ - \int_0^t \lambda_P(u \mid \mathcal{S}_P)\, du \right\}$$

and *PDF*

$$f_P(t \mid \mathcal{S}_P) = \lambda_P(t \mid \mathcal{S}_P) \cdot S_P(t \mid \mathcal{S}_P).$$

The $\mathcal{S}_P$ denotes the set of conditioning variables.

The joint distribution of $(T_J, T_W) \mid (X, \Xi)$ is factored as

$$f_{(T_J,T_W)\mid\left(X,T_{O_k},\Xi\right)}(t_J, t_W \mid x, t_{O_k}, \xi) = f_{T_J\mid\left(X,T_{O_k},\Xi_J\right)}(t_J \mid x, t_{O_k}, \sigma_W \xi) \cdot f_{T_W\mid(X,\Xi_W)}(t_W \mid x, \xi),$$

where ξ represents the (individual-specific) unobserved heterogeneity or frailty. Conditional on this frailty and on the set of observed conditioning variables, including receiving, failed or successfully obtained training, waiting time and the unemployment duration are independent.

For convenience, we specify each conditional hazard as a proportional hazards Lancaster (1990, p.42) model[12]

$$\lambda_P(t \mid \mathcal{S}_P) = \lambda_{P0}(t) \cdot \exp[\alpha_P' x_i + \sigma_p \cdot \xi_i]$$

in which the functional form of the baseline hazard $\lambda_{P0}(t)$ is common to all individuals. Additionally, one of the parameters σ_P, $P = J, W$ needs to be removed, i.e. fixed to the value 1.

During a complete sequence of events, an individual enters unemployment at $s = t_J = t_W = 0$ and leaves the risk set at $s = s^f$. At some point $s = s^1$, (s)he starts some training $\left(Q^k = 1\right)$, which lasts until $s = s^2 > s^1$ and ends in failure $\left(F^k = 1\right)$ or success $\left(S^k = 1\right)$. At some point in time $s = s^E$, (s)he finds a suitable job and leaves unemployment. Depending on the timing of s^E, s^1 and s^f, different contributions to the likelihood apply:

[12] Sometimes it is also called the relative risk or the Cox model (Kalbfleisch and Prentice, 2002, p.95).

- When $s_i^f < s_i^1$, we do not observe the start of any training, hence spell J results in success ($d_J = 1$), i.e. $\left(s_i^E < s_i^f\right)$ or censoring ($d_J = 0$), i.e. $\left(s_i^f < s_i^E\right)$, but W is always censored[13]. In case of success, we have left truncated and interval censored data for the unemployment spell (Kalbfleisch and Prentice, 2002, section 3.9.3, p.83), otherwise, the employment spells are left truncated and right censored

$$\begin{aligned}
& \mathcal{L}_I\left(s_i \mid x_i, \xi_i\right) \\
= \quad & \left\{S_J\left(s_i - 1 \mid x_i, q_i^k = 0, s_i^k = 0, f_i^k = 0, \xi_i\right) - S_J\left(s_i \mid x_i, q_i^k = 0, s_i^k = 0, f_i^k = 0, \xi_i\right)\right\}^{d_J} \\
& \times S_J^{(1-d_J)}\left(s_i \mid x_i, q_i^k = 0, s_i^k = 0, f_i^k = 0, \xi_i\right) \\
& \times S_W\left(s_i \mid x_i, \xi_i\right)
\end{aligned}$$

- When $s_i^1 < s_i^E < s_i^2$, the individual has *started* some training k, after which spell J results in success ($d_J = 1$) or censoring ($d_J = 0$), spell W results in success,

$$\begin{aligned}
& \mathcal{L}_{II}\left(s_i \mid x_i, \xi_i\right) \\
= \quad & \left\{S_J\left(s_i - 1 \mid x_i, q_i^k = 1, s_i^k = 0, f_i^k = 0, \xi_i\right) - S_J\left(s_i \mid x_i, q_i^k = 1, s_i^k = 0, f_i^k = 0, \xi_i\right)\right\}^{d_J} \\
& \times S_J^{(1-d_J)}\left(s_i \mid x_i, q_i^k = 1, s_i^k = 0, f_i^k = 0, \xi_i\right) \\
& \times \left\{S_W\left(s_i^1 - 1 \mid x_i, \xi_i\right) - S_W\left(s_i^1 \mid x_i, \xi_i\right)\right\}
\end{aligned}$$

- When $s_i^2 < s_i^E$, the individual has finished the training k and thus spell J results in success ($d_J = 1$) or censoring ($d_J = 0$), and spell W results in success

$$\begin{aligned}
& \mathcal{L}_{III}\left(s_i \mid x_i, \xi_i\right) \\
= \quad & \left\{S_J\left(s_i - 1 \mid x_i, q_i^k = 1, s_i^k, f_i^k, \xi_i\right) - S_J\left(s_i \mid x_i, q_i^k = 1, s_i^k, f_i^k, \xi_i\right)\right\}^{d_J} \\
& \times S_J^{(1-d_J)}\left(s_i \mid x_i, q_i^k = 1, s_i^k, f_i^k, \xi_i\right) \\
& \times \left\{S_W\left(s_i^1 - 1 \mid x_i, \xi_i\right) - S_W\left(s_i^1 \mid x_i, \xi_i\right)\right\}
\end{aligned}$$

3.2 Computational details

Since our data are only constructed on a monthly basis, we have to take this interval-censoring into account (Gaure et al., 2007). As a consequence people appearing for the first time in the data have already experienced an unemployment spell between one and 30 days. We do not take into account the fact that very short spells[14] are not observed in the data.

Both out of convenience considerations and to avoid restrictions imposed by parametric models (Heckman and Singer, 1984), we both specify the baseline hazard as well as the unobserved heterogeneity (or frailty) non-parametrically. Both baseline hazards are thus specified as piecewise-constant (PWC) with time intervals of one month[15], ie. $k = 1, \ldots, 45$. A convenient characteristic of a PWC specification is that the rather annoying integrals in the likelihood are transformed into sums. Hence we have that

$$\mathcal{L}_I\left(k_i \mid x_i, \xi_i, d_i\right) \quad = \quad \mathcal{L}_{I;J}\left(k_i \mid x_i, \xi_i, d_i\right) \cdot \mathcal{L}_{I;W}\left(k_i \mid x_i, \xi_i, d_i\right)$$

[13]Note that the usual contribution of an observed success to the Likelihood $\lambda^d(s) \cdot S(s)$ is replaced by $\{S(s-1) - S(s)\}^d$, due to the one month interval censoring, i.e. the individual survived until (the end of) month $s-1$, but left the state before (the end of) month s.

[14]Spells of less than one month which do not include the "recording date".

[15]For the exact form of the intervals, see section 4.

where

$$\mathcal{L}_{I;J}\left(k_i \mid x_i, \xi_i, d_i\right) = \left\{\exp\left[\exp\left[\mathcal{I}_{J;i} + \sigma_J \cdot \xi_i\right] \cdot \left\{\lambda_{J0}(1) - \sum_{l=1}^{k_i-1} \lambda_{J0}(l)\right\}\right]\right\}^{d_J} + (-1)^{d_J} \exp\left[\exp\left[\mathcal{I}_{J;i} + \sigma_J \cdot \xi_i\right] \cdot \left\{\lambda_{J0}(1) - \sum_{l=1}^{k_i} \lambda_{J0}(l)\right\}\right]$$

$$\mathcal{L}_{I;W}\left(k_i \mid x_i, \xi_i, d_i\right) = \exp\left[\exp\left[\mathcal{I}_{W;i} + \xi_i\right] \cdot \left\{\lambda_{W0}(1) - \sum_{l=1}^{k_i} \lambda_{W0}(l)\right\}\right]$$

for individuals never entering training, with $\mathcal{I}_{P;i} = \alpha'_P x_i,\ P = J, W$; that

$$\mathcal{L}_{II}\left(k_i \mid x_i, \xi_i, d_i\right) = \mathcal{L}_{II;J}\left(k_i \mid x_i, \xi_i, d_i\right) \cdot \mathcal{L}_{II;W}\left(k_i \mid x_i, \xi_i, d_i\right)$$

with

$$\begin{aligned}
&\mathcal{L}_{II;J}\left(k_i \mid x_i, \xi_i, d_i\right) \\
= &\left\{\left\{\exp\left[\exp\left[\mathcal{I}_{J;i} + \sigma_J \cdot \xi_i\right] \cdot \left\{\lambda_{J0}(1) - \sum_{l=1}^{k_i^s-1} \lambda_{J0}(l)\right\} - \exp\left[\mathcal{I}_{J;i} + \beta_I q_i + \sigma_J \cdot \xi_i\right] \cdot \left\{\sum_{l=k_i^s}^{k_i-1} \lambda_{J0}(l)\right\}\right]\right\}^{d_J}\right. \\
&\left. + (-1)^{d_J} \exp\left[\exp\left[\mathcal{I}_{J;i} + \sigma_J \cdot \xi_i\right] \cdot \left\{\lambda_{J0}(1) - \sum_{l=1}^{k_i^s-1} \lambda_{J0}(l)\right\} - \exp\left[\mathcal{I}_{J;i} + \beta_I q_i + \sigma_J \cdot \xi_i\right] \cdot \left\{\sum_{l=k_i^s}^{k_i} \lambda_{J0}(l) \cdot\right\}\right]\right\}
\end{aligned}$$

$$\begin{aligned}
&\mathcal{L}_{II;W}\left(k_i \mid x_i, \xi_i, d_i\right) \\
= &\exp\left[\exp\left[\mathcal{I}_{W;i} + \xi_i\right] \cdot \left\{\lambda_{W0}(1) - \sum_{l=1}^{k_i^s-1} \lambda_{W0}(l)\right\}\right] - \exp\left[\exp\left[\mathcal{I}_{W;i} + \xi_i\right] \cdot \left\{\lambda_{W0}(1) - \sum_{l=2}^{k_i^s} \lambda_{W0}(l)\right\}\right]
\end{aligned}$$

for individuals starting training; and that

$$\mathcal{L}_{III}\left(k_i \mid x_i, \xi_i, d_i\right) = \mathcal{L}_{III;J}\left(k_i \mid x_i, \xi_i, d_i\right) \cdot \mathcal{L}_{III;W}\left(k_i \mid x_i, \xi_i, d_i\right)$$

where

$$
\begin{aligned}
&\mathcal{L}_{III;J}\left(k_i \mid x_i, \xi_i, d_i\right) \\
= \quad & \left\{ \exp\left[\exp\left[\mathcal{I}_{J;i} + \sigma_J \cdot \xi_i\right] \cdot \left\{ \lambda_{J0}(1) - \sum_{l=1}^{k_i^s - 1} \lambda_{J0}(l) \right\} - \exp\left[\mathcal{I}_{J;i} + \beta_I q_i + \sigma_J \cdot \xi_i\right] \cdot \left\{ \sum_{l=k_i^s}^{k_i^e - 1} \lambda_{J0}(l) \right\} \right.\right. \\
& \quad - \sum_{m=1}^{M_S} s_i \cdot \left\{ \exp\left[\mathcal{I}_{J;i} + \gamma_{AS;m} + \sigma_J \cdot \xi_i\right] \cdot \left\{ \sum_{l=k_i^e + b_m}^{k_i^e + e_m} \left(l \leq k_i - 1\right) \cdot \lambda_{J0}(l) \right\} \right\} \\
& \quad \left.\left. - \sum_{m=1}^{M_F} f_i \cdot \left\{ \exp\left[\mathcal{I}_{J;i} + \gamma_{AF;m} + \sigma_J \cdot \xi_i\right] \cdot \left\{ \sum_{l=k_i^e + b_m}^{k_i^e + e_m} \left(l \leq k_i - 1\right) \cdot \lambda_{J0}(l) \right\} \right\} \right] \right\}^{d_J} \\
& + (-1)^{d_J} \exp\left[\exp\left[\mathcal{I}_{J;i} + \sigma_J \cdot \xi_i\right] \cdot \left\{ \lambda_{J0}(1) - \sum_{l=1}^{k_i^s - 1} \lambda_{J0}(l) \cdot \right\} - \exp\left[\mathcal{I}_{J;i} + \beta_I q_i + \sigma_J \cdot \xi_i\right] \cdot \left\{ \sum_{l=k_i^s}^{k_i^e - 1} \lambda_{J0}(l) \right\} \right. \\
& \quad - \sum_{m=1}^{M_S} s_i \cdot \left\{ \exp\left[\mathcal{I}_{J;i} + \gamma_{AS;m} + \sigma_J \cdot \xi_i\right] \cdot \left\{ \sum_{l=k_i^e + b_m}^{k_i^e + e_m} \left(l \leq k_i\right) \cdot \lambda_{J0}(l) \right\} \right\} \\
& \quad \left. - \sum_{m=1}^{M_F} f_i \cdot \left\{ \exp\left[\mathcal{I}_{J;i} + \gamma_{AF;m} + \sigma_J \cdot \xi_i\right] \cdot \left\{ \sum_{l=k_i^e + b_m}^{k_i^e + e_m} \left(l \leq k_i\right) \cdot \lambda_{J0}(l) \right\} \right\} \right]
\end{aligned}
$$

$$
\mathcal{L}_{III;W}\left(k_i \mid x_i, \xi_i, d_i\right) = \mathcal{L}_{II;W}\left(k_i \mid x_i, \xi_i, d_i\right)
$$

for individuals finishing training. The log-likelihood is now given by

$$
\begin{aligned}
\log \mathcal{L}\left(k_i \mid x_i\right) = \; & \log \mathcal{L}_I\left(k_i \mid x_i, d_i = 0\right) + \log \mathcal{L}_I\left(k_i \mid x_i, d_i = 1\right) \\
& + \log \mathcal{L}_{II}\left(k_i \mid x_i, d_i = 0\right) + \log \mathcal{L}_{II}\left(k_i \mid x_i, d_i = 1\right) \\
& + \log \mathcal{L}_{III}\left(k_i \mid x_i, d_i = 0\right) + \log \mathcal{L}_{III}\left(k_i \mid x_i, d_i = 1\right),
\end{aligned}
$$

since the 6 distinct cases are mutually exclusive, where

$$
\log \mathcal{L}_X\left(k_i \mid x_i, d_i\right) = \log\left[\sum_{p=1}^{P} \mathcal{L}_X\left(k_i \mid x_i, \xi_p, d_i\right) \times \pi_p\right],
$$

with (ξ_p, π_p) denoting the p^{th} mass point of the non-parametrically specified frailty. Note that the zero-mean frailty is thus specified by $2(P-1)$ free parameters, since $\sum_{p=1}^{P} \pi_p = 1$, which can be obtained by the reparametrization

$$
\pi_p = \begin{cases} \frac{\exp[\kappa_p]}{1 + \sum_{p=1}^{P-1} \exp[\kappa_p]} & p < P \\ \frac{1}{1 + \sum_{p=1}^{P-1} \exp[\kappa_p]} & p = P \end{cases},
$$

and since $\sum_{p=1}^{P} \xi_p \pi_p = 0$, which can be obtained by eliminating ξ_P

$$
\xi_P = -\sum_{p=1}^{P-1} \xi_p \exp\left[\kappa_p\right].
$$

In addition, in order for the $\lambda_{P0}(l) \in [0,1]$, with $P = J, W$, we reparametrize them as

$$
\lambda_{P0}(l) = \frac{\exp\left[\mu_{P0}(l)\right]}{\exp\left[\mu_{P0}(l)\right] + \exp\left[-\mu_{P0}(l)\right]},
$$

with inverse transform

$$\mu_{P0}(l) = \frac{1}{2}\ln\left[\frac{\lambda_{P0}(l)}{1-\lambda_{P0}(l)}\right].$$

From this we have that

$$\begin{aligned}
\frac{\partial\lambda_{P0}(l)}{\partial\mu_{P0}(l)} &= \frac{2}{\left(\exp\left[\mu_{P0}(l)\right]+\exp\left[-\mu_{P0}(l)\right]\right)^2}, \\
\frac{\partial^2\lambda_{P0}(l)}{\left(\partial\mu_{P0}(l)\right)^2} &= \frac{-4\left(\exp\left[\mu_{P0}(l)\right]-\exp\left[-\mu_{P0}(l)\right]\right)}{\left(\exp\left[\mu_{P0}(l)\right]+\exp\left[-\mu_{P0}(l)\right]\right)^3},
\end{aligned}$$

hence

$$\begin{aligned}
\frac{\partial F\left(\lambda_{P0}(l)\right)}{\partial\mu_{P0}(l)} &= \frac{\partial F\left(\lambda_{P0}(l)\right)}{\partial\lambda_{P0}(l)}\cdot\frac{\partial\lambda_{P0}(l)}{\partial\mu_{P0}(l)} \\
\frac{\partial^2 F\left(\lambda_{P0}(l)\right)}{\left(\partial\mu_{P0}(l)\right)^2} &= \frac{\partial^2 F\left(\lambda_{P0}(l)\right)}{\left(\partial\lambda_{P0}(l)\right)^2}\cdot\left(\frac{\partial\lambda_{P0}(l)}{\partial\mu_{P0}(l)}\right)^2+\frac{\partial F\left(\lambda_{P0}(l)\right)}{\partial\lambda_{P0}(l)}\cdot\frac{\partial^2\lambda_{P0}(l)}{\left(\partial\mu_{P0}(l)\right)^2} \\
\frac{\partial^2 F\left(\lambda_{P0}(l),\lambda_{P0}(l')\right)}{\left(\partial\mu_{P0}(l)\right)\left(\partial\mu_{P0}(l')\right)} &= \frac{\partial^2 F\left(\lambda_{P0}(l),\lambda_{P0}(l')\right)}{\left(\partial\lambda_{P0}(l)\right)\left(\partial\lambda_{P0}(l')\right)}\cdot\left(\frac{\partial\lambda_{P0}(l)}{\partial\mu_{P0}(l)}\right)\cdot\left(\frac{\partial\lambda_{P0}(l')}{\partial\mu_{P0}(l')}\right)
\end{aligned}$$

Further details can be found in Appendix A.

4 Results

A first issue that needs to be discussed is the choice of the intervals within which the hazard rates are constant. Since the data are observed with monthly intervals, this would be a natural choice. However, in order to avoid identification problems caused by empty cells, we choose one long final interval. In order to decide the length of this final interval, we computed monthly hazard rates (see figure 1) and their variances unconditional on any covariate. We then looked for the first month with less than 20 exits or with a lower 95% confidence limit < 0. The last interval starts at said firs month (and runs to the end of the observation period). For the exit to training, we thus obtain 21 intervals and for the exit to work 40.

Table 3 summarises parameters of interest[16]: the different effects of following a training on the exit to work and σ_J, the factor by which the waiting time frailty enters the unemployment duration equation. The left part of table 3 gives the results for a benchmark model without unobserved heterogeneity, the the right part gives the results for a model including a 3-point frailty. Focussing on the latter results, we conclude that there exists a significant locking-in of training: during the training the probability of exit to work diminshes with 8%. After a successful training, the probability of exit to work increases to 1.81 times the probability without training during the first six months after the training. The next six months it is even raised to 1.99 times, while the second year the effect seems to wear of slightly. After two years, the effect of the training seems to have returned to zero. Remarkably, even an unsuccesful training has a positive impact on the exit to work. The parameter σ_J, is significantly negative, meaning that people who start earlier with a training on average have a longer unemployment duration. The frailty thus does not capture unobserved motivation, but seems to reflect the fact that during the period of observation, the PES explicitly targeted individuals deemed weakly employable, who were less inclined to turn down training offers.

[16]The full tables of results can be found in Appendix B

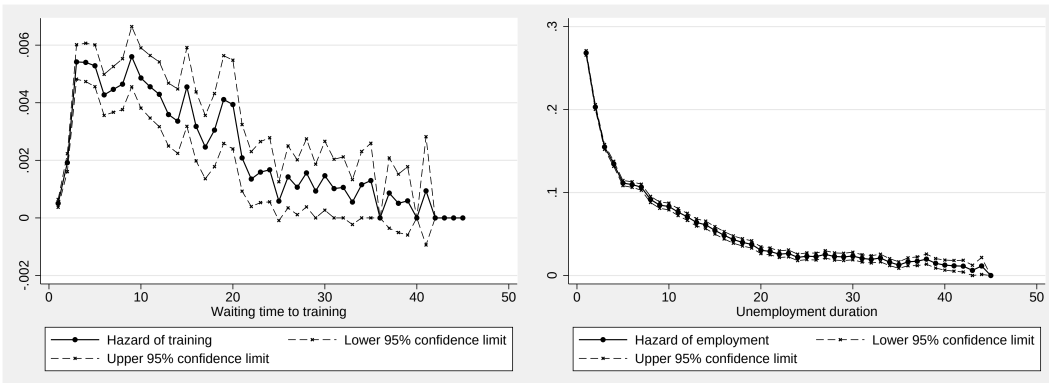

Figure 1: Non-parametric unconditional estimates of the hazards for the exit to training (left) and the exit to work (right)

	Piecewise Constant, no frailty				Piecewise constant, 3 point frailty			
	b	sd	p	factor	b	sd	p	factor
during training	-0.11	0.03	0.001	0.90	-0.08	0.03	0.009	0.92
after training, month $1-6$	0.58	0.06	0.000	1.79	0.60	0.06	0.000	1.81
after training, month $7-12$	0.68	0.11	0.000	1.98	0.69	0.11	0.000	1.99
after training, month $13-24$	0.52	0.17	0.003	1.68	0.53	0.17	0.002	1.69
after training, month ≥ 25	-0.26	0.71	0.713	0.77	-0.23	0.71	0.744	0.79
after unsuccessful training	0.24	0.09	0.006	1.27	0.25	0.09	0.005	1.28
σ_J	/	/	/	/	-0.05	0.01	0.000	/

Table 3: The effect of training on the probability of employment

5 Conclusion

In this paper we estimated the effect of training on the exit to work hazard using the timing of events approach. Next to a significant but small locking in effect, we find a positive effect of training on the exit to work probability, which is almost doubled[17] in the first year after the training and still 70% higher in the second year.

References

Abbring, J. H., van den Berg, G. J., 2003. The nonparametric identification of treatment effects in duration models. Econometrica 71 (5), 1491–1517.
URL `http://dx.doi.org/10.1111/1468-0262.00456`

Abbring, J. H., van den Berg, G. J., 2004. Analyzing the effect of dynamically assigned treatments using duration models, binary treatment models, and panel data models. Empirical Economics 29 (1), 5–20.

Attanasio, O., Kugler, A., Meghir, C., 2011. Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a randomized trial. American Economic Journal: Applied Economics 3 (3), 188–220.

Bollens, J., 2007. De meting van de effectiviteit van arbeidsmarktprogramma's. WSE Report 7-2007, HIVA - Katholieke Universiteit Leuven.

Bollens, J., 2014. Effectiviteit van opleiding voor werklozen. Over.Werk. Tijdschrift van het Steunpunt WSE 24 (2), 46–50.

Card, D., Ibarrarán, P., Regalia, F., Rosas-Shady, D., Soares, Y., 2011. The labor market impacts of youth training in the dominican republic. Journal of Labor Economics 29 (2), 267–300.

Card, D., Kluve, J., Weber, A., 2010. Active labour market policy evaluations: A meta-analysis. Economic Journal 120 (548), F452–F477.

Cho, Y., Kalomba, D., Mobarak, A. M., Orozco, V., 2013. Gender differences in the effects of vocational training: Constraints on women and drop-out behavior. World Bank Policy Research Working Paper 6545, World Bank.

Cockx, B., 2002. Les formations professionnelles du FOREM accélèrent-elles la sortie du chômage en Wallonie? In: Docquier, F., de La Croix, D., Mainguet, C., Perelman, S., Wasmer, E. (Eds.), Capital humain et dualisme sur le marché du travail. De Boeck Supérieur, Ch. 4, pp. 71–94.
URL `http://hdl.handle.net/1854/LU-1177412`

Crépon, B., Ferracci, M., Fougère, D., 2012. Training the unemployed in France: How does it affect unemployment duration and recurrence? Annals of Economics and Statistics / Annales d'Économie et de Statistique 107/108, 175–199.

De Rick, K., De Cuyper, P., 2014. Van een VDAB-opleiding naar werk? hefbomen voor tewerkstelling. Over.Werk. Tijdschrift van het Steunpunt WSE 24 (2), 51–57.

Eberwein, C., Ham, J. C., Lalonde, R. J., 1997. The impact of being offered and receiving classroom training on the employment histories of disadvantaged women: Evidence from experimental data. Review of Economic Studies 64 (4), 655–682.

Gaure, S., Røed, K., Zhang, T., 2007. Time and causality: A Monte Carlo assessment of the timing-of-events approach. Journal of Econometrics 141 (2), 1159–1195.

[17] Compared to the no training exit to work probability.

Greenberg, D. H., Michalopoulos, C., Robins, P. K., 2003. A meta-analysis of government-sponsored training programs. Industrial & Labor Relations Review 57 (1), 31–53.

Ham, J. C., Lalonde, R. J., 1996. The effect of sample selection and initial conditions in duration models: Evidence from experimental data on training. Econometrica 64 (1), 175–205.

Heckman, J., Navarro-Lozano, S., 2004. Using matching, instrumental variables and control functions to estimate economic choice models. Review of Economics and Statistics 86, 30–57, also (2003) NBER Working Paper, 9497.

Heckman, J., Singer, B., 1984. The identifiability of the proportional hazard model. Review of Economic Studies 51 (2), 231–241.

Heckman, J. J., Lalonde, R. J., Smith, J. A., 1999. The economics and econometrics of active labor market programs. In: Ashenfelter, O. C., Card, D. (Eds.), Handbook of Labor Economics. Vol. 3 of Handbook of Labor Economics. Elsevier, Ch. 31, pp. 1865–2097.

Hirshleifer, S., McKenzie, D., Almeida, R., Ridao-Cano, C., 2015. The impact of vocational training for the unemployed: Experimental evidence from Turkey. Economic Journal n/a (n/a), n/a–n/a.
URL http://dx.doi.org/10.1111/ecoj.12211

Hujer, R., Thomsen, S. L., Zeiss, C., 2006. The effects of vocational training programmes on the duration of unemployment in Eastern Germany. Allgemeines Statistisches Archiv – Advances in Statistical analysis 90 (2), 299–321.
URL http://dx.doi.org/10.1007/s10182-006-0235-z

Kalbfleisch, J. D., Prentice, R. L., 2002. The Statistical Analysis of Failure Time Data, 2nd Edition. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc.

Kluve, J., 2010. The effectiveness of European active labor market programs. Labour Economics 17 (6), 904–918.

Lalive, R., Van Ours, J. C., Zweimüller, J., 2008. The impact of active labour market programmes on the duration of unemployment in Switzerland. Economic Journal 118 (525), 235–257.
URL http://dx.doi.org/10.1111/j.1468-0297.2007.02111.x

Lancaster, T., 1990. The Econometric Analysis of Transition Data. Vol. 17 of Econometric Society Monographs. Cambridge University Press, Cambridge, UK.

Osikominu, A., 2013. Quick job entry or long-term human capital development? the dynamic effects of alternative training schemes. Review of Economic Studies 80 (1), 313–342.
URL http://restud.oxfordjournals.org/content/80/1/313.abstract

Richardson, K., van den Berg, G. J., 2013. Duration dependence versus unobserved heterogeneity in treatment effects: Swedish labor market training and the transition rate to employment. Journal of Applied Econometrics 28 (2), 325–351.
URL http://dx.doi.org/10.1002/jae.2263

A Derivatives of the Log-likelihood

The unconditional likelihood is obtained by mixing out the frailty

$$\mathcal{L}\left(k_i \mid x_i, d_i; \alpha\right) \quad = \quad \sum_{p=1}^{P} \mathcal{L}_J\left(k_i \mid x_i, \xi_p, d_i; \alpha_J\right) \cdot \mathcal{L}_W\left(k_i \mid x_i, \xi_p, d_i; \alpha_W\right) \times \pi_p$$

Since the cases we consider are mutually exclusive, it holds that

$$\mathcal{L}\left(k_i \mid x_i, d_i\right) \quad = \prod_{\substack{c = I, II, III \\ d = 0, 1}} \mathcal{L}_c\left(k_i \mid x_i, d_i\right),$$

where

$$\mathcal{L}_c\left(k_i \mid x_i, d_i\right) \quad = \quad \sum_{p=1}^{P} \mathcal{L}_{c;J}\left(k_i \mid x_i, \xi_p, d_i; \alpha_J\right) \cdot \mathcal{L}_{c;W}\left(k_i \mid x_i, \xi_p, d_i; \alpha_W\right) \times \pi_p$$

with $c = I, II, III$. Hence, we have that

$$\ln \mathcal{L}\left(k_i \mid x_i, d_i; \alpha\right) \quad = \sum_{\substack{c = I, II, III \\ d = 0, 1}} \ln \mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right),$$

where

$$\ln \mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right) \quad = \quad \ln \left[\sum_{p=1}^{P} \mathcal{L}_{c;J}\left(k_i \mid x_i, \xi_p, d_i; \alpha_J\right) \cdot \mathcal{L}_{c;W}\left(k_i \mid x_i, \xi_p, d_i; \alpha_W\right) \times \pi_p\right].$$

A.0.1 Gradient

So, in general, we have that

$$\frac{\partial \ln \mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right)}{\partial \alpha_J'} \quad = \quad \left\{\mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right)\right\}^{-1} \cdot \left\{\sum_{p=1}^{P} \frac{\partial \mathcal{L}_{c;J}\left(k_i \mid x_i, \xi_p, d_i; \alpha_J\right)}{\partial \alpha_J'} \cdot \mathcal{L}_{c;W}\left(k_i \mid x_i, \xi_p, d_i; \alpha_W\right) \times \pi_p\right\}$$

$$\frac{\partial \ln \mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right)}{\partial \alpha_W'} \quad = \quad \left\{\mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right)\right\}^{-1} \cdot \left\{\sum_{p=1}^{P} \mathcal{L}_{c;J}\left(k_i \mid x_i, \xi_p, d_i; \alpha_J\right) \cdot \frac{\partial \mathcal{L}_{c;W}\left(k_i \mid x_i, \xi_p, d_i; \alpha_W\right)}{\partial \alpha_W'} \times \pi_p\right\}$$

$$\frac{\partial \ln \mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right)}{\partial \pi_p} \quad = \quad \left\{\mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right)\right\}^{-1} \cdot \mathcal{L}_{c;J}\left(k_i \mid x_i, \xi_p, d_i; \alpha_J\right) \cdot \mathcal{L}_{c;W}\left(k_i \mid x_i, \xi_p, d_i; \alpha_W\right)$$

$$\begin{aligned}
\frac{\partial \ln \mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right)}{\partial \xi_p} \quad &= \quad \left\{\mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right)\right\}^{-1} \cdot \left\{\frac{\partial \mathcal{L}_{c;J}\left(k_i \mid x_i, \xi_p, d_i; \alpha_J\right)}{\partial \xi_p} \cdot \mathcal{L}_{c;W}\left(k_i \mid x_i, \xi_p, d_i; \alpha_W\right)\right. \\
&\qquad \left. \cdot\, \mathcal{L}_{c;J}\left(k_i \mid x_i, \xi_p, d_i; \alpha_J\right) \cdot \frac{\partial \mathcal{L}_{c;W}\left(k_i \mid x_i, \xi_p, d_i; \alpha_W\right)}{\partial \xi_p}\right\} \cdot \pi_p \\
&= \quad \left\{\mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right)\right\}^{-1} \cdot \left\{\frac{\partial \mathcal{L}_{c;J}\left(k_i \mid x_i, \xi_p, d_i; \alpha_J\right)}{\partial \mathcal{I}_{J;i}} \cdot \sigma_J \cdot \mathcal{L}_{c;W}\left(k_i \mid x_i, \xi_p, d_i; \alpha_W\right)\right. \\
&\qquad \left. \cdot +\mathcal{L}_{c;J}\left(k_i \mid x_i, \xi_p, d_i; \alpha_J\right) \cdot \frac{\partial \mathcal{L}_{c;W}\left(k_i \mid x_i, \xi_p, d_i; \alpha_W\right)}{\partial \mathcal{I}_{W;i}}\right\} \cdot \pi_p
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \ln \mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right)}{\partial \sigma_J} \quad &= \quad \left\{\mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right)\right\}^{-1} \cdot \left\{\sum_{p=1}^{P} \frac{\partial \mathcal{L}_{c;J}\left(k_i \mid x_i, \xi_p, d_i; \alpha_J\right)}{\partial \sigma_J} \cdot \mathcal{L}_{c;W}\left(k_i \mid x_i, \xi_p, d_i; \alpha_W\right) \times \pi_p\right\} \\
&= \quad \left\{\mathcal{L}_c\left(k_i \mid x_i, d_i; \alpha\right)\right\}^{-1} \cdot \left\{\sum_{p=1}^{P} \frac{\partial \mathcal{L}_{c;J}\left(k_i \mid x_i, \xi_p, d_i; \alpha_J\right)}{\partial \mathcal{I}_{J;i}} \cdot \xi_p \cdot \mathcal{L}_{c;W}\left(k_i \mid x_i, \xi_p, d_i; \alpha_W\right) \times \pi_p\right\}.
\end{aligned}$$

We further have to take into account the re-parametrization $\pi_p \to \kappa_q$ and the elimination of ξ_P

$$\xi_P = -\sum_{p=1}^{P-1} \xi_p \exp\left[\kappa_p\right].$$

B Full Estimation Results

Equation	Variable	b	sd	p
Exit to training	Sex	0.04	0.05	0.434
	Age (standardized)	0.37	0.03	0.000
	Age squared	-0.64	0.03	0.000
	Months unemployed (past 2 years)	-0.01	0.00	0.095
	Primary education	0.58	0.14	0.000
	Secondary education	0.48	0.15	0.001
	Tertiary education: college	0.31	0.14	0.029
	Educational track: general	0.09	0.10	0.337
	Educational track: technical	0.18	0.08	0.023
	Educational track: vocational	0.12	0.07	0.087
	Vlaams Brabant	0.16	0.08	0.032
	West Vlaanderen	0.35	0.07	0.000
	Oost Vlaanderen	0.20	0.06	0.001
	Limburg	0.22	0.07	0.002
	Drivers license	0.18	0.06	0.003
	Language Dutch	0.75	0.07	0.000
	Belgian nationality	0.24	0.09	0.010
Exit to work	Sex	-0.10	0.01	0.000
	Age (standardized)	-0.21	0.00	0.000
	Age squared	-0.07	0.00	0.000
	Months unemployed (past 2 years)	0.00	0.00	0.449
	Primary education	-0.11	0.01	0.000
	Secondary education	-0.08	0.02	0.000
	Tertiary education: college	0.08	0.01	0.000
	Educational track: general	-0.07	0.02	0.000
	Educational track: technical	0.00	0.01	0.989
	Educational track: vocational	0.02	0.01	0.065
	Vlaams Brabant	0.02	0.01	0.142
	West Vlaanderen	0.18	0.01	0.000
	Oost Vlaanderen	0.03	0.01	0.004
	Limburg	0.07	0.01	0.000
	Drivers license	0.22	0.01	0.000
	Language Dutch	0.26	0.01	0.000
	Belgian nationality	0.08	0.01	0.000
Training effects	during training	-0.11	0.03	0.001
	after training, month $1-6$	0.58	0.06	0.000
	after training, month $7-12$	0.68	0.11	0.000
	after training, month $13-24$	0.52	0.17	0.003
	after training, month ≥ 25	-0.26	0.71	0.713
	after unsuccessful training	0.24	0.09	0.006

Table 4: Results PWC without frailty, baseline hazard parameters not shown

Equation	Variable	b	sd	p
Exit to training	Sex	0.04	0.05	0.430
	Age (standardized)	0.38	0.03	0.000
	Age squared	-0.65	0.03	0.000
	Months unemployed (past 2 years)	-0.01	0.00	0.062
	Primary education	0.59	0.14	0.000
	Secondary education	0.48	0.15	0.001
	Tertiary education: college	0.32	0.14	0.028
	Educational track: general	0.11	0.10	0.272
	Educational track: technical	0.17	0.08	0.033
	Educational track: vocational	0.12	0.07	0.098
	Vlaams Brabant	0.16	0.08	0.034
	West Vlaanderen	0.34	0.07	0.000
	Oost Vlaanderen	0.20	0.06	0.001
	Limburg	0.23	0.07	0.002
	Drivers license	0.19	0.06	0.002
	Language Dutch	0.76	0.07	0.000
	Belgian nationality	0.25	0.09	0.008
Exit to work	Sex	-0.10	0.01	0.000
	Age (standardized)	-0.21	0.00	0.000
	Age squared	-0.07	0.00	0.000
	Months unemployed (past 2 years)	0.00	0.00	0.434
	Primary education	-0.11	0.01	0.000
	Secondary education	-0.08	0.02	0.000
	Tertiary education: college	0.08	0.01	0.000
	Educational track: general	-0.07	0.02	0.000
	Educational track: technical	0.00	0.01	0.985
	Educational track: vocational	0.02	0.01	0.065
	Vlaams Brabant	0.02	0.01	0.141
	West Vlaanderen	0.18	0.01	0.000
	Oost Vlaanderen	0.03	0.01	0.005
	Limburg	0.07	0.01	0.000
	Drivers license	0.22	0.01	0.000
	Language Dutch	0.26	0.01	0.000
	Belgian nationality	0.08	0.01	0.000
Training effects	during training	-0.08	0.03	0.009
	after training, month $1-6$	0.60	0.06	0.000
	after training, month $7-12$	0.69	0.11	0.000
	after training, month $13-24$	0.53	0.17	0.002
	after training, month ≥ 25	-0.23	0.71	0.744
	after unsuccessful training	0.25	0.09	0.005
Frailty parameters	Probability 1	2.46	0.31	0.000
	Probability 2	7.92	0.32	0.000
	Value 1	-2.51	0.00	0.000
	Value 2	0.01	0.00	0.000
	σ_J	-0.05	0.01	0.000

Table 5: Results PWC with 3 point frailty, baseline hazard parameters not shown