"Active labour market policies and job tenure"[1],

Bart Cockx[2] ,
Bruno Van der Linden[3],
Adel Karaa[4]

November 1996

ABSTRACT

In this paper, we study the effect of subsidised on-the-job training, training for the unemployed and pure wage subsidies on the probability of leaving an employer. We base the analysis on a sample of unemployed workers who have been hired during the 1991-92 period. Since individuals benefiting from the policy were over represented in the sample, we face an endogenous sampling problem apart from the well known selectivity problem in evaluation analysis. The analysis deals with these two issues simultaneously. We find that each of the labour market policies increases the length of job tenure. Yet, in line with the literature (Lalonde 1986, Fraker and Maynard 1987), the magnitude and the significance of this effect depend crucially on the parametric assumptions in the model. Nevertheless, one robust conclusion is that subsidised on-the-job training schemes significantly decrease the incidence of job termination. We claim that this result supports Stevens' (1994, 1996) hypothesis of *transferable* training and consequently of the underprovision of training by the market. Training programs for the unemployed and pure wage subsidies only have an important positive but non significant effect on job tenure. We argue that the stated effect of training programs provides weak support for human capital theories as opposed to matching theories in the explanation of job turnover.

---

[1] While the research is joint work of the three individuals mentioned above, the wording has been carried out by Bart Cockx and Bruno Van der Linden. The authors are to be accounted for any errors made. This research has benefited from a grant ERB CHB GCT 940681 from the Human Capital and Mobility Program of the European Commission. This research has also been supported by a grant "Actions de Recherches Concertées" n° 93/98-162 of the Ministry of Scientific Research of the Belgian French Speaking Community. In addition, research funding has been provided by the "Services Fédéraux des Affaires Scientifiques, Techniques et Culturelles" under contract n° SE/10/027. This paper has benefited from the comments of I. Bardoulat.

[2] Institut de Recherches Economiques et Sociales, Département des Sciences Economiques, Université Catholique de Louvain, Belgium.

[3] Fonds National de la Recherche Scientifique and Institut de Recherches Economiques et Sociales, Département des Sciences Economiques, Université Catholique de Louvain, Belgium.
Tel : 32.10.47.34.33. Fax : 32.10.47.39.45. E-mail : vanderlinden@ires.ucl.ac.be.

[4] Human Capital and Mobility fellow at the Institut de Recherches Economiques et Sociales during the research presented in this paper.

1. Introduction

In this paper we study the effect of subsidised on-the-job training, classroom training for the unemployed and pure wage subsidies to employers (i.e. without a compulsory training content). We analyse the effect of these programs on the probability of leaving an employer. The empirical analysis is based on a sample of workers who were hired from the unemployment pool.

Studying the effect of labour market programs on job turnover may improve our understanding of the functioning of the labour market. In particular, it can throw some light on the relative importance of human capital and matching theories in explaining job tenure and on the importance of skill mismatch as an explanation of unemployment. It can also clarify to what extent government intervention in on-the-job training programs can improve allocative efficiency. We start with a discussion of the latter.

Recently, Stevens (1994, 1996) argued that all training cannot be considered as the sum of a general and a specific component. She introduces the notion of *transferable training.* Instead of being useful to either all firms in the economy or only to one specific firm, this type of training is useful to a limited number of firms, and usually to an unequal degree. Stevens argues that human capital theory typically disproves that the market under provides training (see Stevens, 1996, p. 21-22). She shows, however, that this argument crucially depends on whether training is a mix of general and specific human capital (Becker, 1962, 1975). If instead training is *transferable,* Stevens argues that the labour market may be imperfectly competitive, leading to socially suboptimal investment in human capital.

A first objective of this paper is to test the validity of Stevens's argument in the following way. Any firm typically provides a certain amount of training to its workers. Assume a sample of occupied workers where this amount of training is unobserved, but where we know whether a worker has benefited from a *subsidised* on-the-job training or not. If on-the-job training is a mix of general and specific training, then any subsidy supporting that type of training is unlikely to sort any effects on job-turnover. For the subsidy is unlikely to influence the amount of training provided to a given worker. This assertion is valid if the investment required to be eligible to the subsidy is lower than the optimal level in the absence of the subsidy or if the firm and the worker can comply with the formal requirements (e.g. a certain duration of training) without adapting the genuine investment in human capital. On the contrary, if on-the-job training is of the *transferable* type, the subsidy will probably increase the suboptimal level of investment and therefore job tenure.

According to human capital theory, the extent to which job tenure is increased crucially depends on the nature of the training investment. If the training is firm-specific (on-the-job) or weakly *transferable*, job tenure will increase more than if it is general (in a classroom). As the acquired productivity increase is specific to the firm, the wage will rise relative to the alternative wage. The probability of leaving an employer will therefore fall[5]. In addition, employers will be less likely to lay off workers in whom they have invested in specific skills. If, on the other hand,

[5] Note that the argument is unconditional on the present wage. This accommodates to the empirical study of this paper as well as to the studies of Lynch (1991) and Elias (1994). The relationship is less clear-cut if the effect of (on-the-job) training is considered conditional on the present wage (cf. Devine and Kiefer, 1991, p. 230-31).

training is general, the increased productivity is useful to all firms and the worker has less or no incentives to stay on the same job. The same distinction holds *mutatis mutandis* as far as weakly *transferable* and widely *transferable* training are concerned. If these arguments are correct, one would expect that on-the-job training favours more job tenure than classroom training for the unemployed, the latter being more *transferable.* Testing this hypothesis is a second objective of the paper.

Matching theories (Spence, 1973, Stiglitz 1975, Jovanovic, 1979a, 1979b, 1984) provide an alternative explanation for the relationship between training and job tenure. According to these theories, training does not enhance the productivity of the worker but merely improves the quality of the match between the worker and the firm. The quality of the match is improved because participation to the training program signals information about the productivity of the worker that previously was private information. If participation signals that the trainee is more productive than the average worker, training will typically reduce the rate of job turnover. This is likely to occur if workers can freely choose to participate in the program and/or if program participation is only determined by easily observable characteristics such as age, unemployment duration or diploma. On the other hand, if the selection of trainees favours workers whose characteristics are negatively correlated with productivity (see e.g. Burtless, 1985 or Dubin and Rivers, 1993) and if these characteristics are not readily observable to employers, then participation to the training program discloses that the worker is of low productivity, increasing therefore the job turnover of trainees.

According to matching theories, the relationship between training and job tenure can therefore be positive or negative. Irrespective of the direction of the relationship, on-the-job training programs as compared to classroom programs are likely to intensify the relationship, since obviously more information will be disclosed if it is the employer himself who operates the program (Barron et al, 1989).

The validity of the matching theories as opposed to human capital theories can only be tested therefore if participation to training signals low productivity. For then the two theories predict an opposite relationship between participation and job tenure. We will argue below that participation to classroom training for the unemployed might indeed signal low productivity. A third objective of the paper is therefore to test whether matching theories or human capital theories turn out to be more in accordance with our data.

Until now we have been silent about the potential impact of pure wage subsidies to employers on job turnover. We now turn to a discussion of this effect. It will appear that it can provide evidence on the relative importance of skill mismatch as an explanation of unemployment.

Even if it is widely accepted that employment is enhanced during the period in which the subsidy is granted, one might question whether a temporary subsidy can have a long run impact on employment. On the one hand we have authors, such as Layard, Nickell and Jackman (1991), who emphasise the role of genuine duration dependence in any explanation of the evolution of European unemployment exit rates during the last decades. Genuine duration dependence occurs if, for example, an increase in unemployment duration creates some loss of accumulated human capital or generates a discouraged worker effect. In this case, a temporary wage subsidy and/or training program can induce permanent effects as it inverts the process of depreciation and discouragement. On the other hand, authors such as Sneessens and Shadman-Mehta (1995) argue that the causes of unemployment are more structural to the disadvantage of low skilled workers. In view of these models of unemployment, a temporary subsidy program with little training content cannot sort permanent effects, as there exists a fundamental mismatch

between the characteristics of jobs and the qualifications of the unemployed. It is therefore possible to discriminate between these two theories to the extent that we find a significantly positive effect of a temporary employment subsidy on the length of the job spell after the subsidised period. The available data do unfortunately not disentangle the period during which the worker is subsidised from the after-subsidy spell. Hence, the results provided by this paper will necessarily be tentative.

One of the major problems in estimating the effects of the mentioned labour market policies is to correct adequately for the selection bias. The selection bias arises whenever participants to the labour market program are non-randomly selected. For, in this case a differential impact of treatments as compared to controls may merely reflect differences in (un)observed characteristics.

Most commonly one tries to control for the selection bias in a non-experimental setting by modelling jointly the participation decision and its impact on the variable of interest, such as earnings or the exit rate out of (un)employment. Recent research (Lalonde, 1986, Fraker and Maynard, 1987) has thrown doubt on the capacity of non-experimental methods to correct for selection bias. Estimates are found to be sensitive to model specification and estimation method. This line of research asserts that the selection bias can only be controlled for if one disposes of data emerging from an experiment, be it controlled or natural (cf. Angrist, 1992, Meyer, 1995).

This paper remains in the tradition of the non-experimental literature. Heckman and Hotz (1989) defend this approach to the extent that adequate specification tests allow to discard the inappropriate models. In this paper we follow this approach and test for the sensitivity of the results to the model specification. However, due to data limitations and to the computational complexity of the optimisation problem, the sensitivity analysis is limited in nature.

Despite the widespread use of these active labour market policies, there exists relatively little micro-econometric evidence on the effects of these measures. Most progress has been made in the evaluation of training programs. However, most studies concentrate on the effect of training on subsequent earnings (cf. e.g. Ashenfelter, 1978, Bassi, 1984, Ashenfelter and Card, 1985, Barnow, 1987, Edin, 1988, Björklund, 1990, Ackum, 1991, Jensen *et al.*, 1993) or on reemployment probabilities (cf. e.g. Greenhalgh and Stewart, 1987, Main and Shelly, 1990, Main, 1991, Mealli *et al.*, 1996). Only recent studies have considered the effect on the length of the subsequent employment spell as well (cf. e.g. Kaitz, 1979, Ridder, 1986, Card and Sullivan, 1988, Ham and Lalonde, 1991, Gritz, 1993, Bonnal, Fougère and Sérandon, 1994, Torp, 1994, Zweimüller and Winter-Ebmer, 1996). One should distinguish employment duration from the length of job tenure, the variable of interest in this paper. To the extent that there are job-to-job transitions without an intervening spell out of employment, the two concepts differ. Lynch (1991) and Elias (1994) estimate the effect of on-the-job training on the probability of job termination, the variable of interest of this paper. Booth and Satchell (1994) do the same in the case of apprenticeships. These authors do not correct for selection bias, however.

We now turn to a brief discussion of this empirical evidence. We focus on the literature studying job tenure rather than employment duration. It should be emphasised that the literature has not yet adopted an homogeneous vocabulary. In this overview, job-related and private sector (respectively, off-the-job and government sector) training programs have been assimilated with on-the-job (respectively, classroom) training programs. Lynch (1991), Elias (1994) and Booth and Satchell (1994) find that on-the-job training lengthens the job spell. Elias

(1994), however, reports an insignificant impact for men. Regarding classroom training, Lynch (1991) reports an increased turnover for participants in classroom training programs. One should keep in mind, however, that none of the mentioned studies corrected for the selectivity bias. As far as we know, there is no empirical analysis of the effect of pure wage subsidies on job tenure[6].

We now turn to a description of the programs and the institutional context. Section 3 describes the data as well as the sampling scheme. In Section 4 we present the statistical model. Section 5 presents the results and Section 6 concludes.

2. A Description of the Programs and the Institutional Context

Table 1 presents some basic information about expenditures on labour market policies in Belgium and in some neighbour countries. According to OECD data, the total amount spent on active labour market policies is in % of GDP similar in Belgium, France and The Netherlands. Yet, it is much lower than in Germany. In Belgium, although rising, the relative importance of training programs for the unemployed is comparatively low. Conversely, public expenditures on recruitment subsidies and on training programs for occupied adults are relatively high in Belgium. This contrasts with the importance of long-term unemployment and low-skill unemployment in Belgium.

In Belgium, various institutions organise training programs for the unemployed. In accordance with the literature, a distinction should be made between subsidised on-the-job training and training schemes that are not directly related to a specific job in a given firm, designated below simply by the label 'training'. In both cases, given the relatively low number of training slots compared to the level of unemployment, unemployed applicants are selected on the basis of various characteristics (e.g. educational attainment). On-the-job training can be subsidised by the public authorities if the trainee is hired from the unemployment pool. Essentially two such schemes were in action during the 1991-1992 period and are still implemented today. In the first one (the "convention emploi-formation", henceforth CEF), a young low-skilled unemployed has to be hired for an indeterminate duration and during a period that ranges between one and three years he or she has to participate in a training program for at least 240 hours per year. This program is not organised by the firm but it must be closely related to the job. During this period of one to three years, social insurance contributions are only collected on the part of the wage which exceeds minimum wages. In the second case (the "formation professionnelle individuelle en entreprise", henceforth FPI), an unemployed is selected by a firm and trained on-the-job. A training contract is signed and its implementation is controlled by the Employment Agency. The training period lasts typically six months. During this period, the trainee keeps his unemployment benefit and the firm pays only an increasing share of the difference between the normal wage and the unemployment benefit. Meanwhile, no social insurance contributions are paid. If the trainee completes the program successfully, the firm has to hire this worker for a period at least as long as the duration of the training program.

Another category of subsidies, called pure wage subsidies, do not include these training requirements. During the 1991-1992 period, there existed six different pure wage subsidy schemes in Belgium. All of them were temporary wage subsidies paid to the employer. Yet, they differed in various ways : the type of contract that had to be signed (fixed-term or not); the

[6] Bonnal *et al.* (1994) and Ridder (1986) study the effect on employment duration.

categories of unemployed and of firms eligible and the wage subsidy level[7]. In Belgium, pure wage subsidies are often targeted on specific disadvantaged groups defined on the basis of criteria such as age, skill and unemployment duration. With one exception, these pure wage subsidies were paid on a quarterly or monthly basis. The maximum duration of payment of the subsidy varied between 12 and 24 months.

3. The Data and the Sampling Scheme

This analysis uses a sample of workers based on a survey of employers. This section explains first how the sample was selected. This is important, because the selection rule was based on endogenous variables influencing therefore the statistical model presented in Section 4. Second, we will describe the information available in our sample. Before we proceed, notice that we henceforth denote by 'wage subsidies' both subsidised on-the-job training and pure wage subsidies (i.e. without a training content).

The survey on which this research is based was designed for another purpose. It aimed at collecting direct evidence on the importance of dead-weight and substitution effects of wage subsidies and training programs. This was realised on the basis of interview responses given by employers on recently hired workers. As this data set contains (imperfect) information on the length of job tenure, we have exploited this data source differently in order to estimate the impact of active labour market measures on the length of job tenure.

Since the labour market policies considered were often targeted on specific groups of workers, it was not a good idea to take a random sample of employers: In order to observe a sufficient number of beneficiaries of the policies, the required sample size would have been too large and therefore too costly. The budget allowed only about 400 firms, spread all over Belgium, to be interviewed. For this reason, first, a stratified random sample was taken of employers that were registered at the social insurance administration. Subsequently, firms were contacted by phone. To be sampled a firm had to satisfy a number of criteria (see below), aimed at over representing firms that hired workers from a target group defined below. If the firm met these criteria, the interviewer asked for an appointment with its human resource manager[8]. In case of a positive answer, a face-to-face interview was held. Interviews were carried out between May and July 1993. The questionnaire raised questions on the firm and on (at most) five hirings during the 1991-1992 period. We now detail more on the sampling of firms, on the selection of individual hirings and on the available information in the questionnaire.

Firms were selected in the following way:

- First, the sample was restricted to private firms of the service, manufacturing and construction sectors with more than five employees. Firms of the public sector were not considered, because they were typically not entitled to the wage subsidies. Smaller firms were left aside because of their idiosyncrasies.
- The employers were stratified in nine strata. In one dimension the three sectors, service, manufacturing and construction, were distinguished. In the other dimension the strata were defined according to firm size: 5 to 19 workers, 20 to 499 and 500 or more workers. A target number of interviews was fixed for each stratum.

---

[7] The wage subsidy levels ranged from about 10% of the wage cost and more than 50% in particular cases (like disabled workers).

[8] In firms with no human resource manager, the person responsible for hiring and layoff decisions was contacted.

- A firm was selected for interview if at least one of the following criteria were met during the 1991-1992 period : (i) to have hired jobless job-searchers who participated in a training program (during their previous unemployment spell); (ii) to have hired unemployed people with a wage subsidy; or (iii) to have recruited people from the 'hard core' of unemployed[9] who did not participate in a training program and whose hire had not been subsidised. The design of the sample required that the total sample was approximately equally divided among these three categories. In practice, many firms shared two or three of these characteristics. This is emphasised in Figure 1.

Next, the interviewer asked each human resource manager to choose (randomly) five recruitments[10] that occurred in 1991 or 1992[11]. The following constraints were imposed on this selection. First, the selected workers should have been hired from the pool of unemployed (including welfare recipients). Second, recruitments from the 'target group' were over represented. The target group consisted of workers who either had been trained during their previous unemployment spell, or for whom the employer obtained a wage subsidy, or who belonged to the 'hard core' of unemployed. Since Belgian wage subsidies are often restricted to categories of the 'hard core' of unemployed, this third group of workers has some features of a comparison group. The human resource manager was asked to select (if possible[12]) at least three recruitments from this target group.

This description clarifies that we face two endogenous sampling problems. First, we only have information on hired workers. Unemployed workers with a low hiring probability are therefore under represented in the sample. Consequently, to the extent that program participation increases (decreases) the hiring probability, participants will be over (under) represented in the sample and this may bias the estimator of the program effect, unless the statistical model explicitly takes this feature into account[13]. Secondly, it is clear that the sample over represents workers belonging to the 'target group' and therefore workers benefiting from the policy measures. A further discussion of these problems is delayed to the next section.

We now turn to a description of the available information as provided by the human resource manager. First, it should be stressed that we are only imperfectly informed about the length of the job tenure. For, we only know the month in which the employee was hired and whether he or she still occupied a position in the firm at the end of March 1993. Therefore, if someone left before March 1993, we only know that the job duration is lower than the number of months elapsed between the recruitment date and March 1993. In that case, we ignore whether the exit is a quit or a layoff. Nor do we know whether the individual is subsequently employed. If someone is still occupied at the end of March 1993, the job duration is right censored.

---

[9] Rather conventionally, in Belgium, someone is a member of this 'hard core' if he is a young low-skilled worker (less of 25 years old and having at most successfully completed lower secondary education), a long-term unemployed, a disabled jobless or someone entitled to the Minimum Income Guarantee (i.e. a welfare recipient).

[10] Or all the recruitments if their number was lower than five.

[11] Some interviewees selected hirings that occurred earlier. Given the relatively small size of our sample, we decided to consider them, too.

[12] There were firms who did not hire that many workers from the target group within the prescribed period.

[13] See also Ham and Lalonde (1991) for a discussion of this problem. Note that Torp (1994), and Zweimüller and Winter-Ebmer (1996) have a similar sampling scheme. They do not discuss the potential danger of bias, however.

Table 2 lists the characteristics of the sampled workers and of the firms in which they were hired. The reported summary statistics of the firm characteristics are calculated with respect to the sampled individuals and not with respect to the sampled firms. Note that, in view of the non-random nature of the sample, these statistics are a reflection of the sample and not of the Belgian economy.

As explained in Section 2, the various types of training schemes and wage subsidy programs are aggregated in three : training, subsidised on-the-job training and pure wage subsidy. It is important to realise two characteristics of this information. First, as to on-the-job training, we are only informed on the presence of this type of training to the extent that it was subsidised. Consequently, an individual trained on-the-job, but for whom the firm does not collect a subsidy, is regarded as someone who does not participate to any of the programs. This is crucial for a good understanding of our procedure proposed in the Introduction to test the validity of Stevens's transferable training hypothesis. Second, we need to be aware that only time-invariant indicators of the participation in these programs are available. This implies that, for pure wage subsidies, we are not able to distinguish between the probability of job termination during and after program participation. However, for the training program, we measure the post-program participation effect only, since training occurs by definition prior to employment and therefore prior to the selection into the sample. For subsidised on-the-job training we measure a mix of the effect during and post program participation, since for the CEF scheme introduced in Section 2 the hiring date, and therefore the instant at which the worker is sampled, coincides with the beginning of the training period; for the FPI scheme this date coincides with the beginning of the post-training period. We will come back to this issue when we interpret our results.

4. The Statistical Model

In this section we explain how we model job duration and how we will try to correct for the selection bias and for the selectivity introduced by the sampling scheme. First, to the extent that the mentioned selectivity is related to observable variables only, a correctly specified duration model conditioned on these observable covariates will yield consistent estimates of the program participation variables on the probability of job termination. This is the model we will present in the first subsection.

In the second subsection we propose a model that also corrects for selection on unobservables. This model is largely inspired by Gritz (1993), although that Gritz has far more complete information on the labour market history as we do. As is shown below, this requires a specification of the joint probability distribution of the job duration and of the participation decision. The 'participation decision' refers to the joint decision of the unemployed and the Employment Agency[14] to enter a training scheme, and to the joint decision of the firm, the job-searcher and the Administration in charge of the payment of the subsidies that leads to a subsidised employment contract.

In subsection 3 we discuss the adaptations required to allow for the selectivity induced by the endogenous sampling. We demonstrate why we are unable to correct for the selectivity induced

[14] During the 1991-1992 period, unemployed people typically asked for a training. Given the restricted number of training slots, they often waited quite a long time before they got the opportunity to start the training. The kind of criteria an unemployed had to satisfy in order to be accepted varied according to the type of training asked.

by the sampling of hired individuals only. We propose a method to correct for the bias introduced by over representing hires from the 'target group' (cf. Section 3).

Finally, we discuss our procedure of model selection. We justify the use of two criteria: Aikaike's information criterion and Jeffreys-Bayes posterior-probability criterion.

4.1 The Duration Model

For purposes of simplification, we will assume throughout that the stochastic process that has generated the data is stationary with respect to calendar time[15]. In Section 3 we explained that we only have incomplete information on the job duration. On the one hand, we observe for each worker i a non-stochastic variable d_i denoting the number of months that have elapsed between the hiring date and March 1993. On the other hand, we observe the realisation of a stochastic indicator variable J_i that is equal to one if the job duration is right censored at the end of March 1993 and equal to zero otherwise. So if T_i denotes the stochastic job duration of individual i, we have :

$$J_i = \begin{cases} 1 & if \quad T_i > d_i \\ 0 & otherwise \end{cases} \qquad (1)$$

It follows that the probability of observing J_i =1 conditional on d_i, i.e. the individual's contribution to the likelihood function, is equal to the survival probability :

$$prob(J_i = 1 | d_i) = prob(T_i > d_i) = \bar{F}_i(d_i) \qquad (2)$$

Similarly, contribution to the likelihood function of a worker for whom we observe J_i =0 conditional on d_i, is equal to one minus the survival probability :

$$prob(J_i = 0 | d_i) = prob(T_i \leq d_i) = 1 - \bar{F}_i(d_i) \qquad (3)$$

We now turn to the specification of the survival function. This amounts to specifying the hazard rate out of job tenure, as the survival function can be easily expressed in terms of these hazard rates (cf. Kiefer 1988 or Lancaster 1990). We will assume that the hazard for individual i at time t has the following proportional parameterisation :

$$h_i(t) = h_0(t) v \, exp(-z_i' \beta) \qquad (4)$$

where $h_0(t)$ is the baseline hazard at time t and v is the value of an unobserved individual specific effect. For the time being, we will assume that, conditional on the observable covariates z_i, neither T_i nor the unobserved specific effect are correlated with the participation decision into one of the three programs. This assumption will be relaxed in section 4.2.

---

[15] During the 1991-1992 period, both the employment level and the rate of vacancies were broadly constant in Belgium. Yet, the unemployment rate was increasing from 7.2% in 1991 to 7.7% in 1992 (according to OECD data).

To avoid too restrictive distributional assumptions (cf. Lalonde 1986 and Fraker and Maynard 1987), we choose the flexible piecewise constant specification of the baseline hazard as introduced by Prentice and Gloeckler (1978). For this purpose we divide the time axis into *K* intervals: $[t_0,t_1),...,[t_{k-1},t_k),...,[t_{K-1},t_K)$ of (possibly unequal) length $\Delta_k = t_k - t_{k-1}$ with $t_K = \max\{d_1,...,d_i,...,d_n\}$, where *n* is the number of individuals in the sample. By our assumption of stationarity and since we have a sample of individuals flowing into employment, we can normalise $t_0 = 0$. If we now follow Meyer (1990) and define the average integrated baseline hazard on the *k*th interval as

$$\theta_k = \int_{t_{k-1}}^{t_k} h_0(t) \Big/ \Delta_k \tag{5}$$

we can derive the following expression for the survivor function:

$$\bar{F}_i(t) = \exp\left\{-\nu \exp\left(-z_i'\beta\right)\left[\theta_k\left(t - t_{k-1}\right) + \sum_{m=1}^{k-1}\theta_m\Delta_m\right]\right\}, \qquad for \quad t_{k-1} \le t < t_k \tag{6}$$

Let $\phi(.)$ denote the density function of the standard normal distribution. If we assume that the unobserved individual specific effects are normally distributed with variance[16], σ^2, then Equations (2) and (3) allow us to derive the following expression for the likelihood function :

$$L(\varphi) = \prod_{i=1}^{n}\int_{-\infty}^{+\infty}\left\{\left[\bar{F}_i(d_i)\right]^{j_i}\left[1-\bar{F}_i(d_i)\right]^{(1-j_i)}\frac{1}{\sigma}\phi\left(\frac{\nu-1}{\sigma}\right)\right\}d\nu \tag{7}$$

where φ is an unknown parameter vector including σ^2. This equation together with equation (6) completes our specification of the likelihood function.

4.2 Correcting for Selection on Unobservables

Insofar as neither T_i nor the unobserved specific effect are correlated with the participation decision into one of the three programs, the maximisation of the likelihood function in (7) will yield a consistent estimate of the program effects of program participation. This is, however, not very realistic. We will therefore now allow this unobserved individual specific effect to be correlated with the participation decisions. On the other hand, we maintain the assumption that conditional on the observed and unobserved covariates, T_i is independent of the participation decision.

As in Gritz (1993), Bonnal et al. (1994) or Ham and Lalonde (1991) we introduce correlation between the unobserved effects by a one-factor specification (Flinn and Heckman, 1982). The advantage of this specification is its parsimony. This is important in view of the complexity of the optimisation problem (see below). Its disadvantage is that the pattern of correlation it allows is very restrictive. The correlations are either perfectly positive, negative or equal to zero.

For purposes of numerical tractability, we simplify the modelling of the participation decisions to the three programs. In the analysis we will not distinguish the participation decision in subsidised on-the-job training from the one in pure wage subsidies. We therefore distinguish

[16] The mean can be normalised to unity.

only participation in training and in a wage subsidy program. Participation in a training scheme will be denoted by the binary random variable Y_i^T. The binary random variable Y_i^S will designate whether a worker i is subsidised or not. Let p_{1i} be the probability $prob[Y_i^T=1]$ and p_{2i} be the probability $prob[Y_i^S=1]$. Since the decision to participate in a training program occurs before the one of signing a subsidised employment contract, we model the corresponding sequential processes through a nested logit specification:

$$p_{1i} = \frac{\exp(\eta_T v)\exp(z_{1i}'\gamma_1)}{1+\exp(\eta_T v)\exp(z_{1i}'\gamma_1)} \qquad \text{and} \qquad p_{2i} = \frac{\exp(\eta_S v)\exp(z_{2i}'\gamma_2)}{1+\exp(\eta_S v)\exp(z_{2i}'\gamma_2)} \tag{8}$$

where z_{1i} and z_{2i} are two vectors of time-invariant individual characteristics, both encompassed by z_i, and η_T and η_S are coefficients allowing the unobserved individual specific variables to affect the training, respectively wage subsidy decision differently than the hazard out of job tenure. Note that y_i^T is included in z_{2i}. The joint probability of participation is

$$prob(Y_i^T = y_i^T, Y_i^S = y_i^S | z_{1i}, z_{2i}, v) = p_{1i}^{\,y_i^T}(1-p_{1i})^{(1-y_i^T)}\, p_{2i}^{\,y_i^S}(1-p_{2i})^{(1-y_i^S)} \tag{9}$$

Ignoring the selectivity caused by endogenous sampling, the likelihood function therefore takes the following form :

$$L(\varphi) = \prod_{i=1}^{n}\int_{-\infty}^{+\infty}\left\{prob(Y_i^T = y_i^T, Y_i^S = y_i^S | z_{1i}, z_{2i}, v)\left[\overline{F}_i(d_i)\right]^{j_i}\left[1-\overline{F}_i(d_i)\right]^{(1-j_i)}\frac{1}{\sigma}\phi\left(\frac{v-1}{\sigma}\right)\right\}dv \tag{10}$$

4.3 Correcting for Selectivity due to Endogenous Sampling

In Section 3 we already alluded to the endogenous sampling problem. On the one hand, we highlighted the potential selectivity problem induced by only sampling hired workers. On the other hand, there is a potential bias due to the over representation of workers belonging to the 'target group' defined in Section 3. In this section we propose an estimation procedure to correct for the latter bias. We also sketch how to solve the former problem but argue that it is impossible to implement this approach with the present data.

Let S_i be a random binary variable that takes the value 1 if and only if individual i is sampled (i.e. is selected by the interviewee given the requirements of the questionnaire). Otherwise S_i equals zero. To take the sampling design into account, we now write the vector of time-invariant individual characteristics z_i as $[y_i^T, y_i^S, r_i, x_i]$ and supplement it with f_i, where r_i is an indicator equal to 1 if and only if individual i belongs to the 'hard core' unemployed, x_i is the vector of the other individual observable characteristics (including those of the employer), and f_i indicates the firm employing worker i. The joint probability of the participation process should then be written as $prob(Y_i^T = y_i^T, Y_i^S = y_i^S | S_i = 1, x_i, r_i, f_i, v)$. Strictly speaking, this probability is also conditional on the event that individual i has been hired. To simplify the notation, we will disregard this conditioning event for the time being. We will discuss the implications of this conditioning below. Applying Bayes formula leads to the following expression :

$$prob(Y_i^T = y_i^T, Y_i^S = y_i^S | S_i = 1, x_i, r_i, f_i, v) =$$

(11)

$$\frac{prob(S_i = 1 | y_i^S, y_i^T, r_i, f_i)\, prob(Y_i^T = y_i^T, Y_i^S = y_i^S | x_i, r_i, v)}{\sum_{y^S=0}^{1} \sum_{y^T=0}^{1} prob(S_i = 1 | y^S, y^T, r_i, f_i).prob(Y_i^T = y^T, Y_i^S = y^S | x_i, r_i, v)}$$

This shows that the influence of the sampling scheme is channelled through the term $prob(S_i = 1 | \ y_i^S, y_i^T, r_i, f_i)$. Note that in (11) we implicitly assume that the sampling is independent of x_i and v. We believe that this is a reasonable assumption for this sampling scheme (cf. Section 3). On the other hand, the sampling probability is explicitly conditioned on the firm employing the worker. This conditioning is important. For instance, the probability of hiring at least three individuals from the 'target group' will clearly be different between firms who have hired many workers from this group and those who have hired only a few of them.

Let us now precisely consider the consequences for the statistical analysis of over representing individuals out of the 'target group'. Let C_i be a random binary variable which takes the value 1 if and only if i is a member of the target group. Otherwise, C_i is zero. We will assume that, conditional on being (or not) sampled from the target group, i.e. conditional on $C_i = c_i$, the sampling is random. Consequently, we have that

$$prob\left(S_i = 1 | y_i^T, y_i^S, r_i, f_i\right) = prob\left(S_i = 1 | c_i, f_i\right) \qquad (12)$$

If $r_i = 1$, i.e. if i is hired from the 'hard core' of the unemployed, then $C_i = 1$. Conditional on $C_i = 1$, the sampling is random. Hence, the conditional probability of being sampled is equal across groups and expression (11) simplifies to :

$$prob(Y_i^T = y_i^T, Y_i^S = y_i^S | S_i = 1, x_i, r_i = 1, f_i, v) = prob(Y_i^T = y_i^T, Y_i^S = y_i^S | x_i, r_i = 1, v) \qquad (13)$$

and for these individuals the contribution of the likelihood corresponds to those defined in Section 4.2.

If $r_i = 0$, things are more complicated. For a given firm, f, let N_1^f (respectively, N_2^f) be the total number of hirings in (respectively, out of) the target group ($N^f = N_1^f + N_2^f$) during the 1991-92 period. If for this firm the sample is made of respectively n_1^f and n_2^f individuals, then neglecting the subscript i in f_i to avoid multiple levels of indices in the notation

$$prob\left(S_i = 1 | C_i = 1, f_i\right) = \frac{n_1^f}{N_1^f} \quad and \quad prob\left(S_i = 1 | C_i = 0, f_i\right) = \frac{n_2^f}{N_2^f} = \frac{n_2^f}{N^f - N_1^f} \qquad (14)$$

N^f is known and can be conditioned upon, but we only know that $N_1^f \in \left[n_1^f, N^f - n_2^f\right]$[17].

[17] If the human resource managers followed the interview instructions very precisely, we could have imposed the following a priori information : $\forall n_1^f < 3, N_1^f = n_1^f$. For, he was regarded to provide at least three individuals out of the target group, if possible. So, if he has provided information on less individuals, one could infer that no more individuals were hired from the target group in this firm. However, we do not utilise this prior

Hence,

$$prob\left(S_i = 1 \middle| C_i = 1, f_i\right) = \sum_{j=n_1^f}^{N^f - n_2^f} prob\left(S_i = 1 \middle| C_i = 1, N_1^f = j\right) prob\left(N_1^f = j\right)$$

$$= \sum_{j=n_1^f}^{N^f - n_2^f} \frac{n_1^f}{j} C_{N^f - n_2^f}^{j} p^j (1-p)^{N^f - n_2^f - j} \qquad (15)$$

where p is the probability of being in the target group in a given firm and where we used that the distribution of N_1^f, conditional on N^f, is binomial. Since the probability p is unknown, but varies over firms, we assume that p is distributed according to a Beta distribution $B(a,b)$ (i.e. p has a mean equal to $\frac{a}{a+b}$ and a variance equal to $\frac{ab}{(a+b+1)(a+b)^2}$)[18]. It follows that :

$$prob\left(S_i = 1 \middle| C_i = 1, f_i\right) = \frac{1}{\beta(a,b)} \sum_{j=n_1^f}^{N^f - n_2^f} \frac{n_1^f}{j} C_{N^f - n_2^f}^{j} \int_0^1 p^j (1-p)^{N^f - n_2^f - j} p^{a-1} (1-p)^{b-1} dp \qquad (16)$$

Similarly :

$$prob\left(S_i = 1 \middle| C_i = 0, f_i\right) = \frac{1}{\beta(a,b)} \sum_{j=n_2^f}^{N^f - n_1^f} \frac{n_2^f}{j} C_{N^f - n_1^f}^{j} \int_0^1 p^{N^f - n_1^f - j} (1-p)^j p^{a-1} (1-p)^{b-1} dp \qquad (17)$$

For those not hired from the 'hard core' of the unemployed (r_i = 0), these equations provide expressions for the probability of being sampled, conditional on being (or not) in the 'target group', and conditional on being employed in a firm that hired N^f workers during the 1991-92 period.

In order to study the sensitivity of the results to the specification of the distribution of individual specific effects, we will estimate the model presented in this section assuming both, a normal distribution and a discrete distribution with two points of support[19]. We provide in (18) the expression for the likelihood function under the latter assumption. The reader can easily find the expression of the likelihood function assuming a normal distribution.

$$L = \prod_{i=1}^{n} \sum_{m=1}^{2} q_m \left\{ \begin{array}{l} \left[\overline{F}\left(d_i \middle| x_i, r_i, y_i^S, y_i^T, \nu_m\right)\right]^{j_i} \left[1 - \overline{F}\left(d_i \middle| x_i, r_i, y_i^S, y_i^T, \nu_m\right)\right]^{1-j_i} \\ \left[prob\left(Y_i^T = y_i^T, Y_i^S = y_i^S \middle| x_i, r_i, \nu_m\right)\right]^{r_i} \\ \left[prob\left(Y_i^T = y_i^T, Y_i^S = y_i^S \middle| S_i = 1, x_i, r_i, f_i, \nu_m\right)\right]^{(1-r_i)} \end{array} \right\} \qquad (18)$$

---

information as to avoid a specification error whenever the resource managers did not follow up these instructions that precisely.

[18] This assumption is admittedly quite arbitrary. In order to test for the restrictiveness of this assumption, we estimated one of the models with a constant p and found the estimation results to be robust.

[19] Due to computational problems, we were not able to increase the number of points of support.

where $q_2 = 1 - q_1$ and $v_2 = (1 - q_1 v_1)/(1 - q_1)$ by normalisation. The two first bracketed terms are completely specified by Equation (6). The specification of the third bracketed term is given by (8) and (9). The last term is found, on the one hand by substituting (16) and (17), in (12) and subsequently in (11), and on the other hand by substituting (8) and (9) in (11).

We now turn to the discussion of the implication that only workers that have been hired are sampled. Consider an indicator variable, H_i, that is equal to one if individual i has been hired in the 1991-92 period and zero otherwise. By Bayes rule the probability of program participation is then given by the following expression:

$$prob\left(Y_i = y_i \middle| H_i = 1, x_i, r_i, v\right) = \frac{prob(H_i = 1 | y_i, x_i, r_i, v)\, prob\left(Y_i = y_i \middle| x_i, r_i, v\right)}{\sum_y prob(H_i = 1 | y, x_i, r_i, v)\, prob\left(Y_i = y \middle| x_i, r_i, v\right)} \qquad (19)$$

where $Y_i = \left[Y_i^T, Y_i^S\right]$. Clearly, neglecting the conditioning on hiring is only justified to the extent that program participation does not affect the probability of being hired : $prob(H_i = 1 | y, x_i, r_i, v) = prob(H_i = 1 | x_i, r_i, v)$. This is unlikely to be true. However, without any data on unemployed workers who were not hired, we cannot correct for this sampling selectivity. It can be argued that the neglect of this endogenous sampling biases the estimators of program participation downwards (cf. Section 5).

4.4 Criteria of Model Selection

The importance of specification testing has been underlined in the literature (Heckman and Hotz 1989). In this section we discuss criteria of model selection. We wish to investigate to what extent correcting for unobserved individual specific effects is important and to determine the optimal specification of the mixing distribution.

The standard likelihood ratio test is inappropriate when testing the relevance of unobserved heterogeneity. For, the parameter values (e.g. σ^2 of the normal distribution) lie on the boundary of the null hypothesis of no heterogeneity (e.g. σ^2=0). The problem is that this results in a violation of the regularity conditions required for the likelihood ratio to be asymptotically chi-squared. We follow Gritz's (1993) proposal to use Akaike's (1973) information criterion as a guide for model selection. The preferred model according to this criterion is the one that minimises the following statistic:

$$-2 \ln\left[L(\hat{\varphi})\right] + 2N_\varphi \qquad (20)$$

where $L(\hat{\varphi})$ is the optimised value of the likelihood function and N_φ is the number of estimated parameters.

Akaike's criterion is derived on the assumption that there exists a general model from which all competing models can be derived by imposing various restrictions on the vector of unknown parameters. When comparing a model with a normally distributed mixture to one in which the mixture is specified discrete with two points of support, this assumption is not satisfied. For these cases, we follow Chow's (1985) proposition to base model selection on the Jeffrey-Bayes

posterior-probability criterion. In particular, we will use Schwarz's (1978) approximation of this criterion:

$$-2\ln\left[L(\varphi)\right] + N_{\varphi}\ln(n) \qquad (21)$$

Again, the preferred model is the one that minimises this statistic.

5. The Empirical Results

In this section, we first explain how the number of time intervals has been selected. Next, we present the maximum likelihood estimates of the different statistical models presented in Section 4.

With a piecewise-constant specification of the hazard, the total number of intervals has to be chosen. We have adopted the following procedure. Neglecting individual observed and unobserved characteristics, we started with a large number of intervals: $K = 27$. Figure 2 shows that the hazard peaks in the first month, subsequently it decreases sharply and stays at a low level until the end of the first year of employment. From then on, it fluctuates around a higher level with peaks around an employment duration of one year, one year and a half and two years. The latter peaks did not prove to be statistically significant, however. On the basis of the likelihood ratio test we could reduce, at a significance level of 89%, the number of intervals to four (see Figure 2)[20].

Table 3 presents the results ignoring the selection and the endogenous sampling problems (cf. Section 4.1). A negative estimated parameter means that the corresponding variable increases the hazard rate out of the firm and hence reduces job duration. This table shows the estimated vector of parameters with and without unobserved heterogeneity. According to Akaike's Information Criterion the mixture model is to be preferred. From table 3, it appears that the young and the former long-term unemployed have a shorter tenure. The parameters of the firm's characteristics have the expected sign. As far as the policies are concerned, only subsidised on-the-job training has a significant positive effect on job duration.

Before we discuss the results of the model that takes the selection bias (but not the endogenous sampling bias) into account, we try to pinpoint the direction of the selection bias. From the discussion in Section 2, it is clear that the selection into subsidised employment with or without on-the-job training is mainly governed by the employer and the public authorities. First, consider the selection by the public authorities. In Belgium subsidies were predominantly targeted at the 'hard core' of unemployed. To the extent that we imperfectly controlled for this selectivity by including indicator variables such as 'long-term unemployed' or 'low-skilled worker', the measured effect tends to be underestimated. Secondly, given that it is the employer who determines selection in a subsidy scheme, it is clear that he will be ready to enrol less productive workers who are entitled to a subsidy. For, the wage subsidy can compensate for the lower productivity. To the extent that this lower productivity is correlated with characteristics that are not observed by the researcher, but observed by the employer, we therefore expect the selectivity bias to be negative. So, without correcting for the selectivity, we tend to underestimate the effect of the wage subsidies on job duration[21]. However, the selectivity bias

[20] The corresponding table is available upon request.

[21] For subsidised training programs (as opposed to the pure subsidies) matters are more complicated. For the training program permits the employer to select the most productive trainees. The training program therefore

will be smaller to the extent that dead-weight effects (i.e. cases where a subsidised worker would have been hired even in the absence of the wage subsidy) are important. For, in that case the employer mainly hires workers he would have hired anyway and who therefore will not be less productive.

This prediction is not necessarily true for hired trainees. For, the selection of unemployed individuals in a training program is less likely to be captured by the included explanatory variables, not in the least because participation depends on test scores which we ignore. So, to the extent that program administrators cream the most productive workers for participation in training, the selection bias may reverse. Creaming is unlikely to occur in Belgium, however. First, given the relatively low number of training slots, applicants have to wait rather a long time before they enter the training program. Only those with sufficiently bad unobserved characteristics will still be unemployed at that time. Second, in Belgium, financial rewards of training administrators are in no way related to placement ratios or to other labour market outcomes. Hence, program administrators have no financial incentive to cream the pool of applicants.

In Table 4 we report the results when the duration process is modelled jointly with the participation process into training and wage subsidy programs (cf. Section 4.2). A positive estimated parameter means that the corresponding variable increases the probability of entering a program. Yet, only the parameters of the duration process are of interest. Therefore, we henceforth focus on them. The age, long-term unemployment and firm's effects are not deeply modified. However, the three policies have now a much larger positive and significant effect on job duration. This confirms our prediction of a downward bias for the effect of wage subsidies and of training in Table 3. We now investigate whether there is any further bias induced by the endogenous sampling.

Before turning to the results that take the endogenous sampling problem into account (cf. Section 4.3), we again try to predict the direction of the bias. First, as workers from the 'target group' are over represented and since the 'target group' typically consists of less productive workers, we expect that a neglect of the endogenous sampling problem biases the effects of program participation further downwards. Secondly, we claim that even after correction for the latter bias, we will only obtain a lower bound of the estimated effect. This is a consequence of observing hired individuals only. For, if the program compensates for productivity differentials between workers, hired program participants will on average be less productive (in the absence of program participation) than hired non-participants[22].

In Table 5 the estimated effects of program participation are even larger than in Table 4, as expected. Moreover, recall that by conditioning on hired workers only, we still underestimate these effects. Before inferring any conclusion from these estimates, we need to investigate their robustness. Specifically, we investigated the sensitivity of these estimates to the distributional assumption on the mixing distribution. In Table 6, we tabulated the vector of parameters estimated under the assumption that the mixing distribution was discrete with two points of support instead of normal[23]. The labour market policies still turn out to enhance job tenure. Yet,

---

reverses the direction of the selection bias induced by the subsidy. Which of either biases will eventually dominate is impossible to predict.

[22] Note that if the program is a pure selection device the participants may rather be more productive prior to entry to the program than non-participants and the direction of the bias may therefore reverse. We suggest below, however, that the data are inconsistent with the pure matching theories.

[23] In order to simplify the computation, note that we no longer included the variables 'previously welfare

as compared to the results of Table 5, the estimated effects are much smaller and their estimated standard deviation have increased. Moreover, according to the criteria of model selection discussed in Section 4.4, the discrete mixture specification is the preferred one.

How should these results be interpreted? First, note that even if only the effect of subsidised on-the-job training is significantly different from zero, the point estimates of the effects of program participation are all important. Their insignificance reflects our difficulty to disentangle the true effect from the selectivity effect. In order to increase our intuition in the magnitude of the effects, we calculated for an individual with average characteristics in the sample the first quartile duration of job tenure with and without program participation. We chose to report the first quartile duration, i.e. the duration at which the first 25% of hired workers with average characteristics left the job, instead of median duration, because the latter required to extrapolate results beyond the observation period and risked therefore not to be very robust. The first quartile duration of a non-participant with average characteristics is estimated to be 13.8 months. According to the preferred specification (cf. Table 6), if he or she participates to a pure subsidy program this duration increases by 30% to 17.9 months, for on-the-job trainees by 68% to 23.2 months and to training by 15% to 15.9 months. This shows that participation indeed substantially increases the length of a job spell. Moreover, recall that we argued that by considering hired individuals only these estimates are lower bounds. However, the results have proven to be sensitive to the chosen specification and we cannot therefore guarantee that these estimates are the most conservative.

The effect of (classroom) training for the unemployed on job tenure is positive but insignificantly different from zero. We argued that participation signals low productivity. This result therefore provides some weak evidence in favour of human capital theories as opposed to matching theories. According to matching theories, the turnover of trainees should indeed have been higher instead of lower.

We find that participation in the pure subsidy program increases job duration quite dramatically. The effect is far from precisely determined though and therefore careful interpretation is required. Moreover, it is important to recall that our data do not allow us to disentangle the subsidised period from the after subsidy period. The measured effect could therefore just reflect that employers are less likely to sack subsidised workers than unsubsidised ones and that in the post-subsidy period the two types of workers leave the job at the same rate. This interpretation is to some extent supported by the analysis of Van der Linden (1995) based on the same sample as the present research. This paper uses the qualitative information available in the questionnaire to estimate the importance of dead-weight. The estimated frequency is about 60%. It is therefore plausible that the subsidy only has a minor and non significant effect on the after subsidy job tenure. This offers some support to the view that mismatch is a major explanation of unemployment in Belgium. But this is to a large extent a conjecture given the difficulty to disentangle the true effect from the selectivity effect and the temporary from the permanent effect.

Subsidised on-the-job training significantly and importantly favours longer job tenure. This result is in accordance with human capital theory and with the empirical literature surveyed in the introduction. On the one hand it confirms that on-the-job training seems to be more specific (less *transferable*) than classroom training. On the other hand, the mere fact that the effect is significantly positive accords with Stevens' (1994, 1996) hypothesis of *transferable* training

---

recipient' and 'increasing business activity' in the participation model. This is unlikely to affect the results as the coefficients of these variables were not significantly different from zero in the model with the normal mixing distribution.

and that the market under provides training. This conclusion should nevertheless be interpreted with some caution, since the group of subsidised on-the-job training programs encompasses different policies. In the CEF program introduced in Section 2, as for the pure subsidy schemes, we measure the effect of the program on job tenure both during and after participation. On the contrary, in FPI program, we estimate the post-participation effect. The effect on post-program job tenure is therefore over estimated. On the other hand, since we sampled hired workers only, our estimates are a lower bound to the true effect of program participation (cf. supra).

6. Conclusions

This paper has analysed whether active labour market policies have an influence on job tenure. The evaluation has focused on three Belgian policies : Pure wage subsidies (more specifically, temporary subsidies paid to employers who recruit eligible unemployed people with no additional requirement to train the worker), subsidised on-the-job training (followed by or embedded in a labour contract) and training programs organised for the unemployed by public institutions. The main problem in estimating the effect of these policies is the correction for selection bias. Given the data, this paper has adopted a non-experimental approach. To correct for selection on unobservables, we have selected a specification of the joint probability distribution of the job duration and the participation decision. A one-factor loading specification has been used for the mixing distribution. As it is now well established, such an approach does not guarantee that the correction for selection bias is appropriate. A sensitivity analysis was therefore required. In this paper, we compared the estimates resulting from a normal and a discrete with two points of support specification of the mixing distribution.

This analysis was based on a sample of hired workers obtained from a survey of employers. To the extent that program participation affects the hiring probability, the fact of basing the analysis on a sample of hired workers only, may bias downwards the estimation of the program effects. This bias was unavoidable. On the other hand, the problem that the sample over represented program participants and was therefore endogenous was explicitly taken into account in the statistical model.

A first conclusion of the paper is that the labour market policies all lengthen job tenure. Yet, the magnitude of the estimates is sensitive to the model specification, a result already emphasised by Lalonde (1986) and Fraker and Maynard (1987). Within the limits of this paper, the assumption about the distribution of unobserved heterogeneity turns out to be crucial. Facing different specifications, we have implemented a procedure of model selection based on two criteria (the Akaike information criterion and the Jeffrey-Bayes posterior-probability criterion). According to these criteria, the preferred specification corrects for unobserved heterogeneity by means of a discrete mixing distribution with two points of support.

Second, from the selected specification, a rather clear-cut conclusion can been drawn about subsidised on-the-job training which significantly favours longer job tenure. This conclusion is in accordance with theoretical predictions of human capital theory and with previous evaluations of the same policy. On the one hand this result confirms that on-the-job training seems to be more specific (less *transferable*) than classroom training. On the other hand, the mere fact that the effect is significantly positive accords with Stevens' (1994, 1996) hypothesis of *transferable* training and that the market under provides training. Belgian training programs for the unemployed and pure wage subsidies have important positive but non significant effects on job tenure. Given that participation to these training programs seems to signal low productivity, the

result provides weak evidence in favour of the human capital theory as opposed to matching theories in explaining job turnover. Our result on the effect of participation in the pure subsidy program is difficult to interpret, since we could not disentangle the effect during the program participation from the post-program effect. Yet, it is plausible that the subsidy only has a minor and non significant effect on the after subsidy job tenure. A tentative conclusion would be that our results offer some support to the view that mismatch is a major explanation of unemployment in Belgium.

References

ACKUM, S. (1991), "Youth unemployment, labour market programmes and subsequent earnings.", *Scandinavian Journal of Economics*, Vol. 93, pp. 531-453.

AKAIKE, H. (1973), "Information theory and an extension of the maximum likelihood principle", in PETROV, B. and CSAKI, F. (eds.), *Proceedings of the 2d international symposium on information theory*, Akademiai Kiado, Budapest, pp. 267-281.

ANGRIST, J.D., (1992), "What Can be Learned from a Natural Experiment", paper prepared for the Amsterdam Conference on Causality, November.

ASHENFELTER, O. (1978), "Estimating the effect of training programs on earnings.", *Review of Economics and Statistics*, Vol. 60, pp. 47-57.

ASHENFELTER, O. and CARD, D. (1985), "Using the longitudinal structure of earnings to estimate the effect of training programs", *Review of Economics and Statistics*, Vol. 67, pp. 648-660.

BARNOW, B.S. (1987), "The impact of CETA programs on earnings : A review of the literature.", *Journal of Human Resources, XXII*, pp. 157-93.

BARRON, J.M., BLACK, D.A. and DAVIES, R. (1989), "Job Matching and On-the-Job Training", *Journal of Labor Economics*, Vol. 7, pp. 1-19.

BASSI, L. (1984), "Estimating the effect of training programs with non-random selection.", *Review of Economics and Statistics*, Vol. 66, pp. 36-43.

BECKER, G.S. (1962), "Investment in Human Capital: A Theoretical Analysis", *Journal of Political Economy*, 70 (5, part 2), pp.9_49.

BECKER, G.S. (1975), *Human Capital*, 2nd edition, New York, Columbia University Press.

BJöRKLUND, A., (1990), "Evaluation of swedish labor market policy", *Finnish Economic Papers, 3*, pp. 3-13.

BONNAL, L., FOUGERE, D. et SERANDON, A. (1994), "L'impact des dispositifs d'emploi sur le devenir des jeunes chômeurs : une évaluation économétrique sur données longitudinales", *Economie et Prévision*, n° 115, pp. 1-28.`

BOOTH, A. L. and SATCHELL, S.E. (1994), "Apprenticeships and job tenuer", *Oxford Economic Papers*, 46, pp. 676-695.

BURTLESS, G. (1985), "Are Targeted Wage Subsidies Harmful? Evidence from a Wage Voucher Experiment", *Industrial and Labor Relations Review, October, 39, pp.105-114.*

CARD, D. (1985), "Using the longitudinal structure of earnings to estimate the effect of training programs.", *Review of Economics and Statistics*, Vol. 67, pp. 648-60.

CARD, D. and SULLIVAN, D. (1988), "Measuring the effect of subsidized training programs on movements in and out of employment.", *Econometrica*, Vol. 56, pp.497-530.

CHOW, G. C. (1985), *Econometrics*, Mc Graw-Hill,.

DEVINE, T.J. AND KIEFER N.M. (1991), *Empirical Labor Economics: The Search Approach*, New York, Oxford University Press.

DREZE, J.H. and SNEESSENS H.R. (1994), "Technical development, competition from low-wage economies and low-skill unemployment", *Swedish Economic Policy Review*, 1, 1-2, pp. 185-214.

DUBIN, J.A. and RIVERS, D. R. (1993), "Experimental Estimates of the Impact of Wage Subsidies", *Journal of Econometrics*, Vol. 56, n°1-2, pp. 219-42.

EDIN, P.A. (1988), "Individual consequences of plant closures.", Mimeo, Department of Economics, Uppsala University.

ELIAS, P. (1994), "Job-related training, trade union membship, and labour mobility : a longitudinal study", *Oxford Economic Papers*, Vol. 46, pp. 563-578.

FLINN, C.J. and HECKMAN, J.J. (1982), "Models for the analysis of Labor Force Dynamics", in : BASSMANN, R. and RHODES, G. (eds), *Advances in Econometrics*, Vol. 1, pp. 35-95.

FRAKER, T. and MAYNARD, R. (1987), "The Adequacy of Comparison Group Design for Evaluations of Employment-Related Programs", *Journal of Human Resources*, n° 22, pp.194-227.
GREENHALGH, C. and STEWART, M. (1987), "The effects and determinants of training.", *Oxford Bulletin of Economics and Statistics*, Vol. 49, 171-90.
GRITZ, M.R. (1993), "The impact of training on the frequency and duration of employment", *Journal of Econometrics*, Vol. 21, pp. 21-52.
HAM, J.C. and LALONDE, R.J. (1991), "Estimating the effect of training on employment and unemployment durations : evidence from experimental data", NBER Working Paper, no. 3912, forthcoming in *Econometrica.*
HECKMAN, J.J. and HOTZ, V.J. (1989), "Choosing among alternative nonexperimental methods for estimating the impact of social programs : the case of manpower training." *Journal of the American Statistical Association*, Vol. 84, pp. 862-74.
JENSEN, P., PEDERSEN, P.J., SMITH, N. and WESTERGARD-NIELSEN, N. (1993), "The effects of labour market training on wages and unemployment. Some Danish results.", In BUNZEL, J.H., JENSEN P. and WESTERGARD-NIELSEN N. (eds), *Panel Data and Labour Market Dynamics*, North-Holland, Amsterdam.
JOVANOVIC, B. (1979a), "Job Matching and the Theory of Turnover", *Journal of Political Economy*, Vol. 87, pp. 972-90.
JOVANOVIC, B. (1979b), "Firm specific capital and the turnover", *Journal of Political Economy*, Vol. 87, pp. 1246-1260.
JOVANOVIC, B. (1984), "Matching, turnover and unemployment", *Journal of Political Economy*, Vol. 92, pp. 108-122.
KAITZ, H.B. (1979)," Potential use of Markov process models to determine program impact." in BLOCH, F.E. (ed.), *Research in Labor Economics*, Greenwich, Conn. : JAI Press.
KIEFER, N.M. (1988), "Economic Duration Data and Hazard Functions", *Journal of Economic Literature*, Vol. 26, pp. 646-79.
LALONDE, R. (1986), "Evaluating the econometric evaluations of training programs with experimental data". *American Economic Review*, Vol. 76, pp. 604-20.
LANCASTER, T. (1990), *The Econometric Analysis of Transition Data*, Cambridge University Press, Cambridge.
LAYARD, R., NICKELL, S. and JACKMAN, R. (1991) *Unemployment*, Oxford, University Press.
LYNCH L. M., (1991), "The Role of Off-the-Job vs. On-the-Job Training for the Mobility of Women Workers", *American Economic Review, Papers and proceedings*, vol. 81, pp. 151-156.
MAIN, B.G.M. (1991), "The effect of youth training on the employment probability", *Applied Economics*, Vol. 23, n° 2, pp. 3761-72.
MAIN, B.G.M. and SHELLY, M.A. (1990), "The effectiveness of the Youth Training Scheme as a manpower policy" *Economica*, Vol. 57, pp. 495-514.
MEALLI, F., PUDNEY, S. and THOMAS, J. (1996), "Training Duration and Post-Training Outcomes: A Duration-Limited Competing Risk Model", *Economic Journal*, Vol. 106, pp. 422-433.
MEYER, B.D. (1990), " Unemployment Insurance and Unemployment Spells", *Econometrica*, 58(4), pp. 557-82.
MEYER, B.D., (1995), "Natural and Quasi-Experiments in Economics", *Journal of Business and Economic Statistics*, Vol. 13, n°2 pp. 151-161.
PRENTICE, R.L. and GLOECKLER L.A., (1978), "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data", *Biometrics*, Vol. 34, pp. 57-67.
RIDDER, G. (1986), "An event history approach to the evaluation of training, recruitment and employment programmes.", *Journal of Applied Econometrics*, Vol. 11, pp. 109-26.
SCHWARZ, G. (1978), "Estimating the dimension of a problem", *Ann. Statisti.*, vol 6, pp. 461-4.
SNEESSENS, H.R. and SHADMAN-MEHTA, F (1995), "Real wages, skill mismatch and unemployment persistence. France, 1962-1989", *Annales d'Economie et de Statistique*, n°37/38, pp. 255-292.
SPENCE, M. (1973), "Job Market Signalling", *Quaterly Journal of Economics*, Vol. 87, pp.255-74.
STEVENS, M. (1994), "A theoretical model of on-the-job training with imperfect competition", *Oxford Economic papers*, 46, 537-562.
STEVENS, M. (1996), "Transferable training and poaching externalities" in BOOTH, A. L. and SNOWER, D. J., *Acquiring skills. Market failures, theur symptoms and policy responses*, Cambridge University Press, pp. 19-40.
STIGLITZ, J.E. (1975), "The Theory of Screening, Education and the Distribution of Income", *American Economic Review*, Vol. 65, pp. 285-300.

TORP, H. (1994), "The impact of training on employment : assessing a Norwegian labour market programme.", *Scandinavian Journal of Economics*, Vol. 96, 531-50.
VAN DER LINDEN, B. 1995, "Effets de perte sèche et de substitution des formations professionnelles et des aides à l'embauche : une évaluation par enquête auprès d'employeurs", *Bulletin de l'IRES n° 180*, Université Catholique de Louvain.
ZWEIMULLER, J. and WINTER-EBMER, (1996), "Manpower Training Programmes and Employment Stability", *Economica*, Vol. 63, pp. 113-30.

Figure 1: The Sample Composition (376 firms)

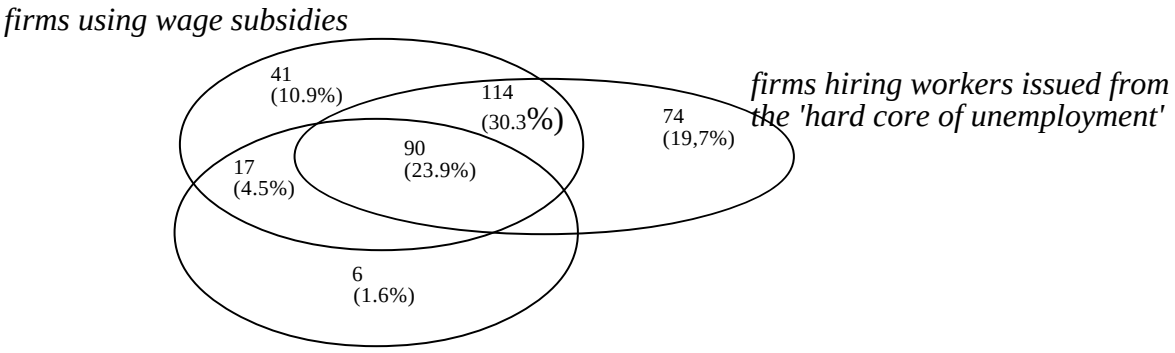

Figure 2: The shape of the θ_k parameters for three partitions of job tenure.

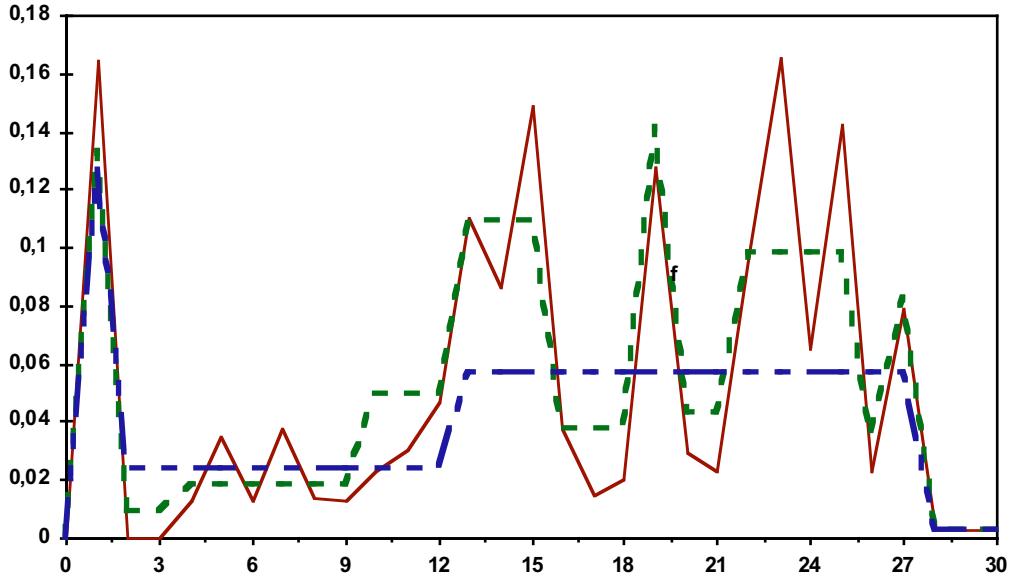

Table 1: Expenditures on labour market programs (% of G.D.P.) and basic unemployment indicators in selected countries, 1992.

	Belgium	Germany	France	The Netherlands
Expenditures				
Passive measures	2.9	2.0	2.0	2.2
Active measures	1.1	1.7	1.0	1.1
- training for unemployed adults	0.14	0.63	0.32	0.19
- training for occupied adults	0.09	0.03	0.06	-
- wage subsidies in the private sector	0.07	0.07	0.04	0.03
Passive + active measures	4.0	3.7	3.0	3.4
Unemployment rate (%)	7.9	4.6	10.4	6.7
Share of long-term unemployment (%)	59.0	33.5	36.1	44.0
Relative unemployment rate by level of educational attainment : primary and, between parentheses, lower secondary compared to the higher education - university (in 1989 or 1990)	7.0 (4.6)	_ (3.1)	3.9 (3.5)	2.7 (1.5)

Source : O.C.D.E. (1994), Employment outlook and Drèze and Sneessens (1994).

Table 2. Summary statistics

Variable	Mean
Hiring date	
* before 1991	3%
* in 1991	39%
* in 1992	57%
* in 1993 (first months)	2%
Still occupied at the end of March 1993?	
Yes	77%
No	23%
Age	
* < 25 years	57%
*≥25 years and < 45 years	28%
* ≥45 years	15%
Women	39%
Low-skilled worker	54%
Previously long-term unemployed	40%
Previously welfare recipient	9%
Disabled	4%
Type of worker	
* production worker	74%
* administrative personnel	22%
* foremen, supervisory or management	4%
Type of program:	
*'pure' wage subsidy	22%
*subsidized on-the-job training	12%
*training	35%
Sector :	
* Manufacturing	36%
* Construction	18%
* Trade	12%
* other services	34%
Brussels & the Walloon Provinces	50%
Increasing business activity during the 1991-1992 period	45%
Firm size:	
less than 50 employees at the end of 1992	55%
at least 50 employees at the end of 1992	45%
Sample size	864

Table 3. The parameter estimates ignoring the selection and the endogenous sampling problems[1]

(standard errors between parentheses)

	Without unobserved heterogeneity	**With unobserved heterogeneity** (normal distribution)
θ_1 ($0 \leq t < 1$ month)	0.114 (0.057)**	0.073 (0.045)*
θ_2 ($1 \leq t < 12$)	0.004 (0.002)*	0.004 (0.002)*
θ_3 ($12 \leq t < 27$)	0.022 (0.011)**	0.018 (0.009)**
θ_4 ($27 \leq t < 61$)	0.002 (0.001)**	0.004 (0.002)**
Age		
* < 25 years	-0.351 (0.173)**	-0.365 (0.199)*
* ≥ 45 years	0.118 (0.255)	0.204 (0.301)
sex (women = 1)	-0.237 (0.180)	-0.293 (0.213)
Low-skilled worker	-0.220 (0.164)	-0.303 (0.195)
Previously long-term unemployed	-0.390 (0.159)**	-0.476 (0.189)**
Previously welfare recipient	0.174 (0.291)	0.213 (0.335)
Disabled	-0.303 (0.349)	-0.259 (0.419))
Production worker	-0.050 (0.419)	-0.106 (0.509)
Administrative personnel	-0.013 (0.391)	-0.073 (0.475)
Pure wage subsidy	0.120 (0.189)	0.147 (0.225)
Subsidized on-the-job training	0.489 (0.254)**	0.596 (0.305)**
Training	0.107 (0.165)	0.127 (0.197)
* Construction sector	-0.031 (0.209)	-0.085 (0.252)
* Trade sector	-0.347 (0.292)	-0.393 (0.334)
* other services	-0.010 (0.194)	-0.015 (0.228)
Brussels & the Walloon region	0.205 (0.157)	0.325 (0.190)*
Increasing business activity	0.247 (0.123)**	0.392 (0.185)**
Firm-specific exit rate	-0.165 (0.060)***	-0.197 (0.078)**
Firm size	0.243 (0.122)**	0.281 (0.138)**
σ^2	_	1.012 (0.129)***
Sample size	864	
Number of parameters	23	24
Mean ln-L	-0.534	-0.531
Akaike Information Criterion	968.299	965.796

*** means not significantly different from zero at the 1 % level; ** means not significantly different from zero at the 5 % level; * means not significantly different from zero at the 10 % level.

[1] The reference case is a middle-aged man who did not exit from the 'hard core of unemployment', who works in the supervisory or management personnel of a small manufacturing firm in Flandres and who did not participate in a program.

Table 4. The parameter estimates of the joint duration and participation process ignoring endogenous sampling - Normal mixing distribution[1].

(standard errors between parentheses)

	Duration process	Participation process : training	Participation process : wage subsidies
θ_1 ($0 \leq t < 1$ month)	0.155 (0.066)**		
θ_2 ($1 \leq t < 12$)	0.009 (0.005)*		
θ_3 ($12 \leq t < 27$)	0.048 (0.020)**		
θ_4 ($27 \leq t < 61$)	0.002 (0.001)*		
Constant		-0.635 (0.128)***	-1.257 (0.540)**
Age			
* < 25 years	-0.464 (0.207)**	0.063 (0.221)	0.814(0.225)***
* ≥ 45 years	0.235 (0.298)	-0.100 (0.300)	0.368 (0.314)
sex (women = 1)	-0.284 (0.211)	-0.515 (0.219)**	-0.241 (0.216)
Low-skilled worker	-0.327 (0.193)*	-0.063 (0.202)	0.255 (0.200)
Previously long-term unemployed	-0.555 (0.189)***	0.369 (0.200)*	0.222 (0.121)*
Previously welfare recipient	0.317 (0.330)	-0.317 (0.340)	0.121 (0.323)
Disabled	-0.452 (0.415)	-0.352 (0.472)	1.582 (0.430)***
Production worker	-0.154 (0.473)	-0.886 (0.449)*	0.593 (0.453)
Administrative personnel	-0.096 (0.505)	-0.708 (0.478)	0.057 (0.480)
Pure wage subsidy	0.943 (0.253)***		
Subsidized on-the-job training	1.409 (0.322)***		
Training	0.758 (0.242)***		-0.334 (0.267)
* Construction sector	-0.153 (0.255)	-0.031 (0.102)	-0.540 (0.274)**
* Trade sector	-0.514 (0.334)	-0.293 (0.278)	-0.725 (0.318)**
* other services	-0.133 (0.228)	-0.513 (0.343)	-0.834 (0.232)***
Brussels & the Walloon region	0.485 (0.209)**	-1.256 (0.197)***	0.727 (0.197)***
Increasing business activity	0.377 (0.180)**	0.041 (0.145)	0.143 (0.112)
Firm-specific exit rate	-0.200 (0.077)***	-0.031 (0.102)	0.248 (0.093)***
Firm size	0.264 (0.132)**	0.043 (0.091)	0.379 (0.095)***
σ^2	1.045 (0.116)***		
η_T	0.198 (0.187)		
η_S	0.234 (0.188)		
Sample size	864	Number of parameters	61
Mean ln-L	-0.404		

*** means not significantly different from zero at the 1 % level; ** means not significantly different from zero at the 5 % level; * means not significantly different from zero at the 10 % level.

[1] The reference case is a middle-aged man who did not exit from the 'hard core of unemployment', who works in the supervisory or management personnel of a small manufacturing firm in the north of the country (Flandres) and who did not participate in a program.

Table 5. The parameter estimates of the joint duration and participation process taking endogenous sampling into account - Normal mixing distribution[1]

(standard errors between parentheses)

	Duration process	Participation process : training	Participation process : wage subsidies
θ_1 $(0 \leq t < 1$ month)	0.168 (0.086)**		
θ_2 $(1 \leq t < 12)$	0.006 (0.004)*		
θ_3 $(12 \leq t < 27)$	0.045 (0.021)**		
θ_4 $(27 \leq t < 61)$	0.001 (0.0005)*		
Constant		-0.632 (0.113)***	-0.979 (0.483)**
Age			
* < 25 years	-0.540 (0.217)**	0.153 (0.239)	1.011 (0.242)***
* ≥ 45 years	0.198 (0.316)	-0.114 (0.324)	0.437 (0.337)
sex (women = 1)	-0.267 (0.223)	-0.510 (0.236)**	-0.227 (0.233)
Low-skilled worker	-0.361 (0.205)*	-0.085 (0.218)	0.105 (0.217)
Previously long-term unemployed	-0.613 (0.198)***	0.421 (0.213)**	0.166 (0.213)
Previously welfare recipient	0.208 (0.337)	0.041 (0.087)	0.042 (0.076)
Disabled	-0.557 (0.445)	-0.369 (0.501)	1.624 (0.465)***
Production worker	-0.237 (0.497)	-1.094 (0.490)**	0.894 (0.498)*
Administrative personnel	-0.034 (0.532)	-0.919 (0.524)*	0.192 (0.520)
Pure wage subsidy	1.338 (0.234)***		
Subsidized on-the-job training	1.754 (0.306)***		
Training	1.167 (0.231)***		-0.363 (0.229)
* Construction sector	-0.187 (0.267)	-0.102 (0.297)	-0.733 (0.294)**
* Trade sector	-0.678 (0.351)*	-0.667 (0.365)*	-0.936 (0.340)***
* other services	-0.212 (0.240)	-0.108 (0.253)	-0.940 (0.251)***
Brussels & the Walloon region	0.384 (0.200)*	-1.097 (0.210)***	1.035 (0.214)***
Increasing business activity	0.341 (0.166)**	0.056 (0.084)	0.132 (0.421)
Firm-specific exit rate	-0.249 (0.083)***	-0.002 (0.105)	0.228 (0.098)**
Firm size	0.281 (0.139)**	0.088 (0.100)	0.617 (0.117)***
σ^2	1.351 (0.067)***		
η_T	0.045 (0.054)		
η_S	0.0642 (0.054)		
a	0.678 (0.009)***		
b	0.764 (0.104)***		
Sample size	864	Number of parameters	63
Mean ln-L	-0.399	Jeffrey-Bayes posterior probability criterion	1116.004

*** means not significantly different from zero at the 1 % level; ** means not significantly different from zero at the 5 % level; * means not significantly different from zero at the 10 % level.

[1] The reference case is a middle-aged man who did not exit from the 'hard core of unemployment', who works in the supervisory or management personnel of a small manufacturing firm in the north of the country (Flandres) and who did not participate in a program.

Table 6. The parameter estimates of the joint duration and participation process taking endogenous sampling into account - Discrete mixing distribution[1]

(standard errors between parentheses)

	Duration process	Participation process : training	Participation process: wage subsidies
θ_1 ($0 \leq t < 1$ month)	0.104 (0.057)*		
θ_2 ($1 \leq t < 12$)	0.004 (0.003)		
θ_3 ($12 \leq t < 27$)	0.022 (0.012)*		
θ_4 ($27 \leq t < 61$)	0.001 (0.001)		
Constant		-0.326 (0.635)	-1.762 (0.619)**
Age			
* < 25 years	-0.409 (0.199)**	0.412 (0.290)	0.451 (0.282)
* ≥ 45 years	0.135 (0.269)	-0.150 (0.340)	0.137 (0.342)
sex (women = 1)	-0.252 (0.192)	-0.658 (0.261)**	-0.475 (0.243)**
Low-skilled worker	-0.284 (0.192)	0.051 (0.260)	0.315 (0.263)
Previously long-term unemployed	-0.468 (0.197)**	0.778 (0.262)**	0.157 (0.246)
Previously welfare recipient	0.161 (0.302)		
Disabled	-0.319 (0.370)	-0.091 (0.455)	1.584 (0.447)***
Production worker	-0.027 (0.423)	-0.456 (0.501)	-0.184 (0.492)
Administrative personnel	-0.104 (0.456)	0.148 (0.555)	0.340 (0.544)
Pure wage subsidy	0.395 (0.383)		
Subsidized on-the-job training	0.751 (0.411)*		
Training	0.220 (0.306)		0.454 (0.429)
* Construction sector	-0.051 (0.224)	0.715 (0.315)**	-0.214 (0.296)
* Trade sector	-0.407 (0.312)	-0.602 (0.396)	-0.071 (0.343)
* other services	-0.046 (0.210)	-0.251 (0.271)	-0.663 (0.269)**
Brussels & the Walloon region	0.244 (0.177)	-1.380 (0.248)***	0.541 (0.225)**
Increasing business activity	0.283 (0.167)*		
Firm-specific exit rate	-0.181 (0.070)***	0.015 (0.101)	0.143 (0.096)
Firm size	0.230 (0.127)*	0.062 (0.101)	0.249 (0.105)**
q_1	0.736 (0.566)		
ν_1	0.444 (0.508)		
η_T	0.119 (0.222)		
η_S	0.052 (0.366)		
a	0.627 (0.327)**		
b	0.943 (0.354)***		
Sample size	864	Number of parameters	60
Mean ln-L	-0.401	Jeffrey-Bayes posterior probability criterion	1098.726

*** means not significantly different from zero at the 1 % level; ** means not significantly different from zero at the 5 % level; * means not significantly different from zero at the 10 % level.

[1] The reference case is a middle-aged man who did not exit from the 'hard core of unemployment', who works in the supervisory or management personnel of a small manufacturing firm in the north of the country (Flandres) and who did not participate in a program.