

IZA DP No. 682

## Vocational Training of Unemployed Workers in Belgium

Bart Cockx

January 2003

# Vocational Training of Unemployed Workers in Belgium

**Bart Cockx**

*IRES, Catholic University of Louvain  
and IZA Bonn*

Discussion Paper No. 682  
January 2003

IZA

P.O. Box 7240  
D-53072 Bonn  
Germany

Tel.: +49-228-3894-0  
Fax: +49-228-3894-210  
Email: [iza@iza.org](mailto:iza@iza.org)

This Discussion Paper is issued within the framework of IZA's research area *Internationalization of Labor Markets*. Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent, nonprofit limited liability company (Gesellschaft mit beschränkter Haftung) supported by the Deutsche Post AG. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public. The current research program deals with (1) mobility and flexibility of labor, (2) internationalization of labor markets, (3) welfare state and labor market, (4) labor markets in transition countries, (5) the future of labor, (6) evaluation of labor market policies and projects and (7) general labor economics.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available on the IZA website ([www.iza.org](http://www.iza.org)) or directly from the author.

## ABSTRACT

### Vocational Training of Unemployed Workers in Belgium\*

In this paper we estimate, for the 1989-93 period in Belgium, the effect of vocational classroom training on the rate of transition from unemployment. We propose a “control function” estimator accounting for variable treatment effects. In the absence of interaction effects between explanatory variables this estimator identifies treatment effects free from selection bias. A natural experiment induces exogenous sub-regional variation in the training supply. This provides over-identifying restrictions that cannot be rejected. During participation, the transition rate decreases by 23% to 30%. Afterwards it increases by 47% to 73%. Making training available for a broader population would, however, reduce the effectiveness of the programme.

JEL Classification: C41, J24, J64, J68

Keywords: vocational training, evaluation, transition models, job-search

Bart Cockx  
Institut de Recherches Economiques et Sociales (IRES)  
Université Catholique de Louvain  
Place Montesquieu 3  
1348 Louvain-la-Neuve  
Belgium  
Tel.: +32 10 473439  
Fax: +32 10 473945  
Email: [cockx@ires.ucl.ac.be](mailto:cockx@ires.ucl.ac.be)

---

\* I am grateful to Isabelle Bardoulat for the research assistance provided. This project has been financed by the following grants: SSTC n° DB/10/035, PAI P4 /01 and Tournesol n° 98.067. We thank the Office National de l'Emploi and the Office communautaire et Régional de la Formation Professionnelle et de l'Emploi for providing the data. This research project has especially benefited from the support of and the discussions with Michel Mouchart and Bruno Van der Linden. We thank Jan van Ours, the participants of the conference on the Econometric Evaluation of Active Labour Market Policies in Europe, June 25-26, 1999 in Mannheim and the participants of the research seminars of the Laboratoire de microéconométrie du CREST in Paris and at the Tinbergen Institute in Amsterdam for their comments.

## 1. Introduction

Since public authorities have realised that macro-economic policies no longer suffice to cope with high and persistent unemployment, active labour market policies have been increasingly used in OECD countries (OECD 1993). Unemployment persistence is especially important in Belgium. The share of long-term<sup>3</sup> unemployment in Belgium is among the highest in the OECD. During the 1983-1996 period this share has only been (slightly) below 60% in the 1991-93 recession (OCDE 1997, p.8). Despite this persistent unemployment, we observe that Belgium has spent a higher share of its GDP on active labour market policies than other countries of the European Union (EU) on average. In 1995 this share amounted to 1.4% for Belgium, as compared to 1.1% for the UE (OCDE 1997, p.99). The high share of expenditures, oriented to direct job creation in the public and the non-profit sector, largely explains this finding. If we exclude expenditures on direct job creation, the figures on shares drop to 0.8% for Belgium and 0.9% for the EU. Moreover, if we consider the outlays on training schemes for unemployed workers or workers at risk of losing their jobs, Belgium spends only 0.16% of its GDP, whereas this share is 0.27% for the EU on average.

We study the impact of classroom training on the transition rate from unemployment in Wallonia, the French speaking region in the South of Belgium<sup>4</sup> during the 1989-93 period. We distinguish between the impact of training during participation and afterwards. We explicitly account for the variability of the impact across trainees. In particular, typically only unemployed workers with the highest returns to training enrol. Consequently, we find that increasing the coverage of the programme to a broader population, such as encouraged by the European Employment Strategy agreed at the Luxembourg Jobs summit in November 1997 (Commission of the EU 1997), is likely to be much less effective than the training programme that we analyse: this broader population will contain workers with lower returns to training pulling down the average return of participants. This claim is confirmed in the empirical analysis below.

In Europe, the literature<sup>5</sup> generally reports mixed evidence. If effects are found to be positive, participation in classroom training increases the chances of being employed and decreases the risk of unemployment. The effects of classroom training on wages and earnings are often reported to be insignificant or even negative. Taking the relatively small amount of time spent in training, some weeks or at most some months, into account, the latter result may not come as a surprise. Compare it to the annual rate of return to education, which is of the order of 10% (Ashenfelter and Rouse 1995).

The literature on the evaluation of training programmes in Europe has only recently started to accumulate significantly. Given that some of these studies may not have adequately corrected for selection bias, we compare these findings to the U.S. findings. In the U.S. there exists a vast literature evaluating the impact of public sector sponsored classroom training programmes (see Lalonde 1995, Friedlander, Greenberg and Robins 1997, and Heckman et al. 1999 for recent surveys). In contrast to the European studies, the most common outcome of interest is earnings, rather than employment or unemployment. In general, this literature finds significantly small, but significantly positive effects on earnings of adults, especially for women. No training programme has been found effective for youth, however. Studies that decompose these earnings gains in its components usually find that they result more from increased weeks of employment than from increased hourly wages or increased hours per

---

<sup>3</sup> Defined as being unemployed for more than one year.

<sup>4</sup> Plasman (1993, 94) and Bollens and Nicaise (1994) also studied this relationship. The former finds that training decreases the transition rate from unemployment, while the latter authors report shorter unemployment spells for participants in Flanders, the region in the North of Belgium. These studies suffer from serious methodological shortcomings, however (see Cockx 1999 for a discussion). Cockx et al. (1998) find that vocational training for unemployed workers increases job *tenure*, but the effect is statistically insignificant.

<sup>5</sup> See Heckman, Lalonde and Smith 1999, Martin and Grubb 2001 and Calmfors *et al.* 2002 for surveys.

week for those who are employed. This is consistent with the findings of the European studies. Eberwein, Ham and Lalonde (1997) report that participation in classroom training increases the frequency of employment, but not the length of employment spells.

Each micro-econometric evaluation study of a labour market policy must deal with potential biases of the programme effect due to selective participation. In this paper, we use administrative data with a small number of explanatory variables. We cannot therefore rely on matching estimators as recently advocated in the literature (Dehejia and Wahba 1999, Heckman, Ichimura and Todd 1997, Lechner 1999, and Brodaty *et al.* 1999 in the context of transition models). In this study, we follow a quasi-experimental approach (see Meyer 1995, Angrist and Krueger 1999). In particular, we follow Cockx and Ridder (2001) who show that the selection bias may be eliminated by aggregation if the variation of the participation rate at the aggregate level is exogenous (see also Wald 1940 and Angrist 1991). As to enhance the credibility of our findings, we contrast those resulting from two different identifying assumptions. The first hinges on the absence of interaction effects. The second exploits a natural experiment generating sub-regional variation in the participation rates in training that is unrelated to the outcome variable.

The method proposed by Cockx and Ridder (2001) boils down to an Instrumental Variable (IV) estimator. If the effect of training is constant over the population or if the programme effect does not influence the participation decision in training, this method identifies the average effect of treatment on the treated (Heckman and Robb 1985 and Heckman 1997). Following Angrist and Imbens (1991), we call this the *selected average treatment effect* (SATE). If the above-mentioned conditions are not satisfied, the IV estimator identifies the local average treatment effect (LATE), i.e. the average impact of training for those unemployed workers who would be induced to participate if they moved from a sub-region with low training capacity to a high one (Imbens and Angrist 1994, Angrist, Imbens and Rubin 1996).

Participation in the training programme considered is voluntary. Since the application process is costly, only unemployed workers with high returns participate. Therefore, if the treatment effect differs between individuals, the IV estimator identifies LATE. In this paper we extend the method of Cockx and Ridder (2001) by proposing a "control function" estimator that identifies the distribution of programme effects of participants and therefore SATE.

Heckman and Honoré (1990) demonstrate that the distribution of the treatment effects can be identified if participation depends only on the gain of the programme, such as in Roy's model (1951). They show that with sufficient price variation across different markets of "occupational choice" one can identify the population skill distribution from aggregate data. In this paper, since the demand for training is rationed by supply, variation in the sub-regional training capacities induces sub-regional variation in the returns to training, i.e. in "prices". We explain how this variation can identify the distribution of treatment effects non-parametrically. In the empirical application, however, we impose a parametric distribution, because there is insufficient variation in the training capacity to estimate the distribution non-parametrically.

In the following section we present the institutional framework. In Section 3 we describe the data and present the statistical model. Section 4 discusses the estimation results. A final section concludes.

## 2. The Institutional Framework

In Wallonia the regional public employment agency, FOREM, is the main operator of vocational classroom training for adult unemployed workers<sup>6</sup>. For this purpose, it has a network of training centres at its disposal widely dispersed over the territory of the Walloon

---

<sup>6</sup> More details of the institutional framework can be found in OECD (1997).

region. FOREM provides vocational training for a wide range of professions of the secondary (FP II) and tertiary sector (FP III). Apart from vocational training, it also offers services information, vocational guidance and work experience to a more disadvantaged population of unemployed workers in its Centres of Reception, and of Guidance and Socio-Professional Initiation (C.A./C.O.I.S.P). During the 1989-93 period, about two thirds of the available training slots are of the vocational type, each sector being equally important. In terms of hours of training, the share of vocational training increases to 85%<sup>7</sup>. The level at which vocational training is offered varies, but the majority of the programmes consists in the development of basic skills required in particular vocations. The data that were used for the evaluation study did not allow distinguishing between the different programme types. The estimated effect of participation to training is therefore necessarily an average effect over the wide variety of programmes offered.

What determines participation into these training programmes? In principle, no restrictions are imposed. Participation is voluntary and any unemployed worker can apply. Moreover, a trainee remains entitled to unemployment benefits during participation, is paid 1 euro for each effective training hour and is reimbursed transportation costs. Nevertheless, a candidate trainee must bear two main costs: the application costs and the job opportunities forgone during the waiting period before the effective beginning of the programme.

The application costs are the cost of gathering information on the programme, as well as the time, effort and transportation costs required for passing a number of tests and interviews. The candidate is tested on basic reading and writing for FP III and on calculation for FP II. In addition, they need to pass medical and psycho-technical tests. These tests are uniform over training centres and types. Subsequently, programme administrators interview candidates and select them on the basis of “motivation”. Second, as a consequence of the limited training capacity the demand for training always exceeds the supply.<sup>8</sup> The candidate must therefore wait some time before entering training. This introduces a second “hurdle” in the application process. Even if the previously mentioned application costs are sunk, there remains the opportunity cost of waiting: participation requires the applicant to reject job offers during this waiting period. He/she will do so as long as the expected returns to participation exceed those of accepting a job prior to entering the programme.

These participation costs induce a process of self-selection. Only workers with returns to training above a certain threshold can compensate for these costs and will apply and wait. The less training slots are available, the longer is the waiting period and the higher must be this threshold. Variation in the degree of rationing between sub-regions induces therefore variation in the threshold return to training. It's the variation of this threshold that we use in the empirical analysis to identify the distribution of treatment effects.

The selection of programme administrators reinforces this self-selection process. Even if we cannot evaluate for certain the implicit selection criteria of the programme administrators, these criteria favour the selection of workers who most benefit from training. In this respect it's important to note that neither training centres, nor administrators or instructors are in any way rewarded on the basis of placement ratios or to other labour market outcomes. Consequently, direct incentives to cream-skim candidates, likely to find jobs even without participation in the programme, are absent.

---

<sup>7</sup> In the sequel figures for which no source are mentioned are taken from FOREM (1989-94) or from unpublished documents of the employment agency.

<sup>8</sup> See Section 3.4 for further details.

### 3. The Data and the Statistical Model

#### 3.1. The Data

We analyse administrative data on unemployment spells beginning<sup>9</sup> between May 1989 and March 1993 in Wallonia, the French speaking region of Belgium. We consider all officially registered full-time unemployed workers. On the one hand, 23% of these spells consist of school-leavers entitled to unemployment benefits after a waiting period of 6 months. The remainder of these spells are of workers entitled to benefits through their past work experience.<sup>10</sup> In the sequel, the school-leavers will be referred to as the ‘young’ and the other group as the ‘old’. In principle, the entitlement duration is indefinite in Belgium. There is one important exception to this rule.<sup>11</sup> Cohabitants, who are not a head of a family, can lose entitlements on grounds of “excessive” unemployment duration. This duration is, however, well beyond the maximal duration, i.e. 28 months, considered in the empirical analysis below.

The data are grouped into monthly intervals. However, the recorded transitions of school-leavers are not reliable throughout the waiting period, i.e. throughout the first 6 months. We therefore need to group the first 6 months in a single interval. Moreover, for purposes of a symmetric treatment, we also group the data of the ‘old’ unemployed workers in a similar way. Transitions within a monthly interval are recorded, but the timing is not. Unfortunately, as a consequence of administrative errors there is a proliferation of movements within the month, rendering this information unreliable. Another flaw of the data is that the destination for which unemployment is left is not known. We therefore cannot determine whether a worker leaves unemployment for a job or whether he/she has left the labour force.

In order to avoid the measurement of spurious exits and exits from the labour force, we follow Plasman (1993, 1994) by imposing that an exit should last more than two months to be recorded as such. In addition, we only consider workers younger than fifty at the start of the unemployment spell. In that way, we eliminate a large share of the withdrawals from the labour force. Besides, for obvious reasons, older workers hardly participate in training.

We can identify the months in which a worker participated in a training programme organised or subsidised by the Walloon employment agency, FOREM. We do not know, however, whether the trainee completed the programme successfully or whether he/she dropped out. Neither are we informed about the nature of training offered.

Unemployed workers may also participate in training programmes organised on-the-job or by other institutions not subsidised by FOREM. Unemployed workers participating in such programmes are assimilated to non-participants. This contaminates our “control group”. However, in view of the marginal number of workers involved in such programmes as a proportion of the total number of unemployed (2.1% in 1991), this will not significantly bias our results.

Table 1 summarises some descriptive statistics on the population retained. The population consists of 1,361,660 spells. One individual may have experienced several periods of unemployment during the observation period. This means that there haven’t been as many individuals entering unemployment during this period. 23% of this population is ‘young’. For both, the young and the old, Table 1 provides statistics with respect to all spells and with respect to spells during which some time was spent in training. According to these statistics trainees are higher educated. The average age of trainees is not very different from any unemployed worker. Among the old, males are more likely to participate in training. Among

---

<sup>9</sup> We do not sample workers who are unemployed at the beginning of the observation period. As the date at which their spell began is not known, these spells are left censored. It is well known that the analysis of left censored spells would complicate the statistical analysis substantially.

<sup>10</sup> Workers are entitled to unemployment benefits if they have been employed for at least 75 days within a prescribed period prior to their claim if younger than 18 years, and up to at least 600 days if older than 50 years (Van Langendonck 1991, p.450)

<sup>11</sup> Leaving out administrative sanctions, such as imposed on workers refusing job offers or declaring false information.

the young, it is rather the women who are over-represented. Old workers who have been unemployed previously also have a slightly higher likelihood of programme participation.

### INSERT TABLE 1 APPROXIMATELY HERE

The administration of FOREM is decentralised in 11 sub-regional departments. For purposes of the empirical analysis it is important to group these sub-regions in ones that are sufficiently homogenous with respect to the mix of training schemes offered. We therefore only report statistics on the 7 re-grouped sub-regions.

Median unemployment duration is 3 months, both for the young and the old. This is not very long, in view of Belgium's reputation of being one of the OECD countries with the highest share of long-term unemployment. However, the latter statistics are based on stock samples suffering from the well known "length bias" (Salant 1977). Median unemployment duration is much longer if we consider spells with some time in training: 9 months for the old and 12 months for the young. This may have several explanations. First, very few workers enter a training programme immediately when entering unemployment. The median duration until a worker enters a training programme is 3 months for the old and 6 months for the young. This may be the consequence of both, the rationing of the demand for training and the time required for an unemployed worker to be convinced that training may increase his/her job finding rate. Second, if the transition from unemployment exhibits negative duration dependence, trainees will experience longer spells as a consequence of the elapsed unemployment duration at the moment of programme entry. Finally, trainees may be a selective sub-population of the unemployed. These explanations will be disentangled in the empirical analysis below. Note that the median length of a training spell is 2 months.

## 3.2 Heterogeneous treatment effects

We allow the overall returns to training,  $\Delta$ , to vary between unemployed workers. Workers know their own returns, but to the researcher they are randomly drawn from a distribution. In this section we discuss the relationship between these returns and the treatment effects considered in the empirical analysis and explain how we identify the distribution of these treatment effects.

Up to this point we did not specify what these overall returns  $\Delta$  measure. In fact, training may have positive returns for several reasons. It may accelerate the speed at which unemployment is left for employment, increase the length of the subsequent employment spell, increase the average wage offer, etc.  $\Delta$  measures the combined impact of training via these outcomes on the lifetime utility of the ex-participant in training. Our data do not measure this overall return directly. In particular, we only observe the return in terms of the enhancement of the transition rate from unemployment. Hence, identification of the distribution of these returns requires a monotonic relationship between  $\Delta$  and the effect of training on the transition rate. Note, however, the nature of this monotonic relation may differ according to the phase in the unemployment spell. We distinguish between three phases: the waiting period (subscript  $w$ ), the period during training (subscript  $d$ ) and the post-training period (subscript  $p$ ). We denote the proportional effect of training on the transition rate from unemployment for an individual with return  $\Delta$  respectively by  $\alpha_w(\Delta)$ ,  $\alpha_d(\Delta)$  and  $\alpha_p(\Delta)$ .

From a theoretical perspective,<sup>12</sup> training does not necessarily enhance the job finding rate. By affecting the productivity of the worker training may induce the worker to become choosier in his job acceptance behaviour and therefore increase unemployment duration. Nevertheless, in the benchmark model we assume that this effect is never dominant:  $\forall \Delta: \alpha_p(\Delta) > 0$ . This assumption is in line with the findings reported in the literature: classroom

---

<sup>12</sup> In Cockx and Bardoulat (1999) we formally model the training participation within a job search framework and show conditions under which the assumptions in this section are justified.



training increases the employment probability and not so much the earnings or the length of employment spells (cf. the review in the Introduction). Moreover, this assumption cannot be rejected in the empirical analysis reported below. In addition, as explained above, for purpose of identification, we assume that the transition rate increases with the return:  $\partial\alpha_p(\Delta)/\partial\Delta>0$ . This is justified as long as higher  $\Delta$ 's are predominantly induced by higher job arrival rates and not by enhanced productivity gains of training, pushing up reservation wages and decreasing thereby the exit rate from unemployment.

Throughout participation in the training programme the worker enhances his productivity. Nevertheless, the time devoted to job search decreases. Moreover, the extra return associated to completing the programme increases the reservation wage during participation. We therefore expect the transition rate from unemployment to decrease during participation and more so, the higher the overall return to training:  $\forall\Delta: \alpha_d(\Delta)<0$  and  $\partial\alpha_d(\Delta)/\partial\Delta<0$ . Finally, during the waiting period the worker would forego the option value of training by accepting a job. Consequently, the prediction of the impact during training is reinforced during this period:  $\forall\Delta: \alpha_w(\Delta)<0$  and  $\partial\alpha_w(\Delta)/\partial\Delta<0$ .

How can we now identify the distribution of  $\alpha_w(\Delta)$ ,  $\alpha_d(\Delta)$  and  $\alpha_p(\Delta)$  among trainees? As explained in Section 2, the (self-) selection process is such that only workers with returns above a certain threshold value  $\Delta^*$  participate in training. As already mentioned, the threshold value  $\Delta^*$  increases with the extent of excess demand for training: workers enter training only if the return more than compensates for the increased waiting cost. The training capacity can be measured by the fraction of unemployed workers of a certain type that in a particular region participate ( $\pi_d$ ) or have participated in training ( $\pi_p$ ). Since the training capacity is inversely related to the extent of rationing, we have that  $\partial\Delta^*(\pi_j)/\partial\pi_j<0$  for  $j=d, p$ .

Since the treatment effects are monotonously increasing ( $\alpha_p(\Delta)$ ) or decreasing ( $\alpha_w(\Delta)$  and  $\alpha_d(\Delta)$ ) in  $\Delta$ ,  $\Delta^*$  also defines threshold values for the treatment effects: the treatment effects of all trainees will be above or below the corresponding threshold. Moreover, since these thresholds vary monotonously with the fraction of trainees and as the fraction of trainees also measures the probability that returns are above (or below) these values, we can identify the distribution function of the returns by the variation in the fraction of trainees, i.e. by the variation in the training capacity.

The following example may clarify the argument. Let us assume that  $|\alpha_d(\Delta)|$  has an exponential distribution with mean  $|\bar{\alpha}_d|$ . We take the absolute value, because, during participation, the impact of training on the transition rate from unemployment is negative. Since  $\partial\alpha_d(\Delta)/\partial\Delta<0$ , the treatment effects of all trainees are below the threshold value  $\alpha_d^*$ . The fraction of participants  $\pi_d$  is therefore equal to the probability that the treatment effect is below  $\alpha_d^*$ . Together with our distributional assumption we therefore have:

$$(1) \quad \pi_d = \Pr[\alpha_d \leq \alpha_d^*] = \exp[-\alpha_d^*/\bar{\alpha}_d]$$

By inverting this relationship we find

$$(2) \quad \alpha_d^* = -\bar{\alpha}_d \ln(\pi_d)$$

This allows us to find an expression for the effect of treatment on the treated (SATE):

$$(3) \quad E[\alpha_d(\Delta) \text{ during participation}] = -\frac{\int_{-\infty}^{-\bar{\alpha}_d \ln(\pi_d)} \alpha \exp[-\alpha/\bar{\alpha}_d] d\alpha}{\bar{\alpha}_d \pi_d} = -\bar{\alpha}_d [1 - \ln(\pi_d)]$$

Variation of  $\pi_d$  identifies SATE and consequently ATE ( $\bar{\alpha}_d \equiv E[\alpha_d(\Delta)]$ ) and the whole distribution of returns.

In principle, we can identify the full distribution of treatment effects non-parametrically.<sup>13</sup> However, in our data  $\pi_d$  varies over a limited range. We therefore impose more structure and impose a parametric form. In fact, in the benchmark model the distribution is exponential, as in the example. We also tried out other specifications, but do not report them in this paper (see Cockx and Bardoulat, 1999).

By similar arguments we can identify the distribution of post-programme impacts and the corresponding SATE and ATE. If  $\alpha_p(\Delta)$  is exponentially distributed, the expression for SATE is  $\bar{\alpha}_p[1 - \ln(\pi_p)]$ . However, we face the following problem. The variation in  $\pi_p$  does not only reflect variation in the training capacity, but also in the transition rate from unemployment in the post-training period. In the empirical analysis we account for this endogeneity by using  $\pi_d$  as an instrument for  $\pi_p$ .

Finally, consider the pre-programme effect. Due to data imperfections, we cannot proceed analogously. The data cannot identify workers queuing for programme participation from non-applicants. However, we can test for the existence of a pre-programme effect, since the aggregate transition rate from unemployment can be shown to decrease with the proportion of trainees  $\pi_d$  (see Cockx and Bardoulat, 1999). This is because the larger number of training slots increases both, the number of candidate trainees and the reservation wage of those who are already waiting for participation.

### 3.3. The Statistical Model

We now develop the statistical model on the basis of which we will estimate the effect of participation in training on the transition intensity from unemployment. We choose Minimum Chi-Square (MCS) as the method of estimation (cf. Cockx 1997 and Cockx and Ridder 2001). This method has two main advantages. First, it does not require strong assumptions on duration dependence<sup>14</sup> and the distribution of unobserved heterogeneity, so that estimates are less sensitive to specification errors. Second, since it transforms the transition model to a linear regression model, we can use simple methods of aggregation to correct for the selection on unobservables. Cockx and Ridder (2001) develop a grouping/IV estimator within this framework. We propose a control function estimator (Heckman and Robb 1985, p. 172).

The MCS method requires a grouping of the data not only by duration, but also by explanatory variables. A drawback of this method is that it is difficult to take fixed unobserved individual heterogeneity within the groups into account.<sup>15</sup> As a consequence, we must also assume that individual returns to training are drawn at the beginning of each duration interval from the distribution of returns independently from previous duration intervals. We believe, however, that this assumption is not so unrealistic in this application, since the time spent in training is relatively short and only few participants will remain in training during more than one duration interval.<sup>16</sup>

In order to reduce biases induced by within-group heterogeneity, we choose, in the benchmark model, to analyse only spells with an elapsed duration of 6 months or more. As such, the sorting process has already rendered groups much more homogenous. Afterwards we will compare the benchmark model with one in which data on the first 6 months are

<sup>13</sup> See Cockx and Bardoulat (1999) for proof of this.

<sup>14</sup> The baseline transition intensities are assumed to be piece-wise constant.

<sup>15</sup> Treatment of fixed individual heterogeneity introduces a non-linearity (for details see Cockx, 1997), invalidating thereby the IV approach below and imposing stronger assumptions on the control function to correct for the selection bias. In addition, the time-varying nature of training participation would complicate estimation even further.

<sup>16</sup> The median duration in training is 2 months and the length of the duration intervals always exceeds 2 months.

included. We will also test whether within-group heterogeneity can indeed be neglected in the benchmark model.

There is another reason to analyse only spells with a certain elapsed duration. We argued that only workers with the returns above a certain threshold participate in training. The reason is that there exist a number of barriers in the application process implying that entering a training programme is costly and takes time. Nevertheless, we observe that within the group of the “old” unemployed workers, there are workers who are trained immediately after entering unemployment. These participants in training cannot have followed the same application procedure or they must have done so while they were still employed. For instance, these trainees may be displaced workers for whom the participation in training may have been a condition negotiated by unions to accept a collective dismissal of employees in the framework of the restructuring of a firm. However, if such an alternative selection procedure applies to these types of workers, the average programme effects are not necessarily the same.

The indicator variable of the training status is time varying. This means that we need to make assumptions on the evolution of the indicator variable within the duration intervals of the grouped data. We follow the solution proposed by Cockx and Ridder (2001). Define the following three states:

1. Unemployment without participation in training
2. Unemployment with participation in training
3. Out of unemployment

If one is willing to assume that the base-line transition intensities are constant within the intervals (as in Prentice and Gloeckler 1979) and that, within a duration interval, there is at most one transition between the states defined above, then Cockx and Ridder show that the treatment of a time-varying indicator variable boils down to jointly analysing two competing risks models, one with origin state 1 and destination states 2 and 3, and another one in which state 2 is the origin state and states 1 and 3 the destinations. State 3 is an absorbing state, because we can no longer observe a worker who left unemployment. In the sequel we will denote origin and destination states by superscripts  $u$  and  $v$  respectively.

The MCS method requires the data to be grouped in homogeneous cells. We define these cells by crossing four criteria: the elapsed unemployment duration,  $k$ <sup>17</sup>, the sub-region,  $m$ , the eligibility criterion to unemployment benefits (‘old’ or ‘young’),  $s$ , and the training status,  $(u, t)$ . The training status can be one of the following three:

1.  $(u, t)=(1, 0)$ : no participation in training, neither presently or in the past
2.  $(u, t)=(2, 1)$ : presently participating in training
3.  $(u, t)=(1, 1)$ : past participation in training

The data are grouped in  $K+1$  intervals  $[\tau_0, \tau_1), \dots, [\tau_{k-1}, \tau_k), \dots, [\tau_{K-1}, \tau_K), [\tau_K, \tau_{K+1})$  of (possibly unequal) length  $\Delta_k = \tau_k - \tau_{k-1}$  with  $\tau_0=0$  and  $\tau_{K+1}=\infty$ . Training capacity may vary between segments  $(k, m, s)$ . This requires some change to the notation introduced in Section 3.2:  $\pi_{km1s}^2 \equiv \pi_d$  refers to the proportion of training slots during participation and  $\pi_{km1s}^1 \equiv \pi_p$  after participation within segment  $(k, m, s)$ . In addition,  $\bar{\alpha}_{t=1}^2 \equiv \bar{\alpha}_d$  denotes ATE during participation in training,  $(u, t)=(2, 1)$ , and  $\bar{\alpha}_{t=1}^1 \equiv \bar{\alpha}_p$  ATE in the post training period,  $(u, t)=(1, 1)$ .

We assume the following proportional specification of the transition intensities,  $h_{km1s}^{uv}$ , between the states:

$$(4) \quad \ln(h_{km1s}^{uv}) = \gamma_k^v + \beta_m^v + \psi_s^v + \lambda_{t=0}^{uv} \pi_{km1s}^{u=2} + \bar{\alpha}_{t=1}^{uv} [1 - \ln(\pi_{km1s}^u)] + \varepsilon_{km1s}^{uv}$$

<sup>17</sup> The data are assumed to be homogeneous over calendar time.

where  $u=1,2;v=1,2,3\neq u;k=1,2,\dots,K;m=1,2,\dots,M;t=0,1;s=0,y$ . The duration effect  $\gamma_k^v$ , the sub-regional effect  $\beta_m^v$  and the effect of the eligibility status  $\psi_s^v$  depend only on the destination state  $v$ . This is crucial for identification: like in differences-in-differences estimators, we exploit the absence of interaction effects between trainees and non-trainees to identify the selection bias (see below). We normalise the type effects and set  $\beta^v=\psi^v=0$ .

The specification retained in (4) assumes that the treatment effects during and post participation follow an exponential distribution: SATE, the average effect of treatment within the considered group, is equal to  $\bar{\alpha}_{t=1}^{uv}[1-\ln(\pi_{km1s}^u)]$  (cf. Equation (3)). The specification also allows testing for the presence of a pre-programme:  $\lambda_{t=0}^3<0$  (cf. Section 3.2). Note, by definition,  $\lambda_{t=0}^{2v}=\lambda_{t=1}^{2v}=0, u=1,2, v=1,3, \bar{\alpha}_{t=0}^{12}=\bar{\alpha}_{t=0}^{21}=0$  and by normalisation,  $\bar{\alpha}_{t=0}^{13}=0$ .

Finally,  $\varepsilon_{km1s}^{uv}$  are unobserved group effects. Note that these reflect between-group heterogeneity and not within-group heterogeneity. Selection on unobservables means that these unobserved group effects are correlated with the indicator variables of participation in training.

Under these assumptions the transition probability  $P_{km1s}^{uv}$ , i.e. the conditional probability of a transition of an individual of type  $(m,t,s)$  from state  $u$  to state  $v$  in duration interval  $k$ , is equal to

$$(5) \quad P_{km1s}^{uv} = \frac{h_{km1s}^{uv}}{\sum_{w=1,w\neq u}^3 h_{km1s}^{uw}} \left[ 1 - \exp\left(-\sum_{w=1,w\neq u}^3 h_{km1s}^{uw} \Delta_k\right) \right] \quad u=1, 2 \text{ and } v=1, 2, 3\neq u$$

This expression maps transition intensities to transition probabilities. The inverse mapping from transition probabilities to transition intensities is

$$(6) \quad z_{km1s}^{uv} \equiv \ln \left[ \frac{-P_{km1s}^{uv} \log(P_{km1s}^{uu})}{\Delta_k (1 - P_{km1s}^{uu})} \right] = \ln(h_{km1s}^{uv}) \quad u=1, 2 \text{ and } v=1, 2, 3\neq u$$

with  $P_{km1s}^{uu} \equiv 1 - \sum_{w=1,w\neq u}^3 P_{km1s}^{uw}$  is the complementary probability of staying in the origin state  $u$  during the  $k^{\text{th}}$  interval.

Let  $r_{km1s}^u$  be the number of individuals of type  $(m,t,s)$  in state  $u$  at the start of duration interval  $k$ , i.e. at time  $\tau_{k-1}$ , and let  $q_{km1s}^{uv}$  of these individuals make a transition from state  $u$  to  $v$  in the  $k^{\text{th}}$  duration interval, then we estimate  $P_{km1s}^{uv}$  by

$$(7) \quad \hat{P}_{km1s}^{uv} = \frac{q_{km1s}^{uv}}{r_{km1s}^u} \quad u=1, 2 \text{ and } v=1, 2, 3\neq u$$

If we replace in (6) the transition probabilities by their estimates, then the second equality does not hold exactly. However, a Taylor series expansion around  $P_{km1s}^{uv}$  yields

$$(8) \quad \hat{z}_{km1s}^{uv} = \ln(h_{km1s}^{uv}) + \omega_{km1s}^{uv}$$

with  $\omega_{km1s}^{uv} \equiv \sum_{w=1,w\neq u}^3 b_{km1s}^{uvw} (\hat{P}_{km1s}^{uw} - P_{km1s}^{uw})$ .<sup>18</sup> We can neglect the remainder of the Taylor expansion (Amemiya 1985, p.276-77).

Conditional on  $r_{km1s}^u$  the random vector of the number of transitions of individuals of type  $(m,t,s)$  from  $u$  to states  $v\neq u$  in the  $k^{\text{th}}$  interval has a multinomial distribution with

<sup>18</sup>  $b_{km1s}^{uvw}$  is defined in Cockx and Bardoulat (1999), Appendix C.1.

parameters  $r_{kmts}^u$  and  $P_{kmts}^u$ , which is the vector of transition probabilities from state  $u$  to states  $v \neq u$  in interval  $k$  for type  $(m, t, s)$ . The errors of the system of regression equations (8) have mean zero, but are heteroskedastic and correlated.<sup>19</sup>

Upon substitution of (4) in (8) we obtain a linear regression model. More precisely, we obtain for each origin state  $u=1,2$  two regression equations corresponding to the two destination states  $v=1,2,3 \neq u$

$$(9) \quad \hat{z}_{kmts}^{uv} = \gamma_k^v + \beta_m^v + \psi_s^v + \lambda_{t=0}^{uv} \pi_{km1s}^{u=2} + \bar{\alpha}_{t=1}^{uv} [1 - \ln(\pi_{km1s}^u)] + \varepsilon_{kmts}^{uv} + \omega_{kmts}^{uv},$$

$u=1, 2$  and  $v=1, 2, 3 \neq u$ . In the sequel we use the notation  $v_{kmts}^{uv} = \varepsilon_{kmts}^{uv} + \omega_{kmts}^{uv}$ .

We now proceed in two steps. First, we assume that the unobserved group effects,  $\varepsilon_{kmts}^{uv}$ , are not correlated with the training status, i.e. we only allow for selection on observables. Next, we propose a control estimator to correct for selection on unobservables.

The benchmark assumption is that the unobserved group effects are not correlated with the training status:

$$(10) \quad E(\varepsilon_{kmts}^{uv} | k, m, s) = 0 \quad \text{for } t=0,1$$

In this case Generalised Least Squares (GLS) can efficiently estimate the parameters in regression equation (9).<sup>20</sup> If, as suggested in Section 3.2, we use  $\pi_t$  as an instrument for  $\pi_p$  the estimator is a Generalised Instrumental Variable (GIV) estimator. Because we have grouped data, we can use a  $\chi^2$ -goodness-of-fit test to evaluate the specification of the model. This statistic can be interpreted as a GMM test for the null hypothesis of error-regressor orthogonality and therefore of selectivity bias in the training impact.

We now allow for the unobserved group effects  $\varepsilon_{kmts}^{u3}$  for destination state 3 to be correlated with the training status. Without loss of generality, we can write the unobserved group effects for destination state 3 in the following way:

$$(11) \quad \varepsilon_{kmts}^{u3} = \eta_k + \zeta_m + \kappa_s + \rho^u + \zeta_{kmts}^{u3}$$

where we assume

$$(12) \quad E(\zeta_{kmts}^{u3}) = 0$$

This means that the regressors of Equation (9) may depend on the unobserved group effects in duration interval  $k$ ,  $\eta_k$ , in sub-region  $m$ ,  $\zeta_m$ , of eligibility type  $s$ ,  $\kappa_s$  and of the training status  $(u, t)$ ,  $\rho^u$ , but not on their interaction  $\zeta_{kmts}^{u3}$ .

The first three terms in (11) can be neglected by adding them to the corresponding regressor coefficients, i.e. respectively  $\gamma_k^v$ ,  $\beta_m^v$  and  $\psi_s^v$ . This will bias the estimators of these coefficients, but since these are not the parameters of interest, this does not matter. Aggregation of the unobserved group effects over the training status and taking expectations then yields

$$(13) \quad \sum_{t=0,1} [f_{kmts}^1 E(\varepsilon_{kmts}^{13}) + f_{km1s}^2 E(\varepsilon_{km1s}^{23})] = 0$$

<sup>19</sup> The variance-covariance matrix is given in Cockx and Bardoulat (1999), Appendix C.1.

<sup>20</sup> See Cockx and Bardoulat (1999) for details.

where  $f_{kms}^u$  is, for any given  $(k,m,s)$ , the population fraction of spells of training type  $(u,t)$ . This equation provides the identifying restriction of the treatment effects. Note that it uses (12), implying that there may not be any systematic interaction effects. It allows us to express the mean bias of non-participants in training in terms of the mean bias of participants:

$$(14) \quad E(\varepsilon_{kms}^{u3} | k,m,s) = \rho_0^1 = \frac{f_{kms}^1}{f_{kms}^1} \rho_1^1 - \frac{f_{kms}^2}{f_{kms}^1} \rho_1^2$$

for all  $(k,m,s)$ . Because we observe all unemployment spells starting between May 1989 and March 1993, the population fractions  $f_{kms}^u$  are known exactly. The aggregation in (13) therefore does not result in an errors-in-variables problem as in Deaton (1985). If we impose the KxMx2 restrictions implied by (14) on the data, we can estimate  $\rho^u$ . The mean selection bias  $\rho^u$  identified as such, is a control function which, inserted in the regression equation (9), purges the equation of the covariance between  $\varepsilon_{kms}^{u3}$  and  $\bar{\alpha}^{u3}$ :

$$(15) \quad \hat{z}_{kms}^{uv} = \gamma_k^v + \beta_m^v + \psi_s^v + \lambda_{t=0}^{uv} \pi_{kms}^{u=2} + \bar{\alpha}_{t=1}^{uv} [1 - \ln(\pi_{kms}^u)] + \rho_t^u + \zeta_{kms}^{uv} + \omega_{kms}^{uv}$$

$u=1,2$  and  $v=1, 2, 3 \neq u$ . The control function estimator of the parameters of (15), restricted by (14) and taking the endogeneity of  $[1 - \ln(\pi_{kms}^u)]$  into account, is therefore a GIV estimator. In the benchmark model we impose the additional restriction that, conditional on the observable variables  $(k,m,s)$ , the bias of participating trainees is on average equal to the post-training bias:  $\rho_t^2 = \rho_t^1$ .

Finally, in the beginning of this sub-section we mentioned that the control function estimator of the impacts of training is biased if within-group unobserved heterogeneity is important. Within-group heterogeneity will generate an interaction between duration and the other explanatory variables. For, negative duration dependence induced by sorting is more important for heterogeneous groups with a high average propensity to leave unemployment than for groups with a low average propensity.<sup>21</sup> In a proportional hazard specification such an interaction is absent. Consequently, if within-group heterogeneity is important, groups with a high propensity to leave unemployment will systematically leave unemployment at a lower rate than the one predicted by the proportional hazard model. Groups with a low propensity will leave at a higher rate than predicted. These deviations should therefore generate, for any given group, positive auto-correlation in the residuals over duration. We therefore follow Cockx and Ridder's (2001) suggestion and calculate a Durbin-Watson ( $d$ ) test statistic that accounts for the panel structure of the data.

### 3.4. Over-Identifying Restrictions: Exogenous Sub-Regional Variation in the Training Supply

In regression equation (15) the treatment effects are identified on the assumption that, after aggregation, no systematic interactions between the treatment and the other explanatory variables exist. In this subsection we argue that we can exploit a natural experiment to test for this assumption. This natural experiment generates sub-regional variation in the participation rates that is not, directly or indirectly, caused by the sub-regional variation in the transition rates from unemployment. Consequently, we need not control for the sub-regional effects in the regression, i.e. we may impose the over-identifying restrictions  $\beta_m^m = 0$  for  $m = 2, \dots, M$ . These restrictions enhance the efficiency of the estimator without bias if there is indeed no

<sup>21</sup> Abbring, van den Berg and van Ours (1994) contrast an economic upturn, when workers have a high propensity to leave unemployment, to a downturn. They prove that unobserved heterogeneity indeed generates a more rapid decline of the hazard in an upturn than in a downturn.

systematic correlation between the sub-regional participation rate in training and the rate at which both trainees and non-trainees leave unemployment in a region. The validity of this assumption can be checked by a Hausman (1978) test.

The natural experiment is justified in two steps. First, we argue that the participation rate is completely determined by supply. Second, we claim that the sub-regional variation in training capacity is exogenous. The rationing of demand for training by the supply is an important stage in our argument. For, suppose that the demand were not rationed. Then, if the demand for training is correlated with the transition rate from unemployment, variations in the transition rate imply variations in the participation rate in training. If there is excess demand, variations in the transition rate will only affect the length of the queue of workers waiting for training and not the participation rate.

Interviews of employees of FOREM revealed that hardly any training programme had problems in filling the available training slots during the 1989-1993 period. Objective statistics on the importance of rationing are only partially available. We do have information on the importance of rationing of vocational training for secondary sector jobs. These represent roughly one third of the available supply of training slots. For these programmes FOREM (1991) reports the average number of workers, who have applied and are waiting either for a decision in the application procedure, or for a training slot to become available. If we report these figures in proportion to the average number of trainees entering a programme in the secondary sector each month, then we obtain a ratio of 4.1 in 1989, 5.4 in 1990, and 7.6 in 1991. This confirms that rationing is important. Note that the demand for training seems to increase significantly when labour market conditions worsen.

As a consequence of rationing, variation in the budget of the local training centres determines variation in participation rates. We mentioned higher that this budget is in no direct way related to placement ratios and other labour market outcomes. Interviews of employees of the employment agency, FOREM, confirm that the sub-regional variation in training budgets does not depend on objective criteria, but rather reflects the balance of political power of the respective sub-regions. There is one major exception to this rule. In Arlon the number of training slots in proportion to the number of unemployed is much higher than in other sub-regions (see Table 1). This is a consequence of the economic restructuring of this region. For, in response to the crisis in the steel industry, the region has benefited from important subsidies of the European Structural Funds, among which subsidies to develop training initiatives. We will therefore test below whether the results are sensitive to setting the regional effect of Arlon to zero, i.e.  $\beta_6=0$ .

Even if there is no direct effect, the sub-regional level of the transition rate from unemployment could have indirectly influenced the number of available training slots. In order to preclude this possibility, we examined whether variations in training capacities over time were in any way systematically related to the sub-regional state of the labour market. We first considered this question for the 1989-1993 observation period. In a second step, we analysed this relationship for the 1980-1988 period. This is important, because the levels in the observation period of the variables under consideration are largely determined by the past. From this analysis<sup>22</sup> we conclude that there is no indication of a systematic relationship between the regional training capacity and unemployment over time.

#### 4. The Estimation Results

In Table 2 we report the estimation results of the benchmark model in which we did not correct for selection on unobservables. Durations are measured in months. For the duration intervals we have chosen  $\{[6,8), [8,10), [10,12), [12,15), [15,18), [18,22), [22,28)\}$ . Note that we have left out the first duration interval for reasons mentioned in sub-section 3.3.

---

<sup>22</sup> See Appendix B in Cockx and Bardoulat (1999) for a more detailed discussion.

## INSERT TABLE 2 APPROXIMATELY HERE

We find negative estimates for the variances of the unobserved group effects for the transitions from origin state 2 to the two destination states ( $v = 1, 3$ ). Following Parks (1980), we set these (co-)variances equal to zero. We only find unobserved group effects for the first origin state ( $u=1$ ).

If these unobserved group effects are correlated with the training status, the estimates of the training impact are biased. However, according to the GMM test the null hypothesis of regressor-error orthogonality cannot be rejected at a significance level of 16%. We will verify below whether this claim is confirmed when we allow for selection on unobservables.

The impacts of training are the following. While participating in training the unemployed worker with average returns reduces his/her transition rate from unemployment by 7% ( $= 1 - \exp(-.07)$ ). After he/she completed training the transition rate increases by 11% ( $= \exp(.104) - 1$ ). Both effects are very significantly different from zero. However, these effects are extrapolations for the population ( $=ATE$ ), on the assumption that the treatment effects are exponentially distributed. In reality only a small minority of the unemployed workers participate in training: 2.1% of the spells starting in the 1989-93 period. The average effect of training on the treated (SATE) is therefore more reliable. We calculate that during participation the transition intensity decreases by 27%<sup>23</sup>, while the post-training effect increases this transition rate by 62%<sup>24</sup>. These results suggest that training programmes can already yield sizeable positive effects in the short-run, within the unemployment spell in which one is trained. On the basis of these findings, we simulate that the median unemployment duration<sup>25</sup> of an unemployed worker, who entered training after 6 months and who was trained during 2 months<sup>26</sup>, decreases from 21 months to 16 months. The higher speed at which unemployment is left after training more than compensates the lower exit rate during participation. In this setting, since we do not allow the impact vary with duration, the effectiveness decreases with the length of the training programme: increasing training duration to 4 or 6 months, without altering the return, increases unemployment duration of a trainee to respectively 17.5 and 19 months. However, if the training duration increases with a fixed training capacity, this should decrease the fraction of trainees and, hence, increase the returns to training.

As average returns of the unemployed population are much lower, these results also show that an extension of the training programme to more workers would significantly reduce its effectiveness. This is because the return to training decreases with the training capacity. For instance, at the 5%, 25% and 50% percentiles of the returns to training distribution the average duration of a trainee who is trained during two months after being 6 months unemployed is respectively equal to 18, 19 and 20 months, approaching thereby the average duration of 21 months in the absence of training.

In the theoretical analysis we predicted that a training programme could lengthen the spell of unemployment of non-trainees. In particular, if demand for training is rationed, sub-regions with a higher training capacity should have both, a higher proportion of unemployed workers waiting to participate in training and a higher reluctance of these workers to accept job offers. The estimated coefficient,  $\lambda^3$ , has the correct sign, but is estimated very imprecisely. The data do therefore not allow evaluating the importance of this adverse impact.

Finally, on the basis of the Durbin-Watson statistic we cannot reject the hypothesis that duration is positively auto-correlated. We suggested that this is sign of within-group

<sup>23</sup> Equation (3) contains the formula to calculate SATE. The value depends on the fraction of trainees. Averaging SATE over all cells yields  $-.313$  and  $1 - \exp(-0.313) = 0.27$ . The smallest SATE within our grouped data is  $-0.377$  and the largest  $-0.250$ .

<sup>24</sup> Averaging SATE over all cells yields  $.480$  and  $\exp(0.480) - 1 = 0.62$ . The smallest SATE within our grouped data is  $0.340$  and the largest is  $0.702$ .

<sup>25</sup> Conditional on the elapsed duration of 6 months.

<sup>26</sup> Which corresponds to the median training duration (see Table 1).



heterogeneity. In that case, the training effects are likely to be biased downwards in absolute value (see Lancaster 1990, p.65). A model that takes heterogeneity into account is matter for further research.

### INSERT TABLE 3 APPROXIMATELY HERE

In the first column of Table 3 we report the results of the GIV estimator, correcting for selection on unobservable between-group effects. We only provide information on the parameters of interest, the mean selection bias, and the GMM statistic. We did not report the other estimates, because the point estimates of all the coefficients are nearly identical. The coefficient corresponding to the mean selection bias,  $\rho^1 = \rho^2$ , is small and insignificant, indicating that there is no significant selection on unobservables. This is in accordance with the finding that the preceding model, reported in Table 2, could not be rejected on the basis of the GMM test statistic. Nevertheless, according to the Hausman test statistic (=19.48), distributed  $\chi^2$  with two degrees of freedom, the two impact effects of training,  $\bar{\alpha}_i^{u3}$ ,  $u=1,2$  of the asymptotically more efficient benchmark model reported in Table 2 are jointly to be rejected against the consistent GIV estimates (P-value=0.000).

The standard errors of coefficients measuring the effect of training increase considerably. The participation effect is no longer significantly different of zero. The post-training impact remains significantly positive. At the point estimates, the median unemployment duration of an unemployed worker, who entered training after 6 months and who was trained during 2 months, decreases now by somewhat more: by 6 instead of 5 months. The relative imprecision of the estimates calls for a sensitivity analysis of these results.

In Section 4.3 we argued that the training capacity varies exogenously between the sub-regions. The second column of Table 3 reports the GIV estimates setting the regional effects to zero:  $\beta_m^3=0$  for  $m = 2, \dots, M$ . According to the Hausman test statistic (= 21.44) these over-identifying restrictions should be rejected. However, we argued that the natural experiment might not apply to the sub-region Arlon. The third column of Table 3 shows that the point estimates are now much closer to the ones of benchmark GIV model if all, but the regional effect of Arlon, are set to zero. The Hausman test statistic (= 2.62) no longer rejects the identifying restrictions at a P-value of 27%. At the point estimates, the median unemployment duration decreases less: by 4 months, from 22 to 18 months. The robustness of the findings provides support for our modelling assumptions.<sup>27</sup>

## 5. Conclusion

On the basis of data for the 1989-93 period in Wallonia, the French-speaking region of Belgium, we find that participation in vocational classroom training programmes has on average reduced the rate of transition from unemployment by 23% to 30%. However, if the worker is still unemployed after participation in the programme, the rate at which unemployment is left is 47% to 73% higher than in the absence of training.<sup>28</sup> On the basis of these findings, we simulate that the median unemployment duration of an unemployed worker, who entered training after 6 months and who was trained during 2 months, decreases by 4 to 6 months, i.e. by 18% to 29%. Training is therefore found to considerably speed up the transition rate from unemployment. However, the analysis warns that the returns to training may be rapidly diminishing and that increasing the coverage of the programme, as proposed in the European Employment Strategy, is unlikely to produce as favourable results. For instance, at the 5%, 25% and 50% percentiles of the returns to training distribution the

<sup>27</sup> We refer the reader to Cockx and Bardoulat (1999) for a more extensive sensitivity analysis.

<sup>28</sup> These effects correspond to SATE (cf. formula in Equation (3)) calculated on the basis of the impact effects reported in columns (1) and (3) of Table 3.

median unemployment duration after 2 months of training decreases respectively by 14%, 10% and 5%.<sup>29</sup> Moreover, we argued that the programme induces non-participants to postpone their exit from unemployment. The data did not, however, allow estimating the importance of the latter effect with sufficient precision.

These estimates of the training effects are conservative, because we found that within group heterogeneity remains important. Moreover, the long-run effects, the effects of training on the incidence and duration of subsequent employment spells and on wages and/or earnings are not included. On the other hand, the effects the analysis ignored general equilibrium effects of training. These could downsize the reported impacts. However, since only 2% of the unemployed participate in training, these effects must be small.

## References

Abbring, J.H., G.J. van den Berg, and J.C. van Ours (1994), "The anatomy of unemployment dynamics", working paper n°9437, Amsterdam: Tinbergen Institute.

Amemiya Takeshi (1985), *Advanced Econometrics*, Oxford: Basil Blackwell.

Angrist, Joshua D. (1991), "Grouped-data estimation and testing simple labor-supply models", *Journal of Econometrics*, 47, 243-66.

Angrist, J. and G. Imbens (1991), "Sources of Identifying Information in Evaluation Models", NBER Technical Working Paper No. 117.

Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin (1996), "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association*, 91 (434), 444-55.

Angrist J.D. and A.B. Krueger (1999), "Empirical strategies in labor economics", in O.C. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics, Volume 3A*, Amsterdam: Elsevier Science, 1278-1366.

Ashenfelter, O. and C. Rouse (1995), "Schooling, intelligence, and income in America: Cracks in the Bell Curve", unpublished mimeograph, Princeton University.

Bollens J. and Nicaise I. (1994), "The Medium-Term Impact of Vocational Training on Employment and Unemployment Duration", paper prepared for the European Association of Labour Economists Conference, Warsaw, September.

Brodaty, Thomas, Bruno Crépon and Denis Fougère (2001), "Using Matching Estimators to Evaluate Alternative Youth Employment Programmes: Evidence from France, 1986-1988", in Lechner and Pfeiffer (eds.), *Econometric Evaluation of Labour Market Policies*, Heidelberg: Physica-Verlag, 85-123.

Calmfors, L., A. Forslund and M. Hemström (2002), "Does Active Labour Market Policy Work? Lessons from the Swedish Experiences", Working Paper 2002:4, IFAU, Institute for Labour Market Evaluation.

---

<sup>29</sup> The effects at the x% are calculated by inserting the impact effects reported in Table 2 into Equation (2). These impact estimates lie between the two extremes reported in columns (1) and (3) of Table 3. Note also that these are extrapolated effects.

Cockx, Bart (1997), "Analysis of Transition Data by the Minimum Chi-Square Method. An Application to Welfare Spells in Belgium", *The Review of Economics and Statistics*, 79(3), 392-405.

Cockx, Bart (1999), "Les formations professionnelles du FOREM accélèrent-elles la sortie du chômage en Wallonie ?", *Etude spéciale du Service d'Analyse économique de l'IRES*, septembre, Louvain-la-Neuve: IRES, Université catholique de Louvain, 105-124.

Cockx, Bart and Isabelle Bardoulat (1999), "Vocational Training: Does it Speed up the Transition Rate out of Unemployment?", Discussion Paper n° 9932, IRES, Université catholique de Louvain.

Cockx Bart and Geert Ridder (2001), "Social Employment of Welfare Recipients in Belgium: An Evaluation", *The Economic Journal*, 111 (470), 322-352.

Cockx, Bart, Bruno Van der Linden, and Adel Karaa (1998), "Active labour market policies and job tenure", *Oxford Economic Papers*, 50, 685-708.

Commission of the European Union (1997), *Conclusions de la Présidence – Luxembourg, 20/21 novembre 1997. Les "lignes directrices" pour 1998*, Bruxelles.

Deaton, A. (1985), "Panel Data from Time Series of Cross-sections", *Journal of Econometrics*, 30, 109-26.

Dehejia, Rajeev H. and Sadek Wahba (1999), "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs", *The Journal of the American Statistical Society*, 94 (448), December, 1053-62.

Eberwein, C., J. Ham and R. Lalonde (1997), "The impact of classroom training on the employment histories of disadvantaged women: evidence from experimental data", *Review of Economic Studies*, 64(4), 655-82.

Friedlander, D., D. Greenberg and P. Robins (1997), "Evaluating government training programs for the economically disadvantaged", *Journal of Economic Literature*, 35(4), 1809-55.

FOREM (1989-1994), Rapport Annuel, Bruxelles.

Hausman, J. A. (1978), "Specification Tests in Econometrics", *Econometrica*, 46, 1251-72.

Heckman, James J. (1997), "Instrumental Variables. A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations", *Journal of Human Resources*, 32(3), 441-62.

Heckman, James J. and B. Honoré (1990), "The empirical content of the Roy model", *Econometrica*, 58(5), 1121-49.

Heckman, James J., H. Ichimura, J. Smith and P. Todd (1997), "Matching as an econometric evaluation estimator: evidence from evaluating a job training program", *Review of Economic Studies*, 64(4), 605-54.

Heckman, James J., R. J. Lalonde and J.A. Smith (1999), "The Economics and Econometrics of Active Labor Market Programs", in Ashenfelter, O. and D. Card (eds.), *The Handbook of Labor Economics, Volume 3A*, Amsterdam: Elsevier Science, 1865-2097.

Heckman, James J. and Richard Robb, Jr. (1985), "Alternative Methods for Evaluating the Impact of Interventions", in Heckman, James J. and B. Singer (eds.), *Longitudinal Analysis of Labor Market Data*, Econometric Society Monograph series, New-York: Cambridge University Press.

Imbens Guido W. and Joshua D. Angrist (1994), "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, 62(2), 467-75.

Lalonde, R. (1995), "The Promise of Public Sector-Sponsored Training Programs", *Journal of Economic Perspectives*, 9(2), Spring, 149-68.

Lancaster, Tony (1990), *The Econometric Analysis of Transition Data*, Cambridge: Cambridge University Press.

Lechner, M. (1999), "Earnings and Employment Effects of Continuous Off-the-Job Training in East-Germany after Unification", *Journal of Business and Economic Statistics*, 17, 74-90.

Martin, J. P. et D. Grubb (2001), "What works and for whom: a review of OECD countries' experiences with active labour market policies", Working Paper 2001:14, IFAU Office of Labour Market Policy Evaluation.

Meyer, Bruce D. (1995), "Natural and Quasi-Experiments in Economics", *Journal of Business and Economic Statistics*, 13(2), 151-61.

OECD (1993), *Employment Outlook*, Paris.

OECD (1997), *Le service publique de l'emploi. Belgique*, Paris.

Parks, Richard W. (1980), "On the Estimation of Multinomial Logit Models with Relative Frequency Data", *The Journal of Econometrics*, 13, 293-303.

Plasman (1993), "Estimation de Durée de Chômage et Rôle des Politiques d'Emploi", in Point d'Appui Emploi Formation (ed.), *Chômage de Longue Durée – Comment en parler? Comment le combattre?*, dossier 4, Bruxelles.

Plasman (1994), *Les Politiques du Marché du Travail: Analyse et Comparaisons européennes; Procédure d'Evaluation (micro et macroéconomique); Evaluation des Politiques de Resorption du Chômage et des Politiques du Temps de Travail en Belgique*, Thèse de Doctorat, Université Libre de Bruxelles, Bruxelles.

Prentice, R. L. and L.A. Gloeckler (1978), "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data", *Biometrics*, 34, 57-67.

Roy, A.D. (1951), "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers (New Series)*, 3, 135-46.

Salant, S.W. (1977) "Search Theory and Duration Data: A Theory of Sorts", *Quarterly Journal of Economics*, 91, 39-57.

Van Langendonck, J. (1991), *Handboek Sociale Zekerheid*, Acco, Leuven.

Wald, A. (1940), "The fitting of straight lines if both variables are subject to error", *Annals of Mathematical Statistics*, 11, 284-300.

**Table 1.:** Population Characteristics

|                                      | Kind of unemployment spells |  |            |  |
|--------------------------------------|-----------------------------|--|------------|--|
|                                      | All Spells                  | Old<br>Spells with<br>some time in<br>training | All Spells | young<br>Spells with<br>some time in<br>training |
| <i>Number of spells</i>              | 1 048 252                   | 23 407   | 313 408    | 5 284  |
| <i>Eligibility Type</i>              |                             |  |            |  |
| Old                                  | 77.0%                       | 81.6%  | 0%         | 0%   |
| <i>Sex</i>                           |                             |  |            |  |
| Male                                 | 46.2%                       | 55.4%  | 42.2%      | 39.2%  |
| <i>Age group</i>                     |                             |  |            |  |
| <=20 years                           | 6.0%                        | 5.9%   | 60.2%      | 49.8%  |
| 21-25 years                          | 31.3%                       | 29.8%  | 39.8%      | 48.8%  |
| 26-30 years                          | 24.0%                       | 25.3%  | 0.9%       | 1.4%   |
| 31-40 years                          | 26.2%                       | 29.7%  | -          | -  |
| 41-50 years                          | 12.6%                       | 9.3%   | -          | -  |
| Average age (years)                  | 29.8                        | 29.5   | 20.3       | 20.8   |
| <i>1st unemployment registration</i> |                             |  |            |  |
| Yes                                  | 15.8%                       | 11.6%  | 90.3%      | 90.8%  |
| <i>Educational attainment</i>        |                             |  |            |  |
| Primary (6years)                     | 31.3%                       | 22.5%  | 11.8%      | 8.0%   |
| Lower Secondary (9 years)            | 29.6%                       | 26.7%  | 29.0%      | 21.0%  |
| Higher Secondary (12 years)          | 22.2%                       | 26.1%  | 36.3%      | 45.1%  |
| Higher Education (>12 years)         | 11.0%                       | 16.5%  | 18.9%      | 24.0%  |
| Other or unknown                     | 5.8%                        | 8.2%   | 4.0%       | 1.9%   |
| <i>Sub-regional department</i>       |                             |  |            |  |
| Nivelles and La Louvière (m=1)       | 17.0%                       | 14.0%  | 16.1%      | 17.2%  |
| Charleroi (m=2)                      | 16.8%                       | 14.2%  | 18.8%      | 12.2%  |
| Mons (m=3)                           | 8.9%                        | 7.3%   | 8.9%       | 9.9%   |
| Mouscron and Tournai (m=4)           | 9.4%                        | 12.4%  | 8.9%       | 11.9%  |
| Liège and Verviers (m=5)             | 27.9%                       | 27.1%  | 25.8%      | 24.4%  |
| Arlon (m=6)                          | 5.4%                        | 9.5%   | 6.00%      | 12.1%  |
| Namur and huy (m=7)                  | 14.6%                       | 15.6%  | 15.4%      | 12.1%  |
| <i>Median Duration in months</i>     |                             |  |            |  |
| Unemployment Duration                | 3                           | 9  | 3          | 12   |
| Duration until Training              | -                           | 3  | -          | 6  |
| Time Spent on Training               | -                           | 2  | -          | 2  |

**Table 2:** Estimates of the Benchmark Model –No Correction for Selection on Unobservables  
(standard error in parenthesis)

|  | Destination state $v$           |                             |                        |
|--|---------------------------------|-----------------------------|------------------------|
|  | 1: Unemployed not participating | 2: Unemployed participating | 3: Out of Unemployment |
| <i>Duration interval (months)</i>                |                                 |                             |                        |
| 6-8 ( $\gamma_2^v$ )                             | -2.173*<br>(.041)               | -5.951*<br>(0.127)          | -3.094*<br>(.061)      |
| 8-10 ( $\gamma_3^v - \gamma_2^v$ )               | .365*<br>(.040)                 | .154*<br>(.075)             | .132*<br>(.041)        |
| 10-12 ( $\gamma_4^v - \gamma_2^v$ )              | .258*<br>(.043)                 | -.100<br>(.078)             | -.154*<br>(.042)       |
| 12-15 ( $\gamma_5^v - \gamma_2^v$ )              | .552*<br>(.040)                 | .013<br>(.075)              | .044<br>(.041)         |
| 15-18 ( $\gamma_6^v - \gamma_2^v$ )              | .497*<br>(.043)                 | -.177*<br>(.076)            | -.202*<br>(.042)       |
| 18-22 ( $\gamma_7^v - \gamma_2^v$ )              | .688*<br>(.045)                 | -.273*<br>(.077)            | -.125*<br>(.043)       |
| 22-28 ( $\gamma_8^v - \gamma_2^v$ )              | .686*<br>(.056)                 | -.586*<br>(.092)            | -.364*<br>(.051)       |
| <i>Sub-region<sup>1</sup></i>                    |                                 |                             |                        |
| Charleroi ( $\beta_2^v - \beta_1^v$ )            | -.072*<br>(.043)                | .143*<br>(.077)             | -.120*<br>(.044)       |
| Mons ( $\beta_3^v - \beta_1^v$ )                 | -.050<br>(.050)                 | .200*<br>(.081)             | -.081*<br>(.046)       |
| Mouscron and Tournai ( $\beta_4^v - \beta_1^v$ ) | -.093*<br>(.045)                | .296*<br>(.092)             | .030<br>(.046)         |
| Liège and Verviers ( $\beta_5^v - \beta_1^v$ )   | -.082*<br>(.038)                | .161*<br>(.073)             | .040<br>(.040)         |
| Arlon ( $\beta_6^v - \beta_1^v$ )                | .245*<br>(.046)                 | .451*<br>(.151)             | .327*<br>(.058)        |
| Namur and Huy ( $\beta_7^v - \beta_1^v$ )        | .057<br>(.042)                  | .331*<br>(.076)             | .045<br>(.043)         |
| <i>Eligibility type</i>                          |                                 |                             |                        |
| Young ( $\psi^v$ )                               | .101*<br>(.028)                 | .227*<br>(.045)             | .241*<br>(.024)        |
| <i>Impact of Training</i>                        |                                 |                             |                        |
| Before participation ( $\lambda_v$ )             |                                 | 4.782<br>(4.188)            | -.242<br>(1.681)       |
| During participation ( $\bar{\alpha}_1^{23}$ )   |                                 | -                           | -.070*<br>(.014)       |
| Post participation ( $\bar{\alpha}_1^v$ )        |                                 | 2.324*<br>(.043)            | .104*<br>(.006)        |
| <i>Unobserved group effects</i>                  |                                 |                             |                        |
| $\hat{\sigma}_{1vw}(\hat{\sigma}_{1vw})$         | -                               | .0404 (.0034)               | .0189 (.0034)          |
| $\hat{\sigma}_{2vw}(\hat{\sigma}_{2vw})$         | 0 (0)                           | -                           | 0 (0)                  |
| WSSR   |                                 | 527.73                      |                        |
| (degrees of freedom)                             |                                 | (499)                       |                        |
| P-value  |                                 | .180                        |                        |
| Durbin-Watson statistic ( $d$ )                  |                                 | 1.2597                      |                        |
| (number of cells)                                |                                 | (546)                       |                        |

<sup>1</sup> The reference sub-region is Nivelles and La Louvière

\*significant at the 5% level

**Table 3:** Testing Over-Identifying Restrictions  
(standard errors in parentheses)

|   | (1)<br>The Benchmark<br>GIV | (2)<br>Without region<br>dummies<br>$\forall i :$<br>$\beta^3=0$ | (3)<br>Without region<br>dummies except for<br>Arlon $\beta^3$ |
|---|-----------------------------|--|--|
| <b>Impact of Training</b>                                 |                             |  |  |
| <i>Before participation</i> ( $\lambda^{13}$ )            | - .416<br>(1.899)           | 8.548*<br>(1.112)  | 2.034<br>(1.500)   |
| <i>During participation</i> ( $\alpha_1^{23}$ )           | -.058<br>(.053)             | -.120*<br>(.053)   | -.081<br>(.050)  |
| <i>Post participation</i> ( $\alpha_1^{13}$ )             | .118*<br>(.058)             | .001<br>(.052)   | .083<br>(.053)   |
| <b>Mean selection bias</b>                                |                             |  |  |
| <i>During = post participation</i><br>( $\rho^2=\rho^1$ ) | -.048<br>(.243)             | .450*<br>(.219)  | .111<br>(.220)   |
| <b>WSSR</b><br>(degrees of freedom)                       | 530.84<br>(498)             | 531.62<br>(504)  | 548.49<br>(503)  |
| <b>P-value</b>  | .149                        | .191   | .079   |
| <b>Hausman test<sup>1</sup>: <math>\chi^2(2)</math></b>   | 19.48                       | 21.44  | 2.62   |
| <b>P-value</b>  | .000                        | .000   | .270   |

\*significant at the 5% level

<sup>1</sup> Tests whether the asymptotically more efficient estimators of the 2 impact effects of training,  $\alpha_1^{23}$ , of model (2) and (3) are jointly to be rejected against the consistent estimators of the corresponding parameters in the baseline model (1). The statistic in column (1) tests whether the model without correction for selection bias of Table 2 is to be rejected against model (1).

## IZA Discussion Papers

| No. | Author(s)                                  | Title   | Area | Date  |
|-----|--|---|------|-------|
| 666 | P. Arnds<br>H. Bonin                       | Frühverrentung in Deutschland: Ökonomische Anreize und institutionelle Strukturen                               | 7    | 12/02 |
| 667 | P. Arnds<br>H. Bonin                       | Arbeitsmarkteffekte und finanzpolitische Folgen der demographischen Alterung in Deutschland                     | 7    | 12/02 |
| 668 | J. Meckl<br>S. Zink                        | Solow and Heterogeneous Labor: A Neoclassical Explanation of Wage Inequality                                    | 3    | 12/02 |
| 669 | A. C. D'Addio<br>I. De Greef<br>M. Rosholm | Assessing Unemployment Traps in Belgium Using Panel Data Sample Selection Models                                | 2    | 12/02 |
| 670 | M. Botticini<br>Z. Eckstein                | From Farmers to Merchants: A Human Capital Interpretation of Jewish Economic History                            | 5    | 12/02 |
| 671 | A. Constant<br>S. Konstantopoulos          | School Effects and Labor Market Outcomes for Young Adults in the 1980s and 1990s                                | 1    | 12/02 |
| 672 | A. Constant<br>D. Massey                   | Self-Selection, Earnings, and Out-Migration: A Longitudinal Study of Immigrants to Germany                      | 1    | 12/02 |
| 673 | A. L. Booth<br>M. Francesconi<br>J. Frank  | Labour as a Buffer: Do Temporary Workers Suffer?  | 2    | 12/02 |
| 674 | D. A. Cobb-Clark<br>V. Hildebrand          | The Wealth and Asset Holdings of U.S.-Born and Foreign-Born Households: Evidence from SIPP Data                 | 6    | 12/02 |
| 675 | J. Konings<br>O. Kupets<br>H. Lehmann      | Gross Job Flows in Ukraine: Size, Ownership and Trade Effects   | 4    | 12/02 |
| 676 | A. Voicu                                   | Agriculture: Transition Buffer or Black Hole? A Three-State Model of Employment Dynamics                        | 4    | 12/02 |
| 677 | F. Galindo-Rueda<br>A. Vignoles            | Class Ridden or Meritocratic? An Economic Analysis of Recent Changes in Britain                                 | 2    | 12/02 |
| 678 | M. Barbie<br>M. Hagedorn<br>A. Kaul        | Fostering Within-Family Human Capital Investment: An Intragenerational Insurance Perspective of Social Security | 3    | 12/02 |
| 679 | A. Kölling<br>C. Schnabel<br>J. Wagner     | Establishment Age and Wages: Evidence from German Linked Employer-Employee Data                                 | 1    | 12/02 |
| 680 | M. Hagedorn<br>A. Kaul                     | Langzeitarbeitslosigkeit in Deutschland: Fakten, Ursachen und Bekämpfung  | 7    | 12/02 |
| 681 | M. Hagedorn<br>A. Kaul<br>T. Mennel        | An Adverse Selection Model of Optimal Unemployment Insurance  | 7    | 12/02 |
| 682 | B. Cockx                                   | Vocational Training of Unemployed Workers in Belgium  | 2    | 01/03 |