Evaluation of the Impact of Training on Individual Labor Market Transitions

Sylvie Blasco[†] Bruno Crépon[‡] Thierry Kamionka[§]

January 2008

Preliminary.

Abstract

In this paper we evaluate the effect of a participation in a training program on the employment and unemployment duration distributions and, more generally, on the mobility between states of the labor market. We use a French survey representative of the labor force population to estimate a multi-spell multi-state transition model. We distinguish two categories of training: training on unemployment and on employment. Participation in such programs and their duration are taken as endogenous variables. We allow training participation to have an impact on the labor market transitions up to 12 months after completion, so that we study both current and past duration and state dependences. We model unobserved heterogeneity to distinguish between true and spurious dependences. We find that past participation in training programs increases the conditional probability of return to employment.

Keywords : Training, Transition Models, Unobserved Heterogeneity.

JEL Classification : J64, C41, C10, J69.

[†] CREST and EHESS (sylvie.blasco@ensae.fr).
[‡] CREST-INSEE, CEPR and IZA (bruno.crepon@ensae.fr).
[§] CNRS and CREST (kamionka@ensae.fr).
Telephone : (33) (0)1 41 17 35 51. Address: CREST. Timbre J390. 15, Boulevard Gabriel Péri. 92245 Malakoff Cedex. France.

1 Introduction

Professional training is a privileged tool used to improve the human capital of the less qualified workers and to facilitate their mobilities on the labor market. The financial efforts made by the public services and firms to improve the ability of workers are realized assuming that participation in such programs has a positive impact on the productivity and improves the employment prospects of the participants. However the existing econometric evaluations of training do not always support such an assumption (see Heckman and al., 1999 and Crépon and al., for a survey).

It is uneasy to obtain clear results on the impact of trainings. First, there is no homogeneity in training programs. Several categories of training should be distinguished, among which training programs offered by public services of employment to unemployed workers, training offered by firms to their employees, apprenticeship, training dedicated to young or non qualified workers, for instance. The effect of training depends on the kind of training evaluated, on the targeted population and on the output considered. It also differs according to the characteristics of the program, i.e. according its duration or whether it is qualifying or not. In the USA and in Great Britain, the wage returns of continuous training appears to be significant and positive (Parent, 1999 and 2003; Blundell and al. 1999). The same occurs in Germany where apprenticeship is shown to have a strongly positive impact on the wages of the beneficiaries, especially if participation concludes to a diploma. In other European countries on the contrary, econometric evaluations tend to show small or insignificant effects of training on wages (Pischke, 2001; Gerfin, 2003). In the particular case of France, training does not significantly increase the wages of participants (Goux and al., 2000 and Fougère and al., 2001)[1]. The fewer econometric evaluations of the impact of training participation on employment prospects and labor market history also lead to unclear conclusions. The impact of training on the labor market trajectories has been shown to have strong positive effect for young workers (Mealli and al; 1999). However, it has not univocal effect for the unemployed workers as it increases the unemployment duration but also increases the duration of the job obtained following the unemployment spell (Crépon and al; 2007).

Another difficulty encountered when one aims at evaluating the impact of training on labor market transitions is that training has time-varying effects, as it clearly appears in the case of unemployment training (Crépon and al.; 2007). Last but not least, the entry in training programs is governed by a selection process. It is well known that the probability of access in training programs differs according to observable and unobservable characteristics (Bassanini and al., 2005; OECD, 2003). The selection issue is all the more important in France because the selection of participants is given to the firms (Crépon and al.; Goux and al., 1997). As a re-

[1]However, Fougère and al. show that training diminishes the wage loss which occurs on average following an inter-firm mobility

sults, participation in training programs is an endogenous variable and one should consider the selectivity phenomena to obtain unbiased estimates of the effect of treatment on the treated.

In this paper, we take into account these three issues to evaluate the impact of training on the situation of the individuals on the labor market. We use a rich French survey devoted to the study of training and professional qualifications (FQP survey). The FQP survey was collected in 2003 by the French institute for statistics and economic studies (INSEE). This data set allows us to evaluate the impact of various training programs on the labor market transitions for a sample individuals representative of the French labor force. As the interviewed individuals describe their labor market history over 60 months across several states, we estimate a multi-state multi-spell transition model and identify the effects of the participation in training program on the unemployment duration, on the job conditional duration distribution and, more generally, on the mobilities on the labor market. We model all the transitions, so that the participation in training programs and the duration of the programs are explicitly taken as endogenous. We allow participation to have an impact on the labor market transitions up to 12 months after completion. We thus consider both current and lagged state dependences, as well as current and lagged duration dependences (Heckman and al., 1980). As we model unobserved heterogeneity and initial conditions *à la* Wooldridge (2005), we distinguish true and spurious dependences. We specify the transition probabilities using a model directly related to mixed proportional hazard duration model in order to be able to interpret straightforwardly the results. Using interval-censored data, it appeared to be difficult to use a timing-of-events approach (see Abbring and van den Berg, 2003). The difficulty comes from the fact that, in order to evaluate the impact of the training on the current state duration, we should consider training as a sub-spell of the employment or unemployment spell (Crépon and al.; 2007), and not as a separate state as it is commonly defined. The presence of interval-censoring, complicates greatly the expression of likelihood function and makes the estimation uneasy, even if we postulate constant hazards. The specification we use, allows to consider directly the relation between conditional hazard function and transition probabilities.

This paper is organized as follow. The section 2 contains a descriptive statistics. The third section includes a presentation of the econometric model. A discussion of the results is presented in the fourth section. The last section concludes.

2 The data set

We use the French survey FQP[2] collected in 2003 by the French national institute (INSEE). These data are nationally representative of the French population aged

[2] *"Formation et Qualification Professionnelles"*.

between 17 and 65 years old. It is realized in order to tackle topics such as training, qualification and professional, social and geographical mobilities. Through retrospective calendars, it gives, on a monthly basis, rich information on the situation on the labor market at the moment of the interview, that is on the period going from March to July 2003, and during the 5 previous years. As a result, we have a 60 months follow-up for each interviewed individual.

This work mostly relies on the professional and training calendars. The professional calendar lists all the individual mobilities on the labor market during the period of observation and allows the reconstruction of a detailed history. A new period in the calendar is indeed motivated by a transition from/to non employment, but also by a change in the characteristics of the job: any change in the contract, any change of firm, establishment or post, involves the creation of a new employment spell. As a result, the transitions from/to unemployment and inactivity, but also the transitions from job to job are registered in the calendar. We have a rich information on the characteristics of the employment spell, among which the precise motivations for the end of a spell of employment (resignation, lay off, end of the term of the contract, person at the origin of the separation) and the type of contract (short, temporary or permanent). The wage associated with each job is not available. We only know whether remuneration stagnates, decreases or increases with the transition. We can deduce the duration of each spell.

The part of the calendar relative to training gives information on all the trainings which duration is higher than 30 hours followed by the individual from 1998 to 2003[3]. Several categories of training can be distinguished, such as training on employment, on unemployment, apprenticeship and self-training. We know the effective duration and length of each training listed in the calendar, and it is possible to determine the category and the main characteristics for the great majority of them. Table 1 gives a brief description of the considered trainings.

Table 1: **Descriptive statistics by training category**

	Unemployment trainings	Employment trainings	All trainings
number of trainings	1404	6282	7686
type of training			
work-training	462	1206	1668
internship/seminars	451	2199	2650
self-training	46	83	129
unknown	318	1580	1580
% qualifying	26.84	12.38	15,09
% specialized	83.84	98,92	97,96

[3]The survey gives also information on shorter trainings, but we do not use it because of its non-exhaustiveness

As we aim at, among others, evaluating the impact of training on the risk and duration of unemployment, we select a sample composed by individuals aged between 17 and 64 years old, who have finished school before May 1998, are not retired on May 2003 and who are not civil servant, nor self-employed[4]. We do not consider students and retired people to clarify the definition of the state "non employment". Indeed, these kinds of inactivity are very specific and are not at the center of our interest. We do not consider the observations corresponding to civil servant and self-employed because their labor market transition processes are specific. We obtain a sample of 26239 individuals.

2.1 Transition matrix

We consider 4 states: employment, non employment, employment training and unemployment training. We classify training as an employment training or an unemployment training depending on the state held at the moment of the training period. We aggregate the unemployment and out-of-labor-force spells in a "non employment" state because the data do not allow to distinguish the transitions between unemployment and nonparticipation[5]. We first consider employment as an aggregate spell, without taking into account the job to job transitions: if an individual holds for example 5 different jobs without experiencing any transition to non employment, then we assume that he occupies only one state. As a consequence, we evaluate the impact of training on the persistence of employment. A further research could consist in allowing for the transitions from job to job, in order to investigate the issue of interfirm mobilities.

Table 2: **Unconstrained transition matrix (number and %)**

$t \rightarrow$ $\downarrow (t-1)$	Empl.	Non Empl.	Empl. Training	Unempl. Training	Total (row)
Empl.	1042109 98,62%	8759 0,83%	5692 0,54%	132 0,01%	1056692 100%
Non Empl.	7450 1,58%	463961 98,17%	124 0,03%	1055 0,22%	472590 100%
Empl. Training	6347 18,54%	53 0,15%	27555 80,49%	277 0,81%	34232 100%
Unempl. Training	166 1,53%	1143 10,56%	285 2,63%	9230 85,27%	10824 100%

[4] But we keep non civil servant employed by the State.

[5] We know whether the individual enters into unemployment or leaves the labor force at the end of the job spell, but we cannot determine whether he stayed in the same state until his next employment spell.

As expected, the transition matrix exhibits strong state inertia. The probability of exiting the occupied state the following month is however greater when the individual is in a training program. The monthly probability of entering in an on employment (resp. unemployment) training program is quite low, accounting for less than 1%, given that the individual occupies a job (resp. is unemployed). 1,5% of the unemployed workers who participate in an unemployment training program a given month are employed the following month and 10,6% return to open unemployment. Only less than 1% of the worker who are in training become unemployment the following month and 18,5% are still on the job.

As the calendar is filled on a monthly basis, we encounter an interval-censoring issue. This can be corrected using an appropriate specification of the transition probabilities (see section 3). The main issue involved by this interval-censoring is that very short spells - which duration is beyond one or two weeks - may not be listed. This may explain why we observe in the data some unusual transitions (Table 2). For example, 132 individuals are observed making a transition from employment to unemployment training. It is unlikely than these workers entered unemployment training the very first day of their unemployment spell. Such observed transitions may rather stand for a transition from employment to unemployment and then a transition from unemployment to training within the same month. When it is possible, we have corrected these observations to reveal the actual trajectories. We thus imposed some constraints on the transitions. The constrained transition matrix (Table 3) is quite similar to the unconstrained one described above. The following econometric analysis is realized on the data summarized by this transition matrix below.

Table 3: **Constrained transition matrix (number and %)**

$t \rightarrow$ $\downarrow (t-1)$	Empl.	Non Empl.	Empl. Training	Unempl. Training	Total (row)
Empl.	1041877 98,55%	9203 0,87%	6029 0,57%	-	1057109 100%
Non Empl.	8001 1,69%	464235 98,02%	-	1390 0,29%	473626 100%
Empl. Training	6609 19,68%	-	26965 80,32%	-	33574 100%
Unempl. Training	-	1519 15,12%	-	8530 84,88%	10049 100%

Table 4 in appendix A shows the transition matrix conditional on individual observed characteristics. The share of people entering in training programs each month differs depending on gender, age and level of education. The same, the monthly probability of staying in the on employment training program varies with the individual characteristics. This runs along with the ideas that participation in training is ruled by a selection process (at least on observable characteristics) and that the duration of participation is not distributed equally among individuals with different characteristics.

2.2 Duration analysis

We first run stratified Kaplan-Meier estimates of the survival function of the durations of employment and unemployment for each spell. We assume here that training participation is exogenous.

Now, to take into account the possible endogeneity of training and reveal the existing correlation between participation in a training program and the duration of a given spell, we estimate bivariate probit models. These models allows to consider the endogeneity of the participation to training programs. Let us consider the first spell. In order to simplify the notations, we are going to omit the index of the individual. The individual occupies state k on May 1998. It is an employment spell if $k = e$, or an unemployment spell if $k = u$. Let us consider a subsample of individual occupying state k ($k = e, u$). Let us denote U_k^* the duration of the spell occupied by the individual on may 1998. We are interested in the impact of a participation in a training program on the conditional probability that this spell lasts at least D months. We note T_k the dichotomic variable describing the participation to training. Let $U_k^* = X_k'\beta_k + T_k\delta_k + \epsilon_k$ and $T_k^* = Z_k'\gamma_k + \nu_k$ be the corresponding latent variables.

$$U_k = \begin{cases} 1, \text{ if } U_k^* = X_k' \, \beta_k + T_k \, \delta_k + \epsilon_k \geq D, \\ 0, \text{ otherwise.} \end{cases}$$

and

$$T_k = \begin{cases} 1, \text{ if } T_k^* = Z_k'\gamma_k + \nu_k \geq 0 \\ 0, \text{ otherwise,} \end{cases}$$

where $(\epsilon_k; \nu_k)'$ are i.i.d. $\mathrm{N}(0, \Sigma)$, $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, X_k and Z_k are vectors of explanatory variables ($k = e, u$).

The corresponding contribution to the likelihood function is

$$\ell(\theta) = \Phi_2(y_1 \, (X_k'\beta_k + T_k\delta_k), y_2 \, Z_k'\gamma_k; y_1 y_2 \rho)$$

where $y_1 = (2U_k - 1)$ and $y_2 = (2T_k - 1)$. We instrument the endogenous treatment variable T_k, using the local share of trained people who are in state k in May

1998[6].

Results are displayed in Appendix A. They show that participation in training programs significantly increases the employment persistence (columns (3) and (4)), but not the length of the spell where training occurs (columns (1) and (2)). **[...]** The selection equation reveals that women are less likely to be trained than men. The higher the level of education, the greater the probability of being trained. The probability of participate in training increases with the local share of individuals who are trained. This reveals the existence of some unobservable local specificities in the use of training programs.

According to this preliminary transition and duration analysis, training participation and the duration of the employment and unemployment spells are positively correlated. Training seems to affect labor market history on the long term, and does not seem to only affect the duration of the spell where participation takes place. The probability of participation depends on individual and local characteristics. In this section we do not use the panel structure of our data. In the following, we use all the information the data contain and use the multi-spells to correct the selection bias we are facing with. Moreover, until now we have focused on the impact of training participation on the duration of the spell where participation occurred. We now present an econometric analysis which allows to evaluate the impact of training on a longer-term.

3 Modeling transitions

3.1 Labor market participation process

A history of a given individual can be represented by a sequence of realizations of a discrete time stochastic process Y_t, $t \in \{1, \dots, 60\}$, taking its value in a discrete-state space $E = \{1, 2, 3, 4\}$. Y_t is the state occupied by the individual at time t. Let us assume that the realizations of the process are independent and identically distributed. We omit the index of the individuals. Let $\{y_t, 1 \leq t \leq 60\}$ be a realization of the process. Let us assume that

$$y_t = \begin{cases} 1, & \text{if the individual is employed at time } t, \\ 2, & \text{if the individual is non employed at time } t, \\ 3, & \text{if the individual is on employment training at time } t, \\ 4, & \text{if the individual is on unemployment training at time } t, \end{cases}$$

where $t \in \{1, \dots, 60\}$.

[6]Because of a too small number of observations, we are not able to calculate these shares on the sole individuals who have common relevant observable characteristics with the individual i, i.e. gender, level of education or age.

This is a discrete-time discrete-state labor market participation process (see, for instance, Fougère and Kamionka, 2008, Heckman 1981, Lancaster 1990).

The initial time $t = 1$ does not correspond to the date of entry into the labor market for all the individuals in the sample. At the beginning of the period of observation, individuals are not all localized at the same point in their transition process. The beginning of this process, from the end of schooling up to the state occupied on May 1998, is unobserved.

Consequently, we have to consider carefully the problem of initial conditions in estimating this discrete time stochastic process. Two approaches can be used in order to solve this problem. The first approach is proposed by Heckman (1981) and consists in modeling the initial conditions. The other method is proposed by Wooldridge (2005) and consists in modeling the unobserved heterogeneity conditionally on the initial conditions of this labor force participation process. In this paper we consider the method proposed by Wooldridge (2005), because this method is flexible and simple to implement[7].

In order to consider past dependence, we assume that training plays a role up to 12 months. This assumption has an economic interpretation : it consists in assuming that the human capital acquired in a training depreciates over time and is lost after a year. This may be restrictive, but one may argue that strictly speaking, the knowledge learned in a training period has short-term effects on the productivity and that the longer-term effects are rather due to learning-by-doing.

Let $\ell(\theta \mid y_1, \ldots, y_{12}, x; \nu)$ denote the conditional contribution of the individual to the likelihood function. x is a vector of exogenous variables and ν is a vector of unobserved heterogeneity terms. The typical conditional contribution has the following form:

$$\ell(\theta \mid y_1, \ldots, y_{12}, x; \nu) = \prod_{t=13}^{60} P(Y_t = y_t \mid y_{t-1}, \ldots, y_{t-12}, x, \nu; \theta),$$

where θ is a vector of parameters ($\theta \in \mathbb{R}^p$).

Let $\phi_t^{NE} = \frac{1}{12} \sum_{k=0}^{11} \mathbb{1}[y_{t-k} = 2]$ denote the fraction of time the individual has occupied the non employment state during the last twelve months. Similarly, let $\phi_t^{ET} = \frac{1}{12} \sum_{k=0}^{11} \mathbb{1}[y_{t-k} = 3]$ denote the fraction of time spent in the employment training state during the last twelve months and let $\phi_t^{UT} = \frac{1}{12} \sum_{k=0}^{11} \mathbb{1}[y_{t-k} = 4]$ represent the fraction of time spent in the unemployment training state for the same months. The three components ϕ_{t-1}^{NE}, ϕ_{t-1}^{ET} and ϕ_{t-1}^{UT} can be considered as appropriate abstracts of the previous history of the individual on the labor market. Let $\phi_t = (\phi_t^{NE}, \phi_t^{ET}, \phi_t^{UT})'$.

We assume that conditionally on the characteristics of the individual (z, ν), on the previous state occupied by the individual y_{t-1} and given the most recent realizations of the components ϕ_{t-1}^{NE}, ϕ_{t-1}^{ET} and ϕ_{t-1}^{UT}, the state occupied by the

[7] Edon and Kamionka (2008), show that in the case of a dynamic probit model the two methods produce similar results.

individual at time t is independent from older history of the process y_{t-j}, where $j \geq 2$. This assumption can be written as follows:

Consequently, for $t = 13, \ldots, 60$, $j \in I\!N$ and $j \geq 2$,

$$Y_t \perp\!\!\!\perp Y_{t-j} \mid y_{t-1}, \phi_{t-1}, \phi_{12}, x, \nu.$$

The assumption implies that the conditional contribution to the likelihood is

$$\ell(\theta \mid y_1, \ldots, y_{12}, x; \nu) = \prod_{t=13}^{60} P(Y_t = y_t \mid y_{t-1}, \phi_{t-1}, \phi_{12}, x, \nu; \theta),$$

or equivalently as

$$\ell(\theta \mid y_1, \ldots, y_{12}, x; \nu) = \prod_{t=13}^{60} \prod_{j=1}^{4} \prod_{k \in E_j} P(Y_t{=}k \mid y_{t-1}{=}j, \phi_{t-1}, \phi_{12}, x, \nu; \theta)^{\delta_{jkt}}, \tag{1}$$

where

$$\delta_{jkt} = \begin{cases} 1, & \text{if } y_{t-1} = j \text{ and } y_t = k, \\ 0, & \text{otherwise.} \end{cases}$$

and $E_j \subset E$ is a subset of states (all transitions are not possible, see section 2). For instance, if the individual occupies state 1 (employment state), she/he can leaves this state to occupy state 2 (non employment) or state 3 (employment training). Consequently, $E_1 = \{2, 3\}$[8].

Thus, a typical contribution to the likelihood function is

$$\ell(\theta) = \int_{\Omega} \ell(\theta \mid y_1, \ldots, y_{12}, x; \nu)\, g(\nu; \theta)\, \eta(d\nu),$$

where $g(\nu; \theta)$ is a density function[9] of the distribution of the unobserved components vector V with respect to a σ-finite measure $\eta(d\nu)$[10].

We assume that conditionally on the observed characteristics of the individual (x) and given the percentage of time spent in each state of the labor market during the initial year, ϕ_{12}^{NE}, ϕ_{12}^{ET} and ϕ_{12}^{UT}, the unobserved heterogeneity component V is independent from the state occupied by the individual a particular month of the beginning year. This assumption means that, conditionally to the number of months the individual initially has stayed in a given state, the realization of the

[8]Let us assume that $E_2 = \{1, 4\}$, $E_3 = \{1\}$ and $E_4 = \{2\}$. Finally, a total of 6 transitions between distinct states are examined.

[9]Wooldridge (2005) proposes to consider the conditional distribution of a vector $v = (v_{jk})$ of unobserved heterogeneity components given exogenous variables x and initial condition ϕ_{12}. The approach conduces to add the initial conditions to the list of explanatory variables in the expression of the conditional transition probabilities given observed and unobserved heterogeneity and to specify the unconditional distribution of unobserved factors.

[10]If V is continuous, η is the Lebesgue measure and, otherwise, if V is discrete, η is the counting measure.

sequence of states brings no additional information with respect to the distribution of unobserved heterogeneity.

Let V denote the unobserved heterogeneity vector. For $j = 1, \dots, 12$,

$$V \perp\!\!\!\perp Y_j \mid \phi_{12}^{NE}, \phi_{12}^{ET}, \phi_{12}^{UT}, x.$$

The assumption implies that the conditional contribution to the likelihood function is

$$\ell(\theta) = \int_{\Omega} \ell(\theta \mid y_1, \dots, y_{12}, x; \nu)\, g(\nu \mid \phi_{12}, x; \theta)\, \eta(d\nu),$$

where $\Omega \subset \mathbb{R}^q$ is the unobserved term space.

3.2 Modeling unobserved heterogeneity

Wooldridge (2005) proposes to model the unobserved heterogeneity given initial conditions. Moreover, we use a factor loading model in order to correlate in a flexible way transitions probabilities using unobserved heterogeneity.

Let v_{jk} denote the unobserved heterogeneity term specific to transition from state j to state k ($j, k \in E$). Assume that

$$v_{jk} = \lambda_{jk}\, \nu_1 + \mu_{jk}\, \nu_2 + \phi_{12}' \gamma_{jk},$$

where ν_1 and ν_2 are two unobserved random components ($\nu = (\nu_1, \nu_2)'$) and $\phi_{12} = (\phi_{12}^{NE}, \phi_{12}^{ET}, \phi_{12}^{UT})'$. $\lambda_{jk}, \mu_{jk} \in \mathbb{R}$ and $\gamma_{jk} \in \mathbb{R}^3$ are parameters. For identification, μ_{13} is fixed to 0.

We have considered two specifications for the distribution of the unobserved heterogeneity vector $V = (V_1, V_2)'$: a discrete distribution and a normal distribution with two independent factors.

3.2.1 A continuous distribution

V_1 and V_2 are assumed independent and identically distributed. V_1 and V_2 are distributed as standard normal distribution. In this case the unobserved term space is $\Omega = \mathbb{R}^2$ and the conditional contribution to the likelihood function is

$$\ell(\theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \ell(\theta | y_1, \dots, y_{12}, x; \nu)\, \frac{1}{2\pi} \exp(-0.5\, (\nu_1^2 + \nu_2^2))\, d\nu_1\, d\nu_2 \quad (2)$$

3.2.2 A discrete distribution

Let us assume that $\nu_j \in \{-1; 1\}$, for all $j = 1, 2$. The joint distribution of $\nu = (\nu_1, \nu_2)'$ is discrete. We assume that

$$\text{Prob}[V = (\nu_1^0, \nu_2^0)'] = \begin{cases} \pi_{00}, & \text{if } \nu_1^0 = -1 \text{ and } \nu_2^0 = -1, \\ \pi_{01}, & \text{if } \nu_1^0 = -1 \text{ and } \nu_2^0 = 1, \\ \pi_{10}, & \text{if } \nu_1^0 = 1 \text{ and } \nu_2^0 = -1, \\ \pi_{11}, & \text{if } \nu_1^0 = 1 \text{ and } \nu_2^0 = 1, \end{cases}$$

where $0 \leq \pi_{jk} \leq 1$ and $\sum_{j=0}^{1} \sum_{k=0}^{1} \pi_{jk} = 1$.

The conditional contribution to the likelihood function is

$$\ell(\theta) = \sum_{j=0,1} \sum_{k=0,1} \ell(\theta|y_1, \ldots, y_{12}, x; (2j-1, 2k-1))\, \pi_{jk}. \tag{3}$$

Here, we have three additional parameters to estimate: π_{00}, π_{01} and π_{10}.

This approach is similar to the one proposed by Heckman and Singer (1984). The number of points of the mixture is fixed to 4. This approach is often used for the estimation of transition model (see Gilbert, Kamionka and Lacroix 2001, Kamionka 2008).

In practice, in order to estimate the model, we use the following parametrisation of the distribution of the unobserved heterogeneity component

$$\pi_{jk} = \frac{\exp(c_{jk})}{\sum_{j'=0}^{1} \sum_{k'=0}^{1} \exp(c_{j'k'})},$$

where $c_{jk} \in \mathbb{R}$, $j, k \in \{0, 1\}$, are parameters ($c_{11} = 0$).

3.3 Transition probabilities

In the expression (1) of the conditional contribution to the likelihood function, we have to specify the transition probability to occupy state k, $k \in E$, given the past history of the process $y_{t-1} = j$, ϕ_{t-1} and ϕ_{12}. We write this conditional probability as $p_{jkt}(\phi_{t-1}, \phi_{12}, x, \nu; \theta) = P(Y_t{=}k \mid y_{t-1}{=}j, \phi_{t-1}, \phi_{12}, x, \nu; \theta)$.

We assume that the conditional transition probabilities is

$$p_{jkt}(\phi_{t-1}, \phi_{12}, x, \nu; \theta) = \begin{cases} \dfrac{\psi_{jkt}}{\sum\limits_{k' \in E_j} \psi_{jk't}} \left(1 - \exp\left(-\sum\limits_{k' \in E_j} \psi_{jk't}\right)\right), \text{ if } k \neq j, \\ \exp(-\sum\limits_{k' \in E_j} \psi_{jk't}), \text{ if } k = j, \end{cases} \tag{4}$$

where $\psi_{jkt} = \exp(X'_{jkt}\, a_{jk} + \phi'_{t-1} b_{jk} + v_{jk}) = \exp(X'_{jkt}\, a_{jk} + \phi'_{t-1} b_{jk} + \lambda_{jk}\, \nu_1 + \mu_{jk}\, \nu_2 + \phi'_{12}\gamma_{jk})$. X_{jk} is a vector of exogenous variables specific to the transition from state j to state k, $k \in E_j$ and $j \in E$. $a_{jk} \in \mathbb{R}^p$ and $b_{jk} \in \mathbb{R}^3$ are vector of parameters.

There is a direct relation between this specification of the transition probabilities and the econometrics of multi-spell multi-state models (see Flinn and Heckman 1983, Fougère and Kamionka 2008). Indeed, $\exp(-\sum_{k' \in E_j} \psi_{jk't} \times 1)$ represents the conditional probability to stay in state j one month again (or to 'survive' in this state). The expression $1 - \exp\left(-\sum_{k' \in E_j} \psi_{jk't} \times 1\right)$[11] represents the conditional

[11] Let $S_{jt} = \sum_{k' \in E_j} \psi_{jk't}$. Then Prob$[0 \leq U \leq 1 \mid S_{jt}] = \int_0^1 S_{jt}\ \exp(-S_{jt}\ u)\, du = 1 - \exp(-S_{jt}\ u)$. It is the conditional probability that the individual stay at most 1 units of time more in state j. We assume that, at most, one transition can occur within a given month. U represents the forward duration in state j. The conditional distribution of this forward duration is an exponential distribution.

probability to stay in state j exactly one month more. Given the individual leaves the state j, the expression $\psi_{jkt}/(\sum_{k' \in E_j} \psi_{jk't})$ is the conditional probability to enter into state k, $k \in E_j$. Finally, ψ_{jkt} can be interpreted as a conditional hazard function for the transition from state j to state k.

3.4 Estimation

The contribution of a given individual to the likelihood function is

$$\ell(\theta) = \int_{\Omega} \prod_{t=13}^{60} \prod_{j=1}^{4} \prod_{k \in E_j} p_{jkt}(\phi_{t-1}, \phi_{12}, x, \nu; \theta)^{\delta_{jkt}} \, g(\nu; \theta) \, \eta(d\nu), \tag{5}$$

where

$$\delta_{jkt} = \begin{cases} 1, \text{ if } y_{t-1} = j \text{ and } y_t = k, \\ 0, \text{ otherwise.} \end{cases},$$

and the expression of the transition probabilities $p_{jkt}(\phi_{t-1}, \phi_{12}, x, \nu; \theta)^{\delta_{jkt}}$ are given by the equation (4).

If the unobserved heterogeneity factors V_j, $j = 1, 2$, are discrete, the likelihood function can be maximized directly with respect to the parameters. If the unobserved factors are distributed as standard normal variables, then we propose to maximize the simulated likelihood (SML) obtained replacing each contribution (5) by the expression

$$\hat{\ell}(\theta) = \frac{1}{H} \sum_{h=1}^{H} \prod_{t=13}^{60} \prod_{j=1}^{4} \prod_{k \in E_j} p_{jkt}(\phi_{t-1}, \phi_{12}, x, \nu_h; \theta)^{\delta_{jkt}},$$

where the drawings ν_h, $h = 1, \ldots, H$, are i.i.d. $N(0, 1)$ and specific to the individual.

The SML estimator $\hat{\theta}_{HN}$ is asymptotically efficient (see, Gouriéroux and Monfort 1991). If $\frac{N}{H} \longrightarrow 0$, then $\sqrt{n}(\hat{\theta}_{NH} - \theta_0) \longrightarrow N(0, I(\theta_0)^{-1})$, where $I(\theta) = E[\frac{\partial \ln(\ell_i(\theta))}{\partial \theta} \frac{\partial \ln(\ell_i(\theta))}{\partial \theta'}]$ and $\ell_i(\theta)$ is the contribution of individual i to the likelihood function, $i = 1, \ldots, N$. In practice, a limited number of drawings allows to obtain a good approximation for the true value of the parameters (see, Kamionka 1998, Laroque and Salanié, 1993).

The variance-covariance matrix $\hat{I}(\theta_0)$ can be estimated using the following estimator

$$\hat{I}(\theta_0) = \frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{h=1}^{H} \frac{\partial \ell_i(\hat{\theta}_{NH} \mid \nu_{ih})}{\partial \theta}}{\sum_{h=1}^{H} \ell_i(\hat{\theta}_{NH} \mid \nu_{ih})} \frac{\sum_{h=1}^{H} \frac{\partial \ell_i(\hat{\theta}_{NH} \mid \nu_{ih})}{\partial \theta'}}{\sum_{h=1}^{H} \ell_i(\hat{\theta}_{NH} \mid \nu_{ih})},$$

where $\ell_i(\theta \mid \nu) = \ell(\theta \mid y_{i,1}, \ldots, y_{i,12}, x_i, \nu)$ is the conditional contribution of individual i to the likelihood function and ν_{ih} are i.i.d. drawings of the unobserved heterogeneity term distributed according to the density function $g(\nu; \theta)$.

4 Evaluation of the impact of training

4.1 Conditional probabilities to enter into training

The sociodemographic characteristics have a similar impact on the entry in employment and unemployment training programs: the monthly probability to enter into training program, either in unemployment or in employment, increases with the level of education. Workers aged over 46 year-old have a significantly smaller access to training than younger workers. However, gender and nationality determine entry in employment training, but not in unemployment training: women are less likely than men to enter into an employment training program when employed and foreigners have smaller probability than French worker to enter into training.

The more the worker experienced non employment periods in the previous year, the higher is her/his probability to enter into an employment training program. The probability to participate in an employment training program increases with the share of time spent on training (of any kind) in the 12 previous months. Unemployment training appears to be a more important determinant for this transition than employment training (except in the constrained non parametric specification). The more the unemployed worker has experienced employment training in the previous year, the more he is likely to participate in an unemployment training program. As we control for unobservable heterogeneity, we can interpret this result as follows: previous participation in programs may reveal the willingness of the employee (or unemployed worker) to participate in such programs and his ability to benefit from it, so that the employer (or the public service of employment) is more likely to offer training to an individual who has already been trained than to others. Concerning the access to unemployment training, we exhibit a negative duration dependence of non employment. However, there is no recurrence effects, as previous unemployment training participation has no significant effects on the probability to re-enter into this state.

4.2 Duration of training

Women participate in longer training programs than men. The educational level appears to have a insignificant impact on the unemployment training duration (in the parametric unconstrained specification only, high school educated worker have however longer training programs than the non educated people). Nationality and age have a insignificant impact on the unemployment training duration (foreigners have longer training period than French workers only in the parametric unconstrained specification).

The more the worker has spent an important share of the previous year on non employment, the more she/he participates in a longer training program. The same phenomenon is observed for the share of time spent in employment training, which reveals a positive duration dependence of employment training. Finally, the time spent in unemployment training does not significantly affect the employment training duration.

Only very educated unemployed workers enter into longer training programs. The family situation, nationality and age do not affect the length of the unemployment training (except for the individuals aged over 55 years old who have longer unemployment training periods than other).

The time spent previously in non-employment and in unemployment programs does not determine the length of the training program. On the contrary, the exit rate from unemployment training is increasing with the time already spent in employment training.

4.3 Impact of training on the risk to leave employment and on the employment duration

In three of the four estimated specifications, only highly educated workers have a longer employment duration and a smaller risk to leave employment than the worker without any diploma. Nationality and family situation do not influence the employment duration. Concerning age, 26-55 years old workers have a smaller probability to leave employment. The workers aged over 55 years old experience higher non employment risk than the others. This result can be explained by retirement as unemployment and out-of-labor-force states are aggregated.

We exhibit past state dependence: the more the worker stayed in non employment, the more he is facing with the risk of non employment. This may reveal unstable trajectories where non employment periods and short employment spells alternate. Participation in employment training during the previous year increases the probability of exiting employment, and all the more than the worker spent an important share of time in such training. Previous unemployment training participation do not significantly reduce the probability of transition to non employment. The results thus tend to reveal a negative impact of training on job tenure.

4.4 Impact of training on the duration of non employment and the return to employment

As expected, the probability to return into employment increases with the level of education and decreases with the age of the unemployed worker. Family situation and nationality do not affect the length of the non employment spell.

The estimates show the usual negative state and duration dependence of non employment: the share of time spent out of employment in the previous year decreases the monthly probability to return into employment. More interestingly, the more the unemployed worker spent time in employment training program during the previous year, the higher is her/his probability to return into employment. This reveals the necessity of allowing for long term effects of training on the labor market history. Last, the more the individual has spent time in unemployment training during the last 12 months, the higher is her/his probability to find a job quickly. This means that unemployment training reduces the non employment duration once the training is completed.

5 Conclusion

In this paper, we consider the impact of past participation to training programs on the individual labor market mobility. Using a French date set, we consider jointly the effects of training programs dedicated to employed workers and training programs proposed to unemployed individuals. Participation to these training programs is endogenous in order to take into account selectivity phenomena allowing, for instance, those who are *ex ante* the most willing to participate in employment training programs may also be the ones who have *a priori* low exit rates from employment. We model the transitions between the states of the labor market using a multi-state multi-spell transition model. We take into account the existence of unobserved heterogeneity using a factor loading specification. The model we use allows to distinguish true from spurious state dependence. The impact of the participation to training programs is considered *via* the proportion of the past year the individual have devoted to these programs. We find that the conditional probability to return into employment is increasing with the proportion of the previous year the individual has spent in training programs whatever the category of these programs. More surprisingly, past participation to employment training is associated with a greater hazard function for the transition from employment to non employment indicating that firms use employment training programs in order to increase general human capital of workers when they anticipate a lower activity. It is interesting to note that the conditional probability to enter into employment training programs (respectively, unemployment training programs) is increasing with the proportion of the last year the individual participated in unemployment training programs (respectively, employment training programs). The conditional probability to reenter into employment training programs is increasing with the proportion of the last year the individual spent in training programs. As we control for observed and unobserved characteristics of the worker, this result indicates that previous participation in these programs may reveal the willingness of the worker to participate in such programs and his ability to benefit from it. Consequently, the employer or the public service of employment is more likely to offer training to workers who have already been trained.

A further research could consist to distinguish the impact of the training programs according to the characteristics of the workers and to study the existence of a state dependence of a higher order.

References

[1] Bassanini, A., Booth, A., Brunello, G., De Paola, M. and Leuven, E., 2005, "Workplace Training in Europe", IZA Discussion Paper Nř 1640.

[2] Blundell, R., Dearden L., Meghir C., and Sianesi B., 1999, "Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy", *Fiscal Studies*, 20(1), 1–23.

[3] Bonnal, L., Fougère D. and A. Sérandon, 1997, "Evaluating the Impact of French Employment Policies on Individual Labour Market Histories", *Review of Economic Studies*, 64(4), 683–713.

[4] Crépon, B. M. Ferracci and D. Fougère, 2007, "Training the Unemployed in France: How Does it Affect Duration and Recurrence", *IZA Working Paper*, DP 3215.

[5] Edon C. and T. Kamionka, 2008, "Modélisation dynamique de la participation au march du travail des femmes en couple", to be published in *Annales d'économie et de statistique*.

[6] Flinn, C.J. and J.J. Heckman, 1983, "The Likelihood function of the multistate-multiepisode Model", in: Advances in Econometrics, 1, R. Basmann and G. Rhodes Eds., Greenwich, CN: JAI Press, 225–231.

[7] Fougère, D., Goux, D., and Maurin, E., 2001, " Formation continue et carrières salariales. Une évaluation sur données individuelles", *Annales d'Économie et de Statistique*, 62 , 49–69.

[8] Fougère D. and T. Kamionka, 2008, "Econometrics of Individual Labor Market Transitions", to be published in *Econometrics of Panel Data*, third edition, L. Matyas and P. Sevestre Eds., Springer.

[9] Gerfin, M., 2003, "Work-Related Training and Wages - An Empirical Analysis for Male Workers in Switzerland", mimeo, Volkswirtschaftliches Institut, Universität Bern.

[10] Gilbert, L., Kamionka T. and G. Lacroix, 2001, "The Impact of Government-Sponsored Training Programs on the Labor Market Transitions of Disadvantaged Men", *CREST Working Paper*, 2001-15, Pairs.

[11] Goux, D. and Maurin, E., 1997, "Train or Pay: Does It Reduce Inequalities to Encourage Firms to Train their Workers?", INSEE, Document de travail G9703, Mars.

[12] Goux, D. and Maurin, E., 2000, "Returns to Firm Provided Training: Evidence from French Worker-Firm Matched Data", *Labour Economics*, 7, 1–19.

[13] Gouriéroux Ch. and Monfort A., 1991, "Simulated Based in Econometric Models with Heterogeneity", *Annales d'Économie et de Statistique*, 20-21, 69–107.

[14] Heckman J. J., Borjas G. J., 1980, "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence", *Economica*, New Series, 47, 187, 247–283.

[15] Heckman J. J., 1981, "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process", in *Structural Analysis of Discrete Data*, eds. C. Manski et D. McFadden. Cambridge, MA, MIT Press.

[16] Heckman, J. and B. Singer, 1984, "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data", *Econometrica*, 52, 271-320.

[17] Heckman, J., Lalonde, R. and Smith, J. 1999, "The Economics and Econometrics of Active Labor markets Programs", in Handbook of Labor Economics, A. Ashenfelter and D. Card (eds.), Elsevier Science, vol.3A, chap 31, pp. 1865-2097.

[18] Kamionka T., 1998, "Simulated maximum likelihood estimation in transition models", *Econometrics Journal*, 1, C129–153.

[19] Kamionka, T. and G. Lacroix, 2008, "Assessing the External Validity of an Experimental Wage Subsidy", to be published in *Annales d'économie et de Statistique*, 86.

[20] Lancaster T., 1990, The econometric analysis of transition data. Cambridge University Press, New-York, USA, 336p.

[21] Laroque G. and B. Salanié, 1993, "Simulation-based Estimation of Models with Lagged Latent Variables", *Journal of Applied Econometrics*, 8, S119–S133.

[22] Mealli, F. and S. Pudney, 1996, "Occupational pensions and Job Mobility in Britain: Estimation of a Random-effects Competing Risks Model", *Journal of Applied Econometrics*, 11, 293–320.

[23] OECD, 2003, "Upgrading Workers' Skills and Competencies", chapter 5, Employment Outlook, Paris.

[24] Parent, D., 1999, "Wages and Mobility: The Impact of Employer-Provided Training", *Journal of Labor Economics*, 17(2), 298–317.

[25] Parent, D., 2003, "Employer-Supported Training in Canada and its Impact on Mobility and Wages", *Empirical Economics*, 28, 431–459.

[26] Pischke, J., 2001, "Continuous Training in Germany", *Journal of Population Economics*, 14, 523–548.

[27] Wooldridge J.M. (2005). – "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity", *Journal of Applied Econometrics*, 20-1, 39–54.

A Descriptive analysis

Table 4: **Conditional Constrained transition matrix (%)**

$t \rightarrow$ $(t-1) \downarrow$		Empl.	Non Empl.	Empl. Training	Unempl. Training	Total (row)
Empl.	Male	98,67	0,7	0,64	-	100
	Female	98,44	1,07	0,5	-	100
	No diploma	98,62	1,09	0,29	-	100
	High School	98,69	0,84	0,47	-	100
	Undergraduate	98,32	0,93	0,75	-	100
	Graduate	98,36	0,66	0,98	-	100
	< 26	97,02	2,21	0,78	-	100
	$26 - 45$	98,45	0,88	0,67	-	100
	> 45	98,83	0,77	0,4	-	100
Non Empl.	Male	2,2	97,45	-	0,35	100
	Female	1,47	98,26	-	0,27	100
	No diploma	1,28	98,54	-	0,19	100
	High School	1,51	98,23	-	0,26	100
	Undergraduate	2,65	96,8	-	0,54	100
	Graduate	3,12	96,29	-	0,59	100
	< 26	4,57	94,68	-	0,75	100
	$26 - 45$	3,38	96,07	-	0,55	100
	> 45	0,55	99,33	-	0,12	100
Empl. Training	Male	22,99	-	77,01	-	100
	Female	16,17	-	83,83	-	100
	No diploma	23,27	-	76,73	-	100
	High School	21,94	-	78,06	-	100
	Undergraduate	18,75	-	81,25	-	100
	Graduate	17,2	-	82,8	-	100
	< 26	15,64	-	84,36	-	100
	$26 - 45$	19,25	-	80,75	-	100
	> 45	21,69	-	78,31	-	100
Unempl. Training	Male	-	16,68	-	83,32	100
	Female	-	14,4	-	85,6	100
	No diploma	-	15	-	85	100
	High School	-	16,7	-	83,3	100
	Undergraduate	-	14,42	-	85,58	100
	Graduate	-	13,04	-	86,96	100
	< 26	-	12,69	-	87,31	100
	$26 - 45$	-	15,79	-	84,21	100
	> 45	-	14,12	-	85,88	100

Table 5: **Biprobit estimates**

	Output equation			
	(1)	(2)	(3)	(4)
Intercept	-1,5790	0,1018	0,6586	0,3078
	(0,3397)	(0,0327)	(0,0851)	(0,0617)
Local unemployment	0,0454	-	-	-
rate in 98	(0,0236)	-	-	-
Participation	1,3462	0,2485	1,4317	1,0986
in training	(1,9854)	(0,2749)	(0,4270)	(0,3691)
Female	0,0588	0,0363	-0,1879	-0,2328
	(0,1567)	(0,0219)	(0,0486)	(0,0379)
Diploma				
High school	-0,0704	-0,0266	0,0168	0,0762
	(0,1742)	(0,0286)	(0,0495)	(0,0402)
Undergraduate	-0,3532	0,1007	0,0103	0,0601
	(0,2437)	(0,0416)	(0,0948)	(0,0744)
Graduate	-0,3816	-0,1343	0,1476	0,2181
	(0,2174)	(0,04561)	(0,1277)	(0,0981)
Age				
35 - 45	0,2010	0,4597	0,4185	0,4764
	(0,2131)	(0,0265)	(0,0451)	(0,0362)
45+	0,6906	0,6356	0,3215	0,2776
	(0,2213)	(0,0263)	(0,0303)	(0,0276)

Standard errors are given in parentheses.

Models estimated:

(1) Pr(duration of the 1998 unemployment spell $>$24 months).
(2) Pr(duration of the 1998 employment spell $>$24 months).
(3) Pr(employment duration starting from 1998 $>$24 months).
(4) Pr(employment duration starting from 1998 $>$48 months).

Biprobit estimates (continued)

	Selection equation			
	(1)	(2)	(3)	(4)
Intercept	-2,0955	-2,0697	-1,7079	-1,70254
	(0,3911)	(0,1216)	(0,1211)	(0,1228)
Local share of trained people	6,3373	4,3818	2,7047	2,6792
	(3,6685)	(0,8933)	(0,6150)	(0,6220)
Female	0,1335	-0,1822	-0,2690	-0,2692
	(0,1799)	(0,0247)	(0,0229)	(0,0229)
Diploma				
High school	0,0091	0,3125	0,3314	0,3313
	(0,2232)	(0,0361)	(0,0332)	(0,0332)
Undergraduate	0,3227	0,5406	0,6496	0,6496
	(0,2653)	(0,0437)	(0,0399)	(0,0399)
Graduate	0,0135	0,6588	0,8134	0,8135
	(0,2657)	(0,0409)	(0,0375)	(0,0375)
Age				
35 - 45	0,4012	0,2371	0,1173	0,1167
	(0,1904)	(0,0327)	(0,0292)	(0,0292)
45+	0,1431	0,1014	-0,1243	-0,1249
	(0,2615)	(0,0331)	(0,0300)	(0,0302)
Firm size				
10 - 49	-	-0,3728	-	-
	-	(0,0864)	-	-
49 - 99	-	-0,1933	-	-
	-	(0,0679)	-	-
100 - 500	-	-0,2001	-	-
	-	(0,0437)	-	-
>500	-	0,0548	-	-
	-	(0,0415)	-	-
ρ	-0,2510	0,3218	-0,2521	-0,1815
	(1,0742)	(0,1285)	(0,2833)	(0,2265)

Standard errors are given in parentheses.

B Results

Table 6: **Parameters estimates**

	Parametric		Non Parametric	
	Unconstrained	Constrained	Unconstrained	Constrained
	From Employment to Employment Training			
Intercept	-5.8509	-5.8234	-5.3732	-5.3590
	(0.1757)	(0.1380)	(0.1427)	(0.1386)
Female	-0.2200	-0.3995	-0.3408	-0.3367
	(0.0928)	(0.0699)	(0.0714)	(0.0701)
Diploma				
High School	0.3658	0.4941	0.5297	0.5296
	(0.1476)	(0.1148)	(0.1156)	(0.1149)
Undergraduate	0.7277	0.9291	0.8305	0.8864
	(0.1674)	(0.1302)	(0.1313)	(0.1300)
Graduate	1.1602	1.1962	1.1573	1.1694
	(0.1498)	(0.1186)	(0.1204)	(0.1187)
Not French	-0.8086	-0.7763	-0.8311	-0.8485
	(0.2778)	(0.2178)	(0.2191)	(0.2159)
Age				
26-35	*-0.0181*	*0.0035*	*-0.0006*	*0.0285*
	(0.1232)	*(0.0997)*	*(0.0998)*	*(0.1032)*
36-45	*-0.0966*	*-0.0492*	*-0.0993*	*-0.0610*
	(0.1363)	*(0.1064)*	*(0.1073)*	*(0.1086)*
46-55	-0.6351	-0.4344	-0.4338	-0.4297
	(0.1671)	(0.1231)	(0.1243)	(0.1247)
55+	*-0.1105*	*-0.1490*	*-0.1640*	*-0.0646*
	(0.3388)	*(0.2789)*	*(0.2776)*	*(0.2746)*
Lag				
% Non-employment	0.1057	0.0654	0.0892	0.0261
	(0.0276)	(0.0195)	(0.0212)	(0.0120)
% Employment training	0.1561	0.2082	0.1622	0.1219
	(0.0267)	(0.0185)	(0.0221)	(0.0132)
% Unemployment training	0.3654	0.3094	0.2903	*0.0456*
	(0.0428)	(0.0315)	(0.0314)	*(0.0376)*

Standard errors are given in parentheses.

Parameters estimates (continued)

	Parametric		Non Parametric	
	Unconstrained	Constrained	Unconstrained	Constrained
	From Employment Training to Employment			
Intercept	-0.9824 (0.2962)	-0.9164 (0.2182)	-0.6298 (0.1999)	-0.4639 (0.2428)
Female	-0.5415 (0.1426)	-0.6678 (0.1086)	-0.5151 (0.1023)	-0.5480 (0.1165)
Diploma				
High School	0.5585 (0.2381)	*0.1371* *(0.1687)*	*0.1303* *(0.1499)*	*0.0943* *(0.1852)*
Undergraduate	*0.0175* *(0.2526)*	*0.0839* *(0.1909)*	*-0.0554* *(0.1740)*	*0.0557* *(0.1945)*
Graduate	*0.1734* *(0.2229)*	*0.0248* *(0.1773)*	*-0.1239* *(0.1689)*	*-0.1665* *(0.1912)*
Not French	-0.9332 (0.4594)	*-0.3677* *(0.3448)*	*-0.5066* *(0.3300)*	*-0.5462* *(0.3154)*
Age				
26-35	*-0.1767* *(0.1981)*	*0.0871* *(0.1412)*	*0.2174* *(0.1393)*	0.3635 (0.1599)
36-45	*-0.1765* *(0.2091)*	*0.0615* *(0.1625)*	*0.0386* *(0.1520)*	*0.0915* *(0.1699)*
46-55	*-0.2036* *(0.2638)*	*0.2816* *(0.1863)*	*0.2003* *(0.1766)*	*0.3040* *(0.1964)*
55+	*0.1711* *(0.4890)*	*-0.0625* *(0.4602)*	*-0.0194* *(0.3949)*	*-0.3532* *(0.4786)*
State Dependence				
% Non-employment	-0.1357 (0.0482)	-0.0946 (0.0343)	-0.0794 (0.0353)	-0.1057 (0.0215)
% Employment training	-0.0909 (0.0263)	-0.0907 (0.0170)	-0.0881 (0.0197)	-0.0824 (0.0154)
% Unemployment training	*-0.0848* *(0.0766)*	*0.0258* *(0.0523)*	*-0.0593* *(0.0510)*	*0.0604* *(0.0417)*

Standard errors are given in parentheses.

Parameters estimates (continued)

	Parametric		Non Parametric	
	Unconstrained	Constrained	Unconstrained	Constrained
	From Employment to Non Employment			
Intercept	-4.7856	4.6276	-4.3970	-4.7484
	(0.1326)	(0.1017)	(0.1296)	(0.1318)
Diploma				
High School	*-0.1271*	-0.1990	-0.1617	*-0.1815*
	(0.1034)	(0.0768)	(0.0851)	*(0.0926)*
Undergraduate	*-0.1251*	*-0.0983*	*-0.1406*	*-0.1556*
	(-0.1251)	*(0.0992)*	*(0.1097)*	*(0.1176)*
Graduate	-0.3693	0.4027	-0.3809	-0.3736
	(0.1306)	(0.0987)	(0.1079)	(0.1142)
Not French	*0.1772*	*0.0881*	*0.0078*	*0.0264*
	(0.1483)	*(0.1192)*	*(0.1231)*	*(0.1445)*
Age				
26-35	-0.4347	-0.4557	-0.4746	-0.4719
	(0.1108)	(0.0846)	(0.0943)	(0.0974)
36-45	-0.7418	-0.7161	-0.6661	-0.6137
	(0.1269)	(0.0944)	(0.1037)	(0.1058)
46-55	-0.4676	-0.4799	-0.3665	-0.3092
	(0.1296)	(0.0956)	(0.1068)	(0.1096)
55+	1.0637	1.1517	1.2942	1.4580
	(0.2017)	(0.1495)	(0.1510)	(0.1724)
Family situation				
Female with young children	*0.0482*	*0.1982*	*0.1777*	*0.2000*
	(0.1418)	*(0.1050)*	*(0.1146)*	*(0.1161)*
Male with young children	*0.1653*	0.2839	*0.1933*	0.2700
	(0.1490)	(0.1113)	*(0.1184)*	(0.1221)
Female without young children	*-0.0870*	*-0.0603*	*-0.0479*	*-0.0418*
	(0.0874)	*(0.0664)*	*(0.0709)*	*(0.0736)*
State Dependence				
% Non-employment	0.0886	0.1548	0.0578	0.1564
	(0.0175)	(0.0123)	(0.0148)	(0.0114)
% Employment training	0.1435	0.0967	0.1199	*0.0318*
	(0.0349)	(0.0233)	(0.0254)	*(0.0189)*
% Unemployment training	*-0.0188*	*-0.0024*	*-0.0722*	*0.0289*
	(0.0866)	*(0.0667)*	*(0.0678)*	*(0.0347)*

Standard errors are given in parentheses.

Parameters estimates (continued)

	Parametric		Non Parametric	
	Unconstrained	Constrained	Unconstrained	Constrained
	From Non Employment to Employment			
Intercept	-2.4365 (0.1756)	-2.5986 (0.1459)	-2.0002 (0.1457)	-2.0448 (0.1596)
Diploma				
High School	0.3274 (0.1126)	0.2817 (0.0878)	0.2740 (0.0865)	0.3375 (0.0986)
Undergraduate	0.4234 (0.1386)	0.5222 (0.1139)	0.3948 (0.1090)	0.5492 (0.1208)
Graduate	0.4741 (0.1401)	0.4510 (0.1155)	0.4090 (0.1120)	0.5396 (0.1228)
Not French	*-0.0303* *(0.1582)*	*-0.1494* *(0.1314)*	*-0.1418* *(0.1261)*	*-0.2004* *(0.1444)*
Age				
26-35	-0.3653 (0.1122)	-0.3740 (0.0915)	-0.3493 (0.0886)	-0.3504 (0.1105)
36-45	-0.8892 (0.1369)	-0.8479 (0.1074)	-0.7544 (0.1017)	-0.7776 (0.1174)
46-55	-1.9971 (0.1621)	-1.8985 (0.1224)	-1.8546 (0.1201)	-2.0827 (0.1355)
55+	-4.8845 (0.5890)	-4.9180 (0.4556)	-4.8340 (0.4584)	-5.2710 (0.4642)
Family situation				
Female with young children	*0.0897* *(0.1600)*	*0.0334* *(0.1380)*	*0.0229* *(0.1210)*	*-0.0429* *(0.1321)*
Male with young children	*0.3132* *(0.1631)*	*0.2059* *(0.1357)*	*0.1604* *(0.1273)*	*0.2671* *(0.1380)*
Female without young children	*0.0864* *(0.0957)*	*0.0015* *(0.1123)*	*0.0647* *(0.0738)*	*0.0999* *(0.0824)*
State Dependence				
% Non-employment	-0.0849 (0.0144)	-0.1138 (0.0105)	-0.0897 (0.0116)	-0.0910 (0.0100)
% Employment training	0.1979 (0.0614)	*0.0435* *(0.0446)*	0.1268 (0.0458)	*0.0291* *(0.0244)*
% Unemployment training	0.1540 (0.0265)	0.1255 (0.0197)	0.1082 (0.0204)	*-0.0108* *(0.0258)*

Standard errors are given in parentheses.

Parameters estimates (continued)

	Parametric		Non Parametric	
	Unconstrained	Constrained	Unconstrained	Constrained
	From Non Employment to Unemployment Training			
Intercept	-4.1629 (0.3789)	-4.5215 (0.2975)	-4.0971 (0.3112)	-4.1714 (0.3634)
Diploma				
High School	*0.4112* *(0.2583)*	0.4573 (0.1879)	0.4586 (0.1988)	0.5678 (0.2013)
Undergraduate	0.9585 (0.2905)	1.0857 (0.2176)	1.0308 (0.2310)	1.0920 (0.2328)
Graduate	0.9754 (0.2970)	1.0118 (0.2273)	1.1017 (0.2376)	1.0678 (0.2365)
Not French	*-0.5269* *(0.4098)*	*-0.2526* *(0.2861)*	*-0.2362* *(0.2998)*	*-0.3890* *(0.3013)*
Age				
26-35	*-0.1369* *(0.2520)*	*0.0007* *(0.2317)*	*0.0173* *(0.1704)*	*-0.0007* *(0.3754)*
36-45	*-0.1168* *(0.2765)*	*0.2094* *(0.2188)*	*0.2179* *(0.2070)*	*0.0486* *(0.2950)*
46-55	-1.4984 (0.3454)	-1.3589 (0.2772)	-1.4982 (0.2657)	-1.8147 (0.3389)
55+	-2.7282 (0.6288)	-2.9284 (0.5434)	-3.0462 (0.5430)	-3.6966 (0.5835)
Family situation				
Female with young children	*0.1325* *(0.3110)*	*0.1929* *(0.2364)*	*0.1986* *(0.2568)*	*-0.0106* *(0.1601)*
Male with young children	*0.5302* *(0.3074)*	*0.4416* *(0.2345)*	0.6004 (0.2430)	0.5704 (0.2374)
Female without young children	*-0.3199* *(0.2065)*	*-0.2600* *(0.1569)*	*-0.1683* *(0.1658)*	*-0.1302* *(0.1580)*
State Dependence				
% Non-employment	-0.1237 (0.0290)	-0.1674 (0.0192)	-0.0988 (0.0229)	-0.1121 (0.0182)
% Employment training	0.5569 (0.0688)	0.3429 (0.0411)	0.5332 (0.0536)	*0.0579* *(0.0468)*
% Unemployment training	*0.0966* *(0.0520)*	*0.0615* *(0.0368)*	*0.0384* *(0.0430)*	*0.0062* *(0.0477)*

Standard errors are given in parentheses.

Parameters estimates (continued)

	Parametric		Non Parametric	
	Unconstrained	Constrained	Unconstrained	Constrained
	From Unemployment Training to Non Employment			
Intercept	-1.5791	-1.8755	-1.2780	-1.7336
	(0.5303)	(0.4352)	(0.3542)	(0.3295)
Diploma				
High School	*0.0760*	*0.0007*	*-0.0459*	*-0.0668*
	(0.2871)	*(0.2314)*	*(0.1891)*	*(0.1955)*
Undergraduate	*-0.2059*	*-0.1821*	*-0.2347*	*-0.2421*
	(0.3311)	*(0.2561)*	*(0.2273)*	*(0.2351)*
Graduate	-0.8391	-0.6522	-0.7106	-0.7955
	(0.3345)	(0.2611)	(0.2219)	(0.2231)
Not French	*-0.3123*	*-0.3428*	*-0.1968*	*-0.4424*
	(0.4436)	*(0.3180)*	*(0.2886)*	*(0.3005)*
Age				
26-35	*0.2433*	*0.1350*	*0.2745*	0.4052
	(0.2669)	*(0.2090)*	*(0.1993)*	(0.2009)
36-45	*0.0921*	*0.1606*	*0.1443*	*0.1191*
	(0.2790)	*(0.2093)*	*(0.1941)*	*(0.1976)*
46-55	*0.5148*	*-0.0683*	*0.0421*	*0.0810*
	(0.3626)	*(0.2776)*	*(0.2603)*	*(0.2653)*
55+	*-2.0348*	-2.5463	-2.3971	-2.7680
	(1.0756)	(1.0325)	(1.0307)	(1.0418)
Family situation				
Female with young children	*-0.0141*	*0.0888*	*0.0463*	*0.0504*
	(0.3727)	*(0.2630)*	*(0.2470)*	*(0.2750)*
Male with young children	*-0.0580*	*-0.0796*	*-0.2321*	*-0.4069*
	(0.3485)	*(0.2533)*	*(0.2347)*	*(0.2409)*
Female without young children	*-0.0113*	*0.0051*	*-0.0781*	*-0.1194*
	(0.2328)	*(0.1739)*	*(0.1596)*	*(0.1692)*
State dependence				
% Non-employment	*-0.0190*	*-0.0283*	*-0.0353*	*-0.0248*
	(0.0400)	*(0.0271)*	*(0.0294)*	*(0.0221)*
% Employment training	0.2519	0.1190	0.1516	-0.1289
	(0.0803)	(0.0531)	(0.0514)	(0.0490)
% Unemployment training	*0.0017*	*-0.0104*	*-0.0218*	-0.0518
	(0.0459)	*(0.0341)*	*(0.0290)*	(0.0235)

Standard errors are given in parentheses.

Table 7: **Parameters estimates for the unobserved heterogeneity**

	Parametric		Non Parametric	
	Unconstrained	Constrained	Unconstrained	Constrained
		Common initial conditions		
% Non-employment	-	*0.0016*	-	-0.0318
	-	*(0.0053)*	-	(0.0081)
% Employment training	-	0.0219	-	0.0363
	-	(0.0099)	-	(0.0141)
% Unemployment training	-	*0.0058*	-	0.0588
	-	*(0.0168)*	-	(0.0141)
		Probabilities		
c_{00}	-	-	*-0.5275*	0.9248
	-	-	*(0.2772)*	0.0821
c_{01}	-	-	2.0900	*-6.4093*
	-	-	(0.1821)	*(22.1350)*
c_{10}	-	-	*0.2037*	0.1141
	-	-	*(0.2138)*	(0.0131)

Standard errors are given in parentheses.

Parameters estimates for the unobserved heterogeneity (continued)

	Parametric		Non Parametric	
	Unconstrained	Constrained	Unconstrained	Constrained
	From Employment to Employment Training			
% Non-employment	*-0.0126*	-	*-0.0052*	-
	(0.0182)	-	*(0.0137)*	-
% Employment training	0.1082	-	0.0831	-
	(0.0215)	-	(0.0161)	-
% Unemployment training	*-0.0453*	-	*-0.0103*	-
	(0.0575)	-	*(-0.236)*	-
μ	-0.8153	-0.8001	0.8365	0.8248
	(0.0793)	(0.0557)	(0.0493)	(0.0468)
	From Employment Training to Employment			
% Non-employment	*-0.0038*	-	*-0.0012*	-
	(0.0300)	-	*(0.0179)*	-
% Employment training	*0.0143*	-	*-0.0210*	-
	(0.0232)	-	*(0.0170)*	-
% Unemployment training	0.1130	-	*0.0550*	-
	(0.0681)	-	*(0.0477)*	-
λ	-1.0802	-0.9210	1.5020	0.6665
	(0.1275)	(0.0822)	(0.1578)	(0.0784)
μ	-0.5784	0.7572	0.9537	1.2630
	(0.1470)	(0.1047)	(0.1573)	(0.0808)
	From Employment to Non Employment			
% Non-employment	0.0881	-	0.0964	-
	(0.0125)	-	(0.0102)	-
% Employment training	*0.0151*	-	*0.0295*	-
	(0.0266)	-	*(0.0193)*	-
% Unemployment training	*0.0409*	-	*0.0652*	-
	(0.0429)	-	*(0.0339)*	-
λ	*-0.0099*	-0.3476	-0.5901	0.9046
	(0.1675)	(0.0716)	(0.0953)	(0.0551)
μ	-0.8257	-0.6777	-1.1142	-0.8938
	(0.0767)	(0.0699)	(0.0730)	(0.0722)

Standard errors are given in parentheses.

Parameters estimates for the unobserved heterogeneity (continued)

	Parametric		Non Parametric	
	Unconstrained	Constrained	Unconstrained	Constrained
	From Non Employment to Employment			
% Non-employment	-0.0529	-	-0.0418	-
	(0.0108)	-	(0.0083)	-
% Employment training	*-0.0201*	-	*0.0028*	-
	(0.0357)	-	*(0.0257)*	-
% Unemployment training	*-0.0566*	-	*-0.0189*	-
	(0.0311)	-	*(0.0236)*	-
λ	-0.3192	-0.4571	*0.0539*	0.7495
	(0.1090)	(0.0715)	*(0.1080)*	(0.0595)
μ	-0.5139	-0.6279	-0.6397	0.5884
	(0.0745)	(0.0674)	(0.0666)	(0.0849)
	From Non Employment to Unemployment Training			
% Non-employment	-0.0707	-	-0.0571	-
	(0.0226)	-	(0.0184)	-
% Employment training	*-0.1074*	-	-0.1661	-
	(0.0611)	-	(0.0544)	-
% Unemployment training	*0.0054*	-	*0.0294*	-
	(0.0532)	-	*(0.0519)*	-
λ	-0.6983	-0.9028	0.6960	1.1071
	(0.1677)	(0.1365)	(0.1419)	(0.1285)
μ	-0.3745	-0.1768	-0.6804	0.8007
	(0.1608)	(0.1366)	(0.1871)	(0.1163)
	From Unemployment Training to Non Employment			
% Non-employment	*-0.0205*	-	*-0.0175*	-
	(0.0241)	-	*(0.0167)*	-
% Employment training	-0.1356	-	-0.1377	-
	(0.0666)	-	(0.0480)	-
% Unemployment training	*0.0032*	-	*-0.0261*	-
	(0.0348)	-	*(0.0251)*	-
λ	*0.0420*	-0.3587	*-0.1463*	0.3740
	(0.1983)	(0.1388)	*(0.1743)*	(0.1291)
μ	*-0.3489*	-0.4019	*-0.2865*	*-0.0473*
	(0.2218)	(0.1573)	*(0.1556)*	*(0.1759)*

Standard errors are given in parentheses.

C Identification

We consider a Markov Chain with K different possible states. The transition matrix is $K \times K$ matrix with elements $\pi_{i,j}$ which are the probability to transit at date t from state i to state j. The $\pi_{i,j}$ satisfy the constraints that $\forall i \in \{1, \ldots, K\}$ $\sum\limits_{j} \pi_{i,j} = 1$. We denote $\eta_i = \pi_{i,i} = 1 - \sum\limits_{j \neq i} \pi_{i,j}$. We consider the $K-1$ vector $\eta = (\eta_i)_{i<K}$. We assume the $\pi_{i,j}$ are random : $\pi_{i,j} = \pi_{i,j}(\omega), \omega \in \Omega$. The main assumptions are the following:

1. $\forall\, i, j$ $\exists$ a finite path $l_1, l_2, \ldots, l_{i,j}$ such that
   (a) $\pi_{i,l_1}(\omega)\, \pi_{l_1,l_2}(\omega) \cdots \pi_{l_{i,j},j}(\omega) = \Gamma_{i,j}(\omega) \neq 0$ as
   (b) $\exists\, \Pi_{i,j}$ continuous function such that $\pi_{ij} = \Pi_{ij}(\gamma)$
2. We decompose $\Gamma = (\Gamma_+, \Gamma_-)$ with $\Gamma_+ = (\Gamma_{i,i+1})_{i<K}$. Given $\pi_{ij} = \Pi_{ij}(\Gamma)$ and the definition of η there are $K-1$ functions $N(\Gamma_+, \Gamma_-,)$ such that $\eta = N(\Gamma_+, \Gamma_-,)$.

Then the joint distribution of the transition probabilities π is identified non parametrically when the number of observed periods tend to infinity.

Consider individuals initially in a given state. Consider to fix ideas this state is $i = 1$. For these individuals one can observe on the data the probability of specific transitions. We consider transitions $i \to j \to i$ for $j > i$. These transitions have probability $\Theta_{i,j} = \Gamma_{i,j}\Gamma_{j,i}$. We also consider transitions $i \to i+1 \to j \to i$ for $j > i+1$. These transitions have probability $M_{i,j} = \Gamma_{i,i+1}\Gamma_{i+1,j}\Gamma_{j,i}$. If we consider $\gamma_{i,j} = \ln \Gamma_{i,j}$, $\mu_{i,j} = \ln M_{i,j}$ and $\theta_{i,j} = \ln \Theta_{i,j}$ we have the following set of equations

$$\gamma_{i,j} + \gamma_{j,i} = \theta_{i,j} \;\; \forall\, i > j$$

This makes $K-i$ equations for $i < K$, and thus a total of $K(K-1)/2$ equations. We also have

$$\gamma_{i,i+1} + \gamma_{i+1,j} + \gamma_{j,i} = \mu_{i,j} \;\; \forall\, i > j+1$$

This makes $K-i-1$ equations for $i < K-1$, and thus a total of $(K-1)(K-2)/2$ equations. We consider first $\gamma_{i,i+1}$ as given. This is an additional set of $K-1$ trivial equations $\gamma_{i,i+1} = \gamma_{i,i+1}$. Therefore we have a total of $K-1+(K-1)(K-2)/2+K(K-1)/2 = K(K-1)$ equations. We can show easily that this set of equations is invertible and we can express $\gamma_{i,j}$ as a function of $\mu_{i,j}$, $\theta_{i,j}$ and $\gamma_{i,i+1}$:

$$\gamma_- = \Gamma_-(\mu, \theta, \gamma_+)$$

Considering now the probability to stay in the different states, we have the set of $K-1$ equations:

$$\eta = N(\Gamma_-(\mu, \theta, \gamma_+), \gamma_+)$$

We assume that this set of equations can be solved in γ_+:

$$\gamma_+ = \Gamma_+ (\mu, \theta, \eta)$$

Thus we can express

$$\gamma = \Gamma (\mu, \theta, \eta)$$

and therefore

$$\pi = \Pi (\mu, \theta, \eta)$$

This shows that identifying the joint distribution of (μ, θ, η) is enough to identify the distribution of π.

We can identify from the data the probabilities of paths starting in state 1 and then composed of an arbitrary number of transitions $\Theta_{i,j}$ for any $j > i$ and then followed by an arbitrary number of transitions $M_{i,j}$ for any $j > i + 1$ and then followed by a stays in state i of an arbitrary length (when a stay is possible) and then a move to the state $i + 1$ and again the same kind of transitions. For a given state of nature, the probability of these paths can be written as

$$\prod_i \prod_{j>i} \Theta_{i,j}^{q_{i,j}} \prod_{j>i+1} M_{i,j}^{m_{i,j}} \eta_i^{d_i} \Gamma_{i,i+1} = \prod_i \exp \left(\sum_{j>i} \theta_{i,j} q_{i,j} + \sum_{j>i+1} \mu_{i,j} m_{i,j} \right) \eta_i^{d_i} \Gamma_{i,i+1} (\mu, \theta, \eta)$$

When we compute the corresponding probabilities for all possible state of nature we have :

$$P = \int \prod_i \exp \left(\sum_{j>i} \theta_{i,j} q_{i,j} + \sum_{j>i+1} \mu_{i,j} m_{i,j} \right) \eta_i^{d_i} \Gamma_{i,i+1} (\mu, \theta, \eta) f (\mu, \theta, \eta)$$

When the number of observed periods tend to infinity, it is possible to identify these probabilities for any number of transition. This is enough to identify the joint distribution $f (\mu, \theta, \eta)$.

We considered this for individuals initially in state 1. Thus the distribution identified is the distribution of π conditional to initial state is state 1. But we can do the same for individual initially in state 2. Nothing is changed, we just have to make am permutation of indexes. This identifies the distribution of π conditional to initial state is state 2. We therefore identify the distribution π conditional on the initial state. As the distribution of initial state is observed, we can identify the global distribution of of the matrix π.