INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2001-15

The Impact of Government-Sponsored Training Programs on the Labor Market Transitions of Disadvantaged Men*

L. GILBERT[1]
T. KAMIONKA[2]
G. LACROIX[3]

* This research received financial support from le Fonds FCAR, Human Resources Development Canada and CIRANO.

[1] Human Resources Development Canada. Email : lucie.gilbert@spg.org

[2] CNRS, CREST and IDEI. France. Email : kamionka@ensae.fr

[3] Département d'économie, Université Laval, CREFA and CIRANO.Canada. Email : guy.lacroix@ecn.ulaval.ca

The Impact of Government-Sponsored Training Programs on the Labor Market Transitions of Disadvantaged Men[†]

Lucie Gilbert[‡] Thierry Kamionka[§] Guy Lacroix[¶]

March 2001

[†] This research received financial support from le fonds FCAR, Human Resources Development Canada and CIRANO.

[‡] Human Resources Development Canada. Email: lucie.gilbert@spg.org

[§] CNRS, CREST and IDEI. Email: kamionka@ensae.fr

[¶] Département d'économique, Université Laval, CRÉFA and CIRANO. Email: guy.lacroix@ecn.ulaval.ca

Abstract : The analysis focuses on the impact of government-sponsored training programs aimed at disadvantaged male youths on their labour market transitions. The richness of the data at our disposal allows us to recreate very detailed individual histories over a relatively long period.

We use a continuous time duration model to estimate the density of duration times in as many as seven states, controlling for the endogeneity of an individual's training status. We investigate the sensitivity of the parameter estimates by comparing a typical non-parametric specification with a series of parametric two-factor loading models, as well as a parametric three-factor loading model.

Our results show that young, poorly educated males who participate in welfare training programs do worse on the labour market than those who do not participate. On the other hand, participation in unemployment training programs provides them some benefits. In general, we find that duration times in any of the seven states considered are sensitive to variations in program parameters such as welfare benefits, policy variables such as the minimum wage rate, and in the economic environment as proxied by the unemployment rate.

Résumé : Nous examinons les effets des dispositifs publics d'emploi sur les trajectoires entre les états du marché du travail des jeunes hommes défavorisés. La richesse des données à notre disposition nous permet de suivre ces trajectoires individuelles sur une longue période.

Nous considérons un modèle de transition à partir duquel nous estimons les distributions conditionnelles des durées de séjour dans les états du marché du travail en contrôlant le caractère endogène de la participation à ces programmes. Nous étudions la sensibilité des estimations en comparant une estimation semi-paramétrique type avec une série d'estimations paramétriques de modèles à deux et trois facteurs de charge.

Les résultats montrent que la participation aux dispositifs à l'assurance chômage augmente la fréquence des transitions vers l'emploi. Les hommes jeunes peu diplômés qui participent aux programmes à l'aide sociale transitent moins fréquemment en emploi que ceux qui ne participent pas à ces programmes. Nous trouvons que les durées de séjour dans les états du marché du travail sont sensibles aux variations du barème de l'aide sociale, à celles du taux de salaire minimum et à l'état de l'économie lorsqu'il est caractérisé par le taux de chômage.

Keywords : Training programs, Welfare benefit, Unemployment duration

JEL-Code : J68, E24, J64, J65

1 Introduction

The impact of government-sponsored training programs has been extensively studied in the past couple of decades.[1] In many countries, such programs have become an integral part of public policies aiming at enhancing self-sufficiency among vulnerable groups. The program costs have escalated as they have become more comprehensive and more systematically used. Not surprisingly, policy makers have shown renewed interest in obtaining accurate and reliable estimates of their efficacy.

The discussions surrounding the efficacy or desirability of training programs rest on complex methodological issues. The main concern lies with proper treatment of an individual's decision to participate in such programs. Severe biases may arise if unobserved individual characteristics that affect the decision to participate are somehow related to the unobservables that affect outcomes on the labour market. Two approaches have been proposed in the evaluation literature to address the so-called issue of "self-selection". The first is the "experimental approach", based on random assignment of applicants into treatment or control groups. The second is the "non-experimental", or "econometric approach", and relies on non-random samples of participants and non-participants. Each approach tackles the self-selection issue from a different angle, but the relative merit of each is still the subject of debate [see Heckman and Smith (1995), Burtless (1995), Ham and LaLonde (1996)].

Most would argue that the "experimental" approach is best suited to eliminate self-selection biases and provide adequate mean program impacts, however measured. Yet, recently this view has been challenged by Ham and LaLonde (1996) in their important paper. In essence they argue that random assignment between control and experimental groups provides an adequate *short-term* mean program impact. On the other hand, the treatment and controls experiencing subsequent spells of employment and unemployment are most likely not random subsets of the initial groups because the sorting process is very different for the two. In other words, random assignment does not guarantee that long-term mean program impacts are void of any systematic biases.

In most countries, experimental evaluation of training programs is impracticable due to a lack of appropriate data. Analysts must instead concentrate either on survey or administrative data, and rely on multi-state transition models. An additional difficulty in using these data is that program participation must be modeled explicitly. Many recent papers have nevertheless managed to successfully model complex transition patterns using such data (Gritz (1993), Bonnal, Fougère and Sérandon (1997) and Mealli, Pudney and Thomas (1996)). Most papers are limited to three separate states of the labour

[1] See Heckman, LaLonde and Smith (1999) for a recent and detailed survey.

market: employment, unemployment (non-employment) and training.[2] In many cases data limitations do not allow identification of any more states. In other cases, analysts purposely focus on few states to keep the statistical model tractable. Indeed, when the data is drawn from stock samples, as is often the case when using administrative data, the statistical model must account for so-called "initial conditions" problems. This usually adds considerable complexity to an already involved statistical model.[3] On the other hand, many have questioned the appropriateness of focusing of few labour market states (Heckman and Flinn (1983), Jones and Riddell (1999)).

This paper investigates the impact of government training programs aimed at poorly educated Canadian male welfare recipients. It should be stressed at the outset that in Canada, as in many European countries, the welfare system aims at supporting individuals without income and who are not entitled to any other social security benefits, irrespective of age.[4] As such, it acts as a safety net for unemployed workers who do not qualify for benefits, or who have exhausted their unemployment benefits. Many programs are available to assist these long term unemployed and those with few skills increase their employability. Understandably, a considerable proportion of program resources has been targeted towards the youths in the past decade. Yet, many have questioned the ability of traditional programs to address the problem [OECD, 1998]. The aim of this paper is precisely to investigate the impact of these programs in enhancing the self-sufficiency of young males welfare claimants, a particular disadvantaged group (see Beaudry and Green (1997)).

The empirical strategy is similar to that used by Gritz (1993) and Bonnal et al. (1997) in that we explicitly account for selectivity into the training programs. It relies on a rich dataset that tracks the transitions of a large number of individuals on a weekly basis across seven different states of the labour market. These states include employment, unemployment, welfare, out of the labour force (OLF), two separate welfare training programs, and unemployment training programs. In all, as many as 24 different transitions are allowed in the model. The sample is drawn from the population of welfare recipients that experienced a spell at any time between 1987 and 1993 in the province of Québec, Canada. To be included in the sample, individuals had to be aged 18 or

[2] One notable exception is Bonnal et al. (1997) who consider as many as 6 different states: permanent employment, temporary employment, public policy employment (training), unemployment, out-of-labour-force (non-employment), and an absorbing state (attrition).

[3] Two biases are likely to result from stock samples: (1) length-bias; (2) inflow-rate bias. The former may arise because lengthy spells are more likely to be ongoing at the time the sample is chosen. The latter is related to the fact that the probability of being sampled is related to the probability of starting a fresh spell at time the sample is chosen. See Gouriéroux and Monfort (1992) and Van den Berg, Lindeboom and Ridder (1994) for a detailed analysis.

[4] Individuals must be aged over 18 to qualify for benefits, although single parents less than 18 still qualify.

19 at any time during that period and to have less than a high-school degree. Sample stratification is used to avoid over-parameterization of the statistical model that would result if too many exogenous variables had to be controlled for.

By merging various administrative data files we can recreate complete individuals' histories on the labour market back to age 16, the legal school-leaving age in Canada. Consequently, each individual in our sample is necessarily observed in the OLF state at the beginning of his history. This sampling scheme thus removes the necessity to control for stock sample biases and has the additional benefit of providing rich transition patterns over a relatively long sample frame.

The econometric model is built on continuous labour market transitions processes and allows entry rates into each state to depend on observed and unobserved heterogeneity components. Heterogeneity terms can be destination-specific, origin-specific or both. In all cases, correlation across heterogeneity terms is allowed. We further investigate the sensitivity of the parameter estimates to various distributions of the heterogeneity components. When parametric distribution functions are used, the model is estimated by Simulated Maximum Likelihood (SML).

The remainder of the paper is organized as follows. Section 2 provides a detailed description the data. Section 3 discusses the econometric model and the various statistical assumption regarding the distributions of the heterogeneity terms. Section 4 reports our empirical findings. Section 5 concludes the paper.

2 Data Description

The basic data used for this study are drawn from the caseload records of Québec's Ministère de la Solidarité sociale. The files contain information on all individuals having received welfare benefits at some time between January 1987 and December 1993. In particular, the start dates and end dates of each welfare and welfare training spells are recorded in the files. The welfare program contains special provisions for those who are indisposed for work due to mental or physical impediments. These individuals are not included in the sample. Thus the final sample comprises only individuals that have no handicap or only a minor, intermediate, or temporary physical handicap. Furthermore, they are fit to work.

The welfare administrative files contain no information on employment or unemployment spells. Our sample was thus linked to the Status Vector files (SV) and the Record of Employment (ROE) files, both under the aegis of Human Resources Development Canada. These files contain very detailed weekly information on insured unemployment

spells and employment spells, respectively. The start dates and end dates of each spell are recorded in these files. Similar information is available with respect to training spells administered under the Unemployment Insurance (UI) program. Merging all three administrative files allows us to define seven different states on the labour market. Aside from the welfare, unemployment and employment states, we can identify two separate welfare training states and one unemployment training state.[5]

The focus of this paper is on poorly educated young men. Thus to be included in the sample, an individual had to be either 18 or 19 years of age at any time between 1987 and 1993 and have completed less than 11 years of schooling over the sample period. A high-school degree in Québec usually entails at least 12 years of schooling. In principle, then, none of the individuals in our sample has earned a high-school diploma. With these selection criteria the final sample contains 3068 individuals.

The upper panel of Table 1 provides summary statistics for individuals who have not participated in a training program. The lower panel presents similar statistics for program participants. In the latter case, the mean durations in either employment, unemployment or welfare are calculated both before and after training. An examination of the table reveals that the two groups are very similar in terms of their observable characteristics; They both have the same average age and nearly identical schooling levels. Yet, there are significant differences in their respective labour market experiences. For instance, non-trainees have longer spells in each of the three states reported in the table. On the whole, the proportion of time non-trainees spend employed is slightly larger than that of trainees prior to training. On the other hand, once they have had training, the proportion of time trainees spend employed become larger than that of non-trainees. This increase stems from the fact that the average employment duration decreases proportionately less that the average duration of welfare and unemployment spells. Taken at face value, this would suggest training programs benefit somewhat to welfare recipients.

Recall that only individuals who experienced a welfare spell between 1987 and 1993 and who were aged 18 or 19 during that period are included in the sample. Those who are 18 or 19 years of age in January 1987 may have already been on the labour market for 2–3 years at most. In order to recreate their complete labour market histories as of the age of 16, it is necessary in some cases to go back as early as January 1984.[6]

[5]The welfare files contain information dating back to 1979 and ending in December 1993. The SV files contains information beginning in January 1987 and ending in December 1996. Finally, The ROE files contain information ranging from January 1975 to December 1996. The analysis focuses on the 1987–1993 period due to data limitations.

[6]Data concerning unemployment spells are available only as of January 1987. Consequently, a small proportion of unemployment spells occurring prior to 1987 may be wrongly coded as out of the labour force (OLF). Two factors lead us to believe that the proportion of such spells is likely insignificant. First,

The start date and end date of each spell is used to create individual histories on the labour market. Overlaps between states are frequent and are not necessarily the result of coding errors. It may well be, for example, that a welfare spell and a work spell overlap. Program designs do not forbid this. In principle, such overlaps could be redefined as a separate state. Given the number of possible states, it is simply not reasonable to allow these overlaps in the analysis. It was decided that, as a rule, starting dates would have precedence over ongoing spells. Thus an ongoing spell with known end date is truncated whenever a new state starts prior to the end date.[7]

The 3 068 individuals in our sample experienced as many as 31 422 spells over the sample period. Table 2 presents all the transitions that occurred at any given point in the sample period. The table identifies seven separate states on the labour market. Welfare Training includes various job search assistance programs as well as skill enhancing programs aimed at welfare recipients. The Job-Reentry Program (JRP) is an on-the-job training program also aimed at welfare recipients. Under this program, participants do not receive benefits but a (subsidized) salary from a regular employer.[8] JRP is treated separately because contrary to other programs most participants qualify for unemployment benefits upon completion. UI is a state in which individuals receive unemployment benefits. Individuals that do not work and that do not qualify for benefits are treated as out of the labour force (OLF) for the purpose of this study. It must thus be kept in mind that UI is not necessarily akin to unemployment in the usual sense. UI Training comprises a series of training programs aimed at UI claimants. The OLF state is the complement of all other states. It includes full-time students, non-entitled unemployed individuals and individuals that are truly out of the labour force.

Table 2 reveals interesting dynamics on the labour market. For instance, the majority of welfare spells end either in employment, in welfare training or OLF. Likewise, welfare training spells end either in welfare, in employment or in OLF. Interestingly, most JRP

the large majority of individuals who were 18 or 19 years of age in the years 1990 and beyond where in the OLF, the employment or the welfare states between 16 and 19. Second, of those individuals, the majority who had an employment spell would not have qualified for UI benefits given the eligibility rules that prevailed between 1984 and 1987.

A similar problem arises with respect to employment spells. Indeed, spells that were ongoing in December 1993 will not show up in the ROE files until they are terminated. To avoid misclassifying these spells as OLF, the ROE files are searched as late as December 1996. Given the average length of employment spells reported in Table 1, it is very unlikely that many employment spells that were ongoing in December 1993 will still be ongoing as late as December 1996, and thus wrongly classified as OLF.

[7]Preliminary analysis was also conducted giving the end date precedence over the start date of a new spell. The resulting transitions matrices and average durations are very robust to this strategy.

[8]Non-profit organizations have to pay a symbolic 1$ per working day. The participants receive regular benefits.

participants enter regular employment upon completion of their program. Very few enter UI even though most qualify for benefits. Other transitions are as expected, except perhaps for UI training. Indeed, the majority of participants return to UI upon completion of their program and very few find regular employment. A number of cells contain few or no observations. The empty cells are consistent with program or policy parameters that prevent a number of transitions to occur or are a consequence of our definitions of the various states.[9] Only transitions comprising more than 75 observations will be considered in the econometric model. This leaves a total of 24 transitions to be modeled explicitly.

The transitions on the labour market have three essential dimensions: the state of origin, the state of destination and the duration in any a given state. Table 2 provides useful information on the first two dimensions. One way to represent all three dimensions simultaneously is to look at the distribution of the sample across all seven states on a weekly basis. This distribution synthesizes both the transitions across states and the mean duration in each.

Figure 1 plots the proportion of individuals in each of the seven states on a weekly basis. The top portion of the figure traces out the proportion of individuals in non-training states (welfare, unemployment, employment, OLF), and the bottom portion traces out the proportions in training states (UI training, welfare training and JRP). There are two distinct features that arise in January 1987 in the top portion of the figure. First, the proportion of individuals in OLF is relatively high. This partly reflects a cohort effect. In January 1987, our sample comprises only individuals that are 18 or 19 years of age. Not surprisingly, a large proportion of them are either still in school or have not yet entered the labour market. As we move rightward along the time axis, these individuals become older and new 18-19 year old entrants join the sample. By the time we reach December 1993, the oldest individuals are between 25–26 years of age. It does not necessarily follow that the sample's average age increases systematically along the time axis. Proportionately more individuals have entered the sample in the recession years 1989–1992 than previously. Second, the proportion of unemployed individuals is zero. As mentioned earlier, the information on unemployment spells is only available as of January 1987. Consequently, only new spells are identifiable in the data. Spells that were ongoing in January 1987 are classified as OLF in the figure.

[9]For example, the welfare files provide information on a monthly basis. Any interruption lasting between 1-3 weeks will not be recorded in the data. The record will show an uninterrupted sequence of monthly benefits receipt. Thus Welfare-Welfare transitions are not identifiable in the data. On the other hand, UI spells are recorded on a weekly basis. Unemployed workers that work a number of weeks or hours while claiming benefits may qualify for additional benefits once they exhaust their original entitlement. The SV files will indicate a new UI spell starting the week following exhaustion. Thus UI-UI transitions are identifiable in the data.

The bottom portion of the figure also indicates that the proportion of individuals in JRP is zero up until approximately January-February 1990. This program was implemented in August 1989 and had too few participants in the beginning months to show up in the figure. Similarly, participation in UI training programs is essentially zero up until February-March 1987. UI training usually occurs after a number of weeks has been spent unemployed. Not surprisingly, then, a certain laps of time is needed before the proportion of UI trainees is large enough to show up in the figure. Training spells that were ongoing in January 1987 are also classified as OLF.

A close look at Figure 1 reveals interesting patterns. First, the proportion of welfare participants remains relatively constant between 1987 and 1989. The economic downturn of 1989 results in an steady increase in the proportion of welfare claimants until the end of 1993. In fact, the proportion increased from 17.9% in January 1988 to 42.3% in December 1993. Such an increase results from both a more important inflow into welfare and longer spell duration [see Duclos, Fortin, Lacroix and Roberge (1999) for details].

The proportion of employed individuals follows a very distinct seasonal pattern with peaks occurring around June-July and troughs around January of each year. Despite these seasonal fluctuations, the proportion of employed individuals increased from 31.2% in January 1988 to 33.5% in January 1990, and then gradually declined to 18.6% in January 1993. The proportion of unemployed individuals is highly negatively correlated with the proportion of employed individuals. The seasonal fluctuations almost perfectly mirror those of employment. Finally, the proportion of individuals in the OLF state also depicts strong seasonal patterns. In January of each year, the proportion of those in OLF increases by about 5 percentage points. It is likely that many seasonal workers lose their job at the beginning of each year and do not qualify for unemployment benefits.

The bottom portion of the figure shows that the proportion of individuals engaged in government-sponsored training programs fluctuates considerably over time. A number of new welfare training programs have been implemented in 1989. Most of these programs aim at enhancing job search skills and usually last a few weeks. The large increase in the proportion of welfare trainees coincide with the implementation of these programs. A dramatic fall occurs towards the end of 1989 presumably linked to budgetary constraints associated with the economic downturn of 1990. The proportion of participants steadily increases thereafter and reaches its peak at the end of 1993. The proportion of UI trainees is relatively constant throughout the whole period, with the exception of 1992. Both the UI training programs and JRP have relatively few participants at any point in time.

The proportions of participants in the combined programs hardly reach beyond 5% over the sample period. The fact that few individuals are engaged in formal training at any point in time is no indication that training programs are inefficient or unattrac-

tive. Access to programs is often limited because of insufficient resources. This lack of resources raises a fundamental question: who gets selected into training ? To the econometrician, participation in a training program is the result of two separate unidentifiable processes. First, the participant has undertaken the necessary steps to take part in the program. Second, the program manager has deemed the participant as eligible. These two processes are likely to be such that participants have unobservable (to the econometrician) characteristics that are systematically different from those of the non-participants. Fortunately, given the information at our disposal it is possible to devise estimators that, under very general assumptions, will yield unbiased estimates of the programs' impacts. These estimators are presented in the next section.

3 Modeling labour market transitions

The labour market history of a given individual is represented by a sequence of n spells of various lengths in any of K (=7) states. Let x_t be the state in which an individual is observed to be at time t. The sequence starts at calendar time $\tau_0 = 0$ when the individual is 16 years of age and ends at time τ_e (τ_e = December 1993). Figure 2 depicts a hypothetical sequence made up of 3 spells of various length in 3 different states. As depicted, the individual is initially observed in the OLF state. He enters into employment at time τ_1 and eventually moves into unemployment at time τ_2. At time τ_e he is still in the midst of an unemployment spell.

Let τ_ℓ denote the calendar time at which a spell in any given state ends. Each spell ℓ ($1 \leq \ell \leq n$) is thus delimited by the start time $\tau_{\ell-1}$ and the end time τ_ℓ ($\tau_\ell > \tau_{\ell-1}$). Let u_ℓ be the duration of spell ℓ ($u_\ell = \tau_\ell - \tau_{\ell-1}$). Finally, let r denote a complete sequence from time 0 to time τ_e :

$$r = ((u_1, x_{\tau_1}), \ldots, (u_{n-1}, x_{\tau_{n-1}}), (u_n, 0)),$$

where $u_n = \tau_e - \tau_{n-1}$ is the duration of the last spell. The last spell of each individual is right-censored since τ_n and x_{τ_n} are not observed. On the other hand, the last spell must have lasted at least $\tau_e - \tau_{n-1}$ units of time in state $x_{\tau_{n-1}}$. Because x_{τ_n} is not observed we conventionally fix $x_{\tau_n} = 0$.

The sequence may be more compactly rewritten as:

$$r = (y_1, \ldots, y_n),$$

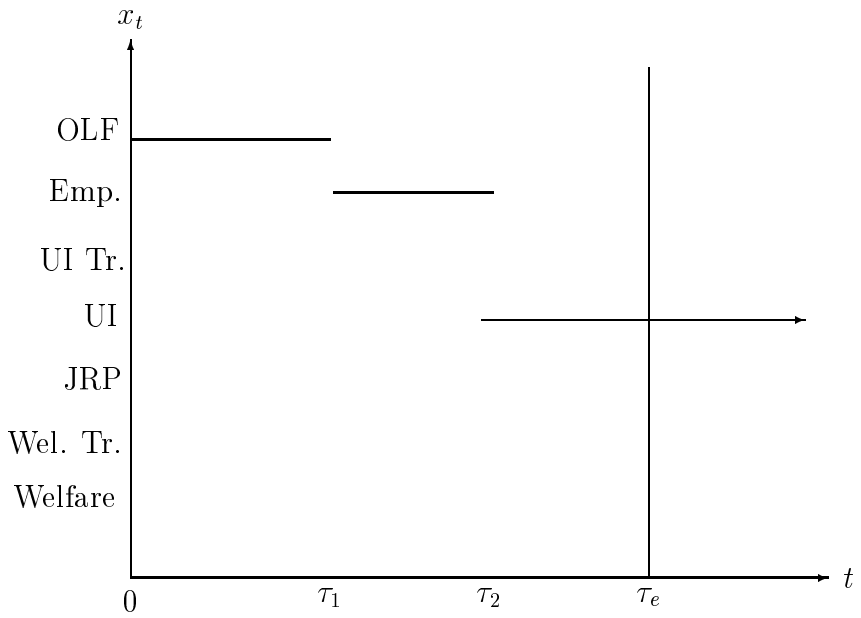


Figure 2: Labour market history of a hypothetical individual.

where

$$y_\ell = \begin{cases} (u_\ell, x_{\tau_\ell}), & \text{if } 1 \leq \ell \leq n-1, \\ (u_n, 0), & \text{if } \ell = n. \end{cases}$$

The initial state, x_0, is the same for each individual in our sample and exogenously determined by school attendance laws. Consequently, there is no need to explicitly model the initial state in which individuals are observed.

3.1 Likelihood function

Each individual contributes a sequence $r = (y_1, \ldots, y_n)$ to the likelihood function. The contribution can be written conditionally on a vector of exogenous variables, z, and an unobserved heterogeneity factor, ν.

Let $l_v(\theta)$ denote the conditional contribution of the sequence r. We have,

$$l_v(\theta) = \prod_{\ell=1}^{n} f(y_\ell \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta),$$

where $f(y_\ell \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta)$ is the conditional density of y_ℓ given y_1, ..., $y_{\ell-1}$, z and ν, and $\theta \in \Theta \subset I\!R^p$ is a vector of parameters. Naturally, the destination state of the

last spell is unknown since the duration is censored. Its contribution to the conditional likelihood function is limited to the survivor function of the observed duration.

The random variable ν is assumed to be independently and identically distributed across individuals, and independent from the exogenous variables z. If the unobserved heterogeneity can take only a finite number of values, $\nu_1, \ldots, \nu_J$, the contribution of a sequence r to the likelihood function is

$$l(\theta) = \sum_{j=1}^{J} \prod_{\ell=1}^{n} f(y_\ell \mid y_1, \ldots, y_{\ell-1}; z; \nu_j; \theta)\ \pi_j, \tag{1}$$

where π_j is the probability that the unobserved heterogeneity term takes the value ν_j ($0 \leq \pi_j \leq 1$, $\sum_{j=1}^{J} \pi_j = 1$).

If ν is a continuous random variable, then

$$l(\theta) = \int_V \prod_{\ell=1}^{n} f(y_\ell \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta)\ g(\nu; \gamma)\ d\,\nu, \tag{2}$$

where $g(\nu; \gamma)$ is a density probability function and V is the support of ν.

Furthermore, if we assume that Y_ℓ is independent of $Y_1, \ldots, Y_{\ell-2}$ given $Y_{\ell-1} = y_{\ell-1}$, $Z = z$ and the value of the unobserved term ν, in which case

$$f(y_\ell \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta) = f(y_\ell \mid y_{\ell-1}; z; \nu; \theta).$$

Given the history of the process, the joint distribution of the duration of spell ℓ and the destination state only depends on the current state on the labour market. This assumption will be relaxed by introducing other characteristics of the history of the process.

3.2 Modeling individual spells

In this section we focus on the conditional distribution of $y_\ell = (u_\ell, x_{\tau_\ell})$, where u_ℓ is the duration of the ℓ^{th} spell in state $x_{\tau_{\ell-1}}$. Define $u^*_{\ell,k}$ as the waiting time before leaving state $x_{\tau_{\ell-1}}$ for state k. At the end of the ℓ^{th} spell, the individual will enter into the state corresponding to the smallest latent duration $u^*_{\ell,k'}$. We will assume that these K latent durations are independently distributed.

Thus the duration of spell ℓ is given by[10]

[10] If the transition from state i to state j cannot be observed, we assume that the corresponding latent distribution is defective and put a set a probability mass equal to one on $+\infty$.

$$u_\ell = \inf_{k'} u^*_{\ell,k'}.$$

Let $f_j(u \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta)$ denote the probability density function (p.d.f.) of the latent duration $u^*_{\ell,j}$, given the history of the process up to time $\tau_{\ell-1}$, ν and covariates z. Let $S_j(u \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta)$ be the corresponding survivor function:

$$S_j(u \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta) = \int_u^{+\infty} f_j(s \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta)\ d\, s.$$

The conditional joint density of the duration of spell ℓ and the destination state k is given by the following expression

$$\begin{aligned} f(u, k|y_1, \ldots, y_{\ell-1}; z; \nu; \theta) &= f_k(u \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta) \\ &\quad \prod_{\substack{j=1 \\ j\neq k}}^{K} S_j(u \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta), \\ &= h_k(u|y_1, \ldots, y_{\ell-1}; z; \nu; \theta)\, S(u|y_1, \ldots, y_{\ell-1}; z; \nu; \theta), \end{aligned}$$

where $h_k(u \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta)$ is the hazard function associated with the latent duration $u^*_{\ell,k}$ and $S(u \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta)$ is the survivor function of the duration of the ℓ^{th} spell. Because the latent durations are assumed to be conditionally independent we have

$$S(u \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta) = \prod_{j=1}^{K} S_j(u \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta),$$

where $u \geq 0$. The expression represents the conditional probability that the duration of spell ℓ is at least equal to u or, equivalently, that all latent durations are at least equal to u. Therefore, the conditional contribution of a given sequence to the likelihood function is:

$$l_v(\theta) = \prod_{\ell=1}^{n} \prod_{k=1}^{K} h_k(u_\ell \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta)^{\delta_{\ell,k}}\ S_k(u_\ell \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta),$$

where $\delta_{\ell,k}$ is equal to 1 if the individual enters into state k at the end of spell ℓ and to 0 otherwise :

$$\delta_{\ell,k} = \begin{cases} 1, & \text{if } x_{\tau_\ell} = k, \\ 0, & \text{otherwise,} \end{cases}$$

$\ell = 1, \ldots, n$.

3.3 Unobserved heterogeneity

So far the discussion surrounding the unobserved heterogeneity components has voluntarily been kept general. The use of maximum likelihood procedures requires that we specify distribution functions for these components. Most applications rely on the work of Heckman and Singer (1984) and approximate arbitrary continuous distributions using a finite number of mass points (see Gritz (1993), Ham and Rea (1987)). More recent papers use richer specifications that allow the heterogeneity terms to be correlated across states (see Bonnal et al. (1997), Ham and LaLonde (1996)). These specifications are sometimes referred to as single or two-factor loading distributions and are also based on a finite set of mass points. In our work, we wish to investigate the robustness of the parameter estimates to various distributional assumptions. We will use two and three-factor loading distributions as in the aforementioned papers. Additionally, we will investigate the consequences on the slope parameters of using various continuous distributions instead of the usual finite sets of mass points.

To fix ideas, let $w = (w_1, \ldots, w_K)$ be a vector of unobserved heterogeneity variables, with w_k a destination-specific component ($k = 1, \ldots, K$). Ideally, the joint distribution of the unobserved heterogeneity terms should not be independent.

Consider first a two-factor loading model (see Van den Berg (1997)) such that

$$w_k = \exp(a_k \; v_1 + b_k \; v_2), \tag{3}$$

where $v_1 \in \{-2, c_2\}$, $v_2 \in \{c_1, c_2\}$, $b_k \in I\!R$, $a_k = 1\!\!1[k \geq 2]$ and $b_1 = 1$. The random variables v_1 and v_2 are assumed to be independent. The constraints imposed on the support of v_1 and v_2 are sufficient for identification and to allow the correlation between $\log(w_k)$ and $\log(w_{k'})$ to span the interval $[-1; 1]$.

Moreover, assume that

$$\text{Prob}[(V_1, V_2) = (v_1^0, v_2^0)] = \begin{cases} p^2, & \text{if } v_1^0 = -2 \text{ and } v_2^0 = c_1, \\ p * (1-p), & \text{if } v_1^0 = -2 \text{ and } v_2^0 = c_2, \\ (1-p) * p, & \text{if } v_1^0 = c_2 \text{ and } v_2^0 = c_1, \\ (1-p)^2, & \text{if } v_1^0 = c_2 \text{ and } v_2^0 = c_2, \end{cases} \tag{4}$$

where $c_1, c_2 \in I\!R$ and the probability p is defined as

$$p = \frac{\exp(d)}{1 + \exp(d)},$$

where $d \in I\!R$ is a parameter.

The correlation between $\log(w_k)$ and $\log(w_{k'})$, denoted $\rho_{k,k'}$, is

$$\rho_{k,k'} = \frac{a_k \; a_{k'} \sigma^2_{v_1} + b_k \; b_{k'} \; \sigma^2_{v_2}}{\sqrt{a_k^2 \; \sigma^2_{v_1} + b_k^2 \; \sigma^2_{v_2}} \sqrt{a_{k'}^2 \; \sigma^2_{v_1} + b_{k'}^2 \; \sigma^2_{v_2}}}, \tag{5}$$

where $k, k' = 1, \ldots, K$ and $\sigma^2_{v_j}$ is the variance of v_j, j=1,2. A positive correlation coefficient between w_j and w_k implies that those who are likely to have high transition rates between any given state and state j will also have high transition rates into state k.

A two-factor loading model with two independent heterogeneity terms with common continuous distribution can also be derived from this specification. As before, let w_k denote the heterogeneity term for destination k:

$$w_k = \exp(a_k \; v_1 + b_k \; v_2),$$

where a_k and b_k are parameters ($a_k = 1\!\!I[k \geq 2]$ and $b_1 = 1$).

Here v_1 and v_2 are assumed to be independently and identically distributed. Let $q(v; \gamma)$ be the p.d.f. of v_1 and v_2. The correlations between $\log(w_k)$ and $\log(w_{k'})$ are given by the same expression as in (5). In principle, $q(v; \gamma)$ represents any well-behaved distribution function.

The above specification can be further generalized to a three-factor loading model with common continuous distribution. In this case the unobserved components depend on the destination state as well as the current state. Let $w_{j,k}$ be specific to the transition between origin j and destination k.

$$w_{j,k} = w'_j \; w_k = \exp(a'_j \; v_3 + b'_j \; v_2) \times \exp(a_k \; v_1 + b_k \; v_2), \tag{6}$$

where a'_j, b'_j, a_k and b_k are parameters ($a'_j = 1\!\!I[j \geq 2]$, $a_k = 1\!\!I[k \geq 2]$, $b_1 = 1$).

In this three-factor loading model, the correlation between destination states k and k' is

$$\rho_{k,k'} = \frac{a_k \; a_{k'} + b_k \; b_{k'}}{\sqrt{a_k^2 + b_k^2} \sqrt{a_{k'}^2 + b_{k'}^2}}. \tag{7}$$

This correlation has the same interpretation as in the two-factor loading model.

On the other hand, the correlation between the two origin states j and j' is given by

$$\rho_{j,j'} = \frac{a'_j \; a'_{j'} + b'_j \; b'_{j'}}{\sqrt{a'^2_j + b'^2_j}\sqrt{a'^2_{j'} + b'^2_{j'}}}. \tag{8}$$

A positive correlation indicates that those who have short spells in state j are likely to have short spell duration in state j' as well.

Finally, the correlation between origin state j and destination state k is given by

$$\rho_{k,j} = \frac{b'_j \; b_k}{\sqrt{a'^2_j + b'^2_j}\sqrt{a^2_k + b^2_k}}, \tag{9}$$

where $j, j', k, k' = 1, \ldots, K$. This correlation is somewhat trickier to interpret. A positive coefficient indicates that those who are likely to have short spell duration in state j are also more likely to enter state k. Conversely, those who are more likely to have short spell duration in state j are less likely to enter state k.

3.4 Specification of conditional hazard functions

Assume an individual is observed in state j during spell ℓ (i.e. $x_{\tau_{\ell-1}} = j$). Let $\psi(j,k)$ denote the heterogeneity term for destination k, given the individual is in state j. There are two possibilities:

$$\psi(j,k) = \begin{cases} w_k, & \text{in the two-factor loading model,} \\ w_{j,k}, & \text{in the three-factor loading model.} \end{cases}$$

The conditional hazard function for transition (j,k) is given by

$$h_{j,k}(u \mid y_1, \ldots, y_{\ell-1}; z; \nu; \theta) = h^0_{j,k}(u; \theta) \; \varphi(y_1, \ldots, y_{\ell-1}; z; \theta) \; \psi(j,k), \tag{10}$$

where φ is a positive function of the exogenous variables and the sequence r, $h^0_{j,k}(u;\theta)$ is the baseline hazard function for transition (j,k), and $\psi(j,k) > 0$.

We have considered three alternative conditional specifications for the baseline hazard functions. For each transition, we have chosen among the following competing specifications on the basis of non-parametric kernel estimations (see Fortin, Fougère and Lacroix (1999a)):

1. *Log-logistic Distribution*
   The baseline hazard function is

$$h^0_{j,k}(u;\theta) = \frac{\beta_{j,k}\ \alpha_{j,k}\ u^{\alpha_{j,k}-1}}{(1+\beta_{j,k}\ u^{\alpha_{j,k}})},$$

   $\alpha_{j,k}, \beta_{j,k} \in I\!R^+$.

   If $\alpha_{j,k} > 1$ then the hazard function is increasing then decreasing with respect of u. If $\alpha_{j,k} \leq 1$ then the hazard function is decreasing.

2. *Piecewise-Constant Hazard Model*
   The expression of the baseline hazard function is

$$h^0_{j,k}(u;\theta) = \alpha_{j,k} I\!I[u < u^0_1] + \beta_{j,k} I\!I[u^0_1 \leq u < u^0_2] + \gamma_{j,k} I\!I[u^0_2 \leq u],$$

   where $\alpha_{j,k}, \beta_{j,k}, \gamma_{j,k} \in I\!R^+$. u^0_1 and u^0_2 are fixed.

   The baseline hazard function can be increasing then decreasing, decreasing then increasing, strictly increasing or strictly decreasing.

3. *Weibull Distribution*
   The baseline hazard function is

$$h^0_{j,k}(u;\theta) = \alpha_{j,k}\ \beta_{j,k}\ u^{\alpha_{j,k}-1},$$

   $\alpha_{j,k}, \beta_{j,k} \in I\!R^+$.

   If $\alpha_{j,k} > 1$ then the hazard function is increasing with respect of u. If $\alpha_{j,k} < 1$ then the hazard function is decreasing with respect of u and if $\alpha_{j,k} = 1$ this conditional hazard function is constant.

3.5 Estimation

We consider three alternative specifications for the unobserved heterogeneity distribution.

1. *Two-Factor Loading and Discrete Distribution*

   The log likelihood is

$$\log(L(\theta)) = \sum_{i=1}^{N} \log(l_i(\theta)), \tag{11}$$

where $l_i(\theta)$ is obtained by substituting the sequence $r_i = (y_{1,i}, \ldots, y_{n_i,i})$ and the observed vector of covariates z_i in (1). N is the size of the sample.

In equation (1) π_j is set equal to[11]

$$\pi_j = \begin{cases} p^2, & \text{if } j = 1, \\ p * (1-p), & \text{if } j = 2, 3, \\ (1-p)^2, & \text{if } j = 4, \end{cases}$$

where $p \in [0; 1]$ is a parameter. The log-likelihood is then maximized with respect of θ ($\theta \in \Theta$).

2. *Two-Factor Loading and Continuous Distribution*

The model includes two unobserved heterogeneity terms v_1 and v_2 ($v_j > 0, j = 1, 2$). We assume these terms to be independently and identically distributed. Let $q(v; \gamma)$ be the p.d.f. of v_j, $j = 1, 2$.

The contribution of a given realization to the likelihood function is given by equation (2), where $\nu = (v_1, v_2)'$, $V = I\!R^+ \times I\!R^+$ and $g(\nu; \gamma) = q(v_1; \gamma)\, q(v_2; \gamma)$. The log-likelihood is given by equation (11), where $l_i(\theta)$ is the contribution to the likelihood of the sequence r_i.[12] Since the integral in $l(\theta)$ cannot generally be analytically computed it must be numerically simulated.

Let $\hat{l}(\theta)$ denote the estimator of the individual contribution to the likelihood function. We assume that

$$\hat{l}(\theta) = \frac{1}{H} \sum_{h=1}^{H} \prod_{\ell=1}^{n} f(y_\ell \mid y_1, \ldots, y_{\ell-1}; z; v_{1,h}, v_{2,h}; \theta),$$

where $v_{1,h}$ and $v_{2,h}$ are drawn independently according to the p.d.f. $q(v; \gamma)$. The drawings $v_{j,h}$ ($j = 1, 2$, $h = 1, \ldots, H$) are assumed to be specific to the individual. The parameter estimates are obtained by maximizing the simulated log-likelihood:

$$\log(L(\theta)) = \sum_{i=1}^{N} \log(\hat{l}_i(\theta)),$$

where $\hat{l}_i(\theta)$ is the simulated contribution of the sequence r_i to the likelihood function.

The maximization of this simulated likelihood yields consistent and efficient parameters estimates if $\frac{\sqrt{N}}{H} \to 0$ when $H \to +\infty$ and $N \to +\infty$ (see Gourriéroux

[11] See section 4.

[12] In what follows, θ includes γ, the parameters of $q(\cdot)$.

and Monfort (1991, 1996)). Under these conditions, this estimator has the same asymptotic distribution as the standard ML estimator. Following Laroque and Salanié (1993) and Kamionka (1998) we have used 20 draws from the random distributions when estimating the models. Using as few as 10 draws yielded essentially the same parameter estimates.

3. *Three-Factor Loading and Continuous Distribution*

   In the three-factor loading model the conditional contribution must be integrated with respect to the distribution of three independent unobserved heterogeneity terms. Let $\hat{l}(\theta)$ denote the estimator of the individual contribution to the likelihood function. Assume further that

$$\hat{l}(\theta) = \frac{1}{H} \sum_{h=1}^{H} \prod_{\ell=1}^{n} f(y_\ell \mid y_1, \ldots, y_{\ell-1}; z; v_{1,h}, v_{2,h}, v_{3,h}; \theta),$$

   where $v_{1,h}$, $v_{2,h}$ and $v_{3,h}$ are drawn independently according to the p.d.f. $q(\nu; \gamma)$. Once again, the parameter estimates obtained from maximizing this function are asymptotically efficient.

4 Estimation Results

This section presents the results of fitting the models outlined in the previous section to the data at our disposal. The estimation of such complex models is computationally demanding. Also, a number of issues must be addressed before dwelling into the results.

4.1 Functional Forms Assumptions

As mentioned in the previous section, it is necessary to specify a baseline distribution function for each transition considered in the model. When selecting a particular functional form, a number of desirable properties should be sought. First, the functional form should allow a number of different shapes of the hazard function so that various combinations of positive and negative duration dependence are possible. Second, it should roughly follow the pattern of transitions times found in the data. Finally, the functional forms should involve as few parameters as possible.

The data at our disposal was analyzed in Fortin et al. (1999a) using non-parametric kernel hazard estimators. The baseline hazard functions were chosen on the basis of their analysis. Table 3 reports the functional form used in each of the 24 transitions

considered in the model. Both the log-logistic and the piecewise constant functions allow non-monotonic hazards. For many transitions, the empirical hazard functions initially increase for a short period of time and then display an extended period of negative duration dependence. The log-logistic function is best suited in these cases. When the empirical hazard function looks relatively flat, it is preferable to use an exponential model with a single parameter. Other non-monotone shapes are best approximated with the piecewise constant hazard function. Monotone increasing or decreasing empirical hazard rates can be satisfactorily approximated with a weibull distribution function.

4.2 Exogenous Covariates

Most studies on labour market transitions include a number of exogenous individual-specific and macroeconomic variables. It is thus customary to include variables such as age, sex, education and minority status to capture behavioural differences across these groups. In this paper we have tried to limit the number of exogenous control variables as much as possible. Given the unusually large number of transitions considered in the analysis, including even as little as 10 exogenous variables would have over-parameterized the likelihood function and rendered its estimation practically infeasible.

An alternative empirical strategy is to circumscribe the sample to relatively homogeneous individuals in terms of observable characteristics. We have elected to concentrate our attention on young and poorly educated men for two reasons: (1) They have fared relatively poorly on the labour market over the past decade (see Beaudry and Green (1997)); (2) As a consequence of their deteriorating labour market outcomes, many have claimed welfare benefits and have been especially targeted for training programs. Having a relatively homogeneous sample in terms of age and education does not remove the need to control for such variables explicitly. Our sampling scheme insures that there is little variance in age at the start of the sample period (see Table 1). As the initial individuals become older, new entrants 18–19 years of age join the sample, thus increasing considerably the variance in age. On the other hand, the sample was chosen so that educational attainment never exceeded 10 years of schooling. Consequently, the variance in education remains relatively constant over the sample period.

We thus explicitly control for age in the regressions. Note that Gritz (1993) has found both education and age to have little impact on any of the transitions considered in his model. The following exogenous variables are included in the model in addition to age: minimum wage, unemployment rate, welfare benefits, and dummy indicators for previous training under either welfare or UI. The minimum wage and the welfare benefits are computed monthly and deflated by the monthly Consumer Price Index (CPI). The monthly unemployment rate is computed for men aged 25-64 for the Province of Québec.

All the variables are computed at the beginning of each spell and are assumed constant throughout the duration of individual spells.

4.3 Parameter Estimates

Table 4 presents the parameter estimates of a three-factor loading model that incorporates a weibull distribution for the heterogeneity variables.[13] The slope parameters of the non-parametric and the (weibull) two-factor loading models are nearly identical to those presented in Table 4 and are not reported for the sake of brevity.

Table 4 is divided into several panels. Each panel contains the parameter estimates for the exit rates of a given state. The parameter estimates of the baseline hazard are presented first followed by those of the control variables. The variable "Wel Tr_1" is a dummy indicator that equals 1 if the individual has experienced a welfare training spell or has participated in JRP at any time prior to the ongoing spell, and 0 otherwise. The variable "Wel Tr_2" is a dummy indicator that equals 1 if the state just prior to the current spell was either welfare training or JRP, and 0 otherwise. The variables "UI Tr_1" and "UI Tr_2" are similarly defined but pertain to UI training programs. The inclusion of "Wel $Tr1_1$" or "UI Tr_1" alone implicitly assumes that the impact of training programs does not wear off with time nor that it accumulate with repeat uses. Including both "Wel Tr_1" and "Wel Tr_2" or "UI Tr_1" and "UI Tr_2" allows to determine whether recent training has more impact than previous training on current spell duration. Both past and recent training variables are included whenever feasible.

4.3.1 Exits from Welfare

The first panel of Table 4 focuses on exits from welfare. Exits to as many as five different states are considered in the model. Parameters related to age indicate that as individuals get older they are more likely to enter employment or OLF upon leaving welfare. In the latter case, this may be an indication that they are more inclined to return to school. Increases in the minimum wage rate increases the transitions towards welfare training,

[13] The model was also estimated using normal, student-t, χ^2 and gamma distributions. The results based on these specifications are not reported here for the sake of brevity, but are available on request. The specification based on the weibull was preferred to all others for two reasons. First, the parameter estimates based on the weibull distribution are very similar to those based on discrete distributions with a finite number of mass points. Given the latter are robust to specification errors on the distribution of the heterogeneity components (see Heckman and Singer (1984)), the weibull distribution appears to depict similar properties. Second, as in Heckman and Singer (1984), the value of likelihood function based on the weibull distribution is larger than those based on other distributions.

JRP and unemployment, but has no impact on transitions into employment. This result is compatible with the results found in a recent paper by Fortin and Lacroix (1997). In that paper it was found using a similar sample that increases in the minimum wage rate increased exits from welfare. Since the transition state was not known, this was interpreted as evidence that firms were not constrained by the minimum wage rate. Instead, an increase in the latter was interpreted as attracting a number of welfare claimants onto the labour market. The results reported here provide a completely different story. Indeed, it appears that increases in the minimum wage rate induce welfare claimants to increase their employability status but does not translate into a larger number being employed. Quite to the contrary, the increased transition rates from welfare to unemployment suggest that a number of individuals that were working while claiming welfare benefits may have lost their job following the increase in the minimum wage rate.

Increases in the unemployment rate translate into smaller transition rates into JRP. This result is compatible with the fact that welfare claimants may be less motivated to increase their employability when job prospects diminish. Alternatively, firms may also be less inclined to hire trainees under the JRP program when the unemployment rate rise.

As expected, increases in welfare benefits decrease the exit rates from welfare. The result is statistically significant in transitions towards training, work and OLF states. A similar finding was reported by Fortin and Lacroix in the aforementioned paper.

Past occurrences of welfare training are generally not very beneficial to the men in our sample. They are associated with higher transition rates into welfare training and lower rates into employment and OLF. The impact is larger for recent occurrences, which suggests that participation in such training programs may convey a bad signal to potential employers. On the other hand, past occurrences of UI training has little impact on the exits from welfare.

4.3.2 Exits from Unemployment

The next panel of the table focuses on the transitions from unemployment. Most parameter estimates that are statistically significant have the expected sign *a priori*. For instance, it is found that as individuals get older they are more likely to exit unemployment for employment and less for welfare. Similarly, increases in the minimum wage rate leads to higher transition rates into UI training but lower rates into employment. These results are consistent with those found with respect to exits from welfare.

Other results presented in the panel indicate that unemployed individuals are more likely to experience a new unemployment spell or to enter welfare and are less likely to

enter employment whenever the unemployment rate increases. Presumably, a number of UI claimants can not find employment and therefore exhaust their benefits. The social security system in Canada entitles them to welfare benefits upon exhaustion of UI benefits. On the other hand, increases in welfare benefits increase the transition rates into welfare and lower those into unemployment and employment. These results suggest that the transitions towards employment are very sensitive to both policy variables, *i.e.* welfare benefits and minimum wages, as well as to the state of the economy as proxied by the unemployment rate.

A number of parameter estimates relating to the training dummy variables are statistically significant. Once again, previous participation in welfare training increases the likelihood of entering welfare upon leaving unemployment and decreases that of entering employment. On the other hand, recent UI training participation appears to have a conflicting impacts. Indeed, UI claimants are more likely to enter either welfare or UI upon leaving unemployment but are also more likely to enter employment. On the whole, these are consistent with those found by Fortin, Fougère and Lacroix (1999b) using different data and econometric estimators and are also consistent to some extent with those of Gritz (1993) and Bonnal et al. (1997). In all three cases it was found that participation in government-sponsored training programs had detrimental effects on the labour market experience of young men. It has been suggested that potential employers may stigmatize participation in such training programs. Because these programs are designed to improve the labour market opportunities of disadvantaged workers, participation in the later may be taken as a signal of unsatisfactory performance in previous employment. Our results indicate that training while on welfare is detrimental to men, but training while on unemployment does not convey the same negative signal.

4.3.3 Exits from Employment

The next panel of the table reports results relating to transitions from employment. Once again, most parameters estimates that are statistically significant have the expected sign. In particular, increases in the minimum wage rate is found to increase the likelihood of leaving employment for either welfare training, and to diminish considerably the likelihood of entering a new job or moving into welfare. Increases in welfare benefits are found to increase the transitions into welfare and to decrease the likelihood of entering welfare training.

The parameter estimates associated with the unemployment rate has the expected sign except perhaps with respect to transitions between employment and unemployment. Indeed, the parameter estimate implies that whenever the unemployment rate increases, workers are less likely to leave employment to enter unemployment. There are several

potential explanations for this result. First, it may well be that when the labour market deteriorates, workers who loose their job have difficulty qualify for UI benefits. Recall from Table 1 that the unconditional mean job duration is approximately 18 weeks. which is roughly equal to the qualifying period. They are thus more likely to turn to welfare, as indicated in the first column of the panel. Second, the deterioration of the labour market may induce some to hold on to their current job longer. The fact that all the parameter estimates are negative, except for welfare, is consistent with this possibility. Finally, increases in welfare benefits increase the transitions from employment to welfare, as expected.

The training variables show interesting results. For instance, those who have participated in welfare training are more likely to enter either welfare or welfare training upon exiting employment, although recent participation makes them less likely to enter welfare anew. Likewise, participation in welfare training translates into less employment–employment transitions. Those who were in UI training just prior to their current employment spell are much more likely to return to UI upon leaving employment and much less likely to experience an employment-employment transition. The likelihood of entering the OLF state following employment decreases substantially if the individual experienced either UI or welfare training in the past.

4.3.4 Exits from OLF

The results presented in the following panel relate to the OLF state. Recall that this state includes individuals that are truly out of the labour force but may also include full-time students and non-entitled unemployed workers. Caution must thus be exercised in interpreting these results.

Surprisingly many parameter estimates turn out to be statistically significant. Of particular interest, transitions from OLF to employment appear to be quite sensitive to the economic environment. Transitions to employment are thus less when the minimum wage rate or the welfare benefits increase. Similarly, the transitions into welfare and welfare training are relatively sensitive to policy variables. As in previous panels, the transitions into welfare training are more likely for those who have previously experienced such training.

For the sake of brevity, the estimation results for training programs are presented in a separate table in an appendix. The econometric model generally does a poorer job at predicting transitions from the training programs compared to those for other states of the labour market,although a number of parameter estimates are statistically significant.

4.3.5 Unobserved Heterogeneity

Table 5 reports the value of the likelihood function for a number of different specifications as well as the parameter estimates related to the unobserved heterogeneity of each. As mentioned earlier, the slope parameters of these specifications are sufficiently similar to omit them from the tables.[14]

The first specification of the table does not control for unobserved heterogeneity and is thus a special case of all the other specifications. A simple likelihood-ratio test strongly rejects the first specification in favour of any specification that includes unobserved heterogeneity. The second specification is a standard non-parametric two-factor loading model and was presented in equation (4). Most parameter estimates are statistically significant, except for b_2 and b_5 which concern transitions into welfare training programs and UI training programs, respectively. Accordingly, these estimates suggest there is little, if any, selectivity into these two training programs.

The third column of the table reports the parameter estimates of a parametric two-factor loading model. As was mentioned earlier, the weibull distribution function was preferred over all other distribution functions that were investigated. Notice that as in the non-parametric specification, only b_2 and b_5 are not statistically significant. The last two lines of the table report the parameter estimates of the weibull distribution, λ and γ.[15] Finally, the last column of the table presents the parameter estimates of the three-factor loading model (see equation(6)), whose slope parameters were presented in Table 4. A simple log-likelihood ratio test rejects the two-factor loading model in favour of the three-factor loading model. Contrary to the two previous specifications, b_2 is now highly statistically significant. Furthermore, nearly all the b'_j parameters are statistically significant. This suggests that the richer specification may be better suited to uncover selection into the different states.

In order to investigate this issue, Table 6 reports the correlation coefficients between the heterogeneity variables that are implicit in each specification along with their standard errors. The first two panels focus on the non-parametric and the weibull two-factor loading models. Recall that these correlation coefficients indicate the extent to which one is as likely to enter state j as state k upon leaving any given state. While a number of coefficients are similar across both panels, there are significant differences. To start with,

[14] Bonnal et al. (1997) also found the slope parameters to be relatively insensitive to the distributional assumptions of the unobserved heterogeneity variables. In their work, they compare a two-factor loading model with a finite number of points of support with a single-factor loading model that draws heterogeneity terms from an i.i.d. $I\!N(0,1)$ distribution. The insensitivity of the slope parameters to the distributional assumption is consistent with the results of Heckman and Singer (1984) using single durations data.

[15] The weibull distribution function of a random variable x is given by $F(x) = 1 - \exp[-\lambda x^{\gamma}]$.

the first line of each panel shows that high transition rates into welfare are associated with lower transition rates into welfare training and higher rates into unemployment. On the other hand, both panels disagree significantly with respect to the correlations between welfare training and the other states, as well as between JRP and other states. The non-parametric model implies that welfare training and UI training are positively correlated whereas the opposite holds true in the parametric model. Similarly, the top panel indicates that JRP is positively correlated to all other states on the labour market, contrary to the parametric model which shows no such relations.

The last panel of Table 6 focuses on the correlation coefficients implicit in the three-factor loading model. Each section of the panel is related to the correlation coefficients in equations (7)–(9), respectively. Hence, the first section has the same interpretation as the correlations of the previous panels. The correlation coefficients reported in this section differ considerably from the previous ones. According to the estimates, it now appears that there is considerable selectivity into welfare training as well as in JRP. Indeed, those who are more likely to participate in the former are also more likely to train under JRP and to find employment. On the other hand, higher transition rates into JRP or welfare training is now associated with lower transition rates into UI and UI training. This is in stark contrast with the previous results. Other correlation coefficients are relatively similar to the previous ones.

The second section of the panel reports the correlation coefficients with respect to the origin states. Large heterogeneity values in the origin state translate into short spell durations. Consequently, the correlations reflect the frequency with which individuals transit across the various states. The estimates show that individuals who are more likely to have long welfare spells are also likely to have short employment spells. The same holds with respect to welfare training and employment, as well as JRP and employment. Those who are more likely to have short unemployment spells are more likely to have long JRP, welfare training or UI training spells.

The last section of the panel reports the implicit correlations between the origin and the destination states. Note that the correlation matrix need not be symmetric nor does the diagonal need be equal to unity. On the other hand, the restrictions that were imposed to achieve identification of the loading parameters imply that the first row of the matrix is equal to the first row of the matrix of the middle section.

For the sake of brevity we will focus our attention on the most interesting correlations. The estimates suggest that those who are likely to have short welfare training spells are also less likely to transit through welfare or JRP and more likely to enter employment. (row 2). Similarly, row 3 indicates that individuals who are likely to have short JRP spells are less likely to return to either welfare or welfare training in the future, and much more likely to enter employment. Finally, those who have short UI training spells (row

5) have higher transitions rates into welfare and welfare training, and lower transitions rates into employment.

These correlations suggest there is considerable selectivity into the training programs. Furthermore, they show that those who are selected into welfare and JRP training programs appear to be different from those who participate in UI training programs. As a matter of fact, all the correlation coefficients of the last section of the panel pertaining to UI training have the opposite sign to those of welfare and JRP training. Consider, for example, those who have unexpectedly long UI training spells and those who have unexpectedly short welfare or JRP spells. According to the last section of the panel, all these individuals are more likely to move into employment upon exiting their respective spells than average. Yet, the middle section indicates that only those on welfare training or JRP are likely to have long employment spells. Those who were on UI training are more likely to have short employment spells.

That those who are likely to have short JRP or welfare training spells are more likely to experiment long employment spells may be somewhat surprising. In fact, these results are consistent with those of Mealli et al. (1996). Indeed, when stydying the impact of various training schemes, they found that some of those who dropped out early did so because they had good labor market opportunities. Our results suggest the same may apply to the welfare and JRP training programs.

5 Conclusion

The analysis has focused on an examination of the impact of government-sponsored training programs aimed at disadvantaged male youths on their labour market transitions. We have elected to concentrate our attention on this group since they have fared relatively poorly on the labour market over the past decade in Canada by all accounts. The richness of the data at our disposal has allowed us to recreate very detailed individual histories over a relatively long period. As many as seven distinct states on the labour market could be identified in the data.

This study has applied a continuous time duration model to estimate the density of duration times in these seven states, controlling for the endogeneity of an individual's training status. Most previous studies have used survey or administrative data that were less amenable to the kind of analysis performed in this paper. Depending on the nature of the data, complex adjustments to the model were often required to account for potential problems related to stock sampling and initial conditions. Fortunately, we were able to avoid these difficulties by recreating each individual's history as early as age

16, the legal school-leaving age in Canada. Consequently, the initial state can be safely considered exogenous, and the subsequent duration times void of any form of bias.

There is no consensus in the literature concerning the appropriate treatment of unobserved heterogeneity in multi-states multi-episodes duration models. When few states are considered, two-factor loading models with a finite set of points of support have become relatively standard. When the analysis focuses on more states, factor loading models require a large number of parameters to be flexible or become relatively restrictive if a parsimonious specification is used. In this paper we have chosen to investigate the sensitivity of the parameter estimates by comparing a typical non-parametric specification and a series of parametric two-factor loading models. These models implicitly assume that the intensity of transitions are related to the state of destination. We have also estimated a parametric three-factor loading model. The novelty of this specification lies in the fact that the intensities of transitions are related to both to the state of destination and the state of origin.

The estimation of the model yields a number of interesting results. As found in previous studies, unobserved heterogeneity appears to play an important role in determining who selects or gets selected in training programs. On the other hand, the slope and baseline hazard parameter estimates are not very sensitive to the choice of a particular distribution function for the unobserved heterogeneity variables. The two-factor loading models, either parametric or non-parametric, yield essentially the same results as the three-factor loading model. These show that the duration times in any of the seven states considered are sensitive to variations in program parameters such as welfare benefits, policy variables such as the minimum wage rate, and in the economic environment as proxied by the unemployment rate. Nearly all the parameter estimates have the expected sign when statistically significant.

The results pertaining to the impact of the welfare training programs and JRP are similar to those found earlier by Gritz (1993), Bonnal et al. (1997) and Fortin et al. (1999a). In essence, young, poorly educated males who participate in these programs do worse on the labour market than those who do not participate, even after controlling for unobserved heterogeneity. On the other hand, participation in training programs while on unemployment insurance provides them some benefits in the form of increased transitions into employment.

References

Beaudry, P., and D. Green (1997) 'Cohort patterns in canadian earnings: Assessing the role of skill premia in inequality trends.' mimeo, University of British Columbia

Bonnal, L., D. Fougère, and A. Sérandon (1997) 'Evaluating the impact of french employment policies on individual labour market histories.' *The Review of Economics Studies* 64(4), 683–718

Burtless, G. (1995) 'The case for randomized field trials in economic and policy research.' *Journal of Economic Perspective* 9(2), 63–84

Duclos, J.-Y., B. Fortin, G. Lacroix, and H. Roberge (1999) 'The dynamics of welfare participation in Québec.' In *Women and Work,* ed. Powell L. and R. Chaykowsky (The John Deutch Institute)

Fortin, B., and G. Lacroix (1997) 'Welfare benefits, minimum wage rate and the duration of welfare spells: Evidence from a natural experiment in canada.' Working Paper # 9708, Department of Economics, Université Laval

Fortin, B., D. Fougère, and G. Lacroix (1999a) 'Hausse des barèmes et sorties de l'aide sociale: Les résultats d'une expérience naturelle au Canada.' *Revue Économique* pp. 451–463

——— (1999b) 'The impact of government-sponsored training programs on labour market transitions.' Prepared for Human Resources Development Canada

Gouriéroux, C., and A. Monfort (1992) 'Modèles de durée et effets de génération.' *Working Paper # 9125*, CREST, Paris

——— (1996) *Simulation-Based Econometric Methods* Core Lectures (Oxford University Press)

Gourriéroux, C., and A. Monfort (1991) 'Simulation based econometrics in models with heterogeneity.' *Annales d'économie et de statistique* 20(1), 69–107

Gritz, R.M. (1993) 'The impact of training on the frequency and the duration of employment.' *Journal of Econometrics* 57, 21–51

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

Ham, J.C., and S.A. Rea (1987) 'Unemployment insurance and male unemployment duration in canada.' *Journal of Human Resources* pp. 325–353

Heckman, J., and B. Singer (1984) 'A method for minimizing the distributional assumptions in econometric models for duration data.' *Econometrica* pp. 271–320

Heckman, J., and C. Flinn (1983) 'Are unemployment and out of the labor force behaviorally distinct labor force states ?' *Journal of Labor Economics*

Heckman, J.J., and J.A. Smith (1995) 'Assessing the case for social experiments.' *Journal of Economic Perspective* 9(2), 85–110

Heckman, J.J., R.J. LaLonde, and J.A. Smith (1999) 'The economics and econometrics of active labor market programs.' In *Handbook of Labor Economics,* ed. O. Ashenfelter and Eds. D. Card (North-Holland) chapter

Jones, S.R.G., and C.W. Riddell (1999) 'The measurement of unemployment: An empirical approach.' *Econometrica* 67(1), 147–61

Kamionka, T. (1998) 'Simulated maximum likelihood estimation in transition models.' *Econometrics Journal* 1, C129–C153

Laroque, G., and B. Salanié (1993) 'Simulation-based estimation of models with lagged latent variables.' *Journal of Applied Econometrics* 8, S119–S133

Mealli, F., S. Pudney, and J. Thomas (1996) 'Training duration and post-training outcomes: A duration-limited competing risks model.' *Economic Journal* 106(435), 422–433

Van den Berg, G.J. (1997) 'Association measures for durations in bivariate hazard rate models.' *Journal of Econometrics* 79(2), 221–245

Van den Berg, G.J., M. Lindeboom, and G. Ridder (1994) 'Attrition in longitudinal panel data, and the empirical analysis of dynamic labour market behaviour.' *Journal of Applied Econometrics* 9, 421–435

Table 1
Sample Characteristics

	Mean	Std. dev.	Mean	Std. dev.
Non Trainees				
Age in January 1987	18.93	0.57		
Education	9.84	1.03		
Duration of employment spells (weeks)†	19.70	26.12		
Duration of welfare spells (weeks)†	48.54	51.47		
Duration of unemployment spells (weeks)†	40.46	14.27		
Proportion of time employed (%)‡	18.12			
Number of observations	1935			
Trainees				
Age in January 1987	18.92	0.57		
Education	9.72	1.03		
	Before training		*After training*	
Duration of employment spells (weeks)†	17.52	24.58	15.98	16.78
Duration of welfare spells (weeks)†	45.57	47.10	32.77	45.47
Duration of unemployment spells (weeks)†	39.55	14.52	31.89	17.53
Proportion of time employed (%)‡	17.06		19.82	
Number of observations	1133			

† *Calculated from non censored spells.*

‡ *Calculated from mean duration in employment, unemployment, welfare and OLF.*

Table 2
Frequency of Transitions Between States

Destination / *Origin*	Welfare	Welfare Training	JRP	U.I.	U.I. Training	Employment	OLF
Welfare	0	1809	140	88	0	1851	1134
Welfare Training	432	0	67	6	0	438	306
JRP	21	4	0	7	0	192	29
U.I.	374	38	2	292	111	1380	1404
U.I. Training	2	1	0	114	0	16	2
Employment	1002	229	35	2918	41	2004	4662
OLF	2614	235	9	523	2	3815	0

Table 3
Baseline Hazard Functional Forms†

Dest. *Origin*	Welfare	Welfare Training	JRP	U.I.	U.I. Training	Emp.	OLF
Welfare		Exp (1)	Exp (1)	Exp (1)		Exp (3)	Exp (1)
Wel Tr	Log-logis.					Log-logis.	Log-logis.
JRP						Exp (1)	
U.I.	Exp (2)			Exp (2)	Exp (1)	Exp (2)	Exp(2)
U.I. Tr				Exp(1)			
Emp	Log-logis.	Weibull		Log-logis.		Log-logis.	Log-logis.
OLF	Exp (2)	Exp (2)		Exp (2)		Exp (2)	

† "Exp" refers to exponential piecewise constant hazard model. The number of parameters are indicated between parentheses.

Table 4
Parameter Estimates
Three-Factor Loading Model – Weibull Distribution
(Standard errors in parentheses)

	Exits from Welfare						
	State of Destination						
	Welfare	Wel Tr	JRP	UI	UI Tr	Emp	OLF
Baseline:		-14.863	-12.632	-18.401		-3.383	-6.563
		(0.487)	(2.188)	(2.124)		(0.370)	(0.443)
						-3.754	
						(0.373)	
						-4.437	
						(0.375)	
Age		-0.203	-0.376	0.613		0.597	1.043
		(0.162)	(0.589)	(0.657)		(0.156)	(0.183)
Min Wage		21.861	16.971	21.284		-1.339	-2.255
		(1.182)	(5.310)	(4.923)		(0.980)	(1.143)
Unemp Rate		-0.281	-1.993	-1.132		-0.182	0.404
		(0.288)	(0.970)	(0.977)		(0.207)	(0.245)
Benefits		-1.878	0.415	0.205		-1.438	-0.791
		(0.361)	(1.111)	(1.081)		(0.249)	(0.264)
Wel Tr_1		0.228	0.515	0.168		-0.123	-0.326
		(0.139)	(0.348)	(0.417)		(0.117)	(0.173)
Wel Tr_2		0.680	-0.177	-2.020		-0.421	-0.119
		(0.142)	(0.459)	(1.216)		(0.165)	(0.255)
UI Tr_1		0.227	0.101	0.609		0.232	0.562
		(0.243)	(0.751)	(0.740)		(0.200)	(0.234)
UI Tr_2							

	Exits from Unemployment						
	State of Destination						
	Welfare	Wel Tr	JRP	UI	UI Tr	Emp	OLF
Baseline:	-7.659			-8.656	-14.295	-3.440	-5.536
	(1.014)			(0.988)	(1.832)	(0.440)	(0.482)
	-4.468			-4.156		-1.241	-2.525
	(0.988)			(0.973)		(0.445)	(0.482)
Age	-0.682			0.317	-0.597	0.472	-0.297
	(0.358)			(0.407)	(0.614)	(0.191)	(0.187)
Min Wage	-0.245			-0.376	16.074	-2.563	1.813
	(2.632)			(2.791)	(4.458)	(1.429)	(1.447)
Unemp Rate	1.556			1.014	-0.063	-0.440	-0.182
	(0.534)			(0.564)	(0.973)	(0.279)	(0.270)
Benefits	1.715			-1.200	0.073	-0.536	-0.272
	(0.465)			(0.592)	(1.142)	(0.264)	(0.340)
Wel Tr_1	0.556			0.330	-0.337	-0.329	-0.107
	(0.273)			(0.345)	(0.474)	(0.184)	(0.184)
Wel Tr_2							
UI Tr_1	-0.091			0.293	-0.126	0.150	0.060
	(0.514)			(0.509)	(0.537)	(0.266)	(0.262)
UI Tr_2	2.261			1.406		0.727	1.401
	(0.593)			(0.631)		(0.372)	(0.345)

Table 4 (continued)
Parameter Estimates
Three-Factor Loading Model – Weibull Distribution
(Standard errors in parentheses)

	Exits from Employment						
	State of Destination						
	Welfare	Wel Tr	JRP	UI	UI Tr	Emp	OLF
Baseline:	-4.891	-24.613		-6.445		-4.744	-6.332
	(0.376)	(1.710)		(0.091)		0.197)	(0.202)
	1.310	-0.032		0.652		1.162	1.636
	(0.093)	(0.096)		(0.023)		(0.055)	(0.037)
Age	-0.485	0.715		0.692		-0.038	-1.369
	(0.177)	(0.345)		(0.111)		(0.125)	(0.078)
Min Wage	-8.842	35.002		-0.583		-2.145	1.317
	(0.972)	(3.316)		(0.625)		(0.720)	(0.450)
Unemp Rate	0.966	-2.229		-0.865		-0.993	-0.851
	(0.264)	(0.658)		(0.174)		(0.205)	(0.130)
Benefits	1.722	-1.619		-0.041		-0.217	0.319
	(0.221)	(0.776)		(0.148)		(0.170)	(0.105)
Wel Tr_1	1.326	1.311		0.104		-0.273	-0.081
	(0.121)	(0.223)		(0.120)		(0.154)	(0.104)
Wel Tr_2	-0.914	0.587		-0.098		-0.126	-0.688
	(0.190)	(0.215)		(0.158)		(0.218)	(0.157)
UI Tr_1	0.143	-1.070		-0.117		0.004	-0.341
	(0.259)	(0.642)		(0.151)		(0.183)	(0.180)
UI Tr_2				1.580		-0.648	-1.222
				(0.339)		(1.190)	(1.326)
	Exits from OLF						
	State of Destination						
	Welfare	Wel Tr	JRP	UI	UI Tr	Emp	OLF
Baseline:	-7.227	-9.808		-0.815		-0.815	
	(0.265)	(0.533)		(0.224)		(0.224)	
	-8.063	-10.690		-1.139		-1.139	
	(0.259)	(0.532)		(0.223)		(0.223)	
Age	-0.274	-0.061		1.035		-0.242	
	(0.124)	(0.331)		(0.275)		(0.103)	
Min Wage	-0.510	9.572		-2.582		-2.946	
	(0.126)	(2.086)		(0.332)		(0.108)	
Unemp Rate	9.031	-1.347		12.763		3.357	
	(0.675)	(0.576)		(1.615)		(0.638)	
Benefits	0.325	-0.173		-1.436		-0.489	
	(0.185)	(0.991)		(0.360)		(0.144)	
Wel Tr_1	0.075	1.673		-1.404		-0.112	
	(0.097)	(0.214)		(0.386)		(0.122)	
Wel Tr_2	-1.537	0.458		-0.240		-0.728	
	(0.222)	(0.233)		(0.553)		(0.218)	
UI Tr_1	0.032	0.533		-0.722		0.151	
	(0.158)	(0.419)		(0.556)		(0.162)	
UI Tr_2							

Table 5
Heterogeneity Parameters for Selected Models
(Standard errors in parentheses)

	No Hetero-geneity	Non Parametric Model	Two Factor Loading Model	Three Factor Loading Model
			Weibull Distribution	*Weibull Distribution*
d		0.899		
		(0.070)		
c_1		-0.753		
		(0.049)		
c_2		-1.566		
		(0.023)		
b_2		0.223	0.242	-2.169
		(0.128)	(0.342)	(0.830)
b_3		2.757	7.510	7.487
		(1.139)	(2.177)	(3.349)
b_4		1.271	5.133	7.196
		(0.135)	(0.700)	(1.187)
b_5		0.419	-0.875	-1.632
		(1.049)	(2.714)	(5.049)
b_6		1.500	5.408	8.073
		(0.096)	(0.632)	(1.087)
b_7		-0.866	-2.410	-6.078
		(0.068)	(0.273)	(1.099)
b'_1				2.974
				(0.812)
b'_2				3.395
				(1.071)
b'_3				-8.281
				(2.512)
b'_4				1.073
				(0.926)
b'_5				-7.296
				(5.031)
b'_6				1.922
				(0.680)
b'_7				1.296
				(0.459)
λ			7.952	13.008
			(1.069)	(2.269)
γ			0.145	0.100
			(0.056)	(0.059)
Log-Likelihood	-150629.1	-149998.7	-150018.4	-149993.9

Table 6
Correlations Between Heterogeneity Variables
(Standard Errors in Parentheses)

	Welfare	Welf. Tr.	JRP	UI	UI Tr.	Employ.	OLF
	Two-Factor Loading Model: Non-Parametric						
Welfare	1.000	-0.398	0.062	0.607	0.332	0.115	0.384
		(0.154)	(0.045)	(0.243)	(0.142)	(0.289)	(0.156)
Wel. Tr.		1.000	0.891	0.487	0.733	0.865	0.694
			(0.091)	(0.390)	(0.216)	(0.181)	(0.242)
JRP			1.000	0.831	0.962	0.999	0.945
				(0.159)	(0.035)	(0.016)	(0.047)
UI				1.000	0.951	0.859	0.967
					(0.067)	(0.202)	(0.053)
UI Tr.					1.000	0.975	0.998
						(0.067)	(0.002)
Emplo.						1.000	0.962
							(0.085)

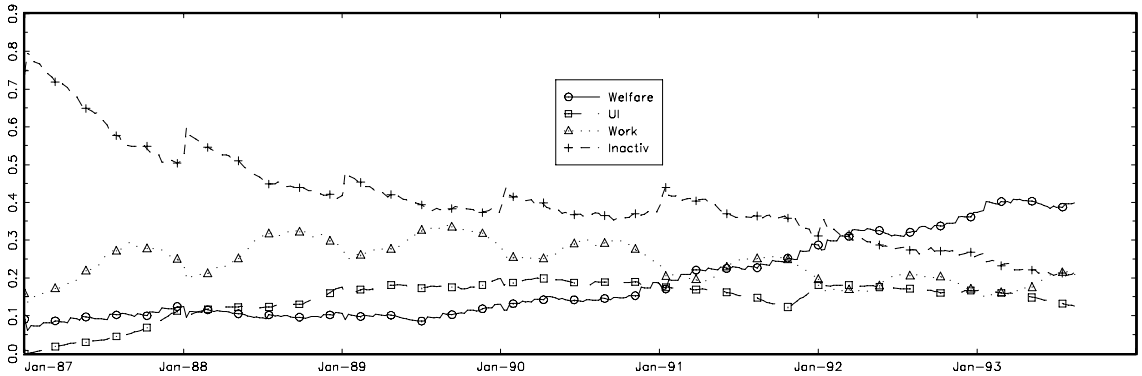
Table 6 (Continued)
Correlations Between Heterogeneity Variables
(Standard Errors in Parentheses)

	Welfare	Welf. Tr.	JRP	UI	UI Tr.	Employ.	OLF
	Two-Factor Loading Model: Weibull Distribution						
Welfare	1.000	-0.443	0.235	0.991	0.982	-0.659	0.983
		(0.108)	(0.315)	(0.005)	(0.005)	(1.157)	(0.004)
Wel. Tr.		1.000	0.767	-0.321	-0.264	0.967	-0.273
			(0.226)	(0.118)	(0.119)	(0.395)	(0.118)
JRP			1.000	0.361	0.416	0.577	0.408
				(0.299)	(0.293)	(1.291)	(0.294)
U.I.				1.000	0.998	-0.553	0.999
					(0.002)	(1.282)	(0.002)
UI Tr.					1.000	-0.502	0.999
						(1.328)	(0.001)
Emplo.						1.000	-0.511
							(1.322)

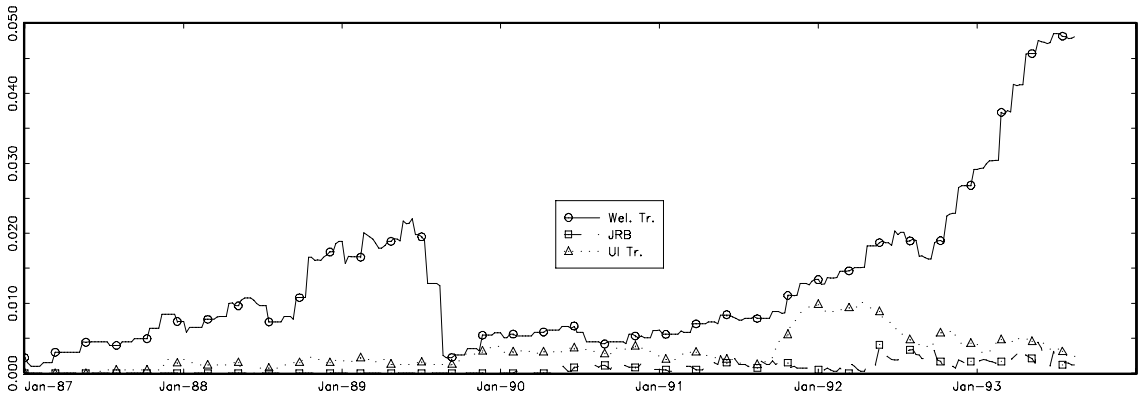
Table 6 (Continued)
Correlations Between Heterogeneity Variables
Three-Factor Loading Model
(Standard Errors in Parentheses)

	Welfare	Welf. Tr.	JRP	UI	UI Tr.	Employ.	OLF
	Correlation Between Destination States						
Welfare	1.000	-0.439	-0.908	0.991	0.990	-0.853	0.992
		(0.104)	(0.061)	(0.008)	(0.003)	(0.720)	(0.002)
Wel. Tr.		1.000	0.775	-0.317	-0.312	0.844	-0.326
			(0.119)	(0.122)	(0.113)	(0.740)	(0.111)
JRP			1.000	-0.845	-0.842	0.993	-0.850
				(0.085)	(0.082)	(0.162)	(0.080)
UI				1.000	0.999	-0.776	0.999
					(0.001)	(0.871)	(0.001)
UI Tr.					1.000	-0.773	0.999
						(0.874)	(0.001)
Emplo.						1.000	-0.782
							(0.859)
	Correlation Between Origin States						
Welfare	1.000	0.948	0.959	-0.993	0.731	-0.991	0.887
		(0.026)	(0.024)	(0.004)	(0.294)	(0.013)	(0.067)
Wel. Tr.		1.000	0.999	-0.903	0.911	-0.896	0.988
			(0.003)	(0.041)	(0.163)	(0.058)	(0.014)
JRP			1.000	-0.918	0.894	-0.912	0.981
				(0.039)	(0.182)	(0.054)	(0.022)
U.I.				1.000	-0.644	0.999	-0.825
					(0.331)	(0.002)	(0.087)
UI Tr.					1.000	-0.632	0.964
						(0.344)	(0.098)
Emplo.						1.000	-0.816
							(0.102)
	Correlation Between Origin-Destination States						
Welfare	1.000	0.948	0.959	-0.993	0.731	-0.991	0.887
		(0.026)	(0.024)	(0.004)	(0.294)	(0.013)	(0.067)
Wel. Tr.	-0.439	-0.416	-0.421	0.436	-0.321	0.435	-0.390
	(0.104)	(0.100)	(0.101)	(0.104)	(0.149)	(0.104)	(0.097)
JRP	-0.908	-0.861	-0.871	0.902	-0.664	0.900	-0.806
	(0.061)	(0.077)	(0.071)	(0.061)	(0.289)	(0.062)	(0.104)
UI	0.991	0.940	0.951	-0.984	0.725	-0.982	0.879
	(0.008)	(0.028)	(0.026)	(0.009)	(0.292)	(0.015)	(0.068)
UI Tr.	0.990	0.939	0.950	-0.983	0.725	-0.981	0.879
	(0.003)	(0.027)	(0.025)	(0.006)	(0.291)	(0.014)	(0.067)
Emplo.	-0.853	-0.808	-0.818	0.847	-0.624	0.845	-0.756
	(0.720)	(0.684)	(0.692)	(0.715)	(0.641)	(0.714)	(0.647)
OLF	0.992	0.941	0.952	-0.985	0.726	-0.983	0.880
	(0.002)	(0.027)	(0.025)	(0.005)	(0.292)	(0.013)	(0.067)

Figure 1
Distribution Across States



Distribution Across Training Programs



Appendix I
Parameter Estimates
Three-Factor Loading Model – Weibull Distribution
(Standard errors in parentheses)

	Exits from Training Programs				
	Exits from Welfare Training			*Exits from JRP*	*Exits from UI Training*
	Welfare	Emp	OLF	Emp	UI
Baseline:	-5.231	-4.221	-5.018	-9.017	1.102
	(0.387)	(0.261)	(0.332)	(2.417)	(2.096)
	0.744	0.232	0.313		
	(0.112)	(0.096)	(0.125)		
Age	0.214	0.365	0.042	0.552	-0.906
	(0.279)	(0.256)	(0.268)	(0.657)	(0.759)
Min Wage	5.319	0.028	-2.857	13.263	-7.681
	(1.745)	(1.751)	(1.983)	(4.174)	(6.215)
Unemp Rate	-3.036	-0.532	-0.062	-1.336	1.562
	(0.469)	(0.511)	(0.553)	(0.930)	(1.138)
Benefits	-3.592	-1.021	0.242	-0.967	0.457
	(0.523)	(0.603)	(0.669)	(1.913)	(1.138)