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

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

Very Preliminary

November 1999

Abstract

The analysis focuses on an examination of 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. As many as seven distinct states on the labour are be identified in the data.

We use 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. 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 far worse on the labour market than those who do not participate. Participation in the Job Re-Entry Program (JRP), a distinct welfare training programs, yields better results in terms of employment. Our estimates clearly indicate that participants in JRP and those in welfare training programs are distinct groups.

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

‡ Human Resources Development Canada. Email: lucie.gilbert@ecn.ulaval.ca

§ CNRS and GRÉMAQ, Université de Toulouse I. Email: kamionka@cict.fr

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

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 are somehow related to the unobservables that affect the outcome of participation. 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 market: employment, unemployment

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

(nonemployment) 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] Many have questioned the appropriateness of focusing of few labour market states (Heckman and Flinn (1983), Jones and Riddell (1999)). It may be even less appropriate to focus on few states when considering the impact of training programs.

This paper investigates the impact of government training programs aimed at poorly educated 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 young Canadian males 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 19 at any time during that period and to have less

[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 (nonemployment), 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 may still qualify.

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) methods.

The remainder of the paper is organized as follows. Section 2 provides a detailed description the data. Section 3.1 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 reasons. These individuals are not included in the sample. Thus the final sample comprises only individuals having 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 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 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 participation. An examination of the table reveals that the two groups are very similar in terms of their observable characteristics. Yet, there are significant differences in their respective labour market experiences. For instance, non-trainees have longer employment and OLF spells than trainees, and have shorter spells in both welfare and unemployment. On the other hand, the average duration of employment, unemployment, welfare and OLF spells decreases significantly following training. On the whole, the average proportion of time spent employed by trainees and non-trainees is remarkably similar.

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 in the files as early as January 1984.[6] 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,

[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 is available only as of January 1987. Consequently, a small proportion of unemployment spells occurring prior to 1987 may be wrongly coded as OLF. Two factors lead us to believe that the proportion of such spells is likely insignificant. First, the large majority of individuals that 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. Similarly, employment 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 will be wrongly coded as OLF.

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 3068 individuals in our sample experienced as many as 31422 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 a 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 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 Out of Labour Force (OLF) state is the complement of all other states. It may include 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 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 50 observations will be considered in the econometric model. This leaves a total of 24 transitions to be modeled explicitly.

[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.

[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 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.

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 (Forthcoming) for details].

The proportion of employed individuals follows a very distinct seasonal pattern with peaks occurring around June-July and troughs around January 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 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 it highest level 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 these programs hardly reaches 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 unattractive. 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 individual responsible for the management of the program 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 ($\tau_e =$ December 1993). Figure 2 depicts a hypothetical sequence

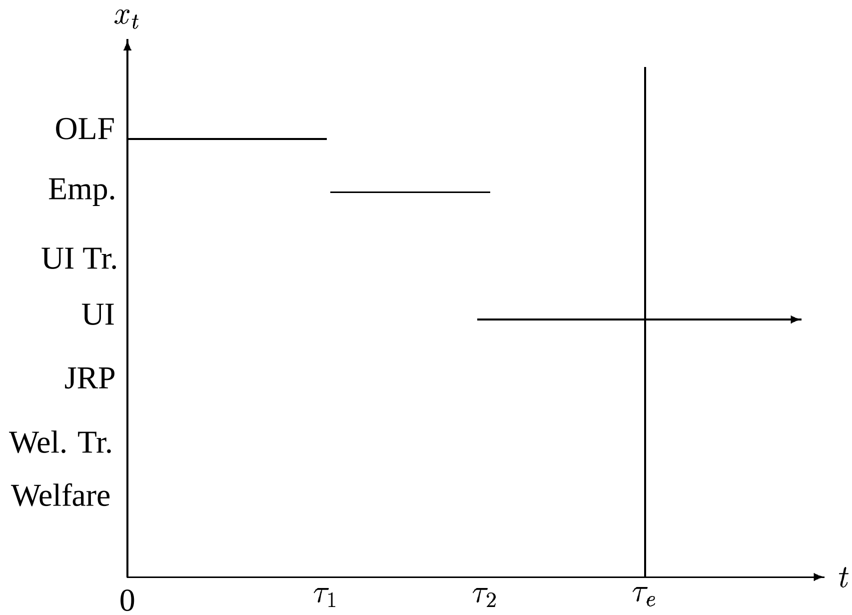

Figure 2: Labour market history of a hypothetical individual.

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 $(1 \leq \ell \leq n+1)$ 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 :

$$r = ((\tau_0, x_{\tau_0}), (u_1, x_{\tau_1}), \ldots, (u_n, x_{\tau_n}), (u_{n+1}, 0)),$$

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

The sequence may be more compactly rewritten as:

$$r = (y_0, y_1, \ldots, y_n, y_{n+1}),$$

where

$$y_\ell = \begin{cases} (\tau_0, x_{\tau_0}), & \text{if } \ell = 0, \\ (u_\ell, x_{u_\ell}), & \text{if } 1 \leq \ell \leq n, \\ (u_{n+1}, 0), & \text{if } \ell = n+1. \end{cases}$$

The initial state, y_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 $r = (y_0, y_1, \ldots, y_n, y_{n+1})$ 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,

$$\ell_v(\theta) = \prod_{\ell=1}^{n+1} f(y_\ell \mid y_0, \ldots, y_{\ell-1}; z; \nu; \theta),$$

where $f(y_\ell \mid y_0, \ldots, y_{\ell-1}; z; \nu; \theta)$ is the conditional density of y_ℓ given y_0, y_1, $\ldots$, $y_{\ell-1}$, z and ν, and $\theta \in \Theta \subset \mathbb{R}^p$ 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

$$l(\theta) = \sum_{j=1}^{J} \prod_{\ell=1}^{n+1} f(y_\ell \mid y_0, \ldots, y_{\ell-1}; z; \nu_j; \theta)\, \pi_j, \tag{1}$$

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

$$l(\theta) = \int_V \prod_{\ell=1}^{n+1} f(y_\ell \mid y_0, \ldots, y_{\ell-1}; z; \nu; \theta)\, g(\nu; \theta)\, d\,\nu, \tag{2}$$

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

Furthermore, we will assume that y_ℓ is independent of $y_0,\ldots,y_{\ell-2}$ given $y_{\ell-1}$, z and ν, in which case

$$f(y_\ell \mid y_0,\ldots,y_{\ell-1};z;\nu;\theta) = f(y_\ell \mid y_{\ell-1};z;\nu;\theta).$$

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 can 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 x_{τ_ℓ}. 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

$$u_\ell = \inf_{k' \neq x_{\tau_{\ell-1}}} u^*_{\ell,k'}.$$

Let $f_j(u \mid y_0,\ldots,y_{\ell-1};z;\nu;\theta)$ 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 $S_j(u \mid y_0,\ldots,y_{\ell-1};z;\nu;\theta)$ be the corresponding survivor function :

$$S_j(u \mid y_0,\ldots,y_{\ell-1};z;\nu;\theta) = \int_u^{+\infty} f_j(s \mid y_0,\ldots,y_{\ell-1};z;\nu;\theta)\, d\, s.$$

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

$$\begin{aligned} f(u,k|y_0,\ldots,y_{\ell-1};z;\nu;\theta) &= f_k(u \mid y_0,\ldots,y_{\ell-1};z;\nu;\theta) \\ &\quad \prod_{\substack{j=1 \\ j\neq k, x_{\tau_{\ell-1}}}}^{K} S_j(u \mid y_0,\ldots,y_{\ell-1};z;\nu;\theta), \\ &= h_k(u|y_0,\ldots,y_{\ell-1};z;\nu;\theta)\, S(u|y_0,\ldots,y_{\ell-1};z;\nu;\theta), \end{aligned}$$

where $h_k(u \mid y_0, \ldots, y_{\ell-1}; z; \nu; \theta)$ is the hazard function associated with the latent duration $u^*_{\ell,k}$ and $S(u \mid y_0, \ldots, y_{\ell-1}; z; \nu; \theta)$ is the survivor function of the duration of the ℓ^{th} spell. Because the latent durations are assumed to be conditionally independent we have

$$S(u \mid y_0, \ldots, y_{\ell-1}; z; \nu; \theta) = \prod_{\substack{j=1 \\ j \neq x_{\tau_{\ell-1}}}}^{K} S_j(u \mid y_0, \ldots, y_{\ell-1}; z; \nu; \theta),$$

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:

$$l_v(\theta) = \prod_{\ell=1}^{n} \prod_{\substack{k=1 \\ k \neq x_{\tau_{\ell-1}}}}^{K} h_k(u \mid y_0, \ldots, y_{\ell-1}; z; \nu; \theta)^{\delta_{\ell,k}} \, S_k(u \mid y_0, \ldots, y_{\ell-1}; z; \nu; \theta),$$

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 double-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

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

where $v_1 \in \{-2, c_2\}$, $v_2 \in \{c_1, c_2\}$, $b_k \in I\!\!R$, $a_k = 1\!\!I[k \geq 2]$ 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 $\log(w_{k'})$ should span the interval $[-1; 1]$.

Moreover, assume that

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

where $c_1, c_2 \in I\!\!R$, $b_k \in I\!\!R$ 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

$$\rho_{k,k'} = \frac{a_k\ a_{k'} \sigma_{v_1}^2 + b_k\ b_{k'}\ \sigma_{v_2}^2}{\sqrt{a_k^2\ \sigma_{v_1}^2 + b_k^2\ \sigma_{v_2}^2}\sqrt{a_{k'}^2\ \sigma_{v_1}^2 + b_{k'}^2\ \sigma_{v_2}^2}}, \tag{3}$$

where $k, k' = 1, \ldots, K$ and $\sigma_{v_j}^2$ is the variance of v_j, j=1,2.

A two-factor loading model with two independent heterogeneity terms with a common **continuous** distribution can also be derived from this specification. Indeed, 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 $b_1 = 1$).

In this version of the model, v_1 and v_2 are assumed to be independently and **identically** distributed. Let $q(v;\theta)$ 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 (3). In the next section, we will present results based on various parametric distribution functions (exponential, weibull, log-normal, normal and student).

The above specifications can be further generalized to a three-factor loading model with a common continuous distribution for the unobserved variables. In this specification, 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.

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

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

In this three-factor loading model, the correlation between $\log(w_k)$ and $\log(w_{k'})$ is

$$\rho_{k,k'} = \frac{a_k \; a_{k'} + b_k \; b_{k'}}{\sqrt{a_k^2 + b_k^2}\sqrt{a_{k'}^2 + b_{k'}^2}}. \tag{4}$$

The correlation between $\log(w'_j)$ and $\log(w'_{j'})$ is

$$\rho_{j,j'} = \frac{a'_j \; a'_{j'} + b'_j \; b'_{j'}}{\sqrt{a'^2_j + b'^2_j}\sqrt{a'^2_{j'} + b'^2_{j'}}}. \tag{5}$$

Finally, the correlation between $\log(w_k)$ and $\log(w'_j)$ is

$$\rho_{k,j} = \frac{b'_j \; b_k}{\sqrt{a'^2_j + b'^2_j}\sqrt{a_k^2 + b_k^2}}, \tag{6}$$

where $j, j', k, k' = 1, \ldots, K$.

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:

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

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

$$h_{j,k}(u \mid y_0,\ldots,y_{\ell-1};z;\nu;\theta) = h^0_{j,k}(u;\theta)\ \varphi(y_0,\ldots,y_{\ell-1};z;\theta)\ \psi(j,k), \tag{7}$$

where φ is a positive function depending on 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

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

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

   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

   $$h^0_{j,k}(u;\theta) = \alpha_{j,k}\,\mathbb{1}[u < u^0_1] + \beta_{j,k}\,\mathbb{1}[u^0_1 \leq u < u^0_2] + \gamma_{j,k}\,\mathbb{1}[u^0_2 < u],$$

   where $\alpha_{j,k}, \beta_{j,k}, \gamma_{j,k} \in \mathbb{R}^+$. 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

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

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

   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 unobserved heterogeneity distribution.

1. *Two-Factor Loading and Discrete Distribution*

   The log likelihood is

$$\log(L(\theta)) = \sum_{i=1}^{N} \log(l_i(\theta)), \tag{8}$$

   where $l_i(\theta)$ is obtained by substituting the sequence $r_i = (y_{0,i}, \ldots, y_{n_i+1,i})$ 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[10]

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

   where $p \in [0; 1]$ is a parameter.

   The log-likelihood is then maximized with respect of θ ($\theta \in \Theta$) in order to obtain the maximum likelihood estimation of the parameters.

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 are independently and identically distributed. Let $q(v; \theta)$ 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 $g(\nu; \theta) = q(v_1; \theta)\, q(v_2; \theta)$. The log-likelihood is given by equation (8) where $l_i(\theta)$ is the contribution to the likelihood of the sequence r_i. 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

$$\hat{l}(\theta) = \frac{1}{H} \sum_{h=1}^{H} \prod_{\ell=1}^{n+1} f(y_\ell \mid y_0, \ldots, y_{\ell-1}; z; v_{1,h}, v_{2,h}; \theta),$$

[10] See section 4.

where $v_{1,h}$ and $v_{2,h}$ are drawn independently according to the p.d.f. $q(v;\theta)$[11]. 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 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 successively 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

$$\hat{l}(\theta) = \frac{1}{H} \sum_{h=1}^{H} \prod_{\ell=1}^{n+1} f(y_\ell \mid y_0, \ldots, y_{\ell-1}; z; v_{1,h}, v_{2,h}, v_{3,h}; \theta),$$

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

4 Estimation Results

This section presents the maximum likelihood estimation results of the hazard function framework outlined in the previous section using administrative data. The estimation of such a complex model is computationally demanding. Also, a number of issues must be addressed before dwelling into the results.

[11] $q(v;\theta)$ is the p.d.f. of one of the distributions we have examined (normal, log-normal, exponential, student and weibull distributions).

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 adopted distribution should allow a number of different shapes of the hazard function so that various combinations of positive and negative duration dependence are possible. Second, the functional form chosen 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, they have been targeted for welfare training programs. Having a relatively homogeneous sample in terms of age and education does not remove the need to control for such variables explicitly. On the other hand, Gritz (1993) has found these variables to have no impact whatsoever on any of the transitions considered in his model. Furthermore, Bonnal et al. (1997) also uses the same strategy to avoid over-parameterizing the likelihood function of their model.

The model includes the following covariates: minimum wage, unemployment rate, welfare benefits, dummy indicators for previous training under either welfare or UI, and the UI basic benefit rate. The basic benefit rate is the proportion of the insurable earnings that is paid as benefits. The rate remained constant at 60% between 1987 and April 1993, when it was reduced to 57%. 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 the spell. While the basic benefit rate only has two distinct values, its identification relies essentially on the fact that spells have different starting dates.

4.3 Parameter Estimates

Table 4 presents the parameter estimates of six different specifications. The table runs over several pages. Each page focuses on the transitions out of a particular state. The first column of the table reports the parameter estimates under the assumption of no unobserved heterogeneity. Under this assumption, the transitions out of each state could be estimated separately as a competing risks model. The specification of the second column is a two-factor loading model based on a finite number of points of support. This specification is typical of those found in the literature. The next three specifications of the table are two-factor loading models that use continuous distributions. The third column uses a log-normal distribution to draw the random components. Columns 4 and 5 both use weibull distributions, but the latter specification also controls for previous training. Finally, the specification of the last column is a tree-factor loading model that uses a weibull distribution to draw the random components and which also controls for past training.[12]

An examination of Table 4 reveals interesting results. Given the large number of parameter estimates, it would be unreasonable to discuss each of them in turn. Instead, we will focus on the main results pertaining to each state considered in the model.

[12]The model was also estimated using normal, student-t, χ^2 and gamma distributions. These results are not reported here for the sake of brevity, but are available on request. The preferred specifications are those based on the weibull distribution for two reasons. First, as shown in Heckman and Singer (1984) (p. 276) 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, the weibull distribution appears to depict similar properties. Second, as in Heckman and Singer (1984), the values of likelihood function based on the weibull distribution are usually larger than those based on other distributions.

4.3.1 Exits from Welfare

As indicated above, the columns of Table 4 are arranged in an increasing order of complexity in terms of unobserved heterogeneity and/or in terms of the number of control variables. The first panel of the table focuses on welfare. Exits to as many as five different states are considered. As a rule, the results are very robust to the type of unobserved heterogeneity considered in the model. The specification of the first column, *i.e.* when unobserved heterogeneity is not accounted for, yields parameter estimates that are significantly different from those of other specifications.

As expected, increases in welfare benefits decreases the exit rates from welfare. The result is statistically significant in transitions towards training, work and OLF states. Increases in the unemployment rate translate into increases in the transitions toward welfare training and OLF, but lower transitions into JRP. This latter result is compatible with the fact that welfare claimants may be motivated to increase their employability when job prospects diminish. Similarly, firms may be less inclined to hire trainees under the JRP program when the unemployment rate rises.

Interestingly, increases in the minimum wage rate increases the transitions towards welfare training, 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 but do 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 basic benefit rate make unemployment, and possibly employment, more attractive alternatives for welfare claimants. The parameter estimates do not support this conjecture since they are either not statistically significant or marginally significant. On the other hand, they are negative and statistically significant with respect to transitions into welfare training and OLF. A priori one would have expected these parameter estimates not to be statistically significant. Consequently, it not clear how to interpret these results.

Columns 4 and 5 are identical specifications except for the inclusion of three dummy indicators for past training in the latter specification. These dummy variables are equal to one whenever a welfare claimant has experienced at least one training program prior to the current welfare spell. Implicitly, it is assumed that the impact of past training spells does not wear off with time nor does it accumulate with repeat uses of training programs.

The inclusion of these dummy variables impacts both the baseline hazard parameters and the slope parameters somewhat. The parameter estimates reveal several interesting effects of training on time spent on welfare. First, past occurrences of JRP increases the likelihood of leaving welfare for welfare training, but decreases the likelihood of entering unemployment or employment. Likewise, past occurrences of welfare training increases the likelihood of transiting from welfare to JRP. In order to transit from welfare to unemployment, an individual must work while claiming benefits. The parameter estimates suggests that these training programs are perceived to some extent as substitutes to regular jobs by welfare claimants.

The last two columns of the table are identical except for the fact that three loading factors are used in the last column instead of two. Using a richer specification for the unobserved heterogeneity components marginally decreases the impact of the dummy training variables.

4.3.2 Exits from Unemployment

The next panel of the table focuses on the transitions out of unemployment. Most parameter estimates that are statistically significant have the expected sign *a priori*. For instance, increases in the minimum wage rate, the unemployment rate and the welfare benefits increase the likelihood of entering welfare upon leaving unemployment. As indicated lower in the table, unemployed individuals experience greater difficulty entering employment when the minimum wage rate and the unemployment rate increase. Presumably, a number of them exhaust their benefits and enter welfare. This is all the more likely if welfare benefits increase as well.

Other results presented in the table indicate that unemployed individuals are more likely to enter a new unemployment spell whenever the unemployment rate increase, but are likely to do so when the welfare benefits increase. Increases in the minimum wage rate increases the likelihood of leaving unemployment for unemployment training. This result is similar to what was found concerning transitions from welfare to welfare training. Finally, the basic benefit rate is found to increase the transitions from unemployment to welfare and to decrease transitions toward employment as expected.

A number of parameter estimates relating to the training dummy variables are statistically significant. They indicate that both past welfare and UI training increase the likelihood of entering welfare upon exiting unemployment. On the other hand, past welfare training decreases the probability of entering employment while past UI training has the opposite effect. This result is 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 y-oung 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. Are 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 exits 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 or unemployment, and to diminish considerably the likelihood of transiting towards another job. Increases in welfare benefits are found to increase the transitions into welfare and to decrease the likelihood of entering welfare training.

The parameter estimate 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. They are thus more likely to turn to welfare, as indicated in the top portion of the panel. Second, the deterioration of the labour market may induce some to hold on to their current jobs longer. The fact that all the parameter estimates are negative, except for welfare, in consistent with this possibility. Consequently, it is very likely that employment spells last longer when the unemployment rate increases.

The training dummy variables in the first two portions of the panel are defined as in previous panels. The next three portions of the panel include four training dummy variables. The first of these, Wel. Tr_1, is equal to one if the individual has experienced welfare or JRP training at any time before the current employment spell. The second dummy variable, Wel. Tr_2, is equal to one if the state just prior to the current employment spell was either welfare or JRP training. The two other dummy variables, *i.e.* UI Tr_1 and UI Tr_2, are similarly defined but relate to UI training. Including these dummy variables enables us to verify the extent to which the training effects taper off with time.

The training variables of the first two portions of the table 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. Having participated in JRP decreases substantially the likelihood of re-entering welfare, but also increases the likelihood of entering welfare training. The last three sections of the panel also reveal interesting results. First, individuals that were in UI training just prior to their current employment spell are much more likely to return to UI upon leaving employment. On the other hand, having gone through UI training programs further in the past has no noticeable effect on this transition. Second, the likelihood of entering the OLF state following employment decreases substantially if the individual experienced either UI or welfare training in the past. The impact is greater when welfare training preceded the employment spell but is not related to the timing of the UI training.

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. Exits from OLF are thus more likely when the benefit rate or the minimum wage rate increase, and less likely when either the unemployment rate or the welfare benefits increase.

The parameter estimates related to the welfare training variables tell a rather disappointing story. Indeed, past participation in these programs increase the likelihood of entering welfare or welfare training anew. On the other hand, past participation in JRP decreases both the likelihood of entering welfare and employment following a spell of inactivity. Incidentally, UI training has no impact on the exit rates out of OLF.

4.3.5 Exits from Training Programs

The penultimate panel of Table 4 reports parameter estimates relating to training programs. Exits from the welfare training programs are sensitive to most exogenous variables of the model. On the other hand, the econometric model does a poorer job at predicting exits from JRP and UI training programs. In these two cases, only the minimum wage rate has any explanatory power, albeit with parameter estimates of the expected sign.

4.3.6 Unobserved Heterogeneity

The last panel of the table reports the value of the likelihood function of each specification as well as the parameter estimates relating to the unobserved heterogeneity. The first specification does not control for unobserved heterogeneity and is thus a special case of all the other specifications. It is strongly rejected on the basis of log-likelihood ratio tests. This is hardly surprising given that the parameter estimates of this specification reported in the previous panels were in some cases strikingly different from the those in which unobserved heterogeneity was controlled for.

The robustness of the parameter estimates with respect to the distribution of the unobserved heterogeneity variables can be investigated by comparing columns 2–4 of the previous panels since these specifications include the same set of exogenous variables. Apart from a few cases, the slope parameters are relatively insensitive to the choice of a particular distribution function. The main differences occur with respect to the baseline hazard parameters. Bonnal et al. (1997) also found their results to be relatively insensitive to the distributional assumptions of the unobserved heterogeneity variables.[13] These results are also consistent with the results of Heckman and Singer (1984) using single durations data.

The specification in column 5 is similar to that in column 4, except for the inclusion of past training variables. A simple likelihood ratio test strongly rejects the model of column 4 in favour of that in column 5. The inclusion of these training dummy variables alters few slope and baseline hazard parameters, the main differences arising with respect to the OLF and employment hazards. Finally, note that the three-factor loading model of column 6 nests the two-loading factor model of column 5.[14] A simple log-likelihood ratio test strongly rejects the two-factor loading model in favour of the three-factor model.

The rejection of the two-factor loading model is relatively surprising given that the slope and baseline hazard parameter estimates of the two specifications are nearly identical. However, introducing a third heterogeneity component considerably alters the loading factor parameters in the rightmost column of the last panel of Table 4. This suggests that the richer specification may be better suited to uncover selection into the different states, if any. In order to investigate this issue, we report in Table 5 the correlation coefficients between the heterogeneity variables that are implicit in each specification along with their standard errors. The first panel is concerned with the non-parametric and the log-normal two-factor loading models. Both specifications yield similar coefficients except for correlations involving UI training. According to these estimates, there is little selectivity into the welfare training programs. Indeed, those who are more likely to experience such programs are no more likely to

[13] In their work, they compare a two-factor loading model with finite points of support with a single-factor loading model that draws heterogeneity terms from an i.i.d. $\mathbb{N}(0,1)$ distribution.

[14] The null assumption is $H_0 : b'_j = 0, j = 1, \ldots, K$.

be employed or unemployed, although they are more likely to be in the OLF state. On the other hand, there appears to be considerable selectivity into JRP. Indeed, the correlation between JRP, employment and unemployment are very large and significant, whereas the correlation between JRP and OLF is large and negative. Such selection may reflect individual preferences but may also reflect selectivity on the part of program administrators.

The next panel of Table 5 reports the correlation coefficients of the two-factor loading models based on the weibull distribution function. The two sections of the panel only differ insofar as the specification of the lower portion includes training dummy variables. Note first that the correlation coefficients are nearly identical to those of the log-normal distribution of the previous panel. The only difference concerns the correlation coefficient between welfare training and OLF which is not statistically different, and which reinforces the idea that there appears to be no selectivity into these programs. Note also that the correlation coefficients that are statistically significant are nearly identical between the two sections of the panel.

The last panel of Table 5 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 (4)–(5), respectively. Hence, the first section is identical to 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. Indeed, those who are more likely to participate in these programs are also less likely to train under JRP and also to find employment. 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 short welfare training spells are also likely to have long JRP spells. Similarly, long welfare training spells are correlated to long OLF spells, whereas to converse applies to JRP 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 more likely to have long welfare spells are less likely to enter welfare training and more likely to enter JRP (row 1). On the other hand, those who are likely to have long welfare training spells are also likely to transit through welfare and to return to welfare training, but much less likely to transit through JRP (row 2). Similarly, row 3 indicates that individuals who are likely to have long JRP spells are likely to return eventually to JRP but are also much less likely to return to either welfare or welfare training in the future.

The correlations presented in Table 5 clearly indicate that participation in either welfare training programs or JRP is strongly correlated to unobservable characteristics. The estimates also suggest that participants in JRP are systematically different from those who participate in welfare training programs. Finally, we find very little evidence of selection bias into the UI 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 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 have assumed that the intensity of transitions were 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 yield essentially the same results. These show that the duration times in any of the seven states considered are sensitive to variations in program parameters such as welfare benefits, minimum wage rate, UI basic benefit rate and the unemployment rate. Nearly all the parameter estimates have the expected sign when statistically significant.

The results pertaining to the impact of the training programs 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 welfare training programs do far worse on the labour market than those who do not participate. Participation in the Job Re-Entry Program, a distinct welfare training programs, yields better results in terms of employment. Participants also appear to return much less to welfare than those who use standard welfare training programs. Finally, participation in a UI related training program yields positive results in terms of employment, but it also raises the likelihood of experiencing both welfare and welfare spells.

Our estimates clearly indicate that participation in the training programs is definitely related to unobservable characteristics. More interestingly, they stress that participants in JRP and those in welfare training programs are distinct groups. To the extent we have adequately accounted for such selectivity, the measured impacts of the various training programs should not be contaminated by selectivity biases.

There are a number of extensions to the empirical results presented above that warrant further investigation. Certainly, it would be preferable eventually to control explicitly for age and schooling levels. Although we have selected our sample so as to limit the extent of variations in these variables, not controlling explicitly for them may amplify the impact of unobserved heterogeneity. Secondly, the exogenous variables are all measured at the beginning of each spell, as is customary in the literature. A more satisfactory strategy presumably would be to make these variables time dependent, given that some spells are relatively lengthy. Finally, it would be interesting to compare these results with those based on a similar sample of women. It is generally acknowledge that young female cohorts have performed relatively better than comparable male cohorts in Canada over our sample period. The results reported in this paper suggest that these extensions may well be worth pursuing.

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 (Forthcoming) '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 beheviorally 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 (Je ne le sais pas) 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.
Individual without training				
Age when entering the sample	19.92	1.84		
Education	9.84	1.03		
Duration of employment episodes (weeks)†	26.15	30.47		
Duration of welfare episodes (weeks)†	35.63	34.43		
Duration of unemployment episodes (weeks)†	39.90	11.99		
Duration of OLF episodes (weeks)†	42.99	50.95		
Proportion of time employed (weeks)‡	0.18			
Number of observations	1165			
Individual with training				
Age when entering the sample	19.77	(1.95)		
Education	9.72	(1.03)		
	Before training		*After training*	
Duration of employment episodes (weeks)†	24.26	37.13	17.44	18.25
Duration of welfare episodes (weeks)†	54.90	54.86	35.14	51.09
Duration of unemployment episodes (weeks)†	41.25	14.41	35.73	16.83
Duration of OLF episodes (weeks)†	30.78	38.86	18.66	21.16
Proportion of time employed (weeks)‡	0.17		0.16	
Number of observations	1903			

† *Calculated from non censored episodes.*
‡ *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 (3)	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 – Exits from Welfare

	No Het	Non-Para	Log-Normal	Weibull 2 Factors	Weibull 2 Factors	Weibull 3 Factors
Welfare to Welfare Training						
Baseline[1]	12.951†	11.769†	9.916†	9.736†	7.437†	7.511†
Replacement	-41.844†	-34.680†	-34.591†	-34.500†	-30.710†	-30.722†
Minimum Wage	11.507†	10.464†	10.555†	10.624†	10.132†	10.127†
Unemp Rate	1.837†	0.641†	0.659†	0.659†	0.855†	0.858†
Welfare Ben.	-0.267	-1.202†	-1.165†	-1.169†	-1.058†	-1.046†
Wel. Tr.[3]					0.133	0.126
JRP[4]					0.486†	0.486†
U.I. Tr.[5]					0.123	0.127
Welfare to JRP						
Baseline[1]	-16.915‡	-12.068	-20.602†	-21.019†	23.330†	-23.004†
Replacement	-2.948	10.486	11.451	11.010	14.679	14.800
Minimum Wage	24.425†	20.061†	20.504†	20.511†	20.177†	19.689†
Unemp Rate	0.114	-2.402†	-2.477†	-2.465†	-2.301†	-2.304†
Welfare Ben.	1.389	0.081	0.093	0.046	0.115	0.095
Wel. Tr.[3]					0.562‡	0.561‡
JRP[4]					-0.011	-0.039
U.I. Tr.[5]					0.112	0.109
Welfare to Unemployment						
Baseline[1]	2.275	0.160	-5.840	-6.134	-3.910	-3.511
Replacement	-30.577†	-16.447	-15.796	-16.035	-19.417‡	-18.808‡
Minimum Wage	12.777	16.201†	16.575†	16.565†	16.483†	15.791†
Unemp Rate	1.538†	-0.399	-0.446	-0.451	-0.568	-0.547
Welfare Ben.	1.579‡	0.672	0.682	0.626	0.571	0.456
Wel. Tr.[3]					0.232	0.210
JRP[4]					-2.164‡	-2.139‡
U.I. Tr.[5]					0.594	0.531
Welfare to Work						
Baseline[2]	-6.268†	5.172†	-1.084	-1.426	-0.781	-0.902
	-6.786†	4.757†	-1.501	-1.836	-1.194	-1.250
	-7.623†	4.039†	-2.253	-2.580	-1.941	-1.910
Replacement	-0.201	-2.715	-2.268	-2.469	-3.464	-2.834
Minimum Wage	6.019†	-0.608	-0.470	-0.447	-0.345	-0.266
Unemp Rate	-0.322†	-0.106	-0.127	-0.126	-0.209	-0.219
Welfare Ben.	-1.107†	-1.214†	-1.204†	-1.231†	-1.272†	-1.342†
Wel. Tr.[3]					0.027	0.020
JRP[4]					-0.454†	-0.436†
U.I. Tr.[5]					0.314	0.273
Welfare to OLF						
Baseline[1]	-2.676	-1.632	-1.661	-1.858	-1.136	-1.025
Replacement	-3.865	-6.066‡	-6.159‡	-5.847	-6.859‡	-6.669‡
Minimum Wage	-1.026	-1.394	-1.424	-1.268	-1.130	-1.055
Unemp Rate	0.396†	0.500†	0.484†	0.508†	0.382	0.425‡
Welfare Ben.	-0.847†	-0.621†	-0.522†	-0.590†	-0.632†	-0.684†
Wel. Tr.[3]					-0.209	-0.242
JRP[4]					-0.276	-0.233
U.I. Tr.[5]					0.708†	0.597†

1 *Exponential hazard.*
2 *Exponential hazard – splines.*
3 *Dummy indicator for any previous welfare training.*
4 *Dummy indicator for any previous JRP.*
5 *Dummy indicator for any previous U.I. training.*

Table 4 (Continued)
Parameter Estimates – Exits from Unemployment

	No Het	Non-Para	Log-Normal	Weibull 2 Factors	Weibull 2 Factors	Weibull 3 Factors
Unemployment to Welfare						
Baseline[2]	-20.617†	-28.243†	-29.821†	-29.848†	-29.184†	-29.120†
	-17.591†	-25.197†	-26.775†	-26.805†	-26.039†	-25.969†
Replacement	22.900†	31.891†	31.726†	31.632†	31.541†	31.616†
Minimum Wage	-5.551	5.312†	5.125†	5.129†	3.266	3.236
Unemp Rate	1.695†	0.907‡	0.921‡	0.921‡	1.084†	1.093†
Welfare Ben.	1.620†	1.289†	1.313†	1.314†	1.403†	1.401†
Wel. Tr.[3]					0.590†	0.583†
JRP[4]					-0.045	-0.046
U.I. Tr.[5]					2.161†	2.171†
Unemployment to Unemployment						
Baseline[2]	-20.463†	-1.743	-7.615	-7.912	-7.524	-7.390
	-16.030†	2.695	-3.168	-3.473	-3.016	-2.884
Replacement	3.966	-1.913	-1.024	-1.382	-1.547	-1.442
Minimum Wage	22.232†	1.669	1.197	1.309	0.371	0.192
Unemp Rate	0.901†	0.947‡	0.981‡	0.973‡	1.042‡	1.065‡
Welfare Ben.	-1.245†	-1.235†	-1.179†	-1.203†	-1.161†	-1.153†
Wel. Tr.[3]					0.315	0.338
JRP[4]					0.305	0.301
U.I. Tr.[5]					1.477†	1.473†
Unemployment to Unemployment Training						
Baseline[1]	-2.325	-6.534	-10.807	-10.988	-10.621	-10.482
Replacement	-15.630‡	-4.039	-4.201	-4.171	-5.389	-5.474
Minimum Wage	4.588	11.050†	11.140†	11.207†	11.962†	11.915†
Unemp Rate	1.758	0.329	0.311	0.299	0.202	0.210
Welfare Ben.	0.934	0.160	0.192	0.203	0.190	0.203
Wel. Tr.[3]					-0.398	-0.408
U.I. Tr.[5]					-0.144	-0.146
Unemployment to Work						
Baseline[2]	-5.570‡	10.517†	4.550‡	3.985	4.577‡	4.730‡
	-3.400	12.708†	6.753†	6.181†	6.788†	6.938†
Replacement	-3.153	-10.825†	-10.570†	-10.526†	-11.610†	-11.669†
Minimum Wage	8.237†	-3.261†	-3.719†	-3.532†	-3.364†	-3.478†
Unemp Rate	-0.801†	-0.266	-0.225	-0.235	-0.277	-0.269
Welfare Ben.	-0.683†	-0.337	-0.299	-0.326	-0.342	-0.330
Wel. Tr.[3]					-0.385†	-0.368†
JRP [4]					0.150	0.156
U.I. Tr.[5]					0.774†	0.784†
Unemployment to OLF						
Baseline[2]	-13.698†	-10.051†	-9.952†	-10.042†	-9.426†	-9.302†
	-10.723†	-7.069†	-6.986†	-7.079†	-6.418†	-6.293†
Replacement	7.977†	6.497‡	6.104‡	6.375‡	5.533	5.353
Minimum Wage	8.652†	2.008	2.358	2.402	2.015	1.984
Unemp Rate	-0.054	-0.236	-0.295	-0.305	-0.271	-0.252
Welfare Ben.	-0.284	-0.390	-0.406	-0.396	-0.378	-0.361
Wel. Tr.					-0.098	-0.127
JRP					0.045	0.061
U.I. Tr.					1.399†	1.353†

1 *Exponential hazard.*
2 *Exponential hazard – splines.*
3 *Dummy indicator for any previous welfare training.*
4 *Dummy indicator for any previous JRP.*
5 *Dummy indicator for any previous U.I. training.*

Table 4 (Continued)
Parameter Estimates – Exits from Employment

	No Het	Non-Para	Log-Normal	Weibull 2 Factors	Weibull 2 Factors	Weibull 3 Factors
Work to Welfare						
Baseline[1]	-7.102†	-7.122†	-7.093†	-7.094†	-7.066†	-7.057†
	1.825†	1.829†	1.823†	1.823†	1.815†	1.812†
Replacement	-10.487†	-7.783†	-10.196†	-10.328†	-10.415†	-10.336†
Minimum Wage	0.165	1.045	0.230	0.207	0.172	0.175
Unemp Rate	1.022†	0.939†	1.002†	1.008†	1.083†	1.087†
Welfare Ben.	1.078†	0.977†	1.050†	1.050†	1.026†	1.038†
Wel. Tr.[2]					0.698†	0.692†
JRP[3]					-1.167†	-1.166†
U.I. Tr.[4]					-0.313	-0.306
Work to Welfare Training						
Baseline[1]	-29.322†	7.320	5.564	5.571	2.880	3.647
	3.352†	-0.012	-0.019	-0.020	-0.004	-0.006
Replacement	-35.936†	-43.508†	-43.520†	-43.638†	-34.009†	-35.355†
Minimum Wage	25.433†	26.226†	26.392†	26.342†	18.898†	18.242†
Unemp Rate	-0.107	-0.610	-0.623	-0.614	-0.282	-0.161
Welfare Ben.	-1.249†	-1.389‡	-1.380‡	-1.374‡	-0.818	-0.629
Wel. Tr.[2]					1.281†	1.267†
JRP[3]					0.510†	0.500†
U.I. Tr.[4]					-1.067‡	-0.996
Work to Unemployment						
Baseline[1]	-6.340†	-6.453†	-6.492†	-6.468†	-6.474†	-6.445†
	0.687†	0.678†	0.671†	0.671†	0.673†	0.672†
Replacement	-2.411†	7.297†	-0.444	-1.208†	-1.309†	-1.171†
Minimum Wage	3.196†	5.624†	3.772†	3.612†	3.734†	3.757†
Unemp Rate	-0.805†	-0.959†	-0.934†	-0.896†	-0.898†	-0.910†
Welfare Ben.	-0.086	-0.436†	-0.228	-0.251	-0.265‡	-0.261‡
Wel. Tr1[2]					0.101	0.111
Wel. Tr2[5]					-0.216	-0.214
U.I. Tr1[3]					-0.130	-0.125
U.I. Tr2[6]					1.524†	1.484†
Work to Work						
Baseline[1]	-4.758†	-4.722†	-4.683†	-4.687†	-4.698†	-4.713†
	1.187†	1.155†	1.136†	1.143†	1.147†	1.162†
Replacement	-0.583	9.910†	1.320†	0.557	0.094	0.123
Minimum Wage	-3.339†	-1.263‡	-2.959†	-3.021†	-2.452†	-2.467†
Unemp Rate	-0.953†	-1.019†	-1.046†	-1.036†	-1.030†	-1.036†
Welfare Ben.	0.013	-0.285‡	-0.122	-0.142	-0.180	-0.159
Wel. Tr1[2]					-0.248	-0.240
Wel. Tr2[5]					-0.122	-0.123
U.I. Tr1[3]					0.014	0.027
U.I. Tr2[6]					-0.667	-0.716
Work to OLF						
Baseline[1]	-6.350‡	-6.283†	-6.332†	-6.336†	-6.356†	-6.310†
	1.650†	1.626†	1.641†	1.642†	1.645†	1.628†
Replacement	-0.060	-0.267	-0.480‡	-0.319	-1.205†	-1.246†
Minimum Wage	-3.623†	-3.903†	-3.774†	-3.811†	-2.606†	-2.717†
Unemp Rate	-0.974†	-0.916†	-0.911†	-0.908†	-0.915†	-0.877†
Welfare Ben.	0.268†	0.321†	0.317†	0.333†	0.260†	0.295†
Wel. Tr1[2]					-0.282†	-0.286†
Wel. Tr2[5]					-0.572†	-0.586†
U.I. Tr1[3]					-0.488†	-0.453†
U.I. Tr2[6]					-1.233	-1.237

[1] *Log-logistic.*
[2] *Dummy indicator for any previous welfare training.*
[3] *Dummy indicator for any previous JRP.*
[4] *Dummy indicator for any previous U.I. training.*

Table 4 (Continued)
Parameter Estimates – Exits from OLF

	No Het	Non-Para	Log-Normal	Weibull 2 Factors	Weibull 2 Factors	Weibull 3 Factors
OLF to Welfare						
Baseline[2]	9.050†	-21.852†	-23.272†	-23.247†	-21.271†	-21.201†
	8.014†	-22.698†	-24.168†	-24.150†	-22.170†	-22.082†
Replacement	-6.865†	24.985†	24.469†	24.301†	21.054†	21.196†
Minimum Wage	-22.989†	10.913†	10.664†	10.630†	10.589†	10.576†
Unemp Rate	0.384†	-0.425†	-0.378†	-0.376†	-0.386†	-0.396†
Welfare Ben.	1.759†	-0.002	0.039	0.049	0.038	0.032
Wel. Tr.[3]					0.261†	0.251†
JRP[4]					-1.236†	-1.229†
U.I. Tr.[5]					0.106	0.101
OLF to Welfare Training						
Baseline[2]	17.926†	-6.003†	-10.899†	-10.984†	-8.832†	-8.868†
	16.876†	-6.818†	-11.706†	-11.791†	-9.634†	-9.662†
Replacement	-64.337†	-29.309†	-25.939†	-25.960†	-25.370†	-26.066†
Minimum Wage	23.722†	34.225†	36.886†	36.923†	30.886†	31.237†
Unemp Rate	1.865†	0.547	0.351	0.341	0.940	0.961
Welfare Ben.	1.757†	-1.772‡	-1.824†	-1.819†	-1.769‡	-1.681‡
Wel. Tr.[3]					0.710†	0.711†
JRP[4]					-0.036	-0.087
U.I. Tr.[5]					0.105	0.201
OLF to Unemployment						
Baseline[2]	17.926†	-6.003†	-10.899†	-10.984†	-8.832†	-8.868†
	16.876†	-6.818†	-11.706†	-11.791†	-9.634†	-9.662†
Replacement	-18.873†	10.249†	9.571†	8.979†	4.909	5.430
Minimum Wage	-24.418†	8.757†	8.562†	8.549†	9.187†	9.175†
Unemp Rate	-1.791†	-2.838†	-2.892†	-2.919†	-2.926†	-2.984†
Welfare Ben.	1.047†	-0.309	-0.321	-0.326	-0.351	-0.359
Wel. Tr.[3]					-0.913†	-0.906†
JRP[4]					-0.046	-0.043
U.I. Tr.[5]					-0.380	-0.375
OLF to Work						
Baseline[2]	11.804†	-2.367	-8.492†	-8.596†	-6.903†	-6.977†
	11.148†	-2.701‡	-8.830†	-8.923†	-7.233†	-7.313†
Replacement	-6.325†	12.069†	12.465†	11.915†	9.093†	9.477†
Minimum Wage	-22.641†	3.718†	3.725†	3.742†	3.719†	3.664†
Unemp Rate	-1.950†	-2.861†	-2.896†	-2.934†	-2.938†	-2.957†
Welfare Ben.	0.098	-0.554†	-0.587†	-0.593†	-0.596†	-0.585†
Wel. Tr.[3]					-0.008	-0.007
JRP[4]					-0.612†	-0.618†
U.I. Tr.[5]					0.130	0.151

[1] *Exponential hazard.*
[2] *Exponential hazard – splines.*
[3] *Dummy indicator for any previous welfare training.*
[4] *Dummy indicator for any previous JRP.*
[5] *Dummy indicator for any previous U.I. training.*

Table 4 (Continued)
Parameter Estimates – Exits from Training

	No Het	Non-Para	Log-Normal	Weibull 2 Factors	Weibull 2 Factors	Weibull 3 Factors
Welfare Training to Welfare						
Baseline[1]	-4.927†	-4.997†	-4.951†	-4.949†	-4.950†	-5.036†
	0.508†	0.504†	0.506†	0.506†	0.506†	0.498†
Replacement	14.105†	16.989†	14.470†	14.324†	14.326†	15.040†
Minimum Wage	-16.050†	-14.842	-15.840†	-15.877†	-15.893†	-15.543†
Unemp Rate	-0.892†	-1.027†	-0.939†	-0.923†	-0.917†	-0.989†
Welfare Ben.	0.325	0.131	0.287	0.277	0.277	0.158
Welfare Training to Work						
Baseline[1]	-3.971†	-4.087†	-4.107†	-4.127†	-4.123†	-4.230†
	0.277†	0.277†	0.270†	0.268†	0.268†	0.262†
Replacement	-3.450†	5.666†	-1.437	-1.959	-2.100	-1.484
Minimum Wage	3.294†	7.772†	3.949‡	3.702‡	3.769‡	3.901‡
Unemp Rate	-0.957†	-1.096†	-0.884‡	-0.864‡	-0.857‡	-0.814
Welfare Ben.	-0.619	-1.233†	-1.019‡	-1.063‡	-1.060‡	-1.234†
Welfare Training to OLF						
Baseline[1]	-5.187†	-5.135†	-5.140†	-5.124†	-5.124†	-5.151†
	0.403†	0.405†	0.401†	0.403†	0.403†	0.404†
Replacement	-16.086†	-16.959†	-16.775†	-16.747†	-16.672†	-16.234†
Minimum Wage	14.358†	14.417†	13.960†	13.989†	13.957†	14.333†
Unemp Rate	-0.483	-0.515	-0.430	-0.430	-0.439	-0.486
Welfare Ben.	-0.219	-0.012	-0.014	0.024	0.021	-0.108
JRP to Work						
Baseline[2]	3.612	5.668	-3.143	-2.889	-3.037	-4.420
Replacement	-14.331†	-8.461	-5.173	-5.844	-5.767	-5.134
Minimum Wage	6.240	9.809	11.792	11.488	11.655	11.941‡
Unemp Rate	-0.276	-0.877	-1.024	-1.035	-1.047	-1.104
Welfare Ben.	-0.581	-1.467	-1.438	-1.731	-1.731	-0.670
Unemployment Training to Unemployment						
Baseline[2]	5.363	7.502†	1.457	0.920	1.233	0.822
Minimum Wage	-20.820†	-11.370†	-9.906†	-9.984†	-10.577†	-10.712†
Unemp Rate	0.480	1.571‡	1.452	1.525	1.559	1.544
Welfare Ben.	-0.084	0.369	0.157	0.231	0.213	0.506

[1] *Log-logistic.*
[2] *Exponential hazard.*

Table 4 (Continued)
Heterogeneity Parameters

	No Het	Non-Para	Log-Normal	Weibull 2 Factors	Weibull 2 Factors	Weibull 3 Factors
Probability		$0.614^{\dagger}$				
ω_1		$-1.866^{\dagger}$				
ω_2		$-2.364^{\dagger}$				
b_2		-0.116	-0.221	-0.365	-0.184	$-2.980^{\dagger}$
b_3		$3.757^{\dagger}$	$7.781^{\dagger}$	$8.304^{\dagger}$	$8.209^{\dagger}$	$8.173^{\dagger}$
b_4		$2.337^{\dagger}$	$5.455^{\dagger}$	$6.201^{\dagger}$	$6.037^{\dagger}$	$8.879^{\dagger}$
b_5		1.323	0.447	0.244	-1.091	-1.691
b_6		$2.585^{\dagger}$	$5.493^{\dagger}$	$6.432^{\dagger}$	$6.168^{\dagger}$	$7.722^{\dagger}$
b_7		$-1.180^{\dagger}$	$-2.501^{\dagger}$	$-2.938^{\dagger}$	$-2.876^{\dagger}$	$-4.902^{\dagger}$
b'_1						$2.768^{\dagger}$
b'_2						$3.093^{\dagger}$
b'_3						$-7.131^{\dagger}$
b'_4						0.333
b'_5						$-8.563^{\dagger}$
b'_6						-0.437
b'_7						$0.738^{\dagger}$
μ			$-1.629^{\dagger}$			
σ			$-0.566^{\dagger}$			
λ				$9.493^{\dagger}$	$9.057^{\dagger}$	$14.685^{\dagger}$
γ				$0.163^{\dagger}$	$0.147^{\dagger}$	$0.216^{\dagger}$
Log-likelihood	-150668.6	-149774.2	-149847.8	-149818.9	-149474.8	149422.6

†*Statistically significant at 5%*
‡*Statistically significant at 10%*

Table 5
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	-0.157	0.982	0.954	0.875	0.962	-0.850
		(0.195)	(0.010)	(0.010)	(0.236)	(0.007)	(0.023)
Wel. Tr.		1.000	0.035	0.145	0.340	0.118	0.653
			(0.200)	(0.197)	(0.495)	(0.196)	(0.150)
JRP			1.000	0.994	0.952	0.997	-0.734
				(0.006)	(0.151)	(0.004)	(0.051)
UI				1.000	0.980	1.000	-0.654
					(0.098)	(0.001)	(0.052)
UI Tr.					1.000	0.974	-0.489
						(0.111)	(0.430)
Emplo.						1.000	-0.675
							(0.048)
	TWO-FACTOR LOADING MODEL – LOG-NORMAL DISTRIBUTION						
Welfare	1.000	-0.216	0.992	0.984	0.408	0.984	-0.929
		(0.263)	(0.004)	(0.004)	(1.680)	(0.004)	(0.015)
Wel. Tr.		1.000	-0.089	-0.036	0.803	-0.037	0.563
			(0.271)	(0.272)	(1.108)	(0.272)	(0.221)
JRP			1.000	0.999	0.521	0.999	-0.874
				(0.002)	(1.571)	(0.002)	(0.029)
U.I.				1.000	0.566	1.000	-0.846
					(1.516)	(0.000)	(0.032)
UI Tr.					1.000	0.565	-0.040
						(1.518)	(1.841)
Emplo.						1.000	-0.847
							(0.031)

Table 5 (Continued)
Correlations Between Heterogeneity Variables
(Standard Errors in Parentheses)

	Welfare	Welf. Tr.	JRP	UI	UI Tr.	Employ.	OLF
	TWO-FACTOR LOADING MODEL WEIBULL DISTRIBUTION - NO DUMMY INDICATORS						
Welfare	1.000	-0.343	0.993	0.987	0.237	0.988	-0.947
		(0.266)	(0.004)	(0.004)	(2.413)	(0.003)	(0.012)
Wel. Tr.		1.000	-0.228	-0.189	0.832	-0.195	0.627
			(0.278)	(0.281)	(1.390)	(0.280)	(0.218)
JRP			1.000	0.999	0.351	0.999	-0.901
				(0.001)	(2.325)	(0.001)	(0.026)
U.I.				1.000	0.388	1.000	-0.883
					(2.288)		(0.027)
UI Tr.					1.000	0.383	0.089
						(2.294)	(2.475)
Emplo.						1.000	-0.886
							(0.026)
	TWO-FACTOR LOADING MODEL WEIBULL DISTRIBUTION - WITH DUMMY INDICATORS						
Welfare	1.000	-0.181	0.993	0.987	-0.737	0.987	-0.945
		(0.337)	(0.004)	(0.004)	(0.943)	(0.003)	(0.013)
Wel. Tr.		1.000	-0.061	-0.018	0.798	-0.021	0.494
			(0.341)	(0.344)	(0.872)	(0.344)	(0.296)
JRP			1.000	0.999	-0.650	0.999	-0.898
				(0.002)	(1.062)	(0.001)	(0.027)
U.I.				1.000	-0.617	1.000	-0.878
					(1.098)	(0.000)	(0.028)
UI Tr.					1.000	-0.619	0.918
						(1.096)	(0.552)
Emplo.						1.000	-0.880
							(0.027)

Table 5 (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	-0.948	0.993	0.994	-0.861	0.992	-0.980
		(0.026)	(0.006)	(0.002)	(0.568)	(0.002)	(0.007)
Wel. Tr.		1.000	-0.902	-0.906	0.978	-0.899	0.992
			(0.042)	(0.038)	(0.232)	(0.040)	(0.007)
JRP			1.000	0.998	-0.792	0.998	-0.948
				(0.001)	(0.682)	(0.000)	(0.020)
UI				1.000	-0.798	0.999	-0.951
					(0.672)	(0.001)	(0.014)
UI Tr.					1.000	-0.788	0.945
						(0.687)	(0.363)
Emplo.						1.000	-0.946
							(0.015)
	CORRELATION BETWEEN ORIGIN STATES						
Welfare	1.000	0.951	-0.990	0.316	-0.993	-0.401	0.594
		(0.027)	(0.006)	(0.756)	(0.006)	(0.369)	(0.197)
Wel. Tr.		1.000	-0.900	0.592	-0.909	-0.099	0.812
			(0.046)	(0.622)	(0.045)	(0.374)	(0.123)
JRP			1.000	-0.181	1.000	0.524	-0.476
				(0.784)	(0.002)	(0.346)	(0.223)
U.I.				1.000	-0.204	0.743	0.951
					(0.782)	(0.467)	(0.225)
UI Tr.					1.000	0.504	-0.496
						(0.350)	(0.221)
Emplo.						1.000	0.500
							(0.310)
	CORRELATION BETWEEN ORIGIN-DESTINATION STATES						
Welfare	1.000	0.951	-0.990	0.316	-0.993	-0.401	0.594
		(0.027)	(0.006)	(0.756)	(0.006)	(0.369)	(0.197)
Wel. Tr.	-0.948	-0.902	0.939	-0.299	0.942	0.380	-0.563
	(0.026)	(0.044)	(0.028)	(0.720)	(0.028)	(0.345)	(0.197)
JRP	0.993	0.944	-0.983	0.313	-0.986	-0.398	0.589
	(0.006)	(0.028)	(0.009)	(0.751)	(0.009)	(0.367)	(0.196)
UI	0.994	0.946	-0.984	0.314	-0.987	-0.398	0.590
	(0.002)	(0.028)	(0.007)	(0.751)	(0.007)	(0.367)	(0.196)
UI Tr.	-0.861	-0.819	0.852	-0.272	0.855	0.345	-0.511
	(0.568)	(0.544)	(0.563)	(0.725)	(0.565)	(0.367)	(0.395)
Emplo.	0.992	0.944	-0.982	0.313	-0.985	-0.397	0.589
	(0.002)	(0.028)	(0.007)	(0.750)	(0.007)	(0.366)	(0.196)
OLF	-0.980	-0.932	0.970	-0.309	0.973	0.392	-0.582
	(0.007)	(0.031)	(0.010)	(0.742)	(0.010)	(0.360)	(0.196)

Figure 1
Distribution Across States

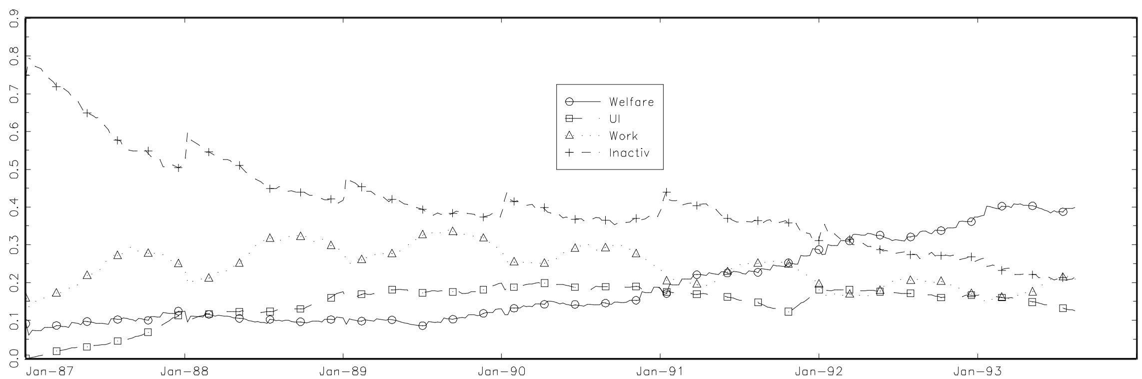

Figure 2
Distribution Across Training Programs