Faculty of Economics, Business Administration and Econometrics
Department of Information Systems

Serie Research Memoranda

Attrition in Panel Data and the Estimation of Dynamic Labor Market Models

Gerard J. van den Berg
Maarten Lindeboom

Research Memorandum 1994-22

12 juli 1994

vrije Universiteit *amsterdam*

ATTRITION IN PANEL DATA
AND THE ESTIMATION OF DYNAMIC LABOR MARKET MODELS

Gerard J. van den Berg*
Maarten Lindeboom**

Abstract

In the empirical analysis of labor market transition models, it is generally assumed that the stochastic processes underlying labor market behavior and the behavior concerning participation in a panel survey are independent. As a consequence, spells that are incomplete due to attrition are treated as spells that are subject to independent right-censoring. Nevertheless, panel survey participants who have a relatively high probability of making a transition may also have a higher probability of dropping out of the panel. In that case the empirical hazard rates underestimate the corresponding transition rates. In this paper we analyze the relation between the durations spent in labor market states and the duration of panel survey participation, by explicitly modeling and estimating both stochastic processes. We use multi-state multi-spell models which allow for stochastically related unobserved determinants.

Date: July 3, 1994.
*Dept. of Econometrics, Free University Amsterdam, De Boelelaan 1105,
1081 HV Amsterdam, The Netherlands, Tinbergen Institute, and INSEE-CREST.
E-mail: vandenberg@sara.nl.
**Dept. of Economics, Faculty of Law, Leiden University, P.O. Box 9521,
2300 RA Leiden, The Netherlands, and Tinbergen Institute.
E-mail: jfaeml@ruljur.leidenuniv.nl.
JEL classification: C41, C33, J64.
Keywords: attrition, panel data, duration models, unemployment, employment, survey participation.
We would like to thank the participants at the PSID Conference on Nonresponse in Washington D.C. (February 1994) for their comments. Robert Moffitt in particular gave many helpful suggestions. The data were provided by The Netherlands Organization for Strategic Labor Market Research (OSA). The research of Gerard J. van den Berg has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences.

1. Introduction.

Models for individual unemployment and employment durations are usually estimated with longitudinal data from panel surveys in which individuals are interviewed a number of times in a certain period. In such empirical analyses it is generally assumed that the stochastic processes underlying labor market behavior and the behavior concerning participation in the panel survey are independent. If this assumption is correct, then attrition from the panel before a duration is completed can be considered as independent right-censoring of that duration variable.

However, it seems plausible that panel survey participants who have a relatively high probability of making a transition on the labor market (like moving from unemployment to employment) also have a higher probability of dropping out of the panel. For example, unemployed individuals may move to another town to work in a job they found, and the agency running the survey may have trouble following them. Also, individuals who spend much time looking for opportunities to make transitions on the labor market may have less time to participate in surveys. In such cases the commonly used procedure to estimate models for the duration spent in a particular state of the labor market (say unemployment) underestimates the rate at which individuals leave that state. In general, if attrition is informative on the occurrence of a transition on the labor market then this should be taken into account when estimating the model. By investigating the former we may thus infer to what extent it is hazardous to estimate particular transitions in the traditional way. Moreover, on the basis of the results, the agency running the survey may want to put more effort in following people who are in a state preceding such a transition.

In this paper we examine whether there is a relation between the durations spent in particular labor market states (notably unemployment and employment) on the one hand, and the duration of panel survey participation on the other. In particular, we will estimate models for the joint distribution of these variables. This means that we have to explicitly model the distribution of survey participation duration and its relation to the distributions of the durations spent in particular states of the labor market. In accordance to the literature on duration analysis we model the distribution of survey participation duration by specifying its hazard rate. This hazard rate is the exit rate out of the panel, and it can be interpreted as the rate at which contact between participating individuals and interviewers is lost. Note that the duration of survey participation is treated as an absolutely continuous random variable. Of course, its realizations can only be observed to lie between two consecutive waves of the panel.

In the next few paragraphs we outline the model framework of the present paper. We use a Mixed Multi–State Semi–Markov Process to describe labor market behavior. This means that the durations spent in different labor market states and the transitions between those states are all modeled simultaneously. The transition rate from one state to another is allowed to depend on the elapsed duration of the spell spent in the former state. In addition, the transition rates are allowed to depend on observed and unobserved explanatory variables. Similarly, the exit rate out of the panel is allowed to depend on the elapsed duration spent in the panel as well as on observed and unobserved explanatory variables. The latter type of variables are also known as unobserved heterogeneity terms or frailties.

There are several ways to model the dependence between a duration spent in a particular labor market state and the duration of survey participation. Here, the emphasis will be on flexible models which allow for such dependence by way of stochastically related unobserved determinants of the durations. Such an approach is in line with the popular modeling setup for sample selection introduced by Heckman (1979). Observation is selective because attrition is informative on the unobserved determinants of the durations spent in particular labor market states. There is a similarity with failure time models in which a duration under continuous monitoring (until failure or right–censoring) is related to the right–censoring time by way of a joint unobserved determinant. (Apparently, Link (1989) was the first to propose such a model and apply it; see also Andersen, Borgan, Gill & Keiding (1993).)

There is a large applied literature in which duration variables are allowed to depend on each other by way of their unobserved determinants (for examples in labor econometrics, see e.g. Flinn & Heckman (1982), Butler, Anderson & Burkhauser (1989) and Coleman (1990)). Some of the models deal with consecutive durations, others with durations occurring simultaneously. In our model, we have multiple consecutive durations in labor market states as well as a duration of survey participation taking place at the same time. A particular advantage of these reduced form models is that they do not a priori restrict the sign of the dependencies if a sufficiently flexible class of distributions is chosen for the unobserved determinants (see Lindeboom & Van den Berg (1994)). Thus, such models can mimic other types of dependencies between the durations. As we will show, the specification we use in this paper for the multivariate distribution of the unobserved heterogeneity terms is more general than typically used in the literature.

One of the major aims of this paper is to offer a comprehensive framework for analyzing the relation between attrition and dynamic labor market behavior. We argue that the use of a multi–state labor market model, in conjunction with multi–spell data, has many advantages. Multi–spell data (i.e. data containing multiple observations per individual of particular types of spells) make identification less dependent on arbitrary functional form restrictions, and they facilitate the estimation of the distributions of the unobserved heterogeneity terms. Multi–state models express the inflow rate into one state in terms of the exit rates out of other states. Moreover, by allowing the exit rate out of the panel to be dependent on the prevailing labor market state, it can be inferred whether attrition is concentrated among individuals in a particular state. Stasny (1988) investigates the latter issue by estimating categorical models for period–to–period gross flows between different labor market states, allowing the probability of nonresponse in a certain period to be dependent on the states occupied in that period and/or in neighboring periods. Her models cannot be used to infer whether occurrence of attrition contains information on the durations spent in labor market states.

Van den Berg, Lindeboom & Ridder (1994) also analyze the implications of attrition in panel data for the empirical analysis of duration models for individual labor market behavior. However, they restrict attention to single–spell duration models for labor market behavior (like unemployment duration models to be estimated with data containing one spell per respondent). The data used in the empirical application were from a sample of the stock of individuals in a certain state. Because of all this, a number of additional assumptions had to be made, and the results may rely heavily on certain functional form restrictions (we will discuss this in detail below). It should be noted that the papers in progress by Lillard & Panis (1994) and Tzeng & Mare (1994) also contain empirical analyses of duration models in the presence of informative attrition, with data containing multiple spells.

As an empirical application of the framework developed in the present paper, we estimate a model distinguishing two labor market states (employment and unemployment), using panel data from The Netherlands. These data are based on interviews between 1985 and 1990 and contain multiple labor market spells for most respondents. We use flexible specifications for the determinants of the hazard rates and the multivariate distributions of the unobserved heterogeneity terms. By comparing the results for different model versions we test the hypothesis of independence of attrition and labor market behavior.

The outline of the paper is as follows. In Section 2 we set up the model. Section 3 deals with the empirical implementation of it. Section 4 contains the empirical application. In Subsection 4.1 we discuss the data. Simple explorative empirical investigations and simple tests on informative attrition are reported in Subsection 4.2. The estimation results for the general model are presented in Subsection 4.3. We check in a number of ways whether the results are sensitive with respect to the model specification. Section 5 concludes.

2. The joint distribution of durations spent in labor market states and the duration of panel survey participation.

2.1. *Labor market behavior.*

We are interested in estimating the distributions of the durations t_j spent by individuals in particular labor market states j as well as the transition rates from one labor market state to another. For ease of exposition, and with the application in Section 4 in mind, we assume that there are only two labor market states, unemployment (u) and employment (e). We assume that all individual differences in the joint distribution of t_u and t_e can be characterized by observed characteristics x and unobserved characteristics v_u and v_e, with x and v_j independent for each j. Conditional on x, v_u and v_e, the variables t_u and t_e are independent. To explain individual differences in t_j, the variable v_i ($i \neq j$) does not give information that is not available in v_j.

The hazard rates of t_u and t_e, given x, v_u and v_e, are of the Mixed Proportional Hazard (MPH) type,

$$\begin{aligned} \theta_u(t_u|v_u,x) &= \lambda_u(t_u).\,v_u.\exp(\beta_u{'}x) \\ \theta_e(t_e|v_e,x) &= \lambda_e(t_e).\,v_e.\exp(\beta_e{'}x) \end{aligned} \tag{2.1}$$

in which some elements of β_u and β_e may be set to zero. The density function $f_j(t_j|v_j,x)$ of $t_j|v_j,x$ can be written as

$$f_j(t_j|v_j,x) = \theta_j(t_j|v_j,x)\ .\exp\left[-\int_0^{t_j} \theta_j(t|v_j,x)\ dt\ \right]$$

(see Lancaster (1990) for a theoretical exposition on MPH models, and Devine & Kiefer (1991) for a survey of the empirical literature in the context of labor market research). The empirical analysis is conditional on x, but unconditional on the unobserved terms v_u and v_e. We assume that x is not time-varying. Note that $t_u|x$ and $t_e|x$ are independent if and only if v_u and v_e are independent.

Individuals experience alternating spells of unemployment and employment. In the empirical application, we observe multiple spells of employment and/or unemployment for some individuals. In general, let a subscript k denote spells of a given type for individual i. We assume in obvious notation that $v_{ijk} \equiv v_{ij}$ and $x_{ik} \equiv x_i$, so the variables v_j and x are fixed across spells for a given individual. Given x, v_u and v_e we then have a Continuous-Time Two-State Semi-Markov Process for dynamic labor market behavior. Since v_u and v_e are unobserved, we observe a mixture of this process with respect to v_u and v_e.

It is well-known that the MPH model for a single type of duration is nonparametrically identified from single-spell data (for a survey, see Lancaster (1990)). However, identification depends crucially on the assumptions underlying the MPH framework. Also, it is generally believed that in practice estimation is next to impossible without (semi-)parametric assumptions, and that the results may depend heavily on these assumptions. As may be clear intuitively, the presence of multi-spell data greatly facilitates the empirical analysis, in particular if the unobserved heterogeneity term is fixed across spells. This is also true for the two-state model, even if the unobserved heterogeneity terms associated with the separate durations are mutually dependent (see Heckman & Honoré (1989) and Honoré (1993) for a number of results).

2.2. *Behavior towards panel survey participation, and its relation to labor market behavior.*

Let z be the length of the period that a randomly chosen individual participates in the panel. We assume that all individual differences in the distribution of z can be characterized by observed characteristics x and unobserved characteristics w, with x and w independent. The hazard of z conditional on x and w is denoted by $\zeta(z|w,x)$ and is assumed to be of the MPH type,

$$(2.2) \qquad \zeta(z|w,x) = \lambda_z(z).w.\exp(\beta_z'x)$$

in which some elements of β_z may be set to zero. If $\lambda_z(z)$ is increasing (decreasing) then, for a given individual, the exit rate out of the panel increases (decreases) as the duration of survey participation proceeds.

There is an extensive literature on conditions under which right–censoring does not affect consistency of the usual estimators in failure time models, when the duration is continuously monitored until failure or right–censoring. Basically, sufficient for this is that the information that an individual is under observation just before t does not improve the prediction of failure at time t (see Andersen, Borgan, Gill & Keiding (1993) and the references therein). However, the relevance of these results for the present context should not be overstated. In panel surveys, individuals are only being interviewed once in a while. This implies that after a transition it may take some time until the next interview, and in case of attrition it is only known that the actual censoring takes place in an interval. If, for example, it is common that respondents decide to leave the panel shortly after a transition, then attrition is informative on labor market behavior and should be taken into account.

As noted in the introduction, the present paper focuses on models in which attrition and labor market behavior are allowed to be related by way of their unobserved determinants. Van den Berg, Lindeboom & Ridder (1994) show that such models are able to mimic other types of dependencies.

We assume that $t_j|v_j,w,x$ and $z|v_j,w,x$ are independent. To explain individual differences in t_j (in z), the variable w (the variable v_j) does not give information that is not available in v_j (in w). As a consequence, $t_j|x$ and $z|x$ are independent if and only if v_j and w are independent. In case of independence of v_j and w, we would have an ordinary duration model for t_j in which attrition can be treated as independent right–censoring.

However, if v_j and w are dependent, then inference on the distribution of t_j has to be based on the joint distribution of $t_j,z|x$. For example, suppose v_j is positively related to w, and suppose that the first wave of the panel consists of a sample of the inflow into labor market state j. Individuals who leave the panel early have on average larger w than individuals who leave the panel later. The former group of individuals will therefore also have on average larger v_j than the latter group. This in turn implies on average larger exit rates out of state j for the former group. Suppose one wishes to estimate $\theta_j(t_j|v_j,x)$ on the basis of data on the first spell in state j. If attrition is treated as independent right–censoring, then the rate at which individuals are observed to leave state j at duration t_j is assumed to equal

$$(2.3) \qquad \theta_j(t_j|x) = \lambda_j(t_j).\exp(\beta_j'x).\mathrm{E}(v_j|\geq t_j,x)$$

in which $\mathrm{E}(v_j|\geq t_j,x)$ denotes the mean of the distribution of v_j conditional on the duration exceeding t_j (see e.g. Lancaster (1990)). However, in our example, the fact that we observe the individual at t_j implies that the rate at which individuals are observed to leave state j at duration t_j is smaller than in (2.3). The distribution of v_j for the group under observation at duration t_j is located left of the distribution of v_j among all individuals (i.e., those under observation and those who left the panel) who are unemployed for t_j units of time. So, if this is ignored then the hazard $\theta_j(t_j|v_j,x)$ will be underestimated.

Note that if v_j and w are unrelated then the event $z \geq t$ may be informative on x, but this has no implications since the analysis is conditional on x.

Let f be a generic symbol for a density of duration variables. The indices and arguments of the density will make clear which variables are considered. Analogously, let g be a generic symbol for a heterogeneity density and let h denote densities of the duration of survey participation. The density $f_{u,e,z}(t_u,t_e,z|x)$ can be expressed as

$$(2.4) \qquad f_{u,e,z}(t_u,t_e,z|x) = \int_{v_u} \int_{v_e} \int_{w} f_u(t_u|v_u,x) \; .f_e(t_e|v_e,x)$$

$$.h(z|w,x) \; .g_{u,e,w}(v_u,v_e,w) \; dw \; dv_e \; dv_u$$

in which the densities on the r.h.s. contain the parameters to be estimated. (Note that $f_j(t_j|v_j,x)$ and $h(z|w,x)$ can be expressed in terms of $\theta_j(t_j|v_j,x)$ and $\zeta(z|w,x)$, respectively, and that $g_{u,e,w}(v_u,v_e,w)$ shows whether there are dependencies of the types we are interested in.) In panel data, we observe versions of t_u, t_e, and z (possibly multiple, possibly censored), and we observe x. The individual likelihood contributions are therefore similar to the expression above.

Suppose we analyze the relation between t_u and z separately from the relation between t_e and z, and suppose we use one duration t_j per respondent (i.e. single-spell data on t_j). (This is basically the approach followed by Van den Berg, Lindeboom & Ridder (1994)). Then the model resembles a dependent MPH competing risks model in which we observe $\min(t_j,z)$, $I(t_j<z)$ and x. Heckman and Honoré (1989) show that then the whole model, including the joint distribution of v_j and w, is nonparametrically identified. However, identification is crucially dependent on the MPH specifications for the exit rates. Also, estimates may be sensitive to parametric assumptions. In the present analysis we use multi-spell data on durations spent in labor market states. As argued in Subsection 2.1, this helps enormously in obtaining good estimates of the distribution of the durations spent in labor market states.

Intuitively, it seems clear that this also helps to estimate the dependence between the durations spent in labor market states and the duration of panel survey participation. For example, we now use information on behavior after the first jump that a respondent makes from one labor market state to another. (See Gottschalk & Moffitt (1992), Lillard & Panis (1994) and Tzeng & Mare (1994) for similar statements on the advantages of multi-spell data for identification of endogenous attrition. Gottschalk & Moffitt (1992) contains an extensive and informative survey of identification issues in the presence of attrition, for a wide range of models.)

The dataset we use contains a variable characterizing for each respondent the identity of the interviewer responsible for dealing with this respondent. One might argue that this variable can be used to design a natural experiment, in which case identification is facilitated. We do not pursue this in our empirical application, for two reasons. First, the number of different interviewers is relatively large. Secondly, and more importantly, the agency running the survey we use assigned the best interviewers to the areas in which the least cooperative respondents were expected to live. Thus, the quality of the interviewer can be expected to be related to unobserved characteristics of the respondent.

It should be noted in advance that, in the data we use, the time intervals between successive interviews are not equal over time, nor are they across individuals. This can be expected to help estimating λ_z (see equation (2.2)). Also, the data we use contain multiple spells of panel survey participation for some individuals (see below), which obviously may be of help as well.

The data used in Van den Berg, Lindeboom & Ridder (1994) as well as the data used in the present paper are from surveys in which the first wave is based on a random sample of the population. As a result, the first wave samples labor market spells in progress. Under certain assumptions, the elapsed duration in the spell at the moment of the first interview contains additional information on the parameters of interest (see the next section for details).

We conclude this subsection by discussing two extensions of the model. First, the level of the exit rate out of the panel may differ between the employed and the unemployed (Stasny (1988) for example finds it is larger for the latter). We may therefore, in addition to the fixed explanatory variables x in $\zeta(z|w,x)$, include a time-varying explanatory dummy variable $\delta(z)$ representing the labor market state prevailing at z.

(2.5) $$\zeta(z|w,x,\delta(z)) = \lambda_z(z).w.\exp(\beta_z'x + \beta_l.\delta(z))$$

Note that this creates a relation between labor market behavior and the exit rate out of the panel even if w is uncorrelated with v_u and v_e.

In one-state one-spell analyses like Van den Berg, Lindeboom & Ridder (1994) it has to be assumed that $\beta_l \equiv 0$, for the following reason. Suppose the individual is observed to leave the panel between the m^{th} and $(m+1)^{th}$ interview. Then it cannot be ruled out that he leaves labor market state j between the m^{th} interview and the moment he leaves the panel. Thus, the likelihood in the model for durations t_j and z depends on the exit rate out of the panel at points of time at which the individual is in labor market state i with $i \neq j$. It is assumed that the latter rate is equal to the rate at points of time at which he is in state j. This indicates another advantage of the multi-state framework. It should however be noted that, to obtain a manageable likelihood function in the latter framework, we also make a simplifying assumption on labor market behavior between the m^{th} interview and the moment the individual leaves the panel (see Subsection 3.2). However, that assumption is much weaker than the assumption above.

The second extension concerns multiple spells of z. Sometimes agencies running a panel survey try to locate individuals who left the panel and/or try to convince them to rejoin the panel. In that case multiple spells of panel survey participation may be observed. The modeling of multiple spells of z is analogous to the modeling of multiple spells of t_j (see Subsection 2.1). Information on additional spells of z helps to identify the elements of $\zeta(z|w,x)$. Alternatively, this information can be used to test the predictive power of the model.

The model may be closed by incorporating the distribution of the duration of not participating in the survey. Because of the lack of information on such spells and the lack of information on what happens on the labor market during such spells, this is not pursued here. In case multiple spells of z are possible, we therefore assume that individuals who left the panel after the m^{th} interview do not return before the $(m+1)^{th}$ interview, and, upon returning, always start their z spells at the date of an interview.

3. Empirical implementation.

3.1. *A random sample.*

In this subsection we examine the empirical implementation of the model in cases in which the first wave of the panel is a random sample of the population.

Since the labor market spells at the date n_1 of the first interview are in progress, we have to deal with the state of the labor market process at the start of the observation period (the so-called initial conditions). We make the following assumption,

Assumption. The labor market process is in equilibrium.

This means that the distributions of the states occupied by individuals at n_1 and the elapsed and residual durations of the corresponding spells can be expressed in terms of the model parameters. Below we derive these distributions by applying results for equilibrium Semi-Markov Processes.

An alternative approach to deal with initial conditions would be to use an ad-hoc specification for the distributions of events directly related to the spell ongoing at n_1, and only relate the distribution of subsequent events to the model parameters (see Flinn & Heckman (1982), and Gritz (1993) for a recent application). Such an approach is potentially more flexible. Here we do not follow it for two reasons. First, the number of observed spells after the spell which is ongoing at n_1 is relatively small. Note that probably the most distinguishing characteristic of European labor markets in comparison to the American labor market is that transitions are much less frequent, so durations are much longer (see e.g. Layard, Nickell & Jackman (1991). Secondly, numerous empirical studies based on the panel data we use have confirmed the assumption above (see e.g. Lindeboom & Theeuwes (1991) and Van den Berg (1992)).

The joint distribution of the observed endogenous variables can be constructed by successive conditioning. We start drawing from the joint distribution $g_{u,e,w,x}(v_u,v_e,w,x)$. From Subsections 2.1 and 2.2, this equals $g_{u,e,w}(v_u,v_e,w).g_x(x)$. Now consider the probability of being in labor market state j at a particular point in time, conditional on v_u, v_e, w and x. Because of the equilibrium assumption, this equals $\mathrm{E}(t_j|v_j,x)/(\mathrm{E}(t_u|v_u,x)+\mathrm{E}(t_e|v_e,x))$ (see e.g. Lancaster (1990)), in which $\mathrm{E}(t_j|v_j,x)$ can be expressed in terms of $\theta_j(t_j|v_j,x)$. Note that this does not depend on w. Let d=1 if the individual is unemployed at n_1, and d=0 otherwise. It follows that the distribution of d conditional on v_u, v_e, w and x equals (with some abuse of notation on the l.h.s.),

$$(3.1) \qquad f_d(d|v_u,v_e,x) = \frac{[\mathrm{E}(t_u|v_u,x)]^d.[\mathrm{E}(t_e|v_e,x)]^{1-d}}{\mathrm{E}(t_u|v_u,x) + \mathrm{E}(t_e|v_e,x)}$$

Now consider the spell in state j which is ongoing at n_1. Let p and r denote the elapsed duration (i.e. the time between inflow into j and n_1) and the residual duration (the time between n_1 and exit out of j), respectively. Given the equilibrium assumption, the joint density of p and r conditional on presence in state j at n_1 and conditional on v_u, v_e, w and x equals

$$(3.2) \qquad f_{p,r}(p,r|d,v_u,v_e,x) = \left[\frac{f_u(p+r|v_u,x)}{\mathrm{E}(t_u|v_u,x)} \right]^d \cdot \left[\frac{f_e(p+r|v_e,x)}{\mathrm{E}(t_e|v_e,x)} \right]^{1-d}$$

(see e.g. Lancaster (1990)). Note that for d=1 (d=0) this density does not depend on v_e (on v_u). The joint distribution of d, p and r conditional on v_u, v_e, w and x follows from multiplication of (3.1) and (3.2). This density can also be easily obtained by successive conditioning starting at the date n_1-p rather than the date n_1 (see Chesher & Lancaster (1983)).

The labor market durations that are realized after the realization of r follow the distributions $f_j(t_j|v_j,x)$ and are mutually independent conditional on v_u, v_e, w and x. So, labor market behavior produces a sequence of endogenous variables $\{d,p,r,t_1,t_2,t_3,...\}$ with $t_1,t_2,t_3,...$ being alternating employment and unemployment durations, and t_1 being an employment duration if and only if d=1. It is straightforward to add information (if available) on spells realized before p.

For simplicity, let us for the moment abstract from the two model generalizations mentioned at the end of Section 2. Then, conditional on v_u, v_e, w and x, the duration of survey participation has density $h(z|w,x)$ and is independent of the labor market durations. If equation (2.5) holds then the density $h(z|w,x)$ can be specified conditional on the labor market process. In any case, the joint distribution of all durations conditional on x is simply obtained by integration w.r.t. $g_{u,e,w}(v_u,v_e,w)$, like in equation (2.4). Note that the only parameters appearing are the parameters of $\theta_j(t_j|v_j,x)$, $\zeta(z|w,x)$ and $g_{u,e,w}(v_u,v_e,w)$.

Before turning to the construction of the likelihood function, we make a remark on the empirical implementation in case of a one-state model when data are used on the labor market spell ongoing at n_1. In that case the first wave provides a sample of the stock of individuals in state j. Such samples are selective. The unobserved heterogeneity distribution in the sample differs from that in the population. Assumptions on the shape of certain functions in the model are needed for tractable empirical inference (see Van den Berg, Lindeboom & Ridder (1991)). In particular, ad-hoc assumptions are needed concerning the inflow into state j, whereas in the present context the inflow into state j is modeled simultaneously with the outflow. In sum, the assumptions needed are stronger than the equilibrium assumption above.

3.2. *The likelihood function.*

The likelihood function is based on the densities derived in Section 2 and Subsection 3.1. Before going into detail, note that the observation period is always finite. Consequently, observation can end for one of two reasons, assuming that there is no return to the panel. Either the individual drops out of the panel before the last or latest interview (Case I), or the individual still participates at the last or latest interview so it is not known what has happened afterwards (Case II). In the latter case there is genuine independent right-censoring of the durations ongoing at the last or latest interview. Similarly, p may be independent right-censored because the retrospective information in the survey only dates back from n_1 to a date n_0 with $-\infty<n_0<n_1$.

Suppose the survey consists of M waves. Let n_m denote the date of the m^{th} interview (m=1,..M). From now on we normalize calendar time by taking n_1=0. In Case I we observe that the individual drops out of the panel between the m^{th} and $(m+1)^{th}$ interview (m=1,..M-1). So, there is some $m\in\{1,..M-1\}$ such that $z\in<n_m,n_{m+1}>$. Also, observation of labor market spells ends at n_m. In Case II $z>n_M$ and the last observed labor market spell is right-censored at n_M.

Consider Case I. Let $\mathcal{LM}$ denote the joint density of the observations on labor market behavior, including the hypothetical observation of the whole duration ongoing at n_m, conditional on v_u, v_e, w and x. It is useful to write $\mathcal{LM}$ as an explicit function of the duration t_j of the last observed labor market spell. We write $\mathcal{LM}$ as $\mathcal{LM}(..,t_j|v_e,v_u,x)$. Further, we denote the starting date of the last observed spell by k.

Let (2.2) describe $\zeta(z|w,x)$. Then $\zeta(z|w,x)$ does not depend on the labor market state at hand. So, if the realization of t_j occurs between n_m and n_{m+1} (i.e., $n_m-k<t_j<n_{m+1}-k$) and if this realization occurs before the realization of z (i.e., $t_j<z-k$) then the value of $\zeta(z|w,x)$ between t_j+k and n_{m+1} is the same as the value of $\zeta(z|w,x)$ between n_m and t_j+k. Consequently, in Case I the likelihood can be written as

$$(3.3)\qquad \int_{t_j=n_m-k}^{\infty} \int_{z=n_m}^{n_{m+1}} \int_{v_u} \int_{v_e} \int_{w} \mathcal{LM}(..,t_j|v_e,v_u,x)\ .h(z|w,x)$$

$$\cdot g_{u,e,w}(v_u,v_e,w)\ \mathrm{d}w\ \mathrm{d}v_e\ \mathrm{d}v_u\ \mathrm{d}z\ \mathrm{d}t_j$$

which, of course, in turn can be completely expressed in terms of $\theta_u(t_u|v_u,x)$,

$\theta_e(t_e|v_e,x)$ and $\zeta(z|w,x)$ on $z\in[0,n_{m+1}>$. Analogously, in Case II we obtain

$$(3.4)\qquad \int_{t_j=n_M-k}^{\infty} \int_{z=n_M}^{\infty} \int_{v_u} \int_{v_e} \int_{w} \mathcal{L}M(..,t_j|v_e,v_u,x)\ .h(z|w,x)$$

$$\cdot g_{u,e,w}(v_u,v_e,w)\ \mathrm{d}w\ \mathrm{d}v_e\ \mathrm{d}v_u\ \mathrm{d}z\ \mathrm{d}t_j$$

which depends on $\theta_u(t_u|v_u,x)$, $\theta_e(t_e|v_e,x)$ and $\zeta(z|w,x)$ on $z\in[0,n_M>$.

If equation (2.5) holds then (3.3) and (3.4) have to be modified. Now the value of the exit rate out of the panel given w and x does depend on the labor market state at hand. We therefore turn from using the density of z to using the exit rate $\zeta(z|w,x,\delta(z))$ in the expressions of the likelihood. More importantly, in Case I, the likelihood depends on the value of $\zeta(z|w,x,\delta(z))$ for $z\in<n_m,n_{m+1}>$, which now in turn depends on labor market behavior between n_m and n_{m+1}, which is unobserved. Any number of labor market transitions can occur between n_m and n_{m+1}. In order to facilitate the analysis, we derive the likelihood contribution of z as if at most one labor market transition can occur between n_m and z (see e.g. Coleman (1990) for the correct expressions for the general case). Then we must distinguish whether the unobserved realization of t_j occurs before or after the realization of z. In the latter case the direct analogue of equation (3.3) applies. In the former case ($n_m<t_j+k<z<n_{m+1}$) the value of $\zeta(z|w,x,\delta(z))$ between n_m and t_j+k differs a factor $\exp(b_l.\delta(n_m))$ from the value between t_j+k and the actual realization of z (see equation (2.5)). In sum, we replace (3.3) by

$$(3.5)\qquad \int_{z=n_m}^{n_{m+1}} \int_{v_u} \int_{v_e} \int_{w} \Bigg[\int_{t_j=z-k}^{\infty} \mathcal{L}M(..,t_j|v_e,v_u,x)\ .\zeta(z|w,x,\delta(n_m))\ .$$

$$\exp\Big[-\int_0^{n_m} \zeta(s|w,x,\delta(s))\ \mathrm{d}s - \int_{n_m}^{z} \zeta(s|w,x,\delta(n_m))\ \mathrm{d}s\ \Big]\ \mathrm{d}t_j \Bigg]\ +$$

$$+\ \Bigg[\int_{t_j=n_m-k}^{z-k} \mathcal{L}M(..,t_j|v_e,v_u,x)\ .\zeta(z|w,x,1-\delta(n_m))\ .\exp\Big[-\int_0^{n_m} \zeta(s|w,x,\delta(s))\ \mathrm{d}s\ -$$

$$\int_{n_m}^{t_j+k} \zeta(s|w,x,\delta(n_m))\ \mathrm{d}s - \int_{t_j+k}^{z} \zeta(s|w,x,1-\delta(n_m))\ \mathrm{d}s\ \Big]\ \mathrm{d}t_j \Bigg]$$

$$\cdot\ g_{u,e,w}(v_u,v_e,w)\ \mathrm{d}w\ \mathrm{d}v_e\ \mathrm{d}v_u\ \mathrm{d}z$$

To understand this expression, note that we distinguish between $t_j+k>z$ and $t_j+k<z$. In the first case (integral over t_j from $z-k$ to ∞) the labor market

state at the moment at which z is realized is still equal to the state at n_m, so $\delta(z)$ at that moment equals $\delta(n_m)$. In the second case (integral over t_j from n_m-k to $z-k$) there is one labor market transition between n_m and the moment at which z is realized. Thus, the labor market state at the latter moment differs from the state at n_m, and equals $1-\delta(n_m)$. Equation (3.4) has to modified analogously.

Extension to the case of multiple spells of panel survey participation is straightforward.

3.3. *The parameterization of* $g_{u,e,w}(v_u,v_e,w)$.

In this subsection we propose a class of discrete distributions for $g_{u,e,w}(v_u,v_e,w)$ and we explore the consequences of this choice for various model characteristics. We argue that our approach is more general than some previously used approaches to model dependence of duration variables by way of stochastically related unobserved explanatory variables.

In the empirical literature on labor market durations, unobserved heterogeneity is often modeled by way of a discrete random variable (see e.g. Nickell (1979) and Ham & Rea (1987)). Usually, if more than two or three points of support are taken then the estimates of some of them coincide. Heckman & Singer (1984) show that in MPH models the non-parametric maximum likelihood estimator of the heterogeneity distribution is a discrete distribution. However, the estimation procedure requires the number of points of support not to be fixed in advance, and estimation of standard errors is not straightforward. Moreover, the procedure is developed for situations in which right-censoring is independent. Nevertheless, this result illustrates the flexibility of discrete distributions as heterogeneity (or mixture) distributions.

In most applied papers in which multiple duration variables depend on each other by way of their unobserved determinants, a one-factor loading specification is used for the multivariate heterogeneity distribution. This means that the log heterogeneity terms are assumed to be linear functions of a single random variable ω, so e.g. $v_j=\exp(c_{0j}+c_{1j}.\omega)$. This restricts the way that the v_j are related. Lindeboom & Van den Berg (1994) show that in such models there may not be enough flexibility in order to obtain correct estimates of the variances of the duration variables as well as of their interrelation. A genuine multivariate specification for the heterogeneity distribution is to be preferred.

Van den Berg (1994) examines the range of values that the correlation of the duration variables can attain in multivariate MPH models, in general as well as for particular parametric families of the multivariate heterogeneity distribution. It turns out that when the heterogeneity terms have a multivariate discrete distribution with two or more points of support for each, and the locations of these points are not fixed in advance, then all possible correlation values can be attained. On the other hand, when e.g. the log heterogeneity terms have a multivariate normal distribution, or when they have a multivariate discrete distribution in which the locations of the points of support are fixed in advance, then the range of values that can be attained is smaller.

Taken together, these results suggest that in the present context, multivariate discrete heterogeneity distributions with unrestricted mass point locations provide maximum flexibility. This is the approach we will follow here. (See Coleman (1990) and Van den Berg, Lindeboom & Ridder (1994) for other examples of the use of multivariate discrete heterogeneity distributions with unrestricted mass point locations.) Note that discrete distributions are also attractive from a computational point of view.

We assume that v_u, v_e and w all have two points of support (v_{u1}, v_{u2}, v_{e1}, v_{e2}, w_1 and w_2, respectively). We take $v_{u1}\geq v_{u2}>0$, $v_{e1}\geq v_{e2}>0$ and $w_1\geq w_2>0$. The associated probabilities are denoted as follows:

$$p_{i_1i_2i_3} = Pr(v_u=v_{ui_2}, v_e=v_{ei_3}, w=w_{i_1}), \text{ with } i_1,i_2,i_3 \in \{1,2\}, \text{ and } \sum p_{i_1i_2i_3} = 1.$$

We now examine properties of the joint distribution of v_u and w and the joint distribution of t_u and z given x. For reasons of symmetry, the results can be directly translated into results for the distributions of v_e,v_u and t_e,t_u given x, or v_e,w and t_e,z given x. The covariance of v_u and w equals

$$(3.6) \qquad \mathrm{COV}(v_u,w) = ((p_{111}+p_{112})(p_{221}+p_{222}) - (p_{121}+p_{122})(p_{211}+p_{212})) \cdot (v_{u1} - v_{u2}) \cdot (w_1 - w_2)$$

It is easy to show that v_u and w are independent if and only if $\mathrm{COV}(v_u,w)=0$. Also, it can be shown that $\mathrm{CORR}(v_u,w)$ does not depend on the magnitudes of v_{u1}, v_{u2}, w_1 and w_2, when $v_{u1}\neq v_{u2}$ and $w_1\neq w_2$.

Since $t_u|x$ and $z|x$ are independent if and only if v_u and w are independent, it follows that $t_u|x$ and $z|x$ are independent if and only if $(p_{111}+p_{112})(p_{221}+p_{222}) = (p_{121}+p_{122})(p_{211}+p_{212})$ (conditional on $v_{u1}\neq v_{u2}$ and $w_1\neq w_2$). This makes it easy to test for independence between the duration of unemployment and the duration of participation in the panel survey. The covariance of $t_u|x$ and $z|x$ equals

$$(3.7) \qquad \mathrm{COV}(t_u,z|x) = ((p_{111}+p_{112})(p_{221}+p_{222}) - (p_{121}+p_{122})(p_{211}+p_{212})) \cdot (t_{u1} - t_{u2}) \cdot (z_1 - z_2)$$

in which $t_{ui} \equiv \mathrm{E}(t_u|x,v_u=v_{ui})$ and $z_i \equiv \mathrm{E}(z|x,w=w_i)$. Thus, $\mathrm{COV}(t_u,z|x)$ and $\mathrm{COV}(v_u,w)$ always have the same sign. The model is flexible in the sense that it allows either sign of the relation between the duration in state j and the duration of survey participation.

Note that (3.7) follows from the joint distribution of $t_u,z|x$. In general it differs from covariances of observable unemployment duration variables with z, like $\mathrm{COV}(r,z|x)$ or the covariance of z with the second observed unemployment duration for each respondent given that the latter is observed. In the latter case the distribution of v_u,w at the moment that the second unemployment duration starts differs from $g_{u,w}(v_u,w)$, and in general depends on x.

3.4. *Some practical issues.*

In Section 4 we report estimates for nine different model specifications. Model 8 is the general model. Model 1 is the model without unobserved heterogeneity, i.e. the model in which it is imposed that $v_{u1}=v_{u2}$, $v_{e1}=v_{e2}$ and $w_1=w_2$. In Models 2 to 4 we allow for unobserved heterogeneity in θ_u and θ_e, in θ_e and ζ, and in θ_u and ζ, respectively, but we impose that the corresponding heterogeneity terms are independent (so e.g. in Model 4 we impose that $(p_{111}+p_{112})(p_{221}+p_{222}) = (p_{121}+p_{122})(p_{211}+p_{212})$ if $v_{u1}\neq v_{u2}$ and $w_1\neq w_2$, in addition to absence of heterogeneity in θ_e). Models 5–7 generalize Models 2–4, respectively, by allowing heterogeneity to be dependent. Finally, Model 9 allows for unobserved heterogeneity in θ_u, θ_e and ζ but restricts the heterogeneity terms to be mutually independent. Together, the results for these models give a fairly complete account of the structure the unobserved heterogeneity distributions, and therefore of the interrelations between the duration variables.

By comparing the results for Model 1 to those for Models 2–4 (or to those for Model 9) it can be tested whether there is unobserved heterogeneity in the exit rates θ_u, θ_e and ζ. Note however that such a comparison is conditional on independence of v_u, v_e and w. Also note that, because of the denominator in equation (3.1), the likelihood does not factorize in terms of parameters related to unemployment duration and parameters related to employment duration, even if v_u, v_e and z are independent. This is a consequence of the fact that the distribution of employment duration determines the initial conditions for the spells of unemployment that are ongoing at the date of the first interview, and vice versa. On the other hand, in Models 1–4, the likelihood does factorize in a part associated with labor market durations and a part associated with the durations of survey participation. Thus, the estimates of the parameters in the latter part are the same for Models 3, 4 and 9, and the test statistic for unobserved heterogeneity in ζ is independent of the statistics for θ_u and θ_e. The LR tests for $H_0{:}v_{j1}=v_{j2}$ and for $H_0{:}w_1=w_2$ are non-standard, because under the null hypothesis fewer parameters are identified than under the alternative. In the literature it is usually assumed that a test in which critical values of the χ^2_2 distribution are used is on the safe side (see e.g. Ham & Rea (1987)).

By comparing results for different models it can also be tested whether the unobserved heterogeneity terms (and therefore the corresponding duration variables given x) are dependent. Conditional on $v_{u1}\neq v_{u2}$ and $w_1\neq w_2$, testing for independence of v_u and w means testing for $(p_{111}+p_{112})(p_{221}+p_{222}) = (p_{121}+p_{122})(p_{211}+p_{212})$. Consequently, conditional on $v_{u1}\neq v_{u2}$, $v_{e1}=v_{e2}$ and $w_1\neq w_2$, and conditional on all terms in brackets in the previous sentence being strictly between zero and one, the LR test for independence based on a comparison of Models 4 and 7 asymptotically has a χ^2_1 distribution under the null hypothesis. Finally, by comparing the results for Models 8 and 9 we can test for joint dependence of the unobserved heterogeneity terms. Conditional on the presence of unobserved heterogeneity in all exit rates this amounts to testing for three restrictions, so we use critical values of the χ^2_3 distribution.

We prefer LR tests to Wald tests, since the results of the former do not depend on the particular parameterization of the model that is estimated, while the results of the latter do. This seems to be particularly relevant for tests on discrete unobserved heterogeneity distributions (see Van den Berg, Lindeboom & Ridder (1994)).

Except for Model 1, we do not include constant terms in x in (2.1), (2.2) and (2.5), since these would be undistinguishable from multiplicative constants in v_u, v_e and w. Further, instead of estimating p_{111}–p_{222} we estimate q_{111}–q_{222} which are implicitly defined by

$$(3.8) \qquad p_{i_1i_2i_3} = \exp(q_{i_1i_2i_3}) \;/\; \sum \exp(q_{i_1i_2i_3}) \qquad\qquad i_1,i_2,i_3 \in \{1,2\}$$

Because the $p_{i_1i_2i_3}$ sum to one, we normalize by taking q_{222}=0. There is a one–to–one mapping between the seven free p parameters on the one hand, and the seven free q parameters on $<-\infty,\infty>^7$ on the other. So, estimating the latter rather than the former parameters has the well–known advantage that no boundary restrictions have to be imposed on the parameter space.

4. Empirical application.

4.1. *The data.*

For the empirical analysis we use data from the OSA (Netherlands Organization for Strategic Labour Market Research) Labour Supply Panel Survey. This panel survey started in 1985. Presently four waves are available (interviews were held in April–May 1985, August–October 1986, August–October 1988 and August–November 1990, respectively).

In the OSA panel a random sample of households in The Netherlands is followed over time. The study concentrates on individuals who are between 15 and 61 years of age, and who are not full–time students. Therefore only households with at least one person in this category are included. All individuals (and in all cases the head of the household) in this category are interviewed. The first wave consists of 4020 individuals (in 2132 households). In 1992, 1384 (34%) of these individuals are still in the panel. In 1986, 1988, and 1990, refreshment samples were drawn, so that in 1990 the sample size was 4438 individuals.

In the OSA panel an effort is made to collect extensive information on the labor market histories of the individuals. From these labor market histories we obtain the sequence of labor market states occupied by the individuals and the durations of the corresponding spells. Part of the information is retrospective. In particular, an attempt was made to determine the elapsed duration of the spell which was ongoing at the date of the first interview. The following labor market positions are distinguished: employment (job–to–job changes are recorded), self–employment, unemployment, and "not in the labor force" (military service, full–time education, etcetera). In addition to these labor market histories, a number of time–constant individual characteristics are recorded.

In this paper we restrict attention to respondents who were at least participating in the first wave of the panel. Individuals who were self-employed for some period during the time span covered by the survey are omitted, since it is likely that the behavior of such individuals, at least in a certain period, deviates substantially from the behavior that the model intends to describe. For similar reasons, we do not use information on respondents who are observed to be working in a part-time job or who are observed to be a nonparticipant for some period. An alternative approach would be to extend the model to include a state of nonparticipation, and allow for transitions to and from this state. This would extend the dimensionality enormously. Moreover, transitions to and from nonparticipation are rare in the data. Therefore, using information on such transitions in an extended model context would, except for a number of imprecisely estimated nuisance parameters, probably not result in any gains. The restrictions reduce the number of labor market states to two: unemployment and full-time employment.

The indicated selection results in a sample of 2336 individuals, of which 239 (2097) were unemployed (employed) at the date of the first interview. Table 1 gives sample averages of explanatory variables. We restrict attention to the effect of the regressors age, education (we distinguish 5 levels), occupation (6 levels), marital status (married), sex (female) and nationality (Dutch).

In our sample, 31% only participates in the first wave, 21% only participates in the first and the second wave, 12% only participates in the first three waves, and 33% participates in all four waves of the panel. Further, 2% only participates in the first and the third wave, and 1% only participates in the first, third and fourth wave. Because the latter two groups are so small, we decided not to use more than one spell of panel survey participation in the analysis. The participation percentages imply that the over-all conditional probability of exit out of the panel between two consecutive waves of the panel is slightly decreasing over time, but is on average close to 30%. Because for most respondents the length of time between the first two interviews is about 75% of the length of time between other consecutive interviews, this means that the over-all exit rate is somewhat decreasing after the first interview and is fairly constant after the second interview. Note that the magnitude of attrition is larger than usually encountered in US panel surveys. This seems to be a typical feature of panel surveys in The Netherlands, and also applies to nonresponse in general. Also recall that in the present survey the time span between two consecutive interviews is relatively large.

Table A1 lists participation numbers for given characteristics of the respondent at the date of the first interview (see the Appendix). It turns out that attrition is relatively large for non-Dutch individuals, for females, for non-married individuals, for individuals with a low level of occupation (meaning a low level of complexity of the work they do) and for young individuals. There is no clear relation to the level of education. As shown below, all these results are confirmed by the estimates of β_z in ζ. Table A1 also shows that attrition is relatively large for individuals who are unemployed at the date of the first interview. Below we report some sample statistics on the percentages of labor market durations that are censored due to attrition.

4.2. *Explorative empirical analysis.*

At this stage we did not estimate the most general model with flexible baseline hazards and unconstrained β_l. The results presented below are based on a model with constant baseline hazards and $\beta_l \equiv 0$. Furthermore, we so far only used at most two labor market spells per individual. The latter constraint is binding for less than 10% of the sample. It turns out that a substantial number of labor market durations are censored due to attrition. For the individuals who are unemployed (employed) at the date n_1 of the first interview, 61% (64%) of the residual durations r is censored due to attrition. Of all (partially) observed subsequent employment (unemployment) durations, 44% (62%) is censored due to attrition.

We perform two different kinds of formal explorative analyses. First of all, we estimate duration models using the p data only, i.e. using only the elapsed (un)employment durations at n_1. These endogenous labor market variables are not subject to attrition. We adopt simple loglinear specifications of the hazard rates θ_u and θ_e as functions of explanatory variables including dummies indicating whether attrition has occurred after the first interview. The test is then the following. If the unobserved heterogeneity term in ζ is (is not) related to the unobserved heterogeneity terms of θ_u and θ_e, then the attrition dummies should be significant (insignificant).

This test resembles standard tests on informative attrition in a discrete-time regression-type model context. As Gottschalk & Moffitt (1992) explain, such tests can only detect a relation between attrition and the permanent unobserved components of labor market behavior. It may well be that in reality there is a relation between attrition and transitory components of labor market behavior. For example, there may be a high probability that individuals drop out of the panel immediately after (and as a consequence of) the occurrence of an actual labor market transition. In that case there is a nonignorable relation between attrition and labor market behavior even if there is no unobserved heterogeneity at all. Such a relation cannot be detected by the test outlined above. On the other hand, it is plausible that estimation of the general model using multi-spell data (see Subsection 4.3) does enable the detection of the latter type of relations. In the end it remains of course possible that there are mutually offsetting relations with both permanent and transitory components.

The estimates of the procedure described above are presented in Table A2 (see the Appendix). (Here, as in the sequel, the unit time period is one month, and t-values are in parentheses.) It turns out that the occurrence of attrition has a significantly positive effect on the estimates of θ_u and θ_e.

As a second formal explorative analysis, we estimated probit models for the occurrence of attrition, including as an additional explanatory variable the (possibly censored) value of p, distinguishing between whether it concerns an employment duration or an unemployment duration. If the individual is employed at n_1, then the value of p referring to unemployment duration is set to zero, and vice versa. Note that this is a very crude procedure. First of all, we do not deal with censoring of p in a sophisticated way. In the next subsection we argue that, in our data, censoring of p in case it refers to an employment duration is particularly awkward. Secondly, note that the estimated effects of p will be influenced by the fact that the exit rate out of the panel is larger for the group of individuals who are unemployed at n_1 (see the previous subsection). These caveats basically reflect the difficulty to design simple explorative analyses on informative attrition in case of multi-state duration models.

In a way, this second explorative analysis is the mirror-image of the analysis above. Not surprisingly, therefore, the results are in accordance to those above (see Table A3 in the Appendix; a positive coefficient means that the probability of attrition is smaller). As shown below, the estimated covariate effects are in accordance to those for β_z in the models estimated in the next subsection, and discussion of them is therefore postponed.

The main conclusion of the explorative analyses is that there is evidence of a relation between the (permanent) unobserved determinants of attrition and labor market durations. As we will see below, this confirms the estimation results of the general model. One might propose additional explorative analyses, like estimating the labor market model with independent attrition with data from different numbers of waves of the panel, to see whether the results are significantly different. However, under the alternative hypothesis of informative attrition the estimates are biased regardless of the number (if >1) of waves used, so such a "test" may have very low power. Therefore we do not perform such analyses. Below we do however estimate the general model with dependent attrition with different numbers of waves, to check the robustness of the results.

4.3. *Estimation results.*

For the estimation of Models 1–9, we did not use elapsed durations p for individuals who are employed at n_1. The reason for this is that they are often censored, in an awkward way. If the respondent is employed at n_1 then in general there is only information on the elapsed duration of the job held at that date. This provides a lower bound on the elapsed duration of employment, but to the extent in which job durations are not randomly distributed within a spell of employment, this lower bound cannot be interpreted as independent right–censoring of p. According to search theory, job durations occurring towards the end of a spell of employment are longer than job durations occurring early in such a spell (see e.g. Mortensen (1986) for a survey). Therefore we have decided to drop such data and integrate p for d=0 out of the likelihood function.

Table 2 contains the parameter estimates for Model 1, the model without unobserved heterogeneity, and Model 8, the general model. Results for the parameters of the mixing distribution for all Models 1–8 are reported in Table 3. Table A4 contains all the estimates for Model 9. We first discuss Table 2, and start with a brief discussion of the results for Model 1.

Most results are as expected. Unemployment duration is strongly affected by age, education and sex. Females and elderly individuals experience longer spells, whereas the more educated individuals experience shorter unemployment spells. For employment duration the results are somewhat different. The effect of sex, nationality and education seem to be negligible, as compared to the strong effects of occupation and marital status. The married and those working on a higher occupational level have longer employment spells.

For survey participation, education appears to be the only variable not of influence. Strong effects are found for the other variables. The fact that the exit rate out of the panel decreases in age is partly due to young adults leaving their parents' household. In surveys in which such individuals are followed, the age effect on attrition is usually opposite and dominated by mortality (see e.g. Fitzgerald, Gottschalk & Moffitt (1994)). In accordance to other studies, attriters have lower skills and belong relatively often to less stable household structures or minorities.

It is remarkable that for each variable with a positive (negative) significant effect on survey participation duration, we also find a positive (negative) effect on employment duration. This suggests that leaving employment may imply a relatively high risk of dropping out of the sample. Stated differently, those with high risk of dropping out of the survey have also shorter employment spells. This is in line with a priori expectations. Another consequence of this relationship may be that, if indeed the processes are governed by the same set of exogenous variables, one may expect these processes to be sensitive to misspecification in either of them. We will return to this below. At first sight there appears to be no obvious relationship between the parameters of survey participation and unemployment duration or employment duration and unemployment duration.

The results for Model 8 (the general model) are reported in the last columns of Table 2. As a general remark one may note that in most cases the significance level of the variables is (slightly) reduced. But more importantly, changes in the parameter estimates occur when one allows for correlated unobserved heterogeneity. This is most prominent for the variables "Married", "Female" and "Dutch". The decrease for "Married" and "Dutch" in ζ is accompanied by a decrease for the same variables in θ_e, whereas the increase for "Female" in ζ is accompanied by a decrease for "Female" in θ_e. In θ_u the change for the variables "Married", "Female" and "Dutch" is in the opposite direction.

Still, it is clear that for all means and purposes the estimates of the covariate effects in β_u and β_e in Models 1 and 8 do not differ a lot. Indeed, in most cases the difference between the estimates for Models 1 and 9 is larger than the difference for Models 8 and 9, so most of the difference for Models 1 and 8 is due to the fact that Model 1 does not allow at all for unobserved heterogeneity. This is important, because in any conventional empirical analysis these covariate effects are the parameters of interest. So, even if in the sequel it will turn out that unobserved heterogeneity terms are significantly dependent, it does not really matter whether one takes account of this or not, since the estimates of what normally are the parameters of interest are insensitive with respect to this.

In Table 3 we report the parameter estimates of the mixing distribution for Models 1–8 along with some correlations. Note that absolute magnitudes of mass points are not very informative by themselves. We will now discuss in more detail what happens when going from Model 1 to more complex models.

Combining the likelihood values for Models 1–4, it can be derived that adding independent unobserved heterogeneity in the unemployment, employment and survey participation duration distributions accounts for an increase of the log likelihood values of 1.6, 27.5 and 37.9, respectively. Using Chi-square critical values it follows that, conditional on independence of heterogeneity, allowing for heterogeneity in employment duration and survey participation duration is a significant improvement of the model. Note that the sum of these increases in log likelihood values equals 67.1, while the difference of the log likelihood values for Models 1 and 9 equals 67.8. These numbers are not exactly the same because the likelihood in these models does not factorize in terms of unemployment and employment duration parameters.

Next, in Models 5–7 we allow for pairwise dependence of employment and unemployment, employment and survey participation, and unemployment and survey participation, respectively. For v_e and w the null hypothesis of independence is strongly rejected. The large differences between the locations of the mass points of the heterogeneity distributions in Model 3 and Model 6 also point in this direction. Note that notably the mass points of v_e change, which indicates that the attrition process is informative for the employment duration. This may be due to the fact that a typical spell of employment consists of three or four consecutive job spells. Thus, within a typical spell of employment, individuals move from one job to another (and therefore potentially from one location to another) a couple of times.

On the basis of a comparison of the results for Models 2–7 we cannot reject mutual independence of v_e and v_u, or of v_u and w. Note that erroneously ignoring the dependence between v_e and w would also affect the estimates of the unemployment duration distribution, even if v_u is independent from w. This is because the steady state employment and unemployment probabilities in the likelihood (see equation (3.1)) do not factorize in terms of the parameters of the unemployment and employment distributions. The latter reflects the fact that the employment duration distribution determines the initial conditions of the unemployment durations that are ongoing at n_1. In case of a sample at n_1 of the inflow into unemployment this effect would not play a role.

The argument of the previous paragraph suggests that allowing for dependence of v_u on v_e and w in a model in which v_e and w are already allowed to be dependent may affect the fit in a different way than it would in a model in which v_e and w are independent. To examine this we test for joint dependence of the unobserved heterogeneity terms conditional on the presence of unobserved heterogeneity in all three distributions. First of all, let us compare the results for Models 8 and 9. Allowing for joint dependence leads to an increase of the log likelihood of 27.8. From this it follows that joint independence of v_u, v_e and w is rejected.

Indeed, the magnitude of this increase is much larger than the sum of the increases when going from Models 2–4 to Models 5–7 (this sum equals 9.6). By comparing the results for Models 5–7 to the results for Model 8, some additional information on this can be obtained. For instance, conditional on the presence of unobserved heterogeneity in the three distributions, and conditional on dependence of v_e and w, we can test informally whether joint dependence with v_u improves the fit of the model. This boils down to imposing two restrictions on Model 8. The value of the "test" statistic equals 21.5 (5407.2 is compared to 5430.3 – 1.6 (being the minus log likelihood value for Model 6 corrected for heterogeneity in v_u; note that this is an approximation as the likelihood does not factorize in employment and unemployment durations; also note that an alternative calculation involving Models 3, 6, 8 and 9 gives a value of 20.1)). It follows that joint dependence with v_u improves the fit of the model substantially. Also note from Table 3 that allowing for joint dependence changes CORR(v_u,w) from 0.54 (in Model 7) to 0.72.

The estimation results for the general model do not change much if the data from the last (fourth) wave of the panel are deleted. If in addition data from the second and third wave are deleted then the results do change. However, in the latter case the number of respondents for which we observe more than one labor market spell is very small, so the latter results will rely heavily on the MPH assumption (see Section 2). In a way, this therefore illustrates the importance of using multi–spell data.

We also experimented with different sets of regressors in the exit rates out of unemployment and employment. It turns out that this does not affect the main conclusions. When we allow for different sets of regressors in θ_e and θ_u then this mainly influences the effects of the regressors that are non–mutual in θ_e and θ_u. This is not surprising given the presence of the steady–state probabilities of employment and unemployment in the likelihood.

A general feature of estimates of models of (informative) attrition and their implications is that they are sensitive to the setup of the panel survey and the efforts of the agency running the survey. This is obvious as far as the level of the exit rate ζ out of the panel is concerned. However, to a certain extent it may also be true for the degree of dependence between labor market behavior and attrition. At the marginal effort, the composition of the sample (and therefore the joint distribution of heterogeneity terms) may change when effort increases. In general one should therefore be cautious when generalizing particular empirical results on informative attrition.

5. Conclusion.

In this paper we have analyzed the relation between individual labor market behavior over time and the duration of participation in panel surveys. We used flexible models which allow for dependence by way of stochastically related unobserved determinants of the duration of survey participation and the durations of being unemployed and being employed.

We argue that from an empirical point of view it is important to analyze these issues in the context of a multi–state labor market model, and to have data containing multiple spells in either state. In a multi–state multi–spell framework, the assumptions needed for empirical inference are much weaker than in a (single–state) single–spell framework. Moreover, the use of multi–spell data facilitates the identification of the (joint) determinants of the various exit rates in the model.

As an application, we estimate a multi–state model using multi–spell data from a panel survey. The empirical analysis shows that unemployment and employment durations are positively related to the duration of survey participation. Tests show in particular a strong significant relation between the unobserved determinants of employment durations and attrition. This alone affects estimation of the unemployment duration distribution, since the distribution of employment durations influences the initial conditions of unemployment durations ongoing at the first wave of the panel. However, we also find evidence for a direct relation between the unobserved determinants of unemployment duration and attrition.

On the other hand, the estimates of the covariate effects in the labor market transition rates do not change a lot when allowing for these relations between labor market durations and attrition. In any standard empirical analysis these covariate effects are the parameters of interest. So, even though we formally find significant dependence between labor market durations and attrition, it does not really matter whether we take account of this or not. In other words, spells that are incomplete due to attrition may be treated as spells that are subject to independent right–censoring.

Some subjects for further research emerge. It is likely that the exit rates out of the panel for different members of the same household are related. This violates the i.i.d. assumption. To deal with this, we can extend the model by allowing the unobserved heterogeneity variables corresponding to these exit rates to be related within a household. Secondly, it would be interesting to analyze a panel survey in which there is substantial return to the panel by individuals who previously attrited. This may clarify whether a relation between labor market behavior and attrition works by way of unobserved heterogeneity ("individual-specific permanent components") or whether attrition is a direct consequence of labor market transitions ("transitory shocks").

REFERENCES

Andersen, P.K., Ø. Borgan, R.D. Gill and N. Keiding (1993), *Statistical models based on counting processes* (Springer-Verlag, New York).

Butler, J.S., K.H. Anderson and R.V. Burkhauser (1989), Work and health after retirement, *Review of Economics and Statistics* 71, 46–53.

Chesher, A. and T. Lancaster (1983), The estimation of models of labour market behaviour, *Review of Economic Studies* 50, 609–624.

Coleman, T.S. (1990), Unemployment behaviour: evidence from the CPS work experience survey, in: Y. Weiss and G. Fishelson, eds., *Advances in the theory and measurement of unemployment* (Macmillan, Basingstoke).

Devine, T.J., and N.M. Kiefer (1991), *Empirical labor economics* (Oxford University Press, New York).

Fitzgerald, J., P. Gottschalk and R. Moffitt (1994), A study of sample attrition in the Michigan Panel Study of Income Dynamics, Working Paper, Institute for Social Research, Ann Arbor.

Flinn, C.J. and J.J. Heckman (1982), Models for the analysis of labor force dynamics, in: R. Basmann and G. Rhodes, eds., *Advances in econometrics* Vol. 1 (JAI Press, Greenwich).

Gottschalk, P. and R. Moffitt (1992), Proposal for a study of sample attrition in the Michigan Panel Study of Income Dynamics, Working Paper, Boston College.

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

Ham, J.C. and S.A. Rea (1987), Unemployment insurance and male unemployment duration in Canada, *Journal of Labor Economics* 5, 325–353.

Heckman, J.J. (1979), Sample selection bias as a specification error, *Econometrica* 47, 153–161.

Heckman, J.J. and B.E. Honoré (1989), The identifiability of the competing risks model, *Biometrika* 76, 325–330.

Heckman, J.J. and B. Singer (1984), A method for minimizing the impact of distributional assumptions in econometric models for duration data, *Econometrica* 52, 271–320.

Honoré, B.E. (1993), Identification results for duration models with multiple spells, *Review of Economic Studies* 60, 241–246.

Lancaster, T. (1990), *The econometric analysis of transition data* (Cambridge University Press, Cambridge).

Layard, R., S. Nickell and R. Jackman (1991), *Unemployment* (Oxford University Press, Oxford).

Lillard, L.A. and C.W.A. Panis (1994), Attrition from the PSID: marital status, income and mortality, Working Paper, RAND, Santa Monica.

Lindeboom, M. and J.J.M. Theeuwes (1991), Job duration in The Netherlands: the co-existence of high turnover and permanent job attachment, *Oxford Bulletin of Economics and Statistics* 53, 243–264.

Lindeboom. M. and G.J. van den Berg (1994), Heterogeneity in bivariate duration models: the importance of the mixing distribution, *Journal of the Royal Statistical Society Series B* 56, 49–60.

Link, W.A. (1989), A model for informative censoring, *Journal of the American Statistical Association* 84, 749–752.

Mortensen, D.T. (1986), Job search and labor market analysis, in: O. Ashenfelter and R. Layard, eds., *Handbook of labor economics* (North-Holland, Amsterdam).

Nickell, S.J. (1979), Estimating the probability of leaving unemployment, *Econometrica* 47, 1249–1266.

Stasny, E.A. (1988), Modeling nonignorable nonresponse in categorical panel data with an example in estimating gross labor-force flows, *Journal of Business and Economic Statistics* 6, 207–219.

Tzeng, M.S. and R.D. Mare (1994), Sibling models for panel attrition bias in the analysis of school transitions, Working Paper, University of Wisconsin, Madison.

Van den Berg, G.J. (1992), A structural dynamic analysis of job turnover and the costs associated with moving to another job, *Economic Journal* 102, 1116–1133.

Van den Berg, G.J., M. Lindeboom and G. Ridder (1991), Attrition in longitudinal panel data, and the empirical analysis of dynamic labour market behaviour, Research Memorandum, Groningen University and Leiden University.

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*, forthcoming.

Van den Berg, G.J. (1994), Association measures for durations in bivariate hazard rate models, Research Memorandum, Free University Amsterdam and INSEE-CREST, Paris.

Table 1 Descriptives of variables used in the analysis

Variable	mean	st. dev	min	max
Age	35.01	10.37	16.0	70.0
Education	1.88	0.85	1	4
Occupational level	2.21	1.02	1	4
Married	0.76	0.43	0	1
Female	0.34	0.47	0	1
Dutch	0.95	0.21	0	1

Table 2. Estimation results

variables/parameters	Model 1 (no heterogeneity)		Model 8 (general model)	
i) Results on unemployment duration				
Log age	-1.605	(8.1)	-1.523	(6.0)
Education	0.180	(2.2)	0.179	(1.9)
Occupational level	-0.107	(1.4)	-0.049	(0.6)
Married	0.015	(0.2)	0.097	(0.6)
Female	-0.359	(3.2)	-0.256	(1.9)
Dutch	0.407	(1.5)	0.384	(1.3)
Constant	1.510	(2.2)	-	-
ii) Results on employment duration				
Log age	-1.417	(5.8)	-1.270	(3.5)
Education	-0.083	(0.8)	-0.125	(1.0)
Occupational level	-0.267	(3.0)	-0.322	(2.9)
Married	-0.933	(6.1)	-1.320	(6.0)
Female	0.069	(0.5)	0.259	(1.4)
Dutch	0.076	(0.2)	-0.291	(0.7)
Constant	0.007	(0.0)	-	-
iii) Results on survey participation duration				
Log age	-0.219	(2.3)	-0.205	(1.6)
Education	0.046	(1.3)	0.036	(0.8)
Occupational level	-0.071	(2.3)	-0.084	(2.1)
Married	-0.359	(5.8)	-0.436	(5.2)
Female	0.168	(3.0)	0.203	(2.9)
Dutch	-0.304	(2.8)	-0.451	(3.1)
Constant	-2.594	(7.5)	-	-
-Log likelihood	5502.74		5407.22	

Table 3 Estimates of the mixing distribution for Models 1-8

i) Model 1: No heterogeneity			
$\log(v_u)$ $\log(v_e)$ $\log(w)$	1.510 0.007 -2.594		
-log likelihood		5502.74	
ii) Model 2: Heterogeneity in t_u and t_e, $v_e \perp v_u$			
	$\log(v_{u1})=2.775$	$\log(v_{u2})=1.602$	
$\log(v_{e1})=2.946$ $\log(v_{e2})=-1.147$	0.046 0.195	0.144 0.615	0.190 0.810
	0.241	0.759	1
-log likelihood		5473.62	
iii) Model 3: Heterogeneity in t_e and z, $v_e \perp w$			
	$\log(w_1)=-1.229$	$\log(w_2)=-2.907$	
$\log(v_{e1})=2.683$ $\log(v_{e2})=-1.397$	0.067 0.260	0.138 0.535	0.205 0.795
	0.327	0.673	1
-log likelihood		5437.30	
iv) Model 4: Heterogeneity in t_u and z, $v_u \perp w$			
	$\log(w_1)=-1.229$	$\log(w_2)=-2.907$	
$\log(v_{u1})=2.470$ $\log(v_{u2})=1.394$	0.065 0.262	0.486 0.187	0.551 0.449
	0.327	0.673	1
-log likelihood		5463.16	
v) Model 5: Heterogeneity in t_u and t_e, $v_e \not\perp v_u$			
	$\log(v_{u1})=2.938$	$\log(v_{u2})=1.669$	
$\log(v_{e1})=3.286$ $\log(v_{e2})=-0.829$	0.051 0.0	0.120 0.829	0.171 0.829
	0.051	0.949	1
$\mathrm{Corr}(v_u,v_e)$ -log likelihood		0.51 5472.93	

Table 3 (continued)

vi) Model 6: Heterogeneity in t_e and z, $v_e \not\perp w$			
	$\log(w_1)=-1.511$	$\log(w_2)=-3.144$	
$\log(v_{e1})=2.045$ $\log(v_{e2})=-18.16$	0.231 0.249	0.114 0.406	0.345 0.655
	0.480	0.520	1
$Corr(v_e,w)$ -log likelihood		0.28 5430.35	
vii) Model 7: Heterogeneity in t_u and z, $v_u \not\perp w$			
	$\log(w_1)=-1.280$	$\log(w_2)=-2.933$	
$\log(v_{u1})=1.599$ $\log(v_{u2})=0.947$	0.363 0.0	0.300 0.337	0.663 0.337
	0.363	0.637	1
$Corr(v_u,w)$ -log likelihood		0.54 5461.27	
viii) Model 8: General model			
$\log(w_1)=-1.408$			
	$\log(v_{u1})=1.816$	$\log(v_{u2})=0.583$	
$\log(v_{e1})=1.727$ $\log(v_{e2})=-10.55$	0.328 0.081	0.0 0.0	0.328 0.081
	0.409	0.0	0.409
$\log(w_2)=-3.022$			
	$\log(v_{u1})=1.816$	$\log(v_{u2})=0.583$	
$\log(v_{e1})=1.727$ $\log(v_{e2})=-10.55$	0.0 0.168	0.112 0.311	0.112 0.479
	0.168	0.423	0.591
$Corr(v_u,v_e)$ $Corr(v_u,w)$ $Corr(v_e,w)$ -log likelihood		0.30 0.72 0.61 5407.22	

Table A1 Cross-tabulations of some variables and survey participation

	Survey participation*				
Variable	1	2	3	4	Total
Age ∈ [15,25]	113	37	103	228	481
Age ∈ [26,35]	297	96	170	247	810
Age ∈ [36,50]	305	118	164	235	822
Age ∈ [51,70]	59	22	50	92	223
Education level 1	294	92	174	319	879
Education level 2	313	112	197	328	950
Education level 3	134	53	92	118	397
Education level 4	33	16	24	37	110
Occupational level 1	184	70	130	209	593
Occupational level 2	354	111	216	396	1077
Occupational level 3	74	33	57	81	245
Occupational level 4	162	59	84	116	421
Married = 0	141	58	118	249	566
Married = 1	633	215	369	553	1770
Female = 0	551	189	305	495	1540
Female = 1	223	84	182	307	796
Dutch = 0	27	16	23	53	119
Dutch = 1	747	257	464	749	2217
Unemployed	62	28	56	93	239
Employed	712	245	431	709	2097

* Survey participation = 1 if individual participates in 85, 86, 88, 90 (Total number = 774)
2 if individual participates in 85, 86, 88 (Total number = 273)
3 if individual participates in 85, 86 (Total number = 487)
4 if individual participates in 85 (Total number = 802)

Table A2 Estimation results for E and U based on p, with future attrition dummies as regressors

	1		2		3		4	
Unemployed								
Log (age)	-1.42	(6.1)	-1.34	(5.6)	-1.35	(5.5)	-1.36	(5.6)
Education	0.19	(1.7)	0.19	(1.7)	0.19	(1.7)	0.19	(1.7)
Occupational level	-0.13	(1.4)	-0.11	(0.5)	-0.11	(1.7)	-0.11	(1.1)
Married	-0.18	(1.2)	-0.15	(1.6)	-0.14	(1.5)	-0.15	(0.9)
Female	-0.53	(3.9)	-0.51	(3.0)	-0.50	(3.1)	-0.50	(3.4)
Dutch	0.36	(1.3)	0.38	(2.7)	0.40	(2.8)	0.40	(1.4)
ATT	-		0.22	(3.2)	0.19	(3.4)	0.23	(0.9)
ATT 1	-		-		0.06	(0.6)	0.08	(0.4)
ATT 2	-		-		-		-0.07	(1.0)
Constant	1.52	(15.7)	1.02	(3.6)	1.01	(3.5)	1.04	(5.6)
Employed								
Log (age)	-5.28	(23.3)	-5.21	(88.0)	-5.12	(54.2)	-5.12	(92.0)
Education	0.24	(3.6)	0.22	(2.1)	0.23	(1.2)	0.23	(2.4)
Occupational level	-0.07	(1.2)	-0.06	(0.6)	-0.07	(0.5)	-0.07	(0.7)
Married	-0.45	(4.1)	-0.42	(3.9)	-0.42	(3.8)	-0.42	(2.2)
Female	0.37	(0.5)	0.34	(0.4)	0.34	(0.4)	0.34	(2.1)
Dutch	-0.82	(1.0)	-0.80	(1.0)	-0.80	(0.9)	-0.80	(6.5)
ATT	-		0.31	(2.0)	0.21	(0.9)	0.25	(1.5)
ATT 1	-		-		0.21	(1.3)	0.23	(0.3)
ATT 2	-		-		-		-0.06	(0.1)
Constant	12.87	(64.1)	12.40	(61.7)	12.13		12.15	(47.1)
-Log likelihood	3527.49		3522.65		3520.97		3520.86	

ATT = 1 if attrition occurs
ATT 1 = 1 if only observed in 1985 (first wave)
ATT 2 = 1 if only observed in 1985 or 1986 (first or second wave)

Table A3: Attrition probits with p as regressor

	1		2		3		4	
Constant	-1.18	(3.3)	0.21	(0.6)	1.26	(2.8)	-0.90	(1.6)
Log (age)	0.07	(0.7)	-0.42	(4.4)	-0.7	(5.4)	-0.17	(1.1)
Education	-0.064	(1.7)	-0.03	(1.0)	0.11	(2.4)	0.07	(1.3)
Dutch	0.32	(2.4)	0.30	(2.3)	0.16	(1.1)	-0.01	(0.1)
Married	0.28	(3.9)	0.28	(4.0)	0.18	(2.0)	0.13	(1.3)
Female	-0.16	(2.7)	-0.08	(1.3)	0.09	(1.2)	-0.05	(0.6)
Occupational level	0.07	(2.2)	0.06	(1.9)	-0.02	(0.4)	0.027	(0.6)
P unemployed			0.0083	(3.8)	0.016	(5.0)	0.0079	(2.4)
P employed			0.0031	(9.9)	0.010	(13.0)	0.0034	(7.6)
-Log likelihood	1454.73		1407.20		908.76		681.71	

Specification 1 and 2: probability of no attrition in any wave
Specification 3: condional on attrition probability of no attrition in wave 1986
Specification 4: conditional on attrition probability of no attrition in 1985 or 1988

Table A4 estimation results for model with Ve⊥Vu, Ve⊥W, Vu⊥W

i) Results on unemployment duration		
Log age	-1.628	(7.5)
Education	0.180	(2.0)
Occupational level	-0.120	(1.4)
Married	-0.021	(0.1)
Female	-0.392	(3.1)
Dutch	0.360	(1.2)
ii) results on unemployment duration		
Log age	-1.280	(3.3)
Education	-0.163	(1.2)
Occupational level	-0.412	(3.2)
Married	-1.666	(5.9)
Female	0.247	(1.1)
Dutch	-0.186	(0.4)
iii) Results on survey participation duration		
Log age	-0.215	(1.8)
Education	0.038	(0.8)
Occupational level	-0.083	(2.1)
Married	-0.441	(5.2)
Female	0.197	(2.7)
Dutch	-0.408	(2.8)
iv) Results on the mixing distribution		
Log (v_{u1})	2.775	(3.0)
Log (v_{u2})	1.602	(2.0)
Log (v_{e1})	2.946	(2.0)
Log (v_{e2})	-1.147	(0.7)
Log (w_1)	-1.230	(7.1)
	-2.909	(28.7)
Pu	0.241	(1.5)
Pe	0.190	(3.6)
Pz	0.327	(3.2)
-Log likelihood	5434.97	

1992-1	R.J. Boucherie N.M. van Dijk	Local Balance in Queueing Networks with Positive and Negative Customers
1992-2	R. van Zijp H. Visser	Mathematical Formalization and the Analysis of Cantillon Effects
1992-3	H.L.M. Kox	Towards International Instruments for Sustainable Development
1992-4	M. Boogaard R.J. Veldwijk	Automatic Relational Database Restructuring
1992-5	J.M. de Graaff R.J. Veldwijk M. Boogaard	Why Views Do Not Provide Logical Data Independence
1992-6	R.J. Veldwijk M. Boogaard E.R.K. Spoor	Assessing the Software Crisis: Why Information Systems are Beyond Control
1992-7	R.L.M. Peeters	Identification on a Manifold of Systems
1992-8	M. Miyazawa H.C. Tijms	Comparison of Two Approximations for the Loss Probability in Finite-Buffer Queues
1992-9	H. Houba	Non-Cooperative Bargaining in Infinitely Repeated Games with Binding Contracts
1992-10	J.C. van Ours G. Ridder	Job Competition by Educational Level
1992-11	L. Broersma P.H. Franses	A model for quarterly unemployment in Canada
1992-12	A.A.M. Boons F.A. Roozen	Symptoms of Dysfunctional Cost Information Systems
1992-13	S.J. Fischer	A Control Perspective on Information Technology
1992-14	J.A. Vijlbrief	Equity and Efficiency in Unemployment Insurance
1992-15	C.P.M. Wilderom J.B. Miner A. Pastor	Organizational Typology: Superficial Foursome of Organization Science?
1992-16	J.C. van Ours G. Ridder	Vacancy Durations: Search or Selection?
1992-17	K. Dzhaparidze P. Spreij	Spectral Characterization of the Optional Quadratic Variation Process
1992-18	J.A. Vijlbrief	Unemployment Insurance in the Netherlands, Sweden, The United Kingdom and Germany
1992-19	J.G.W. Simons	External Benefits of Transport