A bivariate duration model for job mobility of two-earner households

Jos van Ommeren, Piet Rietveld, Peter Nijkamp*

Department of Spatial Economics, Free University, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands

Received 1 December 1999; accepted 28 February 2001

Abstract

In recent years we have observed an increasing interest in the job mobility patterns of employed persons. A new issue in this context is the mutual dependence of job mobility choice of workers who belong to the same household. In the present paper, we focus on bivariate duration models of stock sampled data in which dependence is induced through mixing. We apply the estimation method on the basis of an empirical study on the mutual dependence of job mobility of workers belonging to a two-earner household in the Netherlands. An interesting empirical result is that the marginal willingness to pay for a reduction in commuting time in a two-wage earner household is higher than that usually found for single-earner workers. © 2002 Elsevier Science B.V. All rights reserved.

1. Introduction

This study addresses job mobility in two-earner households. In the past decade the statistical analysis of movements of individuals between distinct states (e.g. employment, unemployment, residence) has become a popular research area. The empirical investigation of the duration of stay in a state is helpful in understanding these movements. In many cases, however, studies on duration of stay have found that unobserved factors, which cannot be identified, play a crucial role (Kiefer, 1988). These unobserved factors lead to the well-known phenomenon that when one follows a cohort of workers who started their job sometime ago, the probability of a move to another job decreases as the duration of the present job increases. The reason is that the cohort is heterogeneous: more flexible workers will change jobs earlier than others. In the course of time, the composition of the cohort is therefore continuously changing. Of course, when data on the individuals in a cohort are available, these are used to take care of the heterogeneity (for example, young people change jobs more frequently than older people), but one will usually find that there remain unobserved factors that lead to a slowdown of the process described above.
The obvious conclusion is that estimation methods of job moving behaviour should take into account these unobserved factors. The aim of the present paper is to indicate how this can be done in the case of two-earner households. We will explore whether the unobserved factors of one wage earner are related to the unobserved factors of the other wage earner belonging to the same household. Therefore, in the present paper, we will focus on the analysis of joint durations of stay in two states, and we will consider bivariate duration models allowing for unobserved variables. During the past decades these models have received some attention in the literature (see, for example, Flinn and Heckman, 1982; Ham and Rea, 1987; Hougaard, 1987; Butler et al., 1989; Lindeboom and van den Berg, 1994 or Van den Berg et al., 1996).

The main contribution of this paper is to analyse job mobility of workers belonging to a two-earner household while allowing for a mutual dependence between the unobserved variables. In order to test the sensitivity of this approach, it seems therefore useful to compare the results with analyses that either ignore unobserved variables or ignore the mutual dependence between unobserved variables.

In the present paper, we apply the estimation method by means of an empirical study of the mutual dependence of job mobility rates of workers belonging to the same household in the Netherlands. In principle, this enables one to obtain more precise estimates than those obtained by univariate methods.

The outline of the paper is as follows. We first derive the distributions that are needed for maximum likelihood estimation of bivariate duration models (Section 2). In Section 3, we derive the distributions to apply conditional likelihood methods, where we condition on the elapsed durations of stay. The second part of the paper (Section 4) contains an illustration of the estimation procedure using observations of job durations of workers belonging to a two-earner household. Section 5 offers some concluding remarks.

2. The distribution of bivariate duration observations

A crucial issue in the empirical application of duration models is the way the data have been collected. Information about the durations of stay can be acquired by means of two ways of sampling: stock sampling and flow sampling. The ‘stock sampling’ approach means that one randomly samples those members of the population at a certain point of time who are in the state of interest. The ‘flow sampling’ approach means that one samples the set of people who enter the state of interest during a certain time interval (see Lancaster, 1990). Often, it is easier (and therefore cheaper) to gather data on stocks. Even more important, most available data sets that contain information on job behaviour are stock samples. Therefore, in this paper we will concentrate on the analysis of individuals sampled from a stock.

Generally, the analysis of moving behaviour based on stock sampled duration observations compared to flow sampled duration observations is rather complicated for two well-known reasons. First, individuals with longer durations are oversampled. This phenomenon is labelled ‘length biased sampling’. Second, the distribution of the unobserved explanatory variables of the individuals in a stock sample may not be equal to the distribution of the unobserved explanatory variables in the population. The latter is labelled ‘density biased sampling’. Therefore, the distribution of unobserved heterogeneity in the stock sample is not the same as the corresponding distribution in the population. In case density and length biased sampling are neglected, the empirical results will generally be biased (or difficult to interpret). Thus, in the derivation of the likelihood function of the durations of stay, these sampling consequences have to be taken into account.

In the current paper, we derive the likelihood that can be used to analyse consistently bivariate duration observations that are stock sampled, while we also state the conditions needed. This method is in line with the estimation method proposed by Chesher and Lancaster (1983) who analyse univariate duration models.

For example, a stock sample may contain individuals who are employed, whereas a flow sample may contain individuals who become employed.
and by Van den Berg et al. (1996) who analyse bivariate duration models for the special case that one type of duration is stock sampled, while the other type is flow sampled.

The analysis of joint durations of stay in two states generally starts from an assumption on the parametric form of the bivariate duration distribution. Currently, there are a wide range of bivariate duration distributions available, which differ regarding the assumptions on the relationship between the two types of durations (see Hougaard, 1987). We consider here bivariate duration models in which the dependence between the two durations is based on stochastically related unobserved components (see Lindeboom and van den Berg, 1994). In the social sciences, this seems a ‘natural’ way of analysing bivariate durations, because not all explanatory variables of heterogeneous agents are observed, the so-called unobserved heterogeneity. In case the unobserved explanatory variables of two single durations are correlated, the two durations are statistically dependent.

We will now derive the joint duration distribution for stock sampled observations while allowing for length and density biased sampling (see, similarly, Van den Berg et al., 1996). To derive functional expressions we assume that individuals are randomly sampled at a fixed moment, say time 0. The bivariate duration data are supposed to be generated as follows. Let the two types of states of interest be denoted by \( S_1 \) and \( S_2 \). So, individuals may occupy two types of states simultaneously. For convenience, we start with assuming in this section that the complete time of being in a certain state is observed. So, two types of durations \( T_1 \) and \( T_2 \) are observed, viz. stock sampled, at time 0; as a consequence, one observes only durations that initiate before 0 and leave the state after 0. The moment an individual enters one of the states of interest is called the entry date. The entry dates \(-p_1\) and \(-p_2\) are distributed with joint density \( q(-p_1, -p_2)\), which is usually called the entry rate. So \( p_1 \) and \( p_2 \) are the observed elapsed durations in states \( S_1 \) and \( S_2 \) at time 0. The complete durations \( t_1 \) and \( t_2 \) are distributed with joint density functions \( f(t_1, t_2)\). The marginal density functions are denoted as \( f(t_1) \) and \( f(t_2) \), respectively. We denote the corresponding distribution functions with \( F \) and the expectation with \( E \). Later on we will see that the likelihood function of the observed joint durations \( t_1 \) and \( t_2 \) is based on the joint distribution of the durations and the joint entry rate.

Suppose that in the population the joint distribution of \( t_1 \) and \( t_2 \), \( f(t_1, t_2) \), depends on the (time-varying) observed variables \( x \) and on unobserved variables \( v_1 \) and \( v_2 \). Moreover, we assume that the durations are independently distributed, given all (observed and unobserved) explanatory variables:

\[
f(t_1, t_2 \mid v_1, v_2, x) = f(t_1 \mid v_1, x) \cdot f(t_2 \mid v_2, x).
\]

So, information about the unobserved variable \( v_2 \) cannot improve the statistical description of \( t_1 \), given the information about \( v_1 \) and \( x \). Similarly, the unobserved variable \( v_1 \) cannot improve the statistical description of \( t_2 \), given the information about \( v_2 \) and \( x \). In other words, the bivariate durations are conditionally independently distributed, the condition being the observed and unobserved explanatory variables. Given these assumptions, identifiability is generally not guaranteed. Therefore, we restrict the analysis to the class of mixed proportional hazard (MPH) models for which nonparametric identifiability can be proven (Honore, 1993). For this class of models, the hazard rates \( \theta(t_1 \mid v_1, x) \) and \( \theta(t_2 \mid v_2, x) \) corresponding to \( f(t_1 \mid v_1, x) \) and \( f(t_2 \mid v_2, x) \), respectively, satisfy the following restriction:

\[\]
\[
\theta(t_1 \mid v_1, x) = \eta(t_1) \cdot \omega(x, \beta_1) \cdot v_1,
\]
\[
\theta(t_2 \mid v_2, x) = \eta(t_2) \cdot \omega(x, \beta_2) \cdot v_2,
\]
where \( \beta_1 \) and \( \beta_2 \) are the parameters.  

The distribution of the unobserved explanatory variables of the individuals in a stock sample is not equal to the distribution of the unobserved explanatory variables in the population, because those with an unobserved high probability of leaving the states of interest are less likely to be sampled. As a consequence, an essential distinction has to be made between the distribution of variables in the stock and in the population. Let the distribution of (observed and unobserved) variables in the stock of individuals in \( S_1 \) and \( S_2 \) be denoted by \( h_s \) and the same distribution of the population by \( h_p \). We denote the mixing or heterogeneity distribution as \( h_p(v_1, v_2, x) \) which characterises the observed and unobserved heterogeneity in the population. We proceed by assuming that the observed explanatory variables in the population are independent of the unobserved explanatory variables in the population:

\[
h_p(v_1, v_2, x) = h_p(v_1, v_2) \cdot h_p(x).
\]

Note that this assumption does generally not imply that the observed explanatory variables in the stock are independent of the unobserved explanatory variables in the stock. We will make clear that such an independence condition is however needed for a feasible maximum likelihood (ML) estimation procedure.

When we observe an individual, we have to take into account that an individual is chosen from the population in the stock (see, similarly, Ridder, 1984). The individual has entered \( S_1 \) and \( S_2 \) at arbitrary \( p_1 \) and \( p_2 \) times before 0, and has survived in \( S_1 \) and \( S_2 \) for at least a duration of \( p_1 \) and \( p_2 \), respectively. Hence, at time 0, the joint distribution \( h_s(p_1, t_1, p_2, t_2, v_1, v_2, x) \) of a selected individual with characteristics \( p_1, t_1, p_2, t_2, v_1, v_2 \) and \( x \) in the stock equals:

\[
h_s(p_1, t_1, p_2, t_2, v_1, v_2, x) = \int_x \int_{v_1} \int_{v_2} h_p(v_1, v_2, x) q(-p_1, -p_2 \mid v_1, v_2, x) f(t_1, t_2 \mid v_1, v_2, x) \, dp_1 \, dp_2 \, h_p(v_1, v_2, x) \, dv_1 \, dv_2 \, dx.
\]

The denominator of (1) can be interpreted as the weighted expected value of \( t_1 \cdot t_2 \). This can easily be seen if one realises that

\[
\int_{p_1} \int_{p_2} F(p_1, p_2) \, dp_1 \, dp_2 = E(t_1 \cdot t_2).
\]

So the denominator of formula (1) extends the well-known result derived for univariate duration observations for which it holds that the probability of sampling an individual in a certain state is proportional to the expected duration of being in that state.

3. Conditional likelihood methods

We propose to analyse the stock sampled bivariate duration observations \( t_1 \) and \( t_2 \), conditional on the elapsed durations \( p_1 \) and \( p_2 \) and the observed explanatory variables. These methods are called conditional
likelihood methods. This contrasts the use of joint likelihood methods that do not condition on \( p_1 \) and \( p_2 \). Several authors have discussed the advantages and disadvantages of using conditional likelihood methods instead of joint likelihood methods (e.g. Ridder, 1984; Van den Berg et al., 1996; Van Ommeren, 1996). One advantage of conditional methods is that they need only very weak assumptions, allowing the entry rates to depend on calendar time. In particular, in the case of employment observations, the advantages outweigh the disadvantage, as it has been shown that job entry rates depend heavily on calendar time due to business cycle effects (see, for example, Van Ommeren et al., 1996; Imbens, 1994). Another advantage is that the functional form of \( f(t_1 \mid v_1, x) \) and \( f(t_2 \mid v_2, x) \) can be freely chosen, whereas joint likelihood methods need to restrict the functional form of the \( f(t_1 \mid v_1, x) \) and \( f(t_2 \mid v_2, x) \) (Van den Berg et al., 1996; Van Ommeren, 1996). Nevertheless, conditional methods are less efficient, as the information of the elapsed durations on the parameters of interest is lost (Ridder, 1984). Moreover, conditional methods require information about the elapsed durations, whereas not all surveys contain this information. In this section we will focus on the typical case that the observations of the elapsed durations are complete.

To keep the relation between \( t_i \) and \( p_i \) clear, we introduce \( r_i \), the residual (remaining) duration of type \( i \), so that \( r_i = t_i - p_i \) (\( i = 1, 2 \)). Given information on \( p_1 \) and \( p_2 \), we use the following (weak) condition that allows the entry rate to depend freely on the elapsed time:

\[
q(-p_1, -p_2 \mid v_1, v_2, x) = k_0(p_1, p_2, x) \cdot k_1(v_1, v_2).
\]

This condition allows \( q \) to depend freely on \( p_1 \) and \( p_2 \). The contribution to the likelihood \( h_s(t_1, t_2 \mid p_1, p_2, x) \) can then be written as

\[
h_s(t_1, t_2 \mid p_1, p_2, x) = \frac{\int_{v_1} \int_{v_2} f(t_1 \mid v_1, x) \cdot f(t_2 \mid v_2, x) \cdot h_f(v_1, v_2) \, dv_1 \, dv_2}{\int_{v_1} \int_{v_2} \overline{F}(p_1 \mid v_1, x) \overline{F}(p_2 \mid v_2, x) h_f(v_1, v_2) \, dv_1 \, dv_2},
\]

where \( h_f(v_1, v_2) \) is defined as

\[
h_f(v_1, v_2) = \frac{k_1(v_1, v_2) \cdot h_p(v_1, v_2)}{\int_{v_1} \int_{v_2} k_1(v_1, v_2) \cdot h_p(v_1, v_2) \, dv_1 \, dv_2}.
\]

Note that \( h_f(v_1, v_2) \) does not depend on \( x \), as \( k_0(p_1, p_2, x) \) cancels from the denominator and numerator of \( h_s(t_1, t_2 \mid p_1, p_2, x) \). Hence, given the choice of a mixing function that approximates \( h_f(v_1, v_2) \), Eq. (2) can be used as a basis for the construction of the likelihood function of \( t_1, t_2 \) conditional on \( p_1, p_2 \) and \( x \). This facilitates the use of ML estimation. As \( h_f(v_1, v_2) \) is an unknown function of \( v_1 \) and \( v_2 \), the choice of a flexible mixing function is recommended to avoid biases. Discrete distributions that allow the masspoints and the corresponding probabilities freely chosen are recommended (see also Nickel, 1979; Ham and Rea, 1987; Heckman and Singer, 1984; Van den Berg et al., 1996).

In conclusion, we have derived the distributions that will enable us to obtain empirical estimates of \( f(t_1 \mid v_1, p_1, x) \), \( f(t_2 \mid v_2, p_2, x) \) and \( h_f(v_1, v_2) \) by means of ML when the observed durations are gathered by means of stock sampling.

4. An empirical application: Job durations of two-earner households in the Netherlands

4.1. Data

The data set used here (called Telepanel), collected in 1992–1993, contains the complete life course pattern of about 3000 Dutch respondents, including the labour career patterns. The data were collected in a retrospective way, mainly for marketing purposes. The data set allows for a distinction between voluntary
moves and involuntary job moves (due to firing). From this data set, we were able to select 57 two-earner households for which all relevant data are observed. By construction, every two-earner household consists of two wage earners of different genders, so we were able to observe 57 males and 57 females. These wage earners worked at least 20 hours per week in the period between 1985 and 1991. On the basis of this data set, we were able to follow these households over time between January 1985 and December 1991 and to observe the job durations of both wage earners. After a job move, the household continues to be included in the analysis, so that we have multiple job duration observations. Within the period of observation, we could observe 46 job moves. In case a worker becomes nonemployed or the household dissolves, the job duration is censored. We will now present the corresponding statistical model.

4.2. The likelihood function

To assess the statistical relationship between the unobserved (person-specific and time-invariant) variables of the job mobility rates of workers belonging to the same household, we aim to estimate a bivariate duration model, given a parametric discrete mixing distribution. In order to avoid overload of technical details we give a rather simple model formulation here. Extensions of the model will be discussed in Section 4.5.

We assume that the job hazard function is mixed exponential, which does not allow for duration-dependency. This assumption implies that job duration \( t_i \), given \( v_i \) and \( x \), is distributed with hazard \( \theta(t_i \mid v_i, x) = v_i \cdot \exp(x \cdot \beta_i), \ i = 1, 2 \). So, the density function and the survival function can therefore be written as

\[
 f(t_i \mid v_i, x) = v_i \exp(x \cdot \beta_i)[1 - \exp^{-t_i \exp(x \cdot \beta_i)}]
\]

and

\[
 F(t_i \mid v_i, x) = 1 - \exp^{-t_i \exp(x \cdot \beta_i)},
\]

respectively.

Given the choice of a conditional likelihood method, formulas (2) and (3) may be used as a basis for the empirical analysis. The likelihood of observations that are right-censored – of which is only known that \( T_i > t_i, \ i = 1, 2 \) – can easily be derived by replacing the densities \( f(t_i) \) by survival functions \( F(t_i) \), \( i = 1, 2 \).

As explained above, the construction of the loglikelihood for time-stationary explanatory variables can be obtained by using formula (2) and given parametric assumptions about \( f(t_i) \) and \( F(t_i) \). However, in our analysis, we wish to incorporate time-varying explanatory variables, since during the period of observation some explanatory variables will certainly change. In our specification, we allow the explanatory variables to change annually (episode splitting). This can be established by rewriting the densities of the durations as a multiplication of annually sampled observations.

In order to construct the likelihood function, we integrate over the unobserved disturbance \( v_1 \) and \( v_2 \), by using the mixing distribution \( h(v_1, v_2) \). The likelihood function \( L \) of \( N (N = 57) \) households followed during \( Y \) years (\( Y = 7 \)) can then be written as follows (\( i = 1, \ldots, Y; \ j = 1, \ldots, N \)):

\[
 L = \prod_{i=1}^{Y} \prod_{j=1}^{N} \int h(v_1, v_2) f(t_i \mid v_1, x) F(t_i \mid v_2, x) \cdot h(v_1, v_2) dv_1 dv_2
\]

We cannot distinguish between voluntary job-to-job moves and voluntary job-to-nonemployment moves.

The number of two-earner households for which all relevant data are observed appeared to be rather small due to missing data in the interview stage, in particular on the work place. Since we wanted to include commuting distance as an explanatory variable, approx. one half of the workers belonging to a two-earner household could be included in our data set.
\[
L = \prod_{j=1}^{N} \int_{t_1}^{t_2} \int_{t_1}^{t_2} \prod_{i=1}^{J} \left[ f(t_{1ij}) \frac{1 - \text{cen}_{1}(ij)}{\bar{F}(t_{1ij})} f(t_{2ij}) \frac{1 - \text{cen}_{2}(ij)}{\bar{F}(t_{2ij})} \right]^{\text{sam}_{ij}} h(v_1, v_2) \, dv_1 \, dv_2,
\]
where \( \text{sam}_{ij} = 1 \), if household \( j \) is sampled in year \( i \), otherwise 0; \( \text{cen}_{1}(ij) = 1 \), if the spell of the first wage earner of household \( j \) in year \( i \) is right-censored, otherwise 0; \( \text{cen}_{2}(ij) = 1 \), if the spell of the second wage earner of household \( j \) in year \( i \) is right-censored. Hence, the loglikelihood of \( t_1, t_2 \), given time-varying hazard rates, can simply be rewritten as the product of stationary conditional densities and survival functions. For further details on this approach, we refer to Lancaster (1990).

The choice of the mixing distribution is based on computational and theoretical considerations. One may approximate \( h_f(v_1, v_2) \) – see formula (3) – by a mixing distribution with four discrete masspoints; so both \( v_1 \) and \( v_2 \) have two points of support (\( v_1 \) has \( v_{11} \) and \( v_{12} \), and \( v_2 \) has \( v_{21} \) and \( v_{22} \)) with probabilities \( P_{11}, P_{12}, P_{21} \) and \( P_{22} \), respectively. So

\[
\begin{align*}
P_{11} &= P(v_1 = v_{11}, v_2 = v_{21}), \\
P_{12} &= P(v_1 = v_{11}, v_2 = v_{22}), \\
P_{21} &= P(v_1 = v_{12}, v_2 = v_{21})
\end{align*}
\]

and

\[
P_{22} = P(v_1 = v_{12}, v_2 = v_{22}).
\]

The masspoints and probabilities of the discrete mixing function can then be estimated (see Van den Berg et al., 1996). In the case of unobserved variables related to the characteristics of workers belonging to the same two-earner households however, it makes sense to assume that the mixing distribution is symmetric. Hence, we suppose that \( v_{11} = v_{21} = u_1 \) and \( v_{12} = v_{22} = u_2 \). Variations on this approach are possible. In Section 4.5 we discuss the effects of using more general formulations such as asymmetry instead of symmetry, three versus two mass points, and stationary versus time-dependent hazard rates.

As hazard functions cannot be negative, we restrict \( u_1 \) and \( u_2 \) to be positive by replacing \( u_1 \) and \( u_2 \) by \( \exp(u_1) \) and \( \exp(u_2) \), respectively. This does not affect the results in any way. In addition, \( P_{11} \) is specified by

\[
\exp(\gamma_1) / [1 + \exp(\gamma_1) + \exp(\gamma_2) + \exp(\gamma_3)],
\]

\( P_{12} \) is specified by

\[
\exp(\gamma_2) / [1 + \exp(\gamma_1) + \exp(\gamma_2) + \exp(\gamma_3)],
\]

\( P_{21} \) is specified by

\[
\exp(\gamma_3) / [1 + \exp(\gamma_1) + \exp(\gamma_2) + \exp(\gamma_3)]
\]

and \( P_{22} \) is specified by

\[
1 / [1 + \exp(\gamma_1) + \exp(\gamma_2) + \exp(\gamma_3)].
\]

Hence, the mixing distribution can be estimated by freely estimating \( \gamma_1, \gamma_2 \) and \( \gamma_3 \), while guaranteeing that the probabilities \( P_{11}, P_{12}, P_{21} \) and \( P_{22} \) are between 0 and 1. The standard deviations of these probabilities are calculated using the delta method. Since we assume that the mixing distribution is symmetric, we use the condition that \( \gamma_1 = \gamma_2 \).

The model has been estimated using a maximum likelihood procedure using the program Gauss. Since the loglikelihood does not have a unique maximum (see Lancaster, 1990), we have estimated the model several times with different starting values for the coefficients.
4.3. The explanatory variables

To analyse job mobility, we have used a range of relevant explanatory variables. Older persons are expected to move less often in the labour market (see, for example, Van den Berg, 1992). Job-to-job mobility is thought to increase with higher educational achievement, because higher education offers more opportunities for upward mobility (Hartog and van Ophem, 1994). Therefore, we include a dummy for workers with an academic degree. Not only formal education, but also the position within the firm affects job mobility: we incorporate, therefore, the number of subordinates and whether or not the person works more than 32 hours per week (Lindeboom and Theeuwes, 1991).

Next, we also suppose that the size of the branch of the firm affects job mobility. Size of branch is defined as the number of persons working at the same workplace location of the firm. This variable is a proxy for the size of the firm. It is generally accepted that larger firms offer more opportunities to grow within the firm and offer better employment conditions, which reduce workers’ job mobility (Barron et al., 1987). It is generally thought that those who work as a civil servant will move less (see Van Ommeren, 1996) and hence we include also a dummy to capture such an effect.

It is assumed that workers prefer to earn a higher wage, so that the hourly wage rate is used in the analysis. We follow the literature by employing the logarithm of the wage in job mobility studies to take into account that the marginal effect of the wage is smaller when the wage is higher (see, for example, Gronberg and Reed, 1994). Household characteristics are captured by including a dummy for the presence of children and a dummy for the tenure of the residence (owner versus renter).

We employ also information on the commuting distance and the commuting distance of the spouse. Since only data on the municipalities of residence and workplace of the individuals are available, we approximate commuting distance by the distance between the centres of the municipalities (measured in kilometres). We expect that a longer commuting distance increases job mobility (see, for example, Van den Berg, 1992). The effect of the present wage on residential mobility is expected to be negative (see, for example, Viscusi, 1980; Robinson, 1990; Van den Berg, 1992). Also, the ratio of the worker’s wage over the spouse’s wage is included. It is hypothesised that a higher wage of the spouse makes a job-to-job move and a job move into nonemployment more likely, since the spouse’s wage diminishes the effect of the risk associated with moving job. As indicated by for example Viscusi (1980) and Van Ommeren et al. (1998a,b), gender may also play a role as a determinant of labour mobility. The time-varying variables are allowed to differ yearly. This statistical model has next been estimated on the basis of our Dutch sample.

4.4. Empirical results

The special feature of this paper is that it addresses job mobility of workers belonging to a two-earner household while allowing for mutual dependence between the unobserved variables. To evaluate the implications of this approach we compare it with analyses that either ignore unobserved variables or ignore the mutual dependence between unobserved variables. So, before we interpret the estimates based on a model using the bivariate mixing function (model III), we will present the results based on a model using a univariate mixing function (model II) and the results based on a model which excludes unobserved variables (model I). In Table 1, the results for models I, II and III are reported.

---

9 The survey does not contain information on commuting time. This is unfortunate, since it seems likely that commuting time has a stronger effect on job search behaviour than commuting distance (Van Ommeren, 1996). Similarly, Dubin (1991) has shown for the United States that workers react stronger to changes in commuting time than to changes in commuting distance.
Let us first focus on the results of model II. It appears that the difference between the two masspoints is 2.27 with a standard deviation of 0.88. Hence, the two masspoints are significantly different and one may conclude that unobserved variables do cause the workers to move job. The results indicate that about 1% of the persons have about 10 times lower hazard rates that are not explained by observed explanatory variables included in the analysis ($\exp^{-2.27} = 0.10$). The difference in the loglikelihood between model I and model II is 6.02. The value of a Likelihood Ratio test is therefore 12.04 and the hypothesis of no unobserved variables is rejected at a 1% significance level ($\chi^2(2) = 9.21$). Hence, we conclude that model II, which allows for unobserved variables, is statistically superior to model I, which ignores unobserved variables.

Let us now focus on the results of model III. In this section we present results for the base case. Variations on the statistical formulation of the model will be discussed in the following section. As ex-
plained in Section 4.2, the bivariate mixing distribution is estimated by estimating \( c_1, c_2 \) and \( c_3 \). It appeared that \( c_1, c_2 \) and \( c_3 \) were rather large so that \( P_{22} \) was practically equal to zero (smaller than 0.00001). Furthermore, the coefficients of \( c_1, c_2 \) and \( c_3 \) were not very significant (i.e., with large standard deviations). We have therefore proceeded by fixing \( P_{22} \) equal to zero. The results associated with the assumption that \( P_{22} \) is equal to zero can be found in Table 1 (model III). Given the estimates, we have also calculated the correlation between the unobserved variables.\(^{10}\)

The difference in the loglikelihood between model II and model III is 4.14. The value of a Likelihood Ratio test is therefore 8.28 and the validity of the model II is rejected at a 1% significance level (\( \chi^2(1) = 6.63 \)). Hence, we conclude that the model that allows for the mutual dependence between the unobserved variables is statistically superior to the model that ignores this dependence. It is remarkable that – despite the statistical superiority of model III – the correlation between the unobserved variables appears to be insignificant, a finding also obtained by Butler et al. (1989).\(^{11}\)

As can be seen from Table 1, the model that allows for mutual dependence between the unobserved variables does not give substantially different results for the coefficients of the explanatory variables when compared to the models which ignore this dependence (compare the results of models I, II and III in Table 1). However, the estimated effects of the explanatory variables on job mobility show some variation and the estimates are more precisely estimated when using model III. In particular, the coefficients of the explanatory variables size of the branch, the number of subordinates, commuting distance and full-time employed appear to be significant according to model III, in contrast to model I.

In conclusion, we arrive at the interesting finding that the use of a bivariate mixing distribution confirms largely the results of less ambitious models, but the precision of the estimates is improved. Interestingly enough, we find similar conclusions in the studies of Butler et al. (1989) and Van den Berg et al. (1996).

We will now focus in greater detail on the effect of the explanatory variables. We find here empirically that older persons move less often in the labour market (this is in conformity with Section 4.3). Also in line with the theoretical considerations we find that employees working for a larger branch move less often. For the number of subordinates a nonmonotone pattern is found.\(^{12}\)

Furthermore, we find that a longer commuting distance increases job mobility (this is consistent with Section 4.3). However, the effect of the spouse’s commuting distance is less clear, according to model III (but present according to models I and II).\(^{13}\) In line with most labour mobility studies, the effect of the wage is negative. According to our results, there are no gender effects and no civil servant effects (both in contrast to the literature mentioned in Section 4.3). It may be argued that most likely the data set at hand is too small to detect specifically these effects.

\(^{10}\) Van den Berg et al. (1996) state the appropriate formula to calculate the correlation given the estimates of the bivariate mixing function.

\(^{11}\) One potential explanation, as suggested by one of the referees, is that the correlation between unobserved variables is restricted to values smaller than one. Potentially, this is a good explanation since the maximum values for the correlation between two discrete variables are often less than one. In particular, the maximum value of the correlation between two Bernoulli random variables with parameters \( p_1 \) and \( p_2 \), respectively, is less than one except when \( p_1 = p_2 \). For example, when \( p_1 = 0.1 \) and \( p_2 = 0.5 \), the maximum correlation is 1/3. This explanation is however not valid here, since we have assumed that the bivariate mixing distribution is symmetric, so the maximum value of the correlation is still one. Finally, it should be noted that the correlation between the duration variables is restricted to values between \(-1/3\) and \(1/2\); see Van den Berg and Steerneman (1991).

\(^{12}\) Those who have zero subordinates have the lowest job mobility, workers with 1–3 subordinates have the highest job mobility and those with more than three subordinates take an intermediate position. The significance varies among the specifications.

\(^{13}\) We refer here to Van Ommeren et al. (1998a), who have examined a theoretical search model of workers who belong to a two-earner household. According to this model, the relationship between spouse’s commuting distance and job mobility is ambiguous.
As expected, we find that a higher education increases job moving behaviour. Interestingly, housing and household effects appear to be absent: we do not find significant effects for residential tenure or for the presence of children. Finally, it is noteworthy (Section 4.3) that the effect of the ratio of the worker’s wage over the spouse’s wage is absent.  

Our general conclusion is that the estimates of the effects of most explanatory variables have signs that are in agreement with other empirical studies that do not explicitly focus on two-earner households. As far as we know, new findings in our analysis are that the spouse’s wage and the spouse’s commuting distance have no effect on job mobility rates of workers.

Finally, we will pay some attention to the marginal willingness to pay for certain job attributes, in particular commuting distance. Gronberg and Reed (1994) have in general shown that the marginal willingness to pay for a job characteristic can be derived from duration data on job mobility. Van Ommeren et al. (1997) have extended their result in order to estimate the marginal willingness to pay for commuting distance.  

\[
\log(\text{wage}) = b_w + b_d \text{commuting distance},
\]

where \(b_w\) is the coefficient of the wage and \(b_d\) the coefficient of the commuting distance.  

Hence, the marginal willingness to pay for commuting distance is proportional to the current wage.

According to model I, the marginal willingness to pay for commuting distance (measured in 10 km, one-way) is \(-0.88\%\) of the hourly wage, whereas its standard deviation is \(0.83\%\). According to model II, the marginal willingness to pay for commuting distance is \(-1.23\%\), whereas its standard deviation is \(0.55\%\). According to model III, the marginal willingness to pay for commuting distance is equal to \(-1.22\%\), whereas its standard deviation is \(0.54\%\). Thus, according to models II and III, the willingness to pay for commuting is significant at a 5\% significance level. Note that the marginal willingness to pay for commuting distance is negative, so one should interpret these results as the marginal willingness to pay to avoid commuting distance.

Assuming that a person works for eight hours a day and travels at a speed of 50 km per hour in both directions, this implies that the marginal willingness to pay for the absence of commuting time is about \(87\%\) of the hourly wage. This percentage is considerably higher than those reported by most other studies that generally find estimates around 30–40\% (see, Small, 1992).  

It should be noted that it is not implausible that a wage earner belonging to a two-earner household has a higher marginal willingness to pay to avoid commuting time than a single wage-earner, since the latter worker is more flexible regarding a residential move and will be able to reduce the commuting time by means of a residential move.

---

14 The spouse’s commuting distance is significant in models I and II, but in the more advanced model the effect is no longer significant.
15 Clearly, commuting distance – in a strict sense – is not an ordinary job characteristic, as the distance can be changed through a residential move.
16 The variance of the estimated MWP is derived using the delta method. So, \(\text{var}(b_d/b_w)\) is calculated as \(\text{var}(\beta_d/\beta_w)^2 \cdot \text{var}(\beta_w) - 2 \cdot \text{cov}(\beta_d, \beta_w)/\beta_w^2\).
17 This number has been calculated as \(\beta_d/\beta_w \times H \times S/2\), where \(H\) is the number of hours per day and \(S\) is the average speed. Note that we have to divide by 2, since the distance has to be travelled twice a day. Furthermore, note that the distance has been measured in 10 km.
18 Assuming a higher average speed, the implied marginal willingness to pay for commuting time is higher.
19 In Van Ommeren et al. (1998b), we have estimated separate models for a set of workers who belong to a two-earner household and for a set of workers that also consists of single wage-earners. The results given in Van Ommeren et al. (1998b) also imply that the marginal willingness to pay for commuting distance is higher for workers who belong to a two-earner household than for single wage-earners.
4.5. The specification revisited

It goes without saying that the estimation results from Section 4.4 take for granted the plausibility of several assumptions made. The results presented above rely on a range of assumptions, which need to be critically reviewed.

First, we have in our analysis originally assumed a time-stationary hazard rate, implying an exponential distribution for the job durations. As an alternative, we have decided to consider also a nonstationary hazard rate, by assuming a Weibull distribution. This implies that the hazard rate $h(t_i | v_i, x)$ can be written as $\alpha t_i^{\alpha-1} v_i \cdot \exp(x \cdot \beta_i)$, $i = 1, 2$. Note that when $\alpha = 1.00$, the Weibull distribution is identical to the exponential distribution. Estimation of the model assuming a Weibull distribution showed next that $\alpha$ equals 0.92 with a standard deviation of 0.17. Hence, the hypothesis that $\alpha = 1.00$ is not refuted. So, there is no reason to reject the assumption that job mobility is not duration dependent.

Second, we have initially assumed that the mixing distribution can be described by means of two masspoints. What are the consequences if a third masspoint is assumed? Including a third masspoint shows that the third masspoint takes on the same value as one of the other two masspoints. Hence, the assumption on the discrete mixing distribution seems valid.

Third, in our initial estimation procedure we had taken for granted that the bivariate mixing distribution is symmetric, by assuming that $v_{11} = v_{12}$, $v_{21} = v_{22}$ and $\gamma_1 = \gamma_2$. We have re-estimated therefore the model relaxing this assumption and tested the restrictions using a standard Likelihood Ratio test. We found empirically that the restrictions are still valid at a 5% significance level and that the point estimates hardly change. In particular, the difference between the value of the masspoints is almost identical for males and females ($v_{11} - v_{12} = -2.23$; $v_{21} - v_{22} = -2.19$).

And finally, we have re-estimated the model while changing the set of the explanatory variables. In particular, we have used the wage of the spouse instead of the ratio of the wage over the spouse’s wage. Again, we found no significant effect. In addition, we have added a variable which measures the distance between the workplaces of worker and the spouse’s workplace. The expected effect of the latter variable is positive according to search theory (see Van Ommeren et al., 1998a). The estimated coefficient showed the right sign but it was not significant.

In conclusion, the statistical results of our bivariate mixing function model are rather satisfactory and also fairly robust vis-à-vis other specifications of the model.

5. Conclusion

In many studies on duration of stay of individuals, it has been found that unobserved variables play a crucial role. The study of joint duration distributions in which the dependence is induced through mixing has resulted in various theoretical and empirical results in the literature. In the present paper, we have analysed stock sampled durations observations. In particular, we have investigated the consequences of the inclusion of unobserved variables that affect a worker’s job mobility jointly with the unobserved variables that affect the job mobility of the worker’s spouse. This is done by using a bivariate mixing distribution.

---

20 In the United States, the Weibull distribution has been rejected as a valid distribution of the job duration. For the Netherlands however, various studies have found the exponential distribution to be adequate (see, Lindeboom and Theeuwes, 1991; Van den Berg, 1992).
21 Also in other contexts it appears that one or two masspoints suffice when estimating the mixing function (see, Van Ommeren, 1996, Chapter 7; Van Ommeren et al., 1998a,b; Van den Berg et al., 1996).
The main conclusion is that the differences compared to the outcomes of a model ignoring heterogeneity and a model employing a univariate or bivariate mixing distribution are not substantial. However, although the estimated effects of the explanatory variables on job mobility are not strongly affected by allowing for a mutual dependence between the unobserved variables, the effects of the explanatory variables have been estimated more precisely by allowing for this dependence. Thus, the model developed and estimated here is statistically superior.

As a by-product of the study on job moving behaviour of workers belonging to two-earner households, we find that the spouse’s commuting distance and the spouse’s wage do not affect drastically job mobility. In addition, we have obtained estimates of the marginal willingness to pay for the avoidance of commuting distance. We arrived at the interesting finding that the implied estimate of the marginal willingness to pay for a reduction in commuting time in two-earner households is higher than those reported in the literature for single-earner workers. Such a finding is consistent with the idea that workers belonging to two-earner households are less flexible in choice situations on the housing market than single wage-earners.

Acknowledgements

This research was sponsored by the Economic Research Foundation, which is part of the Netherlands Organisation for Scientific Research (NWO).

References