

## **A flexible procedure for analysing longitudinal event histories using a multilevel model.**

Harvey Goldstein<sup>1</sup>, Huiqi Pan<sup>2</sup> and J. Bynner<sup>3</sup>

### **Abstract**

Event history or survival models are applicable to outcomes that are measures of duration, for example the length of employment periods or times to death after medical treatment. When individuals are grouped within institutions such as firms or clinics the resulting multilevel structure also needs to be incorporated into the model. An important application is where individuals are the ‘higher level’ units and they experience repeated durations, such as lengths of partnerships. In this paper we show how such repeated measures data can be modelled using a flexible discrete time event history model that incorporates individual level random effects. The model is applied to the analysis of partnership episodes for adult members of the National Child Development Study followed up between the ages of 16 and 33. The exposition will not assume a detailed knowledge of event history modelling.

### **Keywords**

Event history model, longitudinal data, multilevel model, repeated measures.

### **Acknowledgements**

We are very grateful to Fiona Steele and referees for helpful comments.

*1, 3. Bedford Group, Institute of Education, University of London*

*2. Institute of Child health, University of London*

*email: [h.goldstein@joe.ac.uk](mailto:h.goldstein@joe.ac.uk)*

## **Introduction**

Event history or survival models are applicable to outcomes that are measures of duration, for example the length of employment periods or times to death after medical treatment. When individuals are grouped within institutions such as firms or clinics the resulting multilevel structure also needs to be incorporated into the model. An important application is where individuals are the 'higher level' units and they experience repeated durations, such as lengths of partnerships. To make matters more concrete consider a young adult who enters into a partnership, whether or not legalised by marriage, at age 20 years. Suppose that this partnership lasts for 5.5 years and is followed by a non-partnership period of 1 year. This is then followed by another partnership of 2.3 years and a non-partnership period of 2.7 years, then a final partnership whose end point is unknown; either because the person is lost to the study investigator or because at the time of analysis they are still in the partnership. Here there are two 'states', partnership or non-partnership, an event is the transition between states and an episode is the period of being in a particular state. For this person there is some variation in partnership lengths. There is also an apparent difference in the mean time spent in a partnership and the mean time spent out of partnership.

If we now consider a sample of individuals followed up in a similar way we will not only have these kinds of 'within-individual' variation of partnership durations, we will also expect to have a variation 'between-individuals' in the mean length of partnership and non-partnership. The existence of both within - and between - individual variation constitutes a 2-level structure, in this case akin to a standard repeated measures design such as exhibited by individual growth data where body measurements are taken

repeatedly on the same individuals (see Goldstein, 2003). In addition we would expect average durations to be a function of individual characteristics such as social class and educational background.

A large literature on the modelling of event history data exists and a good introduction is the book by Singer and Willett (2003). Nearly all of this, however, deals with single level data, typically where each individual has a single event duration. Goldstein (2003) discusses various generalisations of the various single level approaches to the multilevel case and in this paper we describe just one of these, the so called 'discrete time event history model' that is flexible and requires only a relatively straightforward adaptation of existing methods for handling 2-level data: the data are rearranged so that the 'response' variable is binary (see below). Note also that in the individual example given above there is 'censoring' of the final episode where we do not know the actual duration. Censored data pose particular problems for event history data and we shall describe how these are dealt with later.

This paper provides an exposition of the use of these models with data from the National Child Development Study on partnership durations (Bynner et al., 2002). We are interested in the duration of two states – not being in a partnership and belonging to a partnership. We have chosen as explanatory variables, age at the start of each duration episode (partnership), and social class based upon own occupation since both are known to be associated with the age at which partnerships begin and the number that occur (Bynner et al., 2002).

## Methodology

We begin with a simple 2-level data structure where we have a sample of individuals (level 2 units) and episodes (level 1 units) nested within individuals: a ‘repeated measures’ design. Because duration length distributions are generally positively skewed (negative durations are impossible), we would typically work with the logarithm of the duration length ( $t$ ) as our response. We denote the response by  $y_{ij}$ , where  $i$  indexes the episode and  $j$  the individual. For now we assume also that there is no censoring.

We can classify predictor (covariate) variables into two kinds. The first are characteristics of an individual such as their gender which do not change over time and for such a covariate we use notation of the form  $x_{nj}$  ( $n = 1, 2, \dots$ ) since it is the same for each episode within an individual. The second are ‘time-varying’ effects that will change over time, such as the age at which an episode starts. In this case we will use notation of the form  $x_{nij}$  ( $n = 1, 2, \dots$ ) to indicate that the covariate may change its value from episode to episode within an individual. The modelling procedure will handle both kinds. In addition, as in a standard multilevel framework, each individual will have their own ‘random effect’ (or set of ‘effects’). These can be thought of as describing the mean duration lengths for each individual and they are typically assumed to have a Normal distribution across individuals. We will use the notation  $u_j$  for these and they are described in more detail below. For modelling purposes we assume that, given individual characteristics, any time-varying effects, and the individual level random effects, then the durations for each state are independently distributed. We denote these by  $e_{ij}$ ; they are also termed the level 1 residuals. In other words, within any given individual, once we have described in our

model the average duration length for a particular state, these represent the random variation about that mean value and the successive duration lengths are independently distributed according to some suitable distribution. Before we describe how the data are organised into a suitable form for analysis, we look at the specification of a simple model based on the above assumptions.

Such a simple *2-level* repeated measures model can be written as

$$\begin{aligned}
 y_{ij} &= \beta_0 + \beta_1 x_{1ij} + u_j + e_{ij} \\
 y_{ij} &= \log(t_{ij}), \quad u_j \sim N(0, \sigma_u^2), \quad e_{ij} \sim N(0, \sigma_e^2)
 \end{aligned}
 \tag{1}$$

where  $\beta_0$  is the overall intercept term,  $\sigma_u^2, \sigma_e^2$  respectively refer to the between-individual and between-episode variance, and where we have a single covariate such as age at the start of an episode,  $x_{1ij}$ . For some situations model (1) is perfectly adequate, but problems occur when there is censoring since then the distributional assumption of Normality for the  $e_{ij}$  is important and different assumptions will lead to different estimates. In many cases it will not be easy to establish the most appropriate functional form and for this reason most of the models used for event history data attempt to avoid the use of such strong distributional assumptions. The procedure we now describe does avoid the use of strong distributional assumptions as well as allowing us to incorporate a full multilevel structure.

In discrete time models, instead of modelling the episode duration directly, we divide the time scale into short time intervals: in the case of the partnership data these are 3-months long. We assume that the interval is short enough so that at most one state transition takes place within the interval. The basic data record, which now becomes the lowest level unit

in the multilevel structure, is this time interval and the response is the binary event of whether a transition takes place (=1) or not (=0). The aim of the model, described below, is to predict the probability of a transition as a function of elapsed time from the start of the episode, covariates and random effects. We shall show how this allows us fully to characterise event duration data and to obtain estimates and inferences about covariate effects on duration lengths etc. We refer to the short time intervals as *modelled time intervals* and we discuss below how these are used in the model to account for the elapsed time. To illustrate this structure, data for individual 1 might look as follows:

Individual	Actual time interval (episode)	Modelled time interval	Response	Event state (start of time interval)
1	1	1	0	no partnership
1	2	2	0	no partnership
1	3	3	1	no partnership
1	4	1	0	partnership
1	5	2	0	partnership
1	6	3	0	partnership
1	7	4	1	partnership
1	8	1	0	no partnership
1	9	2	0	no partnership
1	10	3	0	no partnership

Thus, starting in a state of 'no partnership', individual 1 moves in time interval 3 to state 'partnership' and in real time interval 7 and modelled time interval 4 for the partnership episode, to state 'no partnership'. The response variable takes the value zero if no move takes place during an interval and one if a change in partnership status occurs during the interval.

We now set up a model for the probability of moving from one state to another during any given interval. We shall suppose that this depends upon the state that the individual currently is in, the length of time that individual has been in that state, any covariates, and an individual random effect equivalent to the individual random effect term in (1). This probability ( $\pi$ ) is usually referred to as the 'hazard' and we write the probability of moving at modelled time interval  $t$ , that is the probability that the response is a 1, as

$$\pi_{ijk}(t) = P(y_{ijk}(t) = 1 | y_{ijk}(t-1) = 0)$$

where  $k$  indexes individual,  $j$  indexes episode and  $i$  indexes the state. The states (partnership, non-partnership) are modelled by a dummy predictor variable and a simple model for this probability would be

$$\log\left(\frac{\pi_{ijk}(t)}{1 - \pi_{ijk}(t)}\right) = \text{logit}(\pi_{ijk}(t)) = \beta_0 + \beta_1 x_{ijk}(t) + f(t) + u_k + e_{jk} \quad (2)$$

As is conventional for binary responses we use a 'logit' link function for the probability. That is, we work with the natural logarithm of the ratio of the odds that there is a state change, i.e. the ratio of the probability of changing states to the probability of not changing states. The right hand side of (2) has the same standard structure as for a linear multilevel model. The actual binary response ( $y_{ijk}(t)$ ), which is 1 if there is a state change

and 0 if not, is assumed to have a binomial distribution with denominator 1 (a Bernoulli distribution) which we write  $y_{ijk(t)} \sim \text{Binomial}(1, \pi_{ijk(t)})$ .

The terms  $u_k, e_{jk}$  refer to the individual and episode levels respectively,  $x_{1ijk(t)}$  is the dummy variable for the state and  $f(t)$  is a function of the time spent in the current state. In standard multilevel terminology, the component  $\beta_0 + \beta_1 x_{1ijk(t)} + f(t)$  is known as the fixed part of the model and the component  $u_k + e_{jk}$  as the random part.

Clearly, as the time spent in the state increases so the probability of leaving that state increases also. Since time here is measured in terms of discrete (3-monthly) intervals, for  $f(t)$  we could use a set of dummy variables, known as ‘blocking factors’, for intervals  $1, 2, \dots, n$  where  $n$  is the maximum number of intervals observed for any state. The estimated coefficients of these dummy variables would then provide a full prediction of the probability of leaving any current state. One difficulty of such a characterisation is that there may be a small number of very long episodes necessitating a large number of dummy variables, some of which would be poorly estimated. Instead we will usually be able to describe  $f(t)$  by a polynomial function of time, for example by writing  $f(t)$  as, say,  $f(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \dots$ . This is the approach we shall adopt but for completeness we first present the full specification using dummy variables for the modelled time intervals.

Formally then we have a 3-level binary response model where, as we have assumed, the (conditional) responses within episodes within individuals are independent. Level 3 is the individual, level 2 represents variation between repeated episodes within individuals and level 1 refers to the modelled time interval within repeated episodes. Level 2, the episode



level, corresponds to the lowest level of the 2-level repeated measures model originally specified in (1). The model has now become a 3-level one by dividing each episode into a set of lower level units, the modelled time intervals.

Using the conventional logit link function this model can be written in a more general form than (2) as

$$\begin{aligned} \text{logit}(\pi_{ijk(t)}) &= \beta_0 + \sum_{h=1}^p \alpha_h^* z_{hit}^* + \sum_{l=1}^m \beta_l x_{lijk(t)} + u_k^{(i)} + e_{jk} \\ y_{ijk(t)} &\sim \text{Binomial}(1, \pi_{ijk(t)}) \end{aligned} \quad (3)$$

where  $z_{hit}^*$  is the dummy variable for the modelled interval at time  $t$  and  $x_{lijk(t)}$  the covariates, including the dummy variable for partnership state. The level 3 term  $u_k^{(i)}$  is the random effect for individual  $k$  for state  $i$  and the level 2 term  $e_{jk}$  is the random effect associated with the  $j$ -th episode for the  $k$ -th individual. The level 1 random variation is binomial, as described in the second line of (3). In fact, for the present data described in detail in the next section, we can detect no variation at level 2 so that, for simplicity we shall assume just a 2-level model in the following exposition; that is we omit the term  $e_{jk}$ . The reason for the absence of the within-individual variation is partly explained by the fact that there are relatively few individuals with more than one partnership or non-partnership episode, but may also reflect a true high degree of homogeneity of durations within individuals. Thus the model becomes

$$\text{logit}(\pi_{ijk(t)}) = \beta_0 + \sum_{h=1}^p \alpha_h^* z_{hit}^* + \sum_{l=1}^m \beta_l x_{lijk} + u_k^{(i)} \quad (4)$$

It is worth noting that at each level in general there can be both random effects such as  $u_j$  and covariates defined at that level, such as gender. We can also have just a random effect on its own or just covariates without a random effect. Hence, although there is no episode level random effect in the partnership data we retain the subscript  $j$  in the covariate expression to allow for the possibility of episode-level covariates such as age at the start of the episode. As suggested above we use a polynomial of time to approximate the fitting of the full set of dummy variables, one for each modelled time interval. The order of this polynomial is typically 4 or 5 and it describes the underlying hazard. Thus, for a fourth order polynomial (4) becomes

$$\text{logit}(\pi_{ijk(t)}) = \beta_0 + \sum_{h=1}^4 \alpha_h t^h + \sum_{l=1}^m \beta_l x_{lijk} + u_k^{(i)} \quad (5)$$

where the polynomial in time,  $\sum_{h=1}^4 \alpha_h t^h$ , replaces the dummy variable function  $\sum_{h=1}^p \alpha_h^* z_{hit}^*$ .

An alternative is to group the blocking factors into a small number of relatively homogeneous longer intervals but we shall not pursue this.

As with other multilevel structures we may have, at the individual level, several random effects, for example a separate one for each state so that each individual is characterised by a partnership random effect and a separate one for a non-partnership state. These two random variables will both vary and covary across individuals. We shall fit such a model below.

As already pointed out, model (4) is a 2-level model with random variation at the level of the individual and the modelled time interval. The covariates we shall use are age at the start of episode and social class. The response is binary with a logit link function. As

such this model can be fitted by a number of statistical packages including SAS (<http://www.sas.com/>), STATA (<http://www.stata.com/>), and MLwiN (<http://mlwin.com>) and the last of these is used in the present study. An introductory account of such multilevel models can be found in Snijders and Bosker (1999). These packages will provide estimates for the parameters of the model and also estimates of the individual random effects  $u_k$ .

One of the interests in event history models is the estimation of duration length. That is, we wish to estimate the probability ( $S(t_j)$ ) of the duration being equal to  $t$ . In survival time modelling this will be the probability of survival to the end of time interval for which the duration time is  $t$ . This is therefore the product of the probabilities of not making a transition before the end of this interval. Thus, for state  $i$ , the probability of surviving until the end of the interval is

$$S(t) = \prod_{h=1}^{H_t} (1 - \pi_{ijk(t_h)})$$

where  $h$  indexes the modelled time interval,  $t_h$  is the duration length at the end of interval  $h$ , and  $H_t$  is the value of  $h$  corresponding to duration length  $t$ . Thus, if we have 3-month intervals and time is measured in years then  $H_1 = 4$ , etc.

By substituting particular covariate values (and states) into (4) we can obtain the corresponding predicted probabilities and hence duration estimates. Since we can also obtain estimates for the individual random effects we can estimate individual duration estimates as a function of time. Hence, using (4) without including the random effects

gives us population predictions, whereas incorporating the  $u_j$ , say, we obtain predictions for each individual in the sample.

## **Data**

The National Child Development Study (NCDS) is a longitudinal study which takes as its subjects all those living in Great Britain who were born between 3 and 9 March, 1958. The fifth follow-up of the National Child Development Study (NCDS5) took place in 1991 when the cohort members were age 33. 'Your Life Since 1974' was a self-completion questionnaire posted to the cohort members during the course of NCDS5 which asked for retrospective information on relationships, children, jobs and housing from the age of 16 until the time of the 1991 survey. Altogether, 11178 persons filled in either all or some of this section of the survey. In addition the 'Cohort Member Interview' of NCDS5, carried out by trained interviewers, also contained a retrospective partnership history. The final cleaned partnership histories (Bynner et al., 2002) are used to derive the three-monthly duration data for the study.

All but 39 of the cohort members had no more than four partnerships by the age of 33. Partnership involves cohabitation or marriage. Very few cohort members (61) had gone back to partners with whom they had lived before. The present analysis uses only a subset of the variables available. These are the 'start age', the episode level covariate which is the age of the cohort member at the start of the current episode; the social class of their father when the cohort member was aged 11 years (manual or non-manual – 2% had missing data or other codes), which is the individual level covariate coded 1 for non-

manual and 0 for manual, and whether the episode is a partnership (=0) or non-partnership (=1).

The data are for cohort males only and we carry out two analyses using the formulation in (5). The first ignores the initial time to establish a partnership, so that the first state is always a partnership. In this model we have two random effects for each individual. The first is their effect for non-partnerships, which represents the mean duration for that person of their non-partnership episodes: the second is the random effect for their partnership durations. We allow these to be correlated, since for example individuals with short non-partnership intervals may tend to have long partnership intervals, or vice-versa. Thus in (5) we have the two random effects:  $u_k^{(1)}$  which is the effect for individual  $k$  contributing to the response probability when in a partnership, and  $u_k^{(2)}$  which is the effect for individual  $k$  contributing to the response probability when in a non-partnership. Since these random effects vary and covary across individuals, and assuming that they have a bivariate Normal distribution we can write formally

$$\begin{pmatrix} u_k^{(1)} \\ u_k^{(2)} \end{pmatrix} \sim N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u1}^2 & \\ \sigma_{u12} & \sigma_{u2}^2 \end{pmatrix} \right]$$

The second analysis treats the time to first partnership separately, so that we now have covariates associated with partnership and no partnership states as before and an additional set of covariates associated with the first partnership only. At the individual level we therefore will have three random effects, one for the first episode, one for partnership and one for non-partnership durations. In this analysis the start age is omitted as an explanatory variable since, for the first partnership, it is effectively confounded

with the time to first partnership response. The total number of 3-month periods is 140420 with 3737 male cohort members who had partnership histories.

## **Results**

A fifth-order polynomial was used to smooth the time dependency for the hazard, replacing the blocking factors. Main effects for the above factors are fitted, together with interactions between the partnership status dummy variable and all the others, including the polynomial terms. Since the polynomial coefficients are essentially nuisance parameters used only to define the time dependent hazard, we do not interpret them, although they are used in the construction of the survivor or duration length functions. We have retained several ‘non significant’ polynomial coefficients in order to ensure that the underlying hazard is sufficiently accurately estimated; the order was chosen by noting when the predicted probabilities changed only negligibly when a further polynomial term was added.

Table 1 shows the estimates for a variance components model fitted using both quasilielihood estimation (PQL1, see Goldstein, 2003 Chapter 4) and Markov Chain Monte Carlo (MCMC, see Browne and Draper, 2000). The MCMC model uses a gamma prior (see Rasbash et al, 2000) and was run for 10,000 iterations with a burn-in of 1000. It gives somewhat different estimates from PQL1, which can happen with binary response data. It is known that PQL1 procedures can lead to biased estimates for binary responses and therefore in subsequent tables we quote only the MCMC estimates. The coefficient estimate for non-partnership (defined as a dummy variable taking the value 1 for non-partnership and 0 for partnership states) is 1.71 and highly significant when compared with its standard error. The inference is that the probability of exiting a non-partnership

episode, for any given time, is greater than exiting a partnership, that is non-partnership durations tend to be shorter.

The coefficient 'manual' is for a dummy variable set to 1 if manual and zero if not and hence the coefficient estimate refers to the manual – non manual difference. From the MCMC analysis this difference does not quite reach the 5% significance level. The corresponding difference for partnerships is obtained by adding the estimate corresponding to the interaction between partnership and social class to give  $0.113+0.142= -0.029$  for non-partnerships, which again is not significant. The later the starting age the longer the duration for partnerships but there is only a small ( $-0.041+0.026=-0.015$ ) and non significant relationship for non-partnership durations. There is also a relationship with start age where for partnerships the greater the start age the longer the duration tends to be, with a similar relationship for non-partnerships ( $-0.041+0.026 = -0.087$ ). The between individual standard deviation is  $0.61 (\sqrt{0.377})$  so that, assuming Normality, the range for approximately 95% of individuals is 2.44 . This is large compared to the other effects suggesting that while there are average effects of interest most variability occurs between individuals.

To include two random effects for each individual, one for partnership durations and one for non-partnership durations as explained above, we define two (0,1) dummy variables, one equal to 1 for partnership and the other equal to 1 for non-partnership durations. Thus we have two random effects, each varying across individuals and covarying. Table 2 introduces these random coefficients. The covariate coefficient estimates are close to those in table 1 and there is more variation between individuals for partnership durations (1.145) than for non-partnership durations (0.400). We also note that there is a small

negative correlation of -0.16 ( $-0.119/\sqrt{0.400*1.145}$ ), between partnership and non-partnership episodes indicating that long partnerships tend to be weakly associated with short non-partnerships and vice-versa. Thus individuals can, tentatively, be classified on this basis as either long partnership/short non-partnership or long non-partnership/short partnership individuals.

Finally we can look at the duration length distributions. Figures 1, and 2 plot the ‘survivor function’, i.e. the probability of remaining in a state, for different combinations, derived from the estimates in Table 2. This is the median population estimate derived from the model using the fixed part predictor, i.e. at the mean of the random effect distributions. It will be seen from Figure 1 that starting at age 16 years there is a slow decline in the probability of remaining outside a partnership for five years followed by a steep decline and then a tendency to level off at around the age of 30 (14 years after the start age of 16 years), with a fairly constant probability of remaining ‘single’ once that age is reached. For those who have already been in a partnership there is a very steep decline over time in the probability of remaining single but having remained single for ten years or so the probability of remaining single does not decline much with time. If we look at an individual’s position starting at 25 years (figure 2) we have a steep decline for the first ten years for the non-partnership probabilities with a levelling off after that. For the partnerships there is a steady decline in the probability of remaining in a partnership over 15 years.

Table 3 introduces the time to first partnership as a further response. This response is at the individual level, and may covary with the partnership and non-partnership durations. We find, however, that the variance for the first episode duration is small and poorly



estimated with a large standard error. We have therefore omitted it from the analysis so that the random part of the model is the same as in Table 2. We now see that being in a manual occupation greatly increases the probability of leaving the first episode compared to being in a non-manual occupation, the effect associated with this difference being  $0.326 - 0.044 = 0.282$ . This reflects the common finding that those in non-manual occupations acquire partners earlier. In figure 3 we plot the median survival probability of remaining in a non-partnership situation after age 16 separately for each social class. Which shows a slight increase in the social class difference over time, with non-manual individuals having an increasing higher probability of remaining single.

## **Discussion**

We have shown how longitudinal event history data involving repeated episodes for individuals can be modelled as a standard 2-level structure using readily available software. This allows between-individual variation in duration to be estimated and the model can deal with multiple states, although in the present case we have used only two, partnership and non-partnership. In addition to providing an estimate for the between-individual random effect variance for each state, it also allows a correlation to be estimated between the individual level random effects for the different states. While we have only fitted age and social class as explanatory variables in the present case, we can readily extend this to more variables. Education level achieved and parent's partnership status would be some of the most obvious candidates.

An early age of first partnership is associated with having a manual social class. There is a suggestion that for non-partnership episodes, other than the very first, those in the manual social class also have shorter durations, but the coefficient is non-significant. The

earlier the age the partnership starts the shorter the partnership duration, but for non-partnership episodes there is little relationship between age of starting and duration.

The models used in this paper can be extended in a number of directions. For example we can model different kinds of transitions from an episode state, so that partnerships may end in separation, divorce, death etc. We can also fit multivariate models that study the simultaneous durations of different kinds of states such as partnership and employment. Steele et al., (2003) describe how such models can be specified and fitted.

In practice care will be needed in defining the time intervals to be used. If these are too short this can result in extensive data sets, but they should not be long so that more than one transition takes place within an interval. Sometimes the exact durations are unknown and then it will be useful to carry out more than one analysis, varying the time intervals.

**Table 1 Repeated measures models for partnership and non-partnership episodes: starting from first partnership.**

Parameter	Estimate (s.e.). PQL1	Estimate (s.e.) MCMC
<i>Fixed</i>		
intercept	-3.936	-4.065
z	-0.289(0.063)*10 <sup>-1</sup>	-0.283 (0.064) * 10 <sup>-1</sup>
z <sup>2</sup>	0.132(0.070)*10 <sup>-2</sup>	0.160 (0.078) * 10 <sup>-2</sup>
z <sup>3</sup>	-0.510(2.283)*10 <sup>-5</sup>	-0.584 (2.331) * 10 <sup>-5</sup>
z <sup>4</sup>	-0.167(0.160)*10 <sup>-5</sup>	0.230 (0.181) * 10 <sup>-5</sup>
z <sup>5</sup>	0.347(0.357)*10 <sup>-7</sup>	0.451 (0.410) * 10 <sup>-7</sup>
start age	-0.039(0.010)	-0.041 (0.011)
manual	-0.113(0.065)	-0.113 (0.072)
np(non-partnership)	1.612(0.396)	1.710 (0.377)
np*z	0.578(0.164)*10 <sup>-1</sup>	0.616 (0.182) * 10 <sup>-1</sup>
np*z <sup>2</sup>	-0.224(0.135)*10 <sup>-2</sup>	-0.249 (0.147) * 10 <sup>-2</sup>
np* z <sup>3</sup>	-0.145(0.096)*10 <sup>-3</sup>	-0.151 (0.113) * 10 <sup>-3</sup>
np* z <sup>4</sup>	0.695(0.332)*10 <sup>-6</sup>	0.754 (0.383) * 10 <sup>-6</sup>
np* z <sup>5</sup>	0.112(0.132)*10 <sup>-6</sup>	0.090 (0.165) * 10 <sup>-6</sup>
np*start age	0.027(0.016)	0.026 (0.016)
np>manual	0.131(0.105)	0.142 (0.255)
<i>Random</i>		
$\sigma_{v0}^2$	0.289 (0.045)	0.377 (0.064)

Note that z indicates the time interval (1,2...) and is centred at 20. In all tables the dummy variable for non-partnership combines with the hazard polynomial, age and social class variables to form interaction terms. Thus, for example the manual – non manual difference coefficient for non-partnerships is given from the second column as – 0.113+0.142=0.029. Social class is coded Manual=1, non manual =0.

**Table 2. Random coefficient model for partnership and outside partnership. MCMC estimates: starting from first partnership.**

Parameter	Estimate (s.e.)
<i>Fixed</i>	
intercept	-3.709
$z$	$-0.244(0.067)*10^{-1}$
$z^2$	$0.076(0.053)*10^{-2}$
$z^3$	$-0.140(0.240)*10^{-4}$
$z^4$	$-0.089(0.114)*10^{-5}$
$z^5$	$0.093(0.264)*10^{-7}$
start age	-0.063(0.010)
manual	-0.116(0.075)
np(non-partnership)	1.753(0.469)
np*z	$0.499(0.208)*10^{-1}$
np*z <sup>2</sup>	$-0.181(0.134)*10^{-2}$
np* z <sup>3</sup>	$-0.112(0.114)*10^{-3}$
np* z <sup>4</sup>	$0.024(0.332)*10^{-5}$
np* z <sup>5</sup>	$0.055(0.158)*10^{-6}$
np*start age	0.049(0.017)
np>manual	0.152(0.118)
<i>Random</i>	
$\sigma_{v0}^2$ (non-partnership)	0.400(0.211)
$\sigma_{v01}$	-0.119(0.118)
$\sigma_{v1}^2$ (partnership)	1.145(0.171)

**Table 3. Random coefficient model for first partnership, partnership and outside partnership. MCMC estimates.**

Parameter	Estimate (s.e.)
<i>Fixed</i>	
intercept	-5.227
$z$	$-0.316 (0.036) * 10^{-1}$
$z^2$	$0.119 (0.021) * 10^{-2}$
$z^3$	$0.067 (0.083) * 10^{-4}$
$z^4$	$-0.155 (0.025) * 10^{-5}$
$z^5$	$0.270 (0.070) * 10^{-7}$
manual	-0.044 (0.064)
np(non-partnership)	2.746 (0.150)
np*z	$0.682 (0.141) * 10^{-1}$
np*z <sup>2</sup>	$-0.232 (0.144) * 10^{-2}$
np* z <sup>3</sup>	$-0.134 (0.078) * 10^{-3}$
np* z <sup>4</sup>	$0.041 (0.344) * 10^{-5}$
np* z <sup>5</sup>	$0.850 (1.180) * 10^{-7}$
np>manual	0.103 (0.120)
fp (time to first partnership)	1.453 (0.067)
fp*z	0.144 (0.005)
fp*z <sup>2</sup>	$-0.625 (0.034) * 10^{-2}$
fp* z <sup>3</sup>	$0.233 (0.129) * 10^{-4}$
fp* z <sup>4</sup>	$0.479 (0.074) * 10^{-5}$
fp* z <sup>5</sup>	$-0.830 (0.130) * 10^{-7}$
fp>manual	0.326 (0.071)
<i>Random</i>	
$\sigma_{v0}^2$ (non-partnership)	0.462 (0.061)
$\sigma_{v01}$	-0.313 (0.095)
$\sigma_{v1}^2$ (partnership)	0.773 (0.141)

## References

Bynner, J., Elias, P., McKnight, A., Pan, H., et al. (2002). *Young people in transition: changing pathways to employment and independence*. York, Joseph Rowntree Foundation.

Browne, W. and Draper, D. (2000). Implementation and performance issues in the Bayesian and likelihood fitting of multilevel models. *Computational statistics* **15**: 391-420.

Goldstein, H. (2003). *Multilevel Statistical Models*. Third Edition. London, Edward Arnold:

Rasbash, J., Browne, W., Goldstein, H., Yang, M., et al. (2000). *A user's guide to MlwiN (Second Edition)*. London, Institute of Education:

Singer, J. D. and Willett, J. B. (2003). *Applied longitudinal data analysis: modelling change and event occurrence*. New York, Oxford University Press:

Snijders, T. and Bosker, R. (1999). *Multilevel Analysis*. London, Sage:

Steele, F., Goldstein, H. and Browne, W. (2003). A general multilevel multistate competing risks model for event history data. (submitted for publication).

Figure 1.

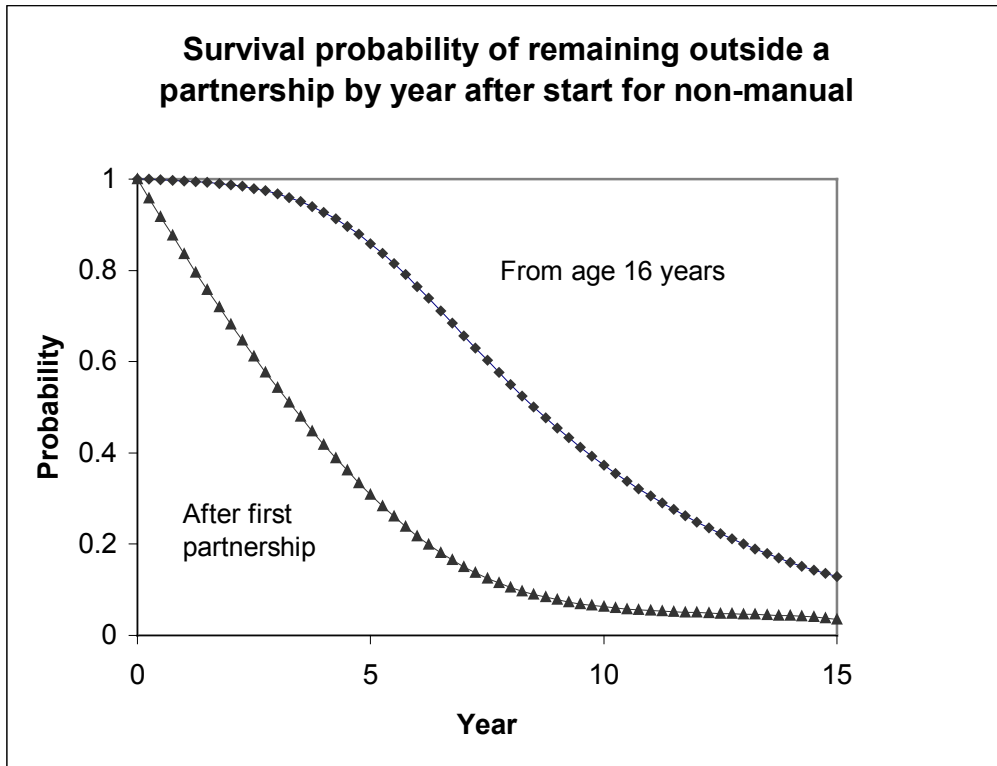


Figure 2.

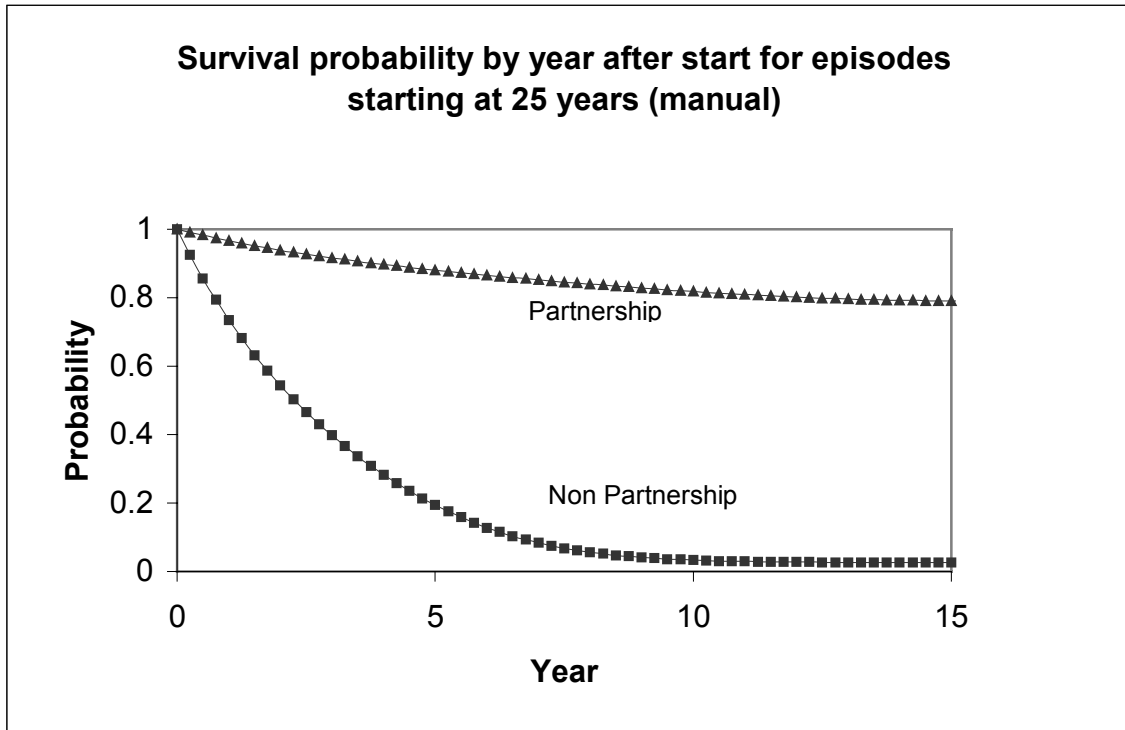




Figure 3.

