Fiona Steele

Multilevel models for longitudinal data

Article (Accepted version) (Refereed)

Original citation:
Steele, Fiona (2008) Multilevel models for longitudinal data. Journal of the Royal Statistical Society: series A (statistics in society), 171 (1). pp. 5-19. ISSN 0964-1998
DOI: 10.1111/j.1467-985X.2007.00509.x

Multilevel Models for Longitudinal Data

Fiona Steele

Centre for Multilevel Modelling
Graduate School of Education
University of Bristol
2 Priory Road
Bristol BS8 1TX

Email: Fiona.Steele@bristol.ac.uk

Summary.

Repeated measures and repeated events data have a hierarchical structure which can be analysed using multilevel models. A growth curve model is an example of a multilevel random coefficients model, while a discrete-time event history model for recurrent events can be fitted as a multilevel logistic regression model. The paper describes extensions to the basic growth curve model to handle autocorrelated residuals, multiple indicator latent variables and correlated growth processes, and event history models for correlated event processes. The multilevel approach to the analysis of repeated measures data is contrasted with structural equation modelling. The methods are illustrated in analyses of children's growth, changes in social and political attitudes, and the interrelationship between partnership transitions and childbearing.

Key words.

Repeated measures; Multilevel models; Structural equation models; Simultaneous equation models; Event history analysis

1. Introduction

Over the past twenty years multilevel modelling has become a standard approach in the analysis of clustered data (Goldstein, 2003). Longitudinal data are one example of a hierarchical structure, with repeated observations over time (at level 1) nested within individuals (level 2). By viewing longitudinal data as a two-level structure, researchers can take advantage of the large body of methodological work in this area, including extensions to more complex hierarchical and non-hierarchical structures, categorical and duration responses and multivariate mixed response types. The aim of this paper is to outline the multilevel modelling approach, demonstrating how traditional growth curve models can be framed as multilevel models, and to describe more recent developments such as multilevel structural equation models for the analysis of repeated hypothetical constructs measured by multiple indicators and for the simultaneous analysis of multiple correlated processes.

Studies using longitudinal data are generally concerned with either the change over time in one or more outcome variables, or the timing of events (Singer and Willett, 2003). Examples of research questions concerned with change include enquiries about child development, changes in the social or economic circumstances of households or areas, and changes in individual attitudes or behaviour. In each case, analysis would be based on repeated measurements on a single outcome or set of outcomes. Examples where the outcome is the duration to the occurrence of an event include studies of the timing of death, births, partnership dissolution or a change in employment status. Event history data may be derived from current status data that are prospectively collected in successive waves of a panel study, e.g. marital or employment status, or from the dates of events that are usually collected retrospectively.

Methods for the analysis of change include growth curve models, also known as latent trajectory models, and autoregressive models. In the growth curve approach the repeated measures are viewed as outcomes that are dependent on some metric of time (e.g. wave or age). In an autoregressive model the outcome at occasion t is a function of lagged outcomes, for example the outcome at $t-1$ in a first-order model. Both types of model can be viewed as special cases of either a multilevel model or a structural equation model. Event history analysis, also known as survival or duration analysis, is used to model the timing of events, allowing for the possibility that durations may be partially observed (censored) for some members of the sample. Multilevel models can be applied when events are repeatable to allow for correlation between the durations to events experienced by the same individual, or when individuals are clustered into higher-level units.

This paper provides an overview of multilevel models for the analysis of change and event processes. The multilevel modelling and structural equation modelling approaches to growth curve analysis, and their relative advantages, are discussed. Generalisations of the basic growth curve model and event history model are described, including growth curve models that allow for autocorrelated residuals, factor analysis models for multiple indicators, and event history models for competing risks and multiple states. Models for multiple change or event processes are also discussed. The multilevel modelling approach is illustrated in analyses of repeated height measurements on children, changes in social and political attitudes, and the interrelationship between partnership transitions and childbearing.

2. Analysing Change

Denote by y_{ti} the response at measurement occasion t $(t = 1,...,T_i)$ for individual $i\,(i = 1,...,n)$. Repeated measures have a two-level hierarchical structure with measurements at level 1 nested within individuals at level 2. The number of measurement occasions may vary across individuals, for example due to attrition. The timing of measurements may also vary, for example if there is variation in the age of children taking an educational test at a given occasion.

In this section, we discuss growth curve models for y_{ti} with extensions to handle autocorrelation, multiple indicators in a measurement model, and multivariate responses.

2.1 Growth curve models

We denote by z_{ti} the time of measurement occasion t for individual i, where the most commonly used time metrics are calendar time and chronological age. In the case of panel data where z refers to calendar time, and variation in the interview date at a given wave can be ignored, then $z_{ti} = z_t$. More generally, and particularly in the context of growth studies where z is age, the timing of measurement at a particular occasion may vary across individuals and we would usually wish to account for this variation in the model.

In the simplest model for a continuous response a linear trajectory is fitted for each individual:

$$\begin{aligned} y_{ti} &= \alpha_{0i} + \alpha_{1i} z_{ti} + \boldsymbol{\beta}^T \mathbf{x}_{ti} + e_{ti} \\ \alpha_{0i} &= \alpha_0 + u_{0i} \\ \alpha_{1i} &= \alpha_1 + u_{1i} \end{aligned} \qquad (1)$$

which is sometimes written in single-equation (or reduced) form as

$$y_{ti} = \alpha_0 + \alpha_1 z_{ti} + \boldsymbol{\beta}^T \mathbf{x}_{ti} + u_{0i} + u_{1i} z_{ti} + e_{ti},$$

where $\mathbf{x}_{ti}$ is a vector of covariates that may be time-varying or individual characteristics, u_{0i} and u_{1i} are individual-specific residuals (or random effects), and e_{ti} are residuals at the measurement occasion level. The time variable z_{ti} is treated as an additional covariate. The average line describing the relationship between y and z at $\mathbf{x}_{ti} = \mathbf{0}$ is given by $\alpha_0 + \alpha_1 z_{ti}$, and u_{0i} and u_{1i} are individual departures from the intercept and slope of this line. It is usually assumed that all residuals are normally distributed, and residuals defined at the same level may be correlated, i.e. $e_{ti} \sim N(0, \sigma_e^2)$ and $\mathbf{u}_i = [u_{0i}\ u_{1i}]^T \sim N(\mathbf{0}, \boldsymbol{\Omega}_u)$ where $\boldsymbol{\Omega}_u = \begin{pmatrix} \sigma_{u0}^2 & \\ \sigma_{u01} & \sigma_{u1}^2 \end{pmatrix}$, and σ_{u0}^2 and σ_{u1}^2 are the between-individual variances in the intercepts and slopes of the individual growth trajectories. It is common practice to centre z_{ti}. For example, if z_{ti} is calendar time and there are five equally spaced measurement occasions, the centred z_{ti} would be coded -2, -1, 0, 1 , 2 and σ_{u0}^2 is then interpreted as the between-individual variance in y at the mid-point. σ_{u01} is the covariance between the intercepts and slopes of the individual trajectories, where a positive (negative) covariance implies that individuals with a high value of y at $z_{ti} = 0$ tend to have a high (low) growth rate.

The between-individual variance in the expected value of y, conditional on covariates $\mathbf{x}_{ti}$, is given by

$$\text{var}(u_{0i} + u_{1i} z_{ti}) = \sigma_{u0}^2 + 2\sigma_{u01} z_{ti} + \sigma_{u1}^2 z_{ti}^2 \qquad (2)$$

i.e. a quadratic function of time.

From (2) it can be seen that, because σ_{u0}^2 and σ_{u1}^2 must both be greater than zero, a positive σ_{u01} implies that the between-individual variance increases after the mid-point $z_{ti} = 0$, i.e. the individual values of y will start to diverge after this time. Conversely, a negative σ_{u01} implies that the between-individual variance decreases (a convergence in individual y-values) for at least some time after $z_{ti} = 0$. Specifically, the quadratic function in (2) reaches its minimum value at $z_{ti} = -\sigma_{u01} / \sigma_{u1}^2$; if such a value lies within the observed range of z_{ti}, the between-individual variance will increase after this point. Thus, individuals with a low y-value at $z_{ti} = 0$ tend to have the highest growth rates and, at some point beyond $z_{ti} = 0$, they may catch up with, or even overtake, individuals who had a high value of y at $z_{ti} = 0$. In the event that an individual with a low y-value at $z_{ti} = 0$ overtakes someone with a higher value at $z_{ti} = 0$, their growth trajectories will cross each other. If this occurs for a sufficient proportion of individuals, the individual y-values will start to diverge and the between-individual variance will increase.

Elaborations to Model (1) include fitting different functions of z_{ti}, and allowing for further levels of clustering. For instance, a polynomial growth curve is specified by including as explanatory variables powers of z_{ti}, and a step function is fitted by treating z_{ti} as categorical. More complex hierarchical or non-hierarchical structures arise when individuals are nested within higher level units or a cross-classification of different types of unit, for example children within schools, or within a cross-classification of schools and neighbourhoods. Further details of the random effects approach to repeated measures analysis can be found in Laird and Ware (1982), Diggle et al. (2002), Raudenbush and Bryk (2002), and Goldstein (2003).

Model (1) can also be framed as a structural equation model (SEM) (Muthén, 1997; Curran, 2003; Bollen and Curran, 2006). The SEM approach to growth curve analysis involves fitting a type of two-factor confirmatory factor model to { y_{ti} }, which are treated as multiple indicators of two latent factors, u_{0i} and u_{1i} :

$$y_{ti} = \mu_{0t} + \boldsymbol{\beta}^T \mathbf{x}_{ti} + \lambda_{0t} u_{0i} + \lambda_{1t} u_{1i} + e_{ti} \qquad (3)$$

where μ_{0t} are occasion-specific intercepts, and λ_{0t} and λ_{1t} are factor loadings. To see the equivalence of (1) and (3) when $z_{ti} = z_t$, we can substitute $\mu_{0t} = \alpha_0 + \alpha_1 z_t$, $\lambda_{0t} = 1$ for all t, and $\lambda_{1t} = z_t$ in (3). Thus the growth curve model is fitted by setting the loadings of the intercept factor u_{0i} to one and, in the case of equally spaced measurements, the loadings of the slope factor u_{1i} to 0, 1, 2 etc (see Bollen and Curran (2006) for further details). A hierarchical level above the individual can be accommodated using multiple-group analysis (see Muthén, 1994).

Where there is individual variation in the timing of measurements at a given occasion, it is more difficult to fit (1) as a SEM. One approach would be to construct an expanded multivariate response vector with an element for each possible value of z_{ti} (observed for any individual) but where, for individual i, all but T_i of these responses are missing. This is a special case of the more general problem of how to incorporate a continuous level 1 predictor in a SEM where not all values of the predictor are observed for all level 2 units. (See Curran (2003) for a brief discussion of a possible solution using definition variables.)

Model (1) can be estimated using maximum likelihood, and the same results would be obtained regardless of whether it is treated as a multilevel model or a structural equation model. However, one approach may be preferred over the other for certain types of data or extensions to (1). It is common in panel studies to have a variable number of responses across individuals, due to attrition or non-monotone patterns of missingness, leading to an unbalanced data structure. If a SEM is used, some method must be used to compensate for missing data, e.g. full information maximum likelihood (Arbuckle, 1996) or multiple imputation (Schafer, 1997). In a multilevel model cluster sizes are not required to be equal and therefore, when applied to repeated measures data, individuals with missing responses can be included without any adjustment provided the data can be assumed missing at random. It is also straightforward to allow for between-individual variation in the timing and spacing of measurements in a multilevel framework because the timing of each measurement occasion z_{ti} is treated as an explanatory variable. We can therefore combine data from individuals with very different measurement patterns, some of whom may have been measured only once and others at several irregularly spaced intervals. Further advantages of the multilevel approach are the facility to allow for more general hierarchical and non-hierarchical structures, non-normal responses and mixed response types in a multivariate setting (see Section 2.4). Finally, multilevel models can now be fitted in a number of specialist and mainstream software packages (a set of software reviews, with syntax for fitting a range of multilevel models, can be downloaded from http://www.cmm.bris.ac.uk/Learning_Training/Software_MM).

The SEM approach is useful when the outcome of interest cannot be directly observed, but is measured indirectly through a set of indicators $\{y_{kti}\}$ at each occasion. A structural equation model for y_{kti} includes a measurement component that links the observed indicators to one or more latent variable, depending on the dimensionality of the latent construct. Other generalisations that might benefit from estimation via SEM are models with predictors measured by multiple indicators and structural models that decompose total effects into direct and indirect effects (Curran, 2003).

<u>Example</u>: Modelling repeated height measurements

We illustrate the application of growth curve modelling in an analysis of height measurements taken on 26 boys on nine occasions, spaced approximately 0.25 years apart between the ages of 11 and 14. (The data are described and analysed in Goldstein et al. (1994).) The height y_{ti} of boy i at occasion t can be modelled as a cubic polynomial function of age, z_{ti}:

$$y_{ti} = \alpha_{0i} + \alpha_{1i} z_{ti} + \alpha_{2i} z_{ti}^2 + \alpha_{3i} z_{ti}^3 + e_{ti}$$
$$\alpha_{ki} = \alpha_k + u_{ki}, \quad k = 0,1,2,3 \qquad (4)$$

where $\text{var}(u_{ki}) = \sigma_{uk}^2$ and $\text{cov}(u_{ki}, u_{k'i}) = \sigma_{ukk'}$, $k \neq k'$.

The analysis was carried out using *MLwiN* (Rasbash et al., 2004). Table 1 shows results from a series of likelihood ratio tests of the nature of variation in boys' growth rates. In Model 1 of Table 1 only the intercept is permitted to vary across boys. This model is clearly rejected in favour of Model 2 which allows for individual variation in growth rates, but only in the linear term α_{1i}. Model 3 is, in turn, found to be a significantly better fit to the data than Model 2. However, allowing the cubic effect to vary across individuals, as in Model 4, shows no significant improvement in model fit. Table 2 shows estimates for the selected model (Model 3) which includes random coefficients for z and z^2, but not for z^3. Age has been centred so that the intercept variance σ_{u0}^2 is interpreted as the between-individual variance in heights at age 12.25 years. The between-individual variance is a fourth-order polynomial function in age, which is a generalisation of (2) where both z_{ti} and z_{ti}^2 have random coefficients. As expected, the variation in boys' heights increases with age (Figure 1).

2.2 Autocorrelation

In Model (1) the occasion-level residuals e_{ti} are assumed to be uncorrelated. In practice, however, measurements that are close together in time will have similar departures from that individual's growth trajectory, leading to autocorrelation between the e_{ti}. We can extend (1) by adding a model for the e_{ti}, leading to a multilevel time series model (Goldstein et al., 1994; Diggle et al., 2002). A general model for measurements spaced s units apart can be written

$$\text{cov}(e_{ti}, e_{t-s,i}) = \sigma_e^2 f(s)$$

where $f(s)$ is a function of the distance between measurements. In most situations the autocorrelation will decrease with s, and it is convenient to characterise the decay process as

$$\text{cov}(e_{ti}, e_{t-s,i}) = \sigma_e^2 \exp(-\gamma\, s) \qquad (5)$$

where $\gamma>0$. Model (5) is a continuous-time analogue of the discrete-time first-order autoregressive, AR(1), model.

Model (5) was fitted to the boys' height data, extending the polynomial growth model (4). Using *MLwiN* we obtain $\hat{\gamma} = 8.56$ (SE=3.28), which implies predicted autocorrelations at lags 0.25, 0.5 and 1 of 0.12, 0.01 and 0.002 respectively. However, allowing for autocorrelation does not significantly improve model fit (Δ -2 log L = 1.1, 1 d.f.).

2.3 Repeated latent variables with multiple indicators

Suppose the outcome of interest is a hypothetical or latent construct y_{ti}^* that cannot be measured directly by a single variable, but is measured indirectly on several occasions by a set of K observed indicators $\{y_{kti}\}$. The multiple indicators y_{kti} may be linked to the latent variable y_{ti}^* through a factor or measurement model:

$$y_{kti} = \lambda_{0k} + \lambda_{1k} y_{ti}^* + v_{ki} + \varepsilon_{kti}, \quad k = 1,\ldots,K, \qquad (6)$$

where λ_{0k} are indicator-specific intercepts and λ_{1k} are factor loadings; $v_{ki} \sim N(0,\sigma_{vk}^2)$ and $\varepsilon_{kti} \sim N(0,\sigma_{\varepsilon k}^2)$ are residuals at the individual and occasion individual level (also called 'uniquenesses') which are assumed to be uncorrelated across indicators. We also assume that y_{ti}^* is normally distributed.

We are usually interested in examining change in the latent variable rather than in its observed indicators, and therefore define a growth curve model for y_{ti}^*, which has the same form as (1) with y_{ti} replaced by y_{ti}^*:

$$y^*_{ti} = \alpha_{0i} + \alpha_{1i} z_{ti} + \boldsymbol{\beta}^T \mathbf{x}_{ti} + e_{ti}$$
$$\alpha_{0i} = \alpha_0 + u_{0i} \qquad (7)$$
$$\alpha_{1i} = \alpha_1 + u_{1i}$$

where e_{ti}, u_{0i} and u_{1i} are normally distributed as before.

Equation (7) is called a structural model, and (6) and (7) together define a multilevel SEM. Extensions to this model include the addition of covariates to (6), and adding further latent variables to the measurement model to explain the association between the $\{y_{kti}\}$. Where there is more than one latent variable, the structural model may be extended to allow for dependencies between them. It is also possible to allow for covariate measurement error by treating covariates as latent variables. See Bollen and Curran (2004; 2006) for further discussion of growth curve models for repeated latent variables and Skrondal and Rabe-Hesketh (2004) for a detailed treatment of more general multilevel SEMs.

Example: Modelling change in social and political attitudes

The multiple indicators growth curve model is applied in an analysis of six social and political attitude items collected at five waves of the British Household Panel Study in 1992, 1994, 1996, 1998 and 2001 (UK Data Archive, 2004). The items are measured on a five-point scale which indicates attitude towards the following statements (coded 1=strongly agree, 2=agree, 3=neither agree nor disagree, 4=disagree, 5=strongly disagree):

1. Ordinary people share the nations wealth
2. There is one law for the rich and one for the poor
3. Private enterprise solves economic problems
4. Public services ought to be state owned
5. Government has an obligation to provide jobs
6. Strong trade unions protect employees

For the purposes of this illustration, we restrict the analysis to the 3787 individuals who responded at each wave and treat the items as if they were measured on a continuous scale.

The SEM described by (6) and (7) is modified in two ways. First, individual change in opinion is modelled as a step function by including as explanatory variables in (7) dummy variables for waves 2-5 with coefficients $\alpha_1,\ldots,\alpha_4$, rather than a linear function in z_{ti}. Second, we simplify (7) to a random intercept model by eliminating the term u_{1i} from the equation for α_{1i}, i.e. we assume that the rate of change is constant across individuals. Two identification constraints are applied in order to fix the scale of the latent variable y^*_{ti}. First, the factor loading for item 1 in (6), λ_{11}, is fixed at one, which constrains the factor to have the same variance as this item. Second, the central location of y^*_{ti} is fixed at the mean response value for the reference year 1992 (wave 1) by constraining the intercept in (7), α_0, to equal zero.

The model was fitted using Gibbs sampling, a Markov chain Monte Carlo (MCMC) method, in WinBUGS (Spiegelhalter et al., 2000). Non-informative priors were assumed for all parameters. Table 3 shows results from 15,000 samples with a burn-in of 1000. Starting with the measurement model, we find that all but items 1 and 3 load negatively on the underlying factor y^*_{ti}. This may be explained by differences in the direction of question wording: compared to the other items, agreement with items 1 and 3 suggests more right-of-centre attitudes. We might therefore interpret y^*_{ti} as a summary measure of social and political attitudes, ranging from right-of-centre (low values of y^*_{ti}) to left-of-centre (high values). All loadings are close to 1 in magnitude, suggesting that the items have approximately equal discriminatory power. After accounting for the common factor y^*_{ti}, there remains a large amount of between and within individual variation in the responses on each item, i.e. the items have low communality. Turning to the structural model, we find evidence of higher values of y^*_{ti} (more left-of-centre attitudes) in 1994 and 1996, with a move towards more right-of-centre attitudes in the waves following the start of the Labour government in 1997.

In this illustrative example, we have omitted respondents with missing data at any wave. Attrition is a pervasive problem in panel studies, and restricting the analysis to complete cases is likely to lead to bias if drop-outs are a non-random sub-sample of the baseline sample. In a Bayesian framework, missing values can be treated as additional parameters and a step can be added to the MCMC algorithm to generate values for the missing responses (Browne, 2004, Chapter 17). An alternative approach is to use multiple imputation, ensuring that the imputation model allows for the dependency between measurements from the same individual (Schafer and Yucel, 2002; Carpenter and Goldstein, 2004).

2.4 Causal models for multivariate responses

Suppose there are longitudinal data on two outcome variables, $y^{(1)}$ and $y^{(2)}$, which we believe are related although the causal direction may be unclear. For example we may have observations on different dimensions of child development, such as cognitive and emotional indicators, measured at several points in time. Model (1) can be elaborated to allow for reciprocal causation between $y^{(1)}$ and $y^{(2)}$ leading to

$$\begin{aligned} y_{ti}^{(1)} &= \alpha_{0i}^{(1)} + \alpha_{1i}^{(1)} z_{ti} + \gamma^{(1)} y_{t-1,i}^{(2)} + \boldsymbol{\beta}^{(1)^T} \mathbf{x}_{ti}^{(1)} + e_{ti}^{(1)} \\ y_{ti}^{(2)} &= \alpha_{0i}^{(2)} + \alpha_{1i}^{(2)} z_{ti} + \gamma^{(2)} y_{t-1,i}^{(1)} + \boldsymbol{\beta}^{(2)^T} \mathbf{x}_{ti}^{(2)} + e_{ti}^{(2)}, \quad t = 2,\ldots,T_i, \end{aligned} \tag{8}$$

where $\alpha_{0i}^{(l)} = \alpha_0^{(l)} + u_{0i}^{(l)}$ and $\alpha_{1i}^{(l)} = \alpha_1^{(l)} + u_{1i}^{(l)}$ for $l = 1,2$, and $\mathbf{x}_{ti}^{(1)}$ and $\mathbf{x}_{ti}^{(2)}$ are response-specific covariate vectors. Model (8) is a simultaneous equation model in which each growth process depends on the lagged outcome of the other process. The two processes are additionally linked by allowing for correlation between residuals across equations. A between-process residual correlation would arise if there were shared or correlated influences on the two processes that were not adequately accounted for by covariates. In the most general model we allow for correlation between the following pairs of residuals: $(e_{ti}^{(1)}, e_{ti}^{(2)})$, $(u_{0i}^{(1)}, u_{0i}^{(2)})$ and $(u_{1i}^{(1)}, u_{1i}^{(2)})$, which allows for correlation between the time-varying or individual-specific unobservables that affect each process. As before, any pair of random effects defined at the same level and appearing in the same equation may be correlated. Thus $\text{cov}(u_{0i}^{(1)}, u_{u1}^{(1)})$ and $\text{cov}(u_{0i}^{(2)}, u_{u1}^{(2)})$ are freely estimated.

The equations in (8) define a multilevel bivariate response model which can be framed as a random slopes model and therefore estimated using multilevel modelling software. The data have a three-level hierarchical structure with responses (level 1) nested within measurement occasions (level 2) within individuals (level 3). Alternatively, the model can be viewed as a confirmatory factor model for a set of $2(T_i - 1)$ responses consisting of the two responses $y_{ti}^{(1)}$ and $y_{ti}^{(2)}$ for occasions $2, 3, \cdots, T_i$. The factors are the random effects, and the model is confirmatory because $(u_{0i}^{(1)}, u_{1i}^{(1)})$ have zero loadings for $y_{ti}^{(2)}$ and, similarly, $(u_{0i}^{(2)}, u_{1i}^{(2)})$ have zero loadings for $y_{ti}^{(1)}$.

A variant of (8) is the commonly used cross-lagged model in which $\alpha_{0i}^{(l)} + \alpha_{1i}^{(l)} z_{ti}$ is replaced by an autoregressive term $\alpha_{0i}^{(l)} + \alpha_{1i}^{(l)} y_{t-1,i}^{(l)}$ $(l = 1, 2)$. Alternatively both latent growth and autoregressive terms can be included, leading to an autoregressive latent trajectory model. (See Bollen and Curran (2004) for further details and a discussion of model identification.) The model can be extended to allow for further levels of clustering. For example, Muthén (1997) applies a simultaneous growth curve model to measures of mathematics achievement and attitudes to mathematics, allowing the intercept of one growth process to affect the slope of the other and controlling for within-school correlation in both outcomes. Measurement error in either or both outcomes can be handled in a multilevel SEM, i.e. a synthesis of (6)-(8).

3. Analysing Event Occurrence

In the previous section we considered models for studying change in an outcome y_{ti} over time. The other main strand of longitudinal research is concerned with the timing of events. Event history data may be in the form of event times, usually collected retrospectively, or a set of current status indicators from waves of a panel study. Both forms of data collection will usually lead to interval-censored rather than continuous duration data because the precise timing of event occurrence is generally unknown. Durations derived from retrospective data are typically recorded to the nearest month or year, depending on the saliency of the event to respondents, while panel data are collected prospectively at infrequent intervals. Thus, although events in the process under study can theoretically occur at any point in time, durations are actually measured in discrete time. We therefore restrict the following discussion to discrete-time models. Another reason for adopting a discrete-time approach is that very general models for repeated events, competing risks, multiple states and multiple processes can be estimated using existing procedures for discrete response data.

3.1 Discrete-time event history analysis

We begin with a brief description of a simple discrete-time model for a single event time (see Allison (1982) for further details). For each individual i we observe a duration y_i which will be right-censored if the event has not yet occurred by the end of the observation period. In addition we observe a censoring indicator δ_i, coded 1 if the duration is fully observed (i.e. an event occurs) and 0 if right-censored. The first step of a discrete-time analysis is to expand the data so that for each time interval t up to y_i, we define a binary response y_{ti} coded as:

$$y_{ti} = \begin{cases} 0 & t < y_i \\ 0 & t = y_i, \delta_i = 0 \\ 1 & t = y_i, \delta_i = 1. \end{cases}$$

For example, if an individual has an event during the third time interval of observation their discrete responses will be (y_{1i}, y_{2i}, y_{3i}) = (0,0,1), while someone who is censored at t=3 will have response vector (0,0,0).

We model the hazard function for interval t, defined as the conditional probability of an event during interval t given that no event has occurred in a previous interval, i.e.

$$h_{ti} = \Pr(y_{ti} = 1 \mid y_{si} = 0, s < t) .$$

The hazard is the usual response probability for a binary variable. Therefore, after restructuring the data, the event indicator y_{ti} can be analysed using any model appropriate for binary responses, such as a logit model:

$$\text{logit}(h_{ti}) = \log\left(\frac{h_{ti}}{1-h_{ti}}\right) = \alpha(t) + \boldsymbol{\beta}^T \mathbf{x}_{ti} \qquad (9)$$

where $\alpha(t)$ captures the nature of the dependency of the hazard function on t, and $\mathbf{x}_{ti}$ is a vector of covariates which may be time-varying or fixed individual characteristics. The baseline logit-hazard $\alpha(t)$ is specified by including some function(s) of t as explanatory variables. For example, a quadratic function is fitted by including t and t^2, and a step function is obtained by treating t as a categorical variable.

3.2 Multilevel event history model for repeated events

Many events that we study in social research may occur more than once to an individual over the observation period. For example, individuals may move in and out of co-residential relationships multiple times, they may have more than one child, and they may have several changes of job. If repeated events are observed we can model the duration of each *episode*, where an episode is defined as a continuous period during which an individual is at risk of experiencing a particular event. When an event occurs, a new episode begins and the duration 'clock' is reset to zero. In discrete-time, we define a binary response y_{tji} for each interval t of episode j for individual i, and denote the corresponding hazard function by h_{tji}.

When events are repeatable, event history data have a two-level hierarchical structure with episodes (level 1) nested within individuals (level 2). Thus repeated events may be analysed using multilevel models. A random effects logit model, also known as a shared frailty model, may be written

$$\text{logit}(h_{tji}) = \alpha(t) + \boldsymbol{\beta}^T \mathbf{x}_{tji} + u_i \qquad (10)$$

where the covariates $\mathbf{x}_{tji}$ may be time-varying, or characteristics of episodes or individuals; and $u_i \sim N(0,\sigma^2)$ is a random effect representing individual-specific unobservables. Model (10) may be extended in a number of ways. Competing risks arise if an episode can end in more than one transition or type of event, in which case y_{tji} is categorical and (10) can be generalised to a multinomial logit model (Steele et al., 1996). Another extension is to simultaneously model transitions between multiple states, for example employment and unemployment. A general multilevel discrete-time model for repeated events, competing risks and multiple states is described by Steele et al. (2004).

3.3 Causal event history models

Although most event history analyses focus on a single event process, it is common to include as time-varying covariates outcomes of another process. For example, a model of marital dissolution might include indicators of the presence and age of children, and studies of the timing of partnership formation typically allow for effects of enrolment in full-time education. In both cases, these covariates are outcomes of a related, contemporaneous event process, and the timing of events in the two processes may be jointly determined. For instance, the number of children by time interval t constitutes an outcome of the fertility process, and childbearing and partnership decisions may be subject to shared influences, some of which will be unobserved. In other words, fertility outcomes may be endogenous with respect to partnership transitions which, if ignored, may lead to biased estimates of the effects of having children on the risk of marital dissolution.

One way to allow for such endogeneity is to estimate a simultaneous equation model, also called a multiprocess model, which is an event history version of model (8) for bivariate repeated measures data. Suppose that there are repeated events in both processes, e.g. multiple marriages and births in the above example. We denote by $h_{tji}^{(1)}$ and $h_{tji}^{(2)}$ the hazard functions for the two correlated processes. The outcomes of processes 1 and 2 by interval t are denoted by $\mathbf{w}_{tji}^{(1)}$ and $\mathbf{w}_{tji}^{(2)}$. These prior outcomes may refer only to episode j (e.g. the number of children with a given partner j), or they be accumulated across all episodes up to and including j (e.g. the total number of children from all partnerships up to time t). A simultaneous equation model which allows for effects of prior outcomes of one process on the timing of events in the other process is

$$\begin{aligned}\text{logit}[h_{tji}^{(1)}] &= \alpha^{(1)}(t) + \boldsymbol{\beta}^{(1)^T}\mathbf{x}_{tji}^{(1)} + \boldsymbol{\gamma}^{(1)^T}\mathbf{w}_{tji}^{(2)} + u_i^{(1)} \\ \text{logit}[h_{tji}^{(2)}] &= \alpha^{(2)}(t) + \boldsymbol{\beta}^{(2)^T}\mathbf{x}_{tji}^{(2)} + \boldsymbol{\gamma}^{(2)^T}\mathbf{w}_{tji}^{(1)} + u_i^{(2)},\end{aligned} \qquad (11)$$

where $\mathbf{u}_i = [u_i^{(1)} \ \ u_i^{(2)}]^T \sim N(\mathbf{0}, \mathbf{\Omega})$, and the random effect covariance is denoted by $\sigma^{(12)}$. A non-zero random effect covariance suggests a correlation between the unobserved individual-specific determinants of each process.

Model (11) can be estimated using methods for multilevel binary response data. The bivariate responses $(y_{tji}^{(1)}, y_{tji}^{(2)})$ for each interval t are stacked into a single response vector and an indicator variable for each response $y_{tji}^{(l)}$ is interacted with $\mathbf{x}_{tji}^{(l)}$ and $\mathbf{w}_{tji}^{(l)}$ $(l = 1, 2)$. Full details are given in Steele et al. (2005). The model is identified by either the presence of individuals with repeated events or covariate exclusions such that $\mathbf{x}_{tji}^{(1)}$ and $\mathbf{x}_{tji}^{(2)}$ each include at least one variable not contained in the other (Lillard and Waite, 1993; Steele et al., 2005). For instance, Lillard and Waite (1993) used data on multiple marriages and births to identify a simultaneous equation model of marital dissolution and childbearing in the USA, and include state-level measures of the ease and acceptability of divorce (which predict the hazard of dissolution but not a conception) to identify the effect of marital stability on the probability of a conception.

<u>Example</u>: Partnership dissolution and fertility

Steele et al. (2005) used a simultaneous equation event history model to study the interrelationship between fertility and partnership transitions among married and cohabiting British women, building on previous work in the US which considered the link between marital fertility and marital dissolution (Lillard and Waite, 1993). The aim of the analysis was to estimate the effect of the presence and age of children on the risk of partnership breakdown, or the conversion of cohabitation into marriage, at time t. A simultaneous equation model was used to allow for the possibility that the decision to have a child with a partner is jointly determined with the decision to end the partnership or to marry a cohabiting partner. If the unobserved factors driving each process are correlated, and this is ignored in the analysis, estimates of the effect of having children will be biased. The model used by Steele et al. (2005) is an extension of (11) with five equations: three for partnership transitions (dissolution of cohabitation and marriage, and conversion of cohabitation to marriage) and two for fertility (distinguishing marriage and cohabitation). Each equation includes a woman-specific random effect and these may be correlated across equations to allow for residual correlation between processes. Of particular interest are the correlations between the hazard of a particular partnership outcome and the hazard of a conception.

The data came from the National Child Development Study which has as its respondents all those born in a particular week in March 1958 (Shepherd, 1997). Partnership and pregnancy histories were collected retrospectively from respondents at ages 33 and 42. The analysis was based on 5142 women who had 7032 partners during the study period. Prior to analysis, the data were restructured to obtain two responses for each six-month interval of each partnership between ages 16 and 42: 1) an indicator of whether the partnership had dissolved or, for cohabitations, been converted to marriage, and 2) an indicator of a conception. A conception date was calculated as the date of birth minus nine months. Still births and pregnancies that ended in abortion or miscarriage were not considered, mainly because these outcomes do not lead to the presence of children which can affect partnership transitions.

Table 4 shows selected elements of the estimated random effects covariance matrix from Steele et al.'s (2005) analysis. For illustration, we focus on the correlations between partnership dissolution and fertility, distinguishing marriage and cohabitation. There are significant, positive correlations between the chance of conceiving in cohabitation and the risk of dissolution from both marriage and cohabitation. However, the correlations between the chance of a marital conception and dissolution from either form of partnership are both small and non-significant. These findings suggest that women with an above-average risk of dissolution (that is, prone to unstable partnerships) tend to have an above-average chance of conceiving during cohabitation.

Estimates of the effects of the presence of children on the logit-hazard that a cohabitation breaks down are given in Table 5. Controls for partnership duration at t and family background are also included in the model, but their coefficients have been suppressed (see Steele et al. (2005) for further results). The results from two model specifications are compared. In the first model, a standard multilevel event history model, the residual correlations between partnership transitions and fertility were constrained to zero which is equivalent to estimating the partnership equations independently of the conception equations. The second model is the simultaneous equation model in which all random effect correlations were freely estimated. From either model, we would conclude that pregnancy or having young children together reduces a cohabiting couple's risk of separation. Nevertheless, the effects obtained from the multiprocess model are slightly more pronounced, due to the positive residual correlation between the chance of a conception and the risk of dissolution (Table 4). In the single-process model, the negative effects of pregnancy and the presence of children are subject to selection bias. The disproportionate presence of women prone to unstable partnerships in the 'pregnancy' and 'having children with the current partner' categories inflates the risk of separation in these categories. Thus, the "true" negative effects of these time-varying indicators of fertility are understated.

The findings for this British cohort contrast with those of Lillard and Waite (1993) for the USA. They found a strong *negative* residual correlation (ρ=-0.86, SE=0.15) between the risk of marital dissolution and the probability of conception within marriage. A negative correlation implies that women with an above average risk of experiencing marital breakdown (on unmeasured time-invariant characteristics) are also less likely to have a child within marriage. Allowing for this source of endogeneity revealed a stabilising effect on marriage of having more than one child.

4. Discussion

It is now widely recognised that observational studies require information on individual change and the relative timing of events in order to investigate questions about causal relationships. Consequently there has been a large amount of investment in the collection of longitudinal data, in the form of both prospective panel data and retrospective event history data. These data have a hierarchical structure which can be analysed using a general class of multilevel models. The aim of this paper has been to show how multilevel modelling – which is fast becoming a standard technique in many social and medical researchers' repertoire – can be used to exploit the richness of longitudinal data on change and event processes.

The simplest model for change fits a growth curve to each individual's repeated measures, and is an example of a two-level random coefficient model. Generalisations to more complex data structures, discrete responses, and simultaneous analysis of multiple change processes are straightforward applications of established multilevel modelling techniques. One example of longitudinal discrete responses is interval-censored event history data. Methods for the analysis of multilevel discrete response models can be applied in the analysis of repeated events, with extensions to handle competing risks, transitions between multiple states, and correlated event processes. All of these analyses can now be performed using mainstream and specialist statistical software. Repeated measures data can also be conceptualised as multiple indicators of underlying latent variables. A structural equation modelling approach is especially fruitful when responses or predictors are measured indirectly by a set of indicators.

Previous authors have demonstrated the equivalence of the multilevel and structural equation modelling approaches to fitting certain types of growth curve models, and in recent years these powerful techniques have converged further. On the multilevel modelling side, early work by McDonald and Goldstein (1989) on multilevel factor analysis has been extended to handle mixtures of continuous, binary and ordinal indicators (Goldstein and Browne, 2005; Steele and Goldstein, 2006) and structural dependencies (Goldstein et al., 2007). Structural equation models have been extended to allow for hierarchical structures using multiple-group analysis (Muthén, 1989). Both types of model can be embedded in the generalised linear latent and mixed modelling (GLLAMM) framework proposed by Rabe-Hesketh et al. (2004) and implemented in the `gllamm` program via Stata (StataCorp, 2005). The GLLAMM approach does not distinguish between random effects in multilevel models and factors in structural equation models, but allows complete flexibility in the specification of the loadings attached to latent variables. Thus a multilevel random effect is fitted by defining a latent variable with all loadings constrained to equal one, and a common factor is fitted by allowing at least one of the loadings to be freely estimated.

Acknowledgements

The extremely helpful comments from two referees are gratefully acknowledged.

References

Allison P.D. (1982) Discrete-time methods for the analysis of event histories. In: *Sociological Methodology* (ed. Leindhardt S), pp. 61-98. Jossey-Bass, San Francisco

Arbuckle J.L. (1996) Full information estimation in the presence of incomplete data. In: *Advanced Structural Equation Modeling* (eds. Marchoulides GA & Schumacker RE), pp. 243-278. Lawrence Erlbaum Associates, Hillsdale, NJ

Bollen K.A. and Curran P.J. (2004) Autoregressive latent trajectory (ALT) models a synthesis of two traditions. *Sociological Methods & Research*, **32**, 336-383.

Bollen K.A. and Curran P.J. (2006) *Latent Curve Models: A Structural Equation Perspective*. John Wiley & Sons, Inc., Hoboken, New Jersey.

Browne W.J. (2004) *MCMC Estimation in MLwiN*. Institute of Education, London.
Carpenter J. and Goldstein H. (2004) Multiple imputation in MLwiN. *Multilevel Modelling Newsletter*, **16**
Curran P.J. (2003) Have multilevel models been structural equation models all along? *Multivariate Behavioral Research*, **38**, 529-568.
Diggle P., Heagerty P., Liang K.-Y. and Zeger S. (2002) *Analysis of Longitudinal Data*. 2nd edn. Oxford University Press, Oxford.
Goldstein H. (2003) *Multilevel Statistical Models*. 3rd edn. Arnold, London.
Goldstein H., Bonnet G. and Rocher T. (2007) Multilevel structural equation models for the analysis of comparative data on educational performance. *Journal of Educational and Behavioral Statistics (to appear)*
Goldstein H. and Browne W.J. (2005) Multilevel factor analysis for continuous and discrete data. In: *Contemporary Psychometrics: A Festschrift for Roderick P. McDonald* (eds. Maydeu-Olivares A & McArdle JJ), pp. 453-475. Lawrence Erlbaum, New Jersey
Goldstein H., Healy M.J.R. and Rasbash J. (1994) Multilevel Time-Series Models with Applications to Repeated-Measures Data. *Statistics in Medicine*, **13**, 1643-1655.
Laird N.M. and Ware J.H. (1982) Random-Effects Models for Longitudinal Data. *Biometrics*, **38**, 963-974.
Lillard L.A. and Waite L.J. (1993) A joint model of marital childbearing and marital disruption. *Demography*, **30**, 653-681.
McDonald R.P. and Goldstein H. (1989) Balanced Versus Unbalanced Designs for Linear Structural Relations in 2-Level Data. *British Journal of Mathematical & Statistical Psychology*, **42**, 215-232.
Muthén B. (1997) Latent variable modeling of longitudinal and multilevel data. In: *Sociological Methodology 1997, Vol 27*, pp. 453-480
Muthén B.O. (1989) Latent Variable Modeling in Heterogeneous Populations. *Psychometrika*, **54**, 557-585.
Muthén B.O. (1994) Multilevel Covariance Structure-Analysis. *Sociological Methods & Research*, **22**, 376-398.
Rabe-Hesketh S., Skrondal A. and Pickles A. (2004) Generalized Linear Structural Equation Modelling. *Psychometrika*, **69**, 167-190.
Rasbash J., Steele F., Browne W.J. and Prosser B. (2004) *A User's Guide to MLwiN, v2.0*. University of Bristol, Bristol.
Raudenbush S.W. and Bryk A.S. (2002) *Hierarchical Linear Models*. Sage, Newbury Park.
Schafer J. (1997) *Analysis of Incomplete Multivariate Data*. Chapman & Hall, New York.
Schafer J.L. and Yucel R.M. (2002) Computational strategies for multivariate linear mixed-effects models with missing values. *Journal of Computational and Graphical Statistics*, **11**, 437-457.
Shepherd P. (1997) The National Child Development Study: An Introduction to the Origins of the Study and the Methods of Data Collection. In. Centre for Longitudinal Studies, Institute of Education, University of London.
Singer J.D. and Willett J.B. (2003) *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford University Press, New York.
Skrondal A. and Rabe-Hesketh S. (2004) *Generalized Latent Variable Modelling: Multilevel, Longitudinal, and Structural Equation Models*. Chapman & Hall/CRC, Boca Raton, FL.
Spiegelhalter D.J., Thomas A. and Best N.G. (2000) WinBUGS Version 1.3 User Manual. In. Medical Research Council Biostatistics Unit, Cambridge
StataCorp (2005) *Stata 9.0 Base Reference Manual.* Stata Press, College Station, TX.

Steele F., Diamond I. and Wang D.L. (1996) The determinants of the duration of contraceptive use in China: A multilevel multinomial discrete-hazards modeling approach. *Demography*, **33**, 12-23.
Steele F. and Goldstein H. (2006) A multilevel factor model for mixed binary and ordinal indicators of women's status. *Sociological Methods & Research*, **35**, 137-153.
Steele F., Goldstein H. and Browne W.J. (2004) A general multistate competing risks model for event history data, with an application to a study of contraceptive use dynamics. *Statistical Modelling*, **4**, 145-159.
Steele F., Kallis C., Goldstein H. and Joshi H. (2005) The relationship between childbearing and transitions from marriage and cohabitation in Great Britain. *Demography*, **42**, 647-673.
UK Data Archive (2004). ESDS Longitudinal, *British Household Panel Survey; Waves 1-11, 1991-2002: Teaching Dataset (Social and Political Attitudes)* [computer file]. University of Essex. Institute for Social and Economic Research, [original data producer(s)]. Colchester, Essex: UK Data Archive [distributor], November 2004. SN: 5038.

Table 1. Likelihood ratio tests, comparing difference growth curves models fitted to boys' heights

Model	No. parameters in $\mathbf{\Omega}_u$	-2 log L	Δ -2 log L	d.f.	p-value
1: Variance only in α_0	1	929.7	-	-	
2: 1+Variance in α_1	3	675.5	254.2	2	<0.001
3: 2+Variance in α_2	6	628.5	47.0	3	<0.001
4: 3+Variance in α_3	10	620.9	7.6	4	0.109

Note: Each model extends the previous model by allowing for an extra random coefficient. For example, Model 1 includes only a random intercept term u_{0i}, while Model 2 has an additional random effect u_{1i} for the coefficient of z_{ti}. Δ -2 log L relates to the decrement in the -2 log-likelihood value between the relevant model and the model in the previous row.

Table 2. Cubic polynomial growth curve fitted to boys' heights

Parameter	Estimate	(SE)
Fixed		
α_0 (intercept)	149.01	(1.54)
α_1 (age)	6.17	(0.35)
α_2 (age^2)	0.75	(0.18)
α_3 (age^3)	0.46	(0.16)
Random		
Between-individual variation		
σ_{u0}^2 (intercept)	61.58	(17.10)
σ_{u10}	8.00	(3.03)
σ_{u1}^2 (age)	2.76	(0.78)
σ_{u20}	1.37	(1.41)
σ_{u21}	0.88	(0.34)
σ_{u2}^2 (age^2)	0.63	(0.22)
Within-individual variation		
σ_e^2	0.22	(0.02)

Table 3. Multilevel structural equation model fitted to social and political items from five waves of the British Household Panel Study

Measurement model	λ_{0k} [a]	(SE) [b]	λ_{1k}	(SE)	σ^2_{uk}	(SE)	$\sigma^2_{\varepsilon k}$	(SE)
1. Ordinary people share wealth	3.55	(0.01)	1 [c]	-	0.18	(0.01)	0.45	(0.01)
2. One law for rich, one for poor	2.40	(0.01)	-1.11	(0.03)	0.22	(0.01)	0.45	(0.01)
3. Private enterprise is solution	2.96	(0.01)	1.04	(0.03)	0.21	(0.01)	0.43	(0.01)
4. Public services to be state owned	3.00	(0.01)	-1.09	(0.03)	0.24	(0.01)	0.57	(0.01)
5. Govt obliged to provide jobs	3.06	(0.01)	-1.05	(0.04)	0.43	(0.01)	0.48	(0.01)
6. Strong unions protect employees	2.84	(0.01)	-1.08	(0.04)	0.37	(0.01)	0.46	(0.01)

Structural model	Est.	(SE)
α_1 (1994 vs. 1992)	0.10	(0.01)
α_2 (1996 vs. 1992)	0.16	(0.01)
α_3 (1998 vs. 1992)	0.02	(0.01)
α_4 (2001 vs. 1992)	0.05	(0.01)
σ^2_{u0}	0.20	(0.01)
σ^2_e	0.03	(0.002)

[a] Point estimates are means of parameter values from 15,000 MCMC samples.
[b] Standard errors are standard deviations of parameter values from MCMC samples.
[c] Constrained parameter.

Table 4. Selected residual covariances from a multiprocess model of partnership transitions and fertility among women of the National Child Development Study, age 16-42

	Dissolution of cohabitation	Dissolution of marriage
Conception in cohabitation	0.131[a]	0.217
	(0.027, 0.243)[b]	(0.074, 0.357)
	0.316[c]	**0.425**
Conception in marriage	-0.009	-0.017
	(-0.0045, 0.025)	(-0.062, 0.027)
	-0.048	**-0.071**

Source: Extract from Table 5 of Steele et al. (2005).

[a] The point estimate of the covariance (the mean of the MCMC samples)
[b] The 95% interval estimate for the covariance
[c] The point estimate of the correlation (the mean of the correlation estimates across samples).

Table 5. Multilevel discrete-time event history analysis of the effects of fertility outcomes on the logit-hazard of dissolution of cohabitation among women of the NCDS, age 16-42

	Single process model		Multiprocess model	
	Estimate[a]	(SE)[b]	Estimate	(SE)
Currently pregnant	-0.639	(0.150)	-0.701	(0.156)
No. preschool with current partner				
1	-0.236	(0.120)	-0.290	(0.120)
2+	-0.753	(0.261)	-0.877	(0.270)
No. older with current partner				
1	-0.032	(0.202)	-0.058	(0.208)
2+	0.239	(0.333)	0.136	(0.341)
Preschool child(ren) with previous partner	-0.330	(0.218)	-0.335	(0.224)
Older child(ren) with previous partner	-0.012	(0.128)	-0.022	(0.130)
Child(ren) with non co-resident partner	-0.019	(0.191)	-0.018	(0.194)

Source: Extract from Table 7 of Steele et al. (2005).

[a] Parameter estimates are means of parameter values from 20,000 MCMC samples, with a burn-in of 5000.
[b] Standard errors are standard deviations of parameter values from MCMC samples.

Figure 1. Between-individual variance in boys' heights as a function of age