Muthén, B. (1996). Growth modeling with binary responses. In A. V. Eye, & C. Clogg (Eds.), Categorical Variables in Developmental Research: Methods of Analysis (pp. 37-54). San Diego, CA: Academic Press. (#64)

2 Growth Modeling with Binary Responses

Bengt O. Muthén
University of California, Los Angeles
Los Angeles, California 90024

1. INTRODUCTION

In developmental research it is natural to place an emphasis on the study of individual differences in change, growth, and decline over time. It is useful to formulate a longitudinal model for these processes to assess the amount of individual variation and to relate the individual variation to background information on the individuals. Random coefficient growth modeling (see, e.g., Laird & Ware, 1982; Rutter & Elashoff, 1994) is suitable for such longitudinal analysis. It goes beyond conventional structural equation modeling of longitudinal data with its focus on autoregressive models (see, e.g., Jöreskog & Sörbom, 1977; Wheaton, Muthén, Alwin, & Summers, 1977) in that it describes individual differences in the longitudinal processes. It is instructive to consider first some examples of longitudinal studies with binary responses in which random coefficient modeling has been used.

1.1. Example 1: Decline in Depression

Gibbons and Bock (1987) reported on the longitudinal analyses of data on Danish psychiatric patients. A total of 100 clinically depressed patients were given one of two drugs and were observed five times weekly. The response variable was recorded as recovered (scored 0) or still depressed (scored 1). Information was collected on time-invariant covariates such as severity of illness and age. Time-varying covariates included the plasma level of the drug. The patients were divided into two groups. One group received the drug imipramine and the other received chlorimipramine. The object of the study was to find out which of the two drugs resulted in the sharpest decline in the probability that the patient still felt depressed.

1.2. Example 2: Change/Stability of Neuroticism

Muthén (1983) studied the data described in Henderson, Byrne, and Duncan-Jones (1981) for 231 Canberra adults interviewed four times at 4-month intervals regarding aspects of "neurotic illness." In a short form of a general health questionnaire, the four questions asked were "In the last month have you suffered from any of the following? Anxiety. Depression. Irritability. Nervousness." A yes response was denoted 1 and no was denoted 0. Time-invariant covariates included gender and a measure of long-term susceptibility to neurosis (the *N* scale from the Eysenck Personality Inventory). Time-varying covariates included life events in the four months prior to the interview. The object of the study was to assess the stability over time of the level of neuroticism of this population of individuals.

1.3. Example 3: Correlated Observations on Asthma Attacks

Stiratelli, Laird, and Ware (1984) studied data on daily observations of 64 asthmatics living in Garden Grove, California. The response variable was the presence or the absence of an asthma attack recorded over a period of about 7 months. Time-invariant covariates included gender, age, and history of hay fever. Time-varying covariates included air pollution and weather conditions. The object of the study was to assess the relative importance of various risk factors for increased probability of asthma attacks.

1.4. Contrasting the Examples

It is interesting to contrast these three examples. Example 3 illustrates the fact that often the longitudinal structure of the data is only a nuisance. Here, the interest is the same as in regression analysis. However, observations over time for the same individual are correlated so that the usual assumption of independent observations does not hold. The focus is on how to do the regression analysis while properly taking into account the nonindependence. In Example 2, the longitudinal structure of the data is not a nuisance but is essential to the analysis. The longitudinal process is one in which no particular trend over time is expected; the interest is in assessing how much responses vary over time for a typical individual. Observations fluctuate up and down over time for each individual and it is the amount of fluctuation that is the focus of the study. Example 1 involves a further elaboration of the longitudinal study. Here, observations do not only fluctuate over time for a given individual but also follow a decreasing trend over time. This chapter focuses on situations illustrated by Examples 1 and 2.

The aim of this chapter is to discuss conventional random effects modeling for binary longitudinal response and to compare that with a generalized random effects model for longitudinal data which draws on techniques used in latent variable modeling. In section 2, the conventional modeling and estimation is presented. Section 3 critiques this approach and gives a more general formulation. Section 4 presents a small Monte Carlo study in which the more general approach is studied and analyses of real data is also presented.

2. CONVENTIONAL MODELING AND ESTIMATION WITH BINARY LONGITUDINAL DATA

Consider a binary variable y and a corresponding continuous latent response variable y^* for which τ is a threshold parameter determining the y outcomes: $y = 1$ when $y^* > \tau$ and $y = 0$ otherwise. Here, the progress over time of the latent response variable y^* is described as

$$y^*_{ik} = \alpha_i + \beta_i t_k + \gamma_k v_{ik} + \zeta_{ik}, \tag{1}$$

where i denotes an individual, t_k denotes a time-related variable with $t_k = k$ (e.g., $k = 0, 1, 2, \ldots, K - 1$), α_i is a random intercept at $t = 0$, β_i is a random slope, γ_k are fixed slopes, v_{ik} is a time-varying covariate, and ζ_{ik} is a residual, $\zeta \sim N(0, \psi_{\zeta\zeta})$. Furthermore,

$$\begin{cases} \alpha_i = \mu_\alpha + \pi_\alpha w_i + \delta_{\alpha_i} \\ \beta_i = \mu_\beta + \pi_\beta w_i + \delta_{\beta_i}, \end{cases} \tag{2}$$

where μ_α, μ_β, π_α, and π_β are parameters, w_i is a time-invariant covariate, and δ_α, δ_β are residuals assumed to have a bivariate normal distribution,

$$\begin{pmatrix} \delta_{\alpha_i} \\ \delta_{\beta_i} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \psi_{\alpha\alpha} & \psi_{\beta\alpha} \\ \psi_{\beta\alpha} & \psi_{\beta\beta} \end{pmatrix}\right]. \tag{3}$$

With $t_k = k$ and a linear function of time, for example, $k = 0, 1, 2, \ldots, K - 1$, the variables α and β can be interpreted as the initial status level and the rate of growth/decline, respectively.

The residuals of ζ_{ik} are commonly assumed to be uncorrelated across time. In line with Gibbons and Bock (1987), however, a first-order autoregressive structure over time for these residuals is presented. Letting $x = (w, v)'$, the model implies multivariate normality for y^* conditional on x with

$$\mathcal{E}\begin{pmatrix} y^*_{i0}\,|\,x \\ y^*_{i1}\,|\,x \\ \vdots \\ y^*_{iK-1}\,|\,x \end{pmatrix} = T\begin{pmatrix} \mu_\alpha + \pi_\alpha w_i \\ \mu_\beta + \pi_\beta w_i \end{pmatrix} + \begin{pmatrix} \gamma_0 v_{i0} \\ \gamma_1 v_{i1} \\ \gamma_2 v_{i2} \\ \vdots \\ \gamma_{K-1} v_{iK-1} \end{pmatrix} \tag{4}$$

and

$$V(y^*\,|\,x) = T\Psi_{\alpha/\beta}T' + \psi_{\zeta\zeta}\begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{K-1} \\ \rho & 1 & \rho & \cdots & \rho^{K-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{K-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{K-1} & \rho^{K-2} & \rho^{K-3} & \cdots & 1 \end{pmatrix}, \tag{5}$$

where with linear growth or decline

$$T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & K-1 \end{pmatrix} \tag{6}$$

and $\Psi_{\alpha/\beta}$ is the 2×2 covariance matrix in Eq. (3).

For given t_k, w, and v, the model expresses the probability of a certain observed response y_{ik} as a function of the random coefficients α and β,

$$\begin{aligned} P(y_{ik} = 1\,|\,\alpha_i, \beta_i, x) &= P(y^*_{ik} > \tau\,|\,\alpha_i, \beta_i, x) \\ &= \int_\tau^\infty \varphi(s\,|\,z, \psi_{\zeta\zeta})ds \\ &= \Phi\left(\frac{-\tau + z}{\psi_\zeta}\right), \end{aligned} \tag{7}$$

where φ is a (univariate) normal density, ψ_ζ denotes the standard deviation of ζ, and

$$z = \alpha_i + \beta_i t_k + \gamma_k v_{ik}. \tag{8}$$

To identify the model, the standardization $\tau = 0$, $\psi_\zeta = 1$ can be used as in conventional probit regression (see, e.g., Gibbons & Bock, 1987).

The probability of a certain response may be expressed as

$$P(y_0, \ldots, y_{K-1}\,|\,x) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} P(y_0, y_1, \ldots, y_{K-1}\,|\,\alpha, \beta, x)\,\varphi(\alpha, \beta\,|\,x)d\alpha\, d\beta, \tag{9}$$

where

$$P(y_0, y_1, \ldots y_{K-1}\,|\,\alpha, \beta, x) = \int_{c^-(y_{i0})} \cdots \int_{c^-(y_{iK-1})} \varphi(y^*_0, y^*_1, \ldots, y^*_{K-1}\,|\,\alpha, \beta, x)dy^*_{i0} \ldots dy^*_{iK-1}, \tag{10}$$

where $c^-(y_{ik})$ denotes the integration domain for y^*_{ik} given that the kth variable takes on the value y_{ik}. Here, the integration domain is either $(-\infty, \tau)$ or $(\tau, +\infty)$.

In the special case of uncorrelated residuals, that is, $\rho = 0$ in Equation (5), the y^* variables are independent when conditioning on α, β, and x so that $P(y_0, y_1, \ldots, y_{K-1}\,|\,\alpha, \beta, x)$ simplifies considerably,

$$P(y_0, y_1, \ldots, y_{K-1}\,|\,\alpha, \beta, x) = \prod_{k=0}^{K-1} \int_{c^-(y_{ik})} \varphi(y^*_{ik}\,|\,\alpha, \beta, x)dy^*_{ik}. \tag{11}$$

In this case, only univariate normal distribution functions are involved so that the essential computations of Equation (9) involve the two-dimensional integral over α and β.

Perhaps because of the computational simplifications, the special case of $\rho = 0$ appears to be the standard model used in growth analysis with binary response. This model was used in Gibbons and Bock (1987; see also Gibbons & Hedeker, 1993). The analogous model with logit link was studied in Stiratelli et al. (1984) and in Zeger and Karim (1991). Gibbons and Bock (1987) considered maximum likelihood estimation using Fisher scoring and EM procedures developed for binary factor analysis in Bock and Lieberman (1970) and Bock and Aitkin (1981). Stiratelli et al. (1984) considered restricted maximum likelihood using the EM algorithm. Gibbons and Bock (1987) used a computational simplification obtained by orthogonalizing the bivariate normal variables α and β using a Cholesky factor so that the bivariate normal density is written as a product of two univariate normal densities. They used numerical integration by Gauss-Hermite quadrature, with the weights being the product of the one-dimensional weights. For the case of $\rho \neq 0$, Gibbons and Bock (1987) used the Clark algorithm to approximate the probabilities of the multivariate normal distribution for y^* in Equation (10). Even when $\rho = 0$ the computations are heavy when there is a large number of distinct x values in the sample. Zeger and Karim (1991) employed a Bayesian approach using the Gibbs sampler algorithm. For recent overviews, see Fitzmaurice, Laird, and Rotnitzky (1993); Longford (1993); Diggle, Liang, and Zeger (1994); and Rutter and Elashoff (1994).

3. MORE GENERAL BINARY GROWTH MODELING

3.1. Critique of Conventional Approaches

In this section, weaknesses in conventional growth modeling with binary data are presented, along with a more general model and its estimation.

The maximum likelihood approach to binary growth modeling leads to heavy computations when $\rho \neq 0$. This seems to have caused a tendency to restrict modeling to an assumption of $\rho = 0$. Experience with continuous response variables, that is, when $y \equiv y^*$, indicates that $\rho = 0$ is not always a realistic assumption. The assumption of a single ρ parameter that is different from zero, as in the Gibbons-Bock first-order autoregressive model, also may not be realistic in some cases. Instead, it appears necessary to include a separate parameter for at least the correlations among residuals that are adjacent in time.

Furthermore, the conventional model specification of $\tau = 0$, $\psi_\zeta = 1$ has no effect when, as in standard probit regression, there is only a single equation that is being estimated. It is important to note, however, that this is not the case in longitudinal analysis. The longitudinal analysis can be characterized as a multivariate probit regression in which the multivariate response consists of the same response variable at different time points. This has the following consequences.

First, the standardization of τ to zero at all time points needs clarification. In the binary case, this does not lead to incorrect results but does not show the generalization to the case of ordered categorical response or to the case of multiple indicators. The threshold τ is a parameter describing a measurement characteristic of the variable y, namely, the level (proportion) of y with zero values on x. Because the same y variable is measured at all time points, equality of this measurement characteristic over time is the natural model specification. In the binary case, however, the equality of the level of y across time points is accomplished by μ_α in Equation (2) affecting y equally over time as a result of the unit coefficient of α_i in Equation (1), which is not explicitly shown. Setting $\tau = 0$ is therefore correct, although an equivalent specification would take τ as a parameter held equal over time points while fixing μ_α at zero. In the ordered categorical case, however, there are several τ parameters involved for a y variable and equality over time of such τ's is called for. In this case, μ_α cannot be separately estimated but may be fixed at zero. The multiple indicator case will be discussed in the next section.

Second, ψ_ζ is the standard deviation of the residual variation of the latent response variable y^*, and fixing it at unity implicitly assumes that the residual variation has the same value over time. This is not realistic because over time different sources of variation not accounted for by the time-varying variable v_{ik} are likely to be introduced. Again, experiences with continuous response variables indicate that the residual variance often changes over time.

In presentations using the logit version of the model, the parameters of τ and ψ_ζ are usually not mentioned (see, e.g., Diggle et al., 1994). This is probably because the threshold formulation, often used in the probit case, is seldom used in the logit case. This has inadvertently lead to an unnecessarily restrictive logit formulation in growth modeling.

3.2. The Approach of Muthén

An important methodological consideration is whether computational difficulties should lead to a simplified model, such as using $\rho = 0$ or $\psi_\zeta = 1$, or whether it is better to maintain a general model and instead use a simpler estimator. Here, I describe the latter approach, building on the model of Equations (1) and (2) to consider a more general model and a limited-information estimator.

First, the ζ_{ik} variables of Equation (1) are allowed to be correlated among themselves and are allowed to have different variances over time. Second, multiple indicators y_{ikj}, $j = 1, 2, \ldots, p$ are allowed at each time point,

$$y^*_{ikj} = \lambda_j \eta_{ik} + \epsilon_{ikj}, \tag{12}$$

where λ_j is a measurement (slope) parameter for indicator j, ϵ_{ikj} is a measurement error residual for variable j at time k, and $y_{ikj} = 1$ if $y^*_{ikj} > \tau_j$. The multiple indicator case is illustrated by Example 2 in which four measurements of a single construct "neurotic illness" (η) were considered. Given that the τ's and the λ's are measurement parameters, a natural model specification would impose equality over time for each of these parameters. Using normality assumptions for all three types of residuals, ζ, δ, and ϵ, again leads to a multivariate probit regression model.

This generalized binary growth model is a special case of the structural equation model of Muthén (1983, 1984). The longitudinal modeling issues just discussed were also brought up in Muthén (1983), where a random intercept model like Equations (1), (2), and (12) was fitted to the Example 2 data. The problems with standardization issues related to τ and ψ_ζ have also been emphasized by Arminger (see, e.g., ch. 3, this volume) and Muthén and Christofferson (1981) in the context of structural equation modeling.

In the approach of Muthén (1983, 1984), conditional mean and covariance matrix expressions corresponding to Equations (4) and (5) are considered. This is sufficient given the conditional normality assumptions. Muthén (1983, 1984) introduced a diagonal scaling matrix Δ containing the inverse of the conditional standard deviations of the latent response variable at each time point,

$$\Delta = diag[V(y^* \mid x)]^{-1/2}. \tag{13}$$

Muthén (1983, 1984) describes three model parts. Using the single-indicator growth model example of Equations (4) and (5),

$$\sigma_1 = \Delta T \mu \tag{14}$$
$$\sigma_2 = \Delta[T \pi \ \gamma] \tag{15}$$
$$\sigma_3 = \Omega = \Delta(T \Psi_{\alpha/\beta} T' + \Psi_{\zeta\zeta})\Delta, \tag{16}$$

where $\Psi_{\zeta\zeta}$ is the $K \times K$ covariance matrix of ζ (cf. Eq. 5). The three parts correspond to the intercepts, slopes, and residual correlation matrix of a multivariate probit regression. Note that Ω is a correlation matrix.

The general model of Muthén (1983, 1984), including multiple indicators as well as multiple factors at each time point, can be expressed as follows. Consider a set of measurement relations for a p-dimensional vector y^*,

$$y^* = \Lambda\eta + \epsilon, \tag{17}$$

and a set of structural relations for an m-dimensional vector of latent variable constructs η,

$$\eta = \alpha + \beta\eta + \Gamma x + \zeta, \tag{18}$$

where Λ, α, B, Γ are parameter arrays, ϵ is a residual (measurement error) vector with mean zero and covariance matrix Θ, and ζ is a residual vector with mean zero and covariance matrix Ψ. The scaling matrix Δ is also included in this general framework, as is the ability to analyze independent samples from multiple populations simultaneously.

Muthén (1983, 1984, 1987) used a least-squares estimator where with $\sigma = (\sigma_1', \sigma_2', \sigma_3')'$,

$$F = (s - \sigma)' \, W^{-1}(s - \sigma), \tag{19}$$

where the s elements are arranged in line with σ and are maximum likelihood estimates of the intercepts, slopes, and residual correlations. Here, s_1 and s_2 are estimates from probit regressions of each y variable on all the x variables, whereas each s_3 element is a residual correlation from a bivariate probit regression of a pair of y variables regressed on all x variables. A generalized least-squares estimator is obtained when the weight matrix W is a consistent estimate of the asymptotic covariance matrix of s. In this case, a chi-square test of model fit is obtained as $n \cdot F$, where n is the sample size and F refers to the minimum value of the function in Equation (19). (For additional technical details on the asymptotic theory behind this approach, see Muthén & Satorra, 1995.) Muthén (1984, 1987) presented a general computer program LISCOMP which carries out these calculations.

3.3. Model Identification

Given that the response variables are categorical, the general binary growth model needs to be studied carefully in terms of parameter identification. Under

the normality assumptions of the model, the number of distinct elements of σ represents the total number of parameters that can be identified, therefore the number of growth model parameters can be no larger than this. The growth model parameters are identified if and only if they are identified in terms of the elements of σ.

It is instructive to consider first the conditional y^* variance in some detail for the case of binary growth modeling. Let $[\Delta]_{kk}^{-2}$ denote the conditional variance of y_{kk}^* given x. For simplicity, the focus is on the case with linear growth. With four time points, the conditional variances of y^* can be expressed in model parameter terms as

$$[\Delta]_{00}^{-2} = \psi_{\alpha\alpha} + \psi_{\zeta_0\zeta_0} \tag{20}$$
$$[\Delta]_{11}^{-2} = \psi_{\alpha\alpha} + 2\psi_{\beta\alpha} + \psi_{\beta\beta} + \psi_{\zeta_1\zeta_1} \tag{21}$$
$$[\Delta]_{22}^{-2} = \psi_{\alpha\alpha} + 4\psi_{\beta\alpha} + 4\psi_{\beta\beta} + \psi_{\zeta_2\zeta_2} \tag{22}$$
$$[\Delta]_{33}^{-2} = \psi_{\alpha\alpha} + 6\psi_{\beta\alpha} + 9\psi_{\beta\beta} + \psi_{\zeta_3\zeta_3}. \tag{23}$$

Note that the Δ elements are different because of across-time differences in contributions from $\Psi_{\alpha/\beta}$ as well as $\psi_{\zeta\zeta}$. In Equation (5), using the Gibbons-Bock standardization of $\psi_\zeta = 1$, there are no free $\psi_{\zeta\zeta}$ parameters to be estimated. Contrary to this conventional approach, there are four different $\psi_{\zeta\zeta}$ parameters in Equations (18) through (21). Because the y^* variables are not directly observed, not all of these parameters are identifiable. Instead of assuming $\psi_\zeta = 1$ for all time points, as in Gibbons-Bock, the first diagonal element of Δ can be fixed to unity, corresponding to the first time point. For the remaining time points, the Δ elements in Equations (19) through (21) are the unrestricted parameters instead of the residual y^* variances of ζ. The residual variances are not taken as free parameters to be estimated, but can be obtained from the other parameters using Equations (18) through (21). Allowing the Δ parameters to be different across time allows the residual variances to be different across time. With four time points, this adds three parameters to the model relative to the conventional model.

The (co)variance-related parameters of the model are in this case the three free Δ elements (not the ψ_ζ's) and the three elements of $\Psi_{\alpha/\beta}$. With covariates, added parameters are μ_α, μ_β, π_β, π_β, and $\gamma_0, \ldots, \gamma_{K-1}$. It can be shown that with four time points, covariances between pairs of residuals at adjacent time points can also be identified in this model (see Muthén & Liu, 1994).

As opposed to the case of a single response variable, multiple indicator models allow for separate identification of the residual variances and the measurement errors of each indicator. In this case, there is the additional advantage that the residual variances for the latent variable constructs η are identified at all time points. Multiple-indicator models would assume equality of the measurement parameters (the τ's and the λ's) for the same response variable across time. In this case, the μ_α intercept is fixed at zero. The Δ matrix scaling is generalized as follows. The scaling factors of Δ are fixed at the first time point for all of the indicators to eliminate the indeterminacy of scale for each different y^* variable (corresponding to each indicator) at this time point. The scaling factors of Δ are free for all indicators at later time points so that the measurement error variances are not restricted to be equal across time points for the same response variable.

3.4. Implementation in Latent Variable Modeling Software

In the case of continuous response variables, Meredith and Tisak (1984, 1990) have shown that the random coefficient model of the previous section can be formulated as a latent variable model. For applications in psychology, see McArdle and Epstein (1987); for applications in education, see Muthén (1993) and Willett and Sayer (1993); and for applications in mental health, see Muthén (1983, 1991). For a pedagogical introduction to the continuous case, see Muthén, Khoo, and Nelson Goff (1994) and Willett and Sayer (1993). Muthén (1983, 1993) pointed out that this idea could be carried over to the binary and ordered categorical case. The basic idea is easy to describe. In Equation 1, α_i is unobserved and varies randomly across individuals. Hence, it is a latent variable. Furthermore, in the product term $\beta_i t_k$, β_i is a latent variable multiplied by a term t_k which is constant over individuals and can therefore be treated as a parameter. The $t_k s$ may be fixed as in Equation (6), but with three or more time points they may be estimated for the third and later time points to represent nonlinear growth. More than one growth factor may also be used.

4. ANALYSES

Simulated and real data will now be used to illustrate analyses using the general growth model with binary data.

4.1. A Monte Carlo Study

A limited Monte Carlo study was carried out to demonstrate the sampling behavior of the generalized least-squares estimator in the binary case. The model chosen for the study has a single binary response variable observed at four time points. There is one time-invariant covariate and one time-varying covariate (one for each time point). The simulated data can be thought of as being in line with the Example 1 situation in which the probability of a problem behavior declines over time. Linear decline is specified with T as in Equation (6). Both the random intercept (α) and the random slope (β) show individual variation as represented both by their common dependence on the time-invariant covariate (w) and their residual variation (represented by $\Psi_{\alpha/\beta}$). The α variable regression has a

positive intercept (μ_α) and a positive slope (π_α) for w, whereas the β variable regression has a negative intercept (μ_β) and a negative slope (π_β) for w. In this way, the time-invariant covariate can be seen as a risk factor, which with increasing value increases α, that is, it increases the initial probability of the problem and decreases β, making the rate of decline larger (this latter means that the higher the risk factor value, the more likely the improvement is in the problem behavior over time). The regression coefficient for the response variable on the time-varying covariates (the v_k's) is positive and the same at all time points. The residual variances ($\psi_{\zeta\zeta}$) are changing over time and there is a nonzero residual covariance between adjacent pairs of residuals that is assumed to be equal. The time-varying covariates are correlated 0.5 and are each correlated 0.25 with the time-invariant covariate. All covariates have means of 0 and variances of 1. The population values of the model parameters are given in Table 1.

TABLE 1 **Monte Carlo Study: 500 Replications Using LISCOMP GLS for Binary Response Variables**

Parameter	True value	n = 1000	n = 250
μ_α	0.50	0.50 (0.05, 0.05)	0.50 (0.10, 0.09)
μ_β	−0.50	−0.51 (0.04, 0.04)	−0.50 (0.10, 0.09)
π_α	0.50	0.50 (0.05, 0.05)	0.50 (0.10, 0.10)
π_β	−0.50	−0.51 (0.04, 0.04)	−0.50 (0.09, 0.09)
$\gamma 1$	0.70	0.70 (0.06, 0.05)	0.70 (0.11, 0.11)
$\gamma 2$	0.70	0.72 (0.11, 0.11)	0.76 (0.40, 0.29)
$\gamma 3$	0.70	0.71 (0.09, 0.09)	0.71 (0.19, 0.18)
$\gamma 4$	0.70	0.71 (0.08, 0.08)	0.71 (0.18, 0.17)
$\psi\alpha\alpha$	0.50	0.49 (0.13, 0.13)	0.48 (0.29, 0.26)
$\psi\beta\alpha$	−0.10	−0.09 (0.06, 0.05)	−0.08 (0.13, 0.11)
$\psi\beta\beta$	0.10	0.10 (0.04, 0.03)	0.10 (0.09, 0.08)
$\psi k+1, k$	0.20	0.22 (0.09, 0.08)	0.26 (0.31, 0.22)
$\Delta 11$	1.00	0.99 (0.13, 0.13)	1.01 (0.29, 0.25)
$\Delta 22$	1.00	1.00 (0.11, 0.11)	1.02 (0.25, 0.22)
$\Delta 33$	1.00	1.00 (0.11, 0.11)	1.03 (0.25, 0.23)
χ^2 Average (df = 15)		14.75	15.51
SD		5.53	5.61
5% Reject proportion		5.2	5.4
1% Reject proportion		1.0	2.0

Note. In parentheses are empirical standard deviations, standard errors. *df*, degrees of freedom; *SD*, standard deviation.

Two sample sizes were used, a larger sample size of $n = 1000$ and a smaller sample of $n = 250$. Multivariate normal data were generated for y^* and x, and the y^* variables were dichotomized at zero. The generalized least-squares estimator was used. The parameter values chosen (see Table 1) imply that the proportion of $y = 1$ at the four time points is .64, .50, .34, and .25. The parameters estimated were μ_α, μ_β, π_α, π_β, γ_0, γ_1, γ_2, γ_3, $\psi_{\alpha\alpha}$, $\psi_{\beta\alpha}$, $\psi_{\beta\beta}$, $\psi_{\zeta_{k+1}\zeta_k}$ (a single parameter), $[\Delta]_{11}$, $[\Delta]_{22}$, and $[\Delta]_{33}$. The threshold parameter τ was fixed at zero and the scaling factor $[\Delta]_{11}$ was fixed at one. As discussed in section 3.3, the four residual variances of $\psi_{\zeta\zeta}$ are not free parameters to be estimated but they are still allowed to differ freely across time (their population values are .5, .6, .5, and .2) because the Δ parameters are free. The degrees of freedom for the chi-square model test is 15. A total of 500 replications were used for both sample sizes in Table 1. Table 1 gives the parameter estimates, the empirical standard deviation of the estimates across the 500 replications, the mean of the estimated standard errors for the 500 replications, and a summary of the chi-square test of model fit for the 500 replications.

As seen in Table 1, the estimates for the $n = 1000$ case show almost no parameter bias, the empirical variation is very close to the mean of the standard errors, and the chi-square test behaves correctly. As expected, the empirical standard deviation is cut in half when reducing the sample size to a quarter, from 1000 to 250. Exceptions to this, however, are the regression slope for the second time point γ_2 and the residual covariance $\psi_{\zeta_{k+1}\zeta_k}$. The cause for these anomalies needs further research. In these cases, the standard errors are also strongly underestimated. In the remaining cases, the standard errors agree rather well with the empirical variation, with perhaps a minor tendency to underestimate the standard errors for the (co)variance-related parameters of ψ and Δ. At $n = 250$, the variation in the regression intercept and slope parameter (μ and π) estimates is low enough for the hypotheses of zero values to be rejected at the 5% level. For the (co)variance-related parameters of ψ, however, this is not the case and the Δ parameters also have relatively large variation. The chi-square test behavior at $n = 250$ is quite good.

4.2. Analysis of Example 2 Data

The model used for the preceding simulation study will now be applied to the Example 2 data of neurotic illness as described previously (for more details, see Henderson et al., 1981). Each of the four response variables will be modeled separately. They can also be analyzed together as multiple indicators of neurotic illness, but this will not be done here. Previous longitudinal analyses of these data were done in Muthén (1983, 1991).

Summaries of the data are given in Table 2. As is shown in Table 2, there is a certain drop from the first to the remaining occasions in the proportion of people answering yes to the neuroticism items. There is also a corresponding drop in the mean of the life event score. Because the latter is used as a time-varying covariate, this means that the data could be fit by a model that does not include a factor for a decline in the response variable, but which instead uses only a random intercept factor model. Given previous analysis results, gender is dropped as a time-invariant covariate. Only the N score, the long-term susceptibility to neurosis, will be used to predict the variation in the random intercept factor.

TABLE 2 Descriptive Statistics for Example 2 Data

Response variables	Time 1	Time 2	Time 3	Time 4
Percentage yes:				
Anxiety	26.4	16.5	15.6	16.0
Depression	25.5	14.7	17.7	13.9
Irritability	40.3	31.2	28.1	29.0
Nervousness	24.2	19.0	16.0	15.6

Covariates	Means	Variances
N	9.31	20.66
$L1$	3.86	6.54
$L2$	3.17	5.89
$L3$	2.58	4.90
$L4$	2.42	5.27

Correlations

N	1.00				
$L1$	0.22	1.00			
$L2$	0.16	0.54	1.00		
$L3$	0.18	0.50	0.49	1.00	
$L4$	0.21	0.53	0.49	0.51	1.00

Two types of models will be fit to the data. First, the general binary growth model will be fit, allowing for across-time variation in the latent response variable residual variance and nonzero covariances between pairs of residuals (the residual covariances are restricted to being equal). Second, the conventional binary growth model, in which these features are not allowed for, will be fit as a comparison. In both cases, the same quantities as in Table 1 will be studied, along with two types of summary statistics. One summary statistic is R^2_α, that is, the proportion of variation is the α factor accounted for by N. A second statistic is the proportion that the α factor variation makes up of the total variation in the latent response variable y^*, calculated at all time points. The α factor represents individual variation in a neurotic illness trait, variation that is present at all time points. In addition to this variation, the y^* variation is also influenced by time-specific variation caused by time-varying, measured covariates (Ls) and time-specific unmeasured residuals (ζs). In this way, the proportion is a time-varying measure of how predominant the trait variation is in the responses.

Table 3 shows the results for the general binary growth model. The model fits each of the four response variables very well. As expected, the N score has a significantly positive influence ($\hat{\pi}_\alpha$) on the random intercept factor and the L scores have significantly positive influences ($\hat{\gamma}s$) on the probability of yes answers for the response variables.

For none of these four response variables is the residual covariance significantly different from zero. Note, however, from the simulation study at $n = 250$ that the variation in this estimate is quite large and that a large sample size is required to reject zero covariance. As shown in the simulation study, the point estimate of the covariance may be of reasonable magnitude. The model with

TABLE 3 **Analysis of Example 2 Data Using the General Growth Model**

	Anxiety	Depression	Irritability	Nervousness
$\mu\alpha$	−1.46 (0.21)	−2.56 (0.34)	−1.21 (0.20)	−2.43 (0.27)
$\pi\alpha$	0.06 (0.01)	0.13 (0.02)	0.06 (0.01)	0.14 (0.02)
$\gamma 1$	0.08 (0.03)	0.14 (0.04)	0.09 (0.03)	0.08 (0.03)
$\gamma 2$	0.05 (0.02)	0.02 (0.03)	0.08 (0.02)	0.03 (0.02)
$\gamma 3$	0.06 (0.03)	0.12 (0.03)	0.07 (0.02)	0.05 (0.03)
$\gamma 4$	0.03 (0.02)	0.04 (0.05)	0.06 (0.02)	0.01 (0.03)
$\psi\alpha\alpha$	0.31 (0.08)	0.39 (0.10)	0.27 (0.07)	0.69 (0.09)
$\psi k, k+1$	−0.01 (0.05)	−0.12 (0.10)	−0.04 (0.04)	−0.11 (0.06)
$\Delta 11$	1.37 (0.19)	0.88 (0.15)	1.41 (0.22)	1.01 (0.11)
$\Delta 22$	1.48 (0.21)	1.08 (0.19)	1.43 (0.24)	1.14 (0.11)
$\Delta 33$	1.20 (0.18)	0.88 (0.14)	1.32 (0.22)	1.00 (0.11)
$\chi^2(19)$	16.49	21.73	26.02	23.30
p-value	.624	.298	.130	.225
$R^2\alpha$	0.19	0.47	0.22	0.37
$P1$	0.34	0.50	0.30	0.75
$P2$	0.31	0.41	0.28	0.52
$P3$	0.31	0.42	0.29	0.54
$P4$	0.30	0.41	0.28	0.52

nonzero residual covariance is therefore maintained. In this particular application, the estimate is small.

The scaling factors of Δ are not significantly different from unity at the 5% level for all but one of the cases. Because in this model there is no β factor, Δ is a function of the α factor residual variance $\psi_{\alpha\alpha}$ and the residual variance $\psi_{\zeta\zeta}$ (cf. Eqs. 18–21). Unit values for the Δ scaling factors would therefore indicate that the residual variances are constant over time in this application. Note, however, from the Table 1 simulation results that the sampling variation in the Δ estimates is quite large at $n = 250$ which makes it difficult to reject equality of residual variances over time. The Table 1 results also indicate that the point estimates for Δ are good.

Table 4 shows the results for the conventional binary growth model. This model cannot be rejected at the 5% level in these applications. The parameter estimates are, in most cases, similar to those for the generalized model of Table 3. Differences do, however, show up in the values for the trait variance proportions, labeled P_1 through P_4 in Tables 3 and 4. Relative to the more general model, the conventional model overestimates these proportions for three out of the four response variables. For example, the conventional model indicates that there is a considerable dominance of trait variation in the response variable Nervousness, with a proportion of .81 for the last three time points (see Table 4).

TABLE 4 **Analysis of Example 2 Data Using the Conventional Growth Model**

	Anxiety	Depression	Irritability	Nervousness
$\mu\alpha$	−1.79 (0.17)	−2.40 (0.16)	−1.56 (0.16)	−2.56 (0.23)
$\pi\alpha$	0.07 (0.01)	0.12 (0.01)	0.08 (0.01)	0.15 (0.02)
$\gamma 1$	0.12 (0.02)	0.13 (0.02)	0.12 (0.02)	0.09 (0.03)
$\gamma 2$	0.06 (0.03)	0.03 (0.03)	0.10 (0.02)	0.03 (0.02)
$\gamma 3$	0.05 (0.04)	0.10 (0.03)	0.09 (0.03)	0.03 (0.03)
$\gamma 4$	0.04 (0.03)	0.05 (0.03)	0.07 (0.03)	0.01 (0.03)
$\psi\alpha\alpha$	0.50 (0.51)	0.31 (0.06)	0.42 (0.06)	0.72 (0.05)
$\chi^2(23)$	26.12	26.29	33.97	29.02
p-value	.295	.287	.066	.180
$R^2\alpha$	0.17	0.49	0.24	0.39
$P1$	0.50	0.43	0.45	0.78
$P2$	0.54	0.47	0.46	0.81
$P3$	0.54	0.45	0.47	0.81
$P4$	0.54	0.46	0.48	0.81

The more general model of Table 3 points to a much lower range of values for the last three time points, .52 to .54.

5. CONCLUSIONS

This chapter has discussed a general framework for longitudinal analysis with binary response variables. As compared with conventional random effects modeling with binary response, this general approach allows for residuals that are correlated over time and variances that vary over time. It also allows for multiple indicators of latent variable constructs, in which case it is possible to identify separately residual variation and measurement error variation. The more general model can be estimated by a limited-information generalized least-squares estimator. The general approach fits into an existing latent variable modeling framework for which software has been developed.

A Monte Carlo study showed that the limited-information generalized least-squares estimator performed well with sample sizes at least as low as $n = 250$. At this sample size, the sampling variability is not unduly large for the regression parameters of the model, but it is rather high for the (co)variance-related parameters of the model. Analyses of a real data set indicated that the differences in key estimates obtained by the conventional model are not always markedly different from those obtained by the more general model, but can lead to quite different conclusions about certain aspects of the phenomenon that is being modeled.

The general approach should be of value for developmental studies in which variables are often binary and in which the variables are often very skewed and essentially binary. The general model allows for a flexible analysis which has so far been used very little with binary responses. A multiple-cohort analysis of this type is carried out in Muthén and Muthén (1995), which describes the development of heavy drinking over age for young adults.

ACKNOWLEDGMENTS

This research was supported by Grant AA 08651-01 from NIAAA for the project "Psychometric Advances for Alcohol and Depression Studies" and Grant 40859 from the National Institute of Mental Health.

REFERENCES

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika, 46,* 443–459.

Bock, R. D., & Lieberman (1970). Fitting a response model for n dichotomously scored items. *Psychometrika, 35,* 179–197.

Diggle, P. J., Liang, K. Y., & Zeger, S. (1994). *Analysis of longitudinal data.* Oxford: Oxford University Press.

Fitzmaurice, G. M., Laird, N. M., & Rotnitzky, A. G. (1993). Regression models for discrete longitudinal data. *Statistical Science, 8,* 284–309.

Gibbons, R. D., & Bock, R. D. (1987). Trend in correlated proportions. *Psychometrika, 52,* 113–124.

Gibbons, R. D., & Hedeker, D. R. (1993). *Application of random-effects probit regression models* (Technical report). Chicago: University of Illinois at Chicago, UIC Biometric Laboratory.

Henderson, A. S., Byrne, D. G., & Duncan-Jones, P. (1981). *Neurosis and the social environment.* Sydney: Academic Press.

Jöreskog, K. G., & Sörbom, D. (1977). Statistical models and methods for analysis of longitudinal data. In D. J. Aigner & A. S. Goldberger (Eds.), *Latent variables in socio-economic models* (pp. 285–325). Amsterdam: North-Holland.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics, 38,* 963–974.

Longford, N. T. (1993). *Random coefficient models.* Oxford: Oxford University Press.

McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development, 58,* 110–133.

Meredith, W., & Tisak, J. (1984). *"Tuckerizing" curves.* Paper presented at the annual meetings of the Psychometric Society, Santa Barbara, CA.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55,* 107–122.

Muthén, B. (1983). Latent variable structural equation modeling with categorical data. *Journal of Econometrics, 22,* 43–65.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika, 49,* 115–132.

Muthén, B. (1987). *LISCOMP. Analysis of linear structural equations with a comprehensive measurement model. Theoretical integration and user's guide.* Mooresville, IN: Scientific Software.

Muthén, B. (1991). Analysis of longitudinal data using latent variable models with varying parameters. In L. Collins & J. Horn (Eds.), *Best methods for the analysis of change. Recent advances, unanswered questions, future directions* (pp. 1–17). Washington, DC: American Psychological Association.

Muthén, B. (1993). Latent variable modeling of growth with missing data and multilevel data. In C. R. Rao & C. M. Cuadras (Eds.), *Multivariate analysis: Future directions 2* (pp. 199–210). Amsterdam: North-Holland.

Muthén, B., & Christofferson, A. (1981). Simultaneous factor analysis of dichotomous variables in several groups. *Psychometrika, 46,* 407–419.

Muthén, B., Khoo, S. T., & Nelson Goff, G. (1994). *Longitudinal studies of achievement growth using latent variable modeling* (Technical report). Los Angeles: UCLA, Graduate School of Education.

Muthén, B., & Liu, G. (1994). *Identification issues related to binary growth modeling (Technical report).* Manuscript in preparation.

Muthén, B., & Muthén, L. (1995). *Longitudinal modeling of non-normal data with latent variable techniques: Applications to developmental curves of heavy drinking among young adults* (Technical report). Manuscript in preparation.

Muthén, B., & Satorra, A. (1995). Technical aspects of Muthén's LISCOMP approach to estimation of latent variable relations with a comprehensive measurement model. *Psychometrika.*

Rutter, C. M., & Elashoff, R. M. (1994). Analysis of longitudinal data: Random coefficient regression modeling. *Statistics in Medicine, 13,* 1211–1231.

Stiratelli, R., Laird, N., & Ware, J. H. (1984). Random-effects models for serial observations with binary response. *Biometrics, 40,* 961–971.

Wheaton, B., Muthén, B., Alwin, D., & Summers, G. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Ed.), *Sociological methodology 1977* (pp. 84–136). San Francisco: Jossey-Bass.

Willett, J. B., & Sayer, A. G. (1993). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin.*

Zeger, S. L., & Karim, M. R. (1991). Generalized linear models with random effects; a Gibbs sampling approach. *Journal of the American Statistical Association, 86,* 79–86.