This article has been accepted for publication in *Structural Equation Modeling: A Multidisciplinary Journal* Published by Taylor & Francis

On the Estimation of Nonlinear Mixed-Effects Models and Latent Curve Models for Longitudinal Data

Shelley A. Blozis

University of California, Davis

Jeffrey R. Harring

University of Maryland

Address correspondence to: Shelley A. Blozis, Psychology Department, University of California, Davis, One Shields Ave, Davis, CA 95616. 530-754-9457. email: sablozis@ucdavis.edu.

Abstract

Nonlinear models are effective tools for the analysis of longitudinal data. These models provide a flexible means for describing data that follow complex forms of change. Exponential and logistic functions that include a parameter to represent an asymptote, for instance, are useful for describing responses that tend to level off with time. There are forms of nonlinear latent curve models and nonlinear mixed-effects model that are equivalent, and so given the same set of data, growth function, distributional assumptions, and method of estimation, the two models yield equivalent results. There are also forms that are strikingly different and can yield different interpretations for a given set of data. This paper discusses cases in which nonlinear mixed-effects models and nonlinear latent curve models are equivalent and those in which they are different and clarifies the estimation needs of the different models. Examples based on empirical data help to illustrate these points.

Keywords: longitudinal data; nonlinear mixed–effects models, nonlinear latent curve models; structured latent curve models

On the Estimation of Nonlinear Mixed-Effects Models and Latent Curve Models for Longitudinal Data

Longitudinal models are essential for the understanding of psychological and behavioral measures that are studied over time. Two major statistical frameworks for longitudinal data analysis in the behavioral sciences are latent curve models and mixed-effects models. Both have been developed to handle response data measured using different scales of measurement, including categorical data, and both can define growth according to many different functional forms. Nonlinear latent curve models and nonlinear mixed-effects models in particular provide a means for describing variables that follow complex forms of change, such as multiphase change (Cudeck & Klebe, 2002; Kohli, Hughes, Wang, Zopluoglu, & Davison, 2015; Ram & Grimm, 2007). Importantly, both frameworks can make use of growth functions that are straightforward in their interpretation and give a parsimonious way of representing complex forms of change (Cudeck, 1996; Fitzmaurice, Laird, & Ware, 2004). To clarify, nonlinear latent curve models and nonlinear mixed-effects models are models for which the growth function that is used to describe change in the response includes one or more coefficients that enter the function in a nonlinear, or non-additive, way (see the Appendix for a discussion on the distinction between linear and nonlinear parameters of a growth model). Examples of nonlinear latent curve models are given in Browne (1993), Browne and Du Toit (1991), Grimm, Ram, and Hamagami (2011), Laursen, Little, and Card, (2012), and Meredith and Tisak (1990); examples of nonlinear mixed-effects models using nonlinear growth functions are given in Pinheiro and Bates (2000) and Davidian and Giltinan (1995, 2003).

Nonlinear mixed-effects models have been given a great deal of attention in the statistical and methodological literature in the past few decades, and many articles and books are devoted to theoretical developments and estimation approaches (Davidian & Gallant, 1993; Davidian & Giltinan, 1995, 2003; Lindstrom & Bates, 1990; Littell, Stroup, Milliken, Wolfinger, & Schabenberger, 2006; Pinheiro & Bates, 1995, 2000; Vonesh, 1996; Vonesh & Chinchilli, 1996). Relatively less attention has been given to nonlinear latent curve models, although early theoretical work on these models occurred within approximately the same time frame (Browne, 1993; Browne & Du Toit, 1991; Meredith & Tisak, 1990). These two nonlinear modeling approaches for the analysis of longitudinal data can differ in important ways, and consequently, can differ in their needs for estimation. Specifically, there is a version of a nonlinear mixed-effects model (i.e., a partially nonlinear mixed-effects model [Davidian & Giltinan, 1995]; a.k.a. a conditionally linear mixed-effects model [Blozis & Cudeck, 1999]) and a version of the nonlinear latent curve model that are equivalent, and thus it can be shown that the estimation requirements are equivalent. There is also a version of a nonlinear latent curve model, namely the structured latent curve model, that differs appreciably from a nonlinear mixed-effects model, and thus the estimation requirements are different. A structured latent curve model in particular offers a unique way of describing longitudinal data, and for some types of longitudinal data, can be preferred over a nonlinear-mixed effects model (Blozis & Harring, 2015). The purpose of this paper is to highlight the similarities and differences between these two important longitudinal frameworks and to clarify their estimation requirements.

The remainder of this paper is as follows: First, nonlinear mixed-effects models and the nonlinear latent curve model that was introduced by Meredith and Tisak (1990) are reviewed. It is then shown how a particular restriction to a nonlinear mixed-effects model results in a model that is equivalent to the nonlinear latent curve model. Specifically, a partially nonlinear mixed-effects model is shown to be equivalent to Meredith and Tisak's nonlinear latent curve model. Next, the structured latent curve model, a special form of a nonlinear latent curve model introduced in Browne (1993) and Browne and Du Toit (1991), is described and contrasted with these models. Estimation requirements of the different models are discussed. Due to its relevance in this discussion, a population-average models is briefly contrasted with the models highlighted here because its distinction from these models helps further understanding of the different options in nonlinear longitudinal models. To aid with the discussions an empirical data set is presented. Recommendations are given for fitting these models with illustrations using SAS statistical software given its flexibility in estimating these different longitudinal models.

Nonlinear Mixed-effects Models and Nonlinear Latent Curve Models

Nonlinear mixed-effects models. Nonlinear mixed-effects models are subject-specific models in which the individual is the focus of study, and so a model for growth is defined at the individual level with one or more of the growth coefficients being specific to the individual (Davidian & Giltinan, 1995, 2003). Many resources provide examples of nonlinear mixed-effects models that are based on a variety of growth functions (Cudeck, 1996; Lindstrom & Bates, 1990; Pinheiro & Bates, 2000). The responses of all individuals in the population are assumed to follow the same growth form, such as assuming that the responses of all individuals follow the same exponential growth function, but the coefficients that describe change in the response can be specific to each individual. Assuming, for instance, the responses of an individual follow a negatively accelerated exponential function over time, a nonlinear mixed-effects model for y_{ti} can be given by

$$y_{ti} = \beta_{1i} - (\beta_{1i} - \beta_{0i}) \exp\{-\beta_{2i} Time_{ti}\} + e_{ti} \tag{1}$$

where y_{ti} is the response at $Time_{ti}$ for individual i, β_{0i} is an individual's expected response at $Time_{ti} = 0$, β_{1i} is an individual's potential response, and β_{2i} combined with values of $Time$ dictates the nonconstant rate of change across time. In the example presented in (1) each of the coefficients is assumed to be a combination of a fixed effect and a random effect, although it is not a requirement of the model that every coefficient has both a fixed effect and corresponding random effect. In (1), for instance, $\beta_{0i} = \beta_0 + b_{0i}$, where β_0 is the fixed intercept common to all individuals, and b_{0i} is the corresponding random effect specific to individual i. Finally, the individual and time-specific residual for the model is given by e_{ti}. This subject-specific model allows the different features that describe change in a response to be unique to the individual, making it possible for the responses of individuals to change at different rates and to have different response levels, both at $Time_{ti} = 0$ and with regard to the potential level. The fixed coefficients of the model represent the typical characteristics of change across the population. For example, the fixed rate of change β_2 represents the average of the change rates across individuals. Thus, one can think of the fixed effects of a nonlinear mixed-effects model as representing the typical parameter values (Cudeck, 1996; Davidian & Giltinan, 1995, 2003).

Within individuals, the set of residuals $\mathbf{e}_i = (e_{1i}, \ldots, e_{n_i})'$, where n_i is the number of observations for person i, is assumed to be multivariate normal with expected value equal to zero and covariance matrix $\mathbf{\Theta}_i$:

$$\mathbf{e}_i \sim N[\mathbf{0}, \mathbf{\Theta}_i]$$

where $\mathbf{\Theta}_i$ is a symmetric matrix of order n_i so that $\mathbf{\Theta}_i$ can vary by individual. For instance, in situations in which individuals are measured at different points in time or possibly a different number of times, allowing $\mathbf{\Theta}_i$ to vary by individual can allow for such differences. With regard to the particular structures assumed for $\mathbf{\Theta}_i$, perhaps the most common assumption is that the residuals are independent with constant variance across time so that $\mathbf{\Theta}_i = \sigma^2\mathbf{I}_{n_i}$, where σ^2 is the common variance, and $\mathbf{I}_{n_i}$ is an identity matrix of order n_i. If the growth model adequately accounts for the dependencies in the data, then the residuals would be expected to be independent across time. Further, if the random coefficients of the growth model account for individual differences in the responses, then the residuals might also be expected to have constant variance across time. In practice, it is recommended that alternative covariance structures be considered for a given problem to evaluate assumptions about the residuals. For instance, even after accounting for change in the responses by fitting a growth model, some dependencies in the residuals may remain and a covariance structure that assumes correlations between the residuals may be more appropriate. In any case, considering different covariance structures for the residuals can be useful in evaluating and interpreting a nonlinear mixed-effects model (Harring & Blozis, 2014).

At the population level the nonlinear mixed-effects model is given by (cf: Davidian & Giltinan, 1995)

$$\boldsymbol{\beta}_i = g(\boldsymbol{\beta}, \mathbf{b}_i) \tag{2}$$

where g is a function that depends on a set of p fixed effects $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$ that are common to all individuals and a set of k individual-specific random effects $\mathbf{b}_i = (b_{1i}, \ldots, b_{ki})'$. The random effects $\mathbf{b}_i$ are assumed to be multivariate normal with expected value equal to zero and covariance matrix $\mathbf{\Phi}$:

$$\mathbf{b}_i \sim N[\mathbf{0}, \mathbf{\Phi}]$$

where $\mathbf{\Phi}$ is a symmetric matrix of order k.

In its simplest form, the function g in (2) denotes the sum of fixed and random effects, and the expected value of $\boldsymbol{\beta}_i$ is equal to the fixed effect $\boldsymbol{\beta}$: $E[\boldsymbol{\beta}_i = \boldsymbol{\beta}]$. In these cases, the fixed effects represent the averages of the individual-level coefficients of the growth model, and so the growth function defined by the fixed effects yields the typical curve (Davidian & Giltinan, 2003). It is important to note that the typical curve is not the same as the curve of the average response. Indeed, the nonlinear mixed-effects model does not specify a model for the average response. Instead, the population-level model is a model for the individual-level coefficients as in (2). The population-level model can also be expanded to include person-level attributes to account for variation in the random effects, with possibly a different set of predictors used to address variation in each of the random effects. In these latter situations, the random effects are conditional on person-level attributes.

Nonlinear latent curve models. Meredith and Tisak (1990) develop the framework for a class of models known as latent curve models. For the sake of the discussion here and without loss in generality, we assume that all individuals in a sample are observed according to the same points in time: $Time_1, \ldots, Time_n$ where n is the number of observations for all individuals. Under a latent curve model, an individual response is assumed to be decomposable into a sum of basis functions and error:

$$\mathbf{y}_i = \mathbf{\Lambda}\boldsymbol{\eta}_i + \mathbf{e}_i \tag{3}$$

where $\mathbf{\Lambda}$ is a matrix of the basis functions assumes to be common across individuals of a population, $\boldsymbol{\eta}_i$ is a set of individual-specific weights, and $\mathbf{e}_i$ is the set of errors. Among different versions of the latent curve model that were presented in their article, one version was based on a growth function that included a nonlinear parameter. Specifically, an individual-level model was given that was based on a negatively accelerated exponential function:

$$y_{ti} = \eta_{1i} - (\eta_{1i} - \eta_{0i})\exp\{-\gamma Time_t\} + e_{ti} \tag{4}$$

where η_{0i} is an individual's expected response at $Time_t = 0$, and η_{1i} is an individual's potential response (Meredith & Tisak, 1990, p 117). The nonlinear coefficient γ was assumed to be fixed across the population. So that the model in (4) follows the form of the latent curve model in (3), a matrix of basis functions $\boldsymbol{\Lambda}$ and the random weight vector are defined as (cf: Browne, 1993)

$$\boldsymbol{\Lambda} = \begin{bmatrix} \exp\{-\gamma Time_1\} & 1 - \exp\{-\gamma Time_1\} \\ \vdots & \vdots \\ \exp\{-\gamma Time_n\} & 1 - \exp\{-\gamma Time_n\} \end{bmatrix} \tag{5}$$

and

$$\boldsymbol{\eta}_i = \begin{bmatrix} \eta_{0i} \\ \eta_{1i} \end{bmatrix},$$

respectively, where $\boldsymbol{\Lambda}$ is evaluated according to $Time_1, \ldots, Time_n$. Thus, the individual-level model in (4) can be equivalently expressed as a sum of basis functions weighted by individual-specific coefficients:

$$y_{ti} = \exp\{-\gamma Time_t\}\eta_{0i} + (1 - \exp\{-\gamma Time_t\})\eta_{1i} + e_{ti} \tag{6}$$

In the nonlinear latent curve model in (6), the two coefficients η_{0i} and η_{1i} enter the model in a linear, or additive, way, and both are assumed to be a sum of a fixed and a random effect. That is, let $\eta_{0i} = \alpha_0 + z_{0i}$ and $\eta_{1i} = \alpha_1 + z_{1i}$, where α_0 and α_1 are the fixed intercept and asymptote, respectively, and z_{0i} and z_{1i} are the corresponding random effects. In contrast to η_{0i} and η_{1i}, the rate parameter γ is fixed across the population and enters each of the basis functions in a nonlinear way. Thus, in this model, only the linear parameters are random and the nonlinear parameter is fixed. Finally, the time-specific residual is given by e_{ti}. Within individuals, the residuals in $\mathbf{e}_i = (e_{1i}, \ldots, e_{n_i})'$ are assumed to be multivariate normal with expected value equal to zero and covariance matrix $\mathbf{\Theta}$:

$$\mathbf{e}_i \sim N[\mathbf{0}, \mathbf{\Theta}]$$

where $\mathbf{\Theta}$ is a symmetric matrix of order n. Similar to the nonlinear mixed-effects model, different structures may be considered for $\mathbf{\Theta}$.

The population level model describes the expected value of the response evaluated according to time and is assumed to have the same functional form as that assumed for the individual-level model. For instance, assuming the exponential growth function in (4) that was assumed at the individual level, the expected value of $\mathbf{y}_i$ under the model is

$$E[\mathbf{y}_i] = \boldsymbol{\mu} = \mathbf{\Lambda}\boldsymbol{\alpha} \tag{7}$$

where $\mathbf{\Lambda}$ is defined in (5) and $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)'$. In the population level model, all coefficients are fixed and describe the mean response across time. That is, α_0 is the intercept of the mean response, α_1 is the potential mean response, and γ governs the rate of change in the mean response. Indeed, the model for the mean response ("curve of the averages") is equal to the average of the subject-specific models ("average of the curves"), a quality of a model that Keats (1983) (also see Singer & Willet, 2003) refers to as dynamic consistency. As in the individual model of (4), the intercept and asymptote parameters of the mean function enter linearly, and the rate parameter enters nonlinearly. Also at the population level, the set of random weights $\mathbf{z}_i$ is assumed to be multivariate normal with expected value equal to zero and covariance matrix $\mathbf{\Phi}$:

$$\mathbf{z}_i \sim N[\mathbf{0}, \mathbf{\Phi}]$$

where $\mathbf{\Phi}$ is a symmetric matrix of order k, and k is the number of random weights. Given the distributional assumptions for $\mathbf{e}_i$ and $\mathbf{z}_i$ and assuming independence between them, the mean structure is defined in (7) and the covariance structure is

$$\mathbf{\Sigma} = \mathbf{\Lambda\Phi\Lambda}' + \mathbf{\Theta},$$

where the orders of both $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$ are n.

A first-order Taylor expansion. Technically the latent curve model of Meredith and Tisak (1990) is based on a first-order Taylor expansion defined as a linear combination of the first-order partial derivatives of a target function taken with respect to the coefficients of the function but for which the expansion is done only with respect to the coefficients that enter the function in a linear way (see Blozis & Harring, 2015). For the exponential function in (4), the first-order partial derivatives of the function evaluated at $Time_t$ are

$$\frac{\partial}{\partial \eta_{0i}}(\eta_{1i} - (\eta_{1i} - \eta_{0i})\exp\{-\gamma Time_t\}) = \exp\{-\gamma Time_t\}$$

and

$$\frac{\partial}{\partial \eta_{1i}}(\eta_{1i} - (\eta_{1i} - \eta_{0i})\exp\{-\gamma Time_t\}) = 1 - \exp\{-\gamma Time_t\}.$$

As noted earlier, these derivatives defined the basis functions of the matrix $\mathbf{\Lambda}$ in (5). The basis functions are then weighted by the set of random weights $\boldsymbol{\eta}_i$ in (6). From this it is clear that the latent curve model defines the response as a decomposable sum of a set of basis functions and error.

Model interpretation. The fixed coefficients in $\mathbf{\Lambda}$ and $\boldsymbol{\alpha}$, namely α_0 and α_1, represent the typical characteristics of change across the population for the two parameters that can vary across individuals. Conversely, the rate parameter γ is fixed across the population, so the model does not allow for the rate parameter to differ across individuals. Similar to the nonlinear mixed-effects model in (1), one can think of the fixed effects of the nonlinear latent curve model as representing the typical function parameter values. It is also true, however, that the fixed effects of a nonlinear latent curve model can be interpreted as the features that describe change in the mean longitudinal response, that is, the response averaged at each measurement occasion. Indeed, the individual-level model and the population-level model of a nonlinear latent curve model assume that both the mean response and the individual responses follow the same functional form. This is not necessarily true of a nonlinear mixed-effects model in which all of the growth coefficients, including those that enter a function in a nonlinear way, vary across individuals. Under a nonlinear mixed-effects model most generally, there is no requirement that the averaged response follow the same functional form as that assumed for the individuals. In fact, a nonlinear mixed-effects model does not include a direct specification of the averaged response (Davidian & Giltinan, 1995, 2003). That is, the population-level model of a nonlinear mixed-effects model concerns the coefficients that define the individual-level model, as given in (2). Similar to nonlinear mixed-effects models, a variety of growth functions can be used to define a latent growth model of (3). For the latent curve model, however, there is the restriction that any growth coefficient that goes into the function in a nonlinear manner be fixed, as described earlier.

Equivalent Model Forms

A particular form of the nonlinear mixed-effects model is one in which the random effects only enter the function in a linear way and fixed effects enter in a linear or nonlinear way. This is known as a partially nonlinear mixed-effects model (Davidian & Giltinan, 1995) (a.k.a. a conditionally linear mixed-effects model [Blozis & Cudeck, 1999]) that differs from a fully nonlinear mixed-effects model that allows for any parameter of a growth function, linear or nonlinear, to vary across individuals. In a partially nonlinear mixed-effects model, the interpretation of the growth coefficients can be done in the same fashion as the nonlinear latent curve model of Meredith and Tisak (1990). That is, the growth coefficients of a partially nonlinear mixed-effects model can be interpreted as the typical coefficients across individuals, as well as the coefficients that describe the average response. To illustrate this, assume that a partially nonlinear mixed-effects model is used to describe longitudinal data using the exponential growth function that was used in (1) but where the rate parameter is fixed across individuals:

$$y_{ti} = \beta_{1i} - (\beta_{1i} - \beta_{0i}) \exp\{-\beta_2 Time_{ti}\} + e_{ti} \tag{8}$$

Letting $\eta_{0i} = \beta_{0i}$, $\eta_{1i} = \beta_{1i}$ and $\gamma = \beta_2$, the individual-level model of the nonlinear latent curve model in (4) is expressed in the same manner as the individual-level model of the partially nonlinear mixed-effects model in (8). As in (4), the individual-level coefficients of the model in (8) are sums of a fixed effect and a random effect: $\beta_{0i} = \beta_0 + b_{0i}$ and $\beta_{1i} = \beta_1 + b_{1i}$, where the expected values of the coefficients are equal to the fixed effects: $E[\beta_{0i}] = \beta_0$ and $E[\beta_{1i}] = \beta_1$. Thus, the intercept and asymptote vary across individuals, and the fixed effects β_0 and β_1 represent the typical growth coefficients. The population-level model is a model for the expected value of the response, analogous to the nonlinear latent curve model:

$$E[y_{ti}] = \mu_t = \beta_1 - (\beta_1 - \beta_0)\exp\{-\beta_2 Time_t\} \tag{9}$$

Letting $\alpha_0 = \beta_0$, $\alpha_1 = \beta_1$, and $\gamma = \beta_2$, the population-level models of the partially nonlinear mixed-effects model in (9) and the nonlinear latent curve model in (4) are equivalent expressions. As for the fixed effects of the nonlinear latent curve model, the fixed effects of the partially nonlinear mixed-effects model can also be interpreted as the coefficients that describe the mean response.

It follows that a partially nonlinear mixed-effects model can also be expressed using a first-order Taylor expansion. At the individual level, let

$$\mathbf{y}_i = \mathbf{\Lambda}\boldsymbol{\beta}_i^* + \mathbf{e}_i$$

where $\mathbf{\Lambda}$ is defined as it was in (5) and $\boldsymbol{\beta}_i^* = (\beta_{0i}, \beta_{1i})'$. At the population level,

$$E[\mathbf{y}_i] = \boldsymbol{\mu} = \mathbf{\Lambda}\boldsymbol{\beta}^* \tag{10}$$

With $\boldsymbol{\beta}^*$ in (10) being equal to $\boldsymbol{\alpha}$ in (7), the mean structure of a partially nonlinear mixed-effects model is equivalent to that of a nonlinear latent curve model. The covariance structure for both models is

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Theta} \tag{11}$$

with the dimensions of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ equal to n.

It is important to note that the shared interpretation of the fixed coefficients of a nonlinear latent curve model and a partially nonlinear mixed-effects model is also shared with those of linear latent curve models and linear mixed-effects models (Davidian & Giltinan, 1995, 2003). That is, if the random coefficients of an individual-level model can only enter the model in a linear way, even if fixed nonlinear coefficients are included in the growth model, then the model’s fixed coefficients have this dual interpretation in describing the typical coefficients and the coefficients of the mean response. Thus, these models possess the property of dynamic consistency.

Estimation

The nonlinear latent curve model and the partially nonlinear mixed-effects model are linear with respect to the random coefficients. Nonlinear coefficients are fixed. A consequence of this is that a closed-form analytic expression for the marginal mean is available, similar to linear mixed-effects models and linear latent curve models (Davidian & Giltinan, 1995). First, let $\boldsymbol{\kappa}$ represent the set of parameters of either a partially nonlinear mixed-effects model or a nonlinear latent curve model. Specifically, $\boldsymbol{\kappa}$ includes the parameters of the mean structure and those of the covariance structure. The mean and covariances structures under either model can then be defined as functions of $\boldsymbol{\kappa}$: $\boldsymbol{\mu}_i(\boldsymbol{\kappa})$ and $\boldsymbol{\Sigma}_i(\boldsymbol{\kappa})$. Assuming more generally now that the number of observations and the timing of observations could vary across individuals, the marginal density of a normally distributed response under a partially nonlinear mixed-effects model and a nonlinear latent curve model can be written equivalently as

$$\mathbf{y}_i \sim N[\boldsymbol{\mu}_i(\boldsymbol{\kappa}), \boldsymbol{\Sigma}_i(\boldsymbol{\kappa})]$$

A loglikelihood function can then be written as

$$\ln L(\boldsymbol{\kappa}; \mathbf{y}_1, \ldots, \mathbf{y}_N) \propto C - \frac{1}{2}\sum_{i=1}^{N}\Big[\ln|\boldsymbol{\Sigma}_i(\boldsymbol{\kappa})| + [\mathbf{y}_i - \boldsymbol{\mu}_i(\boldsymbol{\kappa})]'\boldsymbol{\Sigma}_i(\boldsymbol{\kappa})^{-1}[\mathbf{y}_i - \boldsymbol{\mu}_i(\boldsymbol{\kappa})]\Big] \quad (12)$$

where C is a constant that is independent of $\boldsymbol{\kappa}$. This form of a loglikelihood function can be evaluated using techniques used to fit linear mixed-effects models and linear latent curve models (Jennrich & Schluchter, 1986).

Maximum likelihood (ML) estimates of a nonlinear latent curve model and a partially nonlinear mixed-effects model may be obtained using a software program for nonlinear mixed-effects models, such as SAS PROC NLMIXED. Using NLMIXED, one approach to ML estimation is to apply a first-order Taylor series method of estimation by using the 'METHOD=FIRO' option. The theory of this linearization method is described in Beal and Sheiner (1982). A more general use of a first-order Taylor series expansion is to approximate a nonlinear function. This approach has been used to approximate a nonlinear function under a fully nonlinear mixed-effects models, that is, a nonlinear mixed-effects model in which a random effect enters the function in a nonlinear manner (as was given in (1)). If, however, a function is strictly linear in the random effects (as in a partially nonlinear mixed-effects model or a nonlinear latent curve model), the result of the first-order linearization is a direct decomposition of the nonlinear function defined as a product of basis functions (possibly defined by fixed nonlinear coefficients) and a set of weights. Thus, although a first-order Taylor expansion can be used to approximate a nonlinear function, if a growth model is strictly linear in the random effects, the first-order Taylor expansion is equal to the original function, and thus, is not an approximation. Another approach to obtaining ML estimates is to use a numeric approximation method. The default estimation method for PROC NLMIXED is an adaptive Gaussian quadrature method that is describe in Pinheiro and Bates (1995). By using the procedure's default method of estimation, an approximate marginal likelihood function is evaluated numerically. Thus, it is important to note that these two methods of estimation, the first-order linearization method and the numeric approximation method, will provide equivalent results if a model is strictly linear in the random effects (Davidian & Giltinan, 1995). In other words, regardless of which of these two methods of estimation are chosen, the two methods will yield the same results for models that possess this quality.

Empirical Example

The nonlinear latent curve model can be shown to give equivalent results to that from a partially nonlinear mixed-effects model given the same distributional assumptions for a set of data. This is done using a subset of the data presented in Blozis (2004). The data studied here are performance scores on a quantitative learning task for n = 228 study participants. The scores are the median response times for each of 12 trial blocks. Trial blocks are coded $\mathbf{t} = (0, \ldots, 11)'$ so that the intercept of the model is interpreted as the performance level in the first trial block. Data for a subset of 20 individuals are shown in Figure 1. A negatively accelerated exponential function was used in Blozis to describe the learning responses across trial blocks and is used here for this example. Specifically, the nonlinear latent curve model in (4) was first estimated using the first-order linearization method in SAS PROC NLMIXED with the following syntax:

```
PROC NLMIXED DATA=qrtdata METHOD=FIRO GCONV=0;
```

```
TITLE1 'Nonlinear latent curve model of Meredith & Tisak 1990';
PARMS a0=16 a1=8 gamma=.7 var0=12 cov10=1 var1=10 s2e=3;
theta0=a0+z0; theta1=a1+z1;
basis0=1-exp(-gamma*t); basis1=exp(-gamma*t);
predv= theta0*basis0 + theta1*basis1;
MODEL qrt ~NORMAL(predv,s2e);
RANDOM z0 z1 ~NORMAL([0,0], [var0,cov10,var1]) SUBJECT=subid;
RUN;
```

In the syntax for fitting the nonlinear latent curve model, the person-specific weights are defined as sums of fixed effects and random effects: theta0 = a0 + z0 and theta1 = a1 + z1. The fixed rate parameter is given by gamma. The basis functions of the matrix $\mathbf{\Lambda}$ in (5) are defined by basis0 and basis1. The predicted response, denoted by predv, is defined as the weighted sum of the basis functions as in (6); that is, each column of the basis functions matrix is weighted by its respective individual-specific weight. The observed response, denoted by qrt, is assumed to be normally distributed with predicted value predv and a residual variance s2e. By default, PROC NLMIXED assumes that the time-specific residuals are independent between individuals, as well as between occasions, with constant variance. The person-specific weights z0 and z1 are assumed to be bivariate normal with means equal to 0 and symmetric covariance matrix defined by the elements var0, cov10, and var1, where var0 is the variance of the first random weight, var1 is the variance of the second random weight, and cov10 is their covariance. The nonlinear latent curve model was also fit using the default estimation method of adaptive Gaussian quadrature by removing the METHOD=FIRO option.

To fit a partially nonlinear mixed-effects model using the exponential function to define the individual response in (8), PROC NLMIXED can be used with the following syntax, first using the first-order linearization method (METHOD=FIRO) and again using the default estimation method of adaptive Gaussian quadrature:

```
PROC NLMIXED DATA=qrtdata METHOD=FIRO GCONV=0;
TITLE1 ‘partially nonlinear mixed effects model’;
PARMS B0=16 B1=8 B2=.7 var0=12 cov10=1 var1=10 s2e=3;
theta0i=B0+b0i; theta1i=B1+b1i;
predv= theta0i - (theta0i - theta1i)*exp(-B2*t);
MODEL qrt ~NORMAL(predv,s2e);
RANDOM b0i b1i ~NORMAL ([0,0], [var0, cov10, var1]) SUBJECT=subid;
RUN;
```

In the syntax for fitting the partially nonlinear mixed-effects model, the coefficients theta0i and theta1i are defined as sums of fixed effects and random effects: theta0i = B0+b0i and theta1i = B1+b1i. The fixed rate parameter is given by B2. The predicted value, given by predv, is defined directly by the exponential function as defined in (8). The response, denoted by qrt, is assumed to be normally distributed with predicted value predv and residual variance s2e. Specifically, the trial-specific residual is assumed to be independent between individuals and between trials with constant variance s2e. The random-effects, b0i and b1i, are assumed to be bivariate normal with means equal to 0 and covariance matrix defined by the elements var0, cov10, and var1, where var0 is the variance of the first random effect, var1 is the variance of the second random effect, and cov10 is their covariance.

The results from fitting the two models to the quantitative response time data using the two methods of estimation are given in Table 1. Estimates of the two models using the first-order linearization method are given in the first two columns of results. ML estimates of the two models obtained using the default estimation method, adaptive Gaussian quadrature, are given in the last two columns. As shown, the estimated fixed-effects and the estimated variance and covariance parameters of the random effects and the variance of the residuals are equivalent to at least 4 decimal places across models and estimation methods, as are the estimated standard errors and indices of model fit. In summary, a nonlinear latent curve model and a partially nonlinear mixed-effects model, both of which include a nonlinear parameter that is strictly fixed across the population, yield equivalent models.

Structured Latent Curve Models

A form of a nonlinear latent curve model that differs from the nonlinear latent curve model of Meredith and Tisak (1990) and the nonlinear mixed-effects model is a structured latent curve model. Unlike the nonlinear latent curve model of Meredith and Tisak, a structured latent curve model allows for individual differences with regard to all model coefficients, as will be described later. Under a structured latent curve model the first step is to specify a function that is to describe the mean longitudinal response. An individual-level model is subsequently defined by a first-order Taylor expansion taken with respect to the parameters of the mean growth function and linearly weighted by a set of individual-specific weights (Browne, 1993; Browne & Du Toit, 1991). Thus, the structured latent curve model and the nonlinear latent curve model are both defined using a first-order Taylor expansion, and as was shown previously, the partially nonlinear mixed-effects model can also be expressed using this expansion. The structured latent curve model defines a specific function for the mean response, and as a result, the fixed coefficients of the model are interpreted as the effects that describe change in the mean response. This is a notable difference between a structured latent curve model and a fully nonlinear mixed-effects model. As described earlier for a fully nonlinear mixed-effects model, no assumption is made about the functional form of the mean response. Only the individual responses are assumed to follow a specific function and the population-level model describes the individual-level coefficients across the population.

An important aspect of a structured latent curve model that makes the model different from a nonlinear mixed-effects model (either fully nonlinear or partially nonlinear) and a nonlinear latent curve model (as defined in Meredith & Tisak, 1990) concerns the interpretation of the individual-level model. As described earlier, the individual responses under a nonlinear mixed-effects model, as well as the nonlinear latent curve model, are all assumed to follow the same functional form, such as the exponential function given in (3). So the interpretation of the model at the individual level depends on a common function and the particular coefficients that characterize the responses for the individual. Under a structured latent curve model, the individual-level model has a different interpretation. Specifically, the individual-level model of a structured latent curve model is defined by a first-order Taylor expansion taken with respect to all of the parameters of the mean function, including those that enter the model in a nonlinear way. The particular shape of an individual curve is dictated by the individual-specific weights that are applied to the set of common basis functions from the Taylor expansion, and as a result, an individual curve need not follow the same form as the mean response and may differ in form relative to the curves of other individuals (Browne, 1993; Browne & Du Toit, 1991). As illustrated in Blozis and Harring (2015), the structured latent curve model can result in individual curves that differ markedly from each other and from the averaged response. For instance, in an empirical example presented in Blozis and Harring, the mean response for a set of data was assumed to follow a monotonic function that included an asymptote. Although many of the individual curves in the sample followed this general form, there were several individuals whose response patterns differed, including individuals whose curves were not monotonic. This is a notable aspect of the structured latent curve model because it represents a greater degree of flexibility in how a model can be used to account for individual differences in change, although this also means that the interpretation of the individual curves is unlike the individual-level model of fully nonlinear and partially nonlinear mixed-effects model and the nonlinear latent curve model of Meredith and Tisak (1990).

The structured latent curve model is described in detail here. To help in the description, the empirical learning data that were presented earlier are used again. As was done in the previous examples, the trial blocks were coded as $\mathbf{t}_i = (0, \ldots, 11)'$ so that the intercept of the model represented the performance level in the first trial block. Under a structured latent curve model, the mean response is assumed to follow a particular function. Here, the mean response at trial t, μ_t, is assumed to follow a logistic growth model:

$$\mu_t = f(\boldsymbol{\theta}, t) = \frac{\theta_0 \theta_1}{\theta_0 + (\theta_1 - \theta_0)\exp\{-\theta_2 t\}} \tag{13}$$

where θ_0 is the expected mean performance level at the first trial block, θ_1 is the expected mean potential level, and θ_2 is a parameter that along with time governs the rate of change in the mean response. Given that the model for the mean response is at the population level, the coefficients of the function, θ_0, θ_1 and θ_2, are fixed. Unlike the exponential function in (3) in which two of the coefficients entered the function in a linear way and one coefficient entered in a nonlinear way, the logistic function in (13) includes parameters that all enter the target function in a nonlinear way. The logistic growth function simply represents another possible model to describe the learning responses.

At the individual level, the response y_{ti} at trial t is defined by a first-order Taylor expansion of the mean function $f(\boldsymbol{\theta}, t)$ with random weights applied to the basis functions. Assuming that the mean function is that defined in (13), the expansion that defines the individual-level model is

$$y_{ti} = f(\boldsymbol{\theta}, t) + z_{0i}f_0' + z_{1i}f_1' + z_{2i}f_2' + e_{ti} \tag{14}$$

where f_0', f_1', and f_2' are the first-order partial derivatives of the mean function $f(\boldsymbol{\theta}, t)$ in (13) taken with regard to each of the coefficients in $\boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2)'$, evaluated according to trial t, and weighted by individual-specific weights, z_{0i} , z_{1i} and z_{2i}. Let the k^{th} partial derivative of $f(\boldsymbol{\theta}, t)$ be denoted by $\frac{\partial f(\boldsymbol{\theta}, t)}{\partial \theta_k}$. The first-order partial derivatives of (13) taken with respect to θ_0, θ_1 and θ_2 are (cf: Browne, 1993)

$$\frac{\partial f(\boldsymbol{\theta}, t)}{\partial \theta_0} = f_0' = \frac{\theta_0 - \exp(-\theta_2 t) f(\boldsymbol{\theta}, t)}{\theta_0 + (\theta_1 - \theta_0)\exp(-\theta_2 t)} \tag{15a}$$

$$\frac{\partial f(\boldsymbol{\theta}, t)}{\partial \theta_1} = f_1' = \frac{\theta_1 - (1 - \exp(-\theta_2 t)) f(\boldsymbol{\theta}, t)}{\theta_0 + (\theta_1 - \theta_0)\exp(-\theta_2 t)} \tag{15b}$$

$$\frac{\partial f(\boldsymbol{\theta}, t)}{\partial \theta_2} = f_2' = \frac{(\theta_1 - \theta_0) t \exp(-\theta_2 t) f(\boldsymbol{\theta}, t)}{\theta_0 + (\theta_1 - \theta_0)\exp(-\theta_2 t)}. \tag{15c}$$

Note that the basis functions of (15a)-(15c) are common across all individuals in the population, and so, are functions only of fixed effects. In the individual level model in (14), these nonlinear basis functions are weighted by the random effects z_0, z_1, and z_2, all of which enter the individual-level model in a linear way. The expected values of these random effects are assumed to be zero (e.g., $E[z_{0i}] = 0$) so that the expected value of the individual response is assumed to be equal to the mean function:

$$E[\mathbf{y}_i] = \boldsymbol{\mu} = \mathbf{f}(\boldsymbol{\theta}, \mathbf{t}).$$

Thus, under a structured latent curve model, the mean response can be described by a function that possibly includes all nonlinear parameters, as in (13), all of which are fixed. At the individual level, the model is defined by the first-order Taylor expansion with weights z_0, z_1, and z_2, that are both random across individuals and enter the individual-level model in a linear way. This individual–specific weighting of the basis functions that defines the individual-level model allows for individual differences with respect to each aspect of change. That is, each of the basis functions represents change in the mean response with respect to the particular features that describe the mean response, and the individual-specific weights allow for individuals to differ with regard to each of these features. This aspect of the structured latent curve model distinguishes it from the nonlinear latent curve model of Meredith and Tisak (1990) in which individuals can vary only with regard to the parameters of the mean function that enter in a linear way.

The individual-specific model in (14) is not in the strict form of the latent curve model of Meredith and Tisak (1990). As described in Browne (1993, p. 1977), the structured latent curve model follows the same form as the latent curve model of Meredith and Tisak by making additional assumptions about the model. To describe these assumptions, the empirical learning example is again used. First, let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_{12})'$ be the set of mean responses across trial blocks. Further, let the basis functions in (15a), (15b) and (15c) form the columns of the matrix $\mathbf{\Lambda}$:

$$\mathbf{\Lambda} = \begin{bmatrix} \frac{\partial f(\boldsymbol{\theta},\mathbf{t})}{\partial \theta_0} & \frac{\partial f(\boldsymbol{\theta},\mathbf{t})}{\partial \theta_1} & \frac{\partial f(\boldsymbol{\theta},\mathbf{t})}{\partial \theta_2} \end{bmatrix}$$

where $\mathbf{\Lambda} = \mathbf{\Lambda}(\boldsymbol{\theta}, \mathbf{t})$. Under a structured latent curve model, the mean function is assumed to be invariant to a constant scaling factor (see Shapiro & Browne (1987, Condition 2). Many functions, including the Richards function that subsumes the exponential and logistic functions given here (Richards, 1959), satisfy this requirement. Assume that for the parameter $\boldsymbol{\theta}$ assumed to be in $\mathbf{G}$ and any positive scalar w, there is a parameter set $\boldsymbol{\theta}^*$ also in $\mathbf{G}$ that satisfies the following equality:

$$\mathbf{f}(\boldsymbol{\theta}^*, \mathbf{t}) = w\mathbf{f}(\boldsymbol{\theta}, \mathbf{t}).$$

The implication of this property (Shapiro & Browne, 1987, Lemma 1) of $\mathbf{f}(\boldsymbol{\theta}, \mathbf{t})$ is that there is a set of parameters, denoted here by $\boldsymbol{\alpha}$, such that

$$\mathbf{f}(\boldsymbol{\theta}, \mathbf{t}) = \mathbf{\Lambda}\boldsymbol{\alpha} \tag{16}$$

because $\mathbf{\Lambda}$ is the set of first-order partial derivatives of $\mathbf{f}(\boldsymbol{\theta}, \mathbf{t})$ taken with respect to $\boldsymbol{\theta}$ (see Browne, 1993, p 177). The parameter $\boldsymbol{\alpha}$ can be obtained by solving the linear equation in (16). For example, given that the logistic model for the mean response in (13) is invariant to a constant scaling factor (see Browne & Du Toit, 1991), $\boldsymbol{\alpha}$ can be obtained by solving the linear equation in (16), resulting in $\boldsymbol{\alpha} = (\theta_0, \theta_1, 0)'$. Note that although $\boldsymbol{\alpha}$ contains some of the elements in $\boldsymbol{\theta}$, $\boldsymbol{\alpha} \neq \boldsymbol{\theta}$. Then in the individual-level model, $\mathbf{\Lambda}\boldsymbol{\alpha}$ is used as a substitute for the mean function $\mathbf{f}(\boldsymbol{\theta}, \mathbf{t})$. Thus, by substituting $\mathbf{f}(\boldsymbol{\theta}, \mathbf{t})$ with $\mathbf{\Lambda}\boldsymbol{\alpha}$, the individual-level model in (14) is re-expressed as

$$\mathbf{y}_i = \boldsymbol{\Lambda}\boldsymbol{\alpha} + \boldsymbol{\Lambda}\mathbf{z}_i + \mathbf{e}_i \tag{17}$$

where again $\boldsymbol{\alpha} = (\theta_0, \theta_1, 0)'$ and $\mathbf{z}_i = (z_{0i}, z_{1i}, z_{2i}).'$ Letting $\boldsymbol{\eta}_i = \boldsymbol{\alpha} + \mathbf{z}_i$, the model in (17) further simplifies to

$$\mathbf{y}_i = \boldsymbol{\Lambda}\boldsymbol{\eta}_i + \mathbf{e}_i. \tag{18}$$

Assuming that the mean response follows a nonlinear growth function, such as the logistic growth model in (13) in which all growth coefficients enter the function in a nonlinear way, the random coefficient $\boldsymbol{\eta}_i$ in the individual level model in (18) enters the model in a linear way. The basis functions that define the matrix $\boldsymbol{\Lambda}$ are each based on functions that are nonlinear with regard to all three of the parameters, θ_0, θ_1 and θ_2, but these coefficients are all fixed.

The random weights $\mathbf{z}_i$ are assumed to be multivariate normal with expected values equal to zero and covariance matrix $\boldsymbol{\Phi}$:

$$\mathbf{z}_i \sim N[\mathbf{0}, \boldsymbol{\Phi}]$$

where $\boldsymbol{\Phi}$ is a symmetric matrix of order k, and k is the number of random weights. Within individuals, the residuals $\mathbf{e}_i = (e_{1i}, \ldots, e_{n_i})'$ are assumed to be multivariate normal with expected values equal to zero and covariance matrix $\boldsymbol{\Theta}$:

$$\mathbf{e}_i \sim N[\mathbf{0}, \boldsymbol{\Theta}]$$

where $\boldsymbol{\Theta}$ is a symmetric matrix of order n. Given the distributional assumptions for $\mathbf{e}_i$ and $\mathbf{z}_i$ and assuming independence between them, the marginal mean and covariance structure of the structured latent curve model are

$$\boldsymbol{\mu} = \boldsymbol{\Lambda\alpha}$$

and

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda\Phi\Lambda}' + \boldsymbol{\Theta},$$

respectively. The dimensions of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are n. Similar to all of the models presented previously, the structure for $\boldsymbol{\Theta}$ can take various forms.

Model interpretation. The fixed coefficients of the population-level model in (13), θ_0, θ_1 and θ_2, have two valid interpretations: First, they represent the typical characteristics of change across the population. This is analogous to the fully and partially nonlinear mixed-effects models and the nonlinear latent curve model of Meredith and Tisak (1990). Similar to the nonlinear latent curve model and the partially nonlinear mixed-effects model, the fixed parameters of the structured latent curve model can be also interpreted as the features that describe change in the mean longitudinal response. Thus, the structured latent curve model also possesses the property of dynamic consistency. In this way, the fully nonlinear mixed-effects model is unique to these other models because the fixed coefficients of a fully nonlinear mixed-effects model do not necessarily share this interpretation.

Estimation

The random weights $\boldsymbol{\eta}_i$ that combine with $\boldsymbol{\Lambda}$ to define the individual-level model of a structured latent curve model in (18) enter the model in a linear way. The nonlinear parameters of $\boldsymbol{\Lambda}$, $\boldsymbol{\theta}$, are fixed. Because the model is linear in the random weights, estimation of a structured latent curve model can rely on methods that are also used to fit linear mixed-effects models, partially nonlinear mixed-effects models, and the nonlinear latent curve model of Meredith and

Tisak (1990). This is different from the estimation requirements of a fully nonlinear mixed-effects model that includes a random nonlinear parameter. That is, for a model that includes a random nonlinear parameter, it is not possible to express a likelihood function in such a way as to analytically solve for the model parameters, and alternative methods of estimation are required. Solutions to this problem include methods that use approximations to a loglikelihood function, including a first-order Taylor series approximation, and numeric approximation methods, such as Gaussian quadrature (Davidian & Giltinan, 1995; Pinheiro & Bates, 1995). Conversely, for a structured latent curve model, a loglikelihood can be analytically expressed (as in (12)) and differential calculus applied to obtain an analytic solution for the parameters. This is an important aspect of the structured latent curve model because it indicates that estimation does not require the more complex methods that are typically used to fit fully nonlinear mixed-effects models. Further, the first-order linearization method, as described earlier for estimation of the nonlinear latent curve model of Meredith and Tisak and partially nonlinear mixed-effects model, and methods that are otherwise used to estimate fully nonlinear mixed-effects models, such as Gaussian quadrature, can yield equivalent results if applied in the estimation of a structured latent curve model. This is in contrast to the estimation of a model that includes random nonlinear parameters, as these two estimation approaches can yield different results.

As was briefly noted earlier, a first-order Taylor expansion can be used to approximate a nonlinear function of a fully nonlinear mixed-effects model, but this is not the case for models that are linear in the random effects because the first-order Taylor expansion is equal to the original function. To see this, apply the first-order Taylor series expansion proposed by Beal and Sheiner (1982) to the partially nonlinear mixed-effects model in (8) and the fully nonlinear mixed-effects model in (1), both defined using the exponential growth function. The first-order linearization of the exponential growth model in (8) for individual i would be

$$\begin{aligned} f_i(\boldsymbol{\beta}, \gamma, \mathbf{b}_i, Time_t) &= \beta_{1i} - (\beta_{1i} - \beta_{0i})\exp\{-\gamma Time_t\} \\ &= \beta_1 - (\beta_1 - \beta_0)\exp\{-\gamma Time_t\} + b_{0i}f_0' + b_{1i}f_1' \\ &= f_i(\boldsymbol{\beta}, \gamma, \mathbf{b}_i, Time_t), \end{aligned}$$

where

$$\begin{aligned} \frac{f_i(\boldsymbol{\beta}, \gamma, \mathbf{b}_i, Time_t)}{\partial \beta_{0i}}\bigg|_{\mathbf{b}_i=\mathbf{0}} &= f_0' = \exp\{-\gamma Time_t\}, \\ \frac{f_i(\boldsymbol{\beta}, \gamma, \mathbf{b}_i, Time_t)}{\partial \beta_{1i}}\bigg|_{\mathbf{b}_i=\mathbf{0}} &= f_1' = 1 - \exp\{-\gamma Time_t\} \end{aligned}$$

with the expressions evaluated at $b_{0i} = 0$ and $b_{1i} = 0$. Algebraically, the Taylor expansion is the same as the original function, f_i. This is not the case for the Taylor series expansion of the fully nonlinear mixed-effects model in (1). The Taylor expansion of the exponential function in (1) is no longer exponential but an approximation. Its expansion is

$$\begin{aligned} f_i(\boldsymbol{\beta}, \mathbf{b}_i, Time_t) &= \beta_{1i} - (\beta_{1i} - \beta_{0i})\exp\{-\beta_{2i} Time_t\} \\ &\approx \beta_1 - (\beta_1 - \beta_0)\exp\{-\beta_2 Time_t\} + b_{0i}f_0' + b_{1i}f_1' + b_{2i}f_2' \end{aligned}$$

where the first-order partial derivatives of f_i with respect to each parameter is as it was for the partial nonlinear model with the addition of the partial derivative of f_i with respect to β_{2i}

$$\frac{f_i(\boldsymbol{\beta}, \mathbf{b}_i, Time_t)}{\partial \beta_{2i}}\bigg|_{\mathbf{b}_i=\mathbf{0}} = f_2' = (\beta_1 - \beta_0) Time_t \exp\{-\beta_2 Time_t\}$$

with these expressions evaluated at $b_{0i} = 0$, $b_{1i} = 0$, and $b_{2i} = 0$. Clearly, the Taylor series approximation to the exponential function of (1) is not the same as the original function. Interestingly, the expected value of both Taylor series expansions lead to the same mean exponential function, however, the individual functions will not be the same; and for the fully nonlinear function in (1), the result is not the exponential function.

Empirical Example

To show that fitting a structured latent curve model using either the first-order linearization method or Gaussian quadrature can yield the same ML estimates, the quantitative learning data are again used. The syntax for fitting the structured latent curve model using the logistic growth model given in (13) and using the first-order linearization method of estimation is:

```
PROC NLMIXED DATA=qrtdata GCONV=0 METHOD=firo;
TITLE1 ‘Structured latent curve model using a logistic growth model’;
PARMS theta0=8 theta1=16 theta2=.7 var0=10 cov10=1 var1=12 cov20=.1 cov21=.1
var2=.3 s2e=3;
n0=theta0+z0; n1=theta1+z1; n2=z2;
meanf = (theta0*theta1)/(theta0 + (theta1-theta0)*exp(-theta2*t));
basis0 = (theta0 - exp(-theta2*t)*meanf) / (theta0 + (theta1-theta0)*exp(-theta2*t));
basis1 = (theta1 - (1-exp(-theta2*t))*meanf) / (theta0 + (theta1-theta0)*exp(-theta2*t));
basis2 = ((theta1-theta0)*t*exp(-theta2*t)*meanf) / (theta0 +
(theta1-theta0)*exp(-theta2*t));
predv= n0*basis0 + n1*basis1 + n2*basis2;
MODEL qrt ~NORMAL(predv,s2e);
RANDOM z0 z1 z2 ~NORMAL ([0,0,0], [var0,cov10, var1,cov20,cov21,var2])
SUBJECT=subid;
```

RUN;

In the syntax for fitting the structured latent curve model, the coefficients n0 and n1 are defined as a sum of a fixed effect and a random effect: n0=theta0+z0 and n1=theta1+z1. The coefficient n2 is defined only by the random effect z2. The fixed rate parameter theta2 is only used to define the basis functions, basis0, basis1 and basis2. The predicted value, given by predv, is the weighted sum of the three basis functions. The response, qrt, is assumed to be normally distributed with predicted value predv and a residual variance of s2e. The trial-specific residual is assumed to be independent between individuals and between trials with constant variance s2e. The random-effects, z0, z1 and z2 are assumed to be multivariate normal with means equal to 0 and covariance matrix defined by the elements var0, cov10, var1, cov20, cov21 and var2, where var0, var1 and var2 are the variances of the individual-specific weights, and their covariances are denoted by cov10, cov20 and cov21.

To fit the model using the default method of estimation, adaptive Gaussian quadrature, no METHOD option is specified. The results from fitting the model using the two estimation methods are given in the first two columns of Table 2. As shown in Table 2, the estimates and their standard errors are equivalent (to at least 4 decimal places), and model fit is the same. Thus, no matter the choice of these two estimation methods the same results are obtained, as expected. The computing times for the two methods were comparable. Using the starting values for the parameters in the syntax given earlier, fitting the model using the FIRO method required 2.1 seconds of CPU processing time and using the default method of adaptive Gaussian quadrature required 5.6 seconds of CPU processing time. Although computers can differ in their processing time, and as a result, the time reported here can differ from other computers, a report of the processing times between methods is of value. That is, given the negligible difference in processing times between methods of estimation, either method might be considered reasonable to apply in practice.

Comparing a Structured Latent Curve Model and a Fully Nonlinear Mixed-Effects Model

A structured latent curve model is, of course, one option for describing longitudinal data. By also fitting a nonlinear mixed-effects model to a set of data, it is possible to study the sensitivity of the fixed effects of a growth model to assumptions that are made about the between-subject variation. That is, under a nonlinear mixed-effects model, the curves of all individuals are assumed to follow the same functional form, and between-subject variation is characterized by the variation of the random coefficients about their corresponding fixed values (Davidian & Giltinan, 2003). Under a structured latent curve model, the mean response is assumed to follow a specific function, and the curves of the individuals vary with regard to their dependencies on the set of common basis functions, and so the individual curves may actually depart from the functional form that is assumed for the mean response. For a structured latent curve model the between-subject variation is characterized by the variation of the weights that are applied to the common basis functions (Blozis & Harring, 2015).

To show how these two models can differ in their characterization of longitudinal data, the same logistic growth function that was earlier used in the structured latent curve model to describe the average performance measures was also used to define the individual-level model of a fully nonlinear mixed-effects model. The fixed effects estimates of a nonlinear mixed-effects model are provided in the last two columns of Table 2. Estimates are based on a non-adaptive Gaussian quadrature method because no solution could be obtained using the default adaptive Gaussian quadrature method. The non-adaptive method was used with 30 quadrature points (for a general discussion on specifying the number of quadrature points when using adaptive versus nonadaptive Gaussian quadrature, see Lesaffre & Spiessens, 2001, and Pinheiro & Bates, 1995). As shown in Table 2, the parameter estimates and model fit are the same for the structured latent curve model using both methods of estimation (first-order linearization method and adaptive Gaussian quadrature), as well as the nonlinear mixed-effects model estimated using the first-order linearization method. That is, the fully nonlinear mixed-effects model estimated using the first-order linearization method returned equivalent results to the structured latent curve model. This is due to the fact that the first-order linearization of the nonlinear mixed-effects model is based on the same first-order Taylor expansion that is used to define the structured latent curve model at the individual level. This implies that if one wishes to fit a structured latent curve model and avoid the step of expressing the analytic derivatives that define the elements of the basis functions, then one only needs to express the mean function and use the first-order linearization method for estimation.

Parameter estimates and model fit differ, however, for the fully nonlinear mixed-effects model that was estimated using Gaussian quadrature. Figure 2 displays the fitted curves at the population level based on the structured latent curve model and the fully nonlinear mixed-effects model estimation using Gaussian quadrature, with the curves displayed along with the observed means at each trial block. As shown the figure, the structured latent curve model provides a better representation of the sample means relative to the fully nonlinear mixed-effects model. At the population level, the structured latent curve model specifies that the mean response follows the logistic function; conversely, at the population level of the fully nonlinear mixed-effects model, the curve represents the typical response that is not necessarily equivalent to the mean response.

Differences in model fit and estimated parameters of a fully nonlinear mixed-effects model using these two methods of estimation have been noted in the documentation for PROC NLMIXED (Littell et al., 2006). As shown here, however, the estimates of a nonlinear mixed-effects model obtained using the first-order linearization method can, in fact, provide estimates of a different model. Specifically, if a fully nonlinear mixed-effects model is estimated using a first-order linearization method, then the interpretation of the model aligns with that of a structured latent curve model. Specifically, if a growth function is invariant to a constant scaling factor, a fully nonlinear mixed-effects model estimated using the first-order linearization method will yield results that are identical to that of a structured latent curve model. Note that if the growth function is not invariant to a constant scaling factor, a first-order linearization method of estimation of a nonlinear mixed-effects model will simply serve its purpose of providing an approximation to the function. It is worth noting that the models considered in Preacher and Hancock (2015) are variations of a structured latent curve model. Unlike the structured latent curve model of Browne (1993) and Browne and Du Toit (1991), their models do not require that the chosen mean function be invariant to a constant scaling factor. As a result, the individual-level model provides an approximation to the individual's true response trajectory similar to the approximation that results when using a Taylor expansion to approximate the individual-level response under a fully nonlinear mixed-effects model.

It is also notable that the processing time for the fully nonlinear mixed-effects model estimated using the Gaussian quadrature was relatively intensive. Using the starting values for the parameters in the syntax given earlier, fitting the nonlinear mixed-effects model using

Gaussian quadrature required 2 hours, 11 minutes and 20.7 seconds of CPU processing time. The structured latent curve model fitted using either method of estimation, as well as the fully nonlinear mixed-effects model that was fitted using the first-order linearization method, all required less than 10 seconds of CPU processing time.

A Note on Population-Average Models

Population-average models are used for the analysis of longitudinal data when the focus of study is on the population rather than the individual (Liang & Zeger, 1986; Zeger & Liang, 1986). The objective is generally to describe a variable over time at the population level, that is, to describe the average behavior across the population. Thus, the population averaged response is modeled directly. The covariance structure represents an aggregate of the patterns of association in the responses that are due to the combined influences of within-individual and between-individual sources. So the model does not include random effects that are specific to the individual. Different forms of the covariance structure can be considered to address the patterns of variation and covariation, such as using an unstructured covariance matrix or a first-order autoregressive structure (see Davidian & Giltinan, 2003).

A popular method of estimation of population-average models are generalized estimating equations (GEE) (Liang & Zeger, 1986; Zeger & Liang, 1986). The SAS macro %nlinmix is available for fitting these models using the GEE approach described in Zeger, Liang, and Albert (1988) (see Littell et al., 2006, Chapters 14 and 15). The macro is available in SAS 8 and subsequent versions and is available for download through SAS support. The macro relies on two SAS procedures, PROC NLIN and PROC MIXED. PROC NLIN is a procedure for fitting nonlinear regression models that include only fixed effects. PROC MIXED is a procedure for fitting linear mixed-effects model that can include fixed and random effects that enter the model in a linear way. The %nlinmix macro makes it possible to fit nonlinear mixed-effects models, as well as population-average models, and allows for the same selection of covariance structures that are available as options in PROC MIXED using the REPEATED statement. Because only a single covariance structure is used to account for variances and covariances in the data of a population-average model, no RANDOM statement is used as it is in mixed-effects models to address individual-specific coefficients.

The discussion of population-average models is relevant to the discussion of the structured latent curve model because the mean response is modeled directly in both models. The interpretation of the population-level model of the structured latent curve model is analogous to that of a population-average model in that both describe the average response using a specific function. That is, one interpretation of the fixed effects of a structured latent curve model is that they represent the characteristics of change in the average response, and this is also the case for a population-average model. A key difference between a population-average model and a structured latent curve model is in how the two define the covariance structure. In a population-average approach, a single covariance structure aggregates within- and between-person sources of variation and covariation into a single matrix. Conversely, in a structured latent curve model the covariance structure partitions patterns of association in the response according to within- and between-person sources. Thus, although the averaged response is modeled directly by a growth function, similar to a population-average model, the structured latent curve model also includes an individual-level model that is characterized by individual-specific weights and a separate component to characterize within-individual variation.

To illustrate differences in a population-average model and a structured latent curve model, a population-average model was fit to the learning data assuming three different covariance structures: unstructured, a first-order autoregressive structure with homogeneous variances, and a first-order autoregressive structure with heterogeneous variances. These structures were chosen because they allowed for correlations between the residuals but in different ways. Two of the structures, namely the unstructured covariance matrix and the first-order autoregressive with heterogeneous variances, also allowed for the variances to be different across trials. Other structures could be considered. The unstructured covariance matrix means that all of the variances of the residuals and the covariances between them are estimated without any constraints in their values. Using the %nlinmix macro, this structure is available using the option TYPE=UN. For the learning data, this results in a total of $12(12-1)/2 = 78$ unique variances and covariances to be estimated. A first-order autoregressive residual structure with homogeneous residual variances allows for the residuals between adjacent time points to covary and assumes that the magnitude of the correlation decreases with an increase in distance between time points with the assumption that the variances are constant across time. Specifically, the i,j^{th} element of the covariance matrix $\mathbf{\Sigma}$ is given by $\sigma^2\rho^{|i-j|}$, where σ^2 is the common variance of the residuals across time. Using the %nlinmix macro, this structure is available using the option TYPE=AR(1). For the learning data, this results in a total of 2 parameters (σ^2 and ρ) to be estimated. A first-order autoregressive residual structure with heterogeneous residual variances allows for the residuals between adjacent time points to covary and assumes that the magnitude of the correlation decreases with an increase in distance between time points, and in addition, allows the residual variances to differ across time. Specifically, the i,j^{th} element of the covariance matrix $\mathbf{\Sigma}$ is given by $\sigma_i\sigma_j\rho^{|i-j|}$, where the covariance is given by the product $\sigma_i\sigma_j$. Using the %nlinmix macro, this structure is available using the option TYPE=ARH(1). For the learning data, this results in a total of 12 + 1 parameters (σ_i for $i = 1, \ldots, 12$ and ρ) to be estimated.

These options for the covariance matrix are available using the REPEATED statement when employing the SAS macro %nlinmix. The estimated fixed effects of the population-average model are given in the last three columns of estimates in Table 3. For comparison, the structured latent curve model was fit again to the learning data assuming alternative within-person covariance structures that were described in Harring and Blozis (2014): a first-order autoregressive structure with homogeneous variances and independent residuals with heterogeneous variances. Estimates reported in Table 2 for the structured latent curve model assuming that the residuals are independent with constant variance are repeated in the first column of results in Table 3 for comparison purposes.

With regard to model fit, the structured latent curve model that assumes that the residuals have different variances across the trial blocks has the best fit, as indicated by this model having the smallest values of the AIC and BIC. Thus, this is an improvement over the first version of the model that was applied in which the residual variances were assumed to have constant variance across trial blocks and also provides a better fit than a model that assumes that the residuals follow a first-order autoregressive structure. The best fitting population-average model is that which assumes that the overall covariance structure is unstructured. As shown, the estimated fixed effects between this model and the best-fitting structured latent curve model are not appreciably different, and so similar conclusions about the mean performance scores might be made. What is different about the two models is that the structured latent curve model breaks down the covariance structure into within-individual and between-individual components, whereas the population-average model pools this information into a single matrix. Taking into account model parsimony, the structured latent curve model gives superior fit overall and may be preferred because the model specifically addresses change in the response that is unique to the individual.

Discussion

Understanding individual differences in psychological and behavioral measures is central to behavioral research. Longitudinal data allow researchers to expand the study of individual differences to also allow for the study of within- and between-person change. Longitudinal data provide the means to study how, within person, behaviors may change over time and how individuals may vary from one another in change. Depending on the scale of measurement of the response, as well as the form of change in the response over time, there are several options for longitudinal data analysis. Mixed-effects models and latent curve models offer many options for analyzing longitudinal data, including options for analyzing both normal and non-normal response data, as well as various forms of change.

Mixed-effects models and latent curve models share a common goal of providing for the simultaneous study of within-person variation and between-person variation. If a mixed-effects model and a latent curve model are linear in their parameters, then the two approaches can be used to specify models that are equivalent and parameter estimates will be equal given the same set of data and the same model assumptions (Bauer, 2003). If a model is nonlinear in its parameters, however, there can be important differences between the two approaches. This paper was informative in showing that a partially nonlinear mixed-effects model, a particular version of a nonlinear mixed-effects model that restricts the random effects to enter the model in a linear manner (nonlinear effects must be fixed), and the latent curve model of Meredith and Tisak (1990) are equivalent, and given that the random effects of the models can only enter the model in a linear manner, the estimation requirements of the models match those of linear mixed-effects models. This paper also showed that a structured latent curve model, a special case of a nonlinear latent curve model that allows for individual differences in all model coefficients, is in fact quite different from a nonlinear mixed-effects model, both in its interpretation and with regard to its estimation requirements. This paper also showed that if a nonlinear mixed-effects model is defined using a growth function that is invariant to a constant scaling factor and a first-order linearization method of estimation is applied, the resulting estimates will actually be those of a structured latent curve model.

The individual-level model of a structured latent curve model is represented by a decomposition of a set of common basis functions and a set of random effects that are unique to the individual. This is done by using a first-order Taylor expansion of a common function taken with respect to the coefficients of the function that is then weighted by individual-specific coefficients. The first-order linearization method of estimation relies on this expansion. The first-order linearization method that is employed to estimate nonlinear mixed-effects models is essentially this approach, and so the estimates of a nonlinear mixed-effects model using this linearization method will be equivalent to defining a structured latent curve model, assuming that the function is invariant to a constant scaling factor. This is an important results because it indicates that the choice of estimation method for a nonlinear mixed-effects model can yield estimates of a model that was not intended for the data. This is not the case for the nonlinear latent curve model of Meredith and Tisak (1990) and a partially nonlinear mixed-effects model, as well as the structured latent curve model, because these models can all be expressed using a first-order Taylor expansion to define the individual-level model, and the expansion does not represent an approximation to the function but rather is a direct re-expression of the function.

References

Bates, D. M., & Watts, D. G. (1988). *Nonlinear Regression Analysis and Its Applications*. New York: Wiley.

Bauer, D. J. (2003). Estimating multilevel linear models as structural equation models. *Journal of Educational and Behavioral Statistics*, *28*(2), 134-67.

Beal, S. L., & Sheiner, L. B. (1982). Estimating population kinetics. *CRC Critical Reviews in Biomedical Engineering*, *8*, 195-222.

Blozis, S. A. (2004). Structured latent curve models for the study of change in multivariate repeated measures. *Psychological Methods*, *9*, 334-353. doi: 10.1037/1082-989X.9.3.334

Blozis, S. A., & Cudeck, R. (1999). Conditionally linear mixed-effects models with latent variable covariates. *Journal of Educational & Behavioral Statistics*, *24*(3), 245-270.

Blozis, S. A., & Harring, J. (2015). Understanding individual-level change through the basis functions of a latent curve model. *Sociological Methods and Research*, 1-28.

Browne, M. W. (1993). *Structured latent curve models*. In C. M. Cuadras & C. R. Rao (Eds.), *Multivariate analysis: Future directions 2* (pp. 171-198). Amsterdam: North-Holland.

Browne, M. W., & Du Toit, S. H. C. (1991). Models for learning data." In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 47-68). Washington, D.C.: American Psychological Association.

Burke, C. T., Shrout, P. E., & Bolger, N. (2007). Individual differences in adjustment to spousal loss: A nonlinear mixed model analysis. *International Journal of Behavioral Development*, *31*(4), 405-415.

Cudeck, R. (1996). Mixed-effects models in the study of individual differences with repeated measures data. *Multivariate Behavioral Research*, *31*(3), 371-403.

Cudeck, R., & Klebe, K. J. (2002). Multiphase mixed-effects models for repeated measures data. *Psychological Methods*, *7*(1), 41-63.

Davidian, M., & Gallant, A. R. (1993). The nonlinear mixed effects model with a smooth random effects density. *Biometrika*, *80*(3), 475-488.

Davidian, M., & Giltinan, D. M. (1995). *Nonlinear models for repeated measurement data*. Chapman & Hall/CRC Press.

Davidian, M., & Giltinan, D. M. (2003). Nonlinear models for repeated measurement data: An overview and update. *Journal of Agricultural, Biological, and Environmental Statistics*, *8*, 387-419.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis* (2nd ed.). New York: John Wiley & Sons, Inc.

Grimm, K. J., Ram, N., & Hamagami, F. (2011). Nonlinear growth curves in developmental research. *Child Development*, *82*, 1357-1371.

Harring, J. R., & Blozis, S. A. (2014). Fitting correlated residual error structures in nonlinear mixed-effects models using SAS PROC NLMIXED. *Behavioral Research Methods*, *46*(2), 372-384.

Jennrich, R. I., & Schluchter, M. D. (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics*, *42*(4), 805-820.

Keats, J. A. (1983). Ability measures and theories of cognitive development. *Principles of modern psychological measurement: A festschrift for Frederick M. Lord*, pp. 81-101.

Kohli, N., Hughes, J., Wang, C., Zopluoglu, C., & Davison, M. L. (2015). Fitting a linear-linear piecewise growth mixture model with unknown knots: A comparison of two common approaches to inference. *Psychological Methods*, *20*(2), 259-275.

Laursen, B., Little, T. D., & Card, N. A. (Eds.) (2012). *Handbook of developmental research methods*. New York: Guilford.

Lesaffre, E., & Spiessens, B. (2001). On the number of quadrature points in a logistic random effects model: An example. *Applied Statistics*, *50*(3), 325-335.

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13-22.

Lindstrom, M. J., & Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*, *46*(3), 673-687.

Littell, R. C., Stroup, W. W., Milliken, G. A., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for mixed models*. SAS institute.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*(1), 107-122.

Pinheiro, J. C., & Bates, D. M. (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of computational and Graphical Statistics*, *4*(1), 12-35.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer Science & Business Media.

Preacher, K. J., & Hancock, G. R. (2015). Meaningful aspects of change as novel random coefficients: A general method for reparameterizing longitudinal models. *Psychological Methods*, *20*(1), 84-101.

Ram, N., & Grimm, K. J. (2007). Using simple and complex growth models to articulate developmental change: Matching method to theory. *International Journal of Behavioral Development*, *31*(4), 303-316.

Vonesh, E. F. (1996). A note on the use of Laplace's approximation for nonlinear mixed-effects models. *Biometrika*, *83*(2), 447-452.

Vonesh, E., & Chinchilli, V. M. (1996). *Linear and nonlinear models for the analysis of repeated measurements*. CRC press.

Zeger, S. L., & Liang, K.-Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, *42*(1), 121-130.

Zeger, S. L., Liang, K.-Y., & Albert, P. S. (1988). Models for longitudinal data: A generalized estimating equations approach. *Biometrics*, *44*(4), 1049-1060.

Appendix

Distinguishing Between Linear and Nonlinear Parameters of a Growth Model

The distinction between linear and nonlinear parameters of a growth model is provided here. A parameter enters a function in a nonlinear way if the first-order partial derivative of the function taken with regard to that parameter results in a function that is nonlinear with regard to that parameter (Bates & Watts, 1988). Conversely, a parameter enters a function in a linear way if the first-order partial derivative of the function taken with regard to that parameter results in a function that is linear with regard to that parameter.

From this it is easy to show that polynomial functions, often used to approximate nonlinear relationships, are considered within the class of linear models because polynomial functions are linear with regard to their parameters. Consider, for instance, a cubic function, a polynomial function of degree 3, that can be used to model a nonlinear relationship between *X* and *Y*:

$$Y = f(X) = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$$

Although the relationship between *X* and *Y* is nonlinear in shape, the function is linear in its parameters because all first-order partial derivatives of the function return functions that are linear with respect to each of the parameters of the function, even though the last of the four functions is nonlinear with regard to *X*:

$$\frac{\partial f(X)}{\partial \beta_0} = 1$$
$$\frac{\partial f(X)}{\partial \beta_1} = X$$
$$\frac{\partial f(X)}{\partial \beta_2} = 2X$$
$$\frac{\partial f(X)}{\partial \beta_3} = 3X^2.$$

In a nonlinear model, at least one of coefficients of the model enters in a nonlinear way. Burke, Shrout, and Bolger (2007), for instance, applied a nonlinear mixed-effects model to measures of adjustment to spousal loss. In one model, a measure of depressive symptoms measured over time was assumed to follow a two-level model (a modified form of the function presented in Burke et al.):

$$y_{ij} = \beta_{1i} - (\beta_{1i} - \beta_{0i})\exp\{-\beta_{2i}t_{ij}\} + \varepsilon_{ij} \qquad \text{(a)}$$

where β_{1i} was the asymptotic response representing the level of depressive symptoms assumed to remain long after the loss, β_{0i} was the level of depressive symptoms at the time of loss (with t, a measure of time, centered at the time of loss), and β_{2i} along with t_{ij} assumed to govern the rate of change over time. Each of the coefficients, β_{0i}, β_{1i}, and β_{2i}, was a sum of a fix and random effect (e.g., $\beta_{0i} = \beta_0 + b_{0i}$) with the latter allowing the particular features relating to the bereavement response to among between individuals.

To show that the model for depressive symptoms in (a) is nonlinear with regard to one of its parameters, the first-order partial derivatives of the function taken with respect to each of the coefficients are given here:

$$\frac{\partial f(\boldsymbol{\beta}, t_{ij})}{\partial \beta_{0i}} = \exp\{-\beta_{2i} t_{ij}\}$$
$$\frac{\partial f(\boldsymbol{\beta}, t_{ij})}{\partial \beta_{1i}} = 1 - \exp\{-\beta_{2i} t_{ij}\}$$
$$\frac{\partial f(\boldsymbol{\beta}, t_{ij})}{\partial \beta_{2i}} = (\beta_{1i} - \beta_{0i})(t_{ij}) \exp\{-\beta_{2i} t_{ij}\}$$

As shown, the first two of the first-order partial derivatives taken with regard to β_{0i} and β_{1i}, namely $\frac{\partial f(\boldsymbol{\beta}, t_{ij})}{\partial \beta_0}$ and $\frac{\partial f(\boldsymbol{\beta}, t_{ij})}{\partial \beta_1}$, result in nonlinear functions with regard to the third coefficient, β_{2i}, but not the two coefficients themselves. Thus, the function in (a) is linear with regard to both β_{0i} and β_{1i}. The last of the first-order partial derivatives that is taken with regard to β_{2i}, namely $\frac{\partial f(\boldsymbol{\beta}, t_{ij})}{\partial \beta_2}$, results in a nonlinear function of β_2. Thus, the function in (a) is nonlinear with regard to β_{2i} but not β_{0i} and β_{1i}.