A General Multivariate Latent Growth Model With Applications to Student Achievement

Silvia Bianconcini
Silvia Cagnone
Department of Statistics, University of Bologna

The evaluation of the formative process in the University system has been assuming an ever increasing importance in the European countries. Within this context, the analysis of student performance and capabilities plays a fundamental role. In this work, the authors propose a multivariate latent growth model for studying the performances of a cohort of students of the University of Bologna. The model proposed is innovative since it is composed by (a) multivariate growth models that allow the capture of different dynamics of student performance indicators over time and (b) a factor model that allows measurement of the general latent student capability. The flexibility of the model proposed allows its applications in several fields such as socioeconomic settings in which personal behaviors are studied using panel data.

Keywords: generalized linear latent variable models; longitudinal and mixed data; EM algorithm

Introduction

The Bologna Process started in 1999 with the aim of creating a European Higher Education Area, in which students could choose from a wide and transparent range of high-quality courses and benefit from smooth recognition procedures. The Bologna Declaration has initiated a series of reforms needed to make European Higher Education more compatible and comparable, more competitive and more attractive than before, both for Europeans and for students from other continents. Hence, the evaluation of formative processes has received growing attention from policy makers and public agents in order to identify critical factors to improve curricula, instruction strategies, and learning conditions. To this purpose, several universities have created Data WareHouse (DWH) systems to collect detailed multivariate individual responses over time, which consist of mixtures of count, categorical, and continuous observations. These longitudinal data enable questions to be answered regarding student progress, evaluate how each individual performs over time (within-individual change), and predict changes in the differences between individuals (interindividual differences in change). However, in the presence of
multidimensional observations, the characterization of both temporal and cross-sectional dependencies among response variables with different measurement scales is a challenging problem. In such cases, it is natural to consider models in which dependency among responses is due to the presence of both latent variables and random effects, as shown in several approaches developed in the literature.

Roy and Lin (2000) proposed a two-step linear mixed model applied to multiple continuous outcomes. These authors use time-dependent factors to account for correlations of items within time. In addition to this, random effects are introduced to explain correlations across time of items and time-dependent latent variables. An extension to such models is provided by Dunson (2003), who introduced a dynamic latent trait model for multidimensional longitudinal data in the context of the Generalized Linear Latent Variable Model (GLLVM) so that different kinds of observed variables can be considered. An autoregressive structure that allows for covariates is used to model the dynamic of the time-dependent latent variables. The model is estimated with Markov Chain Monte Carlo (MCMC). Within the same framework, a full information likelihood estimation method via the Expectation–Maximization (EM) algorithm is developed by Cagnone, Moustaki, and Vasdekis (2009) in the specific case of ordinal data.

A general approach is presented by multilevel models (Skrondal & Rabe-Hesketh, 2004) that allow longitudinal and/or multidimensional mixed data to be dealt with. In repeated measures, occasions are viewed as first level units whereas respondents are second level units. With multidimensional data, first level units are represented by items nested within individuals. When data are both multidimensional and longitudinal, more complex hierarchical structures have to be taken into account.

Multidimensional and longitudinal data can be analyzed within the Structural Equation Model (SEM) framework following two different approaches.

According to the first approach (Jöreskog & Sörbom, 2001), a standard confirmatory factor model is considered, the main feature being that the measurement errors are correlated over time. Moreover, the latent variables are identified by setting the same loading over time to 1.

The second approach is represented by latent growth models, widely applied in the analysis of change (Singer & Willett, 2003). Random effects are included in the model to account for individual differences in both the initial status and in the rate of growth. The particular feature of such models is that random coefficients are treated as latent variables within the traditional SEM approach (Muthén & Khoo, 1998). Univariate analyses are usually performed by studying the temporal dynamics of a single indicator, considered as a proxy of the individual performance. Multivariate analysis essentially consists of modeling the trajectories of several items separately, and then allowing for correlations among random coefficients (Bollen & Curran, 2006; Raykov, 2007).

In this article, we propose a new general class of models, which combines (a) multivariate latent curves that describe the time-specific dependencies of the
responses, and (b) a factor model that specifies the relationship between manifest and latent variables. The use of this new methodology has been motivated by data from the DWH of the University of Bologna. We focus on student achievements measured through two items observed at different time points. These achievements can depend on skills that evolve over time (due, e.g., to the student learning process) as well as individual capabilities that are time-independent. The former can be evaluated by analyzing the dependencies of the same and different items across time. The latter can be detected by studying the atemporal relationships between the two items. In the proposed approach, the first component is modeled by means of multivariate latent curves, whereas the second component is modeled by including common factors that do not change over time.

In the literature there are two main approaches that include common factors in multivariate latent growth models: the “factor-of-curves” and the “curve-of-factors,” both discussed by McArdle (1988). In the former, univariate growth models are specified for each item, and correlations between the growth coefficients are explained by means of higher order factors. In the latter, a growth curve is fitted to factor scores representing what the items have in common at each time point. Our proposal shares similarities with the “factor-of-curves,” but also presents significant differences. Both allow a flexible treatment of the longitudinal component by fitting a multivariate growth model. However, in our approach, correlations between curves are freely estimated and common factors are included to explain the atemporal variability between items. One of the main advantages of this modelization is that we can carry out a longitudinal analysis on different kinds of data and, at the same time, evaluate whether atemporal common factors are present. The model is developed within the GLLVM framework (Bartholomew & Knott, 1999). Following this approach, the response variables are assumed to follow different distributions of the exponential family, with item-specific linear predictors depending on latent variables and measurement errors. Furthermore, we extend the GLLVM by including item-specific random coefficients so that each item has its own trajectory over time.

The article is organized as follows. In the Data section, we present the data source and perform an exploratory analysis to demonstrate the potential of our approach in describing student performances over time. Modeling and Estimation section describes the proposed methodology in terms of model specification, identification, and estimation. In the Results section, we present the results of model estimation for the overall data set and for different temporal patterns observed in the sample. We conclude with a Discussion section.

Data

The data set analyzed was extracted from the DWH of the University of Bologna, which is a system that collects and constantly updates information by
integrating data coming from different sources. The project started in 2002 in order to support planning, control, and decision-making processes.

The DWH contains a great deal of information on each student and enables overall university student careers to be examined. Sociodemographic information (gender, country/region of origin, etc.) and the final Secondary School exam mark is available. We decided to analyze the cohort of \( n = 821 \) students enrolled at the Faculty of Economics in the academic year 2001/2002 as this Faculty is one of the largest at the University of Bologna and the year chosen is the first available in the DWH, so several time points can be observed. We analyzed the performance of the selected students in the academic years: \( t_1 = 2001/2002; \)
\( t_2 = 2002/2003; \) and \( t_3 = 2003/2004 \) as the University system requires a degree to be obtained over 3 years. In the data set, two variables are available for evaluating the student performance over time: the mark obtained in every exam and the number of exams taken per time point. The former is a score given for student performance in each exam and it can range from 18 (minimum mark to pass the exam) to 30 cum laude (maximum mark). For the latter, in each academic year there is a scheduled number of exams that the students should take. This number varies both from Faculty to Faculty and from student to student should they not manage to take all the scheduled exams in time. Good performances are obtained from students who take the scheduled exams at each time point as well as students who increase the number of exams taken over time. The average of the marks (AM) per student in each academic year is considered. In the left side of Figure 1, the AM trajectories from a random sample of 100 students are reported. Such trajectories vary from 18 to 30 even if they are mostly concentrated in the range 21 to 28, and the overall mean (black line) is over 24 for all the observed time points.

In a previous article, Bianconcini, Cagnone, Mignani, and Monari (2007) analyzed these data to detect the presence of different patterns of behavior in terms of student performance. This latter was measured using a continuous composite indicator of the number of exams (NE) and the average marks.
A multigroup analysis was carried out to investigate if these performances were a function of specific characteristics such as gender, passages within and between faculties, and the year of graduation. Only the latter showed significant differences between groups in terms of mean vector and covariance matrix elements. Hence, within the selected sample, we distinguish two different types of temporal behavior: the first represents students who graduated in the first 3 years (Grad) while the second is students that at $t_3$ have not yet managed to graduate (Nograd). Indeed, as shown in the center and in the right side of Figure 1, the Grad presents a higher overall average mark and a lower variability than Nograd. In Table 1, the descriptive statistics of AM for the overall sample, Grad ($n_1 = 195$), and Nograd ($n_2 = 626$) are reported. Grad presents the highest correlations over the three time points. On the other hand, Nograd is very similar to the Overall sample in terms of both first and second order moments. The number of exams (NE) is a count variable whose range is different in the observed time points. Table 2 shows the number of students classified according to NE taken in both the three time points and the groups defined before. The hyphens indicate that in the first academic year, a student can take at most eight exams. We can observe that, in general, Grad students present the same behavior over the 3 years, that is, they take a number of exams greater than three, and concentrate it between six and seven. In contrast, Nograd students take a number of exams equal to or greater than zero, mostly concentrated between four and six. Furthermore, few students from the overall sample take more than eight exams at $t_2$ (only 4.1%) and at $t_3$ (only 8.1%). It can be useful to evaluate whether (a) the variable NE shows a dependence over time and (b) there is an association between the variables AM and NE within the same time. To this aim, the variable AM has been recoded for all the time points into four classes according to the quartiles of the distribution. As for the variable NE, the categories from 1 to 3 and categories from 9 to 14 have been collapsed to avoid the problem of sparseness that affects these data in the extreme categories. In Table 3, the values of the chi-square tests (with associated $p$ values) are reported for all the pairs of NE over time and for all the pairs of AM and NE.
within time. The association between AM and NE for Grad is not significant at time $t_2$, whereas all others are significant, indicating that both the variables can be good indicators of a common factor.

### Modeling and Estimation

**Multivariate Latent Growth Curves**

Suppose that $J$ items are observed for $n$ individuals at $T$ different time points. The measured outcomes for a randomly selected individual are denoted by $y = (y_1',...,y_J',...,y_J')'$, where the elements $y_j = (y_{ij},...,y_{ij})'$, $j = 1,2,...,J$, consist of mixtures of count ($j = 1,...,J_1$) and continuous ($j = J_1+1,...,J$) responses. In analyzing data of this type, a challenging problem is the characterization of both the temporal and the cross-sectional dependency among variables that have different measurement scales. In such cases, it is natural to consider models in which the dependencies are due to the presence of both several latent variables and random effects, stacked into the vector $b$ (Cagnone et al., 2009). The marginal distribution of the overall responses is given by

$$f(y) = \int g(y|\eta)h(\eta)d\eta,$$

(1)
where \( g(y|\eta) \) is the conditional distribution of the responses \( y \), given the latent variables \( \eta \), and \( h(\eta) \) is their prior density function. We refer to the GLLVM framework developed in Moustaki and Knott (2001) for multivariate mixed responses. We extend that framework to allow for multivariate longitudinal data. One of the main assumptions of the GLLVM approach is the conditional independence of the responses (within and over time) given the latent variables, that is

\[
g(y|\eta) = \prod_{j=1}^{J} \prod_{t=1}^{T} g(y_{jt}|\eta), \tag{2}
\]

where \( g(y_{jt}|\eta) \) is a distribution of the exponential family. For count data,

\[
g(y_{jt}|\eta) = \binom{n_t}{y_{jt}} \left( \frac{\exp(\nu_{jt})}{1 + \exp(\nu_{jt})} \right)^{y_{jt}} \left( \frac{1}{1 + \exp(\nu_{jt})} \right)^{n_t-y_{jt}} \quad j = 1, \ldots, J, \tag{3}
\]

where \( \nu_{jt} \) is the linear predictor of the generalized linear model and it is defined below. \( n_t \) is the number of “trials.” Here count data are modeled using a binomial distribution, instead of the classical Poisson distribution, since NE is upper bounded, given that there is a scheduled number of exams for each year. Therefore \( n_t \) is the maximum NE observed at time \( t \).

On the other hand, for continuous data

\[
g(y_{jt}|\eta) = \frac{1}{\sqrt{2\pi}\sigma_{jt}} \exp \left( -\frac{1}{2} \left( \frac{y_{jt} - \nu_{jt}}{\sigma_{jt}} \right)^2 \right) \quad j = J_1 + 1, \ldots, J, \tag{4}
\]

where \( \nu_{jt} \) is again the linear predictor and \( \sigma_{jt}^2 \) is the variance of the continuous responses supposed to be heteroscedastic over time and between items.

As in the classical generalized linear model, \( \nu_{jt} \) is the linear predictor for the \( j \)th outcome at time \( t \) and the link between the linear predictor and the conditional means of the random distributions can be any monotonic differentiable function. In this context, the link is the logit of the probability associated to each count for
the binomial distribution defined in Equation 3, and the identity function in the case of the normal distribution defined in Equation 4.

For both kinds of observed variables, the linear predictor is defined as

$$v_{ij} = \sum_{r=0}^{p} \lambda^i_r \beta^j_r + \sum_{k=1}^{q} \lambda^i_k z_k, \quad t = 1, \ldots, T, \quad j = 1, \ldots, J,$$

where \( p \) is the degree of the fitted trajectory. In matrix form

$$v_{ij} = w_{ij}' \eta, \quad t = 1, \ldots, T, \quad j = 1, \ldots, J,$$

where, for each \( t \) and \( j \), \( w_{ij} \) is an \( m \)-dimensional vector, with \( m = J( p + 1 ) + q \), such that

- \( w_{i1}' = [1, \lambda_1, \ldots, \lambda_p^i, 0, \ldots, 0, \lambda_{11}, \lambda_{21}, \ldots, \lambda_{q1}] \)
- \( \ldots \)
- \( w_{iJ}' = [0, \ldots, 0, 1, \lambda_1, \ldots, \lambda_p^i, \lambda_{1J}, \lambda_{2J}, \ldots, \lambda_{qJ}] \)

The latent variables \( \eta' = (\beta_{01}, \ldots, \beta_{p1}, \ldots, \beta_{0J}, \ldots, \beta_{pJ}, z_1, \ldots, z_q) \) are random effects and latent traits that account for both the temporal and the cross-sectional dependence between items. As is done in “classical” univariate growth models, the random coefficients \( \beta^j_r = (\beta_{0j}, \ldots, \beta_{pj}), j = 1, \ldots, J \), and the corresponding loadings \( \lambda^i_r, t = 1, \ldots, T, r = 0, \ldots, p \), are introduced in order to describe the temporal behavior of each item. The \( \lambda^i_r \)'s can either be fixed, as in the case of linear polynomials, or can be parameters to be estimated if a non-linear function is more appropriate. The model is very flexible since it allows the specification of different temporal dynamics for each item.

The common factors \( z' = (z_1, \ldots, z_q) \) account for the time-independent correlation between multiple responses, since all the temporal dependencies have already been accounted for by the multivariate growth model. The repeated measures related to the same common factor are constrained to having equal loadings. We refer to it as measurement invariance assumption. Indeed, assuming for simplicity \( q = 1 \), a model where the factor is equivalently measured at different time points (measurement invariance over time) is equivalent to a more parsimonious model where all the repeated measures load on a single common factor, and the loadings of those measures related to the same item are constrained to be equal. Hence, \( \lambda^i_j, j = 1, \ldots, J \), does not depend on \( t \).

By defining the linear predictor in this way, the temporal dependence between items as well as the autocorrelation of each item is explained by variance and covariance elements related to the random growth parameters \( \beta^j_j, j = 1, \ldots, J \). On the other hand, the cross-correlation between items despite their temporal behavior is caught by the factor model via the loadings \( \lambda^i_j, j = 1, \ldots, J \). These assumptions are contained in the prior density function of the latent variables, \( h(\eta) \), supposed to be a multivariate normal density with a
mean vector \( \mathbf{\mu}_n = (\mu_{\beta_0}, \ldots, \mu_{\beta_{p-1}}, \mu_{\beta_{p}}, \ldots, \mu_{\beta_{p}}, 0, \ldots, 0) \) and a covariance matrix \( \mathbf{\Psi} = \begin{pmatrix} \mathbf{\Psi}_\beta & 0 \\ 0 & \mathbf{\Psi}_z \end{pmatrix} \), where \( \mathbf{\Psi}_\beta \) is the full covariance submatrix related to the random effects \( \mathbf{\beta}_j, j = 1, \ldots, J \), and \( \mathbf{\Psi}_z \) is the submatrix related to the latent factors \( \mathbf{z} \). We assume that the random coefficients \( \mathbf{\beta}_j \) and the factors \( \mathbf{z} \) are independent. Furthermore, some constraints must be placed to ensure identifiability of the model based on the observed data. In particular, as in classical latent variable models, there is indeterminacy related to the scales of the latent factors \( \mathbf{z} \) (Jöreskog, 1969). This indeterminacy can be eliminated by either setting \( \lambda_{ij} = 1 \) or letting the variances of the factors be one, with at least one of the loadings constrained to be positive for each factor. As a consequence of the structural specification of the model via \( h(\mathbf{\eta}) \), the covariances between different responses, seen in the scale provided by the link function, are given by

\[
\begin{align*}
\text{Cov}(u_{ij}, u_{ij}) &= w_{ij\beta}^t \mathbf{\Psi}_\beta w_{ij\beta} + w_{ij}^t \mathbf{\Psi}_z w_{ij} \\
\text{Cov}(u_{ij}, u_{ij'}) &= w_{ij\beta}^t \mathbf{\Psi}_\beta w_{ij'\beta} + w_{ij}^t \mathbf{\Psi}_z w_{ij} \\
\text{Cov}(u_{ij}, v_{ij'}) &= w_{ij\beta}^t \mathbf{\Psi}_\beta w_{ij'\beta} + w_{ij}^t \mathbf{\Psi}_z w_{ij},
\end{align*}
\]

where \( w_{ij\beta} \) and \( w_{ij}, t = 1, 2, \ldots, T, j = 1, 2, \ldots, J \), indicate the coefficients in \( w_{ij} \) related to the random effects \( \mathbf{\beta}_j, j = 1, \ldots, J \), and the latent traits \( \mathbf{z} \), respectively.

The implied covariance matrix is decomposed into the sum of two components: one that accounts for the variability over time and the other one that accounts for the atemporal associations between items. This additive structure enables a longitudinal analysis of the data to be carried out, and at the same time, we can evaluate if the residual variability can be due to common factors underlying the items. Hence, even if these latter components are not present, a longitudinal analysis of the data can be performed by estimating a general multivariate latent growth model.

**Estimation**

The parameters of the model are estimated through the Maximum Likelihood (ML) method via the EM algorithm, since the latent variables \( \mathbf{\eta} \) are unobserved. The algorithm consists of an expectation and a maximization step. In the expectation step, the expected score function from the complete likelihood \( (\mathbf{y}, \mathbf{\eta}) \) is computed. In the maximization step, updated parameter estimates are obtained from the equations derived in the E-step. The whole procedure is repeated until convergence.

For a random sample of size \( n \), it follows by Equation 1 that the complete log likelihood is written as:
\[ L = \sum_{i=1}^{n} \log f(y_i; \eta_i) \]
\[ = \sum_{i=1}^{n} [\log g(y_i; \eta_i) + \log h(\eta_i)], \]  

(7)

where \( g \) is the likelihood of the data conditional on the latent variables and the random effects and \( h \) is the common distribution function of the latent traits and the random effects. From Equation 7, we see that the first component depends on both the factor loadings \( \lambda_{jz}; j = 1, \ldots, J \), and the variance parameters \( \sigma_j^2 = (\sigma_{1j}^2, \ldots, \sigma_{T_j}^2), j = J_1 + 1, \ldots, J \), whereas the second component depends on \( \mu_j \) and \( \Psi \).

The steps of the EM algorithm are defined as follows:

**Step 1:** Choose initial values for the model parameters. Starting values for the loadings are obtained by fitting separate confirmatory factor analysis models at each time point. Initial values for the other parameters are chosen arbitrarily.

**Step 2:** Compute the expected score functions for all the parameters (E-step).

**Step 3:** Obtain improved estimates for the parameters by solving the nonlinear ML equations for the parameters corresponding to the count items and explicit solutions for the parameters of the continuous items and the latent distribution (M-step).

**Step 4:** Repeat Steps 2 and 3 until convergence is attained.

Integrals are approximated using Gauss-Hermite quadrature points. More details on the EM algorithm are provided in the Appendix.

The standard errors of the parameter estimates are derived asymptotically by considering the diagonal elements of the inverse of the information matrix at the ML solution. Specifically, if \( \hat{c} \) is the vector of the unknown parameters, an approximation of the information matrix evaluated at the maximum likelihood \( \hat{c} \) is given by (Bartholomew & Knott, 1999):

\[
[I(\hat{c})]_{jk} = \sum_{i=1}^{n} \frac{1}{f^2(y_i)} \frac{\partial f(y_i)}{\partial c_j} \frac{\partial f(y_i)}{\partial c_k}.
\]

**Results**

We start the analysis by estimating a model for the overall data set of students observed at the three different time points. As already discussed, the aim of the analysis is twofold: (a) to analyze student performances over time with respect to the Number of Exams taken in each occasion (NE) and the corresponding Average Marks (AM) and (b) to measure a general individual capability. Therefore, we analyze how the variables NE and AM change over time by means of multivariate latent growth models, and we extend these
A General Multivariate Latent Growth Model

models by including a common factor that can explain the atemporal variability that exists between the two items. In particular, since only three different academic years are considered, a linear polynomial model ($p = 1$) could be appropriate to describe the temporal pattern of both NE and AM. Measurement invariance over time of the loadings in the one-factor model ($q = 1$) is assumed. Thus, the estimated model (denoted as Model A) is characterized by the following linear predictor

$$v_{i,j} = \beta_{0j} + \beta_{1j}(t - 1) + \lambda_j z = w_{i,j}'\eta, \quad t = 1, 2, 3, \quad j = \text{NE, AM}$$

where

$$w_{\text{NE}}' = [1, t - 1, 0, 0, \lambda_{\text{NE}}], \quad t = 1, 2, 3$$
$$w_{\text{AM}}' = [0, 0, 1, t - 1, \lambda_{\text{AM}}], \quad t = 1, 2, 3$$
$$\eta = [\beta_{0\text{NE}}, \beta_{1\text{NE}}, \beta_{0\text{AM}}, \beta_{1\text{AM}}, z]$$.

Binomial-logistic and Normal heteroscedastic linear regression models are estimated for the NE and AM, respectively. The multivariate normal density of the latent variables $h(\eta)$ has mean vector $\mu_\eta' = (\mu_{\beta_{0\text{NE}}}, \mu_{\beta_{1\text{NE}}}, \mu_{\beta_{0\text{AM}}}, \mu_{\beta_{1\text{AM}}}, 0)$ and covariance matrix $\Psi = \begin{pmatrix} \Psi_\beta & 0 \\ 0 & 1 \end{pmatrix}$. For identification reasons, the variance of the common latent factor $z$ is set equal to 1.

A path diagram of the model is illustrated in Figure 2.

FORTRAN and R codes have been implemented to estimate the model. (They are available from the authors upon request.) Parameter estimates of Model A for the overall data set are reported in Table 4. The significance of the parameters is evaluated using the Wald test.

In terms of the population mean trajectory, NE presents a mean initial status equal to 0.249, indicating that in the first year the students take, on average, around 4.5 exams, as expressed in the original scale. Student progress is described by the slope mean parameter $\mu_{\beta_{1}}$, equal to $-0.443$, reflecting the term-by-term worsening in the probabilities of success over time. However, the estimated number of exams are on average equal to 4.48, 4.98, and 5.04 at $t_1$, $t_2$, and $t_3$, respectively, coherently with the descriptive analyses reported in Table 2.

As for the AM variable, at the initial status, students obtain an average mark (in mean) around 23.98, but this mean worsens over time as indicated by the mean slope parameter $\mu_{\beta_{1\text{AM}}}$, equal to $-0.133$.

By looking at the covariances specific to each item in $\Psi_\beta$, students present a higher variability in the initial status than in the rate of growth, with a negative correlation between initial status and slope, for both NE and AM. Multivariate latent curves also allow analysis of the covariation between the temporal dynamics of NE and AM by estimating cross-covariances between random intercepts and slopes of the two curves. There are positive and significant
covariances between $\beta_{0\text{NE}}$ and $\beta_{0\text{AM}}$ as well as between the random slopes, $\beta_{1\text{NE}}$ and $\beta_{1\text{AM}}$, indicating that students with higher (lower) average marks in the first year tend to take a higher (lower) number of exams at $t_1$, and that students with positive (negative) slopes for AM generally present a similar pattern for NE. On the other hand, negative covariances are estimated between $\beta_{0\text{NE}}$ and $\beta_{1\text{AM}}$ as well as between $\beta_{0\text{AM}}$ and $\beta_{1\text{NE}}$. Different from classical multivariate latent growth modeling, these cross-covariances between NE and AM curves are free from the effect of a common latent factor $z$ we estimated via integrating the growth curves with a one-factor model.

When manifest variables are of different types, care is needed in the interpretation of the factor loadings, depending on the scale of the $y_{jt}$. In order to interpret the latent factor $z$, we shall therefore have to ensure that the $\lambda$'s are calibrated so that they may be meaningfully compared across variable types. This may be done in a variety of ways but we follow the approach of Takane and De Leeuw (1987) and Bartholomew and Knott (1999), which provides a parameterization that keeps the interpretation as close as possible to the familiar methods of traditional factor analysis. This approach is based on a standardization of the coefficients of the latent variable $z$.

For the normal item, the standardized $\lambda^*_{1\text{AM}}$ is given by
A General Multivariate Latent Growth Model

TABLE 4
Estimates for the Overall Data Set (Standard Errors in Brackets)

| Coefficients | Estimates          |
|--------------|--------------------|
| Multivariate growth model |                     |
| $\hat{\mu}_{\beta,NE}$ | 0.249 (0.047)       |
| $\hat{\mu}_{\beta,NE}$ | -0.443 (0.026)      |
| $\hat{\mu}_{\beta,AM}$ | 23.98 (0.326)       |
| $\hat{\mu}_{\beta,AM}$ | -0.113 (0.244)      |
| $\Psi_\beta$ | $\begin{pmatrix} 0.231^* \\ -0.107^* \\ 1.004^* \\ 0.363^* \end{pmatrix}$ |
| Factor model |                     |
| $\lambda_{NE}$ | 0.524 (0.073)       |
| $\lambda_{AM}$ | 2.581 (0.449)       |
| $\sigma^2_{\lambda,AM}$ | 10.080 (0.340)     |
| $\sigma^2_{\lambda,AM}$ | 9.953 (0.176)      |
| $\sigma^2_{\lambda,AM}$ | 14.568 (0.395)     |

where the denominator in Equation 8 is given by the square root of the variance of the continuous variable $y_{t,AM}$, $t = 1, 2, 3$. Hence

$$\lambda^*_{t,AM} = \lambda_{AM} \sqrt{\psi^2_{\lambda,AM} + 2(t-1)\psi_{\lambda,AM,\lambda,AM} + (t-1)^2\psi^2_{\lambda,AM} + \lambda^2_{AM} + \sigma^2_{\lambda,AM}}$$

$$t = 1, 2, 3.$$  

(8)

where the denominator in Equation 8 is given by the square root of the variance of the continuous variable $y_{t,AM}$, $t = 1, 2, 3$. Hence

$$\hat{\lambda}_{1AM}^* = 0.554$$
$$\hat{\lambda}_{2AM}^* = 0.586$$
$$\hat{\lambda}_{3AM}^* = 0.505.$$  

The amplitude of the factor loadings is quite similar on all the three occasions and on average is equal to 0.548, indicating that the measurement invariance assumption is appropriate.

On the other hand, for the binomial item, the standardization follows that proposed for binary items (Moustaki & Knott, 2001) and is based on the
equivalence of the response function and underlying variable approaches (Takane & De Leeuw, 1987). In this context, the standardized coefficients are given by

$$\lambda_{NE}^* = \frac{\lambda_{NE}}{\sqrt{\psi_{pNE}^2 + 2(t-1)\psi_{pNE,pNE} + (t-1)^2\psi_{pNE}^2 + \lambda_{NE}^2 + 1}}, \quad t = 1, 2, 3. \quad (9)$$

where the denominator in Equation 9 is the variance of the continuous variable underlying to the binomial item. The estimated standardized binomial loadings are

$$\hat{\lambda}_{1NE}^* = 0.427$$
$$\hat{\lambda}_{2NE}^* = 0.445$$
$$\hat{\lambda}_{3NE}^* = 0.435,$$

which are quite close to each other; the amplitude of the standardized factor loadings is on average 0.436. The standardized coefficients given for normal and binomial variables can be used for a unified interpretation of the loadings, bringing the interpretation close to factor analysis. The common factor $z$ explains the interrelationships between the two observed items net from their temporal dependence. Both variables are significant indicators of this latent construct (Wald based $p$ values are smaller than .01 in both cases) and influence it positively. This means that the more talented students are, that is they present high values of the common factor, the more they tend to pass exams and receive higher marks. Ignoring the presence of a common factor in multivariate latent growth models can lead to an overestimation of the cross-variation among multiple curves.

The goodness of fit of Model A has been checked separately for the count and continuous part (Moustaki & Knott, 2001). As for the count part of the model, significant information concerning the goodness of fit can be found in the margins. In particular, the one-way margins of the differences between the observed ($O$) and expected ($E$) frequencies under the model are investigated; any large discrepancies will suggest that the model does not fit these counts well. The chi-square test as well as high-way margins are not appropriate because of the sparseness of the data (Reiser, 1996). Table 5 gives the GF-fit measures, calculated as $((O - E)^2/E)$, for each Binomial variable (Bartholomew, Steele, Moustaki, & Galbraith, 2002). We can observe that the GF-fits are not good, especially those on count 7 for $y_{1NE}$, on 8 and 9 for $y_{2NE}$, and on counts 6 and 7 for $y_{3NE}$. Reasons for this misfitting of Model A on the overall sample will be investigated next.

For the normal part of the model, we check the discrepancies between the sample correlation matrix and the one estimated from the model, as illustrated in Table 6 for the variables $y_{1AM}$, $y_{2AM}$, and $y_{3AM}$. The discrepancies between observed correlations and those estimated are particularly small, indicating that the fit of the model for the normal variables is good.
In the Data section, we showed that from the 821 students, two different temporal patterns were evident, one related to those students who graduated at \(t_3\) (Grad) and the other to students who have not yet completed their degree by the end of the third year (Nograd). Hence, we shall analyze these two different groups of students in order to investigate the reasons of the poorness of fit for the count part of the Model A in the overall data set. Therefore, in the following, we fit Model A to the Grad and Nograd students, separately.

The Graduate Students

We first consider the 195 students who graduated at \(t_3\), and we start by fitting Model A described above. The results of the estimation are reported in Table 7. It can be noticed that parameter estimates corresponding to both the multivariate latent growth and the factor parts of Model A differ substantially from what

| Counts | \(\gamma_{1NE}\) | \(\gamma_{2NE}\) | \(\gamma_{3NE}\) |
|--------|----------------|----------------|----------------|
| 0      | 11.58          | 11.65          | 11.23          |
| 1      | 0.10           | 1.59           | 3.05           |
| 2      | 5.78           | 0.22           | 10.39          |
| 3      | 5.45           | 0.39           | 15.20          |
| 4      | 8.73           | 4.97           | 5.91           |
| 5      | 1.10           | 16.18          | 0.95           |
| 6      | 7.58           | 5.12           | 34.35          |
| 7      | 24.47          | 0.33           | 37.70          |
| 8      | 3.65           | 23.75          | 2.33           |
| 9      | –              | 31.37          | 0.04           |
| 10     | –              | 13.06          | 0.66           |
| 11     | –              | 0.31           | 7.94           |
| 12     | –              | 0.02           | 0              |
| 13     | –              | –              | 0.41           |
| 14     | –              | –              | 1.23           |

### Table 5

**Count Items: GF-Fit Values for the One-Way Margins**

### Table 6

**Normal Items: Discrepancies Between Sample and Estimated Correlation Matrices**

| \(\gamma_{1AM}\) | \(\gamma_{2AM}\) | \(\gamma_{3AM}\) |
|----------------|----------------|----------------|
| \(\gamma_{1AM}\) | 0.00           | –0.05          | 0.02           |
| \(\gamma_{2AM}\) | –0.05          | 0.00           | –0.05          |
| \(\gamma_{3AM}\) | 0.02           | –0.05          | 0.00           |
### TABLE 7
Estimates for Graduate (Grad) Students

| Model B | Coefficients | Estimates |
|---------|--------------|-----------|
| **Multivariate growth model** | | |
| $\mu_0^{\text{NE}}$ | 1.133 (0.086) |
| $\mu_1^{\text{NE}}$ | -0.577 (0.050) |
| $\mu_0^{\text{AM}}$ | 26.172 (0.149) |
| $\mu_1^{\text{AM}}$ | 0.080 (0.061) |
| $\Psi_\beta$ | $\begin{pmatrix} 0.133^* \\ -0.097^* \\ 0.441 \\ 0.044 \end{pmatrix}$ |
| | $\begin{pmatrix} 0.072^* \\ -0.327 \\ 2.197 \\ -0.026 \end{pmatrix}$ |
| | $\begin{pmatrix} 0.376 \\ -0.086 \end{pmatrix}$ |
| **Factor model** | | |
| $\lambda_{\text{NE}}$ | 0.005 (0.262) |
| $\lambda_{\text{AM}}$ | 0.670 (3.360) |
| $\sigma_{1\text{AM}}^2$ | 0.702 (0.147) |
| $\sigma_{2\text{AM}}^2$ | 1.292 (0.113) |
| $\sigma_{3\text{AM}}^2$ | 0.328 (0.185) |
| AIC = 4,304.090 |
| BIC = 4,309.602 |

---

**Model B**

| Coefficients | Estimates |
|--------------|-----------|
| **Multivariate growth model** | | |
| $\hat{\mu}_0^{\text{NE}}$ | 1.193 (0.093) |
| $\hat{\mu}_1^{\text{NE}}$ | -0.594 (0.066) |
| $\hat{\mu}_0^{\text{AM}}$ | 26.225 (0.192) |
| $\hat{\mu}_1^{\text{AM}}$ | 0.188 (0.079) |
| $\Psi_\beta$ | $\begin{pmatrix} 0.136^* \\ -0.103^* \\ 0.445 \\ 0.025 \end{pmatrix}$ |
| | $\begin{pmatrix} 0.078^* \\ -0.340 \\ 2.727 \\ -0.021 \end{pmatrix}$ |
| | $\begin{pmatrix} 0.317 \end{pmatrix}$ |
| **Factor model** | | |
| $\hat{\lambda}_{1\text{NE}}$ | 0.300 (0.413) |
| $\hat{\lambda}_{2\text{NE}}$ | 0.465 (0.594) |

(continued)
we obtained for the overall sample. As for the former, both NE and AM present a higher mean initial status than the overall sample. They are equal to 6.05 for NE and to 26.17 for AM, as expressed in the original scales. Furthermore, the mean trajectory for NE has a worsening pattern over time in terms of probability of success, but a different pattern in terms of the average number of exams. An increasing not significant mean trajectory corresponds to AM. By looking at the variability around the mean trajectories, this is significant in the initial status (0.133) and in the rate of growth (0.072) corresponding to NE, which also shows a negative correlation between the random intercept and slope. On the other hand, there is not a significant variability for AM with respect to its mean trajectory, indicating that the Grad students show a similar pattern over time. This finding is also evident in all the (not significant) covariances between random coefficients of NE and AM.

As for the factor part, the loading associated to the binomial variable is very close to 0 and not significant, suggesting that the number of exams for these students is not a measure of the latent variable. However, also the loading associated to AM is not significant. This means that for Grad students it makes no sense to specify a common factor related to the variables AM and NE. A justification could be found from the not significant association between AM and NE for Grad in time $t_2$, as shown in the Data section. This evidence can be a hint to test the assumption of measurement invariance. That is, we can evaluate if repeated observations of the same item equivalently measure the latent construct. If we estimate a model where this assumption is relaxed, denote it as Model B (Table 7), we can notice that the time-dependent loadings

\begin{table}[h]
\centering
\begin{tabular}{ll}
\hline
\textbf{Coefficients} & \textbf{Estimates} \\
\hline
$\hat{\lambda}_{3\text{NE}}$ & $-0.103 (0.185)$ \\
$\hat{\lambda}_{1\text{AM}}$ & $0.156 (0.339)$ \\
$\hat{\lambda}_{2\text{AM}}$ & $0.440 (0.571)$ \\
$\hat{\lambda}_{3\text{AM}}$ & $0.734 (0.886)$ \\
$\hat{\sigma}_{1\text{AM}}^2$ & $0.616 (0.129)$ \\
$\hat{\sigma}_{2\text{AM}}^2$ & $1.300 (0.114)$ \\
$\hat{\sigma}_{3\text{AM}}^2$ & $0.351 (0.177)$ \\
AIC & 4,226.847 \\
BIC & 4,273.520 \\
\hline
\end{tabular}
\caption{Model B (continued)}
\end{table}

Note: AIC = Akaike information criterion; BIC = Bayesian information criterion.
*Significant at 5% level according to the Wald based $p$ values.
are very different. This is particularly true for the variable NE, for which the loading does not change greatly in the first two time points but becomes negative in $t_3$. For the variable AM the loading increases over time. Clearly, the measurement invariance cannot be assumed. This result is also confirmed by the Akaike information criterion (AIC) and Bayesian information criterion (BIC), which show that Model B is better than Model A. However, from our viewpoint, Model B is meaningless in the factor part. Moreover, if we look at the one-way margins associated to Model B (Table 8), we can see that again there are goodness-of-fit problems at time points $t_2$ and $t_3$.

These results for Grad highlight two different aspects of the analysis. First of all, we have a slight individual variability around NE and AM mean trajectories, indicating a similar temporal behavior of these students. Moreover, the higher values of these means compared to those obtained from the overall sample show a good performance for this group. Second, in this case the two variables are not measures of a latent construct.

**The Nograd Students**

Coherent with the previous analysis, we then estimated Model A for Nograd students. The results are reported in Table 9. The growth model shows results similar to the overall sample with a higher variability for AM with respect to its mean trajectory. In this case, the results related to the factor model are very interesting. Different from what we found for the Grad group, the loadings are both significant and positively related to the latent variable and indicate that, as in the overall data set, the factor model is appropriate. Different from what we found in all previous analyses, the

---

**TABLE 8**

*Count Items: GF-Fit Values for the One-Way Margins, Model B.*

| Counts | $y_{1\text{NE}}$ | $y_{2\text{NE}}$ | $y_{3\text{NE}}$ |
|--------|-----------------|-----------------|-----------------|
| 2      | 0.34            | 0.05            | –               |
| 3      | 4.65            | –               | –               |
| 4      | 7.46            | 0.03            | –               |
| 5      | 4.20            | 16.28           | 18.16           |
| 6      | 1.37            | 58.32           | 11.10           |
| 7      | 5.29            | 13.48           | 19.52           |
| 8      | 0.86            | 15.12           | 0.00            |
| 9      | –               | 23.81           | 0.40            |
| 10     | –               | 14.41           | 0.50            |
| 11     | –               | 0.02            | 2.35            |
| 12     | –               | 1.38            | –               |
| 13     | –               | –               | –               |
| 14     | –               | –               | 30.51           |
### TABLE 9
Estimates for Undergraduate (Nograd) Students

**Model C**

| Coefficients | Estimates |
|--------------|-----------|
| **Multivariate growth model** | |
| $\mu_{\beta_0}^{2}$ | $0.30$ (0.055) |
| $\mu_{\beta_1}^{2}$ | $0.353$ (0.035) |
| $\mu_{\beta_2}^{2}$ | $23.337$ (0.431) |
| $\mu_{\beta_3}^{2}$ | $-0.183$ (0.326) |
| $\Psi_\beta$ | $\begin{pmatrix} 0.171^* \\ -0.109^* & 0.131^* \\ 0.930^* & -0.610^* & 5.514^* \\ -0.552^* & 0.658^* & -3.304^* & 3.492^* \end{pmatrix}$ |
| **Factor model** | |
| $\lambda_{\text{NE}}$ | $0.473$ (0.119) |
| $\lambda_{\text{AM}}$ | $2.609$ (0.706) |
| $\sigma_{1\text{AM}}^2$ | $12.356$ (0.567) |
| $\sigma_{2\text{AM}}^2$ | $12.484$ (0.269) |
| $\sigma_{3\text{AM}}^2$ | $17.307$ (0.612) |
| AIC | $18,773.348$ |
| BIC | $18,788.483$ |

(continued)
GF-fits of this model are satisfactory for all the observed time points, as reported in Table 10. Thus the count part of the model is well fitted by the binomial distribution. Also the fit is very good for the normal part, the discrepancies between observed and estimated correlations being very low (Table 11). Therefore, Model A fits the Nograd students’ data well.

If we look again at Table 9 Model A, it can be noticed that the values of $\sigma^2_{AM}$ are quite similar over time, thus, it can be interesting to evaluate if AM is homoscedastic over time. The results of the estimation of the model with homoscedastic errors (Model C) are reported in the bottom part of Table 9. Although all the parameter estimates do not change abruptly, the AIC and BIC are slightly better for Model A than for Model C, suggesting that such an assumption does not hold. Thus, the comparison between the loadings of AM and NE for evaluating the influence of each item on the latent variable requires their standardization.

Table 9 (continued)

| Coefficients | Estimates     |
|--------------|--------------|
| $\sigma^2_{AM}$ | 13.642 (0.234) |
|              |              |
| AIC          | 18777.719    |
| BIC          | 18791.260    |

*Significant at 5% level according to the Wald based $p$ values.

Table 10
Count Items: GF-Fit Values for the One-Way Margins, Model A, Nograd.

| Counts | $y^{1\text{NE}}$ | $y^{2\text{NE}}$ | $y^{3\text{NE}}$ |
|--------|-----------------|-----------------|-----------------|
| 0      | 7.69            | 11.56           | 5.01            |
| 1      | 0.86            | 1.67            | 4.91            |
| 2      | 7.91            | 0.18            | 9.11            |
| 3      | 3.48            | 1.33            | 8.28            |
| 4      | 2.16            | 7.46            | 0.17            |
| 5      | 1.01            | 12.83           | 2.96            |
| 6      | 9.86            | 0.33            | 20.67           |
| 7      | 12.61           | 7.89            | 8.47            |
| 8      | 0.27            | 10.92           | 0.41            |
| 9      |                | 14.13           | 0.27            |
| 10     |                | 5.54            | 3.57            |
| 11     |                |                | 3.21            |
| 12     |                | 0.35            |                |
| 13     |                |                | 1.53            |

GF-fits of this model are satisfactory for all the observed time points, as reported in Table 10. Thus the count part of the model is well fitted by the binomial distribution. Also the fit is very good for the normal part, the discrepancies between observed and estimated correlations being very low (Table 11). Therefore, Model A fits the Nograd students’ data well.
according to Equation 8 for AM and to Equation 9 for NE. We get the following standardized loadings: $\lambda_{AM}^* = (0.525, 0.560, 0.534)$, and $\lambda_{NE}^* = (0.401, 0.414, 0.388)$. As it is in the overall data set, the correlation between $z$ and AM is slightly higher than that between $z$ and NE.

**Discussion**

In this article, we extended and applied multivariate latent growth models to the analysis of student record data collected repeatedly in the DWH system of the University of Bologna. The proposed approach is innovative since it allows the simultaneous evaluation of both student performance over time and specific common factors. Key features include (a) a flexible modeling of the temporal dynamics of the observed variables via specific latent curves and (b) an extension of the multivariate growth model that incorporates a factor part. Such a component explains the association between the observed items by means of latent variables.

The complexity of the proposed model lies in different aspects, such as the presence of mixed data and the possibility of both including several latent variables/random effects and estimating specific temporal patterns for the observed variables. Commercial software can be used to simultaneously fit multivariate latent growth models and common factors. However, problems occur in parameter estimation, mainly due to the treatment of binomial data. In this respect, we implemented an EM algorithm in Fortran and R. The computational complexity of our code is related to the numerical multidimensional integrations on the latent variable space. We can estimate a model with a maximum of six latent variables (random effects plus common factors) on an average PC (AMD Athlon (tm) 64 x 2 Dual Core Processor 3.21 Ghz with 1.96 Gb RAM).

We demonstrated, via different specifications of the model, how our general approach can provide insights into the data structure. In particular, the analysis carried out on a cohort of students enrolled at the Faculty of Economics observed at three different time points highlighted a heterogeneity in the overall data set in terms of both average marks and number of
exams. This is due to the presence of different temporal patterns within the cohort, since we have students who regularly graduate at $t_3$ (Grad) and students who have not yet completed their degree by the end of the third year (Nograd). Grad students perform very well in terms of both number of exams and average marks with similar temporal pattern. Nograd students take a lower number of exams with lower average marks, but within this group we have a significant variability both in the initial status and in the rate of growth. The factor part of the model for Grad students fails in measuring a general factor by means of the observed indicators considered. We found that this fact depends on the fundamental assumption of measurement invariance of items over time. Such an assumption does not hold in this case. However, the model fits the data of Nograd students well. What we called the atemporal general factor is well measured by the average mark and the number of exams taken, both being significantly related to the latent variable. The good performance of the model is confirmed by the analysis of various goodness-of-fit statistics.

The heterogeneity observed in the patterns of graduating and nongraduating students has implications on the results of the overall data set. Parameter estimates for the overall sample are quite similar to those of Nograd students, the factor loadings of the two variables both being significant. This is in part due to the larger sample size of Nograd students. However, the presence of different performances of the Grad students reflects on the poorness of fit of the count variable in the overall data set.

The model proposed here was motivated by the study of student achievement and the good results obtained clearly show its appropriateness. However, such methodology can be applied successfully in many other fields, such as socioeconomic settings in which personal behavior is studied using panel data collected through the administration of questionnaires.

In the present example, no covariates have been considered. In practice, we may have useful time-dependent and time-independent covariates such as gender, region of origin, age, and so on, that can be incorporated into the model. In particular, an emerging field of investigation is based on the comparison of the performances of students who completed the degree compared with those who abandoned it (Draper & Gittoes, 2004; Smith & Naylor, 2001). Furthermore, the analysis can be extended to more time points and it can be evaluated whether nonlinear or higher degree polynomial trajectories can describe the temporal behavior of the items. Preliminary studies performed in this direction with the software LISREL (Bianconcini et al., 2007) showed how different latent curves can fit the weighted average marks for different groups of students within the cohort analyzed by performing a multigroup analysis. The multigroup analysis can represent a useful extension of our model since it enables comparison of groups in terms of sequential equality constraints on the structure and on the parameters involved in
the estimation. Such problems are motivation future investigation along these lines of research.

Appendix

Starting from the complete log likelihood given in Equation 7, we derive the ML estimates of the model parameters. It should be noted that the first component of the log likelihood depends on both the factor loadings \( \lambda_{jz} \), \( j = 1, \ldots, J \) and the variance parameters \( \sigma^2_j = (\sigma^2_{1j}, \ldots, \sigma^2_{Tj}), j = J_1 + 1, \ldots, J \), whereas the second component depends on \( \mu_{\eta} \) and \( \Psi \).

Estimation of \( \mu_{\eta} \) and \( \Psi \)

From the normality of \( b \), the second component of the log likelihood given in Equation 7 (up to a constant) for an individual \( i \) is written as

\[
\log h(\eta_i) = -\frac{1}{2} \ln \Psi - \frac{1}{2} (\eta_i - \mu_{\eta}) \Psi^{-1} (\eta_i - \mu_{\eta})',
\]  

(A1)

The expected score function needed for the EM implementation is taken with respect to the posterior distribution of the latent variables \( h(\eta_i|y_i) \). The expected score function for the parameter vector \( \mu_{\eta} \) becomes

\[
\text{ES}_i(\mu_{\eta}) = \int S_i(\mu_{\eta}) h(\eta_i|y_i) d\eta_i, \quad i = 1, \ldots, n,
\]

(A2)

where

\[
S_i(\mu_{\eta}) = \frac{\partial \log h(\eta_i)}{\partial \mu_{\eta}} = \Psi^{-1} (\eta_i - \mu_{\eta}).
\]

Similarly, we obtain the score function for \( \Psi \), that is,

\[
S_i(\Psi) = \frac{\partial \log h(\eta_i)}{\partial \Psi} = -\frac{1}{2} \Psi^{-1} - \frac{1}{2} \Psi^{-1} (\eta_i - \mu_{\eta})(\eta_i - \mu_{\eta})' \Psi^{-1}.
\]

By solving \( \sum_{i=1}^n \text{ES}_i(\mu_{\eta}) = 0 \) and \( \sum_{i=1}^n \text{ES}_i(\Psi) = 0 \) we derive explicit solutions for the ML estimators of \( \mu_{\eta} \) and \( \Psi \).

Estimation of \( \lambda_{jz} \) and \( \sigma^2_j \)

The estimation of parameters \( \lambda_{jz}, j = 1, \ldots, J \), and \( \sigma^2_j, j = J_1 + 1, \ldots, J \), depends on the first component of the log likelihood given in Equation 7. Under the conditional independence assumption, the log likelihood of the count and continuous data can be written as

(continued)
\[
\log g(y_i | \eta_i) = \sum_{j=1}^{J_1} \sum_{t=1}^{T} \left[ \log \left( \frac{n_t}{y_{ji}} \right) + y_{ji} v_{ji} - n_t \log (1 + \exp(v_{ji})) \right] + \\
+ \sum_{j=J_1+1}^{J} \sum_{t=1}^{T} \left[ -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma_{y_j}^2) - \frac{(y_{ji} - v_{ji})^2}{2\sigma_{y_j}^2} \right].
\]

(A3)

The first component refers to count variables and will be used to derive estimates of the factor loadings corresponding to such variables, that is, \( \lambda_{jz}, j = 1, \ldots, J_1 \). The expected score function of the parameter vector \( \lambda_{jz} \) is again taken with respect to the posterior \( h(\eta_i | y_i) \):

\[
ES_i(\lambda_{jz}) = \int S_i(\lambda_{jz}) h(\eta_i | y_i) d\eta_i, \quad i = 1, \ldots, n,
\]

(A4)

where

\[
S_i(\lambda_{jz}) = \frac{\partial \log g(y_{ji} | \eta_i)}{\partial \lambda_{jz}},
\]

and

\[
\frac{\partial \log g(y_{ji} | \eta_i)}{\partial \lambda_{jz}} = \sum_{t=1}^{T} z_i \left( y_{ji} - n_t \frac{\exp(v_{ji})}{(1 + \exp(v_{ji}))} \right), \quad j = 1, \ldots, J_1.
\]

(A5)

By replacing Equation A5 with Equation A4 and solving \( \sum_{i=1}^{n} ES_i(\lambda_{jz}) = 0 \) we derive nonexplicit solutions for the parameter vector \( \lambda_{jz} \). A Newton–Raphson algorithm is used to solve the nonlinear ML equations.

From the second component in the likelihood (Equation A3), we estimate factor loadings and variance components corresponding to continuous items. The expected score functions for the parameters \( \lambda_{jz}, j = J_1 + 1, \ldots, J \), are given by

\[
ES_i(\lambda_{jz}) = \int S_i(\lambda_{jz}) h(\eta_i | y_i) d\eta_i, \quad i = 1, \ldots, n,
\]

(A6)

where

\[
S_i(\lambda_{jz}) = \sum_{t=1}^{T} z_i^2 \lambda_{jzi} - \sum_{i=1}^{P} z_i (y_{ji} - \sum_{r=0}^{P} \beta_{ij} \lambda_{r}^r).
\]

Similarly, we obtain the score function for each element in \( \sigma_{y_j}^2, j = J_1 + 1, \ldots, J \), that is,

\[
S_i(\sigma_{y_j}^2) = \sigma_{y_j}^2 - (y_{ji} - v_{ji})^2 \quad t = 1, \ldots, T, \quad i = 1, \ldots, n.
\]

(continued)
By solving $\sum_{i=1}^{n} ES_i(\lambda_{jz}) = 0, \sum_{i=1}^{n} ES_i(\sigma_j^2) = 0$ we get explicit solutions for the ML estimators of $\lambda_{jz}$ and $\sigma_j^2$ for $j = J_1 + 1, \ldots, J$.

Integrals are approximated using Gauss-Hermite quadrature points. In order to apply the Gauss-Hermite approximation to the integral of Equations A2, A4, and A6 we consider the Cholesky decomposition of the covariance matrix $\Psi$ given by $\Psi = CC'$. As shown by Cagnone et al. (2009), this is necessary because the non-null submatrices of $\Psi$, namely, $\Psi_\beta$ and $\Psi_z$, are not diagonal.

Acknowledgments

The authors thank the Editor, two anonymous referees, and Professor Paola Monari for insightful comments that have led to a significant improvement of the article. This study has been supported by grants from the Italian Ministry of University and Research (MIUR): PRIN2008 “Latent structure analysis: New boundaries in statistical methods and models.”

References

Bartholomew, D. J., & Knott, M. (1999). Latent variable models and factor analysis. London: Hodder Arnold.

Bartholomew, D. J., Steele, F., Moustaki, I., & Galbraith, J. (2002). Analysis and interpretation of multivariate data for social scientists. Boca Raton, FL: Chapman and Hall/CRC.

Bianconcini, S., Cagnone, S., Mignani, S., & Monari, P. (2007). A latent curve analysis of unobserved heterogeneity in university student achievements. Statistica, 1, 40–56.

Bollen, K. A., & Curran, P. J. (2006). Latent curve models: A structural equation perspective. New York, NY: John Wiley.

Cagnone, S., Moustaki, I., & Vasdekis, V. (2009). Latent variable models for multivariate longitudinal ordinal responses. British Journal of Mathematical and Statistical Psychology, 62, 401–415.

Draper, D., & Gittoes, M. (2004). Statistical analysis of performance indicators in UK higher education. Journal of the Royal Statistical Society: Series A (Statistics in Society), 167, 449–474.

Dunson, D. B. (2003). Dynamic latent trait models for multidimensional longitudinal data. Journal of American Statistical Association, 98, 555–563.

Jöreskog, K. (1969). A general approach to confirmatory maximum likelihood factor analysis. Psychometrika, 34, 183–202.

Jöreskog, K., & Sörbom, D. (2001). LISREL 8: Users reference guide (2nd ed.). Chicago, IL: SSI.

McArdle, J. (1988). Dynamic but structural equation modeling of repeated measures data. In R. B. Cattell & J. Nesselroade (Eds.), Handbook of multivariate experimental psychology (pp. 561–614). New York, NY: Plenum.

Moustaki, I., & Knott, M. (2001). Generalized latent trait models. Psychometrika, 65, 391–411.
Bianconcini and Cagnone

Muthén, B., & Khoo, S. T. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences, 10*, 73–101.
Reiser, M. (1996). Analysis of residual for the multinomial item response model. *Psychometrika, 61*, 509–528.
Raykov, T. (2007). Longitudinal analysis with regressions among random effects: A latent variable modeling approach. *Structural Equation Modeling: A Multidisciplinary Journal, 14*, 146–169.
Roy, J., & Lin, X. (2000). Latent variable models for longitudinal data with multiple continuous outcomes. *Biometrics, 56*, 1047–1054.
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.
Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Boca Raton, FL: Chapman and Hall/CRC.
Smith, J. P., & Naylor, R. A. (2001). Dropping out of university: A statistical analysis of the probability of withdrawal for UK university students. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 164*, 389–405.
Takane, Y., & De Leeuw, J. (1987). On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika, 52*, 393–408.

Authors

SILVIA BIANCONCINI is an Assistant Professor at the Department of Statistics, University of Bologna, Italy; e-mail: silvia.bianconcini@unibo.it. Her research focuses on the nonparametric estimation of unobserved time series components, and latent variable models for mixed and longitudinal data.

SILVIA CAGNONE is an Assistant Professor at the Department of Statistics, University of Bologna, Italy; e-mail: silvia.cagnone@unibo.it. Her research interests include latent variable models for categorical, ordinal and longitudinal data, and structural equation models.

Manuscript received April 9, 2009
Revision received September 7, 2010
Accepted November 6, 2010