Bayesian Approach to Non-ignorable Missingness in Latent Growth Models

Zhenqiu (Laura) Lu[1][0000−0001−9482−1368] and Zhiyong Zhang[2][0000−0003−0590−2196]

[1] University of Georgia, Athens, GA 30602, USA
`zlu@uga.edu`
[2] University of Notre Dame, Notre Dame, IN 46530, USA
`zhiyongzhang@nd.edu`

Abstract. Latent growth curve models (LGCMs) are becoming increasingly important among growth models because they can effectively capture individuals' latent growth trajectories and also explain the factors that influence such growth by analyzing the repeatedly measured manifest variables. However, with the increase in complexity of LGCMs, there is an increase in issues on model estimation. This research proposes a Bayesian approach to LGCMs to address the perennial problem of almost all longitudinal research, namely, missing data. First, different missingness models are formulated. We focus on non-ignorable missingness in this article. Specifically, these models include the latent intercept dependent missingness, the latent slope dependent missingness, and the potential outcome dependent missingness. To implement the model estimation, this study proposes a full Bayesian approach through data augmentation algorithm and Gibbs sampling procedure. Simulation studies are conducted and results show that the proposed method accurately recover model parameters and the mis-specified missingness may result in severely misleading conclusions. Finally, the implications of the approach and future research directions are discussed.

Keywords: Bayesian Estimation · Missing Data · Latent Growth Curve Models · Non-ignorable Missingness · Longitudinal Analysis · Multilevel Modeling

1 Introduction

In social and behavioral sciences, there has been great interest in the analysis of change (e.g., Collins, 1991; Lu, Zhang, & Lubke, 2010; Singer & Willett, 2003). Growth modeling is designed to provide direct information of growth by measuring the variables of interest on the same participants repeatedly through time (e.g., Demidenko, 2004; Fitzmaurice, Davidian, Verbeke, & Molenberghs, 2008; Fitzmaurice, Laird, & Ware, 2004; Hedeker & Gibbons, 2006; Singer

& Willett, 2003). Among the most popular growth models, *latent growth curve models (LGCMs)* are becoming increasingly important because they can effectively capture individuals' latent growth trajectories and also explain the latent factors that influence such growth by analyzing the repeatedly measured manifest variables (e.g., Baltes & Nesselroade, 1979). Manifest variables are evident in the data, such as observed scores; latent variables cannot be measured directly and are essentially hidden in the data, such as the latent initial levels and latent growth rates (Singer & Willett, 2003). We use the term "latent" because these variables are not directly observable but rather are assumed to be inferred, although they may be closely related to observed scores. For example, the latent intercept (i.e., the latent initial level) may be related to the test score at the first occasion, the prior knowledge of a course (such as mathematical knowledge), or other similar variables. The latent slope (i.e., the latent growth rate) may be related to the participant's learning ability, the attitude toward the course, the instructor's teaching methods, or other similar types of variables.

However, with an increase in complexity of LGCMs, comes an increase in difficulties estimating such models. First, missing data are almost inevitable with longitudinal data (e.g., Jelicic, Phelps, & Lerner, 2009; Little & Rubin, 2002). Second, conventional likelihood estimation procedures might fail for complex models with complicated data structures.

1.1   Missing Data

As LGCMs involve data collection on the same participants through multiple waves of surveys, tests, or questionnaires, missing data are almost inevitable. Research participants may drop out of a study, or some students may miss a test due to absence or fatigue (e.g., Little & Rubin, 2002; Schafer, 1997). Missing data can be investigated from their mechanisms, or why missing data occur. Little and Rubin (2002) distinguished *ignorable missingness mechanism* and *non-ignorable missingness mechanism*. For ignorable missingness mechanism, estimates are usually asymptotically consistent when the missingness is ignored (Little & Rubin, 2002), because the parameters that govern the missing process either are distinct from the parameters that govern the model outcomes or depend on the observed variables in the model. The non-ignorable missingness mechanism is also referred to as *missing not at random (MNAR)*, in which the missing data probability depends either on unobserved outcomes, or on latent variables that cannot be fully measured by the observed data, in other words, latent variables that depend on the missing values.

With the appearance of missing data comes the challenge in estimating growth model parameters. To address the challenge, statisticians have developed different approaches and models. Although there are a large amount of literature to address this problem in applied behavioral sciences (e.g., Acock, 2005; Schafer & Graham, 2002; Schlomer, Bauman, & Card, 2010), especially in longitudinal studies (e.g., Jelicic et al., 2009; Roth, 1994), the majority of the literature is on ignorable missingness. This is mainly because (1) analysis models or techniques for non-ignorable missing data are traditionally difficult to implement

and not yet easy to use (e.g., Baraldi & Enders, 2010), and (2) missingness mechanisms are not testable (e.g., Little & Rubin, 2002). At the same time, however, non-ignorable missingness analysis is a crucial and a serious concern in applied research areas, in which participants may be dropping out for reasons directly related to the response being measured (e.g., Baraldi & Enders, 2010; Enders, 2011b; Hedeker & Gibbons, 1997). Not attending to the non-ignorable missingness may result in severely biased statistical estimates, standard errors, and associated confidence intervals, and thus poses substantial risk of leading researchers to incorrect conclusions (e.g., Little & Rubin, 2002; Schafer, 1997; Zhang & Wang, 2012).

In a study of latent growth models, Lu, Zhang, and Lubke (2011) investigated non-ignorable missingness in mixture models. However, the missingness in that study was only allowed to depend on latent class membership. In practice, even within one population, the missingness may depend on many other latent variables, such as latent initial levels and latent growth rates. When observed data are not completely informative about these latent variables, the missingness is non-ignorable. Furthermore, Lu et al. (2011) did not examine how to identify the missingness mechanisms. Accordingly, this study extends previous research to more general non-ignorable missingness and also investigates the influences of different types of non-ignorable missingness on model estimation.

1.2   Bayesian Approach

To implement the model estimation, we propose a full Bayesian approach. Traditionally, maximum likelihood methods have been widely used in most studies for estimating parameters of models in the presence of missing data (e.g., Enders, 2011a; Muthén, Asparouhov, Hunter, & Leuchter, 2011), and statistical inferences have been carried out using conventional likelihood procedures (e.g., Yuan & Lu, 2008). Recently, multiple imputation (MI) methods have been proposed as an alternative approach (e.g., Enders, 2011a; Muthén et al., 2011). MI is a Monte Carlo technique that replaces the missing values with multiple simulated values to generate multiple complete datasets. Each of these simulated datasets is then analyzed using methods that do not account for missingness, that is, using standard analytical methods. Results are then combined to produce estimates and confidence intervals that incorporate the uncertainty due to the missing-data (Enders, 2011a; Rubin, 1987; Schafer, 1997). Both ML and MI estimation methods typically assume that missing data mechanisms are MCAR or MAR. Further, using conventional estimation procedures may fail or may provide biased estimates (Yuan & Zhang, 2012) when estimating model parameters in complex models with complicated data structures such as GMMs with missing data and outliers. In addition, MI requires data to be imputed under a particular model (e.g., Allison, 2002; Newman, 2003). And literature also shows that multiple imputation is inappropriate as a general purpose methodology for complex problems or large datasets (e.g., Fay, 1992). When missingness is MNAR, most work augments the basic analysis using with a model that explains the probability of missing data (e.g., Enders, 2011a; Muthén et al., 2011).

In this article, a full Bayesian estimate approach (e.g., Lee, 2007; Muthén & Asparouhov, 2012) is proposed. There are several advantages. First, this approach involves Gibbs sampling methods (Geman & Geman, 1984). Gibbs sampling is especially useful when the joint distribution is complex or unknown but the conditional distribution of each variable is available. The sequence of samples constructs a Markov chain that can be shown to be ergodic (Geman & Geman, 1984). That is, once convergence is obtained, the samples can be assumed to be independent draws from the stationary distribution. Thus, after convergence the generated value is actually from the joint distribution of all parameters. Each variable from the Markov chain has also been shown to converge to the marginal distribution of that variable (Robert & Casella, 2004). Additional advantages of Bayesian methods include their intuitive interpretations of statistical results, their flexibility in incorporating prior information about how data behave in similar contexts and findings from experimental research, their capacity for dealing with small sample sizes (such as occur with special populations), and their flexibility in the analysis of complex statistical models with complicated data structure (e.g., Dunson, 2000; Scheines, Hoijtink, & Boomsma, 1999).

1.3   Goals and Structure

The goals of the paper are to propose latent growth curve models with non-ignorable missingness and to evaluate the performance of Bayesian methods to recover model parameters. The rest of the article consists of four sections. Section 2 presents and formulates three non-ignorable missingness selection models. Section 3 presents a full Bayesian method to estimate the latent growth models through the data augmentation and Gibbs sampling algorithms. Section 4 conducts a simulation study. Estimates from models with different non-ignorable missingness and different sample sizes are summarized, analyzed, and compared. Conclusions based on the simulation study are drawn. Section 5 discusses the implications and future directions of this study. Finally, the appendix presents technical details.

2   Non-ignorable Missingness in Latent Growth Models

In this section, we model the non-ignorable missingness in growth models. Before we introduce the three selection models, we first review the latent growth curve models (LGCMs).

2.1   Latent Growth Curve Models (LGCMs)

The latent growth curve models (LGCMs) can be expressed by a regression equation with latent variables being regressors. Specifically, for a longitudinal study with N subjects and T measurement time points, let $\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{iT})'$ be a $T \times 1$ random vector, where y_{it} stands for the outcome or observation of

individual i at occasion t $(i = 1, 2, ..., N;\ t = 1, 2, ..., T)$, and let $\boldsymbol{\eta}_i$ be a $q \times 1$ random vector containing q continuous latent variables. A latent growth curve model for the outcome $\mathbf{y}_i$ related to the latent $\boldsymbol{\eta}_i$ can be written as

$$\mathbf{y}_i = \boldsymbol{\Lambda}\boldsymbol{\eta}_i + \mathbf{e}_i \tag{1}$$

$$\boldsymbol{\eta}_i\ = \boldsymbol{\beta} + \boldsymbol{\xi}_i, \tag{2}$$

where $\boldsymbol{\Lambda}$ is a $T \times q$ matrix consisting of factor loadings, $\mathbf{e}_i$ is a $T \times 1$ vector of residuals or measurement errors that are assumed to follow multivariate normal distributions, i.e., $\mathbf{e}_i \sim MN_T(\mathbf{0}, \boldsymbol{\Theta})$ [1], and $\boldsymbol{\xi}_i$ is a $q \times 1$ vector that is assumed to follow a multivariate distribution, i.e., $\boldsymbol{\xi}_i \sim MN_q(\mathbf{0}, \boldsymbol{\Psi})$. In LGCMs, $\boldsymbol{\beta}$ is called *fixed effects* and $\boldsymbol{\xi}_i$ is called *random effects* (e.g., Fitzmaurice et al., 2004; Hedges, 1994; Luke, 2004; Singer & Willett, 2003). The vectors $\boldsymbol{\beta}$, $\boldsymbol{\eta}_i$, and the matrix $\boldsymbol{\Lambda}$ determine the growth trajectory of the model. For instance, when $q = 2$, $\boldsymbol{\beta} = (I, S)'$, $\boldsymbol{\eta}_i = (I_i, S_i)'$, and $\boldsymbol{\Lambda}$ is a $T \times 2$ matrix containing the first column of 1s and the second column of $(0, 1, ..., T-1)$. The corresponding model represents a linear growth model in which I is the latent population intercept (or latent random initial level), S is the latent population slope, I_i is individual i's latent random intercept and S_i is individual i's latent random slope. Furthermore, when $q = 3$, $\boldsymbol{\beta} = (I, S, Q)'$, $\boldsymbol{\eta}_i = (I_i, S_i, Q_i)'$, and $\boldsymbol{\Lambda}$ is a $T \times 3$ matrix containing the first column of 1s, the second column of $(0, 1, ..., T - 1)$, and the third column of $(0, 1, ..., (T - 1)^2)$. The corresponding model represents a quadratic growth curve model with Q and Q_i being latent quadratic coefficients for population and individual i, respectively.

2.2   Selection Models for Non-ignorable Missingness

To address the non-ignorable missingness, there are two general approaches, *pattern-mixture models* (Hedeker & Gibbons, 1997; Little & Rubin, 1987) and *selection models* (Glynn, Laird, & Rubin, 1986; Little, 1993, 1995). In both cases, the statistical analysis requires joint modelling of dependent variable and missing data processes. In this research, selection models are used, mainly for two reasons. First, substantively, it seems more natural to consider the behavior of the response variable in the full target population of interests, rather than in the sub-populations defined by missing data patterns (e.g., Fitzmaurice et al., 2008). Second, the selection model formulation leads directly to the joint distribution of both dependent variables and the missingness (e.g., Fitzmaurice et al., 2008) as follows,

$$f(\mathbf{y}_i, \mathbf{m}_i | \boldsymbol{\nu}, \boldsymbol{\phi}, \mathbf{x}_i) = f(\mathbf{y}_i | \boldsymbol{\nu}, \mathbf{x}_i)\, f(\mathbf{m}_i | \mathbf{y}_i, \boldsymbol{\nu}, \boldsymbol{\phi}, \mathbf{x}_i)$$

where $f(.)$ is a density function, $\mathbf{x}_i$ is a vector of covariates for individual i, $\mathbf{y}_i$ is a vector of individual i's outcome scores, $\boldsymbol{\Theta} = (\boldsymbol{\nu}, \boldsymbol{\phi})$ are all parameters in the model, in which $\boldsymbol{\nu}$ are parameters for the growth model and $\boldsymbol{\phi}$ are

---

[1]   Throughout the article, $MN_n(\cdot)$ denotes a n-dimensional multivariate normal distribution, and $Mt_n(\cdot)$ denotes a n-dimensional multivariate t distribution.

parameters for the missingness, and $\mathbf{m}_i$ is a vector $\mathbf{m}_i = (m_{i1}, m_{i2}, ..., m_{iT})'$ that indicates the missingness status for $\mathbf{y}_i$. Specifically, if y_i is missing at time point t, then $m_{it} = 1$; otherwise, $m_{it} = 0$. Here, we assume the missingness is conditionally independent (e.g., Dawid, 1979), which means across different occasions the conditional distributions of missingness are independent with each other. Let $\tau_{it} = f(m_{it} = 1)$ be the probability that y_{it} is missing, then m_{it} follows a Bernoulli distribution of τ_{it}, and the density function of m_{it} is $f(m_{it}) = \tau_{it}^{m_{it}}(1 - \tau_{it})^{1 - m_{it}}$. For different non-ignorable missingness patterns, the expressions of τ_{it} are different. Lu et al. (2011) investigated the non-ignorable missingness in mixture models. The τ_{it} in that article is a function of latent class membership, and thus the missingness is *Latent Class Dependent (LCD)*.

However, *LCD* was proposed in the framework of mixture models. Within each latent population, there is no class membership indicator. Consequently, the missingness is ignorable. In this article, we consider more complex non-ignorable missingness mechanisms within a population. In general, we assume $\mathbf{L}_i$ is a vector of latent variables that depend on the missing values. A general class of selection models for dealing with non-ignorable missing data in latent growth modelling can be formulated as

$$
\begin{aligned}
f(\mathbf{y}_i, \mathbf{m}_i | \boldsymbol{\beta}, \boldsymbol{\xi}_i, \mathbf{L}_i, \boldsymbol{\gamma}_t, \mathbf{x}_i) &= f(\boldsymbol{\eta}_i | \boldsymbol{\beta}, \boldsymbol{\xi}_i) f(\mathbf{y}_i | \boldsymbol{\eta}_i) \, \Phi(\boldsymbol{\omega}_i' \boldsymbol{\gamma}_t)^{m_{it}} [1 - \Phi(\boldsymbol{\omega}_i' \boldsymbol{\gamma}_t)]^{1 - m_{it}} \\
&= f(\boldsymbol{\eta}_i | \boldsymbol{\beta}, \boldsymbol{\xi}_i) f(\mathbf{y}_i | \boldsymbol{\eta}_i) \Phi(\gamma_{0t} + \mathbf{L}_i \boldsymbol{\gamma}_{Lt} + \mathbf{x}_i' \boldsymbol{\gamma}_{xt})^{m_{it}} \\
&\quad \times [1 - \Phi(\gamma_{0t} + \mathbf{L}_i \boldsymbol{\gamma}_{Lt} + \mathbf{x}_i' \boldsymbol{\gamma}_{xt})]^{1 - m_{it}}
\end{aligned}
\tag{3}
$$

where $\mathbf{x}_i$ is an r-dimensional vector, $\boldsymbol{\omega}_i = (1, \mathbf{L}_i', \mathbf{x}_i')'$ and $\boldsymbol{\gamma}_t = (\gamma_{0t}, \boldsymbol{\gamma}_{Lt}', \boldsymbol{\gamma}_{xt}')'$. The missingness is non-ignorable because it depends on the latent variables $\mathbf{L}_i$ in the model and the observed data are not completely informative about these latent variables. Note that the vector $\boldsymbol{\gamma}_{Lt}$ here should be non-zero. Otherwise, the missingness becomes ignorable.

Specific sub-models under different situations can be derived from this general model. For example, missingness may be related to latent intercepts, latent growth rates, or potential outcomes. To show different types of non-ignorable missingness, we draw the path diagrams, as shown in Figures 1, 2, and 3, to illustrate the sub-models. These sub-models are based on three types of latent variables on which the missingness might depend. In these path diagrams, a square/rectangle indicates an observed variable, a circle/oval means a latent variable, a triangle represents a constant, and arrows show the relationship among them. y_t is the outcome at time t, which is influenced by latent effects such as I, S, and η_q. As the value of y_t might be missing, we use both circle and square in the path diagram. If y_t is missing, then the potential outcome cannot be observed and the corresponding missingness indicator m_t becomes 1. The dashed lines between y_t and m_t show the 1-1 relationship. In these sub-models, the value of m_t depends on the observed covariate x_r and some latent variables. The details of these three sub-models are described as follows.

2.2.1   Latent Intercept Dependent (LID) Missingness (Figure 1). It illustrates the situation where the missingness depends on individual's latent

intercept, I_i. For example, a student's latent initial ability level of the knowledge of a course influences the likelihood of that participant dropping out of or staying in that course. If the latent initial ability of a course is not high, a student may choose to drop that course or even drop out a school. In the case of LID, the $\mathbf{L}_i$ in Equation (3) is simplified to a univariate I_i. Suppose that the missingness is also related to some observed covariates $\mathbf{x}_i$, such as parents' education or family income, then τ_{Iit} is expressed as a probit link function of I_i and $\mathbf{x}_i$

$$\tau_{Iit} = \Phi(\gamma_{0t} + I_i\gamma_{It} + \mathbf{x}'_i\boldsymbol{\gamma}_{xt}) = \Phi(\boldsymbol{\omega}'_{Ii}\boldsymbol{\gamma}_{It}), \tag{4}$$

where $\boldsymbol{\omega}_{Ii} = (1, I_i, \mathbf{x}'_i)'$ and $\boldsymbol{\gamma}_{It} = (\gamma_{0t}, \gamma_{It}, \boldsymbol{\gamma}'_{xt})'$.
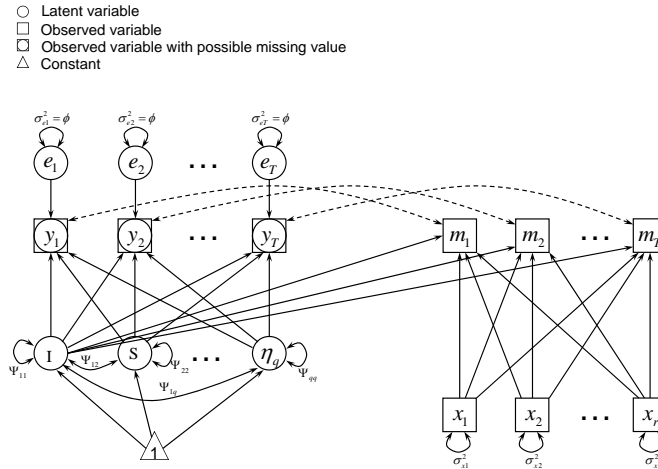
Figure 1. Path diagram of a latent growth model with latent intercept dependent missingness (LID) where $f(m_t)$ depends on covariates x_rs and latent intercept I.

2.2.2 Latent Slope Dependent (LSD) Missingness (Figure 2). It describes the situation where the missingness depends on the latent slope, S_i. For example, a student's latent rate of change in a course influences the likelihood that the participant misses a test in the future. This might be the case, if the student didn't see any improvement over time, at which point he/she might choose to drop out. In the case of LSD, the $\mathbf{L}_i$ in Equation (3) becomes a univariate S_i. Together with some other observed covariates $\mathbf{x}_i$, for example, parents' education or family income, the missing data rate τ_{it} can be expressed as a probit link function of S_i and $\mathbf{x}_i$,

$$\tau_{Sit} = \Phi(\gamma_{0t} + S_i\gamma_{St} + \mathbf{x}'_i\boldsymbol{\gamma}_{xt}) = \Phi(\boldsymbol{\omega}'_{Si}\boldsymbol{\gamma}_{St}), \tag{5}$$

with $\boldsymbol{\omega}_{Si} = (1, S_i, \mathbf{x}_i')'$ and $\boldsymbol{\gamma}_{St} = (\gamma_{0t}, \gamma_{St}, \boldsymbol{\gamma}_{xt}')'$.

Figure 2. Path diagram of a latent growth model with latent slope dependent missingness (LSD) where $f(m_t)$ depends on covariates x_rs and latent slope S.

2.2.3   Latent Outcome Dependent (LOD) Missingness (Figure 3). It assumes that the missingness depends on potential outcomes y_{it}. For example, a student who feels not doing well on a test may be more likely to quit taking the rest of the test. As a result, the missing score is due to the perceived potential outcome of the test. In this case, the $\mathbf{L}_i$ in Equation (3) is the potential outcome y_{it}. With some covariates $\mathbf{x}_i$, we express τ_{it} as a probit link function as follows.

$$\tau_{yit} = \Phi(\gamma_{0t} + y_{it}\gamma_{yt} + \mathbf{x}_i'\boldsymbol{\gamma}_{xt}) = \Phi(\boldsymbol{\omega}_{yit}' \boldsymbol{\gamma}_{yt}), \tag{6}$$

with $\boldsymbol{\omega}_{yit} = (1, y_{it}, \mathbf{x}_i')'$ and $\boldsymbol{\gamma}_{yt} = (\gamma_{0t}, \gamma_{yt}, \boldsymbol{\gamma}_{xt}')'$.

3   Bayesian Estimation

In this article, a full Bayesian estimation approach is used to estimate growth models. The algorithm is described as follows. First, model related latent variables are added via the data augmentation method (Tanner & Wong, 1987). By including auxiliary variables, the likelihood function for each model is obtained. Second, proper priors are adopted. Third, with the likelihood function and the priors, based on the Bayes' Theorem, the posterior distribution of
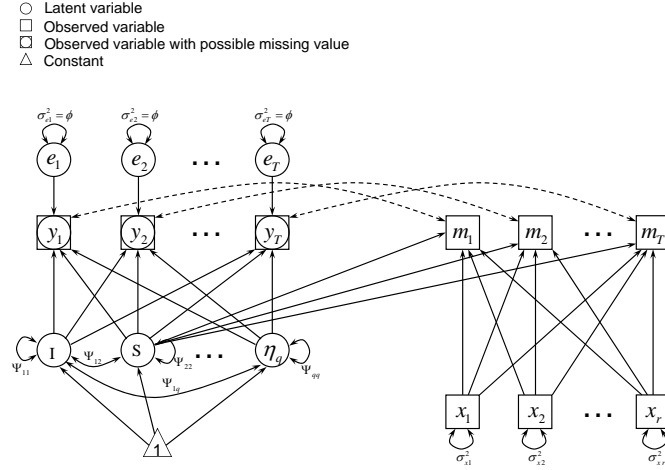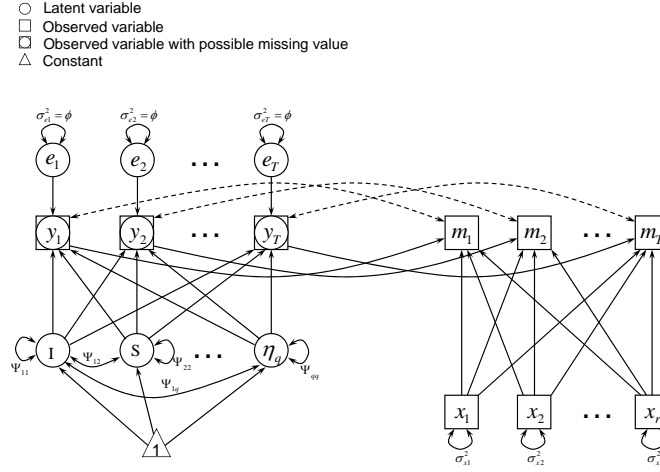
Figure 3. Path diagram of a latent growth model with latent outcome dependent missingness (LOD) where $f(m_t)$ depends on covariates x_rs and potential outcome y.

the unknown parameters is readily available. We obtain conditional posterior distributions instead of the joint posterior because the integrations of marginal posterior distributions of the parameters are usually hard to obtain explicitly for high-dimensional data. Fourth, with conditional posterior distributions, Markov chains are generated for the unknown model parameters by implementing a Gibbs sampling algorithm (Casella & George, 1992; Geman & Geman, 1984). Finally, statistical inference is conducted based on converged Markov chains.

3.1 Data Augmentation and Likelihood Functions

In order to construct the likelihood function explicitly, we use the data augmentation algorithm (Tanner & Wong, 1987). The observed outcomes $\mathbf{y}_i^{obs}$ can be augmented with the missing values $\mathbf{y}_i^{mis}$ such that $\mathbf{y}_i = (\mathbf{y}_i^{obs}, \mathbf{y}_i^{mis})'$ for individual i. Also, the missing data indicator variable $\mathbf{m}_i$ is added to models. Then the joint likelihood function of the selection model for the ith individual can be expressed as $L_i(\boldsymbol{\eta}_i, \mathbf{y}_i, \mathbf{m}_i) = [f(\boldsymbol{\eta}_i) f(\mathbf{y}_i|\boldsymbol{\eta}_i)] \ f(\mathbf{m}_i|\mathbf{y}_i, \boldsymbol{\eta}_i, \mathbf{x}_i)$. For the

whole sample, the likelihood function is specifically expressed as

$$L(\mathbf{y}, \boldsymbol{\eta}, \mathbf{m}) \propto \prod_{i=1}^{N} \left\{ |\boldsymbol{\Psi}|^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{\eta}_i - \boldsymbol{\beta})' \boldsymbol{\Psi}^{-1}(\boldsymbol{\eta}_i - \boldsymbol{\beta}) \right] \right.$$
$$\left. \times |\phi|^{-T/2} \exp\left[-\frac{1}{2\phi}(\mathbf{y}_i - \boldsymbol{\Lambda}\boldsymbol{\eta}_i)'(\mathbf{y}_i - \boldsymbol{\Lambda}\boldsymbol{\eta}_i) \right] \right. \qquad (7)$$
$$\left. \times \prod_{t=1}^{T} \left[\tau_{it}^{m_{it}}(1 - \tau_{it})^{1-m_{it}} \right] \right\},$$

where τ_{it} is defined by Equation (4) for the LID missingness, (5) for the LSD missingness, and (6) for the LOD missingness.

3.2   Prior and Posterior Distributions

The commonly used proper priors (e.g., Lee, 2007) are adopted in the study. Specifically, (1) an inverse Gamma distribution prior is used for $\phi \sim IG(v_0/2, s_0/2)$ where v_0 and s_0 are given hyper-parameters. The density function of an inverse Gamma distribution is $f(\phi) \propto \phi^{-(v_0/2)-1} \exp(-s_0/(2\phi))$. (2) An inverse Wishart distribution prior is used for $\boldsymbol{\Psi}$. With hyper-parameters m_0 and $\mathbf{V}_0$, $\boldsymbol{\Psi} \sim IW(m_0, \mathbf{V}_0)$, where m_0 is a scalar and $\mathbf{V}_0$ is a $q \times q$ matrix. Its density function is $f(\boldsymbol{\Psi}) \propto |\boldsymbol{\Psi}|^{-(m_0+q+1)/2} \exp[-tr(\mathbf{V}_0\boldsymbol{\Psi}^{-1})/2]$. (3) For $\boldsymbol{\beta}$ a multivariate normal prior is used, and $\boldsymbol{\beta} \sim MN_q(\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$ where the hyper-parameter $\boldsymbol{\beta}_0$ is a q-dimensional vector and $\boldsymbol{\Sigma}_0$ is a $q \times q$ matrix. (4) The prior for $\boldsymbol{\gamma}_t$ ($t = 1, 2, \ldots, T$) is chosen to be a multivariate normal distribution $\boldsymbol{\gamma}_t \sim MN_{(2+r)}(\boldsymbol{\gamma}_{t0}, \mathbf{D}_{t0})$, where $\boldsymbol{\gamma}_{t0}$ is a $(2 + r)$-dimensional vector, $\mathbf{D}_{t0}$ is a $(2 + r) \times (2 + r)$ matrix, and both are pre-determined hyper-parameters.

After constructing the likelihood function and assigning the priors, the joint posterior distribution for unknown parameters is readily available. Considering the high-dimensional integration for marginal distributions of parameters, the conditional distribution for each parameter is obtained instead. The derived conditional posteriors are provided in Equations (8) - (11) in the appendix. In addition, the conditional posteriors for the latent variable $\boldsymbol{\eta}_i$ and the augmented missing data $\mathbf{y}_i^{mis}$ ($i = 1, 2, ..., N$) are also provided by Equations (12) and (13), respectively, in the appendix.

3.3   Gibbs Sampling

After obtaining the conditional posteriors, the Markov chain for each model parameter is generated by implementing a Gibbs sampling algorithm (Casella & George, 1992; Geman & Geman, 1984). Specifically, the following algorithm is used in the research.

1. Start with a set of initial values for model parameters $\phi^{(0)}$, $\boldsymbol{\Psi}^{(0)}$, $\boldsymbol{\beta}^{(0)}$, $\boldsymbol{\gamma}^{(0)}$, latent variable $\boldsymbol{\eta}^{(0)}$, and missing values $\mathbf{y}^{mis(0)}$.

2. At the sth iteration, the following parameters are generated: $\phi^{(s)}$, $\mathbf{\Psi}^{(s)}$, $\boldsymbol{\beta}^{(s)}$, $\boldsymbol{\gamma}^{(s)}$, $\boldsymbol{\eta}^{(s)}$, and $\mathbf{y}^{mis(s)}$. To generate $\phi^{(s+1)}$, $\mathbf{\Psi}^{(s+1)}$, $\boldsymbol{\beta}^{(s+1)}$, $\boldsymbol{\gamma}^{(s+1)}$, $\boldsymbol{\eta}^{(s+1)}$, and $\mathbf{y}^{mis(s+1)}$, the following procedure is implemented:

(a) Generate $\phi^{(s+1)}$ from the distribution in Equation (8) in the appendix.

(b) Generate $\mathbf{\Psi}^{(s+1)}$ from the inverse Wishart distribution in Equation (9) in the appendix. iv. Generate $\boldsymbol{\beta}^{(s+1)}$ from the multivariate normal distribution in Equation (10) in the appendix.

(c) Generate $\boldsymbol{\gamma}^{(s+1)}$ from the distribution in Equation (11) in the appendix.

(d) Generate $\boldsymbol{\eta}^{(s+1)}$ from the multivariate normal distribution in Equation (12) in the appendix.

(e) Generate $\mathbf{y}^{mis(s+1)}$ from the normal distribution in Equation (13) in the appendix.

3.4   Statistical Inference

After passing convergence tests, the generated Markov chains can be viewed as from the joint and marginal distributions of all parameters. The statistical inference can then be conducted based on the generated Markov chains.

Suppose θ is an unknown parameter. For different loss functions of θ, the point estimates are different. For example, if a square loss function, $LF = (\theta - \hat{\theta})^2$, is used, then the posterior mean is the estimate of θ; but if an absolute loss function, $LF = |\theta - \hat{\theta}|$, is used, then its estimate is the posterior median. There are other function forms, such as 0-1 loss function, but in this research we use the square loss function.

Let $\mathbf{\Theta} = (\theta_1, \theta_2, ..., \theta_p)'$ denote a vector of all the unknown parameters in the model. Then the converged Markov chains can be recorded as $\mathbf{\Theta}^{(s)}, s = 1, 2, \ldots, S$, and each parameter estimate $\hat{\theta}_j$ $(j = 1, 2, ..., p)$ can be calculated as $\hat{\theta}_j = \sum_{s=1}^{S} \theta_j^{(s)}/S$ with standard error (SE) $s.e.(\hat{\theta}_j) = \sqrt{\sum_{s=1}^{S}(\theta_j^{(s)} - \hat{\theta}_j)^2/(S-1)}$. To get the credible intervals, both percentile intervals and the highest posterior density intervals (HPD, Box & Tiao, 1973) of the Markov chains can be used. Percentile intervals are obtained by sorting $\theta_j^{(s)}$. HPD intervals may also be referred as minimum length confidence intervals for a Bayesian posterior distribution, and for symmetric distributions HPD intervals obtain equal tail area probabilities.

4   Simulation Studies

In this section, simulation studies are conducted to evaluate the performance of the proposed models estimated by the Bayesian method.

4.1   Simulation Design

In the simulation, we focus on linear LGCMs to simplify the presentation. Higher order LGCMs can be easily expanded by adding quadratic or higher order terms.

First, four waves of complete LGCM data $\mathbf{y}_i$ are generated based on Equations (1) and (2). The random effects consist of the intercept I_i and the slope S_i, with $Var(I_i) = 1$, $Var(S_i) = 4$, and $Cov(I_i, S_i) = 0$. The fix-effects are $(I, S) = (1, 3)$. The measurement errors are assumed to follow a normal distribution with mean 0 and standard deviation 1. In the simulation, we also assume there is one covariate X generated from a normal distribution, $X \sim N(1, sd = 0.2)$. Missing data are created based on different pre-designed missingness rates. We assume the true missingness is LSD (also noted as the XS missingness in this study because the missingness depends on the latent individual slope S and covariate X). With LSD, the bigger the slope is, the more the missing data. For the sake of simplicity in the simulation, the missingness rate is set the same for every occasion. Specifically, we set the missingness probit coefficients as $\gamma_0 = (-1, -1, -1, -1)$, $\gamma_x = (-1.5, -1.5, -1.5, -1.5)$, and $\gamma_S = (0.5, 0.5, 0.5, 0.5)$. With the setting, missingness rates are generated based on Equation (5). If a participant has a latent growth slope 3, with a covariate value 1, his or her missingness rate at each wave is $\tau \approx 16\%$; and if the slope is 5, with the same covariate value, the missing rate increases to $\tau \approx 50\%$; but when the slope is 1, the missingness rate decreases to $\tau \approx 2.3\%$.

Next, we fit data with LGCMs with different missingness. Specifically, the model design with different missingness is shown in Table 1, where the symbol "✓" shows the related factors on which the missing data rates depend. For example, when both "X" and "I" are checked, the missingness depends on the individual's latent intercept "I" and the observed covariate "X". Four types of missingness are studied: LID (also noted as XI in Table 1), LSD (XS), LOD (XY), and ignorable (X). The shaded model, LSD (XS), is the true model we used for generating the simulation data. Five levels of sample size (N=1000, N=500, N=300, N=200 and N=100) are investigated, and for each sample size. In total, 4×5=20 summary tables are combined and presented in Tables 2, 3, and 5-9 [2]. Each result table is summarized from 100 converged replications.

4.2   Simulation Implementation

The simulation studies are implemented by the following algorithm. (1) Set the counter $R = 0$. (2) Generate complete longitudinal growth data according to predefined model parameters. (3) Create missing data according to missing data mechanisms and missing data rates. (4) Generate Markov chains for model parameters through the Gibbs sampling procedure. (5) Test the convergence of generated Markov chains. (6) If the Markov chains pass the convergence test, set $R = R + 1$ and calculate and save the parameter estimates. Otherwise, set $R = R$ and discard the current replication of simulation. (7) Repeat the above process till $R = 100$ to obtain 100 replications of valid simulation.

In step 4, priors carrying little prior information are adopted (Congdon, 2003; Gill, 2002; Zhang, Hamagami, Wang, Grimm, & Nesselroade, 2007). Specifically,

---

[2] The summary table for the model with the latent intercept dependent (LID) missingness (XI), for N=100 is not included due to its low convergence rate.

Table 1. Simulation model design. N=1000, 500, 300, 200 and 100

Model	X[5]	I[6]	S[7]	Y[8]
Ignorable (X)	✓			
LID[2] (XI)	✓	✓		
LSD[3] (XS)[1]	✓		✓	
LOD[4] (XY)	✓			✓

Note. [1]The shaded model is the true model XS. [2]LID: Latent Intercept Dependent. [3]LSD: Latent Slope Dependent. [4]LOD: Latent Outcome Dependent. [5]X: Observed covariates. If X is the only item checked, the missingness is ignorable. [6]I: Individual latent intercept. If checked, the missingness is non-ignorable. [7]S: Individual latent slope. If checked, the missingness is non-ignorable. [8]Y: Individual potential outcome y. If checked, the missingness is non-ignorable.

for $\boldsymbol{\varphi}_1$, we set $\boldsymbol{\mu}_{\varphi_1} = \mathbf{0}_2$ and $\boldsymbol{\Sigma}_{\varphi_1} = 10^3\mathbf{I}_2$. For ϕ, we set $v_{0k} = s_{0k} = 0.002$. For $\boldsymbol{\beta}$, it is assumed that $\boldsymbol{\beta}_{k0} = \mathbf{0}_2$ and $\boldsymbol{\Sigma}_{k0} = 10^3\mathbf{I}_2$. For $\boldsymbol{\Psi}$, we define $m_{k0} = 2$ and $\mathbf{V}_{k0} = \mathbf{I}_2$. Finally, for $\boldsymbol{\gamma}_t$, we let $\boldsymbol{\gamma}_{t0} = \mathbf{0}_3$ and $\mathbf{D}_{t0} = 10^3\mathbf{I}_3$, where $\mathbf{0}_d$ and $\mathbf{I}_d$ denote a d-dimensional zero vector and a d-dimensional identity matrix, respectively. In step 5, the iteration number of burn-in period is set. The Geweke convergence criterion indicated that less than 10,000 iterations was adequate for all conditions in the study. Therefore, a conservative burn-in of 20,000 iterations was used for all iterations. And then the Markov chains with a length of $20,000$ iterations are saved for convergence testing and data analysis. After step 7, 12 summary statistics are reported based on 100 sets of converged simulation replications. For the purpose of presentation, let θ_j represent the jth parameter, also the true value in the simulation. Twelve statistics are defined below. (1) The average estimate (est.$_j$) across 100 converged simulation replications of each parameter is obtained as est.$_j = \bar{\hat{\theta}}_j = \sum_{i=1}^{100} \hat{\theta}_{ij}/100$, where $\hat{\theta}_{ij}$ denotes the estimate of θ_j in the ith simulation replication. (2) The simple bias (BIAS.smp$_j$) of each parameter is calculated as BIAS.smp$_j = \bar{\hat{\theta}}_j - \theta_j$. (3) The relative bias (BIAS.rel$_j$) of each parameter is calculated using BIAS.rel$_j = (\bar{\hat{\theta}}_j - \theta_j)/\theta_j$ when $\theta_j \neq 0$ and BIAS.rel$_j = \bar{\hat{\theta}}_j - \theta_j$ when $\theta_j = 0$. (4) The empirical standard error (SE.emp$_j$) of each parameter is obtained as SE.emp$_j = \sqrt{\sum_{i=1}^{100}(\hat{\theta}_{ij} - \bar{\hat{\theta}}_j)^2/99}$, and (5) the average standard error (SE.avg$_j$) of the same parameter is calculated by SE.avg$_j = \sum_{i=1}^{100} \hat{s}_{ij}/100$, where $\hat{s}_{ij}$ denotes the estimated standard error of $\hat{\theta}_{ij}$. (6) The average mean square error (MSE) of each parameter is obtained by MSE$_j = \sum_{i=1}^{100}$ MSE$_{ij}/100$, where MSE$_{ij}$ is the mean square error for the jth parameter in the ith simulation replication, MSE$_{ij} = (\text{Bias}_{ij})^2 + (\hat{s}_{ij})^2$. The average lower (7) and upper (8) limits of the 95% percentile confidence interval (CI.low$_j$ and CI.upper$_j$) are respectively defined as CI.low$_j = \sum_{i=1}^{100} \hat{\theta}_{ij}^l/100$ and CI.upper$_j = \sum_{i=1}^{100} \hat{\theta}_{ij}^u/100$ where $\hat{\theta}_{ij}^l$ and $\hat{\theta}_{ij}^u$ denote the 95% lower and upper

limits of CI for the jth parameter, respectively. (9) The coverage probability of the 95% percentile confidence interval (CI.cover$_j$) of each parameter is obtained using CI.cover$_j = [\#(\hat{\theta}_{ij}^l \leq \theta_j \leq \hat{\theta}_{ij}^u)]/100$. The average lower (10), upper (11) limits, and (12) the coverage probability of the 95% highest posterior density credible interval (HPD, Box & Tiao, 1973) of each parameter are similarly defined by HPD.low$_j$, HPD.upper$_j$, and HPD.cover$_j$, respectively.

4.3   Simulation Results

In this section, we show simulation results for the estimates obtained from the true model and misspecified models.

4.3.1   Estimates from the True Model. First, we investigate the estimates obtained from the true model. Tables 3, 4 and 5 in the appendix show the summarized estimates from the true model for N=1000, N=500, N=300, and N=100. From Tables 3 with the sample size 1000, first, one can see that all the relative estimate biases are very small, with the largest one being 0.067 for γ_{03}. Second, the difference between the empirical SEs and the average SEs is very small, which indicates the SEs are estimated accurately. Third, both CI and HPD interval coverage probabilities are very close to the theoretical percentage 95%, which means the type I error for each parameter is close to the specified 5% so that we can use the estimated confidence intervals to conduct statistical inference. Fourth, this true model has 100% convergence rate. When the sample sizes are smaller, the performance becomes worse as expected.

In order to compare estimates with different sample sizes, we further calculate the five summary statistics across all parameters, which are shown in Table 2. The first statistic is the average absolute relative biases (|Bias.rel|) across all parameters, which is defined as |Bias.rel| $= \sum_{j=1}^{p} |$Bias.rel$_j|/p$, where p is the total number parameters in a model. Second, we obtain the average absolute differences between the empirical SEs and the average Bayesian SEs (|SE.diff|) across all parameters by using |SE.diff| $= \sum_{j=1}^{p} |$SE.emp$_j -$ SE.avg$_j|/p$. Third, we calculate the average percentile coverage probabilities (CI.cover) across all parameters by using CI.cover $= \sum_{j=1}^{p}$ CI.cover$_j/p$. Fourth, we calculate the average HPD coverage probabilities (HPD.cover) across all parameters by using HPD.cover $= \sum_{j=1}^{p}$ HPD.cover$_j/p$. Fifth, the convergence rate is calculated.

Table 2 shows that, except for the case for N=100, the true mode can recover model parameters very well, with small average absolute relative biases of estimates, |Bias.rel|, small average absolute differences between the empirical SEs and the average SEs, |SE.diff|, and almost 95% average percentile coverage probabilities (CI.cover), and the average HPD coverage probabilities (HPD.cover). With the increase of the sample size, both the point estimates and standard errors get more accurate.

4.3.2   Comparison of Different Models. We now compare the estimates obtained from the true model and different misspecified models. In this study,

Table 2. Summary and Comparison of the Results of True Model XS

		$|\text{Bias.rel}|^1$	$|\text{SE.diff}|^2$	MSE^3	CI.cover^4	HPD.cover^5	CVG.rate^6
	1000	0.025	0.007	0.033	0.942	0.942	100%
	500	0.052	0.021	0.079	0.932	0.939	100%
N	300	0.089	0.031	0.150	0.922	0.930	100%
	200	0.160	0.090	0.366	0.909	0.924	94.34%
	100	1.202	2.664	23.743	0.869	0.893	70.42%

Note. [1]The average absolute relative bias across all parameters, defined by $|\text{Bias.rel}| = \sum_{j=1}^{p} |\text{Bias.rel}_j|/p$. The smaller, the better. [2]The average absolute difference between the empirical SEs and the average Bayesian SEs across all parameters, defined by $|\text{SE.diff}| = \sum_{j=1}^{p} |\text{SE.emp}_j - \text{SE.avg}_j|/p$. The smaller, the better. [3]The Mean Square Errors (MSE) across all parameters, defined by $MSE = \sum_{j=1}^{p} [(\text{Bias}_j)^2 + (\hat{s}_j)^2]/p$. The smaller, the better. [4]The average percentile coverage probability across all parameters, defined by $\text{CI.cover} = \sum_{j=1}^{p} \text{CI.cover}_j/p$, with a theoretical value of 0.95. [5]The average highest posterior density (HPD) coverage probability across all parameters, defined by $\text{HPD.cover} = \sum_{j=1}^{p} \text{HPD.cover}_j/p$, with a theoretical value of 0.95. [6]The convergence rate.

the true model is the LGCM with LSD (XS) missingness, and there are three mis-specified models, the LGCM with LID (XI) missingness, the LGCM with LOD (XY) missingness, and the LGCM with ignorable missingness (see Table 1 for the simulation design). Table 6 in the appendix shows the summarized estimates from the mis-specified model with LID (XI) missingness for N=1000, N=500, N=300, and N=200 (the summarized estimates for N=100 are unavailable due to a low convergence rate). Table 8 in the appendix provides the results for the mis-specified model with LOD (XY) missingness for N=1000, N=500, N=300, N=200, and N=100. Table 10 in the appendix is the summary table for the mis-specified model with ignorable (X) missingness for different sample sizes.

To compare estimates from different models, we further summarize and visualize some statistics. Figure 4 (a) compares the point estimates of intercept and slope for all models when N=1000. The true value of slope is 3 but the estimate is 2.711 when the missingness is ignored. Actually, for the model with ignorable missingness, the slope estimates are all less than 2.711 for all sample sizes in our study. Figure 4 (b) focuses on the coverage of slope. When the missingness is ignored, it is as low as 4% for N=1000, and 21% for N=500 (the coverage for N=1000 is lower because the SE for N=1000 is smaller than the SE for N=500). As a result, conclusions based on the model with ignorable missingness can be very misleading. Figure 4 (b) also shows that the slope estimate from the model with the mis-specified missingness, LID (XI), has low coverage, with 76% for N=1000 and 87% for N=500. So the conclusions based on this model may still be incorrect. Figure 4 (c) compares the true model and the model with another type of mis-specified missingness, LOD (XY) for N=1000.

For the wrong model, the coverage is 51% for intercept, and 72% for Cov(I,S). Finally, Figure 4 (d) compares the convergence rates for all models. One can see that the convergence rates of LOD (XY) and LID (XI) models are much lower than those of the true model LSD (XS) and the model with ignorable missingness. When the missingness is ignored, the number of parameters is smaller than that of non-ignorable models, and then convergence rate gets higher.

(a) Intercept and Slope Estimates
(True Int=1 & Slope=3)

(b) Slope Coverage
(Theoretical coverage=95%)

(c) Parameter Coverage for LSD(XS) and
LOD (XY) (Theoretical value=95%)

(d) Convergence Rates
(Closer to 100% is better.)
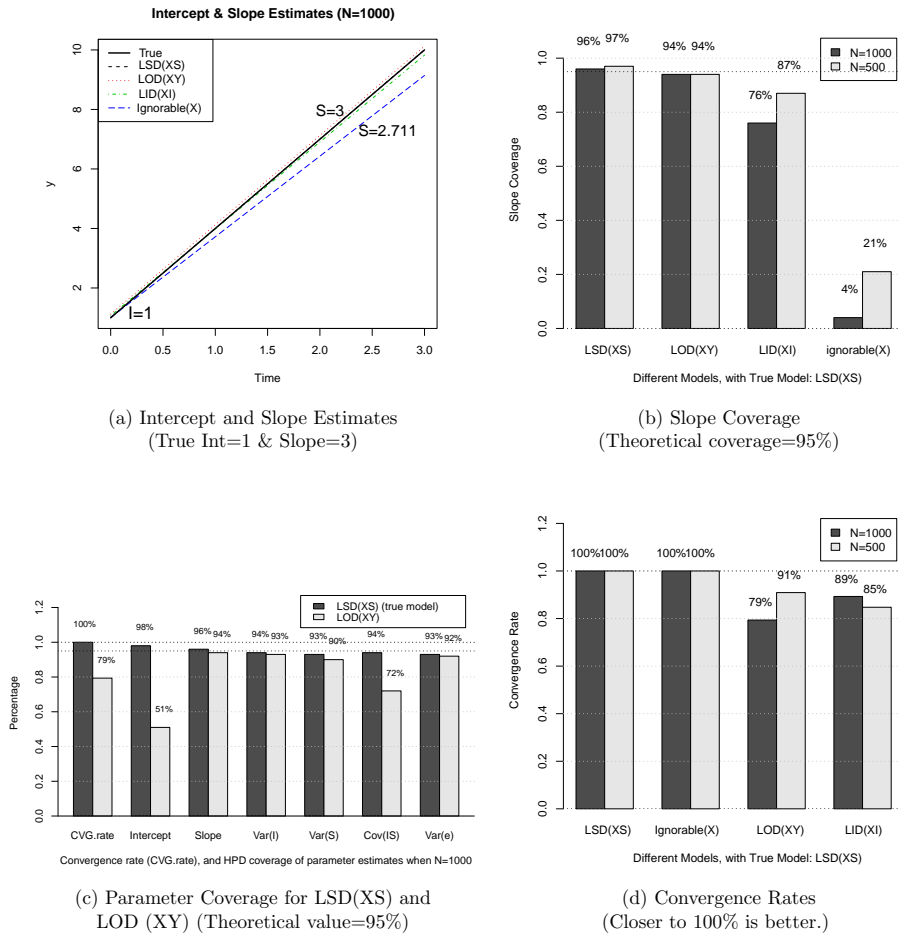
Figure 4. Comparison of four models.

In summary, the estimates from mis-specified models may result in misleading conclusions, especially when the missingness is ignored. Also, the convergence rate of a mis-specified model is usually lower than that of the true model.

4.4   Simulation Conclusions

Based on the simulation studies, we draw the following conclusions: (1) the proposed Bayesian method can accurately recover model parameters (both point estimates and standard errors), (2) the small difference between the empirical SE and the average SE indicates that the Bayesian method used in the study can estimate the standard errors accurately, (3) with the increase of the sample size, estimates get closer to their true values and standard errors become more accurate, (4) ignoring the non-ignorable missingness can lead to incorrect conclusions, (5) mis-specified missingness may also result in misleading conclusions, and (6) the non-convergence of models might be a sign of a misspecified model.

5   Discussion

The models proposed in this article have several implications for future research. First, the missingness in the simulation study is assumed to be independent across different times. If this assumption is violated, likelihood functions might be much more complicated. For example, if the missingness depends on the previous missingness, then the autocorrelation among missingness might be involved. A similar model is the Diggle and Kenward (1994)'s model, in which the probability of missing data at current wave depends directly on the current outcomes as well as on the preceding assessment. Another example is survival analysis (e.g., Klein & Moeschberger, 2003), in which censoring is the common form of missing data problem. In practice, the missingness can come from different sources and can be modeled as a combination of different types of missingness. Second, various model selection criteria could be considered (e.g., Cain & Zhang, 2019). It is an interesting topic for future work to propose new criteria. For example, observed-data and complete-data likelihood functions for random effects models can be used for $f(\mathbf{y}|\theta)$; information criterion can be proposed using other weighted combination of the growth model and the missing data model. Third, the data considered in the study are assumed to be normally distributed. However, in reality, data are seldom normally distributed, particularly in behavioral and educational sciences (e.g., Cain, Zhang, & Yuan, 2017; Micceri, 1989). When data have heavy tails, or contaminated with outliers, robust models (e.g., Hoaglin, Mosteller, & Tukey, 1983; Huber, 1996; Zhang, 2013; Zhang, Lai, Lu, & Tong, 2013) should be adopted to make models insensitive to small deviations from the assumption of normal distribution. Fourth, latent population heterogeneity (e.g., McLachlan & Peel, 2000) may exist in the collected longitudinal data. Growth mixture models (GMMs) can be considered to provide a flexible set of models for analyzing longitudinal data with latent or mixture distributions (e.g., Bartholomew & Knott, 1999).

References

Acock, A. C. (2005). Working with missing values. *Journal of Marriage and Family*, *67*, 1012-1028. doi: https://doi.org/10.1111/j.1741-3737.2005.00191.x

Allison, P. D. (2002). *Missing data (sage university papers series on quantitative applications in the social sciences, 07-136)*. Thousands Oaks, CA: sage: Thousand Oaks, CA: SAGE publications, Inc.

Baltes, P. B., & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (p. 1-39). New York, NY: Academic Press.

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology*, *48*, 5-37. doi: https://doi.org/10.1016/j.jsp.2009.10.001

Bartholomew, D. J., & Knott, M. (1999). *Latent variable models and factor analysis: Kendall's library of statistics* (2nd ed., Vol. 7). New York, NY: Edward Arnold.

Box, G. E. P., & Tiao, G. C. (1973). *Bayesian inference in statistical analysis*. Hoboken, NJ: John Wiley & Sons.

Cain, M. K., & Zhang, Z. (2019). Fit for a bayesian: An evaluation of ppp and dic for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(1), 39–50. doi: https://doi.org/10.1080/10705511.2018.1490648

Cain, M. K., Zhang, Z., & Yuan, K.-H. (2017). Univariate and multivariate skewness and kurtosis for measuring nonnormality: Prevalence, influence and estimation. *Behavior research methods*, *49*(5), 1716–1735. doi: https://doi.org/10.3758/s13428-016-0814-1

Casella, G., & George, E. I. (1992). Explaining the Gibbs sampler. *The American Statistician*, *46*(3), 167-174. doi: https://doi.org/10.1080/00031305.1992.10475878

Collins, L. (1991). Measurement in longitudinal research. In J. L. Horn & L. Collins (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 137–148). Washington, DC: American Psychological Association.

Congdon, P. (2003). *Applied Bayesian modelling*. Chichester, UK: John Wiley & Sons.

Dawid, A. P. (1979). Conditional independence in statistical theory. *Journal of the Royal Statistical Society. Series B (Methodological)*, *41*, 1-31. doi: https://doi.org/10.1111/j.2517-6161.1979.tb01052.x

Demidenko, E. (2004). *Mixed models: Theory and applications*. New York, NY: Wiley-Interscience.

Diggle, P., & Kenward, M. G. (1994). Informative drop-out in longitudinal data analysis. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *43*, 49-93. doi: https://doi.org/10.2307/2986113

Dunson, D. B. (2000). Bayesian latent variable models for clustered mixed outcomes. *Journal of the Royal Statistical Society, B*, *62*, 355-366. doi: https://doi.org/10.1111/1467-9868.00236

Enders, C. K. (2011a). Analyzing longitudinal data with missing values. *Rehabilitation Psychology*, *56*, 1-22. doi: https://doi.org/10.1037/a0025579

Enders, C. K. (2011b). Missing not at random models for latent growth curve analyses. *Psychological Methods*, *16*, 1-16. doi: https://doi.org/10.1037/a0022640

Fay, R. E. (1992). When are inferences from multiple imputation valid. In (Vol. 81, p. 227-232).

Fitzmaurice, G. M., Davidian, M., Verbeke, G., & Molenberghs, G. (Eds.). (2008). *Longitudinal data analysis*. Boca Raton, FL: Chapman & Hall/CRC Press.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis*. Hoboken, New Jersey: John Wiley & sons Inc.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *6*, 721-741. doi: https://doi.org/10.1016/b978-0-08-051581-6.50057-x

Gill, J. (2002). *Bayesian methods: A social and behavioral sciences approach*. Boca Raton, FL: Chapman & Hall/CRC.

Glynn, R. J., Laird, N. M., & Rubin, D. B. (1986). Drawing inferences from self-selected samples. In H. Wainer (Ed.), (p. 115-142). New York: Springer Verlag.

Hedeker, D., & Gibbons, R. D. (1997). Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods*, *2*, 64-78. doi: https://doi.org/10.1037/1082-989x.2.1.64

Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Hedges, L. V. (1994). Fixed effects models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis*. (p. 285-299). NY, US: Russell Sage Foundation.

Hoaglin, D. C., Mosteller, F., & Tukey, J. W. (1983). *Understanding robust and exploratory data analysis*. New York: John Wiley & Sons.

Huber, P. (1996). *Robust statistical procedures* (2nd ed.). Philadelphia: SIAM.

Jelicic, H., Phelps, E., & Lerner, R. M. (2009). Use of missing data methods in longitudinal studies: The persistence of bad practices in developmental psychology. *Developmental Psychology*, *45*, 1195-1199. doi: https://doi.org/10.1037/a0015665

Klein, J. P., & Moeschberger, M. L. (2003). *Survival analysis: Techniques for censored and truncated data* (2nd, Ed.). Springer.

Lee, S. Y. (2007). *Structural equation modeling: A Bayesian approach*. Chichester, UK: John Wiley & Sons.

Little, R. J. A. (1993). Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association*, *88*, 125-134. doi: https://doi.org/10.2307/2290705

Little, R. J. A. (1995). Modelling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association*, *90*, 1112-1121. doi: https://doi.org/10.1080/01621459.1995.10476615

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York, N.Y.: Wiley.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). New York, N.Y.: Wiley-Interscience.

Lu, Z., Zhang, Z., & Lubke, G. (2010). (Abstract) Bayesian inference for growth mixture models with non-ignorable missing data. *Multivariate Behavioral Research*, *45*, 1028-1029. doi: https://doi.org/10.1080/00273171.2010.534381

Lu, Z., Zhang, Z., & Lubke, G. (2011). Bayesian inference for growth mixture models with latent-class-dependent missing data. *Multivariate Behavioral Research*, *46*, 567-597. doi: https://doi.org/10.1080/00273171.2011.589261

Luke, D. A. (2004). *Multilevel modeling (quantitative applications in the social sciences).* Thousand Oaks, CA: Sage Publication, Inc.

McLachlan, G. J., & Peel, D. (2000). *Finite mixture models.* New York, NY: John Wiley & Sons.

Micceri, T. (1989). The unicorn, the normal curve and the other improbable creatures. *Psychological Bulletin*, *105*, 156-166. doi: https://doi.org/10.1037/0033-2909.105.1.156

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychological methods*, *17*(3), 313–335. doi: https://doi.org/10.1037/a0026802

Muthén, B., Asparouhov, T., Hunter, A. M., & Leuchter, A. F. (2011). Growth modeling with nonignorable dropout: Alternative analyses of the STAR*D antidepressant trial. *Psychological Methods*, *16*, 17-33. doi: https://doi.org/10.1037/a0022634

Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods*, *6*, 328-362. doi: https://doi.org/10.1177/1094428103254673

Robert, C. P., & Casella, G. (2004). *Monte Carlo statistical methods.* New York, NY: Springer Science+Business Media Inc.

Roth, P. L. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology*, *47*, 537-560. doi: https://doi.org/10.1111/j.1744-6570.1994.tb01736.x

Rubin, D. (1987). *Multiple imputation for nonresponse in surveys.* Wiley & Sons, New York.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data.* Boca Raton, FL: Chapman & Hall/CRC.

Schafer, J. L., & Graham, J. W.   (2002).   Missing data: Our view of the state of the art.   *Psychological Methods*, *7*(2), 147-177.   doi: https://doi.org/10.1037/1082-989x.7.2.147

Scheines, R., Hoijtink, H., & Boomsma, A. (1999). Bayesian estimation and testing of structural equation models.   *Psychometrika*, *64*, 37-52.   doi: https://doi.org/10.1007/bf02294318

Schlomer, G. L., Bauman, S., & Card, N. A.   (2010).   Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology*, *57*, 1-10. doi: https://doi.org/10.1037/a0018082

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

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, *82*, 528-540. doi: https://doi.org/10.1080/01621459.1987.10478458

Yuan, K.-H., & Lu, Z. (2008). SEM with missing data and unknown population using two-stage ML: theory and its application. *Multivariate Behavioral Research*, *43*, 621-652. doi: https://doi.org/10.1080/00273170802490699

Yuan, K.-H., & Zhang, Z. (2012). Robust structural equation modeling with missing data and auxiliary variables. *Psychometrika*, *77(4)*, 803-826. doi: https://doi.org/10.1007/s11336-012-9282-4

Zhang, Z. (2013). Bayesian growth curve models with the generalized error distribution.   *Journal of Applied Statistics*, *40*(8), 1779–1795.   doi: https://doi.org/10.1080/02664763.2013.796348

Zhang, Z., Hamagami, F., Wang, L., Grimm, K. J., & Nesselroade, J. R. (2007). Bayesian analysis of longitudinal data using growth curve models. *International Journal of Behavioral Development*, *31*(4), 374-383. doi: https://doi.org/10.1177/0165025407077764

Zhang, Z., Lai, K., Lu, Z., & Tong, X.   (2013).   Bayesian inference and application of robust growth curve models using student's t distribution.   *Structural Equation Modeling*, *20(1)*, 47-78.   doi: https://doi.org/10.1080/10705511.2013.742382

Zhang, Z., & Wang, L. (2012). A note on the robustness of a full bayesian method for non-ignorable missing data analysis. *Brazilian Journal of Probability and Statistics*, *26(3)*, 244-264. doi: https://doi.org/10.1214/10-bjps132

Appendix

Appendix A. The Derived Posteriors for LGCMs with Non-ignorable Missingness

Let $\boldsymbol{\eta} = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_N)$, and the conditional posterior distribution for ϕ can be easily derived as an Inverse Gamma distribution,

$$\phi | \boldsymbol{\eta}, \mathbf{y} \sim IG\left(a_1/2, b_1/2\right), \tag{8}$$

where $a_1 = v_0 + NT$, and $b_1 = s_0 + \sum_{i=1}^{N}(\mathbf{y}_i - \mathbf{\Lambda}\boldsymbol{\eta}_i)'(\mathbf{y}_i - \mathbf{\Lambda}\boldsymbol{\eta}_i)$.

Notice that $tr(\mathbf{AB}) = tr(\mathbf{BA})$, so the conditional posterior distribution for $\boldsymbol{\Psi}$ is derived as an Inverse Wishart distribution,

$$\boldsymbol{\Psi}|\boldsymbol{\beta},\boldsymbol{\eta} \sim IW\left(m_1, \mathbf{V}_1\right), \tag{9}$$

where $m_1 = m_0 + N$, and $\mathbf{V}_1 = \mathbf{V}_0 + \sum_{i=1}^{N}(\boldsymbol{\eta}_i - \boldsymbol{\beta})(\boldsymbol{\eta}_i - \boldsymbol{\beta})'$.

By expanding the terms inside the exponential part and combining similar terms, the conditional posterior distribution for $\boldsymbol{\beta}$ is derived as a multivariate normal distribution,

$$\boldsymbol{\beta}|\boldsymbol{\Psi},\boldsymbol{\eta} \sim MN(\boldsymbol{\beta}_1, \boldsymbol{\Sigma}_1), \tag{10}$$

where $\boldsymbol{\beta}_1 = \left(N\boldsymbol{\Psi}^{-1} + \boldsymbol{\Sigma}_0^{-1}\right)^{-1}\left(\boldsymbol{\Psi}^{-1}\sum_{i=1}^{N}\boldsymbol{\eta}_i + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\beta}_0\right)$, and $\boldsymbol{\Sigma}_1 = \left(N\boldsymbol{\Psi}^{-1} + \boldsymbol{\Sigma}_0^{-1}\right)^{-1}$.

The conditional posterior for $\boldsymbol{\gamma}_t$, $(t = 1, 2, \ldots, T)$, is a distribution of

$$\begin{aligned} f(\boldsymbol{\gamma}_t|\boldsymbol{\omega}, \mathbf{x}, \mathbf{m}) \propto \exp\bigg[& -\frac{1}{2}(\boldsymbol{\gamma}_t - \boldsymbol{\gamma}_{t0})'\mathbf{D}_{t0}^{-1}(\boldsymbol{\gamma}_t - \boldsymbol{\gamma}_{t0}) \\ & + \sum_{i=1}^{N}\{m_{it}\log\Phi(\boldsymbol{\omega}_i'\boldsymbol{\gamma}_t) + (1 - m_{it})\log[1 - \Phi(\boldsymbol{\omega}_i'\boldsymbol{\gamma}_t)]\}\bigg]. \end{aligned} \tag{11}$$

where $\Phi(\boldsymbol{\omega}_i'\boldsymbol{\gamma}_t)$ is defined by Equation (4), (5), or (6).

By expanding the terms inside the exponential part and combining similar terms, the conditional posterior distribution for $\boldsymbol{\eta}_i$, $i = 1, 2, \ldots, N$, is derived as a Multivariate Normal distribution,

$$\boldsymbol{\eta}_i|\phi, \boldsymbol{\Psi}, \boldsymbol{\beta}, \mathbf{y}_i \sim MN(\boldsymbol{\mu}_{\eta i}, \boldsymbol{\Sigma}_{\eta i}), \tag{12}$$

where $\boldsymbol{\mu}_{\eta i} = \left(\frac{1}{\phi}\mathbf{\Lambda}'\mathbf{\Lambda} + \boldsymbol{\Psi}^{-1}\right)^{-1}\left(\frac{1}{\phi}\mathbf{\Lambda}'\mathbf{y}_i + \boldsymbol{\Psi}^{-1}\boldsymbol{\beta}\right)$, and $\boldsymbol{\Sigma}_{\eta i} = \left(\frac{1}{\phi}\mathbf{\Lambda}'\mathbf{\Lambda} + \boldsymbol{\Psi}^{-1}\right)^{-1}$.

The conditional posterior distribution for the missing data $\mathbf{y}_i^{mis}$, $i = 1, 2, \ldots, N$, is a normal distribution,

$$\mathbf{y}_i^{mis}|\boldsymbol{\eta}_i, \phi \sim MN\left[\mathbf{\Lambda}\boldsymbol{\eta}_i, \mathbf{I}\phi\right], \tag{13}$$

where $\mathbf{I}$ is a $T \times T$ identity matrix. The dimension and location of $\mathbf{y}_i^{mis}$ depend on the corresponding $\mathbf{m}_i$ value.

Appendix B. Simulation Results

Table 3. Summarized Estimates from True Model: LGCM with LSD Missingness (XS). N=1000 (convergence rate: $100/100 = 100\%$)

	para.[1]	true[2]	est.[3]	BIAS smp.[4]	rel.[5]	SE emp.[6]	avg.[7]	MSE[8]	CI[9] lower[10]	upper[11]	cover[12]	HPD[13] lower	upper	cover
Growth Curve	I	1	0.998	-0.002	-0.002	0.05	0.053	0.005	0.894	1.101	0.99	0.894	1.101	0.98
	S	3	3.003	0.003	0.001	0.079	0.077	0.012	2.853	3.155	0.97	2.853	3.154	0.96
	var(I)	1	1.011	0.011	0.011	0.105	0.102	0.022	0.82	1.22	0.94	0.814	1.213	0.94
	var(S)	4	3.99	-0.01	-0.003	0.232	0.232	0.107	3.56	4.468	0.94	3.545	4.449	0.93
	cov(IS)	0	0.001	0.001	0.001	0.119	0.112	0.026	-0.221	0.217	0.94	-0.218	0.218	0.94
	var(e)	1	1	0	0	0.043	0.042	0.004	0.92	1.086	0.92	0.918	1.084	0.93
Missingness Parameters — Wave 1	γ_{01}	-1	-1.025	-0.025	0.025	0.184	0.174	0.065	-1.375	-0.694	0.93	-1.365	-0.69	0.94
	γ_{x1}	-1.5	-1.541	-0.041	0.027	0.138	0.123	0.036	-1.795	-1.314	0.92	-1.783	-1.307	0.93
	γ_{S1}	0.5	0.515	0.015	0.03	0.066	0.062	0.008	0.4	0.641	0.9	0.397	0.636	0.92
Wave 2	γ_{02}	-1	-1.038	-0.038	0.038	0.191	0.171	0.067	-1.385	-0.714	0.96	-1.376	-0.711	0.97
	γ_{x2}	-1.5	-1.551	-0.051	0.034	0.129	0.119	0.034	-1.798	-1.33	0.95	-1.786	-1.323	0.94
	γ_{S2}	0.5	0.521	0.021	0.042	0.066	0.06	0.008	0.41	0.643	0.95	0.408	0.639	0.94
Wave 3	γ_{03}	-1	-1.067	-0.067	0.067	0.186	0.172	0.069	-1.417	-0.741	0.94	-1.407	-0.737	0.94
	γ_{x3}	-1.5	-1.557	-0.057	0.038	0.117	0.116	0.03	-1.796	-1.341	0.97	-1.785	-1.334	0.97
	γ_{S3}	0.5	0.529	0.029	0.058	0.063	0.058	0.008	0.42	0.648	0.89	0.418	0.643	0.91
Wave 4	γ_{04}	-1	-1.034	-0.034	0.034	0.18	0.173	0.063	-1.384	-0.709	0.94	-1.374	-0.704	0.93
	γ_{x4}	-1.5	-1.539	-0.039	0.026	0.122	0.114	0.029	-1.773	-1.325	0.95	-1.763	-1.319	0.94
	γ_{S4}	0.5	0.514	0.014	0.027	0.058	0.057	0.007	0.407	0.63	0.95	0.405	0.625	0.95

Note. The results are summarized based on 100 converged replications with a convergence rate of $100/100 = 100\%$. [1]The estimated parameter. [2]The true value of the corresponding parameter. [3]The parameter estimate, defined by $\text{est.}_{\cdot j} = \bar{\hat{\theta}}_j = \sum_{i=1}^{100} \hat{\theta}_{ij}/100$. [4]The simple bias, defined by $\text{BIAS.smp}_j = \bar{\hat{\theta}}_j - \theta_j$. [5]The relative bias, defined by $\text{BIAS.rel}_j = (\bar{\hat{\theta}}_j - \theta_j)/\theta_j$ when $\theta_j \neq 0$ and $\text{BIAS.rel}_j = \bar{\hat{\theta}}_j - \theta_j$ when $\theta_j = 0$. [6]The empirical standard errors, defined by $\text{SE.emp}_j = \sqrt{\sum_{i=1}^{100}(\hat{\theta}_{ij} - \bar{\hat{\theta}}_j)^2/99}$. [7]The average standard errors, defined by $\text{SE.avg}_j = \sum_{i=1}^{100} \hat{s}_{ij}/100$. [8]The mean square error, defined by $\text{MSE}_j = \sum_{i=1}^{100} \text{MSE}_{ij}/100$, where $\text{MSE}_{ij} = (\text{Bias}_{ij})^2 + (\hat{s}_{ij})^2$. [9]For percentile confidence interval. [10]The average lower 2.5% percentile. [11]The average upper 97.5% percentile. [12]The average 95% coverage of percentile confidence interval. [13]The lower,upper bounds, and coverage for HPD interval.

Table 4. Summarized Estimates from True Model: LGCM with LSD Missingness (XS) (con't)

para.[1]	true[2]	est.[3]	smp.[4] (BIAS)	rel.[5]	emp.[6] (SE)	avg.[7]	MSE[8]	lower[10] (CI)	upper[11]	cover[12]	lower (HPD[13])	upper	cover
						N=500 (convergence rate: 100/100 = 100%)							
Growth Curve													
I	1	0.986	-0.014	-0.014	0.076	0.074	0.011	0.841	1.132	0.93	0.841	1.132	0.95
S	3	3.001	0.001	0	0.097	0.109	0.021	2.789	3.216	0.97	2.788	3.213	0.97
var(I)	1	0.976	-0.024	-0.024	0.146	0.144	0.042	0.712	1.274	0.97	0.7	1.26	0.97
var(S)	4	4.001	0.001	0	0.388	0.329	0.258	3.403	4.691	0.9	3.373	4.652	0.9
cov(IS)	0	-0.009	-0.009	-0.009	0.155	0.157	0.049	-0.324	0.294	0.96	-0.319	0.297	0.96
var(e)	1	1.014	0.014	0.014	0.06	0.061	0.007	0.901	1.141	0.96	0.897	1.136	0.96
Missingness Parameters — Wave 1													
γ_{01}	-1	-1.082	-0.082	0.082	0.254	0.255	0.137	-1.609	-0.608	0.95	-1.587	-0.596	0.97
γ_{x1}	-1.5	-1.606	-0.106	0.071	0.181	0.186	0.079	-2.002	-1.275	0.95	-1.975	-1.258	0.97
γ_{S1}	0.5	0.54	0.04	0.081	0.083	0.092	0.017	0.375	0.735	0.95	0.368	0.722	0.94
Wave 2													
γ_{02}	-1	-1.096	-0.096	0.096	0.281	0.252	0.152	-1.61	-0.624	0.89	-1.591	-0.615	0.89
γ_{x2}	-1.5	-1.615	-0.115	0.077	0.204	0.18	0.088	-1.996	-1.291	0.91	-1.971	-1.275	0.94
γ_{S2}	0.5	0.546	0.046	0.092	0.104	0.088	0.021	0.385	0.73	0.87	0.379	0.719	0.88
Wave 3													
γ_{03}	-1	-1.068	-0.068	0.068	0.32	0.248	0.169	-1.572	-0.602	0.93	-1.555	-0.594	0.93
γ_{x3}	-1.5	-1.613	-0.113	0.075	0.279	0.174	0.123	-1.978	-1.295	0.9	-1.958	-1.283	0.93
γ_{S3}	0.5	0.536	0.036	0.072	0.116	0.084	0.022	0.381	0.71	0.92	0.378	0.702	0.91
Wave 4													
γ_{04}	-1	-1.123	-0.123	0.123	0.261	0.257	0.15	-1.652	-0.647	0.94	-1.628	-0.633	0.95
γ_{x4}	-1.5	-1.579	-0.079	0.053	0.174	0.168	0.066	-1.933	-1.274	0.95	-1.913	-1.261	0.96
γ_{S4}	0.5	0.543	0.043	0.086	0.089	0.085	0.017	0.388	0.719	0.92	0.382	0.71	0.92
						N=300 (convergence rate: 100/100 = 100%)							
Growth Curve													
I	1	1.001	0.001	0.001	0.104	0.097	0.02	0.81	1.192	0.89	0.811	1.192	0.89
S	3	2.984	-0.016	-0.005	0.149	0.14	0.042	2.712	3.262	0.93	2.71	3.259	0.93
var(I)	1	1.014	0.014	0.014	0.183	0.19	0.07	0.673	1.418	0.96	0.654	1.392	0.96
var(S)	4	3.975	-0.025	-0.006	0.416	0.425	0.354	3.22	4.886	0.96	3.174	4.82	0.96
cov(IS)	0	0.054	0.054	0.054	0.212	0.205	0.09	-0.359	0.449	0.94	-0.351	0.454	0.93
var(e)	1	1.011	0.011	0.011	0.073	0.08	0.012	0.867	1.179	0.96	0.86	1.17	0.96
Missingness Parameters — Wave 1													
γ_{01}	-1	-1.094	-0.094	0.094	0.341	0.345	0.249	-1.822	-0.468	0.97	-1.778	-0.441	0.97
γ_{x1}	-1.5	-1.65	-0.15	0.1	0.265	0.253	0.162	-2.209	-1.217	0.92	-2.155	-1.185	0.94
γ_{S1}	0.5	0.548	0.048	0.097	0.121	0.124	0.033	0.331	0.82	0.97	0.318	0.794	0.97
Wave 2													
γ_{02}	-1	-1.106	-0.106	0.106	0.452	0.34	0.341	-1.819	-0.486	0.93	-1.782	-0.467	0.93
γ_{x2}	-1.5	-1.692	-0.192	0.128	0.345	0.253	0.23	-2.243	-1.254	0.89	-2.196	-1.227	0.9
γ_{S2}	0.5	0.566	0.066	0.132	0.158	0.121	0.046	0.354	0.827	0.93	0.343	0.807	0.92
Wave 3													
γ_{03}	-1	-1.139	-0.139	0.139	0.397	0.335	0.293	-1.845	-0.527	0.91	-1.801	-0.503	0.92
γ_{x3}	-1.5	-1.648	-0.148	0.099	0.305	0.236	0.175	-2.152	-1.233	0.86	-2.115	-1.21	0.92
γ_{S3}	0.5	0.566	0.066	0.132	0.141	0.115	0.038	0.361	0.811	0.9	0.352	0.794	0.91
Wave 4													
γ_{04}	-1	-1.217	-0.217	0.217	0.411	0.356	0.347	-1.976	-0.576	0.9	-1.932	-0.552	0.9
γ_{x4}	-1.5	-1.681	-0.181	0.121	0.263	0.241	0.163	-2.203	-1.257	0.9	-2.161	-1.231	0.92
γ_{S4}	0.5	0.583	0.083	0.165	0.138	0.118	0.041	0.372	0.839	0.88	0.363	0.82	0.91

Note. The same as Table 3

Table 5. Summarized Estimates from True Model: LGCM with LSD Missingness (XS) (con't)

	para.	true	est.	BIAS smp.	BIAS rel.	SE emp.	SE avg.	MSE	CI lower	CI upper	CI cover	HPD lower	HPD upper	HPD cover
							N=200 (convergence rate: $100/106 \approx 94.34\%$)							
Growth Curve	I	1	1.011	0.011	0.011	0.099	0.119	0.024	0.779	1.244	0.98	0.779	1.243	0.98
	S	3	2.975	-0.025	-0.008	0.177	0.171	0.061	2.643	3.314	0.93	2.642	3.312	0.94
	var(I)	1	1.011	0.011	0.011	0.228	0.233	0.107	0.601	1.516	0.94	0.572	1.476	0.92
	var(S)	4	4	0	0	0.474	0.522	0.498	3.095	5.135	0.97	3.029	5.041	0.96
	cov(IS)	0	0.065	0.065	0.065	0.257	0.252	0.134	-0.447	0.549	0.92	-0.436	0.557	0.92
	var(e)	1	1.027	0.027	0.027	0.098	0.099	0.02	0.851	1.238	0.95	0.84	1.224	0.95
Wave 1	γ_{01}	-1	-1.3	-0.3	0.3	0.671	0.5	0.901	-2.399	-0.449	0.93	-2.306	-0.402	0.94
	γ_{x1}	-1.5	-1.874	-0.374	0.249	0.745	0.424	1.113	-2.868	-1.227	0.88	-2.735	-1.169	0.91
	γ_{S1}	0.5	0.647	0.147	0.293	0.323	0.197	0.202	0.334	1.1	0.91	0.311	1.045	0.92
Wave 2	γ_{02}	-1	-1.278	-0.278	0.278	0.69	0.468	0.838	-2.303	-0.463	0.87	-2.227	-0.426	0.89
	γ_{x2}	-1.5	-1.779	-0.279	0.186	0.456	0.349	0.451	-2.578	-1.209	0.91	-2.487	-1.163	0.9
	γ_{S2}	0.5	0.627	0.127	0.254	0.244	0.171	0.117	0.343	1.014	0.9	0.324	0.976	0.91
Wave 3	γ_{03}	-1	-1.191	-0.191	0.191	0.505	0.436	0.5	-2.133	-0.419	0.91	-2.05	-0.377	0.93
	γ_{x3}	-1.5	-1.721	-0.221	0.147	0.502	0.314	0.426	-2.428	-1.193	0.9	-2.348	-1.15	0.94
	γ_{S3}	0.5	0.586	0.086	0.172	0.183	0.152	0.068	0.326	0.926	0.91	0.309	0.889	0.95
Wave 4	γ_{04}	-1	-1.27	-0.27	0.27	0.594	0.467	0.67	-2.304	-0.457	0.86	-2.209	-0.404	0.9
	γ_{x4}	-1.5	-1.808	-0.308	0.205	0.397	0.336	0.382	-2.56	-1.24	0.82	-2.48	-1.195	0.89
	γ_{S4}	0.5	0.618	0.118	0.236	0.204	0.16	0.085	0.345	0.98	0.88	0.325	0.942	0.89
							N=100 (convergence rate: $100/142 \approx 70.42\%$)							
Growth Curve	I	1	1.031	0.031	0.031	0.167	0.168	0.057	0.701	1.359	0.96	0.701	1.359	0.97
	S	3	2.983	-0.017	-0.006	0.236	0.242	0.115	2.514	3.467	0.95	2.51	3.46	0.94
	var(I)	1	0.933	-0.067	-0.067	0.305	0.323	0.206	0.408	1.665	0.93	0.355	1.574	0.91
	var(S)	4	3.965	-0.035	-0.009	0.829	0.747	1.261	2.743	5.656	0.91	2.623	5.458	0.91
	cov(IS)	0	0.069	0.069	0.069	0.333	0.357	0.246	-0.666	0.748	0.93	-0.646	0.762	0.95
	var(e)	1	1.078	0.078	0.078	0.157	0.151	0.054	0.82	1.409	0.93	0.801	1.38	0.94
Wave 1	γ_{01}	-1	-3.257	-2.257	2.257	5.794	1.333	42.792	-6.264	-1.131	0.84	-5.922	-1.018	0.86
	γ_{x1}	-1.5	-4.314	-2.814	1.876	7.492	1.277	69.337	-7.171	-2.396	0.8	-6.739	-2.251	0.85
	γ_{S1}	0.5	1.626	1.126	2.252	2.881	0.55	10.353	0.788	2.857	0.8	0.746	2.698	0.84
Wave 2	γ_{02}	-1	-3.011	-2.011	2.011	5.719	1.322	41.711	-6.062	-1.027	0.85	-5.696	-0.893	0.88
	γ_{x2}	-1.5	-3.772	-2.272	1.515	6.947	1.283	61.237	-6.811	-1.927	0.82	-6.385	-1.774	0.85
	γ_{S2}	0.5	1.436	0.936	1.871	2.57	0.549	8.564	0.653	2.71	0.81	0.586	2.527	0.86
Wave 3	γ_{03}	-1	-2.877	-1.877	1.877	5.93	1.2	42.401	-5.493	-0.898	0.89	-5.233	-0.806	0.91
	γ_{x3}	-1.5	-3.86	-2.36	1.573	6.955	1.153	58.835	-6.508	-2.086	0.83	-6.125	-1.932	0.85
	γ_{S3}	0.5	1.388	0.888	1.776	2.567	0.467	7.977	0.641	2.428	0.85	0.596	2.289	0.89
Wave 4	γ_{04}	-1	-2.831	-1.831	1.831	5.646	1.297	39.835	-5.902	-0.891	0.89	-5.522	-0.753	0.90
	γ_{x4}	-1.5	-3.386	-1.886	1.257	5.379	1.127	37.532	-6.048	-1.745	0.81	-5.622	-1.586	0.88
	γ_{S4}	0.5	1.222	0.722	1.444	1.944	0.457	4.854	0.552	2.312	0.84	0.491	2.152	0.88

Note. Abbreviations are as given in Table 3.

Table 6. Summarized Estimates from LGCM with LID Missingness (XI)

	para.	true	est.	BIAS smp.	BIAS rel.	SE emp.	SE avg.	MSE	CI lower	CI upper	CI cover	HPD lower	HPD upper	HPD cover
					N=1000 (convergence rate: $100/112 \approx 89.29\%$)									
Growth Curve	I	1	1.064	0.064	0.064	0.052	0.044	0.009	0.977	1.151	0.66	0.977	1.150	0.66
	S	3	2.921	-0.079	-0.026	0.082	0.074	0.018	2.776	3.067	0.77	2.777	3.066	0.76
	var(I)	1	0.169	-0.831	-0.831	0.036	0.031	0.693	0.117	0.237	0	0.113	0.230	0
	var(S)	4	3.494	-0.506	-0.126	0.218	0.203	0.344	3.116	3.913	0.40	3.103	3.897	0.37
	cov(IS)	0	0.629	0.629	0.629	0.064	0.064	0.404	0.511	0.762	0	0.507	0.756	0
	var(e)	1	1.439	0.439	0.439	0.049	0.050	0.197	1.343	1.540	0	1.341	1.538	0
Missingness Parameters — Wave 1	γ_{01}	NA	-2.411	NA	NA	0.476	0.442	NA	-3.325	-1.619	NA	-3.27	-1.612	NA
	γ_{x1}	NA	-1.632	NA	NA	0.178	0.152	NA	-1.961	-1.362	NA	-1.939	-1.35	NA
	γ_{I1}	NA	2.794	NA	NA	0.479	0.442	NA	2.011	3.72	NA	2.009	3.661	NA
Wave 2	γ_{02}	NA	-2.439	NA	NA	0.543	0.444	NA	-3.395	-1.667	NA	-3.321	-1.646	NA
	γ_{x2}	NA	-1.644	NA	NA	0.163	0.148	NA	-1.962	-1.382	NA	-1.938	-1.368	NA
	γ_{I2}	NA	2.826	NA	NA	0.546	0.437	NA	2.074	3.762	NA	2.045	3.682	NA
Wave 3	γ_{03}	NA	-2.442	NA	NA	0.492	0.426	NA	-3.336	-1.678	NA	-3.277	-1.662	NA
	γ_{x3}	NA	-1.632	NA	NA	0.137	0.143	NA	-1.934	-1.375	NA	-1.915	-1.364	NA
	γ_{I3}	NA	2.819	NA	NA	0.482	0.422	NA	2.063	3.718	NA	2.048	3.646	NA
Wave 4	γ_{04}	NA	-2.367	NA	NA	0.48	0.427	NA	-3.262	-1.596	NA	-3.199	-1.581	NA
	γ_{x4}	NA	-1.617	NA	NA	0.146	0.141	NA	-1.917	-1.362	NA	-1.896	-1.35	NA
	γ_{I4}	NA	2.733	NA	NA	0.448	0.422	NA	1.98	3.616	NA	1.978	3.569	NA
					N=500 (convergence rate: $100/118 \approx 84.75\%$)									
Growth Curve	I	1	1.060	0.060	0.060	0.076	0.063	0.013	0.938	1.186	0.78	0.937	1.184	0.79
	S	3	2.914	-0.086	-0.029	0.099	0.105	0.028	2.710	3.120	0.88	2.709	3.118	0.87
	var(I)	1	0.197	-0.803	-0.803	0.046	0.048	0.650	0.121	0.309	0	0.114	0.294	0
	var(S)	4	3.448	-0.552	-0.138	0.315	0.284	0.484	2.934	4.043	0.56	2.909	4.012	0.54
	cov(IS)	0	0.633	0.633	0.633	0.074	0.088	0.414	0.474	0.819	0	0.466	0.808	0
	var(e)	1	1.425	0.425	0.425	0.079	0.072	0.192	1.289	1.571	0	1.286	1.567	0
Missingness Parameters — Wave 1	γ_{01}	NA	-2.471	NA	NA	0.726	0.665	NA	-3.903	-1.333	NA	-3.789	-1.292	NA
	γ_{x1}	NA	-1.765	NA	NA	0.298	0.269	NA	-2.385	-1.329	NA	-2.31	-1.294	NA
	γ_{I1}	NA	2.898	NA	NA	0.716	0.671	NA	1.762	4.339	NA	1.718	4.215	NA
Wave 2	γ_{02}	NA	-2.393	NA	NA	0.77	0.631	NA	-3.746	-1.328	NA	-3.63	-1.281	NA
	γ_{x2}	NA	-1.723	NA	NA	0.265	0.239	NA	-2.257	-1.325	NA	-2.206	-1.297	NA
	γ_{I2}	NA	2.815	NA	NA	0.737	0.627	NA	1.759	4.162	NA	1.712	4.048	NA
Wave 3	γ_{03}	NA	-2.425	NA	NA	0.779	0.644	NA	-3.804	-1.337	NA	-3.681	-1.293	NA
	γ_{x3}	NA	-1.761	NA	NA	0.352	0.257	NA	-2.336	-1.343	NA	-2.271	-1.309	NA
	γ_{I3}	NA	2.858	NA	NA	0.796	0.647	NA	1.775	4.235	NA	1.729	4.104	NA
Wave 4	γ_{04}	NA	-2.396	NA	NA	0.782	0.655	NA	-3.86	-1.312	NA	-3.693	-1.24	NA
	γ_{x4}	NA	-1.687	NA	NA	0.294	0.24	NA	-2.223	-1.288	NA	-2.167	-1.259	NA
	γ_{I4}	NA	2.805	NA	NA	0.818	0.662	NA	1.713	4.275	NA	1.657	4.119	NA

Note. Abbreviations are as given in Table 3.

Table 7. Summarized Estimates from LGCM with LID Missingness (XI) (con't)

	para.	true	est.	BIAS smp.	BIAS rel.	SE emp.	SE avg.	MSE	CI lower	CI upper	CI cover	HPD lower	HPD upper	HPD cover
	N=300													

N=300 (convergence rate: 100/148 ≈ 67.57%)

	para.	true	est.	smp.	rel.	emp.	avg.	MSE	lower	upper	cover	lower	upper	cover
Growth Curve	I	1	1.077	0.077	0.077	0.11	0.083	0.025	0.916	1.242	0.78	0.915	1.24	0.81
	S	3	2.864	-0.136	-0.045	0.139	0.135	0.056	2.601	3.131	0.87	2.6	3.129	0.87
	var(I)	1	0.251	-0.749	-0.749	0.084	0.076	0.574	0.136	0.429	0.01	0.123	0.402	0.01
	var(S)	4	3.424	-0.576	-0.144	0.369	0.366	0.601	2.775	4.209	0.71	2.734	4.153	0.67
	cov(IS)	0	0.656	0.656	0.656	0.118	0.119	0.458	0.445	0.909	0	0.433	0.892	0
	var(e)	1	1.413	0.413	0.413	0.101	0.095	0.19	1.237	1.608	0	1.232	1.601	0
Missingness Parameters · Wave 1	γ_{01}	NA	-2.672	NA	NA	1.64	0.984	NA	-4.884	-1.166	NA	-4.637	-1.069	NA
	γ_{x1}	NA	-2.055	NA	NA	0.913	0.487	NA	-3.218	-1.343	NA	-3.058	-1.277	NA
	γ_{S1}	NA	3.108	NA	NA	1.68	1.008	NA	1.582	5.378	NA	1.502	5.129	NA
Wave 2	γ_{02}	NA	-2.768	NA	NA	3.488	0.966	NA	-4.978	-1.265	NA	-4.741	-1.185	NA
	γ_{x2}	NA	-2.243	NA	NA	2.502	0.507	NA	-3.445	-1.474	NA	-3.282	-1.402	NA
	γ_{S2}	NA	3.409	NA	NA	4.711	0.994	NA	1.895	5.668	NA	1.809	5.422	NA
Wave 3	γ_{03}	NA	-2.68	NA	NA	2.057	0.915	NA	-4.769	-1.249	NA	-4.567	-1.176	NA
	γ_{x3}	NA	-1.999	NA	NA	0.878	0.421	NA	-2.989	-1.348	NA	-2.861	-1.289	NA
	γ_{S3}	NA	3.118	NA	NA	1.884	0.936	NA	1.66	5.234	NA	1.59	5.008	NA
Wave 4	γ_{04}	NA	-2.907	NA	NA	2.499	0.941	NA	-4.948	-1.426	NA	-4.744	-1.353	NA
	γ_{x4}	NA	-2.204	NA	NA	1.651	0.498	NA	-3.333	-1.449	NA	-3.196	-1.39	NA
	γ_{S4}	NA	3.371	NA	NA	2.766	0.98	NA	1.875	5.511	NA	1.804	5.296	NA

N=200 (convergence rate: 100/197 ≈ 50.76%)

	para.	true	est.	smp.	rel.	emp.	avg.	MSE	lower	upper	cover	lower	upper	cover
Growth Curve	I	1	1.052	0.082	0.082	0.219	0.1	0.03	0.858	1.248	0.79	0.857	1.247	0.79
	S	3	2.796	-0.114	-0.038	0.525	0.161	0.071	2.484	3.114	0.85	2.483	3.112	0.85
	var(I)	1	0.322	-0.648	-0.648	0.15	0.115	0.469	0.15	0.593	0.1	0.13	0.549	0.07
	var(S)	4	3.353	-0.527	-0.132	0.739	0.435	0.677	2.6	4.302	0.74	2.546	4.225	0.71
	cov(IS)	0	0.617	0.617	0.617	0.276	0.147	0.479	0.352	0.93	0.01	0.338	0.91	0.01
	var(e)	1	1.346	0.376	0.376	0.267	0.115	0.174	1.135	1.586	0.07	1.126	1.574	0.08
Missingness Parameters · Wave 1	γ_{01}	NA	-2.974	NA	NA	5.157	1.352	NA	-6.046	-0.844	NA	-5.701	-0.814	NA
	γ_{x1}	NA	-3.03	NA	NA	3.647	0.986	NA	-5.376	-1.659	NA	-5.055	-1.54	NA
	γ_{S1}	NA	3.622	NA	NA	6.034	1.465	NA	1.414	6.896	NA	1.354	6.57	NA
Wave 2	γ_{02}	NA	-3.094	NA	NA	3.737	1.162	NA	-5.77	-1.267	NA	-5.452	-1.163	NA
	γ_{x2}	NA	-2.551	NA	NA	2.369	0.681	NA	-4.154	-1.547	NA	-3.906	-1.441	NA
	γ_{S2}	NA	3.652	NA	NA	4.116	1.194	NA	1.823	6.378	NA	1.709	6.06	NA
Wave 3	γ_{03}	NA	-2.198	NA	NA	4.971	1.179	NA	-4.869	-0.328	NA	-4.545	-0.211	NA
	γ_{x3}	NA	-2.705	NA	NA	3.501	0.746	NA	-4.405	-1.534	NA	-4.189	-1.452	NA
	γ_{S3}	NA	2.627	NA	NA	5.346	1.21	NA	0.723	5.342	NA	0.631	4.989	NA
Wave 4	γ_{04}	NA	-3.469	NA	NA	4.108	1.285	NA	-6.288	-1.421	NA	-6.014	-1.338	NA
	γ_{x4}	NA	-3.122	NA	NA	3.555	0.912	NA	-5.198	-1.739	NA	-4.895	-1.646	NA
	γ_{S4}	NA	4.199	NA	NA	4.895	1.378	NA	2.059	7.192	NA	1.978	6.86	NA

N=100 (unavailable due to low convergence rate)

Note. Abbreviations are as given in Table 3.

Table 8. Summarized Estimates from LGCM with LOD Missingness (XY)

	para.	true	est.	BIAS smp.	BIAS rel.	SE emp.	SE avg.	MSE	CI lower	CI upper	CI cover	HPD lower	HPD upper	HPD cover
							N=1000 (convergence rate: $100/126 \approx 79.37\%$)							
Growth Curve	I	1	1.12	0.12	0.12	0.062	0.06	0.022	1.002	1.238	0.52	1.002	1.237	0.51
	S	3	3.003	0.003	0.001	0.084	0.078	0.013	2.85	3.158	0.94	2.849	3.156	0.94
	var(I)	1	1.03	0.03	0.03	0.105	0.108	0.024	0.828	1.252	0.93	0.823	1.245	0.93
	var(S)	4	3.994	-0.006	-0.002	0.253	0.235	0.119	3.556	4.479	0.91	3.542	4.46	0.90
	cov(IS)	0	0.112	0.112	0.112	0.146	0.116	0.047	-0.118	0.337	0.74	-0.115	0.338	0.72
	var(e)	1	1.015	0.015	0.015	0.048	0.044	0.004	0.933	1.105	0.91	0.93	1.102	0.92
Wave 1	γ_{01}	NA	0.164	NA	NA	0.134	0.117	NA	-0.072	0.387	NA	-0.066	0.39	NA
	γ_{x1}	NA	-1.106	NA	NA	0.076	0.071	NA	-1.249	-0.973	NA	-1.244	-0.97	NA
	γ_{S1}	NA	0.156	NA	NA	0.093	0.073	NA	0.014	0.299	NA	0.014	0.297	NA
Wave 2	γ_{02}	NA	-1.156	NA	NA	0.185	0.196	NA	-1.557	-0.789	NA	-1.54	-0.781	NA
	γ_{x2}	NA	-1.468	NA	NA	0.127	0.108	NA	-1.69	-1.267	NA	-1.682	-1.262	NA
	γ_{S2}	NA	0.387	NA	NA	0.047	0.044	NA	0.304	0.477	NA	0.302	0.473	NA
Wave 3	γ_{03}	NA	-1.235	NA	NA	0.196	0.186	NA	-1.611	-0.88	NA	-1.602	-0.878	NA
	γ_{x3}	NA	-1.515	NA	NA	0.117	0.109	NA	-1.739	-1.311	NA	-1.731	-1.306	NA
	γ_{S3}	NA	0.241	NA	NA	0.028	0.025	NA	0.193	0.292	NA	0.193	0.291	NA
Wave 4	γ_{04}	NA	-1.179	NA	NA	0.189	0.182	NA	-1.547	-0.833	NA	-1.537	-0.831	NA
	γ_{x4}	NA	-1.511	NA	NA	0.111	0.109	NA	-1.735	-1.308	NA	-1.726	-1.302	NA
	γ_{S4}	NA	0.164	NA	NA	0.018	0.017	NA	0.131	0.2	NA	0.131	0.198	NA
							N=500 (convergence rate: $100/110 \approx 90.91\%$)							
Growth Curve	I	1	1.121	0.121	0.121	0.098	0.085	0.032	0.956	1.288	0.68	0.956	1.287	0.69
	S	3	3.008	0.008	0.003	0.107	0.11	0.024	2.793	3.226	0.95	2.793	3.224	0.94
	var(I)	1	1.004	0.004	0.004	0.146	0.152	0.044	0.725	1.322	0.96	0.714	1.308	0.95
	var(S)	4	3.996	-0.004	-0.001	0.399	0.334	0.27	3.391	4.698	0.86	3.365	4.662	0.86
	cov(IS)	0	0.102	0.102	0.102	0.178	0.163	0.069	-0.223	0.417	0.88	-0.219	0.42	0.89
	var(e)	1	1.026	0.026	0.026	0.062	0.063	0.008	0.91	1.156	0.95	0.906	1.151	0.94
Wave 1	γ_{01}	NA	0.131	NA	NA	0.2	0.175	NA	-0.228	0.459	NA	-0.214	0.467	NA
	γ_{x1}	NA	-1.143	NA	NA	0.119	0.107	NA	-1.366	-0.947	NA	-1.354	-0.939	NA
	γ_{S1}	NA	0.18	NA	NA	0.133	0.109	NA	-0.032	0.396	NA	-0.03	0.395	NA
Wave 2	γ_{02}	NA	-1.246	NA	NA	0.344	0.292	NA	-1.854	-0.709	NA	-1.822	-0.691	NA
	γ_{x2}	NA	-1.52	NA	NA	0.171	0.162	NA	-1.859	-1.225	NA	-1.841	-1.215	NA
	γ_{S2}	NA	0.409	NA	NA	0.078	0.066	NA	0.288	0.547	NA	0.285	0.539	NA
Wave 3	γ_{03}	NA	-1.282	NA	NA	0.332	0.278	NA	-1.858	-0.768	NA	-1.827	-0.753	NA
	γ_{x3}	NA	-1.577	NA	NA	0.257	0.166	NA	-1.922	-1.275	NA	-1.903	-1.262	NA
	γ_{S3}	NA	0.249	NA	NA	0.049	0.038	NA	0.18	0.327	NA	0.178	0.324	NA
Wave 4	γ_{04}	NA	-1.277	NA	NA	0.277	0.269	NA	-1.833	-0.774	NA	-1.808	-0.763	NA
	γ_{x4}	NA	-1.546	NA	NA	0.171	0.159	NA	-1.878	-1.255	NA	-1.861	-1.244	NA
	γ_{S4}	NA	0.174	NA	NA	0.028	0.026	NA	0.126	0.227	NA	0.125	0.224	NA

Note. Abbreviations are as given in Table 3.

Table 9. Summarized Estimates from LGCM with LOD Missingness (XY) (con't)

	para.	true	est.	BIAS smp.	BIAS rel.	SE emp.	SE avg.	MSE	CI lower	CI upper	CI cover	HPD lower	HPD upper	HPD cover
N=300														
Growth Curve	I	1	1.139	0.139	0.139	0.127	0.111	0.047	0.922	1.357	0.70	0.922	1.356	0.69
	S	3	2.988	-0.012	-0.004	0.157	0.144	0.046	2.708	3.274	0.90	2.707	3.272	0.90
	var(I)	1	1.045	0.045	0.045	0.196	0.204	0.082	0.682	1.479	0.94	0.661	1.451	0.95
	var(S)	4	4.04	0.04	0.01	0.463	0.44	0.41	3.261	4.982	0.92	3.212	4.915	0.95
	cov(IS)	0	0.153	0.153	0.153	0.256	0.215	0.135	-0.277	0.569	0.85	-0.27	0.574	0.84
	var(e)	1	1.021	0.021	0.021	0.079	0.082	0.013	0.873	1.195	0.94	0.865	1.184	0.94
	γ_{01}	NA	0.103	NA	NA	0.25	0.235	NA	-0.394	0.533	NA	-0.365	0.549	NA
	γ_{x1}	NA	-1.187	NA	NA	0.163	0.147	NA	-1.499	-0.924	NA	-1.478	-0.91	NA
	γ_{S1}	NA	0.201	NA	NA	0.178	0.144	NA	-0.075	0.491	NA	-0.075	0.487	NA
	γ_{02}	NA	-1.251	NA	NA	0.441	0.392	NA	-2.087	-0.547	NA	-2.028	-0.514	NA
	γ_{x2}	NA	-1.583	NA	NA	0.264	0.224	NA	-2.063	-1.189	NA	-2.026	-1.165	NA
	γ_{S2}	NA	0.416	NA	NA	0.098	0.088	NA	0.258	0.605	NA	0.251	0.591	NA
	γ_{03}	NA	-1.333	NA	NA	0.438	0.368	NA	-2.102	-0.666	NA	-2.058	-0.641	NA
	γ_{x3}	NA	-1.614	NA	NA	0.295	0.222	NA	-2.084	-1.218	NA	-2.054	-1.199	NA
	γ_{S3}	NA	0.259	NA	NA	0.063	0.05	NA	0.169	0.364	NA	0.166	0.358	NA
	γ_{04}	NA	-1.406	NA	NA	0.434	0.387	NA	-2.231	-0.712	NA	-2.169	-0.682	NA
	γ_{x4}	NA	-1.656	NA	NA	0.26	0.232	NA	-2.152	-1.245	NA	-2.117	-1.223	NA
	γ_{S4}	NA	0.188	NA	NA	0.042	0.037	NA	0.122	0.268	NA	0.119	0.261	NA
N=200														
Growth Curve	I	1	1.154	0.154	0.154	0.141	0.135	0.062	0.892	1.421	0.75	0.891	1.419	0.76
	S	3	2.986	-0.014	-0.005	0.187	0.176	0.066	2.648	3.336	0.92	2.644	3.331	0.93
	var(I)	1	1.019	0.019	0.019	0.233	0.25	0.117	0.583	1.562	0.96	0.552	1.517	0.97
	var(S)	4	4.034	0.034	0.008	0.516	0.536	0.557	3.107	5.202	0.96	3.043	5.11	0.96
	cov(IS)	0	0.182	0.182	0.182	0.311	0.263	0.199	-0.347	0.691	0.85	-0.338	0.697	0.85
	var(e)	1	1.047	0.047	0.047	0.103	0.104	0.024	0.863	1.27	0.92	0.852	1.255	0.94
	γ_{01}	NA	0.043	NA	NA	0.375	0.32	NA	-0.654	0.608	NA	-0.594	0.642	NA
	γ_{x1}	NA	-1.269	NA	NA	0.266	0.212	NA	-1.739	-0.911	NA	-1.69	-0.883	NA
	γ_{S1}	NA	0.227	NA	NA	0.272	0.197	NA	-0.148	0.631	NA	-0.15	0.616	NA
	γ_{02}	NA	-1.46	NA	NA	0.675	0.542	NA	-2.674	-0.532	NA	-2.541	-0.471	NA
	γ_{x2}	NA	-1.683	NA	NA	0.373	0.311	NA	-2.387	-1.165	NA	-2.303	-1.125	NA
	γ_{S2}	NA	0.463	NA	NA	0.153	0.123	NA	0.253	0.738	NA	0.24	0.707	NA
	γ_{03}	NA	-1.442	NA	NA	0.608	0.502	NA	-2.541	-0.573	NA	-2.457	-0.535	NA
	γ_{x3}	NA	-1.718	NA	NA	0.632	0.322	NA	-2.445	-1.19	NA	-2.377	-1.153	NA
	γ_{S3}	NA	0.271	NA	NA	0.091	0.069	NA	0.154	0.421	NA	0.148	0.41	NA
	γ_{04}	NA	-1.454	NA	NA	0.597	0.478	NA	-2.471	-0.601	NA	-2.4	-0.563	NA
	γ_{x4}	NA	-1.757	NA	NA	0.37	0.304	NA	-2.414	-1.227	NA	-2.364	-1.194	NA
	γ_{S4}	NA	0.196	NA	NA	0.056	0.046	NA	0.114	0.295	NA	0.11	0.287	NA
N=100														
Growth Curve	I	1	1.161	0.161	0.161	0.252	0.197	0.129	0.776	1.551	0.81	0.775	1.549	0.81
	S	3	3.028	0.028	0.009	0.254	0.259	0.133	2.535	3.548	0.97	2.528	3.539	0.97
	var(I)	1	0.937	-0.063	-0.063	0.332	0.354	0.246	0.375	1.751	0.92	0.315	1.637	0.90
	var(S)	4	4.136	0.136	0.034	0.845	0.809	1.414	2.817	5.971	0.93	2.686	5.757	0.94
	cov(IS)	0	0.15	0.15	0.15	0.453	0.394	0.39	-0.657	0.902	0.88	-0.633	0.918	0.88
	var(e)	1	1.153	0.153	0.153	0.34	0.176	0.184	0.847	1.529	0.86	0.825	1.494	0.89
	γ_{01}	NA	-0.711	NA	NA	5.806	1.29	NA	-3.446	1.305	NA	-3.079	1.381	NA
	γ_{x1}	NA	-3.4	NA	NA	6.975	1.599	NA	-7.14	-1.402	NA	-6.495	-1.259	NA
	γ_{S1}	NA	0.468	NA	NA	4.682	0.964	NA	-1.317	2.211	NA	-1.201	2.072	NA
	γ_{02}	NA	-3.803	NA	NA	9.225	1.602	NA	-7.571	-1.469	NA	-6.978	-1.269	NA
	γ_{x2}	NA	-3.378	NA	NA	6.747	1.042	NA	-5.814	-1.866	NA	-5.463	-1.739	NA
	γ_{S2}	NA	1.029	NA	NA	2.14	0.367	NA	0.49	1.883	NA	0.444	1.762	NA
	γ_{03}	NA	-3.148	NA	NA	6.553	1.128	NA	-5.681	-1.244	NA	-5.386	-1.19	NA
	γ_{x3}	NA	-3.039	NA	NA	4.774	0.759	NA	-4.767	-1.788	NA	-4.554	-1.735	NA
	γ_{S3}	NA	0.534	NA	NA	0.974	0.156	NA	0.275	0.888	NA	0.264	0.839	NA
	γ_{04}	NA	-2.582	NA	NA	4.213	1.27	NA	-5.474	-0.588	NA	-5.149	-0.48	NA
	γ_{x4}	NA	-3.259	NA	NA	5.963	0.989	NA	-5.478	-1.717	NA	-5.176	-1.583	NA
	γ_{S4}	NA	0.297	NA	NA	0.719	0.146	NA	0.064	0.611	NA	0.055	0.579	NA

N=300 (convergence rate: $100/107 \approx 93.46\%$)

N=200 (convergence rate: $100/104 \approx 96.15\%$)

N=100 (convergence rate: $100/138 \approx 72.46\%$)

Row groupings (left margin labels): Growth Curve; Missingness Parameters — Wave 1, Wave 2, Wave 3, Wave 4.

Note. Abbreviations are as given in Table 3 .

Table 10. Summarized Estimates from LGCM with Ignorable Missingness (X)

	para.	true	est.	BIAS smp.	rel.	SE emp.	avg.	MSE	CI lower	upper	cover	HPD lower	upper	cover
	N=1000 (convergence rate: 100/100 = 100%)													
Growth Curve	I	1	1.009	0.009	0.009	0.051	0.052	0.005	0.906	1.111	0.94	0.906	1.111	0.93
	S	3	2.711	-0.289	-0.096	0.078	0.077	0.095	2.56	2.863	0.04	2.561	2.863	0.04
	var(I)	1	1.008	0.008	0.008	0.108	0.104	0.022	0.813	1.221	0.95	0.807	1.214	0.95
	var(S)	4	3.837	-0.163	-0.041	0.232	0.223	0.13	3.422	4.297	0.87	3.409	4.279	0.86
	cov(IS)	0	0.004	0.004	0.004	0.115	0.109	0.025	-0.214	0.214	0.96	-0.21	0.216	0.96
	var(e)	1	0.999	-0.001	-0.001	0.044	0.043	0.004	0.919	1.086	0.92	0.917	1.084	0.92
	N=500 (convergence rate: 100/100 = 100%)													
Growth Curve	I	1	0.999	-0.001	-0.001	0.073	0.074	0.011	0.854	1.143	0.98	0.855	1.143	0.98
	S	3	2.711	-0.289	-0.096	0.099	0.109	0.105	2.497	2.925	0.21	2.497	2.925	0.21
	var(I)	1	0.973	-0.027	-0.027	0.146	0.146	0.043	0.705	1.277	0.98	0.693	1.263	0.98
	var(S)	4	3.852	-0.148	-0.037	0.371	0.317	0.259	3.276	4.518	0.86	3.248	4.48	0.88
	cov(IS)	0	-0.008	-0.008	-0.008	0.154	0.154	0.047	-0.317	0.287	0.96	-0.311	0.292	0.96
	var(e)	1	1.014	0.014	0.014	0.06	0.062	0.008	0.9	1.141	0.96	0.895	1.136	0.95
	N=300 (convergence rate: 100/100 = 100%)													
Growth Curve	I	1	1.009	0.009	0.009	0.103	0.096	0.02	0.821	1.197	0.89	0.821	1.197	0.89
	S	3	2.687	-0.313	-0.104	0.139	0.141	0.137	2.411	2.964	0.34	2.411	2.963	0.35
	var(I)	1	1.006	0.006	0.006	0.189	0.194	0.073	0.657	1.416	0.94	0.639	1.391	0.94
	var(S)	4	3.816	-0.184	-0.046	0.412	0.41	0.372	3.091	4.694	0.93	3.045	4.631	0.92
	cov(IS)	0	0.045	0.045	0.045	0.214	0.2	0.088	-0.359	0.429	0.94	-0.351	0.435	0.94
	var(e)	1	1.01	0.01	0.01	0.075	0.08	0.012	0.864	1.179	0.96	0.857	1.17	0.94
	N=200 (convergence rate: 100/100 = 100%)													
Growth Curve	I	1	1.019	0.019	0.019	0.098	0.116	0.023	0.792	1.247	0.97	0.791	1.246	0.97
	S	3	2.69	-0.31	-0.103	0.178	0.173	0.157	2.352	3.029	0.52	2.351	3.027	0.52
	var(I)	1	0.99	-0.01	-0.01	0.232	0.236	0.11	0.576	1.5	0.94	0.548	1.461	0.95
	var(S)	4	3.884	-0.116	-0.029	0.47	0.509	0.495	3.004	4.992	0.96	2.938	4.898	0.96
	cov(IS)	0	0.066	0.066	0.066	0.25	0.246	0.127	-0.434	0.538	0.91	-0.422	0.546	0.92
	var(e)	1	1.02	0.02	0.02	0.094	0.1	0.019	0.843	1.233	0.95	0.833	1.219	0.97
	N=100 (convergence rate: 100/100 = 100%)													
Growth Curve	I	1	1.031	0.031	0.031	0.174	0.161	0.057	0.714	1.348	0.94	0.715	1.348	0.95
	S	3	2.699	-0.301	-0.1	0.239	0.248	0.209	2.21	3.187	0.78	2.212	3.187	0.78
	var(I)	1	0.863	-0.137	-0.137	0.275	0.315	0.197	0.354	1.579	0.94	0.302	1.487	0.86
	var(S)	4	3.951	-0.049	-0.012	0.815	0.753	1.247	2.726	5.66	0.92	2.601	5.456	0.92
	cov(IS)	0	0.063	0.063	0.063	0.35	0.35	0.25	-0.658	0.728	0.91	-0.637	0.744	0.94
	var(e)	1	1.063	0.063	0.063	0.137	0.149	0.045	0.808	1.39	0.94	0.788	1.361	0.95

Note. Abbreviations are as given in Table 3.