Routledge
Taylor & Francis Group

Multivariate Behavioral Research

ISSN: 0027-3171 (Print) 1532-7906 (Online) Journal homepage: www.tandfonline.com/journals/hmbr20

A Two-Step Estimator for Growth Mixture Models with Covariates in the Presence of Direct Effects

Yuqi Liu, Zsuzsa Bakk, Ethan M. McCormick & Mark de Rooij

To cite this article: Yuqi Liu, Zsuzsa Bakk, Ethan M. McCormick & Mark de Rooij (22 Oct 2025): A Two-Step Estimator for Growth Mixture Models with Covariates in the Presence of Direct Effects, Multivariate Behavioral Research, DOI: 10.1080/00273171.2025.2557275

To link to this article: https://doi.org/10.1080/00273171.2025.2557275

Published online: 22 Oct 2025.

MULTIVARIATE BEHAVIORAL RESEARCH
https://doi.org/10.1080/00273171.2025.2557275

A Two-Step Estimator for Growth Mixture Models with Covariates in the Presence of Direct Effects

Yuqi Liu, Zsuzsa Bakk, Ethan M. McCormick, and Mark de Rooij

Methodology and Statistics Unit, Institute of Psychology, Leiden University, Leiden, The Netherlands

ABSTRACT
Growth mixture models (GMMs) are popular approaches for modeling unobserved population heterogeneity over time. GMMs can be extended with covariates, predicting latent class (LC) membership, the within-class growth trajectories, or both. However, current estimators are sensitive to misspecifications in complex models. We propose extending the two-step estimator for LC models to GMMs, which provides robust estimation against model misspecifications (namely, ignored and overfitted the direct effects) for simpler LC models. We conducted several simulation studies, comparing the performance of the proposed two-step estimator to the commonly-used one- and three-step estimators. Three different population models were considered, including covariates that predicted only the LC membership (I), adding direct effects to the latent intercept (II), or to both growth factors (III). Results show that when predicting LC membership alone, all three estimators are unbiased when the measurement model is strong, with weak measurement model results being more nuanced. Alternatively, when including covariate effects on the growth factors, the two-step, and three-step estimators show consistent robustness against misspecifications with unbiased estimates across simulation conditions while tending to underestimate the standard error estimates while the one-step estimator is most sensitive to misspecifications.

KEYWORDS
Growth mixture model; two-step; estimator; direct effects; covariates

Introduction

Growth mixture models (GMMs; B. Muthén & Shedden, 1999) are statistical models that can be used to identify distinctive growth trajectories within a heterogeneous population, which have been widely used in applied research. For example, Bowers and Sprott (2012) performed GMMs to evaluate the change over time in school achievement profiles of students, and Chen et al. (2024) utilized GMMs to identify distinctive growth trajectories of cognitive function among aging citizens with diabetes in China.

In GMMs, a categorical latent class (LC) variable captures the growth trajectories of unobserved subpopulations, relaxing the single homogeneous population assumption of the simpler latent curve model (LCM; Meredith & Tisak, 1990). GMMs can be used to capture a variety of linear and nonlinear growth trajectories. In this paper, we will focus on the commonly used linear growth pattern, in which the growth trajectory is captured by two continuous growth factors, namely the latent intercept and slope variables. After identifying the sub-populations and their growth trajectories (often called the *measurement model*), researchers are usually interested in understanding the within- and between-person variability by incorporating external variables into GMMs, also known as covariates (known as the *structural model*).

Currently, two main estimators are available for estimating parameters of the GMMs with covariates, namely, the one-step and the bias-adjusted three-step estimators (Diallo & Lu, 2017). For the one-step estimator, also known as the full information maximum likelihood (FIML) estimator, the measurement model and structural model are estimated simultaneously by using all of the available information in the dataset, including the covariates (Huang et al., 2010; McCutcheon, 1987; Vermunt, 2010). This estimator yields efficient estimates when all the model assumptions hold (Bakk et al., 2013). However, simultaneous

CONTACT Yuqi Liu y.liu@fsw.leidenuniv.nl Methodology and Statistics Unit, Institute of Psychology, Leiden University, Leiden, The Netherlands.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/00273171.2025.2557275.

estimation may introduce interpretational confounding, where the latent construct that the applied researchers want to measure can change every time a new covariate is added to the model (Asparouhov & Muthén, 2014; Bakk & Kuha, 2018; Di Mari et al., 2023; Rosseel & Loh, 2024; Vermunt, 2010).

To prevent the interpretational confounding, as an alternative to the one-step estimator, the bias-adjusted three-step estimator has been developed for LC models with external variables (Vermunt, 2010), which estimates the measurement model and structural model separately by breaking down the estimation process into three steps: (1) estimating the measurement model only using repeated measures in the data, (2) classifying individuals to the latent classes based on their estimated posterior class membership probabilities from step-one, and (3) estimating the structural model by relating the class membership to external variables, while correcting the classification errors introduced in step-two. A series of approaches were developed for LC models to correct the classification errors (Bolck et al., 2004; Vermunt, 2010). The three-step maximum likelihood (ML) method proposed by Vermunt (2010) performs well in many LC models with external variables (e.g., latent class analysis [LCA], latent Markov [LM] models), yielding unbiased and efficient parameter estimates (Bakk et al., 2014). In the remainder, we refer to the bias-adjusted three-step estimator as the three-step estimator for simplicity.

The GMMs with covariates

The GMMs have a more complex measurement model than the LCM, as shown in Figure 1. Specifically, it contains a set of repeated measures (i.e., $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T$, where T is the number of time points.) that are regressed on the continuous latent intercept (η_0) and slope (η_1) variables. Furthermore, a categorical LC variable (c) is defined by the latent intercept and slope variables, allowing for population heterogeneity. Covariates can be included to predict the class membership (reflected by the blue line, from covariates x_1 and x_2 to c), but also to directly predict the class-specific growth factors (visualized as the two red lines, from covariate x_2 to η_0; from covariate x_2 to η_1), or both. For example, Chen et al. (2024) related cognitive function trajectories of Chinese respondents 45 years and older with diabetes to a set of baseline covariates (e.g., age, education level, gender, etc.) to discern predictors of cognitive function scores among Chinese elderly.

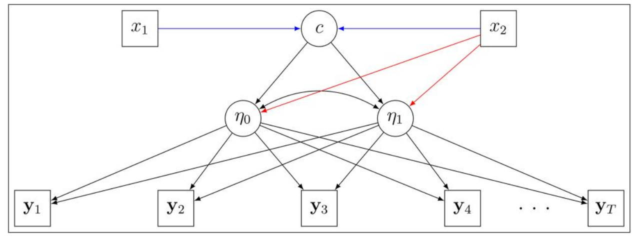


Figure 1. Growth mixture model with covariates x_1 and x_2, where $\mathbf{y}_t (t = 1, \ldots, T)$ is a vector of indicators that are directly regressed on the latent intercept and slope variables measured at t time points (e.g., $y_{t1}, y_{t2}, \ldots, y_{tH}$, where H is the number of indicators), c is the LC variable, η_0 is the latent intercept variable, and η_1 is the latent slope variable.

If the covariates solely predict class membership, the existing estimators yield accurate parameter estimates when the measurement model is correctly specified and the classes are well separated (L. Li & Hser, 2011). However, when the covariates have direct effects (DEs) on the growth factors (the red lines in Figure 1), the situation becomes more complicated. In GMMs, the growth factors, serving as indicators of the latent class variable, constitute a portion of the measurement model. When the covariates have DEs on the growth factors, the association between growth factors is not fully explained by the latent class variables. This situation violates the basic assumption of conditional independence between covariates and the indicators of the measurement model and shows measurement non-invariance or differential item functioning (Kankaraš et al., 2010; Vermunt, 2010). In simple LC models (we refer to the simplest type of a mixture model, e.g., LCA), we assume the indicators of latent classes (i.e., items) to be conditionally independent of the covariates given class membership. If we ignore these DEs, the un-modeled residual correlations between indicators and covariates will lead to bias in the parameter estimates (Masyn, 2017). Likewise, the parameter can be biased when the covariates have DEs on the growth factors that define the LCs in the GMMs.

The performance of existing estimators of GMMs in the presence of DEs on the latent factors

In the presence of DEs, the conditional independence assumption can be relaxed by modeling the DEs of the covariates on the concerned indicators. However, applied researchers may have limited prior evidence confirming which covariates actually exert DEs, and be uncertain in which specific indicators are affected by these DEs. In GMM, such ambiguity may result in model misspecification, either through omitting significant DEs or by erroneously modeling nonexistent effects.

For the three-step estimator, ignoring the DEs of covariates at step one can severely distort the measurement model and lead to substantial parameter bias in LCA and GMM (Asparouhov & Muthén, 2014). To account for this drawback, Asparouhov and Muthén (2014) proposed a possible modification by including the DEs in the measurement model of GMMs, and the modified three-step estimator performs comparably to the one-step when the classes are well separated. However, the number of manipulated factors in this study was limited, and the sample size of 10,000 used in the simulation study is not representative of many applied research settings. Moreover, another more extended simulation study (Diallo & Lu, 2017) showed conflicting results with those of Asparouhov and Muthén (2014). Namely, the modified three-step estimator performs worse than the one-step estimator across all conditions. Furthermore, the modified three-step estimator as proposed by Asparouhov and Muthén performs even worse than the conventional three-step estimator when the sample size is less than 2000, suggesting that the modifications of the three-step estimator should be carefully considered. While this estimator has also been developed in the context of GMMs (Asparouhov & Muthén, 2014; Diallo & Lu, 2017), so far the amount of evidence about its performance is insufficient in more complex setups of GMMs with covariates, particularly when DEs are specified on latent intercept and slope, and in different DE specifications that may occur in applied research contexts.

Recently, Vermunt and Magidson (2021a) proposed a modified version of the three-step estimator for LC models in the presence of DEs, modeling the covariates of interest on indicators and class membership at step one, and re-estimating the effects of concerned covariates on class membership at step three to prevent the overestimation of the DEs due to the unmodeled indirect effects *via* LC variables, and also correcting the classification error that differs across categories of covariates with DEs. This modified method works well, leading to unbiased parameter estimates in LC models. However, their modeling strategy was originally developed for the DEs on the observed indicator in LCA, and its generalization to handle DEs on latent variables—specifically the latent intercept and slope in GMM—has not yet been formally articulated or examined.

For the one-step estimator, the known DEs can be easily specified in the full model. However, despite the interpretational confounding problem, misspecification of DEs can distort the measurement model and thus change the latent class solutions, leading to improper interpretation of results (Vermunt, 2010). Furthermore, the one-step estimator exhibits substantial bias when direct effects (DEs) are ignored within the regression mixture model (RMM) framework (Kim et al., 2016). In contrast, when covariates that contribute to class separation are appropriately included, the one-step estimator performs well under the factor mixture model (FMM) framework. Notably, under conditions of low class separation, the one-step estimator outperforms the three-step estimator during the class enumeration process when relevant covariates are included, and demonstrates greater robustness to misspecification of DEs (i.e., ignoring and overfitting the DEs; Wang et al., 2023).

Regarding the accuracy of parameter estimates, the one-step estimator yields substantial bias in estimated covariate effects when the DEs are ignored in the LCA and the RMM (Janssen et al., 2019; Kim et al., 2016). In addition, when GMMs include numerous covariates for exploratory purposes, the one-step estimator can have convergence issues and local maxima due to the complexity of the likelihood function (Hipp & Bauer, 2006; Vermunt, 2010).

The proposed two-step estimator

A recently developed two-step estimator proposed for LC models (Bakk & Kuha, 2018) can be an alternative to the one-step and three-step estimators. In the two-step estimator, the measurement and structural models are estimated separately, which is similar to the three-step estimator. In step one, the covariates are excluded and only the measurement model is estimated. In step two, all the parameters of the measurement model are fixed at their estimated values from step one, and only the parameters of the remaining structural model are estimated conditioning on the step-one measurement model. Compared to the three-step estimator, the two-step estimator avoids introducing classification errors as in the classification step of the three-step estimator while showing comparable computational efficiency and conceptual advantage. Moreover, the two-step estimator can flexibly model the DEs from covariates to concerned indicators, which is recommended for estimating the single-level and multilevel LCA with external variables (Asparouhov & Muthén, 2014; Bakk et al., 2022; Bakk & Kuha, 2018). Since the measurement model is fixed when estimating the structural model, the two parts of the model do not directly affect each other, which allows the two-step estimator to be more robust than the one-step estimator when the DEs are ignored, which have been implemented in many mixture models, such as in the latent markov model and the multilevel LCA (Di Mari et al., 2023; Di Mari & Bakk, 2018). However, this approach has not yet been extended to GMMs with more complex measurement models that involve both continuous and categorical latent variables. When covariates with DEs are included in GMM, its robustness to misspecification—defined here as either omitting or overfitting the DEs—has not been comprehensively evaluated with regard to bias and efficiency of parameter recovery.

Given the aforementioned problems on the one- and three-step estimators for GMMs as well as the flexibility of the two-step method for modeling covariates and its robustness to misspecification, we propose to extend the two-step estimator for the LCA to the context of GMMs with covariates. Our method is motivated by Asparouhov and Muthén (2014), who proposed the three-step estimator in GMMs and Vermunt and Magidson (2021a) proposed a modeling strategy for the three-step estimator to account for DEs in LC models. Assume we know the possible covariates with DEs. The proposed two-step estimator here separately estimates the measurement and structural models. In step one, we include the DEs of covariates in the measurement model. Note that, to account for the overestimation of the DEs caused by ignoring the association between the latent class variable and covariates (Vermunt & Magidson, 2021a), we model not only the covariate effects on growth factors but also their effects on class membership in the step-one model. In step two, we estimate the structural model with all the interested covariates affecting class membership, conditioning on the step-one model that also includes the covariates with DEs on the latent intercept and slope. Specifically, the regression parameters of the concerned covariate effects on growth factors (i.e., the DEs) are fixed, while its effects on class membership are re-estimated at step two, to ensure obtaining the correct partial regression coefficients when incorporating covariates that solely predict class membership, in line with the recommendations of Vermunt and Magidson (2021a). In addition, we also propose to extend the approach (Vermunt & Magidson, 2021a) of three-step estimator for modeling DEs on the observed indicators in LCA to latent variables in the context of GMM.

In this paper, we introduce the two-step estimator to the context of GMMs and compare the efficiency and reliability of the proposed two-step estimator to the one-step estimator and three-step estimator, in terms of the accuracy of regression parameter estimates and coverage rates. We also examine the robustness of these estimators against misspecification of the covariate effects. Two different ways of misspecification are used, namely (1) we ignore the DEs from covariates to growth factors, and (2) we incorrectly include the DEs on growth factors. Additionally, we also inspect the Type I error rate for models that misspecify the effects between covariates and growth factors.

The remainder of this paper is structured as follows. First, we present the unconditional GMMs and GMMs incorporating covariates, and various estimators in estimating GMMs are given, including the one-step, the proposed two-step, and the three-step estimators. Then, we evaluate the performance of the proposed two-step and the competing one- and three-step estimators *via* extensive simulation studies. We apply the proposed two-step estimator to a real dataset from The China Health and Retirement Longitudinal Study (CHARLS; Zhao et al., 2013). The final section is a discussion of the presented results.

Model specification

The specification of GMMs

GMMs extend the LCM by relaxing the assumption of a single population. Assuming we only have one item at each time point, we first describe the unconditional LCM, which can be defined as

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

where $\mathbf{y}_i$ is a $T \times 1$ vector of repeated measures observed for individual i $(i = 1, 2, \ldots, N)$, T is the number of time points, where $\boldsymbol{\eta}_i$ is a $M \times 1$ vector of latent growth factors, M is the number of growth factors (e.g., for specifying a linear trajectory, M equals 2 and indicates the latent intercept η_{0i} and latent slope η_{1i}.), for capturing individual variation from the average growth trajectory. Finally, $\boldsymbol{\epsilon}_i$ is a $T \times 1$ vector of time-specific residuals. $\mathbf{\Lambda}$ is the $T \times M$ factor loading matrix with fixed coefficients to predetermine the functional form of the growth trajectory. For a linear trajectory with equally spaced time intervals, $\mathbf{\Lambda}$ can be set to

$$\mathbf{\Lambda} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & \ldots \\ 1 & T-1 \end{pmatrix}. \tag{2}$$

In the unconditional LCM, $\boldsymbol{\eta}_i$ can be written as:

$$\boldsymbol{\eta}_i = \boldsymbol{\alpha} + \boldsymbol{\zeta}_i, \tag{3}$$

where $\boldsymbol{\alpha}$ is an $M \times 1$ vector of growth factor means and $\boldsymbol{\zeta}_i$ is an $M \times 1$ vector of growth factor residuals. Under assumptions of independence and multivariate normality, $\boldsymbol{\epsilon}_i \sim N(0, \boldsymbol{\Theta})$, and $\boldsymbol{\zeta}_i \sim N(0, \boldsymbol{\Psi})$, where $\boldsymbol{\Theta}$ is a $T \times T$ variance-covariance matrix of time-specific residuals and $\boldsymbol{\Psi}$ is an $M \times M$ variance-covariance matrix of growth factors. The probability density function f of $\mathbf{y}_i$ is:

$$f(\mathbf{y}_i) \sim \mathrm{MVN}(\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta})), \tag{4}$$

where $\boldsymbol{\mu}(\boldsymbol{\theta})$ is the $T \times 1$ model-implied mean vector and $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is the $T \times T$ model-implied variance-covairance matrix with vector $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\Psi}, \boldsymbol{\Theta})$ of estimated parameters, given by

$$\begin{aligned} \boldsymbol{\mu}(\boldsymbol{\theta}) &= \boldsymbol{\Lambda}\boldsymbol{\alpha} \\ \boldsymbol{\Sigma}(\boldsymbol{\theta}) &= \boldsymbol{\Lambda}\boldsymbol{\Psi}\boldsymbol{\Lambda}' + \boldsymbol{\Theta}. \end{aligned} \tag{5}$$

In GMMs, the single population assumption is relaxed by introducing a categorical latent variable (c) to capture heterogeneity in growth trajectories. Therefore, when there are K distinctive latent classes within the population, assume a vector of repeated measures $\mathbf{y}_i^{(k)}$ is sampled from a multivariate normal distribution for the k_{th} latent class ($k = 1, 2, ..., K$). The marginal distribution of repeated measures for all classes, $\mathbf{y}_i$ is allowed to be non-normally distributed, which can be represented by a finite mixture of K normal distributions, with the probability density function expressed as:

$$f(\mathbf{y}_i) = \sum_{k=1}^{K} \pi(c_i = k) f(\mathbf{y}_i^{(k)}), \tag{6}$$

with class-specific model-implied mean vector $\boldsymbol{\mu}(\boldsymbol{\theta}^{(k)}) = \boldsymbol{\Lambda}\boldsymbol{\alpha}^{(k)}$, and variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta}^{(k)}) = \boldsymbol{\Lambda}\boldsymbol{\Psi}^{(k)}\boldsymbol{\Lambda}' + \boldsymbol{\Theta}^{(k)}$. The $\pi(c_i = k)$ is the class proportion defining the unconditional probability of individual i belonging to class k, where c_i is the class membership for individual i, and $f(\mathbf{y}_i^{(k)})$ presents the class-specific probability density function. The superscript k indicates the parameters are allowed to be class-specific.

The class sizes $\pi(c_i = k)$ are parameterized using a multinomial logistic regression model, given by:

$$\pi(c_i = k) = \frac{\exp(\beta_0^{(k)})}{\sum_{k=1}^{K} \exp(\beta_0^{(k)})} \tag{7}$$

with $\pi(c_i = k) > 0$ and $\sum_{k=1}^{K} \pi(c_i = k) = 1$, where $\beta_0^{(k)}$ is the logit intercept for class k, and this parameter for the reference class ($k = 1$) is standardized to zero for identification ($\beta_0^{(1)} = 0$).

Covariates can be incorporated into the GMMs to predict either the class membership, or the growth factors, or both. To do so, we can extend Equation (6) to be a conditional GMM of the form

$$f(\mathbf{y}_i | \mathbf{x}_i) = \sum_{k=1}^{K} \pi(c_i = k | \mathbf{x}_i) f(\mathbf{y}_i^{(k)} | \mathbf{x}_i), \tag{8}$$

where $\mathbf{x}_i$ denotes a $Q \times 1$ vector of covariates for individual i, and Q is the number of covariates. Therefore, the class proportion $\pi(c_i = k | \mathbf{x}_i)$ become a multinomial logistic function of covariates $\mathbf{x}_i$, which is given by

$$\pi(c_i = k | \mathbf{x}_i) = \frac{\exp(\beta_0^{(k)} + \mathbf{x}_i \boldsymbol{\beta}_{\mathbf{x}}^{(k)})}{\sum_{k=1}^{K} \exp(\beta_0^{(k)} + \mathbf{x}_i \boldsymbol{\beta}_{\mathbf{x}}^{(k)})}, \tag{9}$$

with $\pi(c_i = k | \mathbf{x}_i) > 0$ and $\sum_{k=1}^{K} \pi(c_i = k | \mathbf{x}_i) = 1$. And where $\boldsymbol{\beta}_{\mathbf{x}}^{(k)}$ is a $Q \times 1$ vector of regression slopes for class K.

The $f(\mathbf{y}_i^{(k)} | \mathbf{x}_i)$ is the class-specific probability density function conditioning on $\mathbf{x}_i$. The model-implied class-specific mean vector $\boldsymbol{\mu}(\boldsymbol{\theta}^{(k)})$ and variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta}^{(k)})$ can be expressed by

$$\begin{aligned} \boldsymbol{\mu}(\boldsymbol{\theta}^{(k)}) &= \boldsymbol{\Lambda}\boldsymbol{\alpha}^{(k)} + \boldsymbol{\Lambda}\boldsymbol{\Gamma}^{(k)}\mathbf{x} \\ \boldsymbol{\Sigma}(\boldsymbol{\theta}^{(k)}) &= \boldsymbol{\Lambda}(\boldsymbol{\Psi}^{(k)} + \boldsymbol{\Gamma}^{(k)}\boldsymbol{\Phi}\boldsymbol{\Gamma}^{(k)'})\boldsymbol{\Lambda}' + \boldsymbol{\Theta}^{(k)}. \end{aligned} \tag{10}$$

When there are Q covariates predicting the growth factors, $\boldsymbol{\Phi}$ is a $Q \times Q$ variance-covariance matrix of $\mathbf{x}_i$, and where $\boldsymbol{\Gamma}^{(k)}$ is the $M \times Q$ matrix of coefficients between the growth factors and covariates in class k.

The estimation of GMMs by maximizing the log-likelihood function typically employs the expectation-maximization (EM) algorithm (Dempster et al., 1977).

For illustration, along with the example in Figure 1, assume we have two covariates, x_1 and x_2, where x_1 only predicts the class membership and x_2 predicts both class membership and the growth factors (i.e., the latent intercept and slope for linear trajectory). Thus the $\boldsymbol{\beta}_{\mathbf{x}}^{(k)}$ can be expressed as $\beta_{x_1}^{(k)}$ and $\beta_{x_2}^{(k)}$, representing the logistic regression coefficients for x_1 and x_2 on the latent class variable respectively. And the $\boldsymbol{\Gamma}^{(k)}$ can be extended to $\gamma_I^{(k)}$ and $\gamma_S^{(k)}$, representing the coefficients of x_2 on the latent intercept and slope in Class k respectively. As we consider a linear growth trajectory, the $\boldsymbol{\alpha}^{(k)}$ can be extended as the mean of latent intercept $\alpha_0^{(k)}$ and slope $\alpha_1^{(k)}$ for class k. The $\boldsymbol{\Psi}^{(k)}$ can be expressed as

$$\boldsymbol{\Psi}^{(k)} = \begin{pmatrix} \psi_{00}^{(k)} & \\ \psi_{10}^{(k)} & \psi_{11}^{(k)} \end{pmatrix},$$

where $\psi_{00}^{(k)}$, $\psi_{10}^{(k)}$, and $\psi_{11}^{(k)}$ are the residual variances of the latent intercept and slope and the residual covariance of both growth factors for class k, respectively. All parameters in $\boldsymbol{\theta}_{full}$ that need to be estimated are:

$$\boldsymbol{\theta}_{full} = \{\alpha_0^{(k)}, \alpha_1^{(k)}, \gamma_I^{(k)}, \gamma_S^{(k)}, \psi_{00}^{(k)}, \psi_{10}^{(k)}, \psi_{11}^{(k)}, \beta_0^{(k)}, \beta_{x_1}^{(k)}, \beta_{x_2}^{(k)}, \pi^{(2)}, ..., \pi^{(K)}\}. \tag{11}$$

In what follows, for all proposed estimators, we use the standard Hessian-based standard error estimates as provided by standard software (Vermunt & Magidson, 2021b).

Estimation methods

This section presents the one-step, the proposed two-step, and the three-step estimators for estimating the growth mixture models (GMMs) with covariates defined in the previous section. For simplicity, we use the GMM with covariates presented in Figure 1 as an example. Note that both covariates x_1 without DEs and x_2 with DEs can be extended to a vector of covariates in more complex settings.

The one-step estimator

The one-step FIML estimator estimates all parameters (defined in Equation (11)) of the measurement and the structural models at once (B. Muthén, 2004). When estimating GMM with covariates x_1 and x_2, the EM algorithm is used to maximize the log-likelihood function $\text{Log}L_{FIML}(\boldsymbol{\theta}_{full})$ for $f(\mathbf{y}_i|x_{1i}, x_{2i})$:

$$\begin{aligned}&\text{Log}L_{FIML}(\boldsymbol{\theta}_{full})\\ &= \sum_{i=1}^{N} \log\left(\sum_{k=1}^{K} \pi(c_i = k|x_{1i}, x_{2i}) f\left(\mathbf{y}_i^{(k)}|x_{2i}, \boldsymbol{\mu}(\boldsymbol{\theta}_{full}), \boldsymbol{\Sigma}(\boldsymbol{\theta}_{full})\right)\right)\end{aligned} \tag{12}$$

The proposed two-step estimator

Here, we apply the two-step estimator (Bakk & Kuha, 2018) on the GMMs with covariates, and model the DEs of covariates on the growth factors following the recommendation of Vermunt and Magidson (2021a) on the DEs modeling strategy for LCA.

Step-one. In step one of the proposed two-step estimator, we first estimate the class-specific parameters and class proportions for the GMM with covariates specified before. Note that we include covariate with DEs (i.e., x_2) at step one, containing both the concerned DEs on growth factors and the effect on class membership, in line with the recommendations of Vermunt and Magidson (2021a). The estimated parameters in step one are as follows:

$$\boldsymbol{\theta}_{s1} = \{\alpha_0^{(k)}, \alpha_1^{(k)}, \gamma_I^{(k)}, \gamma_S^{(k)}, \psi_{00}^{(k)}, \psi_{10}^{(k)}, \psi_{11}^{(k)}, \beta_0^{(k)}, \beta_{x_2}^{(k)}, \pi^{(2)}, ..., \pi^{(K)}\} \tag{13}$$

The log-likelihood function of the step-one model can be specified as follows:

$$\begin{aligned}&LogL(\boldsymbol{\theta}_{s1})\\ &= \sum_{i=1}^{N} \log\left(\sum_{k=2}^{K} \pi(c_i = k|x_{2i}) f\left(\mathbf{y}_i^{(k)}|x_{2i}, \boldsymbol{\mu}(\boldsymbol{\theta}_{s1}), \boldsymbol{\Sigma}(\boldsymbol{\theta}_{s1})\right)\right)\end{aligned} \tag{14}$$

Step-two. In step two of the proposed two-step estimator, we examine the association between covariates and class membership by estimating the regression coefficients in Equation (9), and re-estimate the class proportions conditioning on x_1 and x_2. Meanwhile, the remaining parameters of the step-one model are fixed at their estimated value, which is $\hat{\boldsymbol{\theta}}_{s1} = \{\alpha_0^{(k)}, \alpha_1^{(k)}, \gamma_I^{(k)}, \gamma_S^{(k)}, \psi_{00}^{(k)}, \psi_{10}^{(k)}, \psi_{11}^{(k)}\}$. Note that the DEs of x_2 on growth factors were fixed at step-one estimates (i.e, $\gamma_I^{(k)}, \gamma_S^{(k)}$), while the effect of x_2 on class membership is re-estimated at step two which follows the recommendations of Vermunt and Magidson (2021a). Hence, the parameters that need to be estimated in step two are as follows:

$$\boldsymbol{\theta}_{s2} = \{(\beta_0^{(k)}, \beta_{x_1}^{(k)}, \beta_{x_2}^{(k)}, \pi^{(2)}, ..., \pi^{(K)})|\hat{\boldsymbol{\theta}}_{s1}\}. \tag{15}$$

The log-likelihood function of the step-two model can be specified as follows:

$$\begin{aligned}&LogL(\boldsymbol{\theta}_{s2}|\boldsymbol{\theta}_{s1} = \hat{\boldsymbol{\theta}}_{s1}) =\\ &\sum_{i=1}^{N} \log \underbrace{\left(\sum_{k=1}^{K} \pi(c_i = k|x_{1i}, x_{2i}) f\left(\mathbf{y}_i^{(k)}|x_{2i}, \boldsymbol{\mu}(\hat{\boldsymbol{\theta}}_{s1}), \boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}}_{s1})\right)\right)}_{\text{Free(conditional on } \hat{\boldsymbol{\theta}}_{s1}\text{)Fixed } (\hat{\boldsymbol{\theta}}_{s1})}.\end{aligned} \tag{16}$$

Using this approach makes it possible to separate the measurement and structural model, by replacing the FIML approach with a model where a conditional likelihood function is used at step two, allowing for computational efficiency (Bakk & Kuha, 2018).

The three-step estimator

In this paper, we generalize the Vermunt and Magidson (2021a) modeling strategy for the three-step estimator in the presence of known DEs on the observed indicators to the context of GMMs, in which the DEs is on latent variables.

Step-one. The step-one model is equivalent to the step-one model of the proposed two-step estimator, thus, the estimated parameters $\boldsymbol{\theta}_{s1}$ are obtained by maximizing the log-likelihood function of Equation (14).

Step-two. In step-two, individuals are assigned to latent classes based on their posterior probability of class membership given covariate x_{2i}, which is expressed as follows:

$$P(c_i = k|\mathbf{y}_i, x_{2i}) = \frac{\pi(c_i = k|x_{2i})f(\mathbf{y}_i^{(k)}|x_{2i})}{f(\mathbf{y}_i|x_{2i})} \tag{17}$$

Here, we applied modal assignments, in which the individuals are assigned to the class with the highest posterior class probability. The posterior class assignments are denoted by introducing a new categorical variable W, which can take on the values $w = 1, 2, ..., K$. In this step, classification errors are introduced and can be quantified as the posterior class membership conditional on the true class membership and x_2 Vermunt (2010), that is:

$$P(W = w|c = k, x_2) = \frac{1}{N}\sum_{i=1}^{N}\left(\sum_{k=1}^{K} f(\mathbf{y}_i^{(k)}|\boldsymbol{\eta}_i^{(k)})f(\boldsymbol{\eta}_i^{(k)}|x_{2i}) \times P(c = k|\mathbf{y}_i^{(k)}, \boldsymbol{\eta}_i^{(k)}, x_{2i}) P(W = w|\mathbf{y}_i, \boldsymbol{\eta}_i, x_{2i})\right)/P(c = k|x_2\Big). \tag{18}$$

Here, in contrast to the proposal from Bolck et al. (2004), which disregarded x_2 in the classification error matrix, Vermunt and Magidson (2021a) proposed to allow the classification error matrix with elements $P(W = w|c = k, x_2)$ to vary across the level of x_2.

Step-three. The step-three model is an LC model conditional on the x_1 and x_2, with a single indicator W of response probabilities $P(W = w|c = k, x_2)$, that is,

$$P(W = w|x_1, x_2) = \sum_{k=1}^{K} P(c = k|x_1, x_2) \sum_{w \neq k} P(W = w|c = k, x_2). \tag{19}$$

In step three, we only estimate the regression coefficients of the multinominal logistic function relating W and x_1 and x_2, that is, $\boldsymbol{\theta}_{s3} = \{\beta_0^{(k)}, \beta_{x_1}^{(k)}, \beta_{x_2}^{(k)}\}(k = 1, 2, ..., K)$. The $\boldsymbol{\theta}_{s3}$ is obtained by maximizing the following log-likelihood function,

$$LogL(\boldsymbol{\theta}_{s3}) = \sum_{i=1}^{N} log\left(\sum_{w=1}^{K} P(W = w|x_1, x_2)\right) \tag{20}$$

For parameter estimation, we use the ML estimator proposed by Vermunt (2010). For an extended description of the three-step estimator, we refer to Asparouhov and Muthén (2014), Vermunt and Magidson (2021a), and Diallo and Lu (2017).

Simulation settings

Three simulation studies were conducted to assess the performance of the different estimators for the GMMs with covariates. Specifically, we compare the commonly used one-step and the three-step estimators with the proposed two-step estimator among ignored, correctly specified, or overfitted DEs of covariates. The efficiency and accuracy of each estimator are examined, specifically evaluating the parameters of the estimated covariate effects (on the LC variable and DEs on the growth factors) and its corresponding standard errors (*SE*s), in terms of the absolute bias (AB), the mean square error (MSE), the relative efficiency (*SE*/*SD* ratio), and the coverage rates (CRs) of 95% confidence interval (CI). The formulas for these criteria can be found in Appendix A. We expect that the estimators that specify DEs (either correctly specified or overfitted) perform better than the estimators that ignore the DEs, and that the proposed two-step estimator performs comparably to the one-step and the three-step estimators with correctly specified DEs and is more robust than the one-step estimator with misspecified DEs.

Population models

For all three simulation studies, we sampled the data from two-class linear GMMs with 3 observed continuous indicators that directly regress on the latent intercept and slope variables at each time point ($\mathbf{y}_t \sim N(0, 1); t = 1, 2, 3, ..., T$) and with:

- Population model I: only with covariate x_1 predicting class membership (Study 1).
- Population model II: with covariates x_1 predicting class membership and x_2 predicting the latent intercept (i.e., η_0) only (Study 2).
- Population model III: with covariates x_1 predicting class membership and x_2 predicting both growth factors (i.e., η_0, η_1) (Study 3)

The x_1 and x_2 are sampled from $U(1, 5)$. The population models are visualized in Figure 2, in which we only presented the model part of interest (i.e., the association among growth factors, LC variables, and covariates).

The population parameters were chosen based on both previous simulation studies and substantive research on mixture models (e.g., Diallo et al., 2017; M. Li & Harring, 2017; Tofighi & Enders, 2008; Vermunt & Magidson, 2021a). Specifically, within the

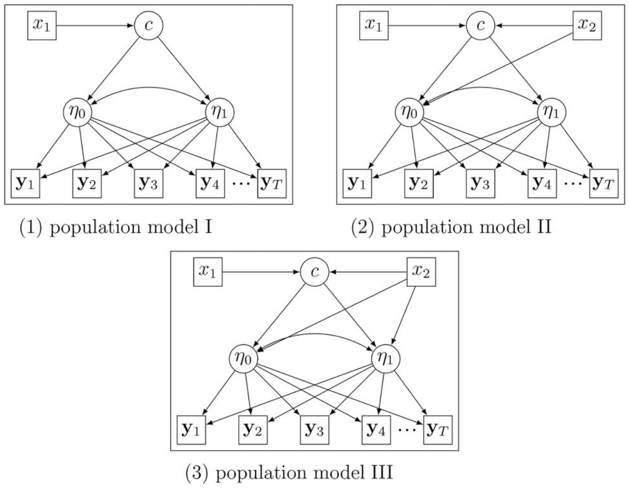


Figure 2. Population models for the three simulation studies, where $\mathbf{y}_t (t = 1, \ldots, T)$ are the indicators measured at t time point, c is the LC variable, η_0 is the latent intercept, and η_1 is the latent slope. The x_1 and x_2 are the covariates.

class k, we assumed the repeated measures $\begin{pmatrix} \mathbf{y}_1^{(k)} \\ \ldots \\ \mathbf{y}_T^{(k)} \end{pmatrix} \sim$ $\mathrm{MVN}(\boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)})$, with

$$\begin{aligned} \boldsymbol{\mu}^{(k)} &= \boldsymbol{\Lambda}(\boldsymbol{\alpha}^{(k)} + \boldsymbol{\gamma}^{(k)} x_2) \\ \boldsymbol{\Sigma}^{(k)} &= \boldsymbol{\Lambda}(\boldsymbol{\Psi} + \boldsymbol{\gamma}^{(k)} \phi \boldsymbol{\gamma}^{(k)'})\boldsymbol{\Lambda}' + \boldsymbol{\Theta}. \end{aligned}$$

The mean vector $\boldsymbol{\mu}^{(k)}$ of repeated measures is defined by $\boldsymbol{\alpha}^{(k)}$, $\boldsymbol{\gamma}^{(k)}$, $\boldsymbol{\Lambda}$ (as defined in Equation (2)), and the mean of x_2. The $\boldsymbol{\alpha}^{(k)}$ includes the mean of latent intercept ($\alpha_0^{(k)}$) and slope ($\alpha_1^{(k)}$), notice that we varied the $\alpha_0^{(k)}$ across two classes and population models to manipulate the level of class separation, the details are presented in the manipulated factor section. For population models I, II, and III, the $\alpha_1^{(k)}$ was defined for class 1 and class 2, as $\alpha_1^{(1)} = -0.20, \alpha_1^{(2)} = 0.20$. As only covariate ($x_2$) predict η_0 in population model II, the $\boldsymbol{\gamma}^{(k)}$ was defined as $\gamma_I^{(1)} = 0.5$, $\gamma_I^{(2)} = -0.5$. In population model III, x_2 predicts both η_0 and η_1, which was defined as $\boldsymbol{\gamma}^{(1)} = (\gamma_I^{(1)}, \gamma_S^{(1)}) = (0.5, 0.25)$, $\boldsymbol{\gamma}^{(2)} = (\gamma_I^{(2)}, \gamma_S^{(2)}) = (-0.5, -0.25)$. To focus on investigating the proper strategy for modeling the DEs in GMMs, we chose a medium size of the DEs from x_2 for growth factors across simulation studies.

The variance-covariance matrix ($\boldsymbol{\Sigma}^{(k)}$) of repeated measures is defined by the $\boldsymbol{\Psi}$, ϕ (the variance of x_2) and $\boldsymbol{\Theta}$. For simplicity, the $\boldsymbol{\Psi}$ and $\boldsymbol{\Theta}$ were set to be invariant across all classes and population models. Specifically, the $\boldsymbol{\Psi}$ was set to

$$\boldsymbol{\Psi} = \begin{pmatrix} \psi_{00} & \\ \psi_{10} & \psi_{11} \end{pmatrix} = \begin{pmatrix} 1.00 & \\ -0.15 & 1.00 \end{pmatrix}.$$

And the $\boldsymbol{\Theta}$ was fixed as a diagonal matrix with all the elements fixed to 1.

We set the logistic regression parameters of x_1 and x_2 on c to determine the class membership (as specified in Equation (9)). Specifically, $\beta_{x_1} = -0.50$ for population models I, II, and III, and $\beta_{x_2} = 0.75$ for population models II and III. The regression intercept (β_0) of c was varied across three population models to manipulate the mixing ratio of classes and to obtain two class size settings, namely equal and unequal classes (see below).

Manipulated factors

According to previous simulation and applied studies on mixture models, we manipulated the following four factors given their important influence on the performance of GMM, including the sample size, the mixing ratios, the degree of class separation, and the number of time points. Previous research showed that the sample size, mixing ratios, and class separation are important factors of model performance in mixture modeling in terms of the class enumeration and parameter recovery (Asparouhov & Muthén, 2014; Diallo & Lu, 2017; L. Li & Hser, 2011; Tofighi & Enders, 2008; Vermunt, 2010; Wang et al., 2023). In addition, the performance of the three-step and two-step estimators highly depends on the class separation and sample size (Bakk & Kuha, 2018; Di Mari et al., 2023; Vermunt, 2010). Moreover, the number of time points also plays a vital role in ensuring the statistical power in GMM (B. O. Muthén & Curran, 1997).

For the manipulated factors in the three simulation studies, (1) two different sample sizes are chosen: $N = 500$ and 1000. (2) Two levels of mixing ratios were applied. (3) Two levels (medium and high) of class separation conditions were applied by tuning the $\alpha_0^{(k)}$ (For details, as shown in Appendix B). The low level is not considered, as this condition is not recommended for step-wise estimators (Vermunt, 2010). The entropy value was used to assess the accuracy of the generated class separation. The entropy value ranged from 0.52 to 0.90 for the medium separation condition and from 0.93 to 0.99 for the high separation condition, averaged from all simulated data sets. (4) We manipulated the number of time points (T) to 3 and 6 across simulation studies. The chosen parameters of manipulated factors are typically used in substantial research and simulation studies in the framework of GMMs.

Data generation and analytical procedure

The three simulation studies consisted of 16 designed conditions (2 levels of sample sizes × 2 levels of mixing ratios × 2 levels of class separation × 2 levels of time points) for each population model, 48 conditions in total. We generated 100 replications for each condition, resulting in 4800 datasets for 3 simulation studies.

In Study 1, the population model I only included covariate x_1 without DEs, we thereby compared the one-step estimator with x_1 to the two-step and the three-step estimator without x_1 in the step-one model, while incorporating x_1 in step-two model of the two-step estimator and step-three model of the three-step estimator. Hence, we employed 3 models to analyze the simulated datasets at each condition.

In Studies 2 and 3, we included covariates x_1, also x_2, that have DEs on the latent intercept for population model II and on both growth factors for population model III. We compared the performance of three alternative estimators in the presence of DEs. For each estimator, we evaluated the impact of three different specifications for the DEs of x_2: ignoring the DEs (Specification A), specifying the DEs on the latent intercept only (Specification B), and specifying the DEs on the latent intercept and slope (Specification C). Note that the correct specification and misspecification of DEs vary depending on the population model of different studies. For study 2, specification B correctly specified DEs, in contrast to study 3, where specification C is correct. In total, each simulated dataset at each condition was analyzed using 9 models (3 estimators × 3 specifications).

Specifically, 3 one-step estimators with varying specifications of DEs were built, which were identical to the three population models, 3 two-step, and 3 three-step estimators, with varying specifications of the DEs at their equivalent step-one models, the corresponding step-wise models are presented in Figure 3a and Figure 3b. All simulated data sets were generated and estimated using LatentGOLD version 6.0 Vermunt and Magidson (2021b), and results were analyzed in R. The LatentGOLD syntax and R code can be found in the author's GitHub repository [1].

To prevent the label-switching problem (Tueller et al., 2011), we provided starting values[2] for all models across three simulation studies in LatentGOLD.

[1]See https://github.com/Yuqi-psy/Two-step-GMM.git.

[2]The starting values are the parameters of population models. A robustness check without specifying the starting values was run, and the results are consistent with our simulation results. The details can be found in the supplementary materials.

Results

Study 1: Population model I with covariate effect only on class membership

In Study 1, we compared the performance of the different estimators on estimating the GMMs with covariate x_1 predicting the class membership only. Table 1 presents the absolute bias (AB) and 95% CI of CRs of parameter β_{x_1} for the three estimators under the 6 time points conditions with varying sample sizes, levels of class separation (measured by the average of entropy), and mixing ratios. The results indicate that all three estimators performed comparably well in the 6 time points conditions, resulting in negligible parameter AB and acceptable CRs over the 8 simulation conditions (2 sample size × 2 class separation × 2 mixing ratio). A similar pattern of results was obtained in the 3 time points conditions as well (see Appendix C).

Figure 4 displays the boxplots for relative efficiency (*SE*/*SD* ratios) with different time points, averaged over the other 3 design factor conditions (i.e., mixing ratio, class separation, and sample size). In the 3 time point conditions, the *SE*/*SD* values are close to 1 over all three estimators, indicating that the standard error (*SE*) estimators are similar to the sampling variance regardless of the sample sizes, levels of class separation, and mixing ratios. In the 6 time point conditions, all three estimators overestimate the *SE*s. Specifically, the proposed two-step estimator performed comparably to the one-step estimator, and better than the three-step estimator, the latter yields a larger magnitude of *SE*s overestimation than the other two estimators. The *SE*/*SD* values at each simulated condition are in Appendix E. We also inspected the mean square value of the β_{x_1}, and the results are similar to the AB (see Appendix D).

Study 2: Population model II with covariate effects on the class membership and on the latent intercept

In study 2, data were generated from the population model II with a DE from covariate x_2 to the latent intercept, and effects of covariates x_1 and x_2 on the LCs. We inspected the estimated parameter bias (AB), coverage rates (CRs), relative efficiency (*SE*/*SD* ratios), and the type I error rates (i.e., the probability of incorrectly accepting a significant effect of x_2 on the latent slope), to evaluate the performance of the one-step, the two-step, and the three-step estimators in the presence of DE.

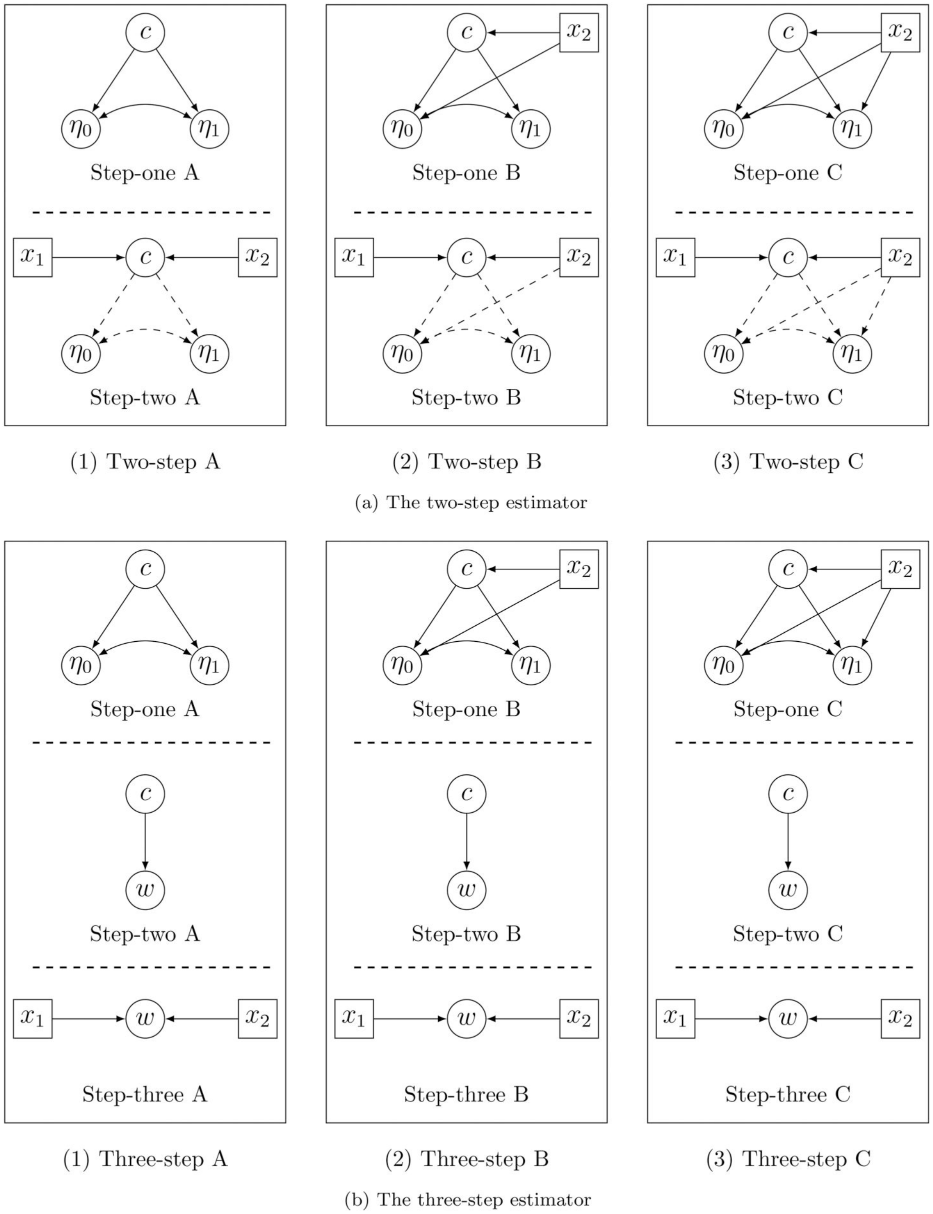


Figure 3. The step-wise models for the two-step and the three-step estimators in simulation studies 2 and 3, with three different specifications of the direct effects from x_2 on c. The dashed lines for the two-step approach represent the parameters on this path that were fixed at the step one estimates. The w is the classification LC variable in the three-step estimator.

Table 2 presents the AB and 95% CI of CRs for covariate effects on LCs (i.e., parameters β_{x_1} and β_{x_2}) in the 6 time points conditions with varying sample sizes, levels of class separation, and mixing ratios. For the β_{x_2}, when we ignored the DEs (specification A), all three estimators show bias, especially when the classes are poorly separated. Specifically, the two-step and the three-step estimators performed comparably, with the AB range from 0.03 to 0.26 and CRs not reaching the nominal level. Not surprisingly, the one-step estimator was the most sensitive estimator to misspecification, with the largest AB range from 0.05 to 1.43 and the lowest CRs. Furthermore, the performance of all three estimators systematically improved when the DEs were specified (either specification B or C), yielding negligible AB (range from 0.00 to 0.01) and acceptable CRs across simulation conditions.

Next, zooming into the different simulation conditions, we see that all three estimators performed better as the classes become more separate, while they are less affected by the sample size and the mixing ratio. When the DEs were specified (specification B or C), the three-step and the two-step estimators were more sensitive to the level of class separation than the one-step estimator. As shown in Table 2, the CRs of these

Table 1. The absolute bias (95% confidence interval of coverage rates) values over 100 replications for the regression coefficient of the latent class variable β_{x_1}, in 6 time points simulation conditions for study 1.

	Mixing ratio = 0.50/0.50				Mixing ratio = 0.30/0.70			
	High entropy		Moderate entropy		High entropy		Moderate entropy	
	$N = 1000$	$N = 500$	$N = 1000$	$N = 500$	$N = 1000$	$N = 500$	$N = 1000$	$N = 500$
$\beta_{x_1} = -0.50$								
One-step	0.01 (0.90-0.98)	0.00 (0.90-0.98)	0.00 (0.93-0.99)	0.04 (0.88-0.97)	0.00 (0.89-0.98)	0.01 (0.92-0.99)	0.00 (0.89-0.98)	0.04 (0.92-0.99)
Two-step	0.01 (0.90-0.98)	0.00 (0.90-0.98)	0.00 (0.95-1.00)	0.02 (0.90-0.98)	0.00 (0.89-0.98)	0.01 (0.92-0.99)	0.00 (0.89-0.98)	0.01 (0.92-0.99)
Three-step	0.01 (0.92-0.99)	0.00 (0.90-0.98)	0.00 (0.96-1.00)	0.02 (0.95-1.00)	0.00 (0.90-0.98)	0.01 (0.92-0.99)	0.01 (0.93-0.99)	0.01 (0.95-1.00)

Note. One-step is the one-step estimator. Two-step is the proposed two-step estimator. Three-step is the three-step estimator. N is the total sample size.

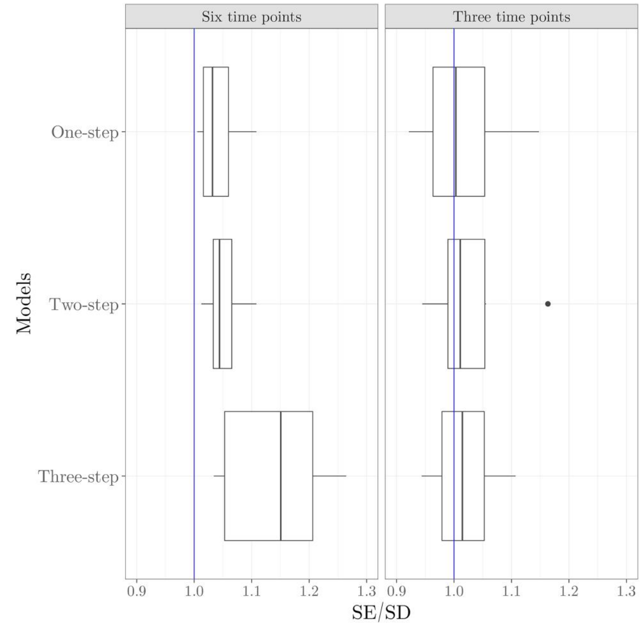


Figure 4. Boxplots of relative efficiency for the regression coefficient of the latent class variable (β_{x_1}) averaged over 8 simulation conditions for study 1. *SE/SD* is the ratio of standard error versus standard deviation. One-step is the one-step estimator. Two-step is the proposed two-step estimator. Three-step is the three-step estimator.

two estimators are substantially reduced but still provide unbiased estimates. For the β_{x_1} which has no DE, all three estimators performed systematically well with AB smaller than 0.03 and CRs close to the nominal level within the CI, across 8 conditions and 3 specifications, except for the one-step estimator with ignored DEs (specification A) under a moderate entropy and equal class size conditions.

Figure 5 depicts the boxplots for *SE*/*SD* ratio values reported at each replication for the β_{x_1} and β_{x_2} averaged over 8 simulation conditions. For the β_{x_2}, all estimators tend to underestimate the *SE*s over different time points conditions. All three estimators with specified DEs (specifications B or C) performed better than estimators with ignored DEs (specification A), with less underestimation of *SE* values. Specifically, the two-step estimator provides slightly more underestimated *SE*s than the one-step and the three-step estimators. Not surprisingly, the one-step estimator with ignored DEs (One-step A) substantially underestimates the *SE*s. When we specified the DEs (specifications B or C), the most efficient estimator was the one-step, followed closely by the three-step and the two-step estimators, the stepwise estimators were less efficient with underestimating the *SE* values, especially in the moderate class separation conditions (for the detailed *SE*/*SD* in each simulated condition, we presented in Appendix E). For the β_{x_1}, the one-step and

Table 2. The absolute bias (AB) (95% confidence interval of coverage rates) values over 100 replications for the effect of x_1 and x_2 on the latent class, latent slope, and latent intercept variables, in 6 time points Conditions for study 2.

	Mixing ratio = 0.50/0.50				Mixing ratio = 0.30/0.70			
	High entropy		Moderate entropy		High entropy		Moderate entropy	
	$N = 1000$	$N = 500$	$N = 1000$	$N = 500$	$N = 1000$	$N = 500$	$N = 1000$	$N = 500$
$\beta_{x_1} = -0.50$								
One-step A	0.03 (0.88–0.97)	0.01 (0.89–0.98)	0.41 (**0.79–0.92**)	0.45 (0.89–0.98)	0.02 (**0.81–0.94**)	0.01 (0.92–0.99)	0.01 (0.89–0.98)	0.04 (0.93–0.99)
One-step B	0.03 (0.89–0.98)	0.01 (0.89–0.98)	0.00 (0.88–0.97)	0.06 (0.93–0.99)	0.02 (0.88–0.97)	0.00 (0.88–0.97)	0.00 (0.89–0.98)	0.02 (0.86–0.97)
One-step C	0.03 (0.89–0.98)	0.01 (0.89–0.98)	0.00 (0.88–0.97)	0.06 (0.93–0.99)	0.02 (0.88–0.97)	0.00 (0.89–0.98)	0.00 (0.89–0.98)	0.01 (0.88–0.97)
Two-step A	0.03 (0.88–0.97)	0.01 (0.89–0.98)	0.00 (0.88–0.97)	0.05 (0.92–0.99)	0.01 (**0.83–0.94**)	0.00 (0.92–0.99)	0.01 (0.92–0.99)	0.01 (0.86–0.97)
Two-step B	0.03 (0.89–0.98)	0.01 (0.89–0.98)	0.00 (0.88–0.97)	0.05 (0.93–0.99)	0.02 (0.88–0.97)	0.00 (0.89–0.98)	0.00 (0.89–0.98)	0.01 (0.86–0.97)
Two-step C	0.03 (0.89–0.98)	0.01 (0.89–0.98)	0.00 (0.88–0.97)	0.05 (0.93–0.99)	0.02 (0.88–0.97)	0.00 (0.89–0.98)	0.00 (0.89–0.98)	0.01 (0.88–0.97)
Three-step A	0.03 (0.89–0.98)	0.01 (0.89–0.98)	0.01 (0.90–0.98)	0.04 (0.93–0.99)	0.02 (0.86–0.97)	0.00 (0.90–0.98)	0.01 (0.92–0.99)	0.01 (0.92–0.99)
Three-step B	0.03 (0.90–0.98)	0.01 (0.89–0.98)	0.00 (0.88–0.97)	0.05 (0.95–1.00)	0.01 (0.88–0.97)	0.00 (0.90–0.98)	0.01 (0.90–0.98)	0.01 (0.90–0.98)
Three-step C	0.03 (0.90–0.98)	0.01 (0.89–0.98)	0.00 (0.88–0.97)	0.05 (0.95–1.00)	0.01 (0.88–0.97)	0.00 (0.90–0.98)	0.00 (0.90–0.98)	0.01 (0.90–0.98)
$\beta_{x_2} = 0.75$								
One-step A	0.07 (**0.66–0.82**)	0.05 (0.85–0.96)	1.43 (**0.00–0.04**)	1.30 (**0.04–0.15**)	0.16 (**0.42–0.62**)	0.18 (**0.58–0.76**)	0.66 (**0.00–0.04**)	0.72 (**0.00–0.04**)
One-step B	0.01 (0.93–0.99)	0.00 (0.86–0.97)	0.00 (0.89–0.98)	0.01 (0.93–0.99)	0.00 (0.86–0.97)	0.01 (0.92–0.99)	0.01 (0.86–0.97)	0.01 (0.90–0.98)
One-step C	0.01 (0.93–0.99)	0.00 (0.86–0.97)	0.00 (0.89–0.98)	0.01 (0.90–0.98)	0.00 (0.86–0.97)	0.01 (0.92–0.99)	0.01 (0.85–0.96)	0.01 (0.92–0.99)
Two-step A	0.04 (**0.73–0.88**)	0.03 (0.88–0.97)	0.15 (**0.39–0.59**)	0.14 (**0.52–0.71**)	0.04 (0.84–0.95)	0.05 (0.88–0.97)	0.18 (**0.30–0.49**)	0.18 (**0.50–0.69**)
Two-step B	0.01 (0.93–0.99)	0.00 (0.86–0.97)	0.00 (**0.78–0.91**)	0.01 (0.86–0.97)	0.00 (0.86–0.97)	0.01 (0.92–0.99)	0.01 (**0.78–0.91**)	0.01 (**0.80–0.93**)
Two-step C	0.01 (0.92–0.99)	0.00 (0.86–0.97)	0.00 (**0.79–0.92**)	0.01 (**0.83–0.94**)	0.00 (0.86–0.97)	0.01 (0.92–0.99)	0.01 (**0.79–0.92**)	0.01 (**0.77–0.91**)
Three-step A	0.02 (0.85–0.96)	0.01 (0.90–0.98)	0.13 (**0.51–0.70**)	0.10 (**0.76–0.90**)	0.04 (**0.81–0.94**)	0.03 (0.92–0.99)	0.22 (**0.36–0.55**)	0.17 (**0.60–0.78**)
Three-step B	0.01 (0.90–0.98)	0.00 (0.86–0.97)	0.00 (0.88–0.97)	0.00 (0.88–0.97)	0.00 (0.88–0.97)	0.01 (0.93–0.99)	0.01 (**0.83–0.94**)	0.01 (**0.83–0.94**)
Three-step C	0.01 (0.90–0.98)	0.00 (0.89–0.98)	0.00 (0.88–0.97)	0.00 (0.86–0.97)	0.00 (0.88–0.97)	0.01 (0.95–1.00)	0.01 (**0.81–0.94**)	0.00 (**0.81–0.94**)
$\gamma_I^{(1)} = 0.50$								
One-step B	0.00 (0.90–0.98)	0.01 (0.95–1.00)	0.00 (0.84–0.95)	0.01 (0.90–0.98)	0.00 (0.92–0.99)	0.00 (0.85–0.96)	0.01 (0.90–0.98)	0.00 (0.88–0.97)
One-step C	0.00 (0.89–0.98)	0.01 (0.93–0.99)	0.00 (**0.83–0.94**)	0.01 (0.90–0.98)	0.00 (0.89–0.98)	0.00 (0.88–0.97)	0.01 (0.90–0.98)	0.00 (**0.83–0.94**)
Step-one B	0.00 (0.90–0.98)	0.01 (0.95–1.00)	0.00 (0.85–0.96)	0.01 (0.90–0.98)	0.00 (0.92–0.99)	0.00 (0.84–0.95)	0.01 (0.90–0.98)	0.00 (0.88–0.97)
Step-one C	0.00 (0.89–0.98)	0.01 (0.93–0.99)	0.00 (**0.81–0.94**)	0.01 (0.89–0.98)	0.00 (0.89–0.98)	0.00 (0.86–0.97)	0.01 (0.88–0.97)	0.00 (**0.81–0.94**)
$\gamma_I^{(2)} = -0.50$								
One-step B	0.01 (0.88–0.97)	0.00 (**0.83–0.94**)	0.00 (0.90–0.98)	0.00 (0.88–0.97)	0.00 (0.89–0.98)	0.00 (0.86–0.97)	0.00 (0.93–0.99)	0.01 (0.88–0.97)
One-step C	0.01 (0.89–0.98)	0.00 (**0.83–0.94**)	0.00 (0.92–0.99)	0.00 (0.86–0.97)	0.00 (0.90–0.98)	0.00 (0.89–0.98)	0.00 (0.92–0.99)	0.00 (0.90–0.98)
Step-one B	0.01 (0.89–0.98)	0.00 (**0.83–0.94**)	0.00 (0.89–0.98)	0.00 (0.88–0.97)	0.00 (0.90–0.98)	0.00 (0.86–0.97)	0.00 (0.93–0.99)	0.01 (0.88–0.97)
Step-one C	0.01 (0.89–0.98)	0.00 (**0.83–0.94**)	0.00 (0.90–0.98)	0.00 (0.88–0.97)	0.00 (0.90–0.98)	0.00 (0.89–0.98)	0.00 (0.92–0.99)	0.00 (0.90–0.98)

Note. One-step A, One-step B, and One-step C are the one-step estimators ignoring the direct effects (DEs; specifications A), specifying the DE on the latent intercept (specification B), and on both growth factors (specification C), respectively. Two-step A, Two-step B, and Two-step C are the two-step estimators with specifications A, B, and C, respectively. Three-step A, Three-step B, and Three-step C are the three-step estimators with specifications A, B, and C, respectively. Step-one B and C are the step-one models of the two-step and three-step estimators with specifications B and C. β_{x_1} and β_{x_2} are regression coefficients of the latent class variable. $\gamma_I^{(1)}$ and $\gamma_I^{(2)}$ are the regression coefficients of the latent intercept in Class 1 and 2. $\gamma_S^{(1)}$ and $\gamma_S^{(2)}$ are the regression coefficients of the latent slope in Class 1 and 2. *N* is the total sample size. The bold numbers reflect conditions that do not contain the 95% confidence interval.

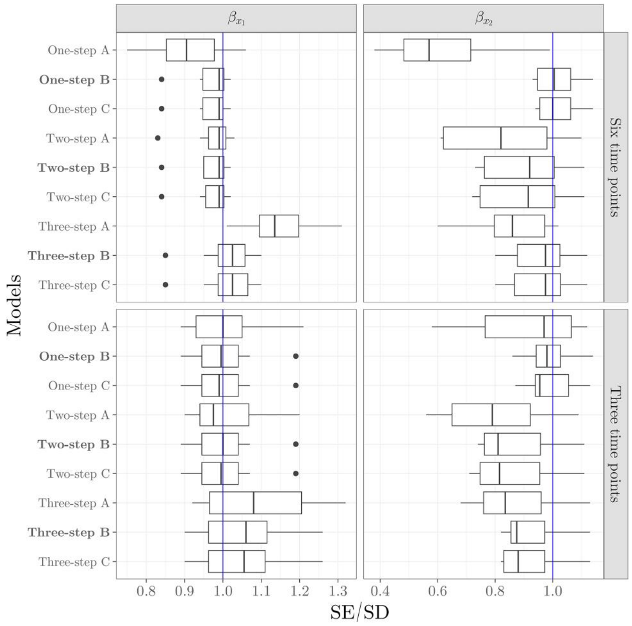


Figure 5. Boxplots of relative efficiency for regression coefficients of the latent class variable (β_{x_1} and β_{x_2}) averaged over 8 simulation conditions and 100 replications for study 2. *SE/SD* is the ratio of standard error versus standard deviation. The bold models are the estimators with correctly specified direct effects in study 2 (specification B).

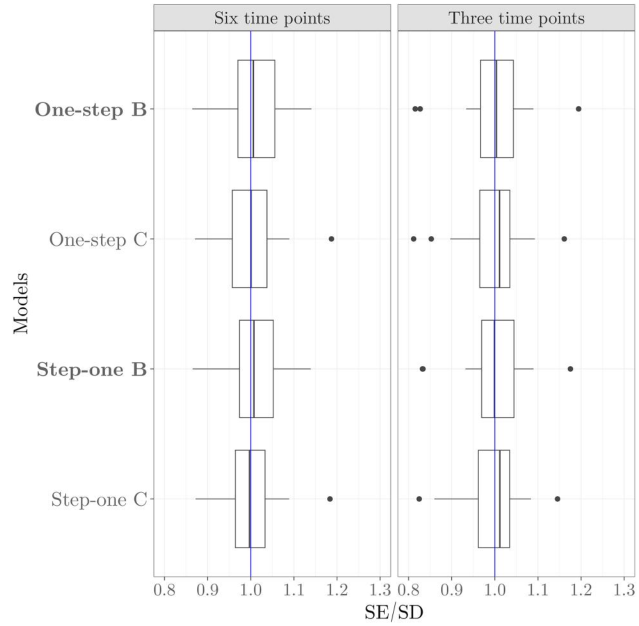


Figure 6. Boxplots of relative efficiency for the regression coefficient of latent intercept variable (γ_I) averaged over 2 classes, 8 simulation conditions, and 100 replications. *SE/SD* is the ratio of standard error versus standard deviation. The bold models are the estimators with correctly specified direct effects in study 2 (specification B).

two-step estimator performed comparably well producing efficient *SE* estimates except for the one-step with ignored DEs (specification A) over different time points conditions. The three-step estimator tends to slightly overestimate the *SE* over three specifications still having the means of *SE*/*SD* less than 1.15.

For the latent intercept coefficients γ_I in 6 time points condition, as we can see, all estimators with modeled DEs (specification B or C) performed systematically well with approximately unbiased parameter estimates across all 8 conditions in Table 2, in terms of the CRs of all three estimators are close to 95% across simulation conditions, except for having a slight undercoverage in the moderate entropy or smaller class size conditions. In Figure 6, we display the boxplots for *SE*/*SD* reported at each simulation condition for γ_I averaged over the two classes. All estimators with specified DEs (specifications B or C) were approximately unbiased under all time point conditions. The results of the AB and CRs for the β_{x_1}, β_{x_2}, and γ_I in 3 time points conditions were similar (see Appendix C).

Table 3 shows the Type I error rates of the γ_S for estimators with misspecified DEs under the 6 time points conditions over 100 replications. As the two-step and the three-step estimators share the same step-one model, we present the results of the step-one model (Step-one C) and the one-step estimator (One-step C) when the DEs are misspecified (specification C). The results show no considerably inflated Type I error rates across the conditions. The type I error rates are close to the expected value of 0.05 except in the unequal mixing ratio, large sample size, and well-separated classes conditions. The type I error rates of the γ_S in 3 time points are similar (as shown in Appendix C). We also inspected the mean square value of all interested parameters, the results are in line with the AB, CRs, and *SE*/*SD* values as shown in Appendix D.

Given the possibility of model misspecification in the applied settings, we ran a robustness check to assess the model performance when we not only misspecify the specific location of DEs but also the covariates with DEs, namely, we model the DEs on the latent intercept and slope from the covariates that only predict the class membership. The results are consistent with study 2, the two-step and three-step estimators are more robust than the one-step estimator when models are misspecified. We refer to the robustness check 2 in the supplementary material for more information.

Study 3: Population model III with covariates effects on class membership and latent intercept and slope

In study 3, data were sampled from the population model III where x_1 affects the LC membership, while

Table 3. The type I error rate values over 100 replications for regression coefficients of the latent slope (γ_S), in 6 time points conditions for study 2.

	Mixing ratio = 0.50/0.50				Mixing ratio = 0.30/0.70			
	High entropy		Moderate entropy		High entropy		Moderate entropy	
	$N = 1000$	$N = 500$	$N = 1000$	$N = 500$	$N = 1000$	$N = 500$	$N = 1000$	$N = 500$
One-step C	0.07	0.07	0.05	0.07	0.10	0.05	0.03	0.07
Step-one C	0.07	0.07	0.05	0.07	0.10	0.05	0.03	0.09

Note. One-step C is the one-step estimator model of the direct effects on both growth factors (specification C). Step-one C is the step-one model of the two-step and three-step estimators with specification C. N is the total sample size.

x_2 predicts the LC membership and also has DEs on the latent intercept and slope factors. Therefore, specification C is the correct way to specify the DEs. Table 4 presents the AB and 95% CI of the CRs for covariate effects on LCs (i.e., the β_{x_1} and β_{x_2}) under the 6 time points conditions. The results were similar to study 2. For the β_{x_2}, compared to ignoring the DEs (specification A), the performance of all three estimators systematically improved when we specified the DEs (specification B and C).

Concerning the design factors, the performance of all three estimators improved in well-separated class conditions, independent of the sample size and mixing ratio. Note that, when the DEs are misspecified (specification B), all three estimators tend to overestimate the β_{x_2} under conditions of moderate class separation and unequal mixing ratio conditions, in terms of AB range from 0.11 to 0.14 and lower CRs. For β_{x_1}, akin to results in study 2, all estimators performed systematically well in each condition, with negligible bias ranging from 0.00 to 0.04, and the nominal level of CRs falls into the CI. The results of the β_{x_1} and β_{x_2} in 3 time points were similar (see Appendix C).

Figure 7 displays the boxplot for *SE*/*SD* ratios reported at each replication in study 3. The results were averaged over the sample size, the mixing ratio, and the class separation conditions. We can observe similar results as in Study 2. Over 3 and 6 time points conditions, all estimators tend to underestimate the *SE*/*SD* values of β_{x_1} except for the one-step estimator with correctly specified DEs (specification C). When we specify DEs, the one-step estimator performed the best, closely followed by the three-step and the two-step estimators. For the β_{x_2}, all estimators were efficient except for the three-step estimator that slightly overestimates the *SE* values in 3 time points conditions (for the detailed *SE*/*SD* in each simulated condition, see Appendix E).

For the γ_I, the results in Table 4 show that all estimators with correctly specified DEs (specification C) performed systematically better than estimators with misspecified DEs (specification B) with regard to approximately unbiased parameter estimates and nominal level of CRs over the sample size, the class separation, and the mixing ratio conditions. The results for the 3 time points are presented in Appendix C. Similar results were observed from the *SE*/*SD* values reported at each replication averaged over 8 simulated conditions and 2 classes in Figure 8a, The estimators with specification C were more efficient concerning *SE* estimates.

Next, for the γ_S, we can see that all three estimators performed well with unbiased parameter estimates and the CRs reaching the nominal level for all simulated conditions and classes. For the *SE* estimators of γ_S, as shown in Figure 8b, all estimators with specification C performed well and tend to slightly underestimate the *SE*, over 8 simulated conditions and 2 classes. The results at 3 time points are presented in Appendix C. We also inspected the mean square value of all interested parameters, the results were in line with the AB, CRs, and *SE*/*SD* values as shown in Appendix D.

Real data example

In this paper, we applied all three estimators with different specifications of DEs on a real data example (9 models in total), which came from The China Health and Retirement Longitudinal Study (CHARLS; Zhao et al., 2013). This study focuses on the Chinese population aged over 45 years, understanding the socioeconomic determinants and outcomes of aging. CHARLS adopted a four-stage probability sampling procedure to ensure its nationally representative sampling. The baseline data were collected from subjects by personal interview in 2011, and the second, third, and fourth data waves were collected in 2013, 2015, and 2018 (Zhao et al., 2014). This analysis uses data or information from the Harmonized CHARLS dataset and Codebook, Version D, as of June 2021 developed by the Gateway to Global Aging Data. The development of the Harmonized CHARLS was funded by the National Institute on Aging (R01 AG030153, RC2 AG036619, R03 AG043052). For more information, we refer to their website.[3]

Table 4. The absolute bias (AB) (95% confidence interval of coverage rates) values over 100 replications for the effect of x_1 and x_2 on the latent class, latent slope, and latent intercept variables, in 6 time points Conditions for study 3.

	Mixing ratio = 0.50/0.50				Mixing ratio = 0.30/0.70			
	High entropy		Moderate entropy		High entropy		Moderate entropy	
	$N=1000$	$N=500$	$N=1000$	$N=500$	$N=1000$	$N=500$	$N=1000$	$N=500$
$\beta_{x_1}=-0.50$								
One-step A	0.00 (0.92–0.99)	0.00 (0.88–0.97)	0.61 (0.89–0.98)	0.41 (0.93–0.99)	0.02 (0.85–0.96)	0.04 (0.86–0.97)	0.01 (0.88–0.97)	0.06 (0.90–0.98)
One-step B	0.01 (0.93–0.99)	0.00 (0.89–0.98)	0.01 (0.92–0.99)	0.02 (0.92–0.99)	0.00 (0.89–0.98)	0.01 (0.88–0.97)	0.03 (0.92–0.99)	0.01 (0.89–0.98)
One-step C	0.01 (0.93–0.99)	0.00 (0.89–0.98)	0.02 (0.89–0.98)	0.02 (0.93–0.99)	0.00 (0.89–0.98)	0.01 (0.86–0.97)	0.01 (0.92–0.99)	0.01 (0.88–0.97)
Two-step A	0.01 (0.92–0.99)	0.00 (0.88–0.97)	0.01 (0.84–0.95)	0.01 (0.89–0.98)	0.00 (0.88–0.97)	0.02 (0.88–0.97)	0.02 (0.90–0.98)	0.00 (0.92–0.99)
Two-step B	0.01 (0.93–0.99)	0.00 (0.89–0.98)	0.01 (0.92–0.99)	0.02 (0.93–0.99)	0.00 (0.89–0.98)	0.01 (0.88–0.97)	0.03 (0.92–0.99)	0.00 (0.89–0.98)
Two-step C	0.01 (0.93–0.99)	0.00 (0.89–0.98)	0.02 (0.89–0.98)	0.02 (0.93–0.99)	0.00 (0.89–0.98)	0.01 (0.86–0.97)	0.01 (0.92–0.99)	0.00 (0.86–0.97)
Three-step A	0.01 (0.95–1.00)	0.00 (0.89–0.98)	0.02 (0.89–0.98)	0.00 (0.93–0.99)	0.00 (0.90–0.98)	0.02 (0.88–0.97)	0.02 (0.95–1.00)	0.01 (0.92–0.99)
Three-step B	0.01 (0.93–0.99)	0.00 (0.89–0.98)	0.02 (0.93–0.99)	0.02 (0.93–0.99)	0.00 (0.90–0.98)	0.01 (0.88–0.97)	0.03 (0.92–0.99)	0.01 (0.92–0.99)
Three-step C	0.01 (0.95–1.00)	0.00 (0.89–0.98)	0.02 (0.92–0.99)	0.02 (0.93–0.99)	0.00 (0.90–0.98)	0.01 (0.88–0.97)	0.02 (0.95–1.00)	0.01 (0.90–0.98)
$\beta_{x_2}=0.75$								
One-step A	0.06 (**0.74–0.89**)	0.06 (0.84–0.95)	2.11 (**0.01–0.07**)	1.84 (**0.02–0.10**)	0.27 (**0.39–0.59**)	0.26 (**0.40–0.60**)	0.85 (**0.00–0.04**)	0.98 (**0.00–0.04**)
One-step B	0.01 (0.93–0.99)	0.00 (0.88–0.97)	0.02 (0.86–0.97)	0.04 (0.86–0.97)	0.01 (0.90–0.98)	0.01 (0.90–0.98)	0.14 (**0.60–0.78**)	0.14 (**0.74–0.89**)
One-step C	0.01 (0.93–0.99)	0.01 (0.88–0.97)	0.00 (0.90–0.98)	0.00 (0.88–0.97)	0.00 (0.90–0.98)	0.00 (0.89–0.98)	0.00 (0.90–0.98)	0.01 (0.85–0.96)
Two-step A	0.04 (0.85–0.96)	0.03 (0.85–0.96)	0.13 (**0.18–0.35**)	0.14 (**0.44–0.63**)	0.04 (0.86–0.97)	0.04 (0.84–0.95)	0.23 (**0.19–0.36**)	0.26 (**0.32–0.51**)
Two-step B	0.01 (0.93–0.99)	0.00 (0.88–0.97)	0.02 (**0.80–0.93**)	0.04 (**0.74–0.89**)	0.01 (0.88–0.97)	0.01 (0.89–0.98)	0.14 (**0.39–0.59**)	0.14 (**0.60–0.78**)
Two-step C	0.01 (0.93–0.99)	0.01 (0.86–0.97)	0.00 (0.84–0.95)	0.00 (**0.78–0.91**)	0.00 (0.90–0.98)	0.00 (0.89–0.98)	0.00 (**0.77–0.91**)	0.01 (**0.79–0.92**)
Three-step A	0.03 (**0.80–0.93**)	0.03 (0.90–0.98)	0.15 (**0.38–0.57**)	0.11 (**0.59–0.77**)	0.04 (0.88–0.97)	0.05 (0.90–0.98)	0.25 (**0.22–0.40**)	0.25 (**0.41–0.61**)
Three-step B	0.01 (0.95–1.00)	0.00 (0.88–0.97)	0.02 (0.88–0.97)	0.03 (**0.81–0.94**)	0.01 (0.89–0.98)	0.01 (0.90–0.98)	0.11 (**0.65–0.82**)	0.11 (**0.78–0.91**)
Three-step C	0.00 (0.95–1.00)	0.01 (0.88–0.97)	0.00 (0.86–0.97)	0.00 (0.84–0.95)	0.00 (0.90–0.98)	0.00 (0.89–0.98)	0.00 (0.86–0.97)	0.00 (**0.83–0.94**)
$\gamma_I^{(1)}=0.50$								
One-step B	0.05 (**0.63–0.80**)	0.06 (**0.70–0.86**)	0.08 (**0.40–0.60**)	0.07 (**0.73–0.88**)	0.06 (**0.64–0.81**)	0.06 (**0.80–0.93**)	0.14 (**0.33–0.52**)	0.14 (**0.59–0.77**)
One-step C	0.00 (0.88–0.97)	0.00 (0.86–0.97)	0.00 (0.95–1.00)	0.00 (0.90–0.98)	0.01 (**0.83–0.94**)	0.00 (0.92–0.99)	0.01 (0.89–0.98)	0.01 (0.84–0.95)
Step-one B	0.05 (**0.61–0.79**)	0.06 (**0.68–0.84**)	0.08 (**0.40–0.60**)	0.07 (**0.73–0.88**)	0.06 (**0.65–0.82**)	0.06 (**0.80–0.93**)	0.14 (**0.32–0.51**)	0.14 (**0.59–0.77**)
Step-one C	0.00 (0.88–0.97)	0.00 (0.86–0.97)	0.00 (0.95–1.00)	0.00 (0.90–0.98)	0.01 (**0.83–0.94**)	0.00 (0.92–0.99)	0.01 (0.89–0.98)	0.01 (0.86–0.97)
$\gamma_I^{(2)}=-0.50$								
One-step B	0.05 (**0.67–0.83**)	0.04 (0.84–0.95)	0.07 (**0.58–0.76**)	0.05 (0.85–0.96)	0.05 (**0.60–0.78**)	0.05 (**0.68–0.84**)	0.04 (**0.76–0.90**)	0.05 (**0.71–0.87**)
One-step C	0.00 (0.86–0.97)	0.01 (0.93–0.99)	0.00 (0.86–0.97)	0.00 (0.89–0.98)	0.00 (0.93–0.99)	0.00 (0.88–0.97)	0.00 (0.88–0.97)	0.01 (0.86–0.97)
Step-one B	0.05 (**0.67–0.83**)	0.05 (**0.83–0.94**)	0.07 (**0.57–0.75**)	0.06 (0.84–0.95)	0.05 (**0.61–0.79**)	0.05 (**0.68–0.84**)	0.04 (**0.73–0.88**)	0.05 (**0.71–0.87**)
Step-one C	0.00 (0.86–0.97)	0.01 (0.93–0.99)	0.00 (0.85–0.96)	0.00 (0.88–0.97)	0.00 (0.93–0.99)	0.00 (0.88–0.97)	0.00 (0.86–0.97)	0.01 (0.85–0.96)
$\gamma_1^{(1)}=0.25$								
One-step C	0.00 (**0.81–0.94**)	0.00 (0.85–0.96)	0.00 (0.90–0.98)	0.01 (0.89–0.98)	0.01 (0.92–0.99)	0.00 (0.89–0.98)	0.00 (0.89–0.98)	0.00 (0.84–0.95)
Step-one C	0.00 (**0.81–0.94**)	0.00 (0.85–0.96)	0.00 (0.90–0.98)	0.01 (0.89–0.98)	0.01 (0.92–0.99)	0.00 (0.89–0.98)	0.00 (0.90–0.98)	0.00 (**0.83–0.94**)
$\gamma_1^{(2)}=-0.25$								
One-step C	0.00 (0.89–0.98)	0.00 (0.92–0.99)	0.00 (0.95–1.00)	0.00 (0.86–0.97)	0.00 (0.88–0.97)	0.01 (0.89–0.98)	0.00 (0.86–0.97)	0.00 (0.88–0.97)
Step-one C	0.00 (0.89–0.98)	0.00 (0.92–0.99)	0.00 (0.95–1.00)	0.00 (0.86–0.97)	0.00 (0.88–0.97)	0.01 (0.89–0.98)	0.00 (0.86–0.97)	0.00 (0.86–0.97)

Note. One-step A, One-step B, and One-step C are the one-step estimators ignoring the direct effects (DEs; specifications A), specifying the DE on the latent intercept (specification B), and on both growth factors (specification C), respectively. Two-step A, Two-step B, and Two-step C are the two-step estimators with specifications A, B, and C, respectively. Three-step A, Three-step B, and Three-step C are the three-step estimators with specifications A, B, and C, respectively. Step-one B and C are the step-one models of the two-step and three-step estimators with specifications B and C. β_{x_1} and β_{x_2} are regression coefficients of the latent class variable. $\gamma_I^{(1)}$ and $\gamma_I^{(2)}$ are the regression coefficients of the latent intercept in Class 1 and 2. $\gamma_S^{(1)}$ and $\gamma_S^{(2)}$ are the regression coefficients of the latent slope in Class 1 and 2. N is the total sample size. The bold numbers reflect conditions that do not contain the 95% confidence interval.

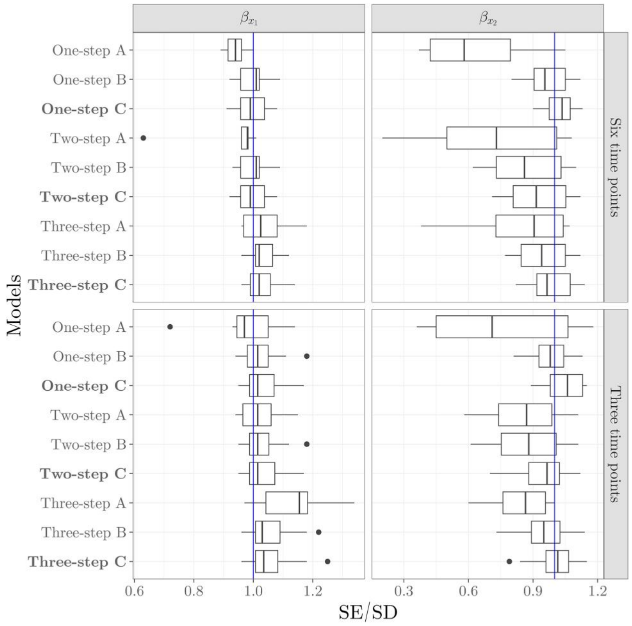
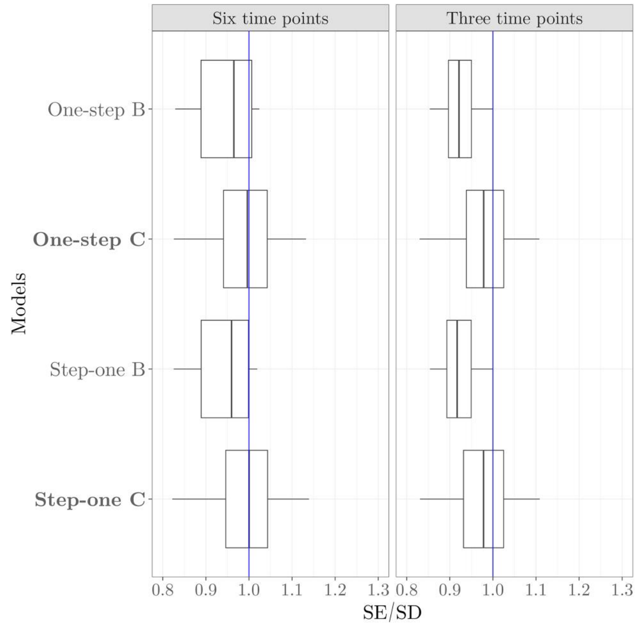


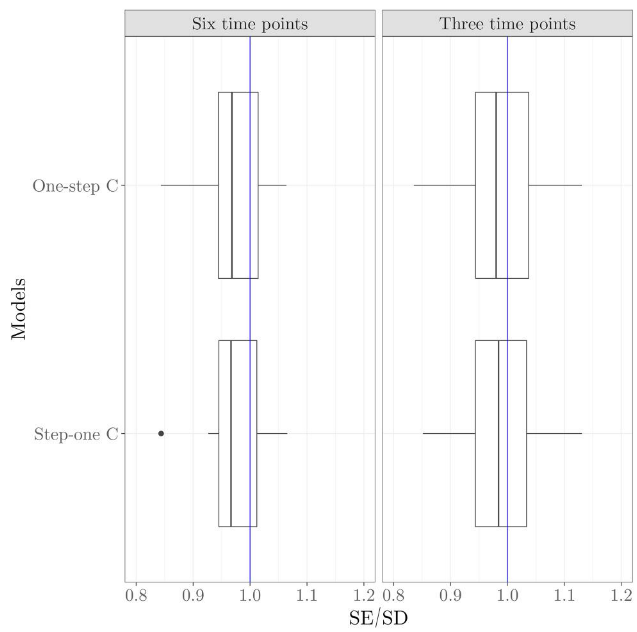
Figure 7. Boxplots of relative efficiency for regression coefficients of the latent class variable (β_{x_1} and β_{x_2}) averaged over 8 simulation conditions and 100 replications for study 3. *SE/SD* is the ratio of standard error versus standard deviation. The bold models are the estimators with correctly specified direct effects in study 3 (specification C).

Previous studies on CHARLS (Chen et al., 2024) detected different cognitive function trajectories of aging people with diabetes. The participants are classified as having diabetes based on their fasting plasma glucose (FPG) $>= 126 mg/dl$ measured after at least 8 h of fasting, or glycosylated hemoglobin (HbAlc) $>= 6.5\%$ (Bai et al., 2021; Roden, 2016) at baseline. We excluded participants with less than two follow-up repeated measures of cognitive function and with missing values in baseline covariates, age and education level. Data from 1259 participants were analyzed as part of the final sample.

For repeated measures, we selected four items to capture two dimensions of cognitive function, namely mental intactness and episodic memory. Specifically, items of date naming, drawing pictures, and serial subtracting 7 from 100 assessed mental intactness, item scores range from 0 to 11, and items of word recall (immediate and delayed) evaluated the episodic memory, which scores from 0 to 10. In line with previous studies, we computed the total score of cognitive function by summing the scores of mental intactness and episodic memory, ranging from 0 to 21. Higher scores reflect better cognitive function. Covariates were also chosen and re-coded following previous applied research. Two covariates were selected,



(a) The direct effect of x_2 on the latent intercept (γ_I).



(b) The direct effect of x_2 on the latent slope (γ_S).

Figure 8. Boxplots of relative efficiency for the regression coefficient of latent intercept (γ_I; Figure 8a) and latent slope (γ_S; Figure 8b) variables averaged over 2 classes, 8 simulation conditions, and 100 replications. *SE/SD* is the ratio of standard error versus standard deviation. The bold models are the estimators with correctly specified direct effects in study 3 (specification C).

including age and education level at baseline. Education level included three categories: no formal education, primary school, and middle school and above. The age was divided into a young group aged between 45 and 59, a middle group aged between 60 and 74, and an old group aged above 74.

[3]See https://g2aging.org/.

Table 5. Model fit statistics for the One-step and the Two-step estimators.

	Log-likelihood	BIC	AIC	AIC3	*df* difference	VLMR	*p*
One-step A	−9378.24	18,884.96	18,792.48	18,810.48			
One-step B	−9325.48	18,808.00	18,694.96	18,716.96	4.00	105.51	0.00***
One-step C	−9325.37	18,836.33	18,702.74	18,728.74	8.00	0.22	0.99
Two-step A	−9384.23	18,804.14	18,778.45	18,783.45			
Two-step B	−9339.03	18,713.75	18,688.06	18,693.06	0.00	90.39	0.00***
Two-step C	−9339.66	18,715.01	18,689.32	18,694.32	0.00	−1.26	0.75

*Note.*One-step A, One-step B, and One-step C are the one-step estimators ignoring the direct effects (DEs; specifications A), specifying the DE on the latent intercept (specification B), and on both growth factors (specification C), respectively. Two-step A, Two-step B, and Two-step C are the two-step estimators with specifications A, B, and C, respectively. BIC is the Bayesian information criterion. AIC is the Akaike information criterion. *df* different is the difference in the degree of freedom between the two models. VLMR is the Vuong-Lo-Mendell-Rubin test score. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

We follow the standard recommendation of Masyn (2017) and Diallo et al. (2017) to fit an unconditional GMM during the class enumeration process. A two-class model was selected in terms of fit measures (Bayesian information criterion [BIC] = 19,328.23, Vuong-Lo-Mendell-Rubin test [VLMR] = 91.18, $p < 0.01$; as shown in Appendix F). The class sizes of the two latent classes are close, and the specific item's score of the two-class model can be found in Appendix F. We labeled the first class as a moderate-decrease class, with a low cognitive function score at baseline and gradually decreased in the following waves. The second class, labeled high-stable, had a high cognitive function at baseline and the function score remained stable. Note that a three-class model was favored in previous literature (Chen et al., 2024). Here, we chose a more parsimonious model for illustration purposes.

To identify the existence of the DEs from covariates to growth factors, we follow the strategy of Masyn (2017) on detecting DEs in LCA. Thus, we estimated a two-class GMM that specified all the potential covariate effects on the latent class variable and growth factors by using the one-step estimator. The results revealed that the education level significantly impacts the intercept of cognitive function trajectories and class membership, while age only impacts class membership. Note that, as there are no optimal methods in detecting DEs in the GMM, we adopted the commonly used method (Masyn, 2017) in the LCA for illustration in this section. In this paper, we focus on the accuracy of parameter recovery and assume the covariate with DEs is prior knowledge. Developing methods for detecting DEs in the GMM remains an important direction for future research.

After identifying the latent classes, we extended the unconditional GMM by incorporating baseline covariates, which were estimated by using three estimators with three different specifications. For all three estimators, we ignored the DEs in specification A, modeled the DEs from the education level on the latent intercept in specification B, and modeled the DEs from the education level on both the latent intercept and slope in specification C. The effect of education level and age on the latent class variables was modeled across all specifications. As shown in Table 5, we present the model comparison of the one-step and the two-step estimators with varying specifications. Compared to estimators with specification A, a significant decrease in the log-likelihood value can be observed from both estimators with specification B, which is also reflected in other fit measures, e.g., BIC, AIC, and AIC3. Moreover, there is no significant improvement in model fit between estimators with specifications B and C, in terms of the presented fit measures. Note that we can not compare the three-step estimator as the models with varying specifications are not nested.

Next, we focused on the estimated regression coefficients of covariates on latent class variables for all estimators with different specifications, as shown in Table 6. The education level has a significant effect on class membership with all three estimators. As we applied effect coding on both covariates, participants with an education level of middle school and above tended to belong to a high-stable class compared to participants with the average education level, with higher baseline cognitive function and remaining stable across all waves. In contrast, participants with no formal education tend to belong to a moderate-decrease class compared to participants with the average education level, with lower baseline cognitive function and gradually decreasing cognitive function. For the one-step, the two-step, and the three-step estimators, the effects of education level are similar, and the effect of education level decreased but was still significant when we modeled DEs. Age has a significant effect on class membership with all three estimators. Specifically, compared to participants with average age, participants aged over 75 tend to belong to a moderate-decrease class with lower cognitive function over time, and younger participants tend to belong to a high-stable class. Using all three

Table 6. The Estimated regression coefficients (standard error) of the education level and age on the latent class variable using different estimators with varying specifications of direct effects.

	Education level			Age		
Models	No formal education	Primary school	Middle school and above	45 – 59	60 – 74	>= 75
One-step A	−1.75*** (0.32)	0.06 (0.17)	1.69*** (0.22)	0.40* (0.19)	−0.14 (0.11)	−0.54*** (0.12)
One-step B	−0.75*** (0.20)	0.07 (0.15)	0.69*** (0.17)	0.47** (0.17)	0.10 (0.10)	−0.58*** (0.11)
One-step C	−0.71** (0.25)	0.05 (0.21)	0.66** (0.22)	0.48** (0.17)	0.11 (0.10)	−0.58*** (0.12)
Two-step A	−2.39** (0.80)	0.37 (0.41)	2.02*** (0.42)	0.41* (0.19)	0.14 (0.12)	−0.55*** (0.13)
Two-step B	−0.99*** (0.22)	0.07 (0.13)	0.91*** (0.13)	0.40* (0.17)	0.05 (0.11)	−0.46*** (0.13)
Two-step C	−0.52*** (0.09)	−0.12 (0.08)	0.64*** (0.09)	0.30 (0.10)	−0.01 (0.09)	−0.29** (0.10)
Three-step A	−2.31*** (0.83)	0.20 (0.45)	2.11*** (0.55)	0.37 (0.27)	0.17 (0.16)	−0.54** (0.18)
Three-step B	−0.97* (0.40)	0.05 (0.22)	0.92*** (0.25)	0.40 (0.34)	0.05 (0.20)	−0.44* (0.23)
Three-step C	−0.51** (0.17)	−0.13 (0.13)	0.65*** (0.15)	0.29 (0.28)	−0.01 (0.17)	−0.28 (0.18)

Note. One-step A, One-step B, and One-step C are the one-step estimators ignoring the direct effects (DEs; specifications A), specifying the DE on the latent intercept (specification B), and on both growth factors (specification C), respectively. Two-step A, Two-step B, and Two-step C are the two-step estimators with specifications A, B, and C, respectively. Three-step A, Three-step B, and Three-step C are the three-step estimators with specifications A, B, and C, respectively. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 7. The estimated regression coefficients (standard error) of education level using different estimators with varying specifications of DEs on latent intercept and slope variables.

	Education level		
Models	No formal education	Primary school	Middle school and above
Latent intercept variable			
High-stable			
One-step B	−1.78*** (0.26)	0.04 (0.23)	1.72*** (0.38)
One-step C	−1.76*** (0.29)	0.05 (0.28)	1.71*** (0.42)
Step-one B	−4.77*** (0.48)	1.76 (0.27) ***	3.00*** (0.25)
Step-one C	−4.05*** (0.44)	1.46 (0.28) ***	2.60*** (0.23)
Moderate-decrease			
One-step B	−0.93 (0.55)	−0.16 (0.28)	1.09*** (0.31)
One-step C	−0.96 (0.58)	−0.16 (0.30)	1.13*** (0.32)
Step-one B	−0.10* (0.39)	−0.07 (0.28)	1.07* (0.45)
Step-one C	−0.56 (0.45)	−0.28 (0.32)	0.84 (0.48)
Latent slope variable			
High-stable			
One-step C	−0.05 (0.13)	0.01 (0.11)	0.05 (0.18)
Step-one C	−0.36 (0.21)	0.14 (0.92)	0.22* (0.10)
Moderate-decrease			
One-step C	−0.03 (0.28)	0.02 (0.15)	0.01 (0.16)
Step-one C	−0.12 (0.18)	−0.03 (0.12)	0.15 (0.19)

Note. One-step B and One-step C are the one-step estimators specifying the DE on the latent intercept (specification B), and on both growth factors (specification C), respectively. Step-one B and C are the step-one models of the two-step and three-step estimators with specifications B and C. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

estimators with varying specifications of DEs, the same overall conclusions are reached.

Table 7 presents the regression coefficients of covariates estimated by different estimators with varying specifications of DEs on latent intercept and slope variables. The education level has a significant effect on the latent intercept in the high-stable class for all three estimators. Within the high-stable class, Participants with no formal education tend to have lower cognitive function at baseline. For both classes, participants with education levels of middle school and above tend to have higher cognitive function at baseline. The estimated regression coefficients of education level on latent intercept are not significant, except for the coefficient estimated by the two-step and the three-step estimators at the education level of middle school and above in the high-stable class. And the estimated regression coefficients are larger using the two-step and the three-step estimators than the coefficients estimated by the one-step estimator. In general, there is no significant difference in the growth rates of cognitive function trajectory for participants with varying education levels across the two classes.

Discussion

In this article, we proposed a two-step estimator for the GMMs with covariates, and we considered a common situation in GMMs where DEs are present between the covariates and the growth factors within each class. We further proposed applying the DE modeling approach as presented by Vermunt and Magidson (2021a) for the stepwise estimators.

We compare our proposed method with currently available estimators in estimating GMMs, namely, the one-step and the three-step estimators. The one-step estimator is the FIML estimator, where all of the covariates are incorporated into the GMMs at once. For the three-step estimator, we also apply Vermunt and Magidson (2021a) proposed modeling strategy for estimating the DEs in latent class analysis (LCA) on the framework of the GMMs. For the three-step estimator, the step-one model is equivalent to the proposed two-step estimator. Then, we compute the conditional posterior probability of subjects given by the covariates with DEs at step two. In the step-three model, we estimate the class proportions conditional on all covariates (including the covariates with DEs) to address the overestimation of the covariate effect on class membership. As there is rarely prior knowledge on how to specify DEs in real settings, we also investigate how different specifications of DEs can influence the performance of different estimators, including ignoring, correctly specifying, and misspecifying the DEs.

The results of our simulation studies show that all three estimators work comparably well in the absence of the DEs from covariates on the growth factors. In the presence of DEs, ignoring the DEs leads to severe bias in the estimated parameters of the three estimators. The two-step and three-step estimators perform comparably with biased parameter estimates and outperform the one-step estimator. The one-step estimator is most sensitive to misspecification with substantial bias and the lowest coverage rates, especially when the level of class separation is moderate. When the DEs are specified, either correctly specified or over-specified, all estimators perform better and comparably well in terms of unbiased parameter estimates and coverage rates close to the nominal level. The results are in contrast to Diallo and Lu (2017), whose results showed that the three-step estimator with specified DEs performed worse than the one-step estimator when estimating the covariate effects on the LC variables. While Diallo and Lu (2017) used a naive approach to modeling DEs, they ignored the indirect effect through the LCs in the step one model. This results in overestimation of DEs on the growth factors. By using the approach of Vermunt and Magidson (2021a), unbiased estimates are obtained using both stepwise estimators. Furthermore, when we specify the plausible DEs path in Study 2, all three estimators are robust to misspecification in terms of acceptable type I error rates. However, in Study 3, when we under-specify the DEs, all estimators tend to overestimate the covariate effects on the LCs, leading to biased parameter estimates. Similar results are found by Di Mari and Bakk (2018), which investigates the performance of estimators on the Latent Markov models with varying specifications of DEs. They recommend that overspecifying the DEs is safer than underspecifying them when there is poor knowledge of which covariates could have DEs in the Latent Markov modeling. A similar strategy can also be applied in the GMMs. Moreover, the one-step estimator has higher type I error rates than the two-step and the three-step estimators when the DEs are mispecified.[4]

In this article, we also investigated the efficiency of different estimators by evaluating the *SE*/*SD* values. It turns out that the two-step and the three-step estimators are less efficient than the one-step estimator when the DEs are specified (either correctly specified or misspecified), tending to underestimate the *SE* values over all simulation conditions. Nevertheless, it is typical for step-wise estimators to slightly underestimate the *SE* values (Bakk & Kuha, 2018; Vermunt, 2010). The step-wise estimators, when obtaining the Hessian matrix of the last step model, ignore the variability in the step one estimates, treating those values as known, and thus ignoring the variance due to uncertainty about the step one parameters. Currently, approaches of correcting *SE*s in step-wise estimators are available for simpler LC models (Bakk & Kuha, 2018), but their extension to multiple LVs in GMM(e.g., latent intercept and slope) is not straightforward. Based on our simulation results, we do not recommend relying on the SE estimates based on the step-two Hessian matrix provided by the software, especially under the low-entropy conditions, and future research is needed to develop an accurate SE estimator. Additionally, introducing Bayesian inferential methods in GMMs to better account for the uncertainty is another possible solution, or alternatively, investigating the use of bootstrap standard errors.

In sum, when there are candidates of covariates with DEs but the specific location of DEs is unknown, we recommend using the one-step estimator with overfitting DEs on both growth factors, in terms of its better performance in efficiency and unbiased estimates, and acceptable type I error rates. When there are no clear candidates of covariates with DEs or for exploratory purposes, we recommend applying the two-step and the three-step estimators in terms of their robustness against model misspecification.

This article considers the number of classes and the DEs as known, and we chose a medium size of DEs on the growth factors. However, this information is often

[4]See also the results of the Robustness Check 2 in the supplementary material.

not available and needs to be investigated in many applied research contexts. In simple LC models, likelihood ratio test and residual statistics (i.e., bivariate residuals[BVR] and expected parameter change[EPC] statistics) are applied to identify the DEs from covariates to indicators (Bakk, 2024; Di Mari et al., 2023; Oberski et al., 2013). However, as the growth factors are latent variables in GMMs, the BVR can not be implemented. Moreover, penalization methods can also be a possible way to select covariates and detect significant DEs. The performance of other indices and methods in the GMM framework with varying effect sizes of DEs has not been studied yet. In this article, the method used to detect DEs in the real data example is not ideally suited for GMM and was employed solely for illustrative purposes. Our main interests focused on parameter estimation accuracy and efficiency over different estimators. Therefore, we do not specifically recommend its use for applied research. For the identification of DEs, the sensitivity of the one-step estimator when both the covariates and location of DEs are misspecified, can be beneficial to identify the presence of DEs, though this needs further investigation in the future.

In addition, this article focuses on a comprehensive evaluation of parameter recovery for different estimators when modeling DEs in GMM. To manage the scope of the simulations, class enumeration is not addressed. Nonetheless, class enumeration in GMM is a crucial and complex issue. Prior research in factor mixture models (FMM) has suggested that class enumeration should be conducted in the presence of covariates (Wang et al., 2023). In contrast, a substantial body of work has recommended excluding covariates during class enumeration, due to the heightened risk of model misspecification (Diallo et al., 2017; Lubke & Muthén, 2005; Nylund et al., 2007; Vermunt, 2010). Moreover, Stegmann and Grimm (2018) noted that while including covariates can improve class enumeration, it may also impair the model's ability to recover the true number of classes. Given these conflicting findings, we believe that further detailed investigation into the role of covariates in class enumeration is warranted, especially in the presence of model misspecification.

Our results also show that the performance of the one-step estimator is substantially affected by the level of class separation when the DEs are ignored, while it is unaffected with specified DEs. The two-step and the three-step estimators are quite sensitive to the level of class separation across varying specifications of DEs. Their performance systematically improved when the classes became more separated, yielding lower bias and higher coverage rates with smaller confidence intervals. As previous research shows the step-wise estimator is not recommended under low class separation conditions (Vermunt, 2010), we only consider moderate and high levels of class separation. However, in the applied setting, weak class separation can occur. For this case, further research can investigate whether using Bayesian methods in step one could improve the performance of two-step estimation in the context of GMMs.

Article information

Conflict of interest disclosures: Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

Ethical Principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: The work of Yuqi Liu was supported by Grant No.202307720051 from the China Scholarship Council.

Role of the funders/sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Acknowledgements: The authors would like to thank the editor, the associate editor Sarah Depaoli, and both reviewers for their comments on prior versions of this manuscript. The ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors' institution or the funding agency is not intended and should not be inferred.

ORCID

Yuqi Liu http://orcid.org/0009-0002-6464-4641
Zsuzsa Bakk http://orcid.org/0000-0001-9352-4812
Ethan M. McCormick http://orcid.org/0000-0002-7919-4340
Mark de Rooij http://orcid.org/0000-0001-7308-6210

References

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural Equation Modeling*, *21*(3), 329–341. https://doi.org/10.1080/10705511.2014.915181

Bai, A., Tao, J., Tao, L., & Liu, J. (2021). Prevalence and risk factors of diabetes among adults aged 45 years or older in China: A national cross-sectional study.

Endocrinology, Diabetes & Metabolism, *4*(3), e00265. https://doi.org/10.1002/edm2.265

Bakk, Z. (2024). Latent class analysis with measurement invariance testing: Simulation study to compare overall likelihood ratio vs residual fit statistics based model selection. *Structural Equation Modeling*, *31*(2), 253–264. https://doi.org/10.1080/10705511.2023.2233115

Bakk, Z., Di Mari, R., Oser, J., & Kuha, J. (2022). Two-stage multilevel latent class analysis with covariates in the presence of direct effects. *Structural Equation Modeling*, *29*(2), 267–277. https://doi.org/10.1080/10705511.2021.1980882

Bakk, Z., & Kuha, J. (2018). Two-step estimation of models between latent classes and external variables. *Psychometrika*, *83*(4), 871–892. https://doi.org/10.1007/s11336-017-9592-7

Bakk, Z., Oberski, D. L., & Vermunt, J. K. (2014). Relating latent class assignments to external variables: Standard errors for correct inference. *Political Analysis*, *22*(4), 520–540. https://doi.org/10.1093/pan/mpu003

Bakk, Z., Tekle, F. B., & Vermunt, J. K. (2013). Estimating the association between latent class membership and external variables using bias-adjusted three-step approaches. *Sociological Methodology*, *43*(1), 272–311. https://doi.org/10.1177/0081175012470644

Bolck, A., Croon, M., & Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis*, *12*(1), 3–27. https://doi.org/10.1093/pan/mph001

Bowers, A. J., & Sprott, R. (2012). Examining the multiple trajectories associated with dropping out of high school: A growth mixture model analysis. *The Journal of Educational Research*, *105*(3), 176–195. https://doi.org/10.1080/00220671.2011.552075

Chen, S., Ling, Y., Zhou, F., Qiao, X., & Reinhardt, J. D. (2024). Trajectories of cognitive function among people aged 45 years and older living with diabetes in China: Results from a nationally representative longitudinal study. *PLOS One*, *19*(5), e0299316. (20112018). https://doi.org/10.1371/journal.pone.0299316

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *39*(1), 1–22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

Diallo, T. M. O., & Lu, H. (2017). On the application of the three-step approach to growth mixture models. *Structural Equation Modeling*, *24*(5), 714–732. https://doi.org/10.1080/10705511.2017.1322516

Diallo, T. M. O., Morin, A. J. S., & Lu, H. (2017). The impact of total and partial inclusion or exclusion of active and inactive time invariant covariates in growth mixture models. *Psychological Methods*, *22*(1), 166–190. https://doi.org/10.1037/met0000084

Di Mari, R., & Bakk, Z. (2018). Mostly harmless direct effects: A comparison of different latent Markov modeling approaches. *Structural Equation Modeling*, *25*(3), 467–483. https://doi.org/10.1080/10705511.2017.1387860

Di Mari, R., Bakk, Z., Oser, J., & Kuha, J. (2023). A two-step estimator for multilevel latent class analysis with covariates. *Psychometrika*, *88*(4), 1144–1170. https://doi.org/10.1007/s11336-023-09929-2

Di Mari, R., Rocci, R., & Gattone, S. A. (2023). LASSO–penalized clusterwise linear regression modelling: A two–step approach. *Journal of Statistical Computation and Simulation*, *93*(18), 3235–3258. https://doi.org/10.1080/00949655.2023.2220058

Hipp, J. R., & Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychological Methods*, *11*(1), 36–53. https://doi.org/10.1037/1082-989X.11.1.36

Huang, D., Brecht, M.-L., Hara, M., & Hser, Y.-I. (2010). Influences of a covariate on growth mixture modeling. *Journal of Drug Issues*, *40*(1), 173–194. https://doi.org/10.1177/002204261004000110

Janssen, J. H. M., van Laar, S., de Rooij, M. J., Kuha, J., & Bakk, Z. (2019). The detection and modeling of direct effects in latent class analysis. *Structural Equation Modeling*, *26*(2), 280–290. https://doi.org/10.1080/10705511.2018.1541745

Kankaraš, M., Moors, G., & Vermunt, J. (2010). Testing for measurement invariance with latent class analysis. In *Cross-cultural analysis. Methods and applications* (pp. 359–384). Routledge.

Kim, M., Vermunt, J., Bakk, Z., Jaki, T., & Van Horn, M. L. (2016). Modeling predictors of latent classes in regression mixture models. *Structural Equation Modeling*, *23*(4), 601–614. https://doi.org/10.1080/10705511.2016.1158655

Li, L., & Hser, Y.-I. (2011). On inclusion of covariates for class enumeration of growth mixture models. *Multivariate Behavioral Research*, *46*(2), 266–302. https://doi.org/10.1080/00273171.2011.556549

Li, M., & Harring, J. R. (2017). Investigating approaches to estimating covariate effects in growth mixture modeling: A simulation study. *Educational and Psychological Measurement*, *77*(5), 766–791. https://doi.org/10.1177/0013164416653789

Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, *10*(1), 21–39. https://doi.org/10.1037/1082-989X.10.1.21

Masyn, K. E. (2017). Measurement invariance and differential item functioning in latent class analysis with stepwise multiple indicator multiple cause modeling. *Structural Equation Modeling*, *24*(2), 180–197. https://doi.org/10.1080/10705511.2016.1254049

McCutcheon, A. L. (1987). *Latent class analysis* (No. 07-064). Sage.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*(1), 107–122. https://doi.org/10.1007/BF02294746

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In *The SAGE handbook of quantitative methodology for the social sciences* (pp. 346–369). SAGE Publications, Inc.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, *55*(2), 463–469. https://doi.org/10.1111/j.0006-341x.1999.00463.x

Muthén, B. O., & Curran, P. J. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods*, *2*(4), 371–402. https://doi.org/10.1037/1082-989X.2.4.371

Nylund, K. L., Tihomir, A., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *14*(4), 535–569. https://doi.org/10.1080/10705510701575396

Oberski, D. L., van Kollenburg, G. H., & Vermunt, J. K. (2013). A Monte Carlo evaluation of three methods to detect local dependence in binary data latent class models. *Advances in Data Analysis and Classification*, *7*(3), 267–279. https://doi.org/10.1007/s11634-013-0146-2

Roden, M. (2016). Diabetes mellitus – Definition, Klassifikation und Diagnose. *Wiener Klinische Wochenschrift*, *128*(S2), 37–40. https://doi.org/10.1007/s00508-015-0931-3

Rosseel, Y., & Loh, W. W. (2024). A structural after measurement approach to structural equation modeling. *Psychological Methods*, *29*(3), 561–588. https://doi.org/10.1037/met0000503

Stegmann, G., & Grimm, K. J. (2018). A new perspective on the effects of covariates in mixture models. *Structural Equation Modeling*, *25*(2), 167–178. https://doi.org/10.1080/10705511.2017.1318070

Tofighi, D., & Enders, C. K. (2008). Identifying the correct number of classes in growth mixture models. *Advances in latent variable mixture models.*, *2007*(1), 317.

Tueller, S. J., Drotar, S., & Lubke, G. H. (2011). Addressing the problem of switched class labels in latent variable mixture model simulation studies. *Structural Equation Modeling*, *18*(1), 110–131. https://doi.org/10.1080/10705511.2011.534695

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, *18*(4), 450–469. https://doi.org/10.1093/pan/mpq025

Vermunt, J. K., & Magidson, J. (2021a). How to perform three-step latent class analysis in the presence of measurement non-invariance or differential item functioning. *Structural Equation Modeling*, *28*(3), 356–364. https://doi.org/10.1080/10705511.2020.1818084

Vermunt, J. K., & Magidson, J. (2021b). *LG-syntax user's guide: Manual for latent gold syntax module version 6.0*. Statistical Innovations Inc.

Wang, Y., Cao, C., & Kim, E. (2023). Covariate inclusion in factor mixture modeling: Evaluating one-step and three-step approaches under model misspecification and overfitting. *Behavior Research Methods*, *55*(6), 3281–3296. https://doi.org/10.3758/s13428-022-01964-8

Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, *22*(158), 209–212. https://doi.org/10.1080/01621459.1927.10502953

Zhao, Y., Hu, Y., Smith, J. P., Strauss, J., & Yang, G. (2014). Cohort profile: The China Health and Retirement Longitudinal Study (CHARLS). *International Journal of Epidemiology*, *43*(1), 61–68. https://doi.org/10.1093/ije/dys203

Zhao, Y., Strauss, J., Yang, G., Giles, J., Hu, P., Hu, Y., Lei, X., Liu, M., Park, A., Smith, J. P., & Wang, Y. (2013). *China health and retirement longitudinal study–2011–2012 national baseline users' guide*. National School of Development, Peking University.

Appendix A. Evaluation criteria

The regression parameters and standard error estimates of covariates were examined in terms of the absolute bias (AB), mean square error (MSE), 95% confidence interval (CI) of coverage rate (CR), and standard error ratio (*SE*/*SD*). The formula of AB, MSE, and the *SE*/*SD* of the interest parameter over 100 replications are as follows,

$$AB(\hat{\theta}) = \frac{\sum_{s=1}^{S} |\hat{\theta}_s - \theta|}{S},$$

$$MSE(\hat{\theta}) = \frac{\sum_{s=1}^{S} (\hat{\theta}_s - \theta)^2}{S}.$$

$$\frac{\sum_{s=1}^{S} SE(\hat{\theta}_s)/S}{SD(\hat{\theta})},$$

where θ, $\hat{\theta}$ and $\hat{\theta}_s$ are population value of the interested parameter, the estimated θ and its value from the s_{th} replication, s is the number of the current replication ($s = 1, 2, \ldots, S$). We expect the AB and MSE values to be close to 0, reflecting an accurate parameter estimator, and the *SE*/*SD* value to be 1, reflecting a robust *SE* estimator. We compute the 95% Wilson (1927) CI of CR, and we expect the nominal level of CR, which is 95%, to fall into its CI.

Appendix B. Manipulated factors

We present the population parameters chosen for manipulating mixing ratios and class separation levels.

- Mixing ratio: For class 1 and class 2: (1) 50%, 50%, by manipulating coefficient $\beta_0 = 0.75$ for Study 1 and $\beta_0 = -1.50$ for studies 2 and 3; and (2) 30%, 70%, by setting $\beta_0 = 1.61$ for Study 1 and $\beta_0 = -0.43$ for studies 2 and 3, respectively.
- Class separation: In study 1, $\alpha_0^{(1)} = 1.5$, $\alpha_0^{(2)} = -1.5$ in medium condition, $\alpha_0^{(1)} = 3$, $\alpha_0^{(2)} = -3$ in high condition. In studies 2 and 3, $\alpha_0^{(1)} = 0.5$, $\alpha_0^{(2)} = -0.5$ in medium condition, $\alpha_0^{(1)} = 1.5$, $\alpha_0^{(2)} = -1.5$ in high condition.