The impact of time coding on the growth mixture model with distal outcome

Liu, Yuqi

Universiteit Leiden

Faculteit der Sociale Wetenschappen

The impact of time coding on the growth mixture model with distal outcome

Liu Yuqi

Master's Thesis Psychology,

Methodology and Statistics Unit, Institute of Psychology

Faculty of Social and Behavioral Sciences, Leiden University

Date: August 2023

Student number: 3634779

Supervisor: Dr. Ethan M. McCormick & Dr. Zsuzsa Bakk

Abstract

Growth mixture models (GMMs) can be used to uncover the heterogeneity in individual growth trajectories, using the discrete latent class variable to predict between class differences of growth patterns in the continuous growth factors. GMMs can incorporate the distal outcomes, defined as the consequence of individuals in latent classes, allowing researchers to explore the consequence of the growth trajectories. Prior work on continuous latent growth variables as predictors of the distal outcome has shown that this prediction is sensitive to the time coding choices in the latent curve models (LCMs). However, the GMM allows categorical latent class variables to directly predict the distal outcome rather than the continuous latent growth factors in the LCMs, and may be less susceptible to the issues in the continuous case. Here, we examine the effect of the time coding approaches on the predictive relationship between categorical latent variables and the distal outcomes, and explore the role of estimation approaches, measurement quality, latent class proportions and sample sizes in determining the reliable recovery of the distal outcome and the latent classes. The results suggest that, the one-step approach, is robust across simulation conditions despite yielding downward biased estimates in standard error, but that the two-step approach results convergence issues, biased parameter recovery, and biased class proportions when the measurement model is weak. We demonstrate that similar issue occurs in both simulated and empirical data. In sum, there is no evidence for an influence of time coding choices on predicting the distal outcomes in the GMMs.

Table of Content

Introduction

Human cognition and behavior evolve in a dynamic process of development. Elaborating on how the developmental process unfolds over time gives more comprehensive understanding of human cognition and behavior. Longitudinal – or repeated measures – data is necessary to characterize individual growth trajectories to explore developmental processes and inter-individual and intra-individual differences. Many statistical approaches are available to help researchers analyze these longitudinal data and understand the differences.

The growth models can be used to capture longitudinal data with many individual growth trajectories over time. Two types of methods are popular for such models (Bauer & Curran, 2003), specifically, the growth models can be fitted within the multilevel modelling framework (Bryk & Raudenbush, 1987) and the structural equation modelling (SEM) framework (McArdle, 1988). The multilevel model was designed to deal with individuals nested within a group, such as staff nested within a company. This model also works for scenarios where repeated measures are nested in each individual, that is, by putting repeated measures at Level 1 and individuals at Level 2 in processing repeated measures data (Bryk & Raudenbush, 1987). In the SEM framework, when dealing with repeated measures data, the models portray the variation between repeated measures by assuming many latent growth trajectories. And these unobserved growth trajectories are characterized by latent variables, which are defined by observed repeated measures (Curran et al., 2010).

A variety of models have been developed within the SEM framework, and the latent curve model (LCM; Meredith & Tisak, 1990) is the most commonly used growth model based on the SEM. The unconditional linear LCM defines two latent growth factors by using repeated-measure observations as indicators, namely the latent intercept factor and the latent slope factor. These growth factors jointly define the pattern of individual growth trajectories, thus allowing to inspect the intraindividual difference during this period. We can also examine the interindividual difference by comparing the intercepts and slopes. In order to precisely model the growth trajectory of individuals, we also need to consider the impact of covariates on the individual growth process, and it is necessary to incorporate covariates into the LCM, namely conditional LCM (Curran et al., 2010). These types of covariates can be specified into time-varying covariates and time-invariant covariates in terms of whether they change over time. After controlling for the effect of covariates and identifying the latent growth patterns, researchers are also interested in exploring the long-term consequence of individuals in terms of their growth patterns, by inspecting the predictive relationship between the latent growth factors and the distal outcomes. In contrast to the covariates, the distal outcome variables are not extra indicators of latent growth patterns but rather a long after consequence after the current assessment waves.

As LCMs model variation in the variable of interest as a function of time, the time coding choices of repeated measures play an important role in the appropriate estimates and interpretability of the parameters in the growth models (Biesanz et al., 2004). When we examine the predictive effect of individual initial states or growth rates by incorporating the distal outcome variables into the LCMs, the effect of time coding is not only limited to the growth factors but also has a crucial influence on parameter estimates and the distal outcome interpretation. Previous work (McCormick et al., 2023) has shown that arbitrarily placed initial time points can impact the estimation of growth factors and the distal outcome, specifically, when we move the intercept from the initial time to a different time point, although the model fit is identical, the estimated regression coefficients of the latent slope factor reverse its sign and significance. Thus, the prediction effect changes from positive to negative and from significant to non-significant, resulting in the opposite interpretation. Given the impact on the substantive interpretation of the results, researchers should be thoughtful about the placement of initial time points when establishing the models. Prior work also proposed two approaches in finding the optimal time encoding method (McCormick et al., 2023), namely the use of aperture points as initial time points is recommended for research in which there is a need to obtain both unstandardized and standardized effects, while for research in which only standardized effects are required, the incremental validity models can be used to find the optimal modelling method.

For conventional LCM, one of the important assumptions is that all individuals are sampled from a single population and share the same set of growth parameters, and the individual variation should vary around the same growth trajectories. However, when the individuals are drawn from a heterogeneous population with unobserved subpopulations, this assumption is compromised. The growth mixture model (GMM) introduced by Muthén and Shedden (1999) can be used to identify underlying homogeneous subpopulations among heterogeneous populations, by incorporating continuous (i.e., latent slope and latent intercept) and categorical (i.e., latent class variable) latent variables to capture growth trajectories of each subpopulation. In contrast to conventional LCM, GMM relaxes the single population assumption, using latent class variables to identify the population heterogeneity by characterizing subpopulations, and latent growth factors to characterize different latent growth trajectories for each subpopulation (Muthén, 2004). Each subpopulation has its specific growth parameters (i.e., intercepts, slopes, error variances), therefor individual variation can vary around class-specific average growth trajectories rather than a common trajectory.

Auxiliary variables, similarly to LGMs, can be incorporated into the GMM models as covariates and distal outcome variables to participate in the latent classes enumeration process (Li & Hser, 2011) and understand the subsequent consequence of different growth trajectories. For covariates, the application research of GMMs can be subdivided into incorporating time-invariant variables and time-varying variables. The former impacts on investigating latent class membership assignment and to explain variations in growth factors within classes (Morin et al., 2011), while the latter has influences on class numeration process and estimation of parameters and standard errors (Diallo et al., 2017). For the distal outcomes, in GMMs, after the measurement model is defined, meaning that the number of latent classes is determined, researchers can incorporate the distal outcome variables to examine the predictive validity of latent trajectories. For example Harring and Hodis (2016) included a college major as a distal outcome to inspect the predictive effect of Domain-specific self-concepts of students' ability (SCA) trajectories in mathematics, they argued that persons with a stable SCA trajectory preferred to choose a mathematics-intensive major in college.

Since GMM is derived from the LCM approach, the choice of time coding in the GMM also should be carefully considered. The current research on the impact of time coding approaches, however, has only been drawn to traditional latent growth models with continuous latent variables. For more complex mixture models, the research on time coding effect has not been extended. In contrast to the LCMs, in GMMs, the discrete latent class variable is introduced to model the fact that individuals are drawn from different sub-populations, and can be used to directly predict distal outcomes. Moreover, as GMMs identify unobserved subpopulations with distinct growth trajectories, in this case, the choice of best time coding approach can be consistent or different across subpopulations and the choice of time coding approaches could be complicated. Considering the unknown effect of time coding approaches on the distal outcome prediction in the GMMs, it is worthy of further exploration.

The goal of this study is to inspect the impact of change in time coding on GMM with distal outcomes. First, we start with briefly presenting the model specification of LCM and GMM, and we expand the unconditional GMM by incorporating the distal outcome. Next, the design of simulation study is described which is followed by a presentation of the simulation result. In addition, an empirical example is presented for demonstration. Finally, we conclude with a summary of the main results, a discussion of the results, limitations, and recommendations for future works.

Models Specification

The latent curve model

The unconditional linear LCM with growth factors, as originally described by McArdle & Epstein, can be represented by using structural equation modelling notation, consisting of a measurement model (1) and structural model (2):

$$y_i = \Lambda\eta_i + \varepsilon_i, (1)$$

$$\eta_i = \alpha + \zeta_i, \quad (2)$$

Where y_i is a $p \times 1$ vector of repeated measure outcomes for individual i and p is the number of time points, η_i is a $q \times 1$ vector of latent growth factors for individual i where q is the number of growth factors, and Λ is a $p \times q$ matrix of factor loadings. The error term ε_i is a $p \times 1$ vector of time-specific errors for individual i , which is assumed to be distributed as $\varepsilon_i \sim MVN(0, \Theta)$, where Θ is a $p \times p$ diagonal covariance matrix, capturing the residuals of repeated measure outcomes at each time points. The equation 2 describes the model for latent variables, where α is a $q \times 1$ vector of the means of latent growth factors for individual i and ζ_i is a $q \times 1$ vector of the residuals with distribution $\zeta_i \sim MVN(0, \Psi)$, where Ψ is a $q \times q$ diagonal covariance matrix, indicating the individual variation from their means. We can combine the above equations into a reduced form:

$$y_i = \Lambda(\alpha + \zeta_i) + \varepsilon_i, \quad (3)$$

The following is the probability density function, the model-implied means and covariance equations:

$$f(y_i) = \Phi[y_i; \mu(\theta)\Sigma(\theta)], \quad (4)$$

$$\mu(\theta) = \Lambda\alpha, \quad (5)$$

$$\Sigma(\theta) = \Lambda\Psi\Lambda' + \Theta, \quad (6)$$

Where Φ is the probability density function for y_i, $\mu(\theta)$ is the model-implied means vector, θ is the number of all estimated parameters. $\Sigma(\theta)$ is the model-implied covariance matrix. And the coding of time is specified by the factor loading matrix Λ , which is defined as:

$$\Lambda = \begin{bmatrix} 1 & t_0 \\ 1 & t_1 \\ \vdots & \vdots \\ 1 & t_{T-1} \end{bmatrix}, (7)$$

where the two columns represent the intercept factor and slope factor at time point t $(t = 1,2, \ldots, T)$, respectively.

The growth mixture model

The growth mixture model relaxes the assumption of LCM that individuals come from a homogeneous population and allows the growth factors (i.e., slope and intercept) to vary across unobserved subpopulations, which is accomplished by combining categorical latent class variables (Bauer & Curran, 2003; Jung & Wickrama, 2008). As the LCM, the GMM also includes a measurement model (8) and a structural model (9):

$$\mathrm{y_{ic}} = \Lambda_\mathrm{c}\eta_\mathrm{ic} + \varepsilon_\mathrm{ic}, (8)$$

$$\eta_\mathrm{ic} = \alpha_\mathrm{c} + \zeta_\mathrm{ic}, (9)$$

Where the c subscripts in these equations indicate that each latent classes have their own parameters $(c = 1,2, \ldots, C)$, and C is represented the number of latent classes. And the i subscripts indicate that parameters differ across individuals. The combination form of Equations 8 and Equations 9 is:

$$\mathrm{y_{ic}} = \Lambda_\mathrm{c}(\alpha_\mathrm{c} + \zeta_\mathrm{ic}) + \varepsilon_\mathrm{ic}, (10)$$

And the probability density function, the model-implied means and covariance equations are:

$$f(y_i) = \sum_{c=1}^{C} \pi_c \, \Phi_c[y_i; \mu_c(\theta_c)\Sigma(\theta_c)], (11)$$

$$\mu_c(\theta_c) = \Lambda_c \alpha_c, (12)$$

$$\Sigma(\theta_c) = \Lambda_c \Psi \Lambda'_c + \Theta_c, (13)$$

Where π_c is the unconditional probability that an individual belongs to latent class c , and Φ_c is the multivariate probability density function for latent class c.

The GMM can be extended by including a distal outcome variable D_i as shown in Figure 1, which is basically an ANOVA model that D_i is the outcome variable, and we compare the difference of D_i among subpopulations. For simplicity, we incorporated a continuous distal outcome. The structural model of the basic GMM, the equation 14, is extended with:

$$\eta_i = \alpha + Bc_i + \zeta_i, (14)$$

The elements of the above equation are:

$$\begin{bmatrix} \eta_{ic} \\ D_i \end{bmatrix} = \begin{bmatrix} \alpha_c \\ \alpha_D \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \beta_c & 0 \end{bmatrix} \begin{bmatrix} \eta_{ic} \\ \eta_{iD} \end{bmatrix} + \begin{bmatrix} \zeta_{ic} \\ \zeta_{iD} \end{bmatrix}, (15)$$

Where D_i is the distal outcome for individual i, α_D is the intercept of the model for the distal outcome predicted from the latent class variable and β_c is the regression coefficient for latent class c , ζ_{iD} is the person-specific residual of the distal outcome.

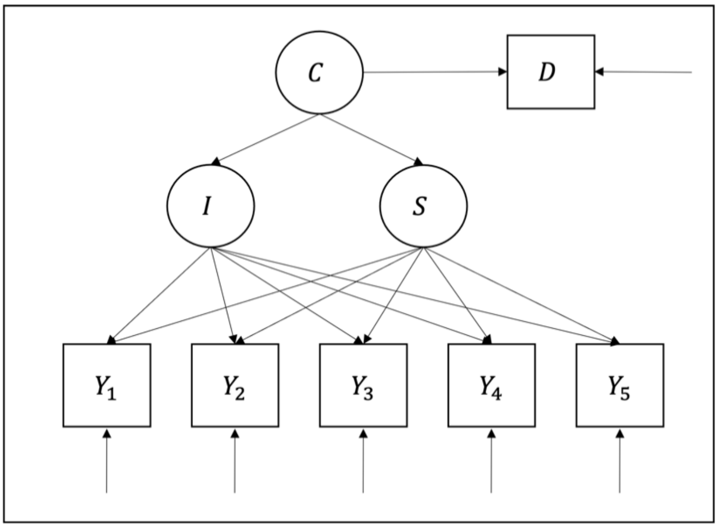
For the GMMs parameter estimation, we perform one-step approach (Bandeen-Roche et al., 1997) and two-step approach (Bakk & Kuha, 2018) in this study. In the one-step approach, the distal outcome variable will be treated as extra observed indicator for latent variables in the measurement model, and the measurement model as well as structural model are estimated at one time by using full information maximum likelihood (FIML) estimation. In two-step approach, we first exclude the distal outcome variable and only estimate the measurement model. After the parameters of measurement model are fixed, we will estimate the remain parameters of the structural model, thus, the pseudo-maximum likelihood (PML) estimation is performed(Gong & Samaniego, 1981).

Simulation Design

In the current simulation study, the population model is a GMM with 3 classes, measured by five repeated measures and explaining a continuous distal outcome variable, as specified in Figure 1. The simulation datasets were generated from a heterogenous population with three groups (i.e., group 1, group 2, group 3). Based on this model, in terms of whether the variable change as time, the main components in the datasets can be specified into time-invariant measures including a continuous distal outcome variable, a categorical variable representing the true group of individuals and a set of intercept and slope variables for each group, and time-varying measures, i.e., the five repeated measure variables.

Figure 1

The Growth Mixture Model with a Distal Outcome.



Note. C = latent class variable; I = latent intercept variable; S = latent slope variable; D = continuous distal outcome variable.

In each group, for time-invariant measures, two uncorrelated growth factors were generated, corresponding to an intercept ($\alpha_I = 3.00, \psi_{II} = 1.00$) and slope ($\alpha_S = 0.20, \psi_{SS} = 0.25$) factor and the continuous distal outcome variable (D_i) was defined with class-specific means ($\mu_{group1} = 0, \mu_{group2} = 3, \mu_{group3} = 5,$) and generated with a normally distributed residual variance.

For time-varying measures in each group, five continuous repeated measures were generated based on the slope and intercept factors with a normally distributed residual variance, the corresponding factor loading matrixes Λ of growth factors for each group are as followed:

$$\Lambda_{group1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \Lambda_{group2} = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \Lambda_{group3} = \begin{bmatrix} 1 & -4 \\ 1 & -3 \\ 1 & -2 \\ 1 & -1 \\ 1 & 0 \end{bmatrix}$$

In other words, we defined the intercept to represent the initial time points for group 1, to represent the middle time points for group 2 and to represent the final time points for group 3. And the effect size of the measurement model was manipulated by setting the R^2 of measurement model to satisfy the design condition of different effect sizes as described below.

We generated data across three different design factors: the sample size (N), the effect size of measurement model and the latent class proportions. We generated small, medium and large sample sizes ($N = 200, 500, 1000$) conditions consistent with typical applications in the literature. And the sample size is combined from three groups within the heterogeneous population.

For the latent class proportions, we manipulated the class proportions to be balanced or imbalanced across latent classes, i.e., the proportion is 33.3% of each class in the balanced class proportion scenario for N =200, 500, 1000 and the proportion of classes is 20%, 30% and 50% in the unbalanced scenario.

The other important design factor, the effect size of the measurement model, was manipulated by changing the strengths of association between latent classes and the observed outcomes, which was quantified in R^2. The effect size implemented in the study was two conditions, R^2 = 0.40 and R^2 = 0.80, reflecting weak and strong measurement quality conditions respectively. Previous literature revealed that improving measurement quality can help to reduce the bias of estimated latent class proportions (Lanza et al., 2013). We would expect better model performance (e.g., less bias) in latent class recovery under strong effect size condition ($R^2 = 0.80$) rather than weak condition ($R^2 = 0.40$).

In summary, three fully crossed design conditions (sample size, effect size and latent class proportion) yield a total of 12 conditions ($3 \times 2 \times 2$) in data generation.

To compare the effect of time coding when performing different estimation approaches, we were using a one-step approach (Bandeen-Roche et al., 1997) and a two-step approach (Bakk & Kuha, 2018) for parameter estimation under the 12 conditions. Moreover, the statistical models that fitted the simulated datasets, used three GMMs with a distal outcome - each with different time coding schemes, specifically, we placed the intercept at different time points: including placing the intercept at an initial time point (model a), middle time point (model b) and final time point (model c). The corresponding factor loading matrixes Λ of growth factors are as followed:

$$\Lambda_a = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \Lambda_b = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \Lambda_c = \begin{bmatrix} 1 & -4 \\ 1 & -3 \\ 1 & -2 \\ 1 & -1 \\ 1 & 0 \end{bmatrix}$$

In total, there are 6 (3×2)statistical models estimated on each simulated dataset.

The effect of the latent class variable on the distal outcome across different time coding approaches will be evaluated by inspecting the recovery of the mean of the distal outcome, the latent class proportions, and latent class memberships. The evaluation criteria for each simulation condition will include the estimated latent class proportions and the mean of the distal outcome, the bias of the estimates, the mean of estimated standard error (SE), the standard deviation (SD) of the estimated parameters, the entropy-based R^2 and the classification error. We will report the ratio of the SE and the SD and expect the value to be 1, which indicates models perform well in SE estimates. For data simulation and setting up statistical models, statistical programs LatentGOLD (Vermunt & Magidson, 2021) and R were used.

Simulation Results

Label switching and coverage rates

The label switching issue can arise during estimating the GMMs, namely the order of latent classes may arbitrarily switch across data sets for each replication. Although this will not affect the model fit, the estimation of means of the distal outcome will switch accordingly. In the study, we can detect the label switching issues by comparing the estimated parameters with the known data generating parameters and re-ordering the estimated latent classes and the means of the distal outcome to force them to match the generating order. Considering the latent class proportions were difficult to reorder under balanced class proportions condition, and since the strength of the effect of the latent class variables on the distal outcome was not strong, resulting in the distal outcome predominated the estimation of latent class parameters, the generating means of the distal outcome were used as reorder references. Thus, we first reordered the estimated means of the distal outcome and used the sequence of reordered distal outcomes to rank the estimated latent class proportions.

Another issue that may arise during fitting the GMM to the data sets is non-convergence. A nonconverging model can be the result of several problems, including model nonidentification, multivariate outliers in the dataset, empirical nonidentification and specifying the wrong model (Tueller & Lubke, 2010). In this study, we excluded the replications that did not converge, and only reported parameters from models which converged successfully. For all conditions, the convergence rates can be found in Table 1. We expect that it will be easier to obtain convergence when we fit the model on data set with a large sample size and strong effect size of the measurement model. For GMMs using the one-step approach, the convergence rates with large effect sizes were higher than with small effect sizes regardless of the time coding scheme and balanced or imbalanced latent class proportions. Using the two-step estimator, the convergence rates at step one was lower across all conditions, except for the conditions combined with $N = 200$ and effect size = 0.8. As we increased the effect size, the convergence rates substantially improved. As there is only 71% of the sample successfully converged on average, this is too low thus we do not look at step two results in more detail.

In general, the large effect size conditions performed better in convergence rate than the small effect size condition. Considering there might be less confidence for the models which have a hard time converging since there are fewer iterations, we only discussed the results of means of the distal outcome that used one-step approaches. And all the results presented in the hereafter tables should be interpreted under the context of convergence rates.

Table 1

Convergence Rates (out of 100 Replications) Collapsed across all Conditions

			One-step		Step one of the Two-step	
N	Effect Size	Time	Balanced	Imbalanced	Balanced	Imbalanced
N=1000	Small (0.4)	initial	0.83	0.81	0.49	0.49
		middle	0.82	0.89	0.43	0.56
		final	0.83	0.84	0.46	0.55
	Large (0.8)	initial	0.97	0.96	0.85	0.87
		middle	0.97	0.97	0.92	0.89
		final	0.99	0.96	0.85	0.82
N=500	Small (0.4)	initial	0.86	0.80	0.50	0.53
		middle	0.84	0.85	0.49	0.51
		final	0.83	0.83	0.50	0.59
	Large (0.8)	initial	0.97	0.94	0.89	0.84
		middle	0.99	0.96	0.94	0.88
		final	0.97	0.96	0.88	0.87
N=200	Small (0.4)	initial	0.88	0.83	0.51	0.59
		middle	0.89	0.84	0.54	0.66
		final	0.81	0.85	0.55	0.64
	Large (0.8)	initial	0.98	0.97	0.97	0.87
		middle	0.99	0.96	0.94	0.92
		final	0.98	0.98	0.93	0.87

Note. N

= Total sample size; Effect Size = The effect size of measurement model; Balanced = Balanced class proportions; Imbalanced = Imbalanced class proportions; Time = Time coding schemes; Initial = Placing the intercept at the initial time point; Middle = Placing the intercept at middle time point; Final = Placing the intercept at final time point; One-step = One step approach; Step one of Two-step = Step one of the two-step approach.

Latent class proportions

In this section, we presented the results of the latent class proportion recovery using the one-step and two-step methods, respectively.

For one-step approach, Table 2 and Table 3 presented the estimated latent class proportions and SEs using the one-step approach for all sample size and effect size conditions, in balanced and imbalanced latent class proportion conditions respectively. We compared the performance of models using different time coding schemes in estimating latent class proportions. It can be easily noticed that models estimated latent class proportions matched the generating values irrespective of time coding approaches used for the balanced class proportion conditions, the value of bias was between 0.00 and 0.04 as shown in Table 2. For the imbalanced conditions, as shown in Table 3, using either of time coding approaches, the value of bias was larger in the 200 sample-size conditions, and compared to the generating class proportion, the models tended to estimate more equal class proportions by overestimating the smallest latent class and underestimating the largest class. As we can see, the values of the ratio of the SE and the SD were lower than 1, indicating that the SE was underestimated regardless of sample size, effect size and latent class proportions. As we increased the effect size, SE estimates increased but remained underestimated. When the sample size was 500 or 1000 and the effect size was set to 0.8, all three models using different time coding schemes performed well. That is, the magnitude of bias in SEs estimates was smallest for each latent class and the latent class proportions properly matched the generated class proportions regardless of balanced or imbalanced class proportions conditions.

Table 2

The latent class proportion recovery for GMMs with different time coding approaches using the one-step approach on simulated data under balanced latent class proportions

			Initial			Middle			Final		
N	Effect Size	Class	*Est.*	*Bias*	SE/SD	*Est.*	*Bias*	SE/SD	*Est.*	*Bias*	SE/SD
Balanced latent class proportion (C1/C2/C3=0.33/0.33/0.33)											
1000	Small (0.4)	C1	0.32	-0.01	0.27	0.30	-0.04	0.24	0.30	-0.04	0.27
		C2	0.36	0.03	0.37	0.36	0.03	0.36	0.36	0.02	0.39
		C3	0.35	-0.01	0.23	0.34	0.01	0.24	0.35	0.01	0.26
	Large (0.8)	C1	0.32	-0.02	0.60	0.32	-0.01	0.57	0.33	-0.01	0.53
		C2	0.34	0.01	0.68	0.34	0.01	0.63	0.33	0.00	0.62
		C3	0.34	0.01	0.53	0.34	0.00	0.54	0.34	0.01	0.58
500	Small (0.4)	C1	0.35	0.02	0.42	0.33	0.00	0.35	0.32	-0.01	0.35
		C2	0.33	-0.01	0.50	0.34	0.01	0.48	0.32	-0.01	0.41
		C3	0.32	-0.01	0.32	0.33	0.00	0.29	0.36	0.02	0.31
	Large (0.8)	C1	0.34	0.01	0.62	0.34	0.01	0.49	0.34	0.01	0.57
		C2	0.33	0.00	0.63	0.34	0.01	0.62	0.33	0.00	0.58
		C3	0.33	0.00	0.65	0.32	-0.01	0.58	0.33	-0.01	0.54
200	Small (0.4)	C1	0.32	-0.01	0.46	0.32	-0.02	0.47	0.31	-0.03	0.46
		C2	0.34	0.01	0.53	0.33	0.00	0.49	0.34	0.01	0.47
		C3	0.34	0.00	0.41	0.35	0.02	0.41	0.35	0.02	0.38
	Large (0.8)	C1	0.34	0.01	0.52	0.34	0.01	0.49	0.33	0.00	0.50
		C2	0.31	-0.02	0.52	0.31	-0.02	0.47	0.34	0.01	0.50
		C3	0.35	0.01	0.45	0.35	0.01	0.42	0.33	-0.01	0.46

Note. $N =$ Sample size; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; S.E.= Mean of Standard Error of the estimates; Effect Size = The effect size of measurement model; Est. = Mean of the model estimates; SD = The standard deviation of the estimated parameter; Bias = Mean estimated minus true value.

Table 3

The latent class proportion recovery for GMMs with different time coding approaches using the one-step approach on simulated data under imbalanced latent class proportions

			Initial			Middle			Final		
N	Effect Size	Class	*Est.*	*Bias*	SE/SD	*Est.*	*Bias*	SE/SD	*Est.*	*Bias*	SE/SD
Imbalanced latent class proportion (C1/C2/C3=0.20/0.30/0.50)											
1000	Small (0.4)	C1	0.21	0.01	0.35	0.22	0.02	0.28	0.20	0.00	0.34
		C2	0.33	0.03	0.38	0.31	0.01	0.36	0.31	0.01	0.38
		C3	0.46	-0.04	0.27	0.48	-0.02	0.26	0.49	-0.01	0.27
	Large (0.8)	C1	0.21	0.01	0.61	0.20	0.00	0.62	0.21	0.01	0.71
		C2	0.33	0.03	0.60	0.33	0.03	0.62	0.34	0.04	0.47
		C3	0.47	-0.03	0.50	0.47	-0.03	0.51	0.45	-0.05	0.41
500	Small (0.4)	C1	0.22	0.02	0.39	0.21	0.01	0.40	0.21	0.01	0.39
		C2	0.31	0.01	0.47	0.31	0.01	0.51	0.29	-0.01	0.34
		C3	0.47	-0.03	0.34	0.49	-0.01	0.38	0.50	0.00	0.25
	Large (0.8)	C1	0.21	0.01	0.59	0.21	0.01	0.53	0.22	0.02	0.55
		C2	0.33	0.03	0.71	0.35	0.05	0.48	0.35	0.05	0.47
		C3	0.46	-0.04	0.60	0.44	-0.06	0.41	0.44	-0.06	0.39
200	Small (0.4)	C1	0.23	0.03	0.51	0.21	0.01	0.53	0.21	0.01	0.55
		C2	0.33	0.03	0.55	0.34	0.04	0.56	0.32	0.02	0.48
		C3	0.44	-0.06	0.41	0.45	-0.05	0.40	0.46	-0.04	0.38
	Large (0.8)	C1	0.24	0.04	0.56	0.24	0.04	0.52	0.25	0.05	0.64
		C2	0.33	0.03	0.44	0.33	0.03	0.39	0.35	0.05	0.44
		C3	0.43	-0.07	0.38	0.43	-0.07	0.34	0.40	-0.10	0.40

Note. N = Sample size; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; S.E.= Mean of Standard Error of the estimates; Effect Size = The effect size of measurement model; Est. = Mean of the model estimates; SD = The standard deviation of the estimated parameter; Bias = Mean estimated minus true value.

For two-step approach, in Table 4 and Table 5, the estimated latent class proportions were reported across all conditions for step one model. For all twelve conditions, on average the bias of estimated latent class proportions was high for the first class and the third class across different time coding approaches, specifically, the models overestimated the smallest class and underestimated the largest class. It is likely that models tended to generate latent classes with higher separation. We inspected model performance in SE estimates using two-step approach by comparing the ratios of the SE and the SD across all conditions of the step one model. As can be seen, on average the value of ratios severely diverged from 1, which shows that the two-step approaches in the first step yielded severe bias in SE estimates across different time coding schemes.

Table 4

The latent class proportion recovery for GMMs with different time coding approaches using the two-step approach on simulated data under balanced latent class proportions

			Initial			Middle			Final		
N	Effect Size	Class	*Est.*	*Bias*	SE/SD	*Est.*	*Bias*	SE/SD	*Est.*	*Bias*	SE/SD
Balanced latent class proportion (C1/C2/C3=0.33/0.33/0.33)											
1000	Small (0.4)	C1	0.15	-0.18	0.58	0.17	-0.16	0.77	0.18	-0.16	0.61
		C2	0.31	-0.03	0.69	0.31	-0.02	0.79	0.30	-0.03	0.86
		C3	0.54	0.21	0.63	0.52	0.19	0.68	0.52	0.19	0.74
	Large (0.8)	C1	0.16	-0.17	1.06	0.16	-0.17	0.91	0.16	-0.17	0.88
		C2	0.34	0.00	1.62	0.35	0.01	1.47	0.35	0.01	1.54
		C3	0.50	0.17	1.33	0.49	0.16	1.35	0.49	0.17	1.13
500	Small (0.4)	C1	0.16	-0.17	0.79	0.16	-0.17	0.74	0.15	-0.19	0.67
		C2	0.30	-0.03	1.02	0.29	-0.04	0.91	0.30	-0.03	0.88
		C3	0.54	0.21	0.73	0.55	0.21	0.67	0.55	0.22	0.69
	Large (0.8)	C1	0.16	-0.17	1.14	0.14	-0.19	0.89	0.15	-0.18	0.96
		C2	0.32	-0.01	1.40	0.33	-0.01	1.32	0.33	0.00	1.47
		C3	0.52	0.18	1.12	0.53	0.20	1.09	0.52	0.19	1.03
200	Small (0.4)	C1	0.16	-0.17	1.02	0.17	-0.16	1.19	0.16	-0.17	1.09
		C2	0.30	-0.03	1.33	0.29	-0.05	1.17	0.28	-0.05	1.24
		C3	0.53	0.20	1.05	0.54	0.21	0.95	0.55	0.22	0.89
	Large (0.8)	C1	0.15	-0.18	1.05	0.15	-0.19	0.99	0.14	-0.19	0.88
		C2	0.32	-0.01	1.49	0.32	-0.01	1.30	0.31	-0.02	1.25
		C3	0.53	0.19	1.19	0.53	0.20	1.04	0.55	0.22	1.00

Note. N = Sample size; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; S.E.= Mean of Standard Error of the estimates; Effect Size = The effect size of measurement model; Est. = Mean of the model estimates; SD = The standard deviation of the estimated parameter; Bias = Mean estimated minus true value.

Table 5

The latent class proportion recovery for GMMs with different time coding approaches using the two-step approach on simulated data under imbalanced latent class proportions

			Initial			Middle			Final		
N	Effect Size	Class	*Est.*	*Bias*	SE/SD	*Est.*	*Bias*	SE/SD	*Est.*	*Bias*	SE/SD
Imbalanced latent class proportion (C1/C2/C3=0.20/0.30/0.50)											
1000	Small (0.4)	C1	0.16	-0.04	0.62	0.16	-0.04	0.55	0.15	-0.05	0.50
		C2	0.31	0.01	0.73	0.30	0.00	0.60	0.30	0.00	0.93
		C3	0.53	0.03	0.55	0.54	0.04	0.46	0.55	0.05	0.68
	Large (0.8)	C1	0.14	-0.06	1.10	0.14	-0.06	0.91	0.15	-0.05	1.00
		C2	0.34	0.04	1.59	0.34	0.04	1.95	0.34	0.04	1.74
		C3	0.52	0.02	1.28	0.52	0.02	1.13	0.51	0.01	1.22
500	Small (0.4)	C1	0.15	-0.05	0.65	0.16	-0.04	0.79	0.13	0.07	0.65
		C2	0.31	0.01	0.90	0.31	0.01	1.10	0.27	0.08	0.91
		C3	0.54	0.04	0.72	0.53	0.03	0.77	0.60	0.12	0.67
	Large (0.8)	C1	0.12	-0.08	0.94	0.12	-0.08	0.83	0.14	-0.06	0.91
		C2	0.32	0.02	1.34	0.32	0.02	1.21	0.33	0.03	1.57
		C3	0.55	0.05	1.04	0.56	0.06	1.02	0.53	0.03	1.06
200	Small (0.4)	C1	0.15	-0.05	1.02	0.15	-0.05	0.80	0.14	-0.06	1.06
		C2	0.28	-0.02	1.16	0.27	-0.03	0.89	0.28	-0.02	1.27
		C3	0.56	0.06	0.91	0.57	0.07	0.73	0.58	0.08	1.07
	Large (0.8)	C1	0.12	-0.07	0.86	0.12	-0.08	0.92	0.14	-0.06	0.98
		C2	0.29	-0.01	1.42	0.29	-0.01	1.23	0.29	-0.01	1.62
		C3	0.58	0.08	1.03	0.59	0.09	0.97	0.57	0.07	1.09

Note. $N =$ Sample size; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; S.E.= Mean of Standard Error of the estimates; Effect Size = The effect size of measurement model; Est. = Mean of the model estimates; SD = The standard deviation of the estimated parameter; Bias = Mean estimated minus true value.

Latent class membership assignment

Two indices measuring latent class assignment uncertainty (i.e., the classification error and the entropy-based R^2) were reported to assess the recovery of latent class membership. In Table 6 and Table 7, the results over two effect sizes and three sample sizes conditions for using one-step approach and two-step approach are presented. It can be observed that the models performed poorly across all the conditions, resulting in low values of the entropy based R^2, especially for models using two-step approaches, as shown in Table 7, the values were lower than 0.60.

Another unexpected/ counterintuitive pattern emerged as can be seen in Table 6, for the conditions with the smallest sample size ($N = 200$), the models performed well across all time coding approaches regardless of measurement quality and latent class proportions, resulting in the highest value of entropy-based R^2 and the lowest classification error. Meanwhile, under the larger sample size ($N = 500\ or\ 1000$) conditions, the model performed poorly, resulting in a higher classification error and the lower entropy-based R^2. A similar pattern can be seen in Table 7, when we were using two-step approach, comparing the entropy values and classification errors at step one of two-step approach under various conditions, the model performed better under small sample size ($N = 200$) than under larger sample sizes ($N = 500$ or 1000).

Table 6

Classification error and Entropy for GMMs with different time coding using the one-step approach

			Classification Error		Entropy-based R^2	
N	Effect Size	Time	Balanced	Imbalanced	Balanced	Imbalanced
N=1000	Small (0.4)	initial	0.18	0.17	0.59	0.60
		middle	0.18	0.16	0.58	0.62
		final	0.19	0.17	0.61	0.60
	Large (0.8)	initial	0.25	0.24	0.47	0.47
		middle	0.25	0.24	0.47	0.46
		final	0.25	0.24	0.47	0.47
N=500	Small (0.4)	initial	0.18	0.16	0.60	0.61
		middle	0.17	0.15	0.61	0.63
		final	0.17	0.15	0.62	0.63
	Large (0.8)	initial	0.23	0.22	0.50	0.50
		middle	0.23	0.21	0.50	0.52
		final	0.23	0.21	0.50	0.51
N=200	Small (0.4)	initial	0.15	0.14	0.65	0.66
		middle	0.15	0.14	0.65	0.66
		final	0.15	0.14	0.65	0.65
	Large (0.8)	initial	0.18	0.17	0.58	0.59
		middle	0.18	0.16	0.59	0.61
		final	0.18	0.17	0.59	0.59

Note. $N =$ Sample size; S.E.= mean of Standard Error of the estimates; Effect Size = the effect size of measurement model; Balanced = Balanced class proportions; Imbalanced = Imbalanced class proportions; Time = Time coding schemes; Initial = Placing the intercept at initial time point; Middle = Placing the intercept at middle time point; Final = Placing the intercept at final time point; One-step = One step approach; Step one of Two-step = Step one of the two-step approach.

Table 7

Classification error and Entropy of step one in two-step approach for GMMs with different time coding

			Classification Error		Entropy-based R^2	
N	Effect Size	Time	Balanced	Imbalanced	Balanced	Imbalanced
N=1000	Small (0.4)	initial	0.24	0.23	0.46	0.48
		middle	0.25	0.22	0.48	0.50
		final	0.26	0.24	0.45	0.45
	Large (0.8)	initial	0.35	0.34	0.27	0.27
		middle	0.35	0.34	0.27	0.26
		final	0.35	0.34	0.26	0.26
N=500	Small (0.4)	initial	0.24	0.24	0.48	0.47
		middle	0.22	0.23	0.52	0.49
		final	0.24	0.22	0.47	0.46
	Large (0.8)	initial	0.31	0.30	0.32	0.32
		middle	0.31	0.29	0.32	0.32
		final	0.32	0.31	0.31	0.31
N=200	Small (0.4)	initial	0.24	0.20	0.47	0.54
		middle	0.22	0.18	0.51	0.59
		final	0.23	0.19	0.48	0.55
	Large (0.8)	initial	0.27	0.24	0.41	0.43
		middle	0.26	0.23	0.43	0.44
		final	0.26	0.25	0.42	0.41

Note. N = Sample size; S.E.= mean of Standard Error of the estimates; Effect Size = the effect size of measurement model; Balanced = Balanced class proportions; Imbalanced = Imbalanced class proportions; Time = Time coding schemes; Initial = Placing the intercept at initial time point; Middle = Placing the intercept at middle time point; Final = Placing the intercept at final time point; One-step = One step approach; Step one of Two-step = Step one of the two-step approach.

Estimated means of the distal outcome

As reported in the above section, in terms of the convergence rate, the entropy and latent class proportions recovery, the two-step estimator performed poorly, showed that measurement model was weak. Thus, the prediction of means of the distal outcome was not reliable. In this section, we focused on the results of the distal outcome recovery using the one-step approach only.

Table 8 and Table 9 show the simulation results of the distal outcome using a one-step approach across different sample sizes and effect sizes, for balanced class proportions condition and unbalanced class proportions condition respectively. We compare the performance of models using different time coding schemes in distal outcome prediction of a given latent class, according to the mean of estimates, the mean bias, the mean of the estimated SE and the SD of the estimated parameters. As can be seen in Table 8 and Table 9, compared to the true value of means of the distal outcome, the estimated means recovered well across all time coding approaches, reflected in the amount of bias, ranging from -0.50 to 0.37. Note that the recovery of the distal outcome was generally more accurate under more ideal conditions, specifically the condition of large sample sizes, effect size and balanced class proportions. Inspection the value of the ratios of the SD to the SE for balanced class proportions condition shows that, there was a large magnitude of bias in the SE estimates and the largest deviant can be observed in the second estimated latent class across all simulated conditions of sample size and effect size, which yielded values of the ratio that severely diverged from 1. And as we increased the sample size and effect size, the values were closer to 1, showing that model performed better for SE estimates under larger sample sizes and strong effect sizes conditions. The same tendencies can be seen in the imbalanced conditions as well.

Table 8

The distal outcome recovery for GMMs with different time coding approaches using the one-step approach on simulated data under balanced latent class proportion

				Initial			Middle			Final		
N	Effect Size	Class	True	*Est.*	*Bias*	SE/SD	*Est.*	*Bias*	SE/SD	*Est.*	*Bias*	SE/SD
Balanced latent class proportion (C1/C2/C3=0.33/0.33/0.33)												
1000	Small (0.4)	C1	0.00	0.10	0.10	0.49	0.07	0.07	0.47	0.03	0.03	0.51
		C2	3.00	2.90	-0.10	0.20	2.63	-0.37	0.19	2.60	-0.40	0.20
		C3	5.00	4.96	-0.04	0.47	4.88	-0.12	0.49	4.95	-0.50	0.45
	Large (0.8)	C1	0.00	-0.08	-0.08	0.73	-0.06	-0.06	0.68	-0.05	-0.05	0.63
		C2	3.00	2.85	-0.15	0.51	2.96	-0.04	0.52	2.94	-0.06	0.52
		C3	5.00	5.01	0.01	0.71	5.02	0.02	0.70	4.97	-0.03	0.79
500	Small (0.4)	C1	0.00	0.19	0.19	0.64	0.16	0.16	0.68	0.07	0.07	0.60
		C2	3.00	2.99	-0.01	0.29	2.85	-0.15	0.24	2.70	-0.30	0.29
		C3	5.00	4.98	-0.02	0.65	4.91	-0.09	0.64	4.89	-0.11	0.60
	Large (0.8)	C1	0.00	0.02	0.02	0.81	0.00	0.00	0.61	0.04	0.04	0.70
		C2	3.00	3.03	0.03	0.61	3.09	0.03	0.51	3.07	0.07	0.53
		C3	5.00	5.02	0.02	0.79	5.00	0.00	0.79	5.00	0.00	0.76
200	Small (0.4)	C1	0.00	0.00	0.00	0.59	0.02	0.02	0.59	-0.01	-0.01	0.62
		C2	3.00	2.89	-0.11	0.37	2.81	-0.19	0.41	2.66	-0.34	0.35
		C3	5.00	4.93	-0.07	0.68	4.92	-0.08	0.65	4.94	-0.06	0.65
	Large (0.8)	C1	0.00	-0.02	-0.02	0.63	-0.05	-0.05	0.64	-0.05	-0.05	0.63
		C2	3.00	2.95	-0.05	0.48	2.92	-0.08	0.47	3.03	0.03	0.50
		C3	5.00	5.09	0.09	0.59	5.07	0.07	0.63	5.12	0.12	0.63

Note. $N =$ Sample size; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; S.E.= Mean of Standard Error of the estimates; Effect Size = The effect size of measurement model; Est. = Mean of the model estimates; SD = The standard deviation of the estimated parameter; Bias = Mean estimated minus true value; True = The generating values of means of the distal outcome.

Table 9

The distal outcome recovery for GMMs with different time coding approaches using the one-step approach on simulated data under imbalanced latent class proportion

				Initial			Middle			Final		
N	Effect Size	Class	True	*Est.*	*Bias*	SE/SD	*Est.*	*Bias*	SE/SD	*Est.*	*Bias*	SE/SD
Imbalance latent class proportion (C1/C2/C3=0.20/0.30/0.50)												
1000	Small (0.4)	C1	0.00	0.11	0.11	0.49	0.20	0.20	0.47	0.13	0.13	0.56
		C2	3.00	3.15	0.15	0.24	3.10	0.10	0.20	2.92	-0.08	0.23
		C3	5.00	4.95	-0.05	0.61	4.90	-0.10	0.61	4.92	-0.08	0.52
	Large (0.8)	C1	0.00	0.01	0.01	0.67	0.00	0.00	0.74	0.03	0.03	0.79
		C2	3.00	3.09	0.09	0.52	3.12	0.12	0.52	3.15	0.15	0.53
		C3	5.00	5.03	0.03	0.72	5.04	0.04	0.69	5.07	0.07	0.62
500	Small (0.4)	C1	0.00	0.16	0.16	0.59	0.16	0.16	0.72	0.12	0.12	0.62
		C2	3.00	3.15	0.15	0.28	2.98	0.02	0.32	2.85	-0.15	0.24
		C3	5.00	4.93	-0.07	0.59	4.87	0.13	0.65	4.85	-0.15	0.53
	Large (0.8)	C1	0.00	0.03	0.03	0.69	0.06	0.06	0.56	0.09	0.09	0.67
		C2	3.00	3.13	0.13	0.52	3.17	0.17	0.45	3.25	0.25	0.46
		C3	5.00	5.06	0.06	0.83	5.18	0.18	0.45	5.20	0.20	0.42
200	Small (0.4)	C1	0.00	0.20	0.20	0.61	0.13	0.13	0.47	0.11	0.11	0.61
		C2	3.00	3.22	0.22	0.39	3.00	0.00	0.20	2.94	-0.06	0.39
		C3	5.00	5.11	0.11	0.61	5.07	0.07	0.61	5.02	0.02	0.71
	Large (0.8)	C1	0.00	0.17	0.17	0.62	0.19	0.19	0.74	0.26	0.26	0.67
		C2	3.00	3.17	0.17	0.41	3.26	0.26	0.52	3.37	0.37	0.46
		C3	5.00	5.21	0.21	0.57	5.22	0.22	0.69	5.26	0.26	0.57

Note. N = Sample size; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; S.E.= Mean of Standard Error of the estimates; Effect Size = The effect size of measurement model; Est. = Mean of the model estimates; SD = The standard deviation of the estimated parameter; Bias = Mean estimated minus true value; True = The generating values of means of the distal outcome.

Empirical Example

To demonstrate the effect of time coding approaches on the growth mixture model with the distal outcome, we used an empirical data from the Midlife in the United States (MIDUS; $N = 3924$; https://www.midus.wisc.edu/), which is a longitudinal research about health and well-being, and contains participant's experienced negative affect measured at three different time points (each 9 years apart), specifically, the response scale of negative affect is range from 1 to 5, indicating participants never had negative affect in this time to felt negative affect all the time, respectively. In addition, the data also include the number of experienced chronic health conditions that participants experienced in the past 12 months, response range from 1 to 20, indicating that participants experienced 1 to 20 times chronic health conditions. (Willroth et al., 2020). First, we excluded the distal outcome variable (i.e., chronic health conditions), and apply unconditional growth mixture models with 1 to 4 latent classes to three waves measured negative affect. For these growth mixture models, we placed the intercept at the initial time point. Based on the following fit statistics for each model: the Bayesian Information Criteria (BIC; Schwarz, 1978), the Entropy (Jedidi et al., 1993), the Log-likelihood (Lo, 2001) and Akaike's information criterion (AIC; Akaike, 1987), as shown in Table 6, we selected a four-class model as the best fitting model with the smallest value of BIC and AIC, and its entropy value was 0.81, which was higher than 0.80, indicating good classification accuracy (Greenbaum et al., 2005; Muthén, 2004).

Table 10

Fit index, entropy, and model comparisons for estimated GMMs on real date example

GMMs	Log-likelihood	Parameters	AIC	BIC	Entropy
One-class	-5602.19	8	11220.38	11267.31	1.00
Two-class	-4360.83	13	8747.66	8823.91	0.70
Three-class	-2747.54	20	5535.08	5652.40	0.83
Four-class	-2444.55	25	4939.10	5085.75	0.81

Note. AIC = Akaike's information criterion; BIC = Bayesian information criterion; GMMs = Growth mixture models

After the optimal number of latent classes was identified, we incorporated the distal outcome variable (i.e., the number of experienced chronic health conditions at the last time point) into the model and inspected the predictive effect of the latent class variable on the distal outcome. As in the simulation study, we changed the time coding approaches by removing the intercept at initial, middle, and final time points and used two parameter estimation approaches. While one-step approach converged successfully with a good value of entropy-based R^2 (entropy-based R^2 = 0.81), two-step approach failed to converge at step 2. Here, we only presented the results estimated using one-step approach. Table 11 presents the estimated profile of latent classes and predicted means of the chronic health conditions when the intercept is placed at initial time point. Specifically, one class is characterized as having relatively higher experienced negative effect at the initial time and the experienced effect remained stable over the last two times. This class is 24% of the sample, and participants in this class tended to experience more chronic health conditions. The second class is 38% of the sample, which is the largest class, and is characterized by having lower initial experienced effect and remained stable in the second and third times. Participants in this class had lower prevalence of chronic conditions. The third class is 29% of the sample, participants with the lowest experienced effect at the initial time and continued to decline at the second and third experiences, in which they tended to experience the fewest chronic health conditions. There is only 9% of the sample in the last class, participants in this class reported to have the largest negative effect at initial time and continued to increase, expected to experience chronic health conditions far beyond the other classes.

Table 11

Classes proportion and class-specific means of participant's subjective negative effect at three time points and predicted means the chronic health conditions

	C1	C2	C3	C4
Class proportion	0.24	0.38	0.29	0.09
Wave 1	1.96	1.34	1.15	1.96
Wave 2	1.95	1.35	1.08	2.07
Wave 3	1.93	1.35	1.00	2.18
Chronic	3.14	2.55	2.04	10.17

Note, N =2559; Chronic = the means of chronic health conditions; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; C4 = Latent Class 4.

As we also compared the model performance across different time coding approaches, table 11 presented the estimated parameters, classification error and entropy-based R^2 when we alter the intercept into three different time points. As can be seen, the model performed well with an Entropy-based R^2 of 0.81 and a classification error of 0.08, the estimated number of chronic health condition and the latent class proportions remains constant when we place the initial status at different time points.

Table 12

Parameter recovery for GMMs with different time coding on real data example

	Latent class proportion				Chronic				Classification error	Entropy-based R^2
	C1	C2	C3	C4	C1	C2	C3	C4		
Initial	0.09	0.24	0.29	0.38	2.04	2.55	3.14	10.17	0.08	0.81
Middle	0.09	0.23	0.29	0.38	2.05	2.55	3.15	10.19	0.08	0.81
Final	0.09	0.23	0.29	0.38	2.05	2.55	3.15	10.19	0.08	0.81

Note. N =2559; Chronic = the number of chronic health conditions; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; C4 = Latent Class 4.

Discussion

In order to properly estimate and interpret the parameters of growth trajectories and the effect of growth parameters on distal outcomes, the impact of time coding choices needs to be better understood. Prior research (McCormick et al., 2023) explored the impact of time coding decisions in linear latent curve model with distal outcome variables, and showed that the predictive effect of linear slope on the distal outcome will change from positive to negative, and from significant to non-significant, as we move the intercept to different time points. In this study, we inspected the effect of time coding approaches on distal outcome regressions in a growth mixture model (GMM) framework, where the categorical latent class variable rather than the continuous latent growth factors (i.e., intercept and slope) is directly related to the distal outcome. The simulation results revealed that the estimated means of the distal outcome remain invariant in the GMM when we model the intercept at different time points. The simulation results are further supported by findings from the empirical demonstration using three waves of data from the midlife in the United States (MIDUS).

The role of time coding

Results suggest that when the categorical latent variable is used to predict the distal outcome directly, the recovery of the latent class proportions and means of the distal outcome does not vary with the change of the time coding approaches. In contrast to the previous research that inspected the predictive effect of continuous latent variables on the distal outcomes by using different time coding approaches (McCormick et al., 2023), we found that there is no evidence for an influence of time coding on the relationship between the categorical latent variable and the distal outcomes.

Differences in the variable type that used to predict the distal outcome, may be the reason that we did not observe variations of the distal outcome from the categorical latent variable prediction. In GMMs, the distal outcome directly relates to and predict by the categorical latent variable. As the latent class variable is discrete and may convey less information than continuous latent growth factors. In this case, the latent class variable may not be sensitive to reflect changes caused by altering the time coding approaches, thus the prediction of the distal outcome remains unchanged under different time coding approaches.

Another reason can be the framework of statistical models is different. In GMMs, the combination of categorical and continuous latent variables is used to capture the qualitatively different growth trajectories. While for the LCMs, we can only capture the quantitively difference by continuous latent variables (Muthén, 2001). When we place the intercept at different time points, we might have expected this to have induced fluctuations in the means, variances, and covariances of the latent growth factors, but it appears that the fluctuation may not make the qualitative difference in the growth trajectories of subpopulations. In other words, the average growth trajectories in each subpopulation are not influenced by this re-specification. The categorical latent class variables indicated by the latent growth factors, do not capture this difference as it may still be quantitative. Therefore, the same subpopulations are identified, and the distal outcome variables predicted by latent class variables may not change significantly either. We could examine the distributional parameters of subpopulations across different time coding approaches to further understand and verify this process in the future.

Impact of the utility of measurement model

In the current study, we manipulated three designed features to inspect model performance under different conditions, including measurement quality, sample size and latent class proportion. Consequently, the utility of measurement model has an important effect on model performance. These effects are combined with variations in sample size, latent class proportions, class separation and measurement quality. When considering the recovery of the class-specific means for the distal outcome, simulation results show that parameter recovery and SE estimates are better with larger sample sizes, balanced class proportions and strong measurement quality. Prior work (Tueller & Lubke, 2010) indicates that increasing sample size is efficient to increase class separation, in that have a significant effect on model performance (e.g., less bias), consistent with our results. Furthermore, improving measurement quality – achieved by increasing the reliability of the growth indicators – results in decreased bias in the estimated parameters of interest. Although the parameter estimates look appropriate, it is important to note that the magnitude of bias in SE estimates is large and underestimated (i.e., ratios of the SE and the SD are small than 1) across all conditions. In this case, the recovered parameter distribution was strongly biased towards underestimating the variability compared with the generated parameters. This tendency leads to significance tests that are too optimistic, resulting in the increased probability of Type I error.

Impact of estimation approaches

To examine model performance in different estimators, we used one-step approach (Bandeen-Roche et al., 1997) and two-step approach (Bakk & Kuha, 2018) to obtain model parameters in this study. However, we only discussed the recovered means of the distal outcome using one-step approach, since the convergence rates for two-step approaches are extremely low across all twelve conditions, there is less confidence than one-step approach in the recovery of means of the distal outcome. As discussed in previous research (Vermunt, 2010), this can be attributed to the poor quality of the measurement model, or the strength of association between the latent class variable and the observed indicators is not strong enough. It is also evidenced by the entropy-based R^2 at step one being lower than 0.60, shows that the separation between classes is low, the estimated latent classes are similar.

The study also examined the prevalence of latent classes. In our simulation set up, the models using one-step approach estimated class proportions that well matched the generating model across all conditions. However, for the two-step approach, the models did not perform well. Compared to the generating model, the models tended to exaggerate the class imbalance, where the smaller class is underestimated, and the larger class is overestimated regardless of conditions. This appears to be driven by the highly reliance of models on the information that the distal outcomes provide in estimating the latent class proportions. Specifically, in one-step approach, the distal outcome is considered an additional indicator of the latent class variable, making a joint effort with the repeated measures variable to estimate the measurement model. Based on the more extensive information provided by the distal outcome, estimation was more accurate, particularly in the conditions of low separation between classes (i.e., smaller sample sizes), the distal outcome may influence model performance more so than the repeated measures themselves (Bakk & Kuha, 2018). However, for the two-step approach, the distal outcome is excluded from the measurement model. In this case, when the measurement model is weak and the class separation is low, two-step estimators may be hard to accurately estimate the measurement model. As reflected in the parameter estimates for latent class proportions, the one-step method outperformed the two-step approach, supported by the additional information provided from the distal outcome variable.

Counterintuitive pattern

By contrast, we found that perhaps counterintuitively, when the sample size is small (*N* = 200), the models perform better than under typically more-ideal conditions (i.e., large sample size) in terms of the classification error and entropy-based R^2. As previous studies observed (Galindo Garre & Vermunt, 2006; Vermunt, 2010), this is because the maximum likelihood estimator overestimates the variability among latent classes when estimating the model (Vermunt, 2010), which leads to the classification error being underestimated. As demonstrated by the results in Table 6 and Table 7, the average classification error (ACE) under conditions of small sample size ($N = 200$, $ACE_{balanced} = 0.16$, $ACE_{imbalanced} = 0.15$) is lower than larger sample size ($N = 500$, $ACE_{balanced} = 0.20$ and $ACE_{imbalanced} = 0.18$; $N = 1000$, $ACE_{balanced} = 0.22$ and $ACE_{imbalanced} = 0.20$) conditions. That is, the models are too optimistic in estimating the difference among latent classes under smaller sample sizes. In particular, when the datasets are generated under the conditions of small sample size and the separation between classes is low, this pattern is more obvious. In this study, we attempted to increase the separation between latent classes by designing imbalanced latent class proportions. However, for using one step approach, the classification error is even lower on small sample size conditions when the latent class proportions are imbalanced (Table 7; *N* =

200, ACE $_{imbalanced}$ = 0.21, ACE $_{balanced}$ = 0.25), showing that the sample size is more influential on the model performance under these conditions.

Empirical data example

To demonstrate our findings, we fitted our models on empirical data from the Midlife in the United States (MIDUS; https://www.midus.wisc.edu/). Through the class enumeration process, we selected the best model with four latent classes for predicting chronic health conditions. After that, we changed the time coding approaches, and examine the effect on the distal outcome prediction by using one-step and two-step estimation approaches. For one-step approach, we found that the prediction of chronic health conditions, the estimated latent class proportions and the estimated latent class membership assignments remain constant across different time coding schemes, consistent with the results in our simulation study. For two-step approach, despite the entropy-based R^2 of the measurement model being 0.81, the model fail to converge at step 2. By inspecting the growth pattern of latent classes in table 11, we can see that the initial status and growth rates of latent trajectories are similar, resulted in a weak measurement model that is insufficient to accurately predict distal outcomes.

Limitations

In this study, we conducted a simulation study to inspect the effect of time coding on the distal outcome prediction in GMMs. While our findings illuminate some of the contrasts between prior work on continuous growth models, are still some limitations of our approach in growth mixture models that should be explored in future work. One such limitation is that we only consider prediction of the distal outcome with the categorical latent class variable, but GMMs sometimes allow the continuous growth latent variables to directly predict the distal outcome (Nylund-Gibson et al., 2019). As prior work (McCormick et al., 2023) found that the effect of latent slope variable on the distal outcome varies across different time coding approaches in the LCMs, this approach merits further consideration. If we inspect the effect of time coding approaches on the predictive association between the distal outcome and the continuous growth latent variables, rather than the latent class variable, the prediction of the distal outcome may be similarly influenced.

In addition, the standard errors of the class-specific means for the distal outcome were underestimated across all conditions for one-step approach, suggesting that an inflated Type I error is likely in these models. It should also be noted that, in this study, we only manipulated the quality of the GMM measurement model, but if the association between the latent class variable and the distal outcome is ignored during class membership assignment stage, the predictive effect of the latent class variable on the distal outcome will be diminished (Lanza et al., 2013). Thus, the magnitude of the effect of the latent class variable and the distal outcome should be considered in future research.

Conclusions

To conclude, incorporating the distal outcome into growth models can help researchers to understand the consequence of differences in growth trajectories. The choices of time coding have a significant influence on the estimation and interpretation of the distal outcome in the linear growth curve model. In this study, we extend these findings to the GMM framework by examining the influence of time coding choices on the latent class recovery and prediction of the distal outcomes using one-step and two-step approaches. Here, in contrast, we found that when the categorical latent class variable is related to the distal outcome, the effect of time coding choices on parameters recovery and distal outcome prediction is absent. These findings contribute to understanding the impact of time coding choices on the distal outcome prediction in growth models and suggest that the growth mixture model is robust to these the choice of time coding.

Reference

Akaike, H. (1987). Factor Analysis and AIC. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected Papers of Hirotugu Akaike* (pp. 371–386). Springer New York. https://doi.org/10.1007/978-1-4612-1694-0_29

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

Bandeen-Roche, K., Miglioretti, D. L., Zeger, S. L., & Rathouz, P. J. (1997). Latent Variable Regression for Multiple Discrete Outcomes. *Journal of the American Statistical Association*, *92*(440), 1375–1386. https://doi.org/10.1080/01621459.1997.10473658

Bauer, D. J., & Curran, P. J. (2003). Distributional Assumptions of Growth Mixture Models: Implications for Overextraction of Latent Trajectory Classes. *Psychological Methods*, *8*, 338–363. https://doi.org/10/b45p2v

Biesanz, J. C., Deeb-Sossa, N., Papadakis, A. A., Bollen, K. A., & Curran, P. J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods*, *9*(1), 30–52. https://doi.org/10/b64bpd

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, *101*(1), 147. https://doi.org/10.1037/0033-2909.101.1.147

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve Frequently Asked Questions About Growth Curve Modeling. *Journal of Cognition and Development*, *11*(2), 121–136. https://doi.org/10.1080/15248371003699969

Diallo, T. M. O., Morin, A. J. S., & Lu, H. (2017). Performance of growth mixture models in the presence of time-varying covariates. *Behavior Research Methods*, *49*(5), 1951–1965. https://doi.org/10/gb4c8g

Galindo Garre, F., & Vermunt, J. K. (2006). Avoiding Boundary Estimates in Latent Class Analysis by Bayesian Posterior Mode Estimation. *Behaviormetrika*, *33*(1), 43–59. https://doi.org/10.2333/bhmk.33.43

Gong, G., & Samaniego, F. J. (1981). Pseudo Maximum Likelihood Estimation: Theory and Applications. *The Annals of Statistics*, *9*(4). https://doi.org/10.1214/aos/1176345526

Greenbaum, P. E., Del Boca, F. K., Darkes, J., Wang, C.-P., & Goldman, M. S. (2005). Variation in the drinking trajectories of freshmen college students. *Journal of Consulting and Clinical Psychology*, *73*(2), 229. https://doi.org/10.1037/0022-006X.73.2.229

Harring, J. R., & Hodis, F. A. (2016). Mixture Modeling: Applications in Educational Psychology. *Educational Psychologist*, *51*(3–4), 354–367. https://doi.org/10.1080/00461520.2016.1207176

Jedidi, K., Ramaswamy, V., & Desarbo, W. S. (1993). A maximum likelihood method for latent class regression involving a censored dependent variable. *Psychometrika*, *58*(3), 375–394. https://doi.org/10.1007/BF02294647

Jung, T., & Wickrama, K. A. S. (2008). An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling: Latent Trajectory Classes. *Social and Personality Psychology Compass*, *2*(1), 302–317. https://doi.org/10/b4zd7x

Lanza, S. T., Tan, X., & Bray, B. C. (2013). Latent Class Analysis With Distal Outcomes: A Flexible Model-Based Approach. *Structural Equation Modeling: A Multidisciplinary*

Journal, *20*(1), 1–26. https://doi.org/10.1080/10705511.2013.742377

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

Lo, Y. (2001). Testing the number of components in a normal mixture. *Biometrika*, *88*(3), 767–778. https://doi.org/10.1093/biomet/88.3.767

McArdle, J. J. (1988). Dynamic but Structural Equation Modeling of Repeated Measures Data. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of Multivariate Experimental Psychology* (pp. 561–614). Springer US. https://doi.org/10.1007/978-1-4613-0893-5_17

McArdle, J. J., & Epstein, D. (1987). Latent Growth Curves within Developmental Structural Equation Models. *Child Development*, *58*(1), 110–133. https://doi.org/10.2307/1130295

McCormick, E. M., Curran, P. J., & Hancock, G. R. (2023). *Latent Growth Factors as Predictors of Distal Outcomes: Completing the Triad*. https://doi.org/10.31234/osf.io/fevra

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

Morin, C. M., Belleville, G., Bélanger, L., & Ivers, H. (2011). The Insomnia Severity Index: Psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep*, *34*(5), 601–608. https://doi.org/10.1093/sleep/34.5.601

Muthén, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class–latent growth modeling. In *New methods for the analysis of change* (pp. 291–322). American Psychological Association. https://doi.org/10.1037/10409-010

Muthén, B. (2004). Latent Variable Analysis: Growth Mixture Modeling and Related Techniques for Longitudinal Data. In D. Kaplan, *The SAGE Handbook of Quantitative Methodology for the Social Sciences* (pp. 346–369). SAGE Publications, Inc. https://doi.org/10.4135/9781412986311.n19

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

Nylund-Gibson, K., Grimm, R. P., & Masyn, K. E. (2019). Prediction from Latent Classes: A Demonstration of Different Approaches to Include Distal Outcomes in Mixture Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(6), 967–985. https://doi.org/10.1080/10705511.2019.1590146

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2). https://doi.org/10.1214/aos/1176344136

Tueller, S., & Lubke, G. (2010). Evaluation of structural equation mixture models: Parameter estimates and correct class assignment. *Structural Equation Modeling*, *17*(2), 165–192. https://doi.org/10.1080/10705511003659318

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. (2021). LG-syntax user's guide: Manual for latent gold syntax module version 6.0. *Arlington, MA: Statistical Innovations Inc*.

Willroth, E. C., Ong, A. D., Graham, E. K., & Mroczek, D. K. (2020). Being happy and becoming happier as independent predictors of physical health and mortality. *Psychosomatic Medicine*, *82*(7), 650. https://doi.org/10.1097/PSY.0000000000000832