Psychological Methods
2013, Vol. 18, No. 2, 186–219

© 2013 American Psychological Association
1082-989X/13/$12.00 DOI: 10.1037/a0031609

# Mixture Class Recovery in GMM Under Varying Degrees of Class Separation: Frequentist Versus Bayesian Estimation

Sarah Depaoli
University of California, Merced

Growth mixture modeling (GMM) represents a technique that is designed to capture change over time for unobserved subgroups (or latent classes) that exhibit qualitatively different patterns of growth. The aim of the current article was to explore the impact of latent class separation (i.e., how similar growth trajectories are across latent classes) on GMM performance. Several estimation conditions were compared: maximum likelihood via the expectation maximization (EM) algorithm and the Bayesian framework implementing diffuse priors, "accurate" informative priors, weakly informative priors, data-driven informative priors, priors reflecting partial-knowledge of parameters, and "inaccurate" (but informative) priors. The main goal was to provide insight about the optimal estimation condition under different degrees of latent class separation for GMM. Results indicated that optimal parameter recovery was obtained though the Bayesian approach using "accurate" informative priors, and partial-knowledge priors showed promise for the recovery of the growth trajectory parameters. Maximum likelihood and the remaining Bayesian estimation conditions yielded poor parameter recovery for the latent class proportions and the growth trajectories.

*Keywords:* growth mixture modeling, latent class separation, Markov chain Monte Carlo, Bayesian estimation, priors

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Many processes in the social and behavioral sciences are dynamic in that they change over time. As a result, we have seen an increase in methodological tools available to model growth and development for repeated measures data. Growth mixture modeling (GMM) represents a technique that is designed to capture change over time for unobserved subgroups (or populations) that exhibit qualitatively different patterns of growth. Specifically, GMM is a tool used to model different growth trajectories (or patterns of growth) by grouping individuals into a finite number of latent classes based on observed data patterns. Each of these latent classes is then characterized by a distinct growth trajectory, representing differences in change or development over time across the latent groups of individuals.

GMM has been implemented at an increasing rate in substantive applications since it was first introduced into the applied literature (see e.g., B. Muthén & Shedden, 1999; Nagin, 1999). For example, Gueorguieva, Mallinckrodt, and Krystal (2011) examined changes in depression patterns for patients with major depression and reported a latent class of "responders" and another of "nonresponders" to depression treatment. In a similar application, Morin et al. (2011) used GMM to assess fluctuations in elevated levels of anxiety for adolescents over time and reported five latent classes representing different patterns of anxiety. Finally, Schaeffer et al. (2006) used GMM to model patterns of aggression and disruptive behavior in children throughout elementary school years; they found three substantively different latent classes for boys and girls.

Despite the growing popularity and increased usage (see e.g., Bauer, 2007), there are still certain limitations and concerns surrounding GMM that should be mentioned. Specifically, Bauer and Curran (2003) found that multiple latent classes could appear optimal when the population was only comprised of one group and when the data were nonnormal. This finding is troubling in that applied researchers may be inclined to retain a mixture solution with multiple latent classes when in fact the underlying data structure contains a single group or population. Alternative (absolute) fit indices have been proposed to distinguish between the case of multiple latent classes and a single-class scenario with nonnormal data (see e.g., B. O. Muthén, 2003; Nylund, Asparouhov, & Muthén, 2007; Rindskopf, 2003). However, these methods for assessing fit may still lead to inappropriate model solutions that do not accurately represent the population model.

Relevant to the current article, one additional issue that arises when estimating mixture models deals with the separation of the latent classes. Mixture class separation refers to the distance (or differences) between two latent classes. In the case of GMM, the separation between latent classes is often characterized by how similar/different the growth trajectories are between the latent classes. If separation between the growth trajectories is poor (i.e., the growth trajectories are similar), then individuals may be inaccurately assigned into latent classes. Likewise, class assignment is particularly impacted with relatively smaller sample sizes (e.g., $n = 100$) under poor class separation (see e.g., Tueller & Lubke, 2010). Therefore, it may be more difficult recovering relatively smaller latent classes, especially under conditions of poor class separation.

The aim of the current article is to further explore the impact of class separation on model results. In particular, the current article

This article was published Online First March 25, 2013.

I am grateful to David Kaplan for valuable advice and feedback on earlier versions of this article.

Correspondence concerning this article should be addressed to Sarah Depaoli, School of Social Sciences, Humanities, and Arts, University of California, Merced, 5200 North Lake Road, Merced, CA 95343. E-mail: sdepaoli@ucmerced.edu

assesses the potential for the Bayesian estimation framework via the Markov chain Monte Carlo (MCMC) algorithm to improve the performance of GMM under cases of poor class separation. It is also of interest to examine the role that Bayesian methods play in recovering relatively small latent classes, especially when class separation is poor. The Bayesian estimation framework is compared to a more conventional estimation procedure for mixture models using maximum likelihood estimation via the expectation maximization (EM) algorithm (ML/EM; see e.g., Enders & Tofighi, 2008; Jedidi, Jagpal, & DeSarbo, 1997; Kaplan, 2002; McLachlan & Peel, 2000; Tueller & Lubke, 2010).

This study is organized as follows. The next section defines class separation in more detail and presents a summary of the recent literature focused on mixture class separation in the context of finite mixture models. Issues surrounding the estimation of mixture models through ML/EM are briefly discussed. This section is followed by a general discussion of Bayesian estimation, where key Bayesian terminology is defined. The specification of GMM is presented next. This section is followed by the specification of the prior distributions used for GMM in the Bayesian framework, as well as the MCMC convergence diagnostic implemented in this study.

Next, the simulation study design is presented which is followed by a presentation of results. This article concludes with a summary of the findings, a discussion of the benefits/risks of using the ML/EM and MCMC estimation conditions under different levels of class separation, and implications for applied mixture model research.

## Defining Class Separation

In GMM, class separation typically refers to the amount of overlap between growth trajectories for the latent classes. Suppose that individual growth trajectories were plotted for two latent classes. Class separation would refer to how much (or little) overlap there was between the trajectories for each of the classes. In particular, separation refers to how much overlap there is between the latent class distributions.

This issue of class separation in GMM can perhaps be best illustrated through a contrived example highlighting the two extreme versions of separation. To give this example a context, suppose that reading level was assessed across four equally spaced time points in first grade classrooms. Each student would then have a trajectory that illustrated the growth (or change) rate in reading level across the time points. Likewise, each latent class could also be summarized by an estimated growth trajectory representing the students comprising that latent class.

The most extreme example of poor class separation would be if two latent classes perfectly overlapped. In this case, a plot of individual growth trajectories for students sampled from both of the latent classes might look like the plot presented on the top of Figure 1. In this plot, growth trajectories from the two classes are perfectly overlaid, and there is no clear visual separation between the classes. In fact, if the estimated growth trajectories for these two classes were plotted, they would be exactly the same. The plot on the bottom of Figure 1 illustrates growth trajectories for the two latent classes that have zero separation. For illustrative purposes, the growth trajectories are slightly separated in the plot to visualize both trajectories; however, the trajectories actually overlap perfectly in this scenario. In this case, correct class assignment for all individuals would be quite

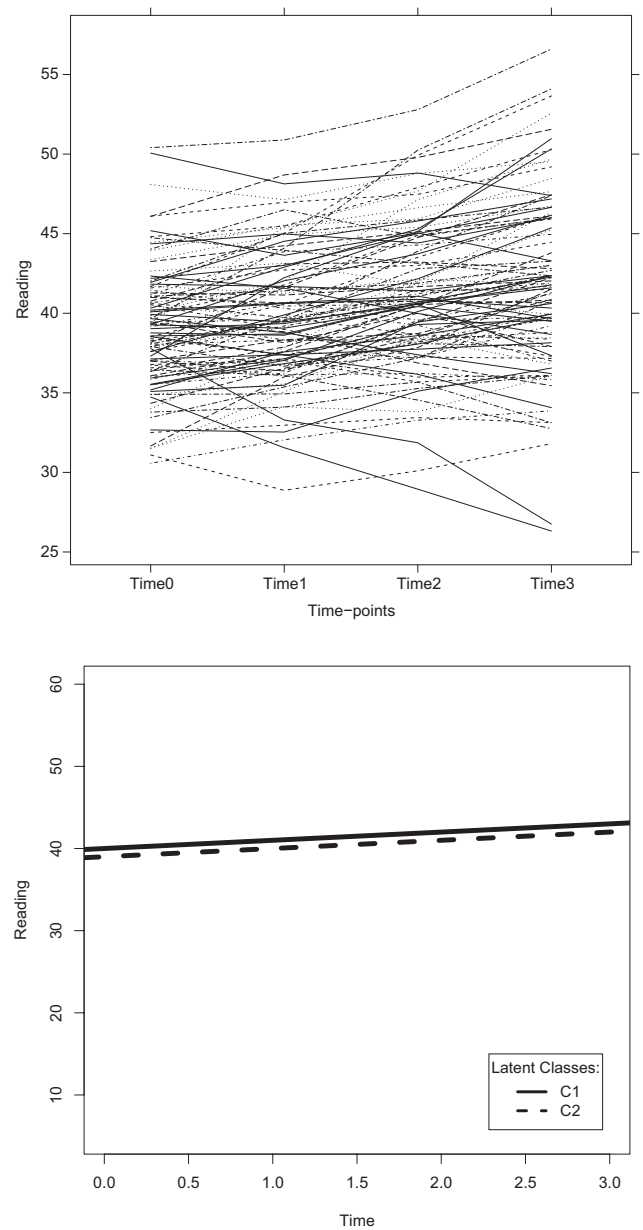


*Figure 1.*. Contrived example illustrating no class separation. C1 = Latent Class 1; C2 = Latent Class 2.

improbable to obtain since the classes perfectly overlap and are therefore indistinguishable from one another.[1]

On the other extreme, one could also imagine a scenario where latent classes have a large amount of separation, meaning that there is zero overlap between the distributions of the latent classes. Using our same reading level example, this scenario would consist of individual growth trajectories that remained separate across the

[1] The purpose of this example was solely to illustrate the concept of poor class separation. The model discussed here would actually result in the two latent classes collapsing into a single class because they are identical. This result would therefore yield a model that is not identified.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

two classes. For example, the top of Figure 2 depicts individual growth trajectories from two extremely separated latent classes. The separation between individuals is very clear to visualize here. Notice how there is very little overlap among individual trajectories from each of the two classes.

The bottom of Figure 2 illustrates the estimated growth trajectories for the two latent classes. It is clear that these trajectories represent latent classes with very different beginning reading levels and very different growth rates over time. In this case of extreme class separation, class assignment for individuals is likely to be accurate due to the high separation between the two latent classes.

Although these two extremes are helpful in illustrating the concept of class separation in the context of GMM, the more realistic (and potentially more interesting) middle ground is the focus of the current study. Within this middle ground of class separation, individual growth trajectories corresponding with each of the two latent classes would represent separation somewhere in between that presented in Figures 1 and 2. In particular, the individual growth trajectories across the two latent classes would have some overlap, but there would still be distinguishing features of growth across the classes.

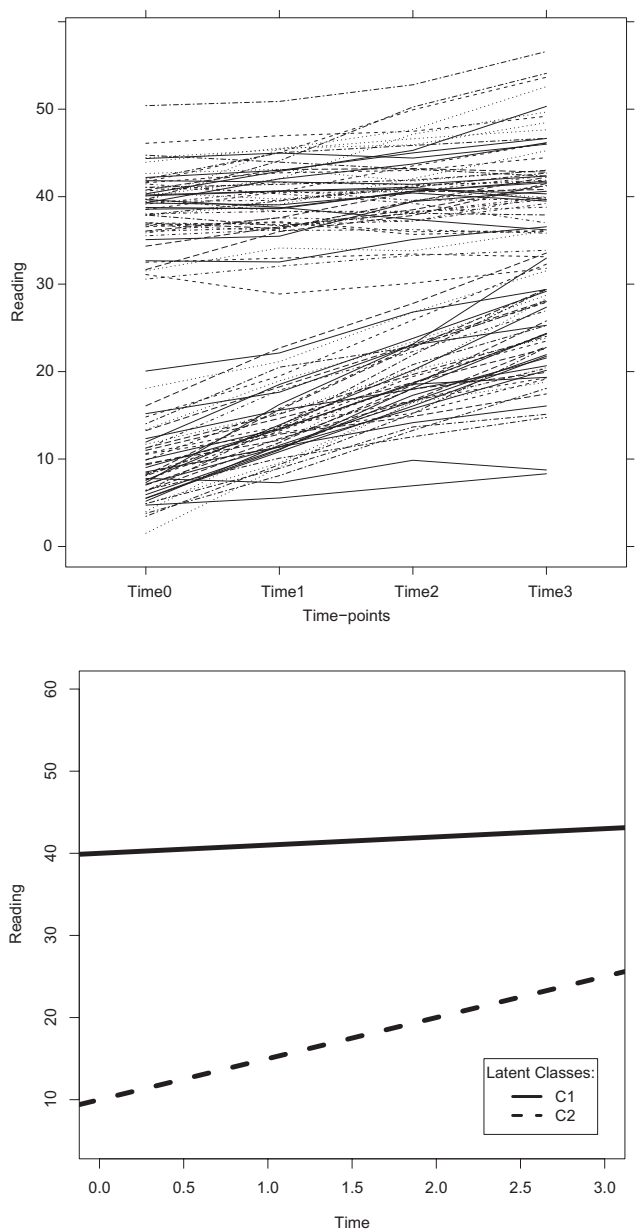


*Figure 2..* Contrived example illustrating high class separation. C1 = Latent Class 1; C2 = Latent Class 2.

As an example, the top plots in Figure 3 illustrate a scenario where the two latent classes represent moderate separation in reading levels. The two classes can be visually distinguished relatively well, but the trajectories certainly overlap across classes, which makes the classes more difficult to differentiate than the scenario presented in Figure 2. The bottom plot in Figure 3 depicts the estimated growth trajectories for these latent classes; note that the two classes differ substantively in the initial reading level and growth rates.

Another issue of concern here that closely relates to class separation is the size of the latent classes in the sample. For instance, the top plots in Figure 3 show the same level of class separation, but the plot on the right illustrates a relatively smaller sample size comprising the second latent class (i.e., there are fewer trajectories grouped together in the bottom half of the right plot). As class separation worsens (i.e., as class trajectories become more similar), higher sample sizes are required to properly recover the latent classes and obtain accurate class assignment (Tueller & Lubke, 2010). In other words, it may be more difficult to identify the second class in the right plot compared to the left plot due to the relatively smaller sample size within that second latent class. Likewise, poor class separation can also create problems in convergence (Tofighi & Enders, 2008; Tueller & Lubke, 2010).

A topic also tied indirectly to latent class separation is the issue of class enumeration (i.e., the number of latent classes retained by the investigator). Class enumeration may be defined in terms of latent classes that are substantively different in some manner (B. O. Muthén & Muthén, 2000). Substantive differences may be determined in the context of latent class separation. For example, higher separated classes may be retained as being substantively different, whereas the investigator may allow poorly separated classes to collapse because the trajectories are similar. However, poorly separated classes may still contain substantive differences that warrant the classes to be retained by the investigator as separate classes. In this case, the substantive differences between classes may be determined on how the classes differ substantively on some covariate, regardless of the separation of the latent class trajectories (B. O. Muthén, 2004). Here, the investigator may choose to retain poorly separate classes since there is still a substantive difference between them. The issue of class enumeration is not directly addressed in the current study; for more information, see for example, Henson, Reise, and Kim (2007), B. O. Muthén (2004), and Nylund et al. (2007). However, the current investigation presents a sensitivity-type analysis of class separation in the context of different estimation conditions to examine the optimal conditions for estimating GMM under different research scenarios.

In order to properly recover the latent classes under conditions of declining class separation and relatively lower sample sizes, more data are required to obtain accurate results. However, it may not always be a viable option to collect additional data—especially with the resources required to collect longitudinal data. Another option to collecting more data would be to embed additional

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

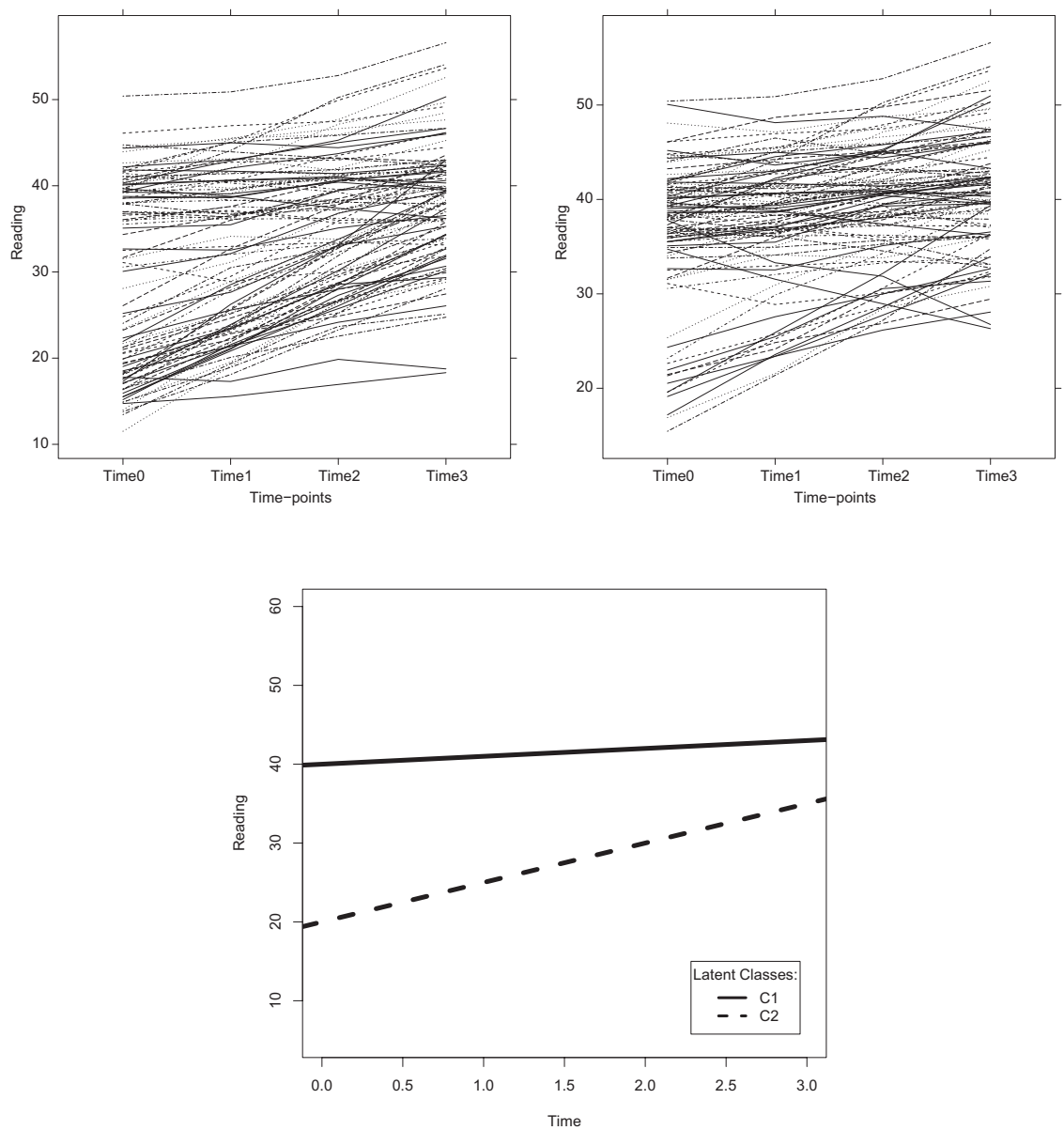


*Figure 3..* Contrived example illustrating moderate class separation. C1 = Latent Class 1; C2 = Latent Class 2.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

knowledge or information about the latent classes and parameters into the estimation process. The ability to include additional information in this manner is one of the main features of the Bayesian estimation framework, which is discussed in detail below.

## Class Separation for Finite Mixture Models

It is important to have an understanding for the impact that class separation can have on estimation within different modeling situations. Very few studies have examined class separation directly within the context of structural equation modeling, but two such studies will be briefly highlighted here. Tueller and Lubke (2010) conducted a study that assessed estimates and class assignment recovery using a structural equation mixture model. Factor means were varied in order to create the four desired Mahalanobis distance (MD) values representing varying degrees of class separation with MDs = 0.5, 1.0, 1.5, and 2.0.[2] Findings indicated that there was larger bias in the estimates when relatively smaller sample sizes (e.g., $n$ = 100) were used with a lower degree of class separation (e.g., MD = 0.5); a study by Tolvanen (2008) found comparable results. Correct class assignment improved as class separation increased, indicating that class assignment becomes

[2] The Mahalanobis distance is one type of measure that is used to quantify/define separation between latent classes. The details of the calculation implemented here are presented in the Design section.

easier to identify when separation is high (e.g., MD = 2.0). However, the overall result was that class recovery was poor for this study due to the complexity of the model being assessed. More complex models appeared to create additional difficulties in proper class recovery.

A study by Tofighi and Enders (2008) used an alternate method of determining class separation in the context of GMM by manipulating the variance components to create three different levels of class separation. To decrease separation between classes, the magnitude of the within-class variance parameters was increased—increasing the variance parameters would thus increase the amount of overlap between the latent classes. Growth factor means were held constant throughout this process, but the spread of the within-class trajectories increased as separation declined. Overall, this study found that various likelihood ratio-based fit tests and information criteria had a difficult time identifying the class structure properly within the poor class separation levels of the study. The Tueller and Lubke (2010) and Tofighi and Enders (2008) studies both indicated that the accuracy of class recovery decreased when only a small proportion of cases were in one of the classes (e.g., only 7% of the cases), regardless of total sample size.

Akin to Tueller and Lubke (2010) and Tolvanen (2008), the current investigation specifically focuses on class separation as defined through the latent growth factors (i.e., the growth trajectory intercept and slope). However, separation can also be characterized in terms of the repeated measures manifest (observed) variables.[3] Given that there is no consensus for the "best" method of defining separation, the researcher should determine what definition of separation best suits the particular research scenario in question.

The previous work presented on class separation was all completed in the context of the ML/EM estimation framework. In general, higher sample sizes are needed to ensure adequate estimates are produced when working with this asymptotic theory-based estimator. Similarly, higher class separation crossed with higher sample sizes is optimal when using the ML/EM estimator. However, these conditions may not be particularly realistic in applied research settings. Large sample sizes are not always practical to obtain. Likewise, the magnitude of class separation is not something that can easily be controlled or predetermined in an applied setting. It is also the case that ML/EM can be sensitive to starting values, particularly within the context of mixture models (Bollen & Curran, 2006; McLachlan & Peel, 2000; B. O. Muthén, 2004). Specifically, Hipp and Bauer (2006) presented findings which suggest that mixture models are particularly susceptible to converging on local solutions, especially as the complexity of the model increases (e.g., with more latent classes or the inclusion of random effects). These issues have led researchers to consider alternative estimation procedures in search of an estimator that can be reliable under conditions more typical of real data (i.e., smaller samples and less defined class separation). One such estimator is the Bayesian estimation framework via the MCMC estimation algorithm, which is used in the current study.

## The Bayesian Estimation Framework and the Impact of Priors

The Bayesian estimation framework has been used at an increasing rate to estimate structural equation models (see e.g., Lee, 2007; B. O. Muthén & Asparouhov, 2012; Yang & Dunson, 2010; Zhang, Hamagami, Wang, Nesselroade, & Grimm, 2002). The major distinction between the Bayesian and more conventional, frequentist (e.g., ML/EM) frameworks is in the elicitation and specification of prior distributions (or priors). Priors represent an expectation or a prior belief about what a parameter value looks like. Specifically, priors are characterized by probability distributions that are specified for each of the parameters in the model. These distributions are then incorporated into the estimation process, and they act as a means to include additional information about the parameter values in the model.

To give an example of what a prior might look like in the context of GMM, let us revisit the contrived example plotted in Figure 3. Before even collecting any data, the investigator would have some sort of prior beliefs about parameter values for the model being estimated. For example, if the investigator wanted to estimate a model with two latent classes, he or she might also have an idea about what the trajectories look like for each latent class. In this scenario, the investigator would specify prior distributions for the intercepts and slopes for each trajectory. If the investigator had very specific beliefs about what those trajectories looked like in the population, then this certainty would be incorporated into the prior.

A prior distribution that contains very specific knowledge about the parameter value is often referred to as an informative prior because the investigator is incorporating a lot of certainty about the distribution of that parameter. For example, if the investigator had strong prior knowledge about the initial reading levels for each of the two latent classes in Figure 3, then informative priors might be specified for the intercept parameters for each latent class. Specifically, the investigator may assume that initial reading levels are normally distributed and that Class 1 (i.e., the higher reading level class) has an average reading level score about 40.

The degree of certainty surrounding this guess of an average value of 40 can also be captured by the amount of variance incorporated into the normal prior. So far, the investigator is assuming (or making a guess) that the intercept for Class 1 follows as distribution of $\mathcal{N}(40, \sigma^2)$, where 40 represents the mean of the normal distribution and $\sigma^2$ represents the variance. If there is a lot of certainty that 40 is the initial reading level, then the variance may be fixed as a relatively low value (e.g., $\sigma^2 = 5$). This would create a prior for the intercept that was $\mathcal{N}(40, 5)$. In this example, the values comprising the normal distribution (i.e., 40 and 5) represent hyperparameters. Hyperparameters are the parameters comprising a probability distribution, and they are manipulated to represent either more or less certainty about a parameter value when specifying the corresponding prior.

If, on the other hand, the investigator was not very certain about the guess of 40 representing the average reading level for Class 1, then this uncertainty would be incorporated into the prior accordingly. In this case, the variance hyperparameter for the normal prior would be increased to incorporate more uncertainty in the prior distribution. The prior might then be specified as $\mathcal{N}(40, 100)$

[3] The author would like to thank an anonymous reviewer for highlighting the fact that separation among latent classes can be determined using multiple techniques. The MD information based on the manifest variables for the study conditions detailed below can be found in the Appendix.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

to reflect a great deal of variation surrounding the initial guess of 40 for the intercept. When there is a good deal of uncertainty in the prior, then this is typically referred to as a noninformative prior (or diffuse prior). The information (or lack of information) integrated into the prior is then combined with the data during the estimation process, and the combination of these two pieces of information are then used to produce the posterior distribution, which can be summarized (typically) by a measure of central tendency that represents the final parameter estimate.

Tueller and Lubke (2010) found that the conventional frequentist framework did not recover latent classes well when sample sizes were smaller (e.g., $n = 100$) and class separation was poor (e.g., MD = 0.5). The Bayesian estimation framework may show improvement over frequentist estimation in modeling situations akin to that presented in Tueller and Lubke (2010). In particular, the specification of informative priors may aid in recovering poorly separated classes, especially when only smaller sample sizes are available. As an example, the two latent classes depicted in Figure 3 may not be able to be properly recovered using ML/EM since there is a moderate amount of trajectory overlap across the classes; this conjecture has been corroborated by results presented in Tolvanen (2008); Tofighi and Enders (2008); and Tueller and Lubke (2010). However, informative priors on the growth trajectories and latent class proportions (i.e., the proportion of cases in each of the two classes) may aid in proper recovery of the latent classes and the corresponding trajectories.

Theory indicates that with large amounts of data, the priors specified in the model do not have a significant impact since the data swamps the prior, thus making it irrelevant (Ghosh & Mukerjee, 1992). However, research has indicated that even noninformative priors can impact estimates using relatively large sample sizes. One such study examining the impact of noninformative priors on sample size was conducted by Natarajan and McCulloch (1998) in the context of a probit model with a single random effect. Findings indicated that diffuse priors impacted the posterior distribution even for relatively moderate/large sample sizes (e.g., $n = 100$) used for this regression model. Likewise, Lambert, Sutton, Burton, Abrams, and Jones (2005) conducted a similar study that addressed the impact of different forms prior distributions (all noninformative) on a random effects meta analysis model. This study found that the choice of the prior distribution had a large impact on the estimated posterior distribution, especially when the number of studies in the simulated meta-analysis data was relatively small (e.g., 5, 10, and 30 studies in a meta-analysis, which acts akin to a small sample size).

The current study aims at addressing the direct impact of both informative and noninformative prior distributions in the context of different sample sizes. This study is, in part, a replication of the work by Natarajan and McCulloch (1998) and Lambert et al. (2005) in that different levels of informativeness (noninformative versus informative) of the priors will be directly assessed in relation to varying levels of sample size. However, this study extends the investigation to mixture models, as well as varying degrees of mixture class separation. Note that conventional ML/EM estimation is also presented.

Along this continuum of different forms of priors (noninformative/informative), a different amount of theory-based knowledge is integrated into the model. When noninformative priors are implemented (or even ML/EM—since no additional knowledge is incorporated into the model), the situation may be viewed as more exploratory in nature because very little theory-based knowledge is being integrated into the model. However, when informative priors are used for model parameters, this may be viewed as more "confirmatory" in the sense that knowledge is directly integrated into the model through the priors.[4] Specifically, informative priors represent particular knowledge about the number (and size) of latent classes, the shape of trajectories, and other parameter values. The continuum of priors is in this sense linked to the model being either more or less "confirmatory" in nature.

## Growth Mixture Modeling Specification

For any mixture model, we first assume that the data are generated from a mixture distribution ($f(\mathbf{y}_i \mid \Psi)$) represented by the following mixture density function for mixture class $c$ such that

$$f(\mathbf{y}_i|\boldsymbol{\Psi}) = \sum_{c=1}^{C} \pi_c f_c(\mathbf{y}_i|\theta_c), \tag{1}$$

where $\mathbf{y}_i$ is a vector of repeated measure outcomes for person i across $T$ time points, $\boldsymbol{\pi}_c$ represents the unknown mixture class proportion for the $c$th mixture class with $c = (1, 2, \ldots, C)$, and $f_c$ are the densities across the $C$ latent classes that are assumed to be multivariate normal: $y|c \sim \mathcal{MN}(\mu_c, \Sigma)$, where $\mu_c$ and $\Sigma$ are defined below. Further,

$$\boldsymbol{\Psi} = (\boldsymbol{\pi}, \boldsymbol{\Theta}')' \tag{2}$$

is a vector of unknown parameters that include the mixing proportions $\pi = (\pi_1, \pi_2, \ldots, \pi_C)$ and the model parameters $\boldsymbol{\Theta} = (\theta_1', \theta_2', \ldots, \theta_c')'$, where $\theta_c$ represents the model parameters for latent class $c$.

The specification of GMM includes both a measurement model and a structural model. The measurement model can be seen as

$$y_i = \Lambda_y \eta_{ic} + \varepsilon_i \tag{3}$$

where $\mathbf{y}_i$ still represents a vector of repeated measure outcomes for person $i$. The $\Lambda_y$ term is essentially a matrix of factor loadings, with $T$ (number of time points) rows and $K$ (number of latent factors) columns. The first column is fixed as all 1s and the remaining $K - 1$ columns represent constant time values. The $\boldsymbol{\eta}_{ic}$ term is a vector of latent factors (growth parameters in this case) that has $K$ elements. Finally, $\varepsilon_i$ represents a vector of normally distributed residuals with a diagonal covariance matrix $\Omega_\varepsilon$. The residual variances are typically assumed equal across time and uncorrelated across time, but these restrictions can be relaxed if desired. Note that the current investigation assumes that residual variances are equal across time, and independence is also assumed between the elements of structured error matrix $\Omega_\varepsilon$, which is further illustrated through the specification of prior distributions presented below.

[4] The use of the word "confirmatory" here is meant to form an analogy for using informative priors. In this case, informative priors are "confirmatory" only in the sense that they reflect specific (theory-driven) knowledge about parameters values. This knowledge is then used in the model estimation process to inform results. Note that the use of the word "confirmatory" here differs from a confirmatory model, where a model contains restrictions based on theory and those restrictions are then evaluated to see if they are consistent with the data.

The growth parameters can still be considered as random variables in this modeling perspective, and this is addressed in the structural part of the model, which is as follows:

$$\eta_{ic} = \alpha_c + \zeta_i, \tag{4}$$

where $\eta_{ic}$ still represents a vector of the growth factors, $\alpha_c$ is a vector of factor means, and $\zeta_i$ is a vector of normally distributed deviations of the parameters from their respective population means with distribution $\mathcal{N}$ (0, $\Omega_\zeta$), where $\Omega_\zeta$ is the variance/covariance matrix. The $c$ subscripts in this equation indicate that the factors are allowed to vary across latent classes such that $c$ = (1, 2, . . ., $C$) represents the mixture class variable, and the $i$ subscripts allow the parameters to vary across individuals. Combining Equations 3 and 4 produces a reduced form, where

$$\boldsymbol{y}_i = \Lambda_y(\alpha_c + \zeta_i) + \varepsilon_i. \tag{5}$$

Further, the model-implied mean and covariance of this reduced form can be written as

$$\mu_c = \Lambda_y \alpha_c \tag{6}$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}_y \boldsymbol{\Omega}_\zeta \boldsymbol{\Lambda}'_y + \boldsymbol{\Omega}_\varepsilon, \tag{7}$$

where $\boldsymbol{\mu}_c$ represents the mean vector of the repeated measure $y$s that varies across latent classes $c$s, and $\boldsymbol{\Sigma}$ represents the covariance matrix of the $y$s. It should be noted here that covariance matrices need not be held homogeneous across classes.

## Bayesian GMM: Specification of Prior Distributions

As is true with any type of model, the key to specifying GMM within a Bayesian estimation framework is properly setting up the prior distributions on the model parameters. This section describes the conjugate priors for each of the parameters estimated in GMM as presented in Lee (2007) and Asparouhov and Muthén (2010).[5] Although it is true that nonconjugate priors can be specified for any model, this is typically not encouraged for mixture models (see e.g., Diebolt & Robert, 1994; Lee, 2007). The use of fully noninformative priors (e.g., uniform) can lead to improper posterior distributions. As a result, it is common for mixture models to be specified with conjugate priors to avoid this problem altogether.

Next, the prior distributions can be specified for the unknown parameters in the model. The parameters estimated in this model are the mixture class proportions ($\pi_c$), the (growth) factor means ($\boldsymbol{\alpha}_c$), the variance of the residuals ($\boldsymbol{\Omega}_\varepsilon$), and the factor variance/covariance matrix ($\boldsymbol{\Omega}_\zeta$). To begin, the process assigning individuals to particular latent classes is assumed to follow a multinomial distribution with a sample size parameter n and a class proportion parameter $\pi$. The conjugate prior for this class proportion parameter $\pi_c$ is the Dirichlet distribution denoted as

$$\pi_c \sim \mathcal{D}[\delta_1 \ldots \delta_C],$$

with the hyperparameter(s) $\delta_1 \ldots \delta_C$, which control how uniform the distribution will be. Specifically, these parameters represent the proportion of cases in the $C$ latent classes.

The next model parameters to receive prior distributions are the (growth) factor means, which are distributed normally:

$$\boldsymbol{\alpha}_c \sim \mathcal{N}[\mu_{\alpha_c}, \boldsymbol{\Omega}_\zeta],$$

where $\mu_{\alpha_c}$ represents the expectation for the factor means, and $\boldsymbol{\Omega}_\zeta$ represents the factor variances and covariances.

The next prior to specify is for the variances of the residuals denoted above as $\Omega_\varepsilon$. Note that in order to specify a prior for an individual cell in the $\Omega_\varepsilon$ matrix, the notation will be expanded out to represent individual elements in the $J \times J$ matrix. Specifically, let $\omega_{\varepsilon_{jj}}$ represent a single cell in the variance/covariance matrix $\boldsymbol{\Omega}_{\varepsilon_{ij}}$. The conjugate prior specified here for the residual variances is the inverse gamma (IG) distribution and can be seen as

$$\omega_{\varepsilon_{jj}} \sim \mathcal{IG}[a_{\omega_{\varepsilon_{jj}}}, b_{\omega_{\varepsilon_{jj}}}],$$

where the hyperparameters $a$ and $b$ represent the shape and scale parameters for the IG distribution, respectively.

The last prior distribution to be specified is for the matrix of factor variances and covariances denoted as $\Omega_\zeta$. Recall that from Equation 5, $\zeta_i$ represents a vector of deviations of the parameters from their respective population means. The conjugate prior specified here for the factor variance/covariance matrix is the Inverse Wishart (IW) distribution and is denoted as

$$\Omega_\zeta \sim \mathcal{IW}[\Omega, d],$$

where $\Omega$ is a positive definite matrix of size $p$, and $d$ is an integer that can vary depending on the informativeness of the prior distribution. Note that for setting up an IW prior distribution, each term in the variance/covariance matrix can receive a prior. As a result, the $\Omega$ hyperparameter can actually be replaced with a constant rather than a matrix.

## Relevant MCMC Convergence Diagnostic

Although there are several different convergence diagnostics used to assess parameter convergence within the MCMC estimation algorithm, one of the most common diagnostics to employ is the Brooks, Gelman, and Rubin diagnostic. This method of assessing convergence originated with the work by Gelman and Rubin (see e.g., Gelman, 1996; Gelman & Rubin, 1992a; Gelman & Rubin, 1992b), who designed a diagnostic called the potential scale reduction (PSR) factor, which is based on the theory of analysis of variance. The PSR factor was intended to assess convergence among several parallel chains with varying starting values. Note that this diagnostic can also be used to assess convergence in a single MCMC chain by comparing the first portion of the post burn-in iterations to the last portion of the chain; see L. K. Muthén and Muthén (2012) for more details of how this can be implemented.[6] The current study employs this method of convergence assessment on a single MCMC chain. Although the Brooks, Gelman, and Rubin diagnostic is the only assessment used to determine parameter convergence here, note that there are several

[5] A conjugate prior is mathematically convenient in that it leads to a posterior belonging to the same distributional family as the prior.

[6] The burn-in phase of a chain refers to the first portion of the chain that is often discarded due to chain fluctuations. The post-burn-in phase is the portion of the chain that has presumably converged, and it is treated as the posterior.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

other diagnostics commonly implemented in the MCMC literature (see e.g., Sinharay, 2004).

## Design

Data for this study were simulated using the Monte Carlo framework within the M*plus* Version 7 software program (L. K. Muthén & Muthén, 2012). A 3-class GMM was specified across four time points in order to study mixture class recovery under four levels of class separation. This section begins with a description of the individual mixture class trajectories generated in the model. Next is a description of the factors included in the study, where all of the levels of the following factors were fully crossed with one another: mixture class separation (MD = 2.0, 1.5, 1.0, and 0.5), trajectory shape (linear and quadratic), mixture class proportions (0.33/0.33/0.33, 0.45/0.45/0.10, and 0.70/0.20/0.10), sample size ($n$ = 150, and $n$ = 800), and estimator (ML/EM and Bayesian estimation with diffuse, informative, data-driven informative, weakly informative, partial informative, and inaccurate priors).

Growth trajectories were fixed within each latent class to represent four levels of mixture class separation. Mixture class separation was defined through a multivariate MD calculation between the growth parameters of the latent classes.[7] To begin, the trajectory of the first latent class (C1) was fixed to model substantive results from Kaplan (2002) that presented a study of reading development across four time points for kindergartners and first graders. The "fast reading development" latent class obtained in Kaplan's analysis was used as the first latent class in the current study. Class separation between the subsequent classes was then determined in relation to this trajectory set for C1.

The second latent class (C2) represented a latent class with a lower initial achievement level than C1. This class was fixed as having high latent class separation from C1 in all conditions. High separation was determined by setting the MD at 2.0, which is commonly viewed as representing high class separation among classes (see e.g., Lubke & Muthén, 2005; Tolvanen, 2008; Tueller & Lubke, 2010).

Finally, the third latent class (C3) was generated to represent a latent class with the relatively lowest initial achievement level compared to C1 and C2. The separation between C2 and C3 was varied to represent four different degrees of mixture class separation. Specifically, MD was fixed at 2.0, 1.5, 1.0, or 0.5 to represent degrees that varied from high to very poor latent class separation. The population values used to form these latent classes can be found in Table 1.

The trajectory shape of C3 was also varied across the cells of this design. Specifically, C3 was either fixed as having a linear trajectory or a quadratic trajectory. When C3 represented a quadratic trajectory, C1 and C2 were both set as linear trajectories. The covariance structures remained the same across classes with the addition of the quadratic term. The purpose of adding a quadratic trajectory for C3 was to vary how similar (or different) the shape of the C3 trajectory was from C2, regardless of mixture class separation. If separation is relatively poor (e.g., MD = 1.0), parameter recovery for a linear trajectory of C3 may be different from parameter recovery for a quadratic trajectory. Not only do the latent classes vary by MD separation, but under the quadratic level, the classes now also vary in growth shape. Note, however, that the addition of the quadratic term altered the MD values for the quadratic conditions to be 2.57, 1.50, 1.00, and 0.50. Separation based on MD calculated on the manifest variables is presented in the Appendix. Population values for the quadratic conditions are presented in Table 1. A depiction of the four levels of linear trajectories can be found in Figure 4, and a depiction of the four quadratic trajectory levels can be found in Figure 5.

The relative size of the three latent classes is also of concern here. Specifically, latent class size was manipulated such that C1, C2, and C3 had sample size proportions of 0.33/0.33/0.33, 0.45/0.45/0.10, or 0.70/0.20/0.10, respectively. The mixture proportion level of 0.33/0.33/0.33 was used to assess the impact of equal sample sizes across all classes, 0.45/0.45/0.10 was chosen to assess the impact of a clear minority class (C3), and 0.70/0.20/0.10 was used to examine the impact of a clear majority latent class (C1). Likewise, total sample sizes typical of applied research of $n$ = 150 and $n$ = 800 were used to examine whether there would be an impact of the latent class sample size on parameter recovery under different levels of mixture class separation.

Seven different estimation levels were compared within each level of mixture class separation: ML/EM using multiple starting values, Bayesian estimation via the MCMC algorithm using M*plus* default diffuse priors (MCMC-diffuse), MCMC using informative and "accurate" priors (MCMC-informative), MCMC using data-driven informative priors (MCMC-data-driven), MCMC using modified-diffuse (or weak) priors (MCMC-weak), MCMC using informative and "accurate" priors on a partial subset of parameters and default diffuse priors on the remaining parameters (MCMC-partial), and MCMC using informative and "inaccurate" priors (MCMC-wrong); specifics of the priors used in this study are detailed below. There were 1,000 replications requested for each cell of this design. Likewise, each MCMC analysis consisted of a single chain with the first 25,000 iterations discarded as the burn-in phase and the last 25,000 iterations used as the post burn-in iterations.

For the MCMC-diffuse level, the default noninformative priors provided by M*plus* (see, L. K. Muthén & Muthén, 2012) were used for all of the parameters. For example, a normally distributed parameter would contain a mean hyperparameter set arbitrarily at zero and a variance hyperparameter specified as 1,010 to indicate no precision in the prior. Likewise, the default Dirichlet prior placed on mixture class proportions assumed equal sample sizes with a prior of $\mathcal{D}(10,10,10)$ for three respective latent classes.

For the MCMC-informative level, informative priors were placed on the growth parameters. The mean hyperparameter for each prior was fixed as the corresponding growth parameter population value, whereas, the variance hyperparameter was determined in a systematic way for each growth parameter. Specifically, each variance hyperparameter was fixed at 5.00% of the corresponding population value. For example, if the population

[7] The multivariate Mahalanobis distance is used to compute the distance between two latent classes. Specifically, the equation to compute this distance is $MD = \Delta = \{(\mu_1 - \mu_2)'\Sigma^{-1}(\mu_1 - \mu_2)\}^{1/2}$, where $\Sigma^{-1}$ represents the inverse of the common covariance matrix, and the $\mu_1$ and $\mu_2$ terms represent the means for the first and second latent classes, respectively (McLachlan & Peel, 2000). In this case, the means would be the intercept and slope growth parameters for each trajectory. For ease of computation, the variance/covariance matrix was the same for all three classes in this study. Note that the Mahalanobis distance can also be computed for manifest variables, if desired.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 1
*Population Model Values*

| Parameter | High separation | Moderate separation | Poor separation | Very poor separation |
|---|---|---|---|---|
| Linear conditions | | | | |
| Growth parameters: C1 | | | | |
| Intercept | 48.000 | 48.000 | 48.000 | 48.000 |
| Slope | 3.000 | 3.000 | 3.000 | 3.000 |
| Growth parameters: C2 | | | | |
| Intercept | 42.802 | 42.802 | 42.802 | 42.802 |
| Slope | 4.000 | 4.000 | 4.000 | 4.000 |
| Growth parameters: C3 | | | | |
| Intercept | 37.600 | 38.558 | 39.790 | 42.530 |
| Slope | 3.000 | 3.000 | 3.000 | 3.000 |
| Variances: All classes | | | | |
| Intercept variance | 18.000 | 18.000 | 18.000 | 18.000 |
| Slope variance | 2.000 | 2.000 | 2.000 | 2.000 |
| All time-specific residuals | 15.000 | 15.000 | 15.000 | 15.000 |
| Fixed covariances: All classes | | | | |
| Slope with intercept | 1.200 | 1.200 | 1.200 | 1.200 |
| Quadratic conditions | | | | |
| Growth parameters: C1 | | | | |
| Intercept | 48.000 | 48.000 | 48.000 | 48.000 |
| Slope | 3.000 | 3.000 | 3.000 | 3.000 |
| Growth parameters: C2 | | | | |
| Intercept | 42.802 | 42.802 | 42.802 | 42.802 |
| Slope | 4.000 | 4.000 | 4.000 | 4.000 |
| Growth parameters: C3 | | | | |
| Intercept | 37.600 | 38.558 | 39.790 | 42.530 |
| Slope | 3.000 | 3.000 | 3.000 | 3.000 |
| Quadratic | 0.900 | 0.900 | 0.900 | 0.900 |
| Variances: All classes | | | | |
| Intercept variance | 18.000 | 18.000 | 18.000 | 18.000 |
| Slope variance | 2.000 | 2.000 | 2.000 | 2.000 |
| Quadratic variance | 0.000 | 0.000 | 0.000 | 0.000 |
| All time-specific residuals | 15.000 | 15.000 | 15.000 | 15.000 |
| Fixed covariances: All classes | | | | |
| Slope with intercept | 1.200 | 1.200 | 1.200 | 1.200 |
| Slope with quadratic | 0.000 | 0.000 | 0.000 | 0.000 |
| Intercept with quadratic | 0.000 | 0.000 | 0.000 | 0.000 |

*Note.* All levels of trajectory shape were fully crossed with the three mixture class proportion levels (0.33/0.33/0.33, 0.70/0.20/0.10, and 0.45/0.45/0.10), the two levels of sample size (150 and 800 cases), and the seven levels of estimation (maximum likelihood via the EM algorithm, MCMC using informative priors, MCMC using data-driven priors, MCMC using partial-knowledge priors, MCMC using Mplus default-diffuse prior, MCMC using weak priors, and MCMC using weak but wrong ["inaccurate"] priors). EM = expectation maximization; MCMC = Markov chain Monte Carlo; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

value for a slope parameter is 3.00, the variance hyperparameter would be fixed at $0.05 * 3 = 0.15$, which produces a prior of $\mathcal{N}(3, 0.15)$ for this slope parameter. This method of determining the priors provided both a systematic way of computing the variance hyperparameter as well as a presumably reasonable, informative prior for each growth parameter. The prior distributions for all growth parameters were determined in a similar fashion under the MCMC-informative level. Likewise, informative priors were also placed on the mixture class proportions, indicating accurate knowledge of class sizes. A Dirichlet prior was placed on mixture class proportions such that for $n = 150$ with equal latent class sizes the prior would be $\mathcal{D}(50,50,50)$, with 0.45/0.45/0.10 the prior would be $\mathcal{D}(67, 67,15)$, and with 0.70/0.20/0.10 the prior would be $\mathcal{D}(105,30,15)$.[8] It should be noted here that the MCMC-informative level is likely not indicative of an applied research scenario since population parameters, which are typically not known, were utilized to define the priors. However, this level should be viewed as representing the upper bound performance of Bayesian GMM that might be expected in an empirical study.

Given that the MCMC-informative level represents the upper bound performance of Bayesian GMM, another (perhaps more applicable) level of informative priors was included here. Specifically, these informative priors were data-driven in that they were defined based on results obtained from ML/EM and did not directly contain information from population values since population

[8] The Dirichlet priors are handled in a slightly different manner in M*plus*. That is, an informative Dirichlet prior is specified on each mixture class threshold [c#i] as $\mathcal{D}(\alpha_i, \alpha_K)$, where $i = 1, \ldots, K - 1$. For example, for $n = 150$ and 0.70/0.20/0.10 mixture class proportions, we would have $[c\#1 * 1.9459] \sim \mathcal{D}(105,15)$ and $[c\#2 * 0.6931] \sim \mathcal{D}(30,15)$.

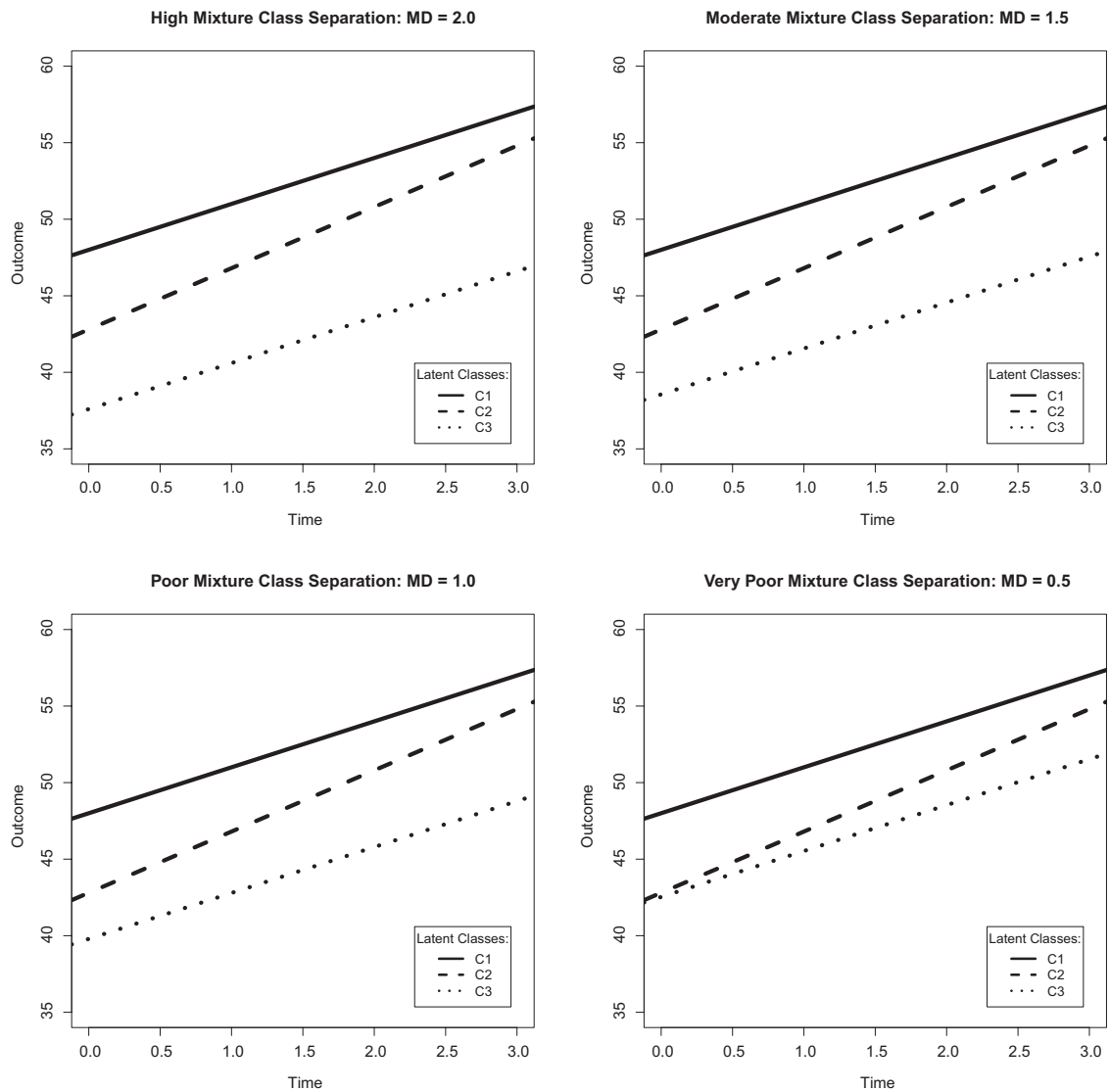


*Figure 4..* Linear growth trajectories for different mixture class separation levels. C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; MD = Mahalanobis distance.

values are rarely known in applied settings. In particular, the ML/EM average parameter estimate across 1,000 replications was used as the mean hyperparameter for the intercept and slope parameters. Likewise, the average variance of the estimates across the 1,000 replications was used as the variance hyperparameter. There were four ML/EM analyses that informed the priors for this estimation level representing each of the four levels of class separation in the population. Each of the ML/EM analyses driving these priors were based on $n = 800$ cases, 0.33/0.33/0.33 class proportions, and 1,000 replications (as mentioned). Prior distributions for this level are presented in Table 2. The aim of using ML/EM results to define priors was to mimic the research scenario where previous GMM results are used to help define informative priors for Bayesian GMM. Class proportions received $\mathcal{D}(1,1,1)$ priors to indicate no knowledge of class sizes. This data-driven prior level is referred to as MCMC-data-driven.

The default diffuse priors in M*plus* that were implemented in the MCMC-diffuse level are not necessarily indicative of diffuse (or weakly informative) priors typically specified in the Bayesian literature. Specifically, there are potential problems with $\mathcal{D}(10,10,10)$ acting as an informative prior—especially when latent class sample sizes are smaller. Likewise, $\mathcal{N}(0,10^{10})$ may in some cases act akin to an improper prior (which is not a known probability density) since the variance hyperparameter is quite large. As a result, a modified-diffuse level that specifies weakly informative priors (MCMC-weak) was also implemented here. In particular, a prior of $\mathcal{D}(1,1,1)$ was used for the latent class proportions, and a normal prior was used for the intercepts and slopes, where the population value was used as the mean hyperparameter and the variance hyperparameter was set as representing 50% of the population value. Both of these priors are weakly informative, and they represent priors that are perhaps more reasonable and

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

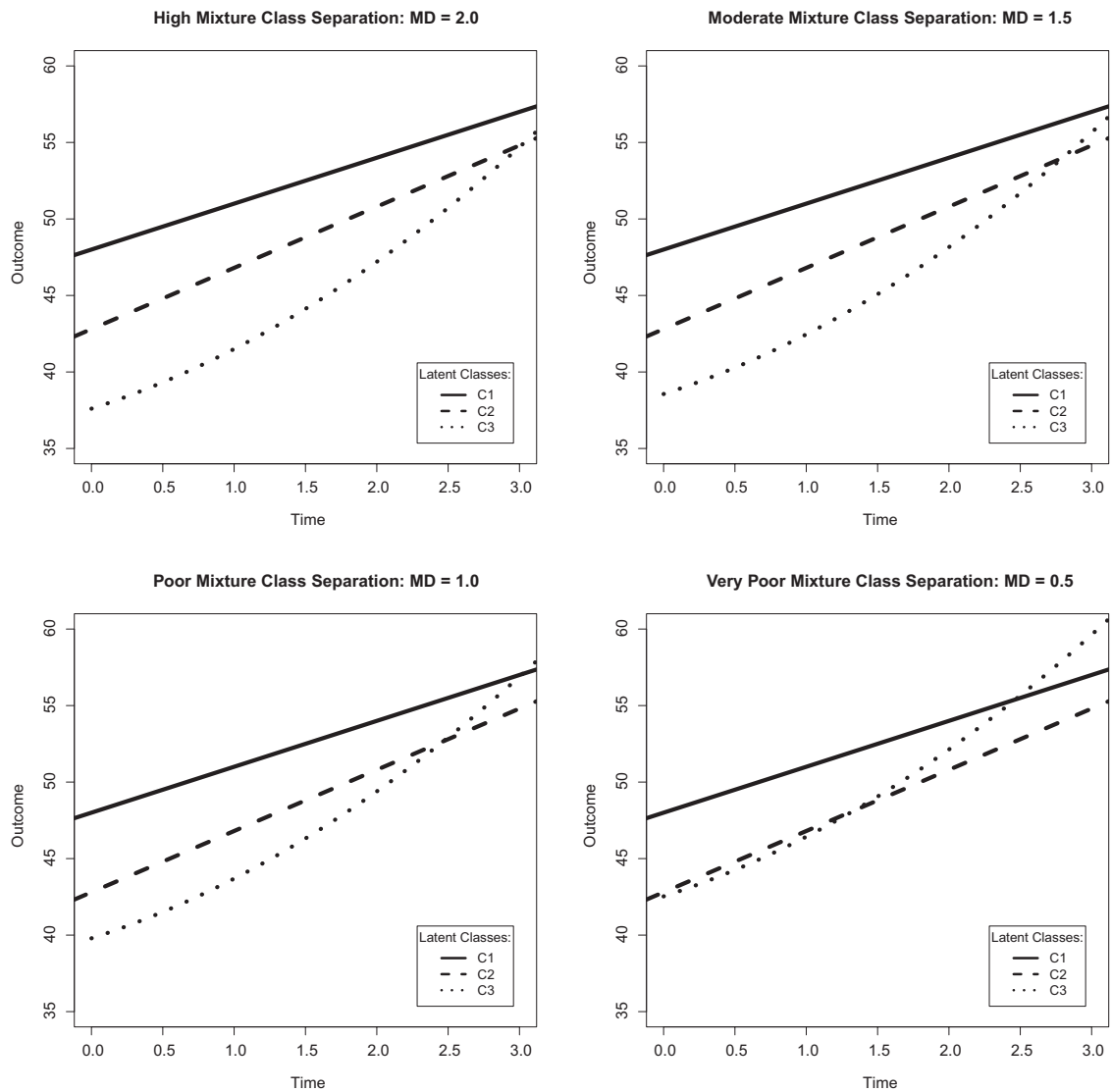


*Figure 5..* Quadratic growth trajectories for different mixture class separation levels. C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; MD = Mahalanobis distance.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

indicative of the applied Bayesian literature than the default diffuse priors in M*plus*.[9]

For the MCMC-partial level, informative priors were placed only on the intercept growth parameters for the three latent classes. Priors for the intercepts were specified in the same manner as in the MCMC-informative level described above. The mixture class proportions and slopes for each class were given default diffuse priors. This Bayesian estimation level was designed to reflect prior knowledge of initial starting levels only, with no prior knowledge of latent class size or growth rate.

For the MCMC-wrong level, the prior distributions on the growth parameters were identified in a systematic manner and were included to assess the impact of "inaccurate" mean hyperparameters and weaker variance hyperparameters. Specifically, the variance hyperparameter term was fixed as 50% of the corresponding population value. Additionally, for each of the growth parameters, the mean hyperparameter was decreased by 3 standard deviations based on this fixed variance hyperparameter. As an example, take the intercept parameter for C1, which is 48. The variance hyperparameter would represent 50% of this value, which is $48 * 0.50 = 24$. Under this variance hyperparameter, the mean hyperparameter would be (population $-$ 3*standard deviation) $= 48 - (3*\sqrt{24}) = 33.30$, producing a prior distribution of

[9] Another level of modified-diffuse priors was preliminarily examined where a $\mathcal{D}(1,1,1)$ prior was used for class proportions and a $\mathcal{N}(0,100)$ prior was used for the intercepts and slopes. Bias levels produced from this condition were incredibly high for all model parameters (e.g., bias levels consistently greater than 100%). It was therefore determined that a more informative prior should be specified for the intercepts and slopes. The MCMC-weak estimation condition presented here reflects this modification.

Table 2
*Prior Distributions for the MCMC-Data-Driven Level*

| Parameter | High separation | Moderate separation | Poor separation | Very poor separation |
|---|---|---|---|---|
| Linear conditions | | | | |
| Growth parameters: C1 | | | | |
| Intercept | N(48.2118,10.9773) | N(48.2553,12.4920) | N(48.5667,14.8025) | N(49.6126,17.8481) |
| Slope | N(2.7811,2.6830) | N(2.7415,3.8255) | N(2.7111,4.1031) | N(2.8431,4.3072) |
| Growth parameters: C2 | | | | |
| Intercept | N(42.4359,5.6136) | N(42.7795,4.4823) | N(43.2457,4.1408) | N(44.4104,3.2141) |
| Slope | N(3.7960,3.7022) | N(3.8210,3.7149) | N(3.5951,4.3322) | N(3.2727,4.6924) |
| Growth parameters: C3 | | | | |
| Intercept | N(37.1984,9.4274) | N(38.1169,11.5254) | N(39.0328,12.2920) | N(40.3106,14.4354) |
| Slope | N(3.0510,3.0199) | N(3.1228,4.2750) | N(3.4936,5.2997) | N(3.8362,5.4452) |
| Quadratic conditions | | | | |
| Growth parameters: C1 | | | | |
| Intercept | N(47.8868,12.5770) | N(48.0601,13.8302) | N(48.4626,15.8356) | N(49.1358,27.8130) |
| Slope | N(2.8234,18.9199) | N(2.8950,22.0336) | N(2.7798,24.2025) | N(2.6979,23.3540) |
| Growth Parameters: C2 | | | | |
| Intercept | N(42.3945,6.9981) | N(42.9233,5.3453) | N(43.4317,4.17222) | N(44.4341,2.9388) |
| Slope | N(3.7859,10.0927) | N(3.6163,11.2923) | N(3.5757,9.7919) | N(3.5131,10.7965) |
| Growth Parameters: C3 | | | | |
| Intercept | N(37.6675,8.1465) | N(38.6984,7.8865) | N(39.5838,8.5656) | N(41.2719,8.5416) |
| Slope | N(2.9678,17.4164) | N(3.1492,16.0769) | N(3.3248,16.8650) | N(3.6289,16.3928) |
| Quadratic | N(0.9141,1.8217) | N(0.7915,1.6750) | N(0.6979,1.8082) | N(0.5489,1.9687) |

*Note*. MCMC = Markov chain Monte Carlo; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; N = normally distributed prior distribution.

$\mathcal{N}(33.30, 24)$ for the intercept parameter for C1 containing a true population value of 48. Class proportions received default diffuse Dirichlet priors.

The current investigation defined the simulations such that the number of $K$ latent factors estimated was prespecified in the population and analysis models. Further, the number of C latent classes was also prespecified.

The purpose of including all of these factors was to assess parameter recovery under very specific trajectory conditions in order to uncover the ability of latent classes to be properly recovered in different modeling situations. Latent class recovery was assessed through the computation of parameter bias. Specifically, bias was computed for each parameter based on the average estimate produced by the simulations.[10] Bias values less than 10.00% were deemed as representing low or negligible parameter bias, whereas anything greater than or equal to 10.00% was considered to be exhibiting high and therefore problematic parameter bias. However, all levels of bias rates are presented in the tables so that the full pattern of results can be interpreted by the reader.

## Results

Only parameter bias results for $n = 150$ are presented in tables here. Specific results for $n = 800$ are summarized here, and tables can be made available upon request. All problematic bias levels greater than 10.00% are presented in bold in the following tables; mean squared error (MSE) values are also presented for each parameter estimated, as well as the estimated proportion of cases in each class.

### Label Switching and Convergence Rates

One issue that can arise when estimating mixture models via MCMC sampling is referred to as label switching. It occurs when the ordering of latent classes arbitrarily changes during the MCMC chain. Not only can this affect the final estimates produced, but label switching can also complicate the assessment of convergence. One common method used for preventing label switching is to specify identifiability constraints within the model. In the current study, identifiability constraints were specified within M*plus* based on the intercept growth factors such that C1 > C2 > C3. Each iteration of the Gibbs sampler was computed such that the specified constraint was satisfied, and thus results were free from label switching. This method can help prevent the chain from arbitrarily sampling an alternative latent class midchain (see e.g., Diebolt & Robert, 1994; Frühwirth-Schnatter, 2001). This type of constraint is typical to use in preventing within-chain label switching from occurring. To prevent between-chain label switching (i.e., when two or more MCMC chains are requested and between-chain convergence is of concern), only one chain was used to assess convergence in the current study. However, if multiple chains had been used, the PSR factor can aid in identifying between-chain label switching. That is, convergence will not be obtained via the PSR factor if two chains do not converge with each other even if they show within-chain convergence.

Likewise, the PSR factor can also capture within-chain label switching since dramatic fluctuations within the chain would result in nonconvergence according to the PSR factor. Employing identifiability constraints is just one method for preventing label switching. For a more comprehensive review see Celeux, Hurn,

[10] Percent of bias is computed by using the following equation: 100*([estimate − population value]/population value).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

and Robert (2000); Jasra, Holmes, and Stephens (2005), and Stephens (2000).[11]

As mentioned above, M*plus* examines chain convergence within the MCMC estimation algorithm using the Brooks, Gelman, and Rubin diagnostic. This diagnostic produces a PSR factor to assess convergence within a given chain. Note that a PSR factor of 1.0 would indicate perfect convergence between the first portion of the post burn-in iterations and the last portion of the chain. The convergence criterion implemented in the current study required the PSR factor to fall within ±.05 of 1.0. In other words, convergence was obtained if the PSR factor fell within the bound of 0.95 and 1.05. If this criterion was not met for a particular chain, then convergence for that replication was not obtained. Note that only replications where convergence was obtained were included in the parameter estimates.

For the ML/EM estimator, the number of starts for the models specified in this study was 100, which follows general guidelines presented in Hipp and Bauer (2006). To prevent across-replication label switching from occurring across Monte Carlo replications, it is common to implement user-specified starting values for all parameters to ensure that the class labeling will be consistent across replications.[12] User-specified starting values that matched the population values were implemented here. Potential issues arising from across-replication label switching were prevented with the use of these user-specified starting values as well as the identifiability constraints described above. The 100 starts that M*plus* generated were perturbations of these user-specified starting values.[13] The number of final stage optimization steps was 10. These settings were used to ensure that estimates were not the result of problems with local maxima.

One issue that can arise when estimating complex models (e.g., finite mixture models) is that of nonconvergence. In an empirical setting, nonconvergence can be a result of several different issues that include incorrectly specifying the model, multivariate outliers, and model nonidentification (Tueller & Lubke, 2010). Of particular note here, nonconvergence can be a consequence of empirical nonidentification where an insufficient number of cases comprise a latent class. In this situation, there are too few cases in a latent class to estimate the parameters that are specified in the model. Empirical nonidentification may occur particularly under conditions containing a very small number of cases in a latent class. In a Monte Carlo setting, convergence rates are typically reported alongside results in order to provide information about the confidence one should have in the simulation findings, and results should always be interpreted in the context of these convergence rates. In a simulation study, if convergence is not obtained for a particular replication, then that replication is not incorporated into the parameter estimates that are reported. Convergence rates were combined for all class separation cells and are presented in Table 3. All results that are presented in the following section should be interpreted in the context of convergence rates. In particular, there is less confidence in results from conditions with relatively low convergence rates (e.g., MCMC-weak) compared to conditions with higher convergence rates (e.g., ML/EM).

Table 3
*Convergence Rates (out of 1,000 Replications) Collapsed Across Class Separation Levels*

| Estimation level | Linear | | Quadratic | |
|---|---|---|---|---|
| | $n$ = 150 | $n$ = 800 | $n$ = 150 | $n$ = 800 |
| ML/EM | ≥999 | ≥999 | ≥999 | ≥999 |
| Info | ≥964 | ≥962 | ≥842[a] | ≥690[b] |
| Default–diffuse | ≥739 | ≥700 | 76–156 | 81–229 |
| Partial | ≥947 | ≥855 | ≥502[c] | ≥350[d] |
| Wrong | ≥867 | ≥826 | 195–415 | 119–332 |
| Weak | ≤31 | ≤216 | ≤10 | ≤16 |
| Data-driven | ≤191 | 132–551 | ≤16 | ≤49 |

*Note.* ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default–diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data-driven = Bayesian estimation using data-driven priors; MD = Mahalanobis distance.
[a] Except for the cell of very poor separation (MD = 0.5) under 0.33/0.33/0.33, which had a convergence rate of 611. [b] Except for three cells under very poor separation (MD = 0.5), which had convergence rates of 371, 494, and 506. [c] Except for three cells under very poor separation (MD = 0.5), which had convergence rates of 298, 335, and 338. [d] Except for three cells under very poor separation (MD = 0.5), which had convergence rates of 162, 168, and 248.

## Parameter Estimate Bias

Tables 4–5 present parameter bias percentages for the condition of the linear growth trajectory and latent class proportions of 0.33/0.33/0.33. Latent class proportions were properly recovered for MCMC-informative, MCMC-diffuse, and MCMC-partial. For growth trajectory parameters, MCMC-informative and MCMC-data-driven produced the lowest bias levels under all class separation levels, thus indicating the trajectory shapes were well recovered. MCMC-informative only produced problematic bias for the growth parameter variances (which received diffuse priors). Likewise, MCMC-data driven only produced higher bias for the growth parameter variances and the C2 slope term. In contrast, MCMC-diffuse and ML/EM produced the highest bias for the model parameters. Specifically, the trajectory shapes under these diffuse conditions were not recovered properly, even under high mixture class separation (i.e., MD = 2.0). This general finding was also the case to an extent under MCMC-wrong, where higher bias was exhibited for many of the model parameters. Notice also that none of the estimation conditions properly recovered the intercept

[11] It should also be mentioned that if label switching does occur, the researcher can use a relabeling algorithm to reassign the class labels across the latent classes by performing a k-means type clustering of the MCMC samples (see e.g., Farrar, 2006; Stephens, 2000).

[12] Across-replication label switching occurs when latent classes are arbitrarily ordered across Monte Carlo replications such that estimates averaged across replication are meaningless for the latent classes. For example, the first replication may have ordered "Class 1" first and "Class 2" second, and the second replication may have reversed this ordering with "Class 2" appearing first and "Class 1" appearing second. Averages across these two replications would be meaningless without first reordering the latent classes to be consistent.

[13] It is important not to use random starts in Monte Carlo studies since this may result in across-replication label switching (Nylund et al., 2007). Note that the same user-specified starting values were used for ML/EM and MCMC conditions for consistency across estimation levels.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 4
*Parameter Estimate Bias: 0.33/0.33/0.33 Mixture Class Proportions, Linear Conditions, 150 Cases*

| | High separation | | | | | | |
|---|---|---|---|---|---|---|---|
| Parameter | ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| C1 proportion | 7.28 (0.36) | −0.32 (0.33) | 0.13 (0.33) | −0.76 (0.33) | **23.63** (0.43) | **15.00**[a] (0.38) | 0.59 (0.34) |
| C2 proportion | **11.20** (0.37) | 1.03 (0.34) | 2.39 (0.34) | 4.43 (0.35) | **12.16**[a] (0.38) | **32.78** (0.44) | **27.17** (0.42) |
| C3 proportion | **−18.45** (0.27) | −0.67 (0.33) | −2.49 (0.33) | −3.65 (0.32) | **−35.76** (0.19) | **−47.75** (0.17) | **−27.72** (0.24) |
| C1 | | | | | | | |
| Intercept (*MSE*) | −0.80 (14.65) | −0.57 (0.50) | −2.56 (2.78) | −0.96 (0.53) | −3.79 (4.25) | −2.33 (3.53) | −1.60 (1.85) |
| Slope (*MSE*) | −0.89 (4.76) | 0.75 (0.02) | 6.80 (0.38) | 5.25 (0.36) | −1.25 (0.24) | −2.12 (0.22) | −1.71 (0.40) |
| C2 | | | | | | | |
| Intercept (*MSE*) | −1.38 (6.72) | 0.04 (0.34) | 0.02 (0.81) | 0.05 (0.22) | −3.73 (3.44) | −2.15 (3.12) | −1.02 (0.43) |
| Slope (*MSE*) | **−16.95**[a] (7.12) | −0.36 (0.03) | **−11.49**[a] (0.62) | **−10.43** (0.56) | **−11.29**[a] (0.43) | −2.50 (0.17) | −8.65 (0.38) |
| C3 | | | | | | | |
| Intercept (*MSE*) | −1.37 (19.07) | 0.67 (0.39) | 3.03 (2.48) | 0.88 (0.35) | −3.67 (3.55) | −8.73 (56.31) | −0.19 (0.76) |
| Slope (*MSE*) | **14.53**[a] (6.84) | 0.33 (0.02) | 8.73 (0.42) | 8.79 (0.45) | **−27.61** (0.97) | −1.29 (0.16) | 1.79 (0.25) |
| Variances | | | | | | | |
| I with S (*MSE*) | **49.30** (9.30) | **−43.29**[a] (2.12) | **−43.00**[a] (2.35) | **−36.23**[a] (2.28) | **−74.68** (2.70) | **−37.69** (2.43) | **−69.88** (3.09) |
| Intercept (*MSE*) | **15.78**[a] (130.69) | **20.44** (45.57) | **46.44** (111.03) | **29.70** (58.88) | **37.98** (83.68) | **48.82** (152.73) | **72.73** (226.24) |
| Slope (*MSE*) | **−86.30** (4.61) | **14.02**[a] (0.61) | −7.45 (0.42) | −3.01 (0.42) | −1.05 (0.49) | **18.18**[a] (0.43) | 2.96 (0.39) |

*Note.* Bold values indicate problematic bias levels greater than 10.00% for $n = 150$ cases. Numbers in parentheses under C1–C3 class proportions represent actual mixture class proportions. Numbers in parentheses for growth parameters and variances represent mean squared error (*MSE*) values. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data driven = Bayesian estimation using data-driven priors; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; I with S = correlation between intercept and slope; *MSE* = mean squared error: The variance of the estimates across the replications plus the square of the bias.
[a] Indicates that percentage bias decreased below 10.00% when sample sizes were increased to $n = 800$.

and slope covariance; this result largely held in all subsequent conditions discussed below.

Tables 4–5 include superscripts denoted to indicate bias levels that decreased below 10.00% when sample sizes were increased to $n = 800$. Notice that several of the high bias levels decreased when a larger total sample size was used. However, note that latent class proportions still remained quite biased under ML/EM, MCMC-data-driven, MCMC-weak, and MCMC-wrong. Overall, under this

Table 5
*Parameter Estimate Bias: 0.33/0.33/0.33 Mixture Class Proportions, Linear Conditions, 150 Cases*

| | Poor separation | | | | | | |
|---|---|---|---|---|---|---|---|
| Parameter | ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| C1 proportion | 0.23 (0.33) | 0.48 (0.33) | −0.98 (0.33) | −0.27 (0.33) | **22.28** (0.40) | **25.23**[a] (0.41) | 1.62 (0.34) |
| C2 proportion | 19.32 (0.39) | 1.72 (0.34) | 4.40 (0.34) | 4.19 (0.34) | **18.01** (0.39) | **35.62** (0.45) | **37.01** (0.45) |
| C3 proportion | **−16.51** (0.28) | 0.83 (0.33) | −0.39 (0.33) | −0.90 (0.33) | **−37.26** (0.21) | **−57.82** (0.14) | **−35.60** (0.21) |
| C1 | | | | | | | |
| Intercept (*MSE*) | 0.50 (15.45) | −0.49 (0.46) | −1.70 (1.84) | −0.82 (0.45) | −3.02 (2.96) | −0.60 (2.56) | 0.56 (2.83) |
| Slope (*MSE*) | −1.88 (5.16) | 0.92 (0.02) | 6.43 (0.37) | 5.15 (0.34) | −3.56 (0.25) | 4.47 (0.28) | −1.86 (0.82) |
| C2 | | | | | | | |
| Intercept (*MSE*) | 0.72 (4.94) | 0.42 (0.30) | 1.48 (0.96) | 0.48 (0.22) | −1.26 (0.92) | −1.81 (1.99) | 0.23 (0.52) |
| Slope (*MSE*) | **−22.47** (7.59) | −0.45 (0.03) | **−14.22**[a] (0.63) | **−12.67**[a] (0.65) | −9.84 (0.37) | −3.87 (0.80) | **−18.64**[a] (0.95) |
| C3 | | | | | | | |
| Intercept (*MSE*) | −2.99 (18.74) | 0.20 (0.26) | 0.74 (0.88) | 0.48 (0.22) | −3.72 (3.84) | −20.22[a] (135.26) | −8.46 (52.19) |
| Slope (*MSE*) | **24.41** (7.77) | 0.36 (0.02) | **12.67**[a] (0.52) | **12.41**[a] (0.54) | **−25.42** (0.89) | **−10.12**[a] (0.91) | 2.65 (0.98) |
| Variances | | | | | | | |
| I with S (*MSE*) | **51.78** (8.62) | **−49.38**[a] (2.06) | **−47.85**[a] (1.95) | **−41.24**[a] (1.96) | **−75.94** (2.53) | **−99.70**[a] (4.09) | **−86.92** (3.56) |
| Intercept (*MSE*) | −0.17 (89.08) | **14.01**[a] (31.70) | **24.20** (47.30) | **21.07** (37.08) | **22.01** (43.27) | **19.15** (73.76) | **42.57** (114.89) |
| Slope (*MSE*) | **−92.16** (5.00) | **14.53**[a] (0.60) | −5.86 (0.38) | −7.39 (0.41) | −2.87 (0.46) | 4.86 (0.72) | 5.23 (0.51) |

*Note.* Bold values indicate problematic bias levels greater than 10.00% for $n = 150$ cases. Numbers in parentheses under C1–C3 class proportions represent actual mixture class proportions. Numbers in parentheses for growth parameters and variances represent mean squared error (*MSE*) values. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data driven = Bayesian estimation using data-driven priors; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; I with S = correlation between intercept and slope; *MSE* = mean squared error: The variance of the estimates across the replications plus the square of the bias.
[a] Indicates that percentage bias decreased below 10.00% when sample sizes were increased to $n = 800$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

| Moderate separation | | | | | | |
|---|---|---|---|---|---|---|
| ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| 5.10 (0.35) | −0.41 (0.33) | −1.31 (0.33) | −1.14 (0.33) | **23.79** (0.41) | 5.28 (0.35) | −2.02 (0.33) |
| **11.46** (0.37) | 0.94 (0.34) | 2.93 (0.34) | 4.16 (0.35) | **16.67**[a] (0.39) | **35.33** (0.45) | **34.80** (0.45) |
| **−16.53** (0.28) | −0.50 (0.33) | −1.59 (0.33) | −2.99 (0.32) | **−40.43** (0.20) | **−40.58** (0.20) | **−32.76** (0.22) |
| −0.27 (15.00) | −0.54 (0.48) | −2.17 (2.34) | −0.88 (0.52) | −3.42 (3.66) | 1.16 (3.29) | −0.78 (2.24) |
| −2.12 (6.61) | 0.85 (0.02) | 5.80 (0.33) | 4.81 (0.33) | −3.33 (0.27) | −5.58 (0.37) | −2.56 (0.41) |
| −0.47 (5.61) | 0.14 (0.31) | 0.64 (0.78) | 0.15 (0.21) | −2.56 (1.94) | −0.72 (3.77) | −0.42 (0.39) |
| **−18.84** (7.64) | −0.38 (0.03) | **−12.70**[a] (0.61) | **−11.51** (0.59) | **−10.25**[a] (0.39) | −1.57 (0.57) | **−12.60**[a] (0.46) |
| −1.92 (18.68) | 0.55 (0.34) | 2.09 (1.62) | 0.81 (0.31) | −3.61 (4.21) | **−20.06**[a] (194.79) | −3.60 (21.13) |
| **18.37**[a] (7.08) | 0.23 (0.02) | **10.37**[a] (0.47) | **10.33** (0.45) | **−26.49** (0.95) | −7.98 (0.35) | 5.96 (0.67) |
| **52.71** (9.16) | **−45.35**[a] (2.02) | **−42.47**[a] (2.13) | **−36.65** (1.96) | **−70.68** (2.63) | **−43.92** (2.80) | **−61.58** (3.06) |
| 7.85 (109.24) | **17.84** (37.59) | **36.40** (76.20) | **25.87** (48.99) | **30.76** (64.12) | **12.36** (77.71) | **62.24** (191.64) |
| **−88.54** (4.78) | **14.00**[a] (0.60) | −7.68 (0.41) | −4.48 (0.39) | −3.19 (0.45) | **17.10** (0.92) | **14.81**[a] (0.59) |

0.33/0.33/0.33 latent class proportion, ML/EM, MCMC-weak, and MCMC-wrong appeared to have the most difficulty properly recovering class proportions and growth parameters. The MCMC-informative condition performed the best, with high bias presented only in the within-class (co)variance parameters (which received diffuse priors). MCMC-data-driven performed well under higher class separation (e.g., MD = 2.0 and 1.5) with the exception of the parameter (co)variances.

| Very poor separation | | | | | | |
|---|---|---|---|---|---|---|
| ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| −6.48 (0.31) | 0.60 (0.33) | −1.11 (0.33) | −0.25 (0.33) | **17.45**[a] (0.39) | **56.31**[a] (0.52) | **22.52**[a] (0.40) |
| **24.85** (0.41) | 1.32 (0.33) | 4.62 (0.35) | 2.93 (0.34) | **18.09** (0.39) | **18.06** (0.39) | **19.65** (0.39) |
| **−15.34** (0.28) | 1.11 (0.33) | −0.48 (0.33) | 0.36 (0.33) | **−32.50** (0.22) | **−71.34** (0.09) | **−39.14** (0.20) |
| 2.45 (17.58) | −0.29 (0.36) | −0.64 (1.16) | −0.61 (0.35) | −1.89 (1.43) | −0.87 (3.62) | 0.23 (6.69) |
| −1.22 (5.26) | 0.95 (0.02) | 7.32 (0.35) | 6.83 (0.38) | −5.84 (0.29) | 5.17 (0.14) | 1.94 (0.55) |
| 3.27 (6.31) | 1.85 (0.81) | 3.57 (2.77) | 1.97 (0.86) | 1.47 (0.90) | −2.10 (5.38) | 2.51 (1.34) |
| **−25.21** (7.46) | −0.49 (0.03) | **−15.56** (0.70) | **−14.29** (0.67) | −9.51 (0.36) | −5.85 (0.33) | **−21.91** (1.51) |
| −6.62 (24.81) | −1.45 (0.56) | −2.80 (2.09) | −1.25 (0.42) | −5.52 (8.24) | −36.88[a] (383.76) | 3.17 (30.85) |
| **31.87** (7.93) | 0.45 (0.02) | **15.02** (0.57) | **13.20** (0.54) | **−21.65**[a] (0.72) | **−13.70**[a] (0.36) | 4.12 (3.64) |
| **45.54** (7.72) | **−61.88** (2.08) | **−67.14** (2.11) | **−74.93** (2.34) | **−91.45** (2.71) | **−96.45** (3.81) | **−73.26** (6.94) |
| **−19.40**[a] (76.33) | 7.33 (21.33) | 4.96 (21.75) | **12.64**[a] (23.84) | 6.53 (21.81) | −0.99 (52.67) | **29.68**[a] (48.43) |
| **−92.37** (4.98) | **13.56**[a] (0.60) | −7.76 (0.40) | −8.96 (0.39) | −4.17 (0.45) | **17.14**[a] (1.03) | **−12.62**[a] (0.19) |

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 6
*Parameter Estimate Bias: 0.33/0.33/0.33 Mixture Class Proportions, Quadratic Conditions, 150 Cases*

| | High separation | | | | | | |
|---|---|---|---|---|---|---|---|
| Parameter | ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| C1 proportion | 7.27 (0.36) | −0.25 (0.33) | 0.46 (0.33) | −1.54 (0.33) | **22.02** (0.41) | **31.78** (0.44) | **−26.40** (0.25) |
| C2 proportion | 5.52 (0.35) | 0.79 (0.34) | 1.48 (0.34) | 4.69 (0.35) | **10.61**[a] (0.37) | **68.00** (0.56) | **81.52**[a] (0.61) |
| C3 proportion | **−12.76** (0.29) | −0.51 (0.33) | −1.91 (0.33) | −3.13 (0.32) | **−32.61** (0.22) | **−99.74** (0.00) | **−55.09** (0.15) |
| C1 | | | | | | | |
| Intercept (*MSE*) | −0.91 (16.14) | −0.67 (0.45) | −2.33 (2.67) | −1.07 (0.57) | −2.83 (2.60) | −0.07 (0.00) | −1.23 (1.70) |
| Slope (*MSE*) | **−18.52** (17.77) | 0.74 (0.02) | **15.75**[a] (2.33) | 8.97 (2.95) | **−23.85**[a] (1.10) | −2.35 (0.01) | 0.14 (1.99) |
| C2 | | | | | | | |
| Intercept (*MSE*) | −0.49 (9.62) | 0.08 (0.33) | −0.05 (0.75) | −0.04 (0.26) | −2.61 (2.52) | **−11.09**[a] (22.52) | −1.63 (1.17) |
| Slope (*MSE*) | −9.08 (25.34) | −0.16 (0.03) | 5.04 (1.69) | −0.71 (2.63) | **−22.46**[a] (1.76) | **40.52**[a] (2.63) | **−14.94**[a] (0.76) |
| C3 | | | | | | | |
| Intercept (*MSE*) | −0.40 (14.44) | 0.65 (0.38) | 2.57 (2.36) | 1.24 (0.48) | −2.36 (5.65) | **−59.34**[a] (497.89) | −5.62 (14.63) |
| Slope (*MSE*) | **16.48**[a] (18.27) | −2.09 (0.02) | **−24.88**[a] (2.22) | −7.31 (2.11) | **−33.59** (1.39) | **18.86**[a] (0.32) | **−40.46** (10.85) |
| Quadratic (*MSE*) | **−35.74** (2.10) | −4.52 (0.02) | 3.02 (0.17) | **−10.31**[a] (0.25) | **−20.19**[a] (0.16) | **−99.61**[a] (0.80) | **−85.33**[a] (0.68) |
| Variances | | | | | | | |
| I with S (*MSE*) | **−163.18** (19.18) | **−142.77** (5.97) | **−286.24** (15.84) | **−246.35** (12.01) | **−320.22** (19.40) | **180.23** (4.68) | **−474.84** (35.58) |
| Intercept (*MSE*) | **32.43** (189.42) | **27.11** (57.27) | **53.50** (143.56) | **42.23** (94.42) | **49.71** (131.63) | **−22.71** (16.70) | **127.79** (609.78) |
| Slope (*MSE*) | **−26.55** (2.61) | **17.27**[a] (0.77) | **22.04**[a] (0.79) | **21.91**[a] (0.77) | **34.25** (1.24) | 0.62 (0.00) | **68.48**[a] (3.37) |

*Note.* Bold values indicate problematic bias levels greater than 10.00% for $n = 150$ cases. Numbers in parentheses under C1–C3 class proportions represent actual mixture class proportions. Numbers in parentheses for growth parameters and variances represent mean squared error (*MSE*) values. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data driven = Bayesian estimation using data-driven priors; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; I with S = correlation between intercept and slope; *MSE* = mean squared error: The variance of the estimates across the replications plus the square of the bias.
[a] Indicates that percentage bias decreased below 10.00% when sample sizes were increased to $n = 800$.

Tables 6–7 include bias results for the condition using 0.33/0.33/0.33 latent class proportions and a quadratic growth trajectory for C3. Results appear largely comparable to the linear conditions presented in Tables 4–5, with one notable difference. Namely, ML/EM appeared to have less difficulty recovering the latent class proportions with the addition of the quadratic trajectory for C3. In general, MCMC-wrong, MCMC-weak, and MCMC-data-driven all showed higher bias levels for the class proportions and the intercept and slope parameters. The other estimation conditions produced lower bias results in

Table 7
*Parameter Estimate Bias: 0.33/0.33/0.33 Mixture Class Proportions, Quadratic Conditions, 150 Cases*

| | Poor separation | | | | | | |
|---|---|---|---|---|---|---|---|
| Parameter | ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| C1 proportion | 2.73 (0.34) | 0.62 (0.33) | 0.55 (0.33) | −1.05 (0.33) | **17.92**[a] (0.39) | −5.12 (0.31) | **−21.05** (0.26) |
| C2 proportion | 7.71 (0.36) | 1.48 (0.33) | 5.39 (0.35) | 5.33 (0.35) | **13.22**[a] (0.37) | 6.74 (0.35) | **110.42**[a] (0.69) |
| C3 proportion | −7.41 (0.31) | 0.93 (0.33) | −2.91 (0.32) | −1.25 (0.33) | **−28.11** (0.24) | 1.41 (0.33) | **−86.34** (0.05) |
| C1 | | | | | | | |
| Intercept (*MSE*) | 0.32 (24.95) | −0.59 (0.43) | −1.77 (1.68) | −0.96 (0.52) | −2.40 (2.04) | −8.05 (19.14) | 4.69 (12.12) |
| Slope (*MSE*) | **−17.84** (18.49) | 0.64 (0.02) | **14.61** (1.49) | 8.67 (2.51) | **−24.94**[a] (1.06) | 3.79 (0.12) | **29.51**[a] (1.25) |
| C2 | | | | | | | |
| Intercept (*MSE*) | 1.72 (9.02) | 0.46 (0.30) | 1.23 (0.86) | 0.39 (0.21) | −0.54 (0.68) | −4.42 (10.36) | 1.11 (1.58) |
| Slope (*MSE*) | **−15.54** (22.46) | 0.31 (0.03) | 6.66 (1.66) | 1.10 (2.42) | **−20.12**[a] (1.35) | −4.26 (0.24) | −6.24 (0.08) |
| C3 | | | | | | | |
| Intercept (*MSE*) | −1.72 (13.35) | 0.22 (0.27) | 0.38 (1.21) | 0.75 (0.29) | −2.96 (7.56) | −2.07 (12.73) | **−18.13**[a] (78.10) |
| Slope (*MSE*) | **23.87** (18.83) | −2.24 (0.02) | **−23.96** (2.35) | **−14.58** (2.26) | **−34.39** (1.49) | 2.27 (0.27) | **−90.42**[a] (8.86) |
| Quadratic (*MSE*) | **−47.52** (2.25) | −6.50 (0.02) | 0.32 (0.21) | **−10.78**[a] (0.29) | **−25.68**[a] (0.21) | **−77.80** (0.55) | **−99.80** (0.81) |
| Variances | | | | | | | |
| I with S (*MSE*) | **−137.78** (15.88) | **−134.87** (5.47) | **−228.64** (10.12) | **−228.38** (10.54) | **−266.37** (13.43) | **−317.91** (17.39) | **−368.18** (20.98) |
| Intercept (*MSE*) | **16.33**[a] (115.58) | **20.33** (40.84) | **29.93** (65.46) | **31.41** (62.97) | **32.54** (69.34) | **55.55**[a] (130.78) | **37.87** (105.88) |
| Slope (*MSE*) | **−31.59** (2.90) | **17.70**[a] (0.81) | **22.63** (0.75) | **18.94** (0.69) | **27.55** (0.97) | **41.90**[a] (1.41) | **73.03** (2.19) |

*Note.* Bold values indicate problematic bias levels greater than 10.00% for $n = 150$ cases. Numbers in parentheses under C1–C3 class proportions represent actual mixture class proportions. Numbers in parentheses for growth parameters and variances represent mean squared error (*MSE*) values. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data driven = Bayesian estimation using data-driven priors; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; I with S = correlation between intercept and slope; *MSE* = mean squared error: The variance of the estimates across the replications plus the square of the bias.
[a] Indicates that percentage bias decreased below 10.00% when sample sizes were increased to $n = 800$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

| Moderate separation | | | | | | |
|---|---|---|---|---|---|---|
| ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| 1.80 (0.34) | −0.26 (0.33) | 0.23 (0.33) | −1.63 (0.33) | **19.70** (0.40) | **30.87**[a] (0.44) | **111.90** (0.71) |
| 9.92 (0.37) | 0.69 (0.34) | 4.45 (0.35) | 4.38 (0.35) | **12.68**[a] (0.38) | **52.41** (0.51) | **−26.22**[a] (0.25) |
| **−11.69**[a] (0.29) | −0.41 (0.33) | −4.65 (0.32) | −2.72 (0.32) | **−32.36** (0.23) | **−83.25** (0.06) | **−85.65** (0.05) |
| −0.03 (20.09) | −0.65 (0.43) | −2.13 (1.97) | −0.98 (0.52) | −2.57 (2.41) | −0.72 (0.94) | −6.74 (11.82) |
| **−20.28** (19.75) | 0.73 (0.02) | **17.34** (1.84) | 2.54 (2.33) | **−24.51**[a] (1.10) | **24.95**[a] (0.66) | **25.79**[a] (1.64) |
| 0.55 (9.18) | 0.20 (0.31) | 0.68 (0.68) | 0.04 (0.22) | −1.89 (1.56) | −4.81 (6.00) | −3.54 (3.17) |
| −9.43 (23.05) | 0.02 (0.04) | 1.46 (1.44) | 2.64 (2.10) | **−21.86**[a] (1.61) | 0.06 (0.60) | **−36.70**[a] (4.27) |
| −1.15 (14.09) | 0.52 (0.33) | 1.70 (1.65) | 1.07 (0.40) | −3.11 (10.30) | **−38.62**[a] (293.44) | **−14.05**[a] (37.83) |
| **23.45**[a] (18.21) | −2.07 (0.02) | **−28.07** (2.37) | −8.80 (2.27) | **−34.00** (1.48) | **−45.26** (3.97) | **−163.90** (35.19) |
| **−42.80** (2.22) | −5.94 (0.02) | 4.16 (0.17) | **−12.07**[a] (0.27) | **−26.29**[a] (0.20) | **−93.37** (0.71) | **−102.54**[a] (0.85) |
| **−145.14** (17.17) | **−143.33** (5.89) | **−272.26** (13.63) | **−245.92** (11.64) | **−304.32** (17.29) | **−334.32** (23.24) | **−568.37** (46.82) |
| **24.31** (147.66) | **24.33** (50.23) | **42.37** (90.77) | **39.27** (86.22) | **45.28** (116.93) | **51.33**[a] (208.15) | **114.74** (467.50) |
| **29.00** (2.79) | **19.09**[a] (0.81) | **24.44**[a] (1.00) | **21.82** (0.74) | **34.60** (1.22) | **35.08**[a] (0.76) | **65.02**[a] (2.06) |

comparison. Most of the latent class proportions were recovered properly when sample sizes increased to $n = 800$ under the quadratic conditions, with the exception of conditions under MCMC-weak. Likewise, trajectory recovery appeared to improve for all estimation conditions when the sample size was increased. However, there was still poor recovery for some growth parameters in all estimation conditions, with the exception of MCMC-informative.

Tables 8–9 present results from the condition with 0.70/0.20/0.10 class proportions where C3 had a linear trajectory shape.

| Very poor separation | | | | | | |
|---|---|---|---|---|---|---|
| ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| 2.37 (0.34) | 0.56 (0.33) | −0.89 (0.33) | −0.07 (0.33) | **15.91**[a] (0.38) | **54.21** (0.51) | **−53.04**[a] (0.15) |
| 6.79 (0.35) | 1.54 (0.34) | 5.09 (0.35) | 3.61 (0.34) | **15.64** (0.38) | **39.83** (0.46) | **61.56**[a] (0.53) |
| −6.13 (0.31) | 0.93 (0.33) | −1.17 (0.33) | −0.51 (0.33) | **−28.52** (0.24) | **−91.02** (0.03) | −5.49 (0.31) |
| 1.15 (17.75) | −0.48 (0.36) | −0.27 (0.87) | −0.76 (0.41) | −1.04 (1.07) | −4.06 (4.05) | **−22.98**[a] (168.36) |
| **−21.59** (20.76) | 0.80 (0.02) | 6.03 (1.40) | **11.85**[a] (1.91) | **−23.85** (1.08) | **12.54** (0.23) | **−50.15** (4.72) |
| 3.99 (9.61) | 1.87 (0.82) | 3.56 (2.67) | 1.96 (0.83) | 1.86 (1.23) | 0.29 (0.26) | 3.87 (2.91) |
| **−17.77** (21.38) | 0.13 (0.03) | 4.76 (1.35) | 0.76 (2.05) | **−18.03**[a] (1.18) | **11.58**[a] (0.31) | −9.62 (0.48) |
| −5.19 (19.97) | −1.32 (0.49) | −2.44 (1.95) | −1.12 (0.38) | −6.54 (25.21) | **−40.97**[a] (321.84) | −2.88 (3.84) |
| **32.54** (18.65) | −2.38 (0.02) | **−23.88**[a] (2.72) | **−12.70** (2.24) | **−38.86** (1.91) | **−55.48** (3.97) | **10.69**[a] (1.03) |
| **−64.17** (2.38) | −6.41 (0.02) | −8.21 (0.25) | **−20.40**[a] (0.31) | **−28.62** (0.22) | **−97.64**[a] (0.77) | **−87.52** (0.67) |
| **−75.55** (11.55) | **−85.42** (3.11) | **−171.56** (6.51) | **−184.60** (7.07) | **−193.47** (8.02) | **−383.90** (22.75) | **−232.63** (10.01) |
| −1.18 (72.92) | **13.23**[a] (25.95) | **13.05**[a] (29.82) | **20.99**[a] (36.73) | **12.04**[a] (31.15) | **40.59** (68.27) | **38.61** (69.66) |
| **−39.68** (3.27) | **18.23**[a] (0.76) | **19.61**[a] (0.71) | **20.69**[a] (0.75) | **25.81** (0.86) | **77.80** (2.95) | **31.64** (1.06) |

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 8
*Parameter Estimate Bias: 0.70/0.20/0.10 Mixture Class Proportions, Linear Conditions, 150 Cases*

| | High separation | | | | | | |
|---|---|---|---|---|---|---|---|
| Parameter | ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| C1 proportion | **−33.14** (0.47) | 0.37 (0.70) | **−50.16** (0.35) | **−32.71** (0.47) | **−37.52** (0.44) | **−21.37** (0.55) | **−14.20**[a] (0.60) |
| C2 proportion | **67.98** (0.34) | 0.36 (0.20) | **76.97** (0.35) | **70.52** (0.34) | **110.04** (0.42) | **88.16** (0.38) | **58.36** (0.32) |
| C3 proportion | **95.99** (0.20) | −3.33 (0.10) | **197.16** (0.30) | **87.95** (0.19) | **42.58** (0.14) | **−26.73**[a] (0.07) | **−17.34**[a] (0.08) |
| C1 | | | | | | | |
| Intercept (*MSE*) | 2.40 (10.32) | −0.30 (0.30) | 2.02 (1.48) | 2.00 (1.19) | −1.12 (1.36) | 1.20 (0.99) | 0.36 (0.78) |
| Slope (*MSE*) | 0.51 (3.55) | 0.83 (0.03) | −1.00 (0.28) | 0.84 (0.21) | −9.47 (0.26) | 1.49 (0.20) | 0.95 (0.19) |
| C2 | | | | | | | |
| Intercept (*MSE*) | 4.46 (12.56) | 0.38 (0.22) | 7.84 (11.75) | 3.98 (3.13) | 5.17 (7.18) | 1.42 (3.99) | 1.52 (1.05) |
| Slope (*MSE*) | **−27.48** (8.23) | −0.56 (0.01) | **−15.99** (0.72) | **−12.59**[a] (0.65) | **−14.27** (0.49) | −9.64 (0.37) | **−12.12**[a] (0.68) |
| C3 | | | | | | | |
| Intercept (*MSE*) | 0.84 (28.33) | **14.27**[a] (0.27) | **10.70**[a] (20.61) | 2.96 (1.60) | −4.34 (39.07) | **−18.52**[a] (136.21) | −4.03 (9.07) |
| Slope (*MSE*) | **13.62**[a] (6.07) | **32.59**[a] (0.01) | 8.53 (0.50) | **13.04**[a] (0.76) | **−36.33**[a] (1.78) | **−13.55**[a] (0.81) | −7.51 (0.71) |
| Variances | | | | | | | |
| I with S (*MSE*) | **19.24**[a] (6.57) | **−49.42** (2.10) | **−72.18** (2.57) | **−53.93**[a] (2.04) | **−116.73** (3.66) | **−96.70** (3.32) | **−99.40** (3.66) |
| Intercept (*MSE*) | −6.43 (88.04) | **16.16**[a] (33.30) | **27.20** (53.14) | 2.52 (22.11) | **16.53**[a] (33.63) | **11.85** (40.96) | **36.96** (76.55) |
| Slope (*MSE*) | **−79.03** (4.08) | **16.04**[a] (0.61) | −8.16 (0.39) | −2.22 (0.39) | −1.10 (0.46) | **24.40**[a] (0.93) | 9.02 (0.59) |

*Note.* Bold values indicate problematic bias levels greater than 10.00% for $n = 150$ cases. Numbers in parentheses under C1–C3 class proportions represent actual mixture class proportions. Numbers in parentheses for growth parameters and variances represent mean squared error (*MSE*) values. Dashes (—) represent estimates that were unreasonably out of bounds (e.g., >10,000) and therefore were not presented in the output. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data driven = Bayesian estimation using data-driven priors; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; I with S = correlation between intercept and slope; *MSE* = mean squared error: The variance of the estimates across the replications plus the square of the bias.
[a] Indicates that percentage bias decreased below 10.00% when sample sizes were increased to $n = 800$.

With the addition of a clear majority latent class (C1), the recovery of latent class proportions was very poor overall with the exception of MCMC-informative. The latent class proportion for C3 was consistently overestimated under ML/EM, MCMC-diffuse, MCMC-partial, and MCMC-wrong. The opposite effect occurred for the other estimation conditions in that C3 was consistently underestimated in the MCMC-weak and MCMC-data-driven conditions. Tables 8–9 also indicated that MCMC-informative produced the lowest bias levels for the growth trajectory parameters across all latent class separation

Table 9
*Parameter Estimate Bias: 0.70/0.20/0.10 Mixture Class Proportions, Linear Conditions, 150 Cases*

| | Poor separation | | | | | | |
|---|---|---|---|---|---|---|---|
| Parameter | ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| C1 proportion | **−39.77** (0.42) | 0.10 (0.70) | **−51.31** (0.34) | **−35.74** (0.45) | **−41.89** (0.41) | **−13.05** (0.61) | **−14.80** (0.60) |
| C2 proportion | **74.06** (0.35) | 0.38 (0.20) | **73.50** (0.35) | **60.28** (0.32) | **96.38** (0.39) | **65.47** (0.33) | **57.96** (0.32) |
| C3 proportion | **130.31** (0.23) | −1.50 (0.10) | **212.17** (0.31) | **129.64** (0.23) | **100.44** (0.20) | **−39.55** (0.06) | **−12.33**[a] (0.09) |
| C1 | | | | | | | |
| Intercept (*MSE*) | 3.18 (12.67) | −0.26 (0.30) | 2.48 (2.07) | 2.17 (1.33) | −1.07 (1.45) | 0.50 (1.45) | 0.62 (0.75) |
| Slope (*MSE*) | 1.10 (3.97) | 0.81 (0.03) | 0.31 (0.31) | 1.78 (0.23) | −9.44 (0.26) | 1.76 (0.29) | 3.61 (0.19) |
| C2 | | | | | | | |
| Intercept (*MSE*) | 5.90 (13.19) | 0.64 (0.23) | 8.06 (12.37) | 4.00 (3.12) | 4.99 (8.57) | −2.73 (29.13) | 2.72 (1.81) |
| Slope (*MSE*) | **−29.31** (8.37) | −0.60 (0.01) | **−17.45** (0.85) | **−14.85** (0.78) | **−14.14** (0.50) | **−14.05** (0.66) | **−17.80** (0.87) |
| C3 | | | | | | | |
| Intercept (*MSE*) | −0.14 (25.73) | 0.06 (0.20) | 6.84 (8.66) | 2.88 (1.58) | 0.50 (25.78) | **−33.24** (262.23) | **−10.23**[a] (39.49) |
| Slope (*MSE*) | — (—) | −0.02 (0.01) | **10.24** (0.51) | **15.11** (0.71) | **−29.11**[a] (1.21) | **−20.08** (0.77) | **−12.25** (1.76) |
| Variances | | | | | | | |
| I with S (*MSE*) | **28.04**[a] (6.66) | **−49.18** (2.04) | **−74.17** (2.36) | **−60.96**[a] (2.16) | **−115.37** (3.49) | **−75.67** (2.64) | **−123.56** (4.41) |
| Intercept (*MSE*) | **−10.55**[a] (81.52) | **13.45**[a] (28.80) | **14.87** (29.49) | 3.27 (20.51) | **11.70** (25.28) | 9.99 (31.72) | **30.22** (50.67) |
| Slope (*MSE*) | **−89.70** (4.76) | **15.83**[a] (0.64) | −8.77 (0.40) | −5.88 (0.40) | −3.93 (0.43) | **12.48** (0.70) | 9.01 (0.52) |

*Note.* Bold values indicate problematic bias levels greater than 10.00% for $n = 150$ cases. Numbers in parentheses under C1–C3 class proportions represent actual mixture class proportions. Numbers in parentheses for growth parameters and variances represent mean squared error (*MSE*) values. Dashes (—) represent estimates that were unreasonably out of bounds (e.g., >10,000) and therefore were not presented in the output. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data driven = Bayesian estimation using data-driven priors; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; I with S = correlation between intercept and slope; *MSE* = mean squared error: The variance of the estimates across the replications plus the square of the bias.
[a] Indicates that percentage bias decreased below 10.00% when sample sizes were increased to $n = 800$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

| Moderate separation | | | | | | |
|---|---|---|---|---|---|---|
| ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| **−35.63** (0.45) | 0.25 (0.70) | **−50.95** (0.34) | **−33.92** (0.46) | **−34.04** (0.46) | **−19.51** (0.56) | **−17.29**[a] (0.58) |
| **71.76** (0.34) | 0.43 (0.20) | **75.58** (0.35) | **66.28** (0.33) | **97.48** (0.39) | **70.79** (0.34) | **65.75** (0.33) |
| **105.87** (0.21) | −2.63 (0.10) | **205.49** (0.31) | **104.90** (0.20) | **43.35** (0.14) | −5.00 (0.10) | **−10.45**[a] (0.09) |
| 2.78 (11.66) | −0.29 (0.31) | 2.30 (1.78) | 2.09 (1.26) | 1.19 (0.83) | 0.57 (1.11) | 0.66 (1.06) |
| 1.72 (3.39) | 0.82 (0.03) | −0.25 (0.30) | 1.20 (0.21) | −6.62 (0.25) | 4.52 (0.12) | 0.83 (0.24) |
| 5.00 (12.21) | 0.46 (0.21) | 8.02 (12.24) | 3.97 (3.10) | 4.10 (4.05) | −0.59 (30.61) | 2.15 (1.34) |
| **−26.64** (8.45) | −0.61 (0.01) | **−17.39** (0.81) | **−13.37** (0.69) | **−14.78** (0.54) | **−29.24** (4.90) | **−13.10**[a] (0.66) |
| 0.22 (28.29) | 0.31 (0.25) | 9.12 (14.22) | 2.99 (1.66) | −5.26 (39.28) | **−17.13**[a] (112.87) | −5.11 (13.06) |
| — (—) | 0.07 (0.01) | **10.37**[a] (0.51) | **14.24** (0.75) | **−35.76**[a] (1.72) | −9.39 (0.71) | −8.78 (0.91) |
| **18.63**[a] (6.42) | **−47.23** (2.05) | **−70.95** (2.32) | **−55.22**[a] (2.05) | **−100.30** (3.16) | **−107.32** (3.40) | **−107.80** (3.75) |
| −8.87 (82.73) | **14.82**[a] (30.93) | **21.47** (40.58) | 2.45 (21.39) | **10.70**[a] (27.57) | **22.56** (34.15) | **33.13** (59.91) |
| −0.83 (4.37) | **15.63**[a] (0.64) | −8.92 (0.40) | −3.60 (0.39) | −0.93 (0.47) | −4.39 (0.22) | **11.14**[a] (0.55) |

conditions. Note, however, that all seven estimation conditions were able to properly recover the trajectory for the majority latent class C1. For all estimation conditions except for MCMC-informative, the slope for C2 was poorly recovered, and C3 showed patterns of problematic bias levels for the intercept and slope. Further, these bias levels under C3 increased as class separation decreased, especially for ML/EM. Increasing the sample size to $n = 800$ only decreased bias levels for select parameters, but several parameters exhibited high bias levels even with this higher sample size.

| Very poor separation | | | | | | |
|---|---|---|---|---|---|---|
| ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| **−47.23** (0.37) | −0.19 (0.70) | **−52.43** (0.33) | **−41.33** (0.41) | **−40.19** (0.42) | **−12.79** (0.61) | **−24.05** (0.53) |
| **84.99** (0.37) | 0.51 (0.20) | **72.71** (0.35) | **57.30** (0.31) | **90.16** (0.38) | **65.81** (0.33) | **104.53** (0.41) |
| **160.64** (0.26) | 0.28 (0.10) | **221.58** (0.32) | **174.70** (0.27) | **101.00** (0.20) | **−42.08** (0.06) | **−40.70** (0.06) |
| 4.49 (17.59) | −0.30 (0.29) | 3.00 (2.76) | 2.07 (1.23) | 1.83 (1.28) | 0.71 (2.00) | 1.81 (2.12) |
| 4.76 (4.28) | 0.70 (0.03) | 1.22 (0.28) | 3.47 (0.26) | −7.23 (0.26) | 0.83 (0.19) | 3.76 (0.24) |
| 7.38 (14.79) | 1.91 (0.77) | 8.55 (13.85) | 4.73 (4.25) | 5.74 (6.59) | −3.52 (58.67) | 4.49 (3.98) |
| **−32.93** (8.66) | −0.4 (0.01) | **−19.81** (0.95) | **−16.60** (0.83) | **−14.26** (0.51) | **−17.89** (1.03) | **−20.97** (1.08) |
| −3.43 (23.59) | −1.50 (0.58) | 1.59 (1.32) | 1.22 (0.42) | −3.97 (23.93) | **−36.80**[a] (332.54) | **−26.92**[a] (192.06) |
| **242.80** (—) | 0.07 (0.01) | **12.49** (0.50) | **13.19** (0.53) | **−26.35**[a] (1.05) | **−20.53**[a] (1.29) | **−25.22** (2.87) |
| **27.74**[a] (6.67) | **−55.53** (2.03) | **−77.80** (2.23) | **−82.38** (2.43) | **−105.82** (3.08) | **−87.26** (3.52) | **−102.13** (3.59) |
| **−18.92**[a] (76.12) | **12.89**[a] (25.41) | 4.74 (20.23) | 7.33 (20.71) | 1.82 (19.36) | 7.93 (36.21) | **14.11**[a] (35.61) |
| **−94.88** (5.11) | **15.48**[a] (0.61) | −8.66 (0.39) | −9.30 (0.40) | −3.64 (0.44) | **15.03**[a] (0.52) | 4.27 (0.39) |

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 10
*Parameter Estimate Bias: 0.70/0.20/0.10 Mixture Class Proportions, Quadratic Conditions, 150 Cases*

| | High separation | | | | | | |
|---|---|---|---|---|---|---|---|
| Parameter | ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| C1 proportion | **−35.23** (0.45) | 0.72 (0.71) | **−47.95** (0.36) | **−32.63** (0.47) | **−30.14** (0.49) | −6.48 (0.65) | **−11.34** (0.62) |
| C2 proportion | **83.42** (0.37) | −0.66 (0.20) | **87.57** (0.38) | **72.73** (0.35) | **90.92** (0.38) | **55.04** (0.31) | **67.64** (0.34) |
| C3 proportion | **79.77** (0.18) | −3.75 (0.10) | **160.49** (0.26) | **82.93** (0.18) | **29.15** (0.13) | **−64.70**[a] (0.04) | **−55.87**[a] (0.04) |
| C1 | | | | | | | |
| Intercept (*MSE*) | 2.50 (11.43) | −0.36 (0.31) | 2.18 (1.78) | 1.88 (1.12) | 1.46 (1.04) | −0.10 (2.35) | 0.37 (1.37) |
| Slope (*MSE*) | **−26.92** (15.13) | 0.63 (0.05) | 5.16 (1.14) | 1.53 (1.59) | **−28.34**[a] (1.33) | 6.73 (0.30) | 9.34 (0.69) |
| C2 | | | | | | | |
| Intercept (*MSE*) | 4.89 (13.03) | 0.44 (0.23) | 7.77 (11.67) | 3.75 (2.86) | 3.54 (4.46) | **−10.20**[a] (80.75) | 0.72 (1.23) |
| Slope (*MSE*) | **−11.79** (21.51) | −0.32 (0.02) | **−12.23** (1.69) | −4.71 (2.58) | **−30.13** (2.30) | **−23.19**[a] (1.75) | **−19.92** (1.29) |
| C3 | | | | | | | |
| Intercept (*MSE*) | 0.78 (25.51) | 0.28 (0.26) | 9.40 (16.18) | 2.92 (1.60) | **−13.61**[a] (111.05) | **−38.26**[a] (293.72) | **−10.50**[a] (27.84) |
| Slope (*MSE*) | **18.83** (19.12) | −0.98 (0.01) | **−34.39**[a] (3.36) | 3.12 (3.95) | **−53.59** (3.52) | **−52.22**[a] (3.85) | **−75.79** (11.90) |
| Quadratic (*MSE*) | **−33.83** (2.24) | −3.47 (0.02) | **−10.99** (0.23) | **−34.76** (0.58) | **−56.14** (0.42) | **−75.28** (0.60) | **−100.38** (0.82) |
| Variances | | | | | | | |
| I with S (*MSE*) | **−40.40**[a] (9.45) | **−88.02** (3.33) | **−172.18** (6.84) | **−133.80** (4.69) | **−196.50** (8.70) | **−267.78** (15.02) | **−320.26** (18.00) |
| Intercept (*MSE*) | 5.16 (87.21) | **19.54**[a] (37.25) | **21.12**[a] (46.14) | 8.11 (28.32) | **19.56**[a] (49.77) | **24.89**[a] (72.33) | **66.07** (174.01) |
| Slope (*MSE*) | **−43.75** (2.76) | **11.58**[a] (0.58) | 9.26 (0.50) | **11.15**[a] (0.51) | **18.66**[a] (0.68) | **18.44**[a] (0.66) | **34.27**[a] (1.11) |

*Note.* Bold values indicate problematic bias levels greater than 10.00% for $n = 150$ cases. Numbers in parentheses under C1–C3 class proportions represent actual mixture class proportions. Numbers in parentheses for growth parameters and variances represent mean squared error (*MSE*) values. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data driven = Bayesian estimation using data-driven priors; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; I with S = correlation between intercept and slope; *MSE* = mean squared error: The variance of the estimates across the replications plus the square of the bias.
[a] Indicates that percentage bias decreased below 10.00% when sample sizes were increased to $n = 800$.

Tables 10–11 present results for the 0.70/0.20/0.10 latent class proportions but with a quadratic trajectory shape for C3. Latent class proportion recovery is relatively comparable to that presented in Tables 8–9 with linear trajectory shapes. All of the estimation conditions with the exception of MCMC-informative were unable to properly recover latent class proportions under any level of class separation. The one main exception to this statement was with MCMC-weak, where the class proportion

Table 11
*Parameter Estimate Bias: 0.70/0.20/0.10 Mixture Class Proportions, Quadratic Conditions, 150 Cases*

| | Poor separation | | | | | | |
|---|---|---|---|---|---|---|---|
| Parameter | ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| C1 proportion | **−40.14** (0.42) | 0.55 (0.70) | **−49.20** (0.36) | **−35.74** (0.45) | **−34.48** (0.46) | −4.38 (0.67) | **−14.55**[a] (0.60) |
| C2 proportion | **83.63** (0.37) | −0.35 (0.20) | **84.53** (0.37) | **65.84** (0.33) | **85.73** (0.37) | **58.16** (0.32) | **85.79** (0.37) |
| C3 proportion | **113.74** (0.21) | −3.16 (0.10) | **175.36** (0.28) | **118.52** (0.22) | **69.90** (0.17) | **−85.63** (0.01) | **−69.68**[a] (0.03) |
| C1 | | | | | | | |
| Intercept (*MSE*) | 2.85 (12.56) | −0.33 (0.30) | 2.47 (2.08) | 2.01 (1.22) | 1.93 (1.42) | −0.23 (2.75) | 0.27 (0.77) |
| Slope (*MSE*) | **−29.97** (17.40) | −0.05 (0.06) | 0.71 (1.65) | 1.92 (1.59) | **−23.63**[a] (1.06) | **15.91**[a] (0.30) | 5.28 (0.85) |
| C2 | | | | | | | |
| Intercept (*MSE*) | 6.40 (14.38) | 0.71 (0.27) | 7.77 (11.41) | 3.91 (3.03) | 4.70 (5.22) | **−13.59**[a] (109.42) | 2.21 (1.47) |
| Slope (*MSE*) | **−17.14** (21.33) | −0.30 (0.02) | −6.04 (0.96) | −2.10 (2.45) | **−28.66** (1.97) | **−29.55**[a] (2.97) | **−11.44** (1.30) |
| C3 | | | | | | | |
| Intercept (*MSE*) | −0.12 (22.10) | −0.01 (0.19) | 6.04 (7.88) | 2.84 (1.62) | −6.80 (72.01) | **−49.22** (446.27) | **−18.23**[a] (64.31) |
| Slope (*MSE*) | **26.89** (21.19) | −1.14 (0.01) | **−35.71**[a] (3.32) | −8.20 (3.45) | **−47.17** (2.67) | **−52.73**[a] (3.83) | **−136.98** (23.71) |
| Quadratic (*MSE*) | **−52.46** (2.54) | −5.34 (0.02) | **−12.20** (0.22) | **−35.83** (0.52) | **−53.83** (0.40) | **−98.94** (0.79) | **−98.29** (0.78) |
| Variances | | | | | | | |
| I with S (*MSE*) | **−56.69**[a] (9.78) | **−87.68** (3.21) | **−152.68** (5.51) | **−146.49** (5.04) | **−186.58** (7.32) | **−258.00** (12.97) | **−343.57** (21.34) |
| Intercept (*MSE*) | 2.21 (76.30) | **17.02**[a] (32.00) | **13.53**[a] (32.88) | 9.23 (27.31) | 9.19 (30.44) | **18.17**[a] (57.49) | **47.25** (85.76) |
| Slope (*MSE*) | **−45.47** (2.95) | **12.50**[a] (0.61) | 7.01 (0.44) | **10.44**[a] (0.50) | **18.31**[a] (0.66) | **24.03**[a] (1.03) | **14.63**[a] (0.77) |

*Note.* Bold values indicate problematic bias levels greater than 10.00% for $n = 150$ cases. Numbers in parentheses under C1–C3 class proportions represent actual mixture class proportions. Numbers in parentheses for growth parameters and variances represent mean squared error (*MSE*) values. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data driven = Bayesian estimation using data-driven priors; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; I with S = correlation between intercept and slope; *MSE* = mean squared error: The variance of the estimates across the replications plus the square of the bias.
[a] Indicates that percentage bias decreased below 10.00% when sample sizes were increased to $n = 800$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

| Moderate separation | | | | | | |
|---|---|---|---|---|---|---|
| ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| **−37.28** (0.44) | 0.73 (0.71) | **−48.40** (0.36) | **−33.98** (0.46) | **−33.06** (0.47) | −2.43 (0.68) | −8.34 (0.64) |
| **81.99** (0.36) | −0.61 (0.20) | **82.40** (0.36) | **69.77** (0.34) | **91.16** (0.38) | **50.53** (0.30) | **57.77** (0.32) |
| **96.96** (0.20) | −3.87 (0.10) | **174.04** (0.27) | **98.35** (0.20) | **49.08** (0.15) | **−84.05** (0.02) | **−57.16** (0.04) |
| 2.59 (12.09) | −0.34 (0.31) | 2.29 (2.00) | 2.01 (1.21) | 1.73 (1.34) | 1.02 (0.51) | 0.23 (0.93) |
| **−23.34** (16.45) | 0.58 (0.05) | 7.57 (1.35) | 3.69 (1.59) | **−27.68**[a] (1.28) | **−15.85**[a] (0.28) | **13.75**[a] (0.96) |
| 5.57 (13.54) | 0.50 (0.23) | 7.40 (10.70) | 3.82 (2.92) | 4.33 (4.64) | −6.36 (17.05) | 0.55 (2.47) |
| **−17.45** (21.90) | −0.42 (0.02) | −8.11 (2.02) | −4.73 (2.50) | **−27.76** (2.04) | **−10.43**[a] (0.79) | **−23.91** (3.33) |
| 0.98 (23.70) | 0.23 (0.23) | 8.62 (13.15) | 3.04 (1.75) | −9.17 (82.54) | **−50.04**[a] (421.73) | **−12.44**[a] (34.98) |
| **20.74** (20.06) | −1.11 (0.01) | **−32.70** (3.20) | −1.62 (3.69) | **−50.95** (3.11) | **−27.87**[a] (2.22) | **−98.85** (16.29) |
| **−43.56** (2.45) | −4.31 (0.02) | **−16.17** (0.24) | **−35.57** (0.54) | **−51.91** (0.40) | **−100.73** (0.82) | **−98.02** (0.78) |
| **−52.24**[a] (9.66) | **−85.80** (3.29) | **−171.43** (6.64) | **−136.75** (4.86) | **−191.34** (7.93) | **−49.10** (4.51) | **−347.46** (19.31) |
| 5.68 (85.08) | **18.37**[a] (35.19) | **18.73**[a] (40.07) | 7.60 (29.98) | **13.71**[a] (38.67) | −3.55 (80.48) | **50.91** (99.21) |
| **−44.51** (2.82) | **10.99**[a] (0.57) | 8.81 (0.56) | 9.10 (0.48) | **19.69**[a] (0.70) | **−12.49**[a] (0.12) | **47.93**[a] (1.79) |

for C1 (the majority class) was properly recovered under high and moderate class separation. Again, ML/EM, MCMC-diffuse, MCMC-partial, and MCMC-wrong consistently overestimated the class size for the minority class C3. In contrast, MCMC-weak and MCMC-data-driven consistently underestimated the size of C3. With respect to the intercept and slope parameters, MCMC-wrong, MCMC-weak, and MCMC-data-driven produced the highest bias levels. Results indicated that when a

| Very poor separation | | | | | | |
|---|---|---|---|---|---|---|
| ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| **−43.79** (0.39) | 0.41 (0.70) | **−50.99** (0.34) | **−40.29** (0.42) | **−38.28** (0.43) | **17.37** (0.82) | **−20.57**[a] (0.56) |
| **76.92** (0.35) | −0.27 (0.20) | **77.69** (0.36) | **63.88** (0.33) | **83.42** (0.37) | **−11.07** (0.18) | **80.65** (0.36) |
| **152.72** (0.25) | −2.35 (0.10) | **201.58** (0.30) | **154.28** (0.25) | **101.15** (0.20) | **−99.48** (0.00) | **−17.29** (0.08) |
| 3.81 (16.18) | −0.41 (0.30) | 3.18 (3.19) | 1.93 (1.16) | 2.46 (1.93) | −1.07 (2.21) | 2.46 (5.44) |
| **−32.27** (19.03) | 0.61 (0.05) | −0.03 (1.34) | 8.10 (1.52) | **−24.71** (1.01) | 3.80 (0.20) | 9.35 (0.54) |
| 7.61 (16.57) | 2.05 (0.87) | 8.22 (13.00) | 4.71 (4.24) | 6.09 (7.53) | **−23.85**[a] (184.70) | 5.02 (4.77) |
| **−15.35** (21.60) | −0.28 (0.02) | −3.47 (1.73) | −0.58 (2.19) | **−28.63** (1.95) | **−11.84**[a] (2.51) | **−22.78** (1.38) |
| −2.96 (19.35) | −1.37 (0.44) | 1.77 (1.72) | 1.22 (0.47) | −5.48 (41.95) | **−59.83** (657.87) | **−12.83**[a] (48.41) |
| **32.92** (20.04) | −1.20 (0.01) | **−42.55** (3.29) | **−27.96** (3.80) | **−46.03** (2.42) | **−35.59**[a] (1.44) | **−55.09** (7.41) |
| **−74.91** (2.69) | −5.06 (0.02) | **−20.07** (0.18) | **−26.29** (0.41) | **−48.94** (0.36) | **−99.57** (0.80) | **−99.08** (0.80) |
| **−43.48** (8.95) | **−75.29** (2.85) | **−143.13** (5.08) | **−139.62** (4.93) | **−143.93** (4.80) | **−142.30** (4.95) | **−214.13** (7.41) |
| −6.33 (67.08) | **17.57** (30.20) | 6.81 (23.40) | **11.04**[a] (27.93) | 0.61 (20.74) | −4.59 (17.17) | **22.09** (22.24) |
| **−47.81** (3.21) | **11.02** (0.61) | 9.81 (0.47) | 8.37 (0.54) | **17.93**[a] (0.69) | **48.83**[a] (2.98) | **23.04**[a] (0.74) |

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 12
*Parameter Estimate Bias: 0.45/0.45/0.10 Mixture Class Proportions, Linear Conditions, 150 Cases*

| | High separation | | | | | | |
|---|---|---|---|---|---|---|---|
| Parameter | ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| C1 proportion | **−12.34**[a] (0.39) | −0.11 (0.45) | **−24.45** (0.34) | −9.93 (0.41) | −1.66 (0.44) | **−16.82**[a] (0.37) | 6.02 (0.48) |
| C2 proportion | **−16.11**[a] (0.38) | 0.22 (0.45) | **−22.63** (0.35) | **−20.86** (0.36) | −9.51 (0.41) | **13.89** (0.51) | −9.80 (0.41) |
| C3 proportion | **128.03** (0.23) | −0.51 (0.10) | **211.86** (0.31) | **138.54** (0.24) | **50.25** (0.15) | **13.16** (0.11) | **17.01** (0.12) |
| C1 | | | | | | | |
| Intercept (*MSE*) | 1.33 (14.62) | −0.45 (0.41) | −0.32 (0.74) | 0.42 (0.31) | −1.32 (0.98) | 1.16 (3.80) | −0.70 (1.32) |
| Slope (*MSE*) | −0.39 (4.57) | 1.36 (0.02) | 6.32 (0.33) | 6.73 (0.27) | −0.93 (0.23) | 0.31 (0.51) | 4.37 (0.39) |
| C2 | | | | | | | |
| Intercept (*MSE*) | 2.13 (7.40) | 0.41 (0.36) | 4.37 (4.02) | 2.26 (1.15) | 0.97 (1.06) | 0.95 (1.66) | 0.96 (0.82) |
| Slope (*MSE*) | **−21.06**[a] (7.48) | −0.70 (0.03) | −9.45 (0.50) | −7.10 (0.46) | −7.11 (0.31) | −3.27 (0.29) | −8.05 (0.51) |
| C3 | | | | | | | |
| Intercept (*MSE*) | 1.16 (26.39) | 0.26 (0.21) | 8.79 (12.23) | 2.99 (1.55) | −4.04 (27.11) | **−18.41**[a] (129.76) | −2.92 (11.88) |
| Slope (*MSE*) | **25.34** (7.39) | 0.02 (0.01) | **16.65** (0.72) | **20.22** (0.92) | **−31.73**[a] (1.45) | −6.41 (0.62) | −2.71 (0.72) |
| Variances | | | | | | | |
| I with S (*MSE*) | **26.50** (7.68) | **−55.83** (2.14) | **−79.03** (2.67) | **−59.81** (2.32) | **−114.58** (3.69) | **−64.98** (3.35) | **−121.93** (4.73) |
| Intercept (*MSE*) | −1.71 (81.19) | **14.46**[a] (32.79) | **28.09** (53.07) | 4.08 (21.72) | **17.99** (34.34) | **13.84** (44.95) | **36.02** (75.86) |
| Slope (*MSE*) | **−84.79** (4.55) | **15.27**[a] (0.61) | −7.58 (0.41) | −1.92 (0.41) | −2.36 (0.48) | **22.42**[a] (0.85) | **10.00**[a] (0.58) |

*Note.* Bold values indicate problematic bias levels greater than 10.00% for $n = 150$ cases. Numbers in parentheses under C1–C3 class proportions represent actual mixture class proportions. Numbers in parentheses for growth parameters and variances represent mean squared error (*MSE*) values. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data driven = Bayesian estimation using data-driven priors; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; I with S = correlation between intercept and slope; *MSE* = mean squared error: The variance of the estimates across the replications plus the square of the bias.
[a] Indicates that percentage bias decreased below 10.00% when sample sizes were increased to $n = 800$.

clear majority latent class was present, recovery appeared to be best when an informative prior was used. By increasing the total sample size to $n = 800$, some of the high bias levels decreased in the trajectory parameters. However, high bias in the latent class proportions tended to remain even with the increased sample size.

Finally, results for the 0.45/0.45/0.10 latent class proportions are presented in Tables 12–13 for the linear trajectory condi-

Table 13
*Parameter Estimate Bias: 0.45/0.45/0.10 Mixture Class Proportions, Linear Conditions, 150 Cases*

| | Poor separation | | | | | | |
|---|---|---|---|---|---|---|---|
| Parameter | ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| C1 proportion | **−22.98** (0.35) | −0.43 (0.45) | **−26.33** (0.33) | **−14.80** (0.38) | −7.87 (0.41) | **−10.49** (0.40) | −1.84 (0.44) |
| C2 proportion | **−13.51**[a] (0.39) | 0.21 (0.45) | **−23.45** (0.34) | **−24.85** (0.34) | **−11.06** (0.40) | 6.23 (0.48) | −0.42 (0.45) |
| C3 proportion | **164.24** (0.26) | 0.99 (0.10) | **224.01** (0.32) | **178.44** (0.28) | **85.21** (0.19) | **19.17** (0.12) | **10.16** (0.11) |
| C1 | | | | | | | |
| Intercept (*MSE*) | 2.68 (18.26) | −0.46 (0.39) | 0.23 (0.88) | 0.50 (0.33) | −0.91 (0.85) | 0.38 (2.09) | 0.26 (1.93) |
| Slope (*MSE*) | −2.22 (4.96) | 1.35 (0.02) | 6.68 (0.35) | 8.53 (0.32) | −2.20 (0.24) | 6.58 (0.45) | 5.18 (0.69) |
| C2 | | | | | | | |
| Intercept (*MSE*) | 3.47 (7.49) | 0.60 (0.35) | 4.68 (4.52) | 2.31 (1.16) | 1.91 (1.31) | 1.66 (1.76) | 2.12 (1.29) |
| Slope (*MSE*) | **−22.19** (7.55) | −0.69 (0.03) | **−10.71**[a] (0.52) | −9.30 (0.54) | −6.26 (0.25) | −5.58 (0.31) | −9.82 (0.38) |
| C3 | | | | | | | |
| Intercept (*MSE*) | −1.09 (22.64) | −0.09 (0.15) | 4.25 (3.68) | 2.08 (0.89) | −3.16 (18.12) | **−21.82**[a] (164.63) | −9.32 (37.74) |
| Slope (*MSE*) | **35.76** (8.17) | −0.02 (0.01) | **18.69** (0.65) | **20.62** (0.81) | **−24.98**[a] (1.02) | **−12.46**[a] (0.83) | −3.71 (1.39) |
| Variances | | | | | | | |
| I with S (*MSE*) | **34.18**[a] (7.50) | **−55.65** (2.12) | **−82.58** (2.55) | **−77.98** (2.44) | **−117.10** (3.58) | **−101.85** (3.66) | **−123.85** (4.75) |
| Intercept (*MSE*) | **−10.10**[a] (82.04) | **12.83**[a] (27.90) | **16.84** (31.80 | 7.13 (22.23) | **12.01** (26.78) | **18.04** (44.05) | **25.79** (51.09) |
| Slope (*MSE*) | **−90.05** (4.89) | **14.68**[a] (0.60) | −7.00 (0.38) | −5.10 (0.41) | −2.08 (0.46) | 9.86 (0.58) | **12.96**[a] (0.54) |

*Note.* Bold values indicate problematic bias levels greater than 10.00% for $n = 150$ cases. Numbers in parentheses under C1–C3 class proportions represent actual mixture class proportions. Numbers in parentheses for growth parameters and variances represent mean squared error (*MSE*) values. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data driven = Bayesian estimation using data-driven priors; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; I with S = correlation between intercept and slope; *MSE* = mean squared error: The variance of the estimates across the replications plus the square of the bias.
[a] Indicates that percentage bias decreased below 10.00% when sample sizes were increased to $n = 800$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

| Moderate separation | | | | | | |
|---|---|---|---|---|---|---|
| ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| **−17.75** (0.37) | −0.16 (0.45) | **−25.34** (0.34) | **−11.94** (0.40) | −4.65 (0.43) | 2.44 (0.46) | **13.25**[a] (0.51) |
| **−14.20**[a] (0.39) | 0.17 (0.45) | **−23.17** (0.35) | **−22.98** (0.35) | **−10.16** (0.40) | −4.94 (0.43) | **−16.67** (0.37) |
| **143.81** (0.24) | −0.04 (0.10) | **218.29** (0.32) | **157.17** (0.26) | **66.65** (0.17) | **11.26** (0.11) | **15.15** (0.12) |
| 2.02 (16.55) | −0.43 (0.38) | −0.08 (0.78) | 0.46 (0.34) | −1.13 (0.94) | −0.73 (2.20) | −1.19 (1.50) |
| −0.36 (4.93) | 1.40 (0.02) | 7.71 (0.32) | 7.08 (0.28) | −1.10 (0.23) | 3.26 (0.17) | 7.58 (0.29) |
| 2.95 (7.34) | 0.43 (0.35) | 4.53 (4.32) | 2.26 (1.14) | 1.42 (1.09) | 0.39 (3.14) | 0.99 (0.66) |
| **−21.20** (7.73) | −0.65 (0.03) | **−10.51**[a] (0.58) | −8.01 (0.47) | −6.78 (0.27) | 3.05 (0.33) | −8.59 (0.34) |
| 0.14 (24.26) | 0.18 (0.18) | 6.77 (7.91) | 2.72 (1.34) | −3.13 (20.22) | **−15.22**[a] (114.88) | −5.13 (14.06) |
| **27.53**[a] (7.42) | −0.07 (0.01) | **17.52** (0.69) | **21.15** (0.91) | **−28.50**[a] (1.23) | −8.45 (0.49) | −2.29 (0.98) |
| **28.10**[a] (7.56) | **−53.26** (2.13) | **−82.46** (2.62) | **−66.79** (2.36) | **−117.18** (3.68) | **−110.88** (3.36) | **−112.21** (4.22) |
| −5.05 (81.00) | **12.99**[a] (30.32) | **22.22** (40.32) | 5.73 (21.52) | **15.50** (31.38) | **17.44** (44.90) | **31.49** (66.08) |
| **−87.73** (4.76) | **14.93**[a] (0.61) | −7.00 (0.40) | −3.51 (0.40) | −1.25 (0.46) | **21.10**[a] (0.47) | 8.83 (0.57) |

tions and in Tables 14–15 for the quadratic conditions. Tables 12–13 indicate that there was a problem in recovering latent class proportions since each class exhibited high bias under six of the seven estimation conditions, with the exception being MCMC-informative where latent class proportions were well recovered. Note that MCMC-weak and MCMC-data-driven both unexpectedly showed lower bias levels under the very poor class separation level (MD = 0.5), however, these

| Very poor separation | | | | | | |
|---|---|---|---|---|---|---|
| ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| **−29.93** (0.32) | −1.07 (0.45) | **−27.32** (0.33) | **−21.08** (0.36) | **−12.51** (0.39) | −6.27 (0.42) | −0.90 (0.45) |
| −9.13 (0.41) | 0.53 (0.45) | **−23.09** (0.35) | **−25.99** (0.33) | **−12.58** (0.39) | 4.20 (0.47) | 1.66 (0.46) |
| **175.76** (0.28) | 2.40 (0.10) | **226.85** (0.33) | **211.82** (0.31) | **112.93** (0.21) | 9.32 (0.11) | −3.42 (0.10) |
| 3.70 (19.53) | −0.68 (0.42) | 0.70 (1.01) | 0.24 (0.27) | −0.46 (0.71) | 0.07 (1.88) | 0.52 (3.79) |
| −0.25 (5.37) | 1.16 (0.02) | 7.86 (0.40) | 9.54 (0.37) | −3.34 (0.24) | 7.49 (0.18) | **11.82**[a] (1.11) |
| 4.80 (8.61) | 1.51 (0.60) | 5.22 (5.49) | 2.93 (1.71) | 2.88 (2.08) | −0.19 (20.79) | 2.94 (2.16) |
| **−21.58** (6.97) | −0.78 (0.03) | **−12.59** (0.56) | **−10.17** (0.51) | −5.93 (0.25) | −6.64 (0.23) | **−19.09** (0.91) |
| −5.82 (25.32) | −1.75 (0.65) | −1.50 (3.12) | −0.34 (0.16) | −5.02 (13.86) | **−28.95**[a] (250.83) | **−22.34** (144.48) |
| **38.54** (7.90) | 0.08 (0.00) | **21.52** (0.78) | **18.50** (0.68) | **−19.68**[a] (0.77) | **−17.94**[a] (0.79) | **−24.39** (3.61) |
| **28.06**[a] (7.41) | **−66.35** (2.28) | **−96.41** (2.79) | **−101.71** (3.03) | **−118.68** (3.58) | **−142.31** (4.36) | **−165.35** (6.73) |
| **−18.90**[a] (77.52) | **15.07**[a] (28.01) | 7.40 (21.49) | **13.98**[a] (25.30) | 6.96 (22.86) | **15.29** (42.20) | **18.51** (34.22) |
| **−90.84** (4.95) | **14.84**[a] (0.64) | −7.07 (0.41) | −8.84 (0.41) | −3.90 (0.44) | **22.07**[a] (0.85) | 8.32 (0.58) |

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 14
*Parameter Estimate Bias: 0.45/0.45/0.10 Mixture Class Proportions, Quadratic Conditions, 150 Cases*

| | High separation | | | | | | |
|---|---|---|---|---|---|---|---|
| Parameter | ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| C1 proportion | **−13.84**[a] (0.39) | 0.14 (0.45) | **−22.59** (0.35) | −9.41 (0.41) | −1.08 (0.45) | **−10.52**[a] (0.40) | 6.33 (0.48) |
| C2 proportion | **−12.65**[a] (0.39) | 0.28 (0.45) | **−20.92** (0.36) | **−19.66** (0.36) | **−13.46** (0.39) | **−18.13** (0.37) | 2.47 (0.46) |
| C3 proportion | **119.19** (0.22) | −1.88 (0.10) | **195.80** (0.30) | **130.83** (0.23) | **65.44** (0.17) | **128.93** (0.23) | **−39.60**[a] (0.06) |
| C1 | | | | | | | |
| Intercept (*MSE*) | 0.69 (13.42) | −0.60 (0.40) | 0.09 (0.83) | 0.24 (0.34) | −0.77 (0.79) | −5.80 (10.96) | −0.27 (1.52) |
| Slope (*MSE*) | **−18.52** (18.26) | 0.67 (0.03) | 9.79 (2.50) | 7.66 (1.89) | **−24.21**[a] (1.15) | 4.90 (0.12) | **30.35** (1.17) |
| C2 | | | | | | | |
| Intercept (*MSE*) | 2.97 (8.71) | 0.43 (0.33) | 4.25 (3.98) | 2.09 (1.07) | 1.27 (1.47) | −1.78 (8.84) | −0.27 (0.98) |
| Slope (*MSE*) | **−11.08**[a] (20.23) | −0.31 (0.05) | −2.46 (2.20) | 1.24 (2.51) | **−22.70**[a] (1.59) | −6.31 (0.22) | **−25.16**[a] (1.67) |
| C3: | | | | | | | |
| Intercept (*MSE*) | 2.50 (21.97) | 0.34 (0.21) | 8.32 (11.71) | 3.16 (1.72) | −7.73 (60.37) | −1.38 (13.02) | **−12.87**[a] (54.44) |
| Slope (*MSE*) | **20.54** (20.90) | −0.99 (0.01) | **−22.75** (2.45) | 2.94 (3.26) | **−43.67** (2.52) | −0.42 (0.23) | **−67.16** (7.29) |
| Quadratic (*MSE*) | **−46.10** (2.60) | −3.39 (0.02) | **−27.12** (0.26) | **−42.31** (0.53) | **−58.78** (0.43) | **−92.67** (0.71) | **−99.17** (0.80) |
| Variances | | | | | | | |
| I with S (*MSE*) | **−77.74** (11.37) | **104.81** (4.04) | **−198.79** (7.51) | **−152.02** (5.53) | **−218.52** (9.73) | **−245.11** (10.93) | **−258.03** (12.92) |
| Intercept (*MSE*) | **14.81** (98.59) | **20.75** (40.73) | **26.54** (51.93) | **13.38**[a] (36.33) | **23.80** (56.89) | **48.16**[a] (104.68) | **35.37** (105.13) |
| Slope (*MSE*) | **−46.49** (2.94) | **11.99**[a] (0.60) | **10.41**[a] (0.45) | 9.60 (0.53) | **19.34**[a] (0.73) | **23.11**[a] (0.81) | **15.15**[a] (0.42) |

*Note.* Bold values indicate problematic bias levels greater than 10.00% for $n = 150$ cases. Numbers in parentheses under C1–C3 class proportions represent actual mixture class proportions. Numbers in parentheses for growth parameters and variances represent mean squared error (*MSE*) values. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data driven = Bayesian estimation using data-driven priors; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; I with S = correlation between intercept and slope; *MSE* = mean squared error: The variance of the estimates across the replications plus the square of the bias.
[a] Indicates that percentage bias decreased below 10.00% when sample sizes were increased to $n = 800$.

results should be interpreted along with the fact that convergence rates were very low for these cells. Overall, the highest bias by far was found in the C3 proportion, which represents the minority latent class. ML/EM, MCMC-diffuse, and MCMC-partial exhibited the most difficulty in accurately recovering the trajectory shapes. Not surprisingly, MCMC-informative was still largely able to properly recover latent class trajectories.

Table 15
*Parameter Estimate Bias: 0.45/0.45/0.10 Mixture Class Proportions, Quadratic Conditions, and 150 Cases*

| | Poor separation | | | | | | |
|---|---|---|---|---|---|---|---|
| Parameter | ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| C1 Proportion | **−18.68** (0.37) | −0.13 (0.45) | **−23.60** (0.34) | **−14.30** (0.39) | −7.46 (0.42) | **29.33**[a] (0.58) | **−67.26**[a] (0.15) |
| C2 Proportion | **−16.64**[a] (0.38) | 0.30 (0.45) | **−20.97** (0.36) | **−22.96** (0.35) | **−13.65** (0.39) | **−12.59**[a] (0.39) | **79.75**[a] (0.81) |
| C3 Proportion | **158.94** (0.26) | −0.75 (0.10) | **200.56** (0.30) | **167.65** (0.27) | **95.00** (0.20) | **−75.34**[a] (0.02) | **−56.23**[a] (0.04) |
| C1 | | | | | | | |
| Intercept (*MSE*) | 1.37 (14.20) | −0.60 (0.38) | 0.10 (0.96) | 0.34 (0.33) | −0.46 (0.74) | −0.58 (38.81) | 9.52 (20.89) |
| Slope (*MSE*) | **−25.27** (18.36) | 0.39 (0.03) | **14.25** (1.72) | **10.03** (2.21) | **−23.69**[a] (1.06) | **18.96** (3.12) | **88.69** (7.08) |
| C2 | | | | | | | |
| Intercept (*MSE*) | 4.09 (10.06) | 0.60 (0.33) | 4.44 (4.20) | 2.17 (1.09) | 2.28 (1.85) | **−54.48**[a] (65.30) | 2.82 (1.46) |
| Slope (*MSE*) | **−12.47** (19.67) | −0.58 (0.05) | −3.64 (2.08) | 2.90 (2.47) | **−20.46**[a] (1.50) | **−64.51**[a] (1.81) | **−18.53** (0.55) |
| C3 | | | | | | | |
| Intercept (*MSE*) | 0.18 (17.22) | −0.01 (0.12) | 4.39 (4.12) | 2.33 (40.72) | −5.11 (48.88) | **−99.99** (439.44) | **−10.36**[a] (16.99) |
| Slope (*MSE*) | **29.48** (21.38) | −1.00 (0.01) | **−29.61** (2.28) | −7.54 (2.77) | **−43.44** (2.43) | **−157.06** (2.88) | **−61.10**[a] (3.36) |
| Quadratic (*MSE*) | **−68.23** (2.71) | −3.89 (0.02) | **−19.01** (0.19) | **−38.32** (0.56) | **−52.48** (0.38) | − (0.80) | **−87.47** (0.62) |
| Variances | | | | | | | |
| I with S (*MSE*) | **−78.01** (10.67) | **−97.36** (3.86) | **−162.58** (5.27) | **−158.17** (5.60) | **−186.19** (7.27) | **113.65** (9.89) | **−286.73** (11.84) |
| Intercept (*MSE*) | 6.94 (84.69) | **17.96**[a] (34.20) | **17.81** (43.23) | **15.17**[a] (34.20) | **15.41** (36.68) | 1.44 (24.98) | **26.26** (22.33) |
| Slope (*MSE*) | **−48.35** (3.16) | **11.58**[a] (0.61) | 6.84 (0.39) | 8.36 (0.45) | **14.40**[a] (0.55) | **813.00** (1.25) | 21.42 (0.18) |

*Note.* Bold values indicate problematic bias levels greater than 10.00% for $n = 150$ cases. Numbers in parentheses under C1–C3 class proportions represent actual mixture class proportions. Numbers in parentheses for growth parameters and variances represent mean squared error (*MSE*) values. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Default diffuse = Bayesian estimation using default Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; Data driven = Bayesian estimation using data-driven priors; C1 = Latent Class 1; C2 = Latent Class 2; C3 = Latent Class 3; I with S = correlation between intercept and slope; *MSE* = mean squared error: The variance of the estimates across the replications plus the square of the bias.
[a] Indicates that percentage bias decreased below 10.00% when sample sizes were increased to $n = 800$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

| Moderate separation | | | | | | |
|---|---|---|---|---|---|---|
| ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| **−16.86** (0.37) | 0.07 (0.45) | **−21.98** (0.35) | **−12.25**[a] (0.39) | −4.28 (0.43) | **−13.39** (0.39) | 4.20 (0.47) |
| **−13.92** (0.39) | 0.27 (0.45) | **−19.68** (0.36) | **−21.14** (0.35) | **−15.60** (0.38) | **27.99**[a] (0.58) | 1.40 (0.46) |
| **138.53** (0.24) | −1.51 (0.10) | **187.48** (0.29) | **150.27** (0.25) | **89.46** (0.19) | **−65.70** (0.03) | **−25.19**[a] (0.07) |
| 1.21 (15.15) | −0.61 (0.38) | −0.19 (0.96) | 0.34 (0.32) | −0.53 (0.78) | 1.80 (1.83) | 1.23 (1.63) |
| **−26.31** (19.00) | 0.60 (0.03) | **13.50** (1.67) | 8.85 (1.91) | **−21.79**[a] (1.01) | **20.87**[a] (0.47) | **24.88**[a] (0.97) |
| 3.68 (9.54) | 0.48 (0.33) | 4.10 (3.67) | 2.09 (1.04) | 1.82 (1.44) | −1.45 (1.20) | −0.39 (1.08) |
| **−13.59** (20.21) | −0.48 (0.05) | −5.28 (1.81) | 0.23 (2.42) | **−23.42**[a] (1.82) | −9.82 (0.30) | **−18.94**[a] (1.45) |
| 2.04 (19.31) | 0.29 (0.17) | 6.19 (7.44) | 2.90 (1.51) | −2.62 (27.05) | **−47.28** (413.92) | **−17.66**[a] (106.56) |
| **26.83** (21.39) | −0.96 (0.01) | **−25.53** (1.96) | 1.34 (2.98) | **−39.94** (2.05) | **−47.66** (3.60) | **−106.37** (20.03) |
| **−56.69** (2.80) | −4.48 (0.02) | **−18.12** (0.17) | **−43.53** (0.51) | **−52.37** (0.39) | **−100.63** (0.82) | **−87.57** (0.68) |
| **−83.07** (11.15) | **−103.90** (4.16) | **−184.54** (7.07) | **−149.44** (5.06) | **−213.27** (9.01) | **−291.88** (17.93) | **−369.68** (24.93) |
| **12.15** (94.07) | **19.69** (37.00) | **25.20** (50.98) | **11.73**[a] (32.96) | **20.76** (47.68) | 8.18 (65.67) | **33.47** (136.04) |
| **−45.47** (3.09) | **12.09**[a] (0.62) | 8.44 (0.47) | 8.24 (0.49) | **15.83**[a] (0.69) | **41.81** (0.79) | **39.52**[a] (1.15) |

Results were relatively comparable for the quadratic trajectory conditions presented in Tables 14–15. Specifically, MCMC-informative was the only estimation level that did not exhibit problematic bias in latent class proportions. MCMC-wrong was able to properly recover the majority class C1 under the two higher class separation levels. ML/EM, MCMC-weak, and MCMC-data-driven all still appeared to have the most difficulty in recovering the trajectory shapes, regardless of class separation. MCMC-

| Very poor separation | | | | | | |
|---|---|---|---|---|---|---|
| ML/EM | Info | Default diffuse | Partial | Wrong | Weak | Data driven |
| **−22.52** (0.35) | −0.80 (0.45) | **−26.57** (0.33) | **−20.05** (0.36) | **−11.48** (0.40) | **34.57**[a] (0.61) | **37.69**[a] (0.62) |
| **−17.92** (0.37) | 0.82 (0.45) | **−19.90** (0.36) | **−23.71** (0.34) | **−16.70** (0.37) | **−15.14**[a] (0.38) | **−35.66** (0.29) |
| **181.97** (0.28) | −0.06 (0.10) | **209.10** (0.31) | **196.91** (0.30) | **126.81** (0.23) | **−87.47** (0.01) | −9.12 (0.09) |
| 2.26 (16.23) | −0.82 (0.44) | 0.57 (1.13) | 0.18 (0.22) | 0.06 (0.52) | 0.03 (3.72) | −1.80 (1.89) |
| **−28.57** (19.76) | 0.22 (0.03) | **14.95** (1.72) | **14.53** (1.76) | **−24.50**[a] (1.09) | **33.03** (1.18) | **44.93** (2.22) |
| 5.45 (12.94) | 1.61 (0.65) | 5.15 (5.47) | 3.04 (1.84) | 3.30 (2.52) | **−17.57**[a] (151.03) | 1.53 (1.03) |
| **−15.96** (21.18) | −0.71 (0.05) | −1.06 (1.14) | 0.84 (2.03) | **−22.20**[a] (1.49) | **−26.97**[a] (2.27) | **−41.85** (4.55) |
| −4.36 (18.93) | −1.62 (0.56) | −1.31 (1.19) | −0.12 (0.15) | −4.23 (7.39) | **−58.55** (673.72) | **−16.13**[a] (87.16) |
| **38.60** (19.76) | −1.06 (0.01) | **−25.13** (2.20) | **−16.09** (2.84) | **−38.33** (1.72) | **−53.73**[a] (3.55) | **−107.14** (30.03) |
| **−80.51** (2.62) | −3.64 (0.02) | **−27.91** (0.26) | **−39.57** (0.45) | **−47.50** (0.31) | **−100.29** (0.81) | **−100.62** (0.82) |
| **−40.76** (9.22) | **−80.61** (3.19) | **−122.93** (4.20) | **−144.83** (4.79) | **−146.05** (4.92) | **−237.71** (10.98) | **−320.01** (17.17) |
| −3.58 (69.31) | **19.38**[a] (33.44) | 2.62 (23.89) | **17.53**[a] (35.07) | 3.45 (22.85) | −4.15 (35.43) | **22.27** (23.80) |
| **−50.76** (3.38) | 8.76 (0.58) | 9.90 (0.48) | 6.32 (0.49) | 7.14 (0.45) | **29.14**[a] (0.90) | −0.49 (0.29) |

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

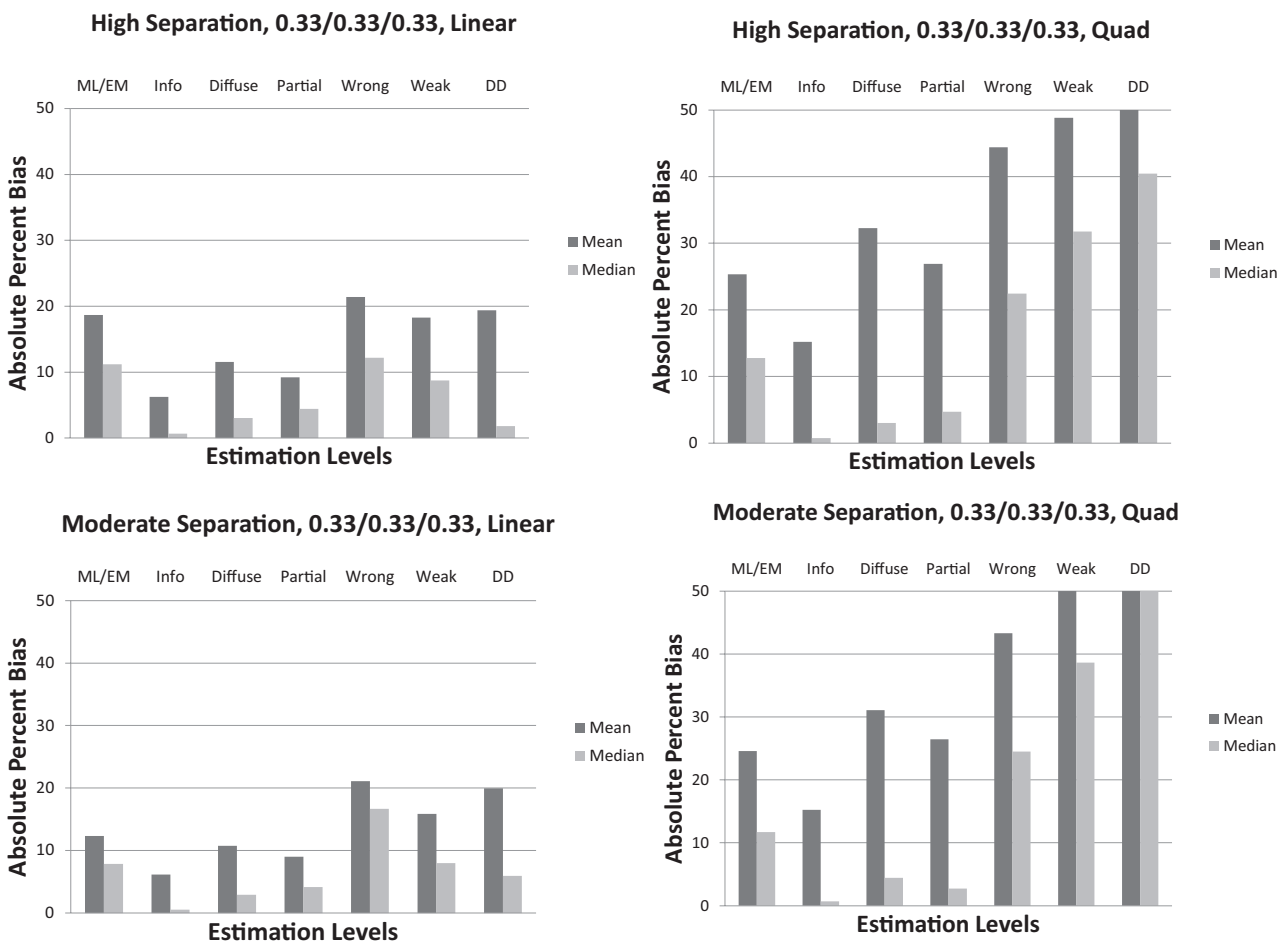


*Figure 6..* Absolute overall bias levels for 0.33/0.33/0.33 latent class proportions, high and moderate separation. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Diffuse = Bayesian estimation using Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; DD = data-driven priors; Quad = quadratic.

diffuse, MCMC-partial, and MCMC-wrong also showed some problematic bias levels in latent class proportions and trajectory shapes; however, the prevalence and degree of bias was still less than some of the other estimation conditions. With respect to total sample size, it is important to note that many of the problematic bias levels remained even under the larger sample size (i.e., $n = 800$).

## Overall Bias Levels for Estimation Conditions

Tables 4–15 detailed information about the recovery of individual parameters within each estimation level. However, it is also of interest to examine how well the estimation levels compare to one another overall. Figures 6, 7, 8, 9, 10, and 11 present a visual depiction of the overall performance of each estimation level with respect to combined bias rates across all parameters in the model. The overall mean and median of the absolute-value bias percent rates are illustrated here for each estimation level under $n = 150$ cases; similar plots for $n = 800$ can be made available upon request.[14]

Findings from Figures 6–11 indicate that there is better overall performance when a linear trajectory is present versus a quadratic trajectory. Under the linear trajectory conditions, ML/EM and MCMC-diffuse produced the higher overall absolute bias levels when there was a true minority or majority latent class. When class sizes were equivalent, MCMC-weak showed the highest overall absolute bias levels under the linear trajectory conditions. When a quadratic trajectory was introduced, only MCMC-informative produced overall bias levels that were within a reasonable range; the other conditions all yielded much higher overall bias rates.

[14] Absolute values of percentage bias were used here to compute overall mean and median bias rates. This was done to prevent two parameters with bias rates of 100% and −100% from canceling each other out in a misleading manner by averaging as if there was no bias present in that estimation level at all. As a result, the direction of bias is not depicted in the figures; however, the figures present the magnitude of overall bias for each estimation level. The mean and median are both reported here as to depict the patterns of outlier bias influence, where the mean bias level would be skewed upward when one or more parameters had extreme bias levels compared to the other parameters in the model.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

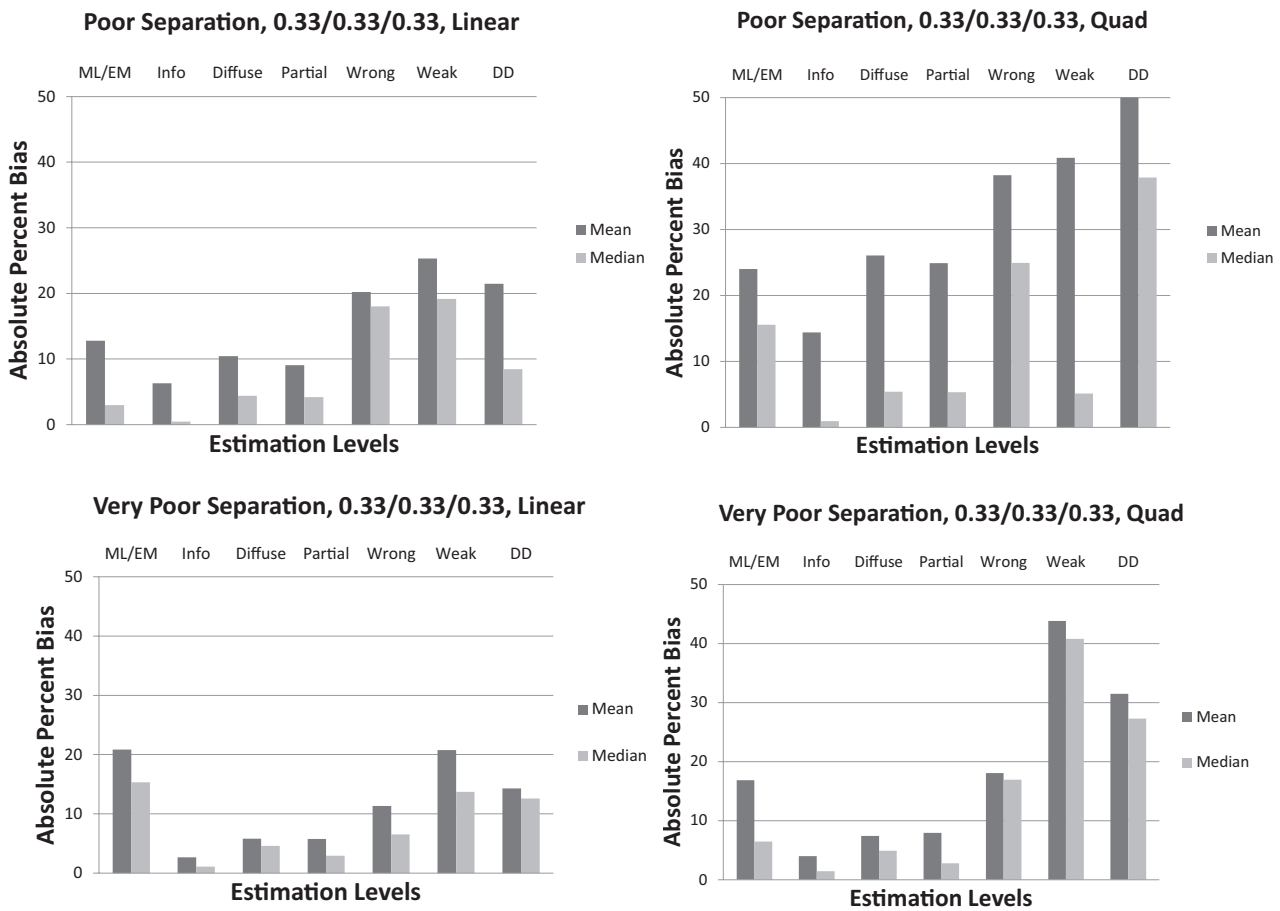


*Figure 7.*. Absolute overall bias levels for 0.33/0.33/0.33 latent class proportions, poor and very poor separation. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Diffuse = Bayesian estimation using Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; DD = data-driven priors; Quad = quadratic.

## Mean Squared Error

As mentioned, Tables 4–15 also include MSE values for each of the parameter estimates. There are some notable patterns of MSE that will be highlighted here. Overall, ML/EM, MCMC-weak, and MCMC-data-driven produced the highest MSE values across all cells of the study. In contrast, MCMC-informative produced noticeably lower MSE value across the different cells of the study. One additional pattern worthy of mention is that MSE values were typically larger for the variance parameters compared to the growth trajectory parameters across all estimation conditions; this was especially the case for the variance of the intercept.

## Concluding Remarks

The goal of this study was to assess the ability for growth trajectories and latent class proportions to be properly recovered under various estimation conditions. It was of particular interest to examine the recovery of trajectories and latent class proportions under the four levels of latent class separation, as well as different relative class proportions across classes.

## Discussion and Summary of Main Findings

MCMC-informative performed quite well overall in that it was largely able to uncover small but substantively different trajectories. As expected, the trajectory shapes were accurately recovered across conditions when informative priors were placed on the growth parameters. Latent class proportions were also well recovered with the use of informative Dirichlet priors on the class proportions. Overall, the parameter recovery results indicated that this estimation level was optimal for identifying small but real latent classes. These results are not too surprising since the use of informative priors is akin to combining data from the current study with the data from the previous studies that are driving the informative prior (Gelman, Carlin, Stern, & Rubin, 2004). This combination of information from previous and current data effectively increases the sample size for the

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

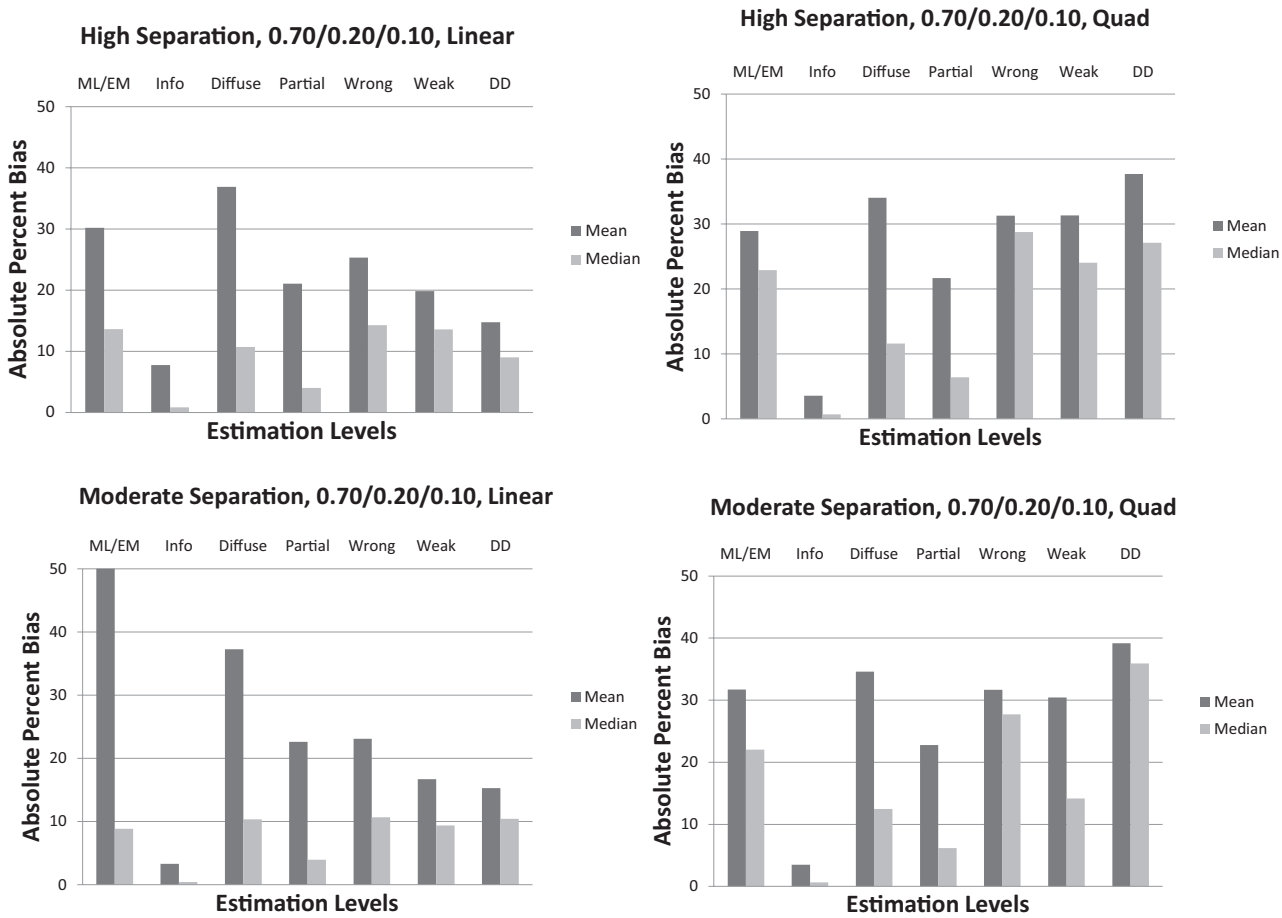


*Figure 8..* Absolute overall bias levels for 0.70/0.20/0.10 latent class proportions, high and moderate separation. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Diffuse = Bayesian estimation using Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; DD = data-driven priors; Quad = quadratic.

analysis. This theoretical increase in sample size increases power, hence, producing more accurate estimates that exhibit less bias. Overall, these findings speak to the upper bound performance of Bayesian GMM when "accurate" informative priors are used.

Even in conditions where the mean hyperparameter was not specified as the true growth parameter value, the Bayesian estimation method resulted in reasonably good performance under some higher class separation levels. In general, this performance would depend on how different the prior mean is from the true value and how much variability is allowed in the prior. This finding indicates that there may be a range of "acceptable" levels of informative priors that can still produce accurate recovery of trajectories. In order to examine this topic further and assess the degree to which (in)accurate priors can affect estimates, results from a sensitivity-type analysis for the accuracy of priors would potentially provide insight to the impact of priors on trajectory recovery.

The poor performance of MCMC-diffuse is in line with Richardson and Green (1997). Specifically, MCMC-diffuse was specified as a prior of $\mathcal{N}$ (0, $10^{10}$) for normally distributed parameters. With such a large variance component, the priors were acting as almost improper noninformative priors. Richardson and Green (1997) discuss the notion that the use of fully noninformative priors will not lead to obtaining proper posterior distributions in the univariate mixture modeling context (see also: Diebolt & Robert, 1994; Roeder & Wasserman, 1997). As a result, it is not surprising to find such high bias levels in the MCMC-diffuse conditions in the current study. Richardson and Green (1997) suggest the use of weakly informative priors as a more appropriate solution. The findings of the current investigation suggest that more informative priors are necessary in the context of mixture modeling.

## Limitations of the Study Design and Conclusions

It should be noted here that there are several limitations to the design and conclusions that can be drawn from the current study. First, the informative and diffuse (as well as ML/EM) conditions are not necessarily directly comparable conceptually speaking. Given that

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

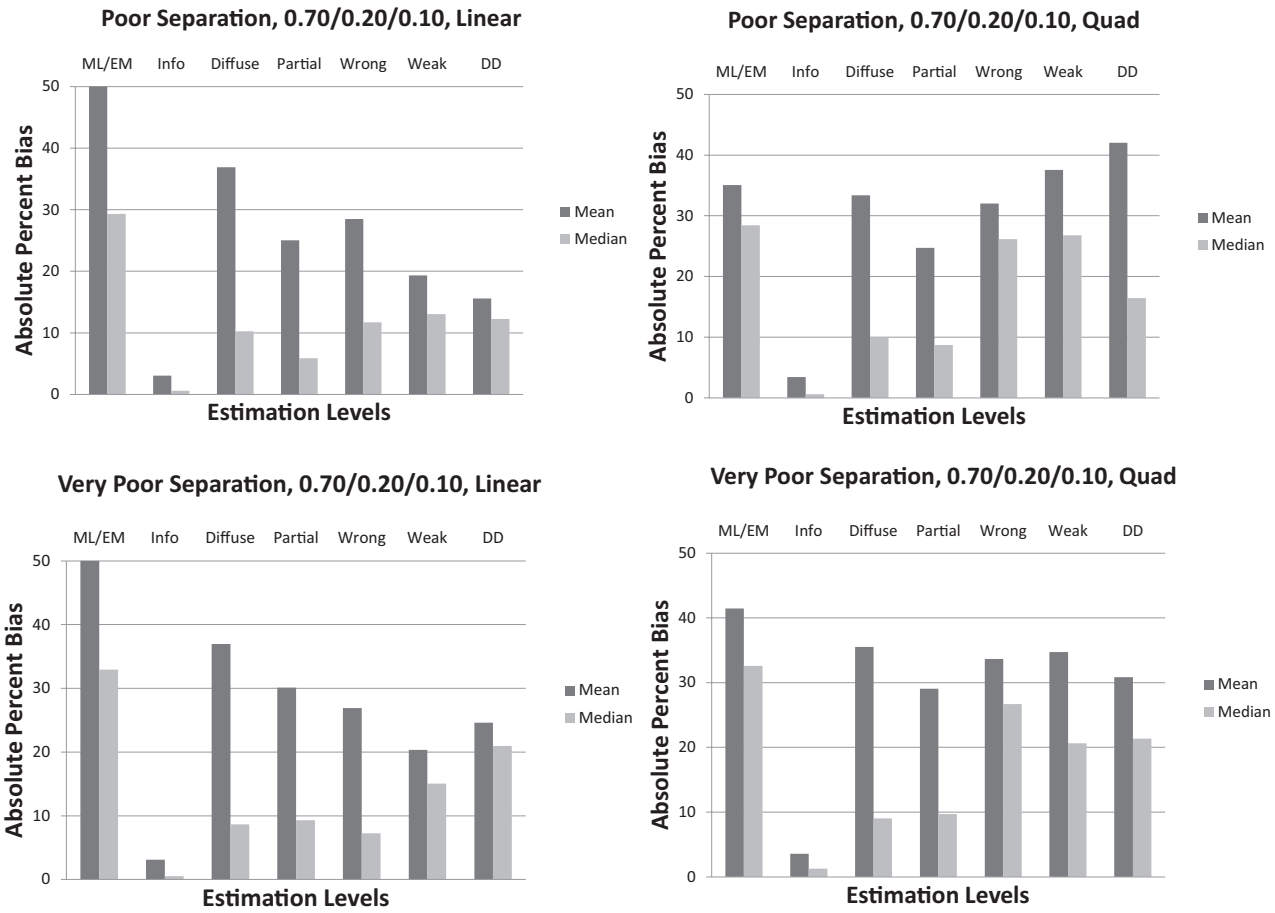


*Figure 9..* Absolute overall bias levels for 0.70/0.20/0.10 latent class proportions, poor and very poor separation. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Diffuse = Bayesian estimation using Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; DD = data-driven priors; Quad = quadratic.

the informative prior conditions are more "confirmatory" in nature (as discussed above), the results are conceptually different compared to the more exploratory conditions that did not implement theory-based knowledge into the model. The difference in conceptualizing these methods may also in part help inform why results between informative and diffuse conditions (and ML/EM) were not comparable. In fact, these two extremes may be representing different types of modeling, with MCMC-diffuse being more exploratory and MCMC-informative being more "confirmatory" in nature. Despite the conceptual differences between the different forms of priors, the main focus of this study was to assess the impact of different forms of priors on the performance of GMM.

In addition, some of the conditions within this fully crossed design might not reflect conditions found in the existing mixture class literature. Specifically, within the 0.70/0.20/0.10 proportion level, a sample size of 150 cases would yield only 15 cases in the smallest class. This particular condition was included to help provide a full picture of parameter recovery in relation to the other conditions included here. Although some empirical mixture model investigations have used comparable sample sizes (see e.g., Chung, Maisto, Cornelius, & Martin, 2004; Szapocznik et al., 2004), working with latent class sample sizes this small may create doubt for the substantive existence of such a small latent class and also potentially increase problems related to empirical underidentification in some cases (Kenny, 1979).

In order to prevent the size of the current simulation study from becoming unwieldy, several potentially interesting conditions were not examined in this investigation. For example, the MCMC-partial estimation level only specified informative priors on a single subset of parameters. A different result may be obtained if, for example, the informative priors were specified for the slopes instead of the intercepts (i.e., there was a priori knowledge about growth rates and not initial performance levels). Likewise, it would also be interesting to assess the performance of Bayesian GMM using different levels of informativeness. Particularly, this study implemented variance hyperparameters for growth trajectory parameters that were 5% of the population value to assess the upper bound of Bayesian GMM performance. However, it may also be interesting to conduct a sensitivity

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

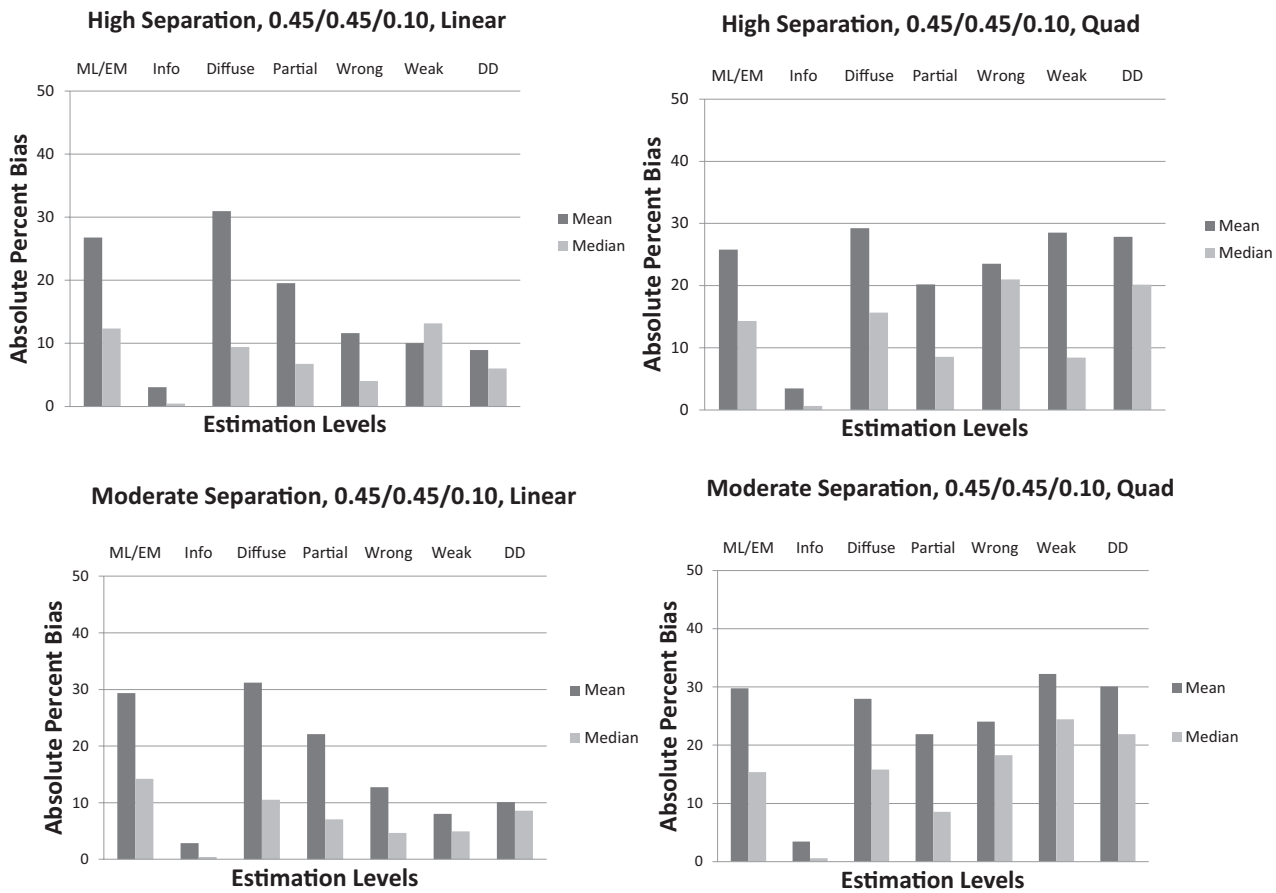


*Figure 10..* Absolute overall bias levels for 0.45/0.45/0.10 latent class proportions, high and moderate separation. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Diffuse = Bayesian estimation using Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; DD = data-driven priors; Quad = quadratic.

analysis where various levels of hyperparameters are examined (e.g., 10%, 20% of the population value) to assess at what point the accuracy of the estimates declines as the variance hyperparameter increases. Depaoli (in press) looked at a similar issue where hyperparameter values for "inaccurate" priors were examined in a similar sensitivity-type analysis; however, it would also be of interest to asses this issue in terms of latent class separation and "accurate" priors.

## Defining Priors

Finally, the elicitation and defining process of priors is a topic that has experienced a prolific presence in the Bayesian literature (see e.g., Kass & Wasserman, 1996; O'Hagan, 1998). There are arguably many different ways of eliciting and defining priors, and each of these methods likely carry different degrees of accuracy/effectiveness within the estimation process. The current investigation employed two main methods for defining priors in the simulations. The first method used the population value to inform the hyperparameter values. Although this method is likely not indicative of an applied research setting, it was included to examine an upper bound to the performance of Bayesian GMM. The second method used a data-driven prior in that estimates from ML/EM were used to inform hyperparameter values. This method using results from another estimation process (e.g., maximum likelihood) is perhaps one of the more common forms of specifying data-driven priors in the applied Bayesian literature (see e.g., Berger, 2006; Brown, 2008; Candel & Winkens, 2003; van der Linden, 2008). However, there are also other methods for defining priors that were not assessed in the current study but that are worth briefly discussing here.

There are several different methods for defining hyperparameter values that are data-driven but that do not rely on initial parameter estimation via maximum likelihood or akin. For example, Raftery and Lewis (1996); Richardson and Green (1997), and Wasserman (2000) have all constructed methods of defining data-driven priors based on summary statistics (e.g., median, mean, variance, range of data) rather than parameter estimates; for an extensive discussion of current data-driven

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

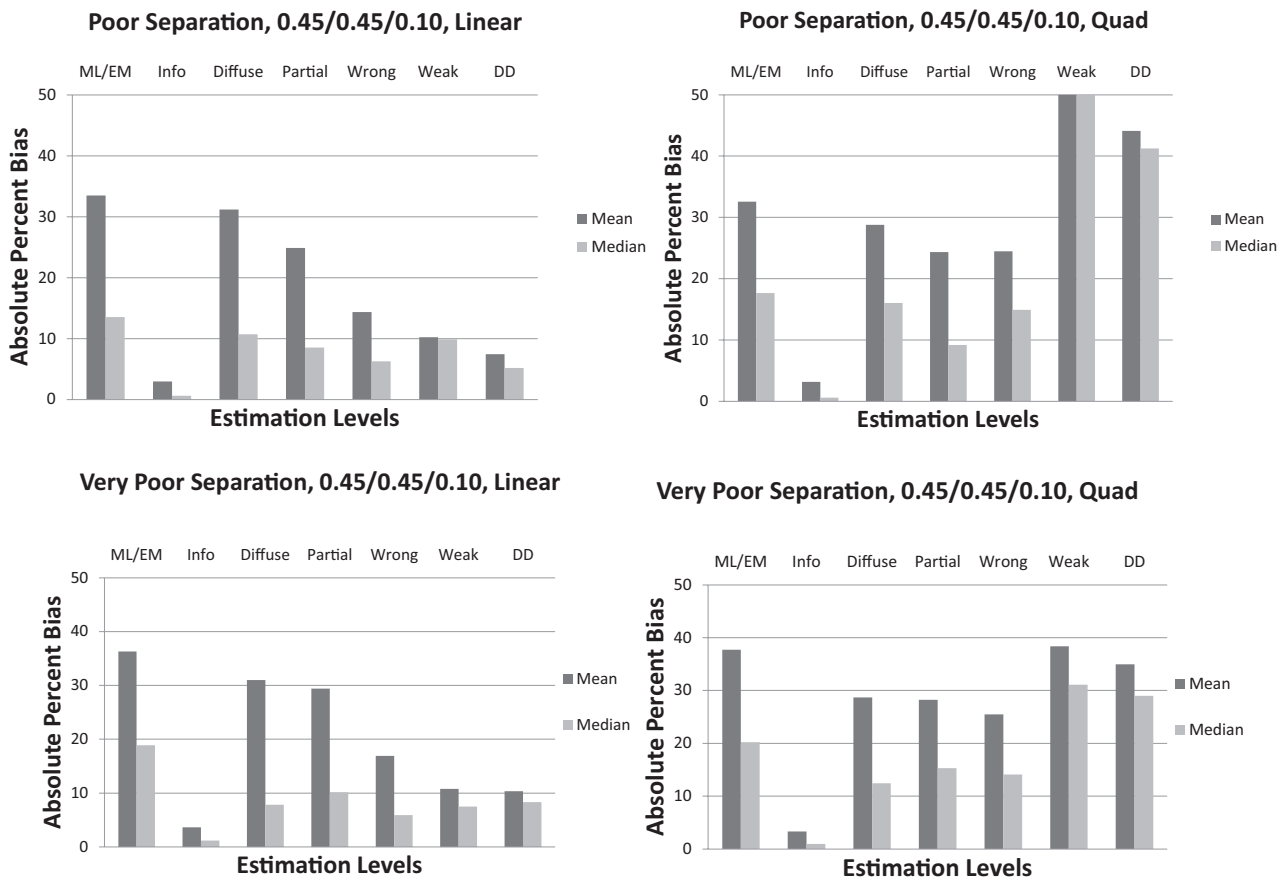


*Figure 11..* Absolute overall bias levels for 0.45/0.45/0.10 latent class proportions, poor and very poor separation. ML/EM = maximum likelihood via the EM algorithm; EM = expectation maximization; Info = Bayesian estimation using informative priors; Diffuse = Bayesian estimation using Mplus default diffuse priors; Partial = Bayesian estimation using informative priors on a partial subset of parameters; Wrong = Bayesian estimation using weak and wrong ("inaccurate") priors; Weak = Bayesian estimation using weak priors; DD = data-driven priors; Quad = quadratic.

methods, see Darniede (2011). These data-driven methods all show promise for applied Bayesian research, but the accuracy of the estimates produced through the various methods has yet to be examined within latent variable models; this is potentially an interesting area for future methodological research.

Another method for defining priors and hyperparameter values is typically referred to as prior elicitation. The elicitation of priors is a process in which experts provide knowledge about an unknown parameter and that knowledge is converted into a prior (probability) distribution. O'Hagan et al. (2006) indicated that elicitation is an important method for defining priors when expert knowledge is the only viable source of information (e.g., assessing the risk of terrorist attacks). Skrondal and Rabe-Hesketh (2004) stated that one of the main reasons for adopting a Bayesian approach using priors would be to specify prior distributions that reflect elicited prior knowledge. They argued that this approach of defining priors is "truly Bayesian" since priors are defined through information gathered from previous research or elicited from content experts.

Regardless of the estimation process or the method used to define priors, the important issue is whether or not the model is useful and able to provide good predictions. In the case of the current study, the Bayesian estimation framework showed promise in accurately estimating GMMs under various conditions. The intention is that these findings prove to be useful for applied researchers assessing growth or change over time.

## References

Asparouhov, T., & Muthén, B. (2010). *Bayesian analysis using Mplus*. Los Angeles, CA: Muthén & Muthén.

Bauer, D. J. (2007). Observations on the use of growth mixture models in psychological research. *Multivariate Behavioral Research, 42,* 757–786.

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.

Berger, J. (2006). The case for objective Bayesian analysis. *Bayesian Analysis, 3,* 385–402.

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

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Brown, L. D. (2008). In-season prediction of batting averages: A field test of empirical Bayes and Bayes methodologies. *The Annals of Applied Statistics, 2,* 113–152.

Candel, J. J. M., & Winkens, B. (2003). Performance of empirical Bayes estimators of level-2 random parameters in multilevel analysis: A Monte Carlo study for longitudinal designs. *Journal of Educational and Behavioral Statistics, 28,* 169–194.

Celeux, G., Hurn, M., & Robert, C. (2000). Computational and inferential difficulties with mixture posterior distributions. *Journal of the American Statistical Association, 95,* 957–970.

Chung, T., Maisto, S. A., Cornelius, J. S., & Martin, C. S. (2004). Adolescents' alcohol and drug use trajectories in the year following treatment. *Journal of Studies on Alcohol, 69,* 105–114.

Darniede, W. F. (2011). *Bayesian methods for data-dependent priors* (Doctoral dissertion). Ohio State University, Columbus.

Depaoli, S. (in press). The impact of "inaccurate" informative priors for growth parameters in Bayesian growth mixture modeling. *Structural Equation Modeling*.

Diebolt, J., & Robert, C. P. (1994). Estimation of finite mixture distributions through Bayesian sampling. *Journal of the Royal Statistical Society, 56,* 363–375.

Enders, C. K., & Tofighi, D. (2008). The impact of misspecifying class-specific residual variances in growth mixture models. *Structural Equation Modeling, 15,* 75–95.

Farrar, D. (2006). *Approaches to the label-switching problem of classification, based on partition-space relabeling and label-invariant visualization*. Blacksburg, VA: Virginia Polytechnic Institute and State University.

Frühwirth-Schnatter, S. (2001). Markov chain Monte Carlo estimation of classical and dynamic switching and mixture models. *Journal of the American Statistical Association, 96,* 194–209.

Gelman, A. (1996). Inference and monitoring convergence. In W. R. Gilks, S. Richardson, & D. J. Spiegelhalter (Eds.), *Markov chain Monte Carlo in practice* (pp. 131–143). New York, NY: Chapman & Hall.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis*. New York, NY: Chapman & Hall/CRC.

Gelman, A., & Rubin, D. B. (1992a). Inference from iterative simulation using multiple sequences. *Statistical Science, 7,* 457–511.

Gelman, A., & Rubin, D. B. (1992b). A single series from the Gibbs sampler provides a false sense of security. In J. M. Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian statistics 4* (pp. 625–631). Oxford, England: Oxford University Press.

Ghosh, J. K., & Mukerjee, R. (1992). Non-informative priors (with discussion). In J. M. Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian statistics 4* (p. 195–210). Oxford, England: Oxford University Press.

Gueorguieva, R., Mallinckrodt, C., & Krystal, J. (2011). Trajectories of depression severity in clinical trials of Duloxetine. *Archives of General Psychiatry, 68,* 1227–1237.

Henson, J. M., Reise, S. P., & Kim, K. H. (2007). Detecting mixtures from structural model differences using latent variable mixture modeling: A comparison of relative model fit statistics. *Structural Equation Modeling, 14,* 202–226.

Hipp, J. R., & Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychological Methods, 11,* 36–53.

Jasra, A., Holmes, C. C., & Stephens, D. A. (2005). Markov chain Monte Carlo methods and the label switching problem in Bayesian mixture modeling. *Statistical Science, 20,* 50–67.

Jedidi, K., Jagpal, H. S., & DeSarbo, W. S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science, 16,* 39–59.

Kaplan, D. (2002). Methodological advances in the analysis of individual growth with relevance to education policy. *Peabody Journal of Education, 77,* 189–215.

Kass, R. E., & Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association, 91,* 1343–1370.

Kenny, D. A. (1979). *Correlation and causality*. New York, NY: Wiley.

Lambert, P. C., Sutton, A. J., Burton, P. R., Abrams, K. R., & Jones, D. R. (2005). How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Statistics in Medicine, 24,* 2401–2428.

Lee, S.-Y. (2007). *Structural equation modeling: A Bayesian approach*. Chichester, England: Wiley.

Lubke, G. H., & Muthén, B. O. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods, 10,* 21–39.

Lubke, G. H., & Muthén, B. O. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling, 14,* 26–47.

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

Morin, A., Mainano, C., Nagengast, B., Marsh, H., Morizot, J., & Janosc, M. (2011). General growth mixture analysis of adolescents' developmental trajectories of anxiety: The impact of untested invariance assumptions on substantive interpretations. *Structural Equation Modeling, 18,* 613–648.

Muthén, B. O. (2003). Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran (2003). *Psychological Methods, 8,* 369–377.

Muthén, B. O. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Newbury Park, CA: Sage.

Muthén, B. O., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods, 17,* 313–335.

Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research, 24,* 882–891.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics, 55,* 463–469.

Muthén, L. K., & Muthén, B. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods, 4,* 139–157.

Natarajan, R., & McCulloch, C. E. (1998). Gibbs sampling with diffuse proper priors: A valid approach to data-driven inference? *Journal of Computational and Graphical Statistics, 7,* 267–277.

Nylund, K., Asparouhov, T., & Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14,* 535–569.

O'Hagan, A. (1998). Eliciting expert beliefs in substantial practical applications. *The Statistician, 47,* 21–35.

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., . . . Rakow, T. (2006). *Uncertain judgments: Eliciting experts' probabilities*. New York, NY: Wiley.

Raftery, A. E., & Lewis, S. M. (1996). Implementing MCMC. In W. R. Gilks, S. Richardson, & D. J. Spiegelhalter (Eds.), *Markov chain Monte Carlo in practice* (pp. 115–130). New York, NY: Chapman & Hall.

Richardson, S., & Green, P. J. (1997). On Bayesian analysis of mixtures with an unknown number of components. *Journal of the Royal Statistical Society, 59,* 731–792.

Rindskopf, D. (2003). Mixture or homogeneous? Comment on Bauer and Curran (2003). *Psychological Methods, 8,* 364–368.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Roeder, K., & Wasserman, L. (1997). Practical Bayesian density estimation using mixtures of normals. *Journal of the American Statistical Association, 92,* 894–902.

Schaeffer, C., Petras, H., Ialongo, N., Masyn, K., Hubbard, S., Poduska, J., & Kellan, S. (2006). A comparison of girl's and boy's aggressive-disruptive behavior trajectories across elementary school: Prediction to young adult antisocial outcomes. *Journal of Consulting and Clinical Psychology, 74,* 500–510.

Sinharay, S. (2004). Experiences with Markov chain Monte Carlo convergence assessment in two psychometric examples. *Journal of Educational and Behavioral Statistics, 29,* 461–488.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Boca Raton, FL: Chapman & Hall/CRC.

Stephens, M. (2000). Dealing with label switching in mixture models. *Journal of the Royal Statistical Society, 62,* 795–809.

Szapocznik, J., Feaster, D. J., Mitrani, V. B., Prado, G., Smith, L., Robinson-Batista, C., . . . Robbins, M. S. (2004). Structural ecosystems therapy for HIV-Seropositive African American women: Effects on psychological distress, family hassles, and family support. *Journal of Consulting and Clinical Psychology, 72,* 288–303.

Tofighi, D., & Enders, C. K. (2008). Identifying the correct number of classes in growth mixture models. In G. R. Hancock & K. M. Samuelson (Eds.), *Advances in latent variable mixture models* (pp. 317–341). Charlotte, NC: Information Age.

Tolvanen, A. (2008). *Latent growth mixture modeling: A simulation study* (Doctoral dissertation). University of Jyväskylä, Jyväskylä, Finland.

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

van der Linden, W. J. (2008). Using response times for item selection in adaptive testing. *Journal of Educational and Behavioral Statistics, 33,* 5–20.

Wasserman, L. (2000). Asymptotic inference for mixture models using data-dependent priors. *Journal of the Royal Statistical Society, 62,* 159–180.

Yang, M., & Dunson, D. B. (2010). Bayesian semiparametric structural equation models with latent variables. *Psychometrika, 75,* 675–693.

Zhang, Z., Hamagami, F., Wang, L., Nesselroade, J. R., & Grimm, K. (2002). Bayesian analysis of longitudinal data using growth curve models. *International Journal of Behavioral Development, 31,* 374–383.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

# Appendix

## Class Separation via Manifest Variables Using Mahalanobis Distance

Mahalanobis distance (MD) using the manifest variables can be computed using the same equation presented in footnote 7 if the implied manifest variable covariance matrix is homogeneous across classes, which was the case here. In the current investigation, the implied (linear-trajectory) manifest variable values for C2 were $y1 - y4 = 42.802$, 46.802, 50.802, and 54.802, respectively. The implied manifest variable values for C3 are presented in Table A1 for each of the four separation conditions named in the study (MD = 2.0, 1.5, 1.0, and 0.5, respectively).

The diagonal covariance matrix for the residuals was *Diag*[15,15,15,15] or

$$\begin{bmatrix} 15 & 0 & 0 & 0 \\ 0 & 15 & 0 & 0 \\ 0 & 0 & 15 & 0 \\ 0 & 0 & 0 & 15 \end{bmatrix}.$$

Using the above values, the MD values based on the manifest variables were produced for each of the four separation conditions named in the study; these MD values based on manifest variables can be found in Table A2.

Note that relatively smaller residual variances will increase the separation between classes based on the manifest variables and relatively larger residual variances will decrease separation (Lubke & Muthén, 2007). For example, Table A3 illustrates how MD varies through changing only the residual variances—the manifest values for $y1$–$y4$ remain constant in this example. Note that MD as implemented here requires the covariance matrix be homogenous across classes, whereas other distance measures do not make this assumption (e.g., the Kullback-Leibler divergence measure).

(*Appendix continues*)

Table A1
*Implied Manifest Values for the C3 Trajectory Under Four Class Separation Levels*

| | Separation: Linear C3 | | | | Separation: Quadratic C3 | | | |
|---|---|---|---|---|---|---|---|---|
| Time | High | Moderate | Poor | Very poor | High | Moderate | Poor | Very poor |
| 1 | 37.6 | 38.558 | 39.79 | 42.53 | 37.6 | 38.558 | 39.79 | 42.53 |
| 2 | 40.6 | 41.558 | 42.79 | 45.53 | 41.5 | 42.458 | 43.69 | 46.43 |
| 3 | 43.6 | 44.558 | 45.79 | 48.53 | 47.2 | 48.158 | 49.39 | 52.13 |
| 4 | 46.6 | 47.558 | 48.79 | 51.53 | 54.7 | 55.658 | 56.89 | 59.63 |

*Note*. C3 = Latent Class 3.

Table A2
*Mahalanobis Distance Values Based on Manifest Variables for Four Separation Conditions*

| | Separation: Linear C3 | | | | Separation: Quadratic C3 | | | |
|---|---|---|---|---|---|---|---|---|
| MD | High | Moderate | Poor | Very poor | High | Moderate | Poor | Very poor |
| MD | 12.31 | 9.13 | 5.76 | 1.17 | 4.54 | 2.97 | 1.67 | 1.68 |
| Sqrt(MD) | 3.51 | 3.02 | 2.40 | 1.08 | 2.13 | 1.72 | 1.29 | 1.30 |

*Note*. C3 = Latent Class 3; Sqrt = square root; MD = Mahalanobis distance.

Table A3
*Illustration of the Impact of Residual Variances on Mahalanobis Distance Values*

| First class: $y1$–$y4$ | Second class: $y1$–$y4$ | Diagonal residual variance | Mahalanobis distance |
|---|---|---|---|
| 48,51,54,57 | 47,49,51,53 | *Diag*[15,15,15,15] | 2.0 |
| 48,51,54,57 | 47,49,51,53 | *Diag*[20,20,20,20] | 1.5 |
| 48,51,54,57 | 47,49,51,53 | *Diag*[30,30,30,30] | 1.0 |
| 48,51,54,57 | 47,49,51,53 | *Diag*[60,60,60,60] | 0.5 |

*Note*. Diag = diagonal.

Received January 21, 2011
Revision received November 6, 2012
Accepted November 21, 2012 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.