ABSTRACT

Title of Dissertation: USING LATENT PROFILE MODELS AND UNSTRUCTURED GROWTH MIXTURE MODELS TO ASSESS THE NUMBER OF LATENT CLASSES IN GROWTH MIXTURE MODELING

Min Liu, Doctor of Philosophy, 2011

Directed By: Professor Gregory R. Hancock
Department of Measurement, Statistics and Evaluation

Growth mixture modeling has gained much attention in applied and methodological social science research recently, but the selection of the number of latent classes for such models remains a challenging issue. This problem becomes more serious when one of the key assumptions of this model, proper model-specification is violated.

The current simulation study compared the performance of a linear growth mixture model in determining the correct number of latent classes against two less parametrically restricted options, a latent profile model and an unstructured growth mixture model. A variety of conditions were examined, both for properly and improperly specified models. Results indicate that prior to the application of linear growth mixture model, the unstructured growth mixture model is a promising way to identify the correct number of unobserved groups underlying the data by using most model fit indices across all the conditions investigated in this study.

USING LATENT PROFILE MODELS AND UNSTRUCTURED GROWTH MIXTURE MODELS TO ASSESS THE NUMBER OF LATENT CLASSES IN GROWTH MIXTURE MODELING

By

Min Liu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:
Professor Gregory R. Hancock, Chair
Professor Robert J. Mislevy
Professor Jeffrey R. Harring
Professor Hong Jiao
Professor Paul J. Smith

Dedication

To my father, Jian Liu (刘坚, 又名会贵，1943.7.20-2010.9.16)

Acknowledgements

Every dissertation is a long journey. Fortunately, I am not alone during this travelling. First, I want to thank my family, my parents, Jian Liu and Cuiyun Liu who always encourage me pursuing a Ph.D degree and never lose faith in me, and my son, Derek I. Mei who has been my biggest motivation and make me become a better person.

My deepest gratitude goes to my advisor, Dr. Gregory Hancock. It was in the intellectual discussion with him that I find out the right direction to proceed. Even when I drifted away from time to time because of my frustrations caused by either academic or family difficulties, it was his generous patience and warmest encouragement that brought me back. His help and support extends far beyond this dissertation.

I'm also thankful to Dr. Robert Mislevy, from whom I learned a lot, including how to be a professional. Without his support, my life must be more difficult. And I also want to thank Dr. Robert Lissitz. It is he who provided me the opportunity to study in EDMS, from where I started my academic life.

I also sincerely appreciate Dr. Jeffrey R. Harring, Dr. Hong Jiao, and Dr. Paul J. Smith, for their valuable comments on my dissertation. And I do think Dr. Harring and Dr. Jiao are good models for young professionals.

Last but not least, I want to thank Dr. Heather Buzick for her nice and sincere help to this dissertation, other EDMS fellows and my friends in both U.S and China who gave me support when I walked in darkness. They include Dr. Yvonne Delores Oslin, Dr. Hua Wei, Dr. Weihua Fan, Junhui Liu, Youngmi Cho, Li Du, Tao Chen, Zhigang Fan, Bo Huang, Yao Tong, Yunting Li, Li Zhang, Limei Zhang, Binbin Xu, Chunxia Huang and Xiaoxia Song.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

Research question that the current study aims to address arises from an empirical research on reading achievement development of elementary students across Kindergarten to 5^{th} grade (Douglas & Liu, 2009). Below Spaghetti Plot illustrates six random samples of students' reading achievement scores from Early Childhood Longitudinal Study- Kindergarten Cohort (ECLS-K). Visual inspection indicates some students have steeper growth in the early years than others. This apparent heterogeneity motivated the need to consider using multiple growth trajectories to model this type of growth for all students. For this research purpose, growth mixture model (GMM), was selected as a suitable tool to investigate unobserved different group-based growth curves in this longitudinal data because GMM, as briefly introduced in the following paragraph, has its advantages over traditional or other statistical methods for studying developmental process.

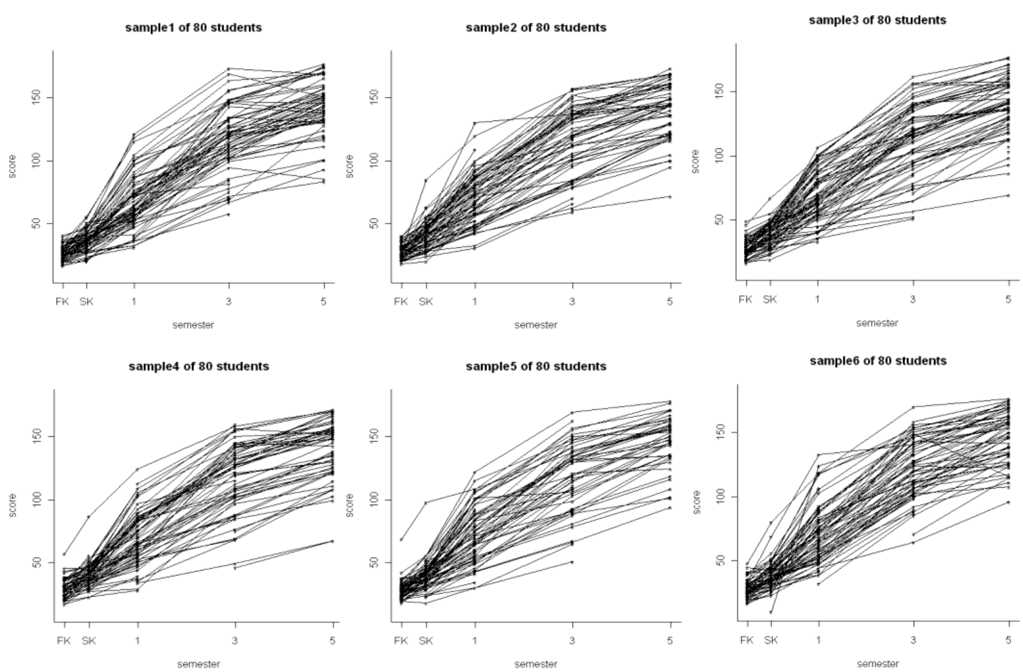

Figure 1. *Spaghetti plots of reading achievement scores across Kindergarten to 5th Grade*

Traditional mean-based methods (e.g., repeated-measures ANOVA) for studying individuals' developmental change assume that all individuals change in a uniform pattern. That is to say, no random variation among individuals is allowed. More advanced statistical techniques proposed in the latter part of the twentieth century made an improvement by incorporating individual variation from the single fixed function into the models, such as hierarchical linear modeling (see, e.g., Raudenbush & Bryk, 2002), random-effect modeling (Laird & Ware, 1982) and latent growth modeling (LGM) in the structural equation modeling context (for review see Hancock & Lawrence, 2006). However, all of these methods assume there is only one population (i.e., one group-based trajectory) underlying the data, which may not be met in practice. Numerous examples can be illustrated in this regard. For example in education, students from kindergarden to 5th grade can be classified into fast and normal readers in terms of their different growth trajectories in learning reading (Douglas & Liu, 2009). Taken another example in marketing application, Jedidi, Jagpal, and Desarbo (1997) illustrated the misleading model estimations due to ignoring the existence of heterogeneity.

Growth mixture modeling (GMM) has gained much attention in the past decade for its capability of exploring and identifying different group-based growth curves in longitudinal data by considering both random effects and population heterogeneity. Therefore, GMM has been widely applied in the social and behavioral sciences. Examples of its application include studies of college alcohol development (e.g., Greenbaum, Del Boca, Darkes, Wang, & Goldman, 2005), depression patterns (e.g., Stoolmiller, Kim, & Capaldi, 2005), reading skills from kindergarten to 5th grade (e.g., Douglas & Liu, 2009), medication effects (e.g., Muthén, Brown, Hunter, Cook & Leuchter, 2011), and criminal behavior trajectories (Kreuter & Muthén, 2008a).

Whenever researchers start their data analysis using GMM, a question arises, how many different growth trajectories should be applied for this data? In other words, how many unobserved groups exhibit distinct growth patterns across time? Give a further reflection, which criteria or method can be used to identify the number of unobserved groups accurately? In fact, the problem of class enumeration has invoked numerous debates on whether and how GMM should be used in practice soon after its appearance. So the main theme throughout the current research work is about how to identify the number of latent class for GMM accurately.

In fact, the enumeration of latent groups (classes) is a problematic issue, not only for GMM, but also for other mixture models (e.g., mixture confirmatory factor analysis models and latent class models). But this problem is particularly challenging when the key assumptions of GMM are violated, as Bauer (2007) and Bauer and Curran (2003, 2004) pointed out. As these authors stated, when the assumption of having a properly specified within-class model is not met, spurious classes may be generated to compensate leading to further inaccurate longitudinal inference. This is especially disconcerting in practice because the true model is never known a *priori*, which is the dilemma that researchers have to deal with in the empirical study for students' reading skills development.

To address this problem, the current work proposes to use less restricted mixture models to determine the number of latent classes prior to applying GMM directly. This idea is theoretically compelling in the sense that fewer restrictions are imposed on the model structure and thus there is less chance that model misspecification would occur. Consequently, the possible spurious latent classes caused by the improperly specified model might, in theory, be avoided. This idea has never been empirically investigated for GMM. As such, the current study is an extensive Monte Carlo study examining the accuracy of the number of latent classes for GMM suggested through a *priori* application of two less-restricted mixture models: the Latent Profile Model (LPM), which is completely unrestricted since no restricted relation is imposed among variables, and the Unstructured Growth Mixture Model (UGMM), which is partially restricted in the sense that the growth function is not restricted to be linear but the correlations among observed variables are still driven by latent growth factors. A wide range of model fit indices were used to choose the number of latent classes for each model and their relative performance was evaluated.

CHAPTER 2: LITERATURE REVIEW

To better understand this work and its contributions to related field, this chapter reviews the related literature as follows: Section 2.1 describes a general theory framework for GMM; Section 2.2 presents key methodological problems and consequence associated with GMM and suggested solutions; Section 2.3 proposes the main idea of the current work and introduces the unrestricted LPM and less restricted UGMM; Section 2.4 introduces three types of model selection indicators for evaluating the number of latent classes in a GMM context and related simulation studies for comparing the efficiency of those indicators.

2.1. Growth Mixture Model

Although some precursor work (e.g., Verbeke & Lesaffre, 1996) had implied the similar idea of a mixture of random effects in linear mixed-effects model, GMM was first formally introduced by Muthén and Shedden (1999), and was extended in later publications by Muthén and his colleague (2001, 2002, 2004, & 2008).

2.1.1 General Function for Growth Mixture Model

According to Muthén and Shedden' (1999) work, the general function for GMM can be written in matrix form as:

$$\mathbf{y} = \mathbf{\Lambda}^k \boldsymbol{\eta}^k + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\eta}^k = \boldsymbol{\alpha}^k + \mathbf{\Gamma}^k \mathbf{x} + \boldsymbol{\zeta}^k$$

where

$$\boldsymbol{\varepsilon} \sim N(0, \mathbf{\Theta}^k)$$

and

$$\boldsymbol{\zeta}^k \sim N(0, \boldsymbol{\Psi}^k)$$

All the symbols with superscript *k* imply that they differ across latent classes.

$\mathbf{y}$ denotes the vector of continuous repeated measures for an individual, $\boldsymbol{\Lambda}^k$ is the matrix of factor loadings, which usually has a fixed pattern reflecting the growth function. For example, $\boldsymbol{\Lambda} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$ indicates a linear function for a GMM with four equally spaced repeated measures. $\boldsymbol{\varepsilon}$ is residual vector at level 1 and it is assumed to be normally distributed with mean zero and a typically diagonal covariance matrix $\boldsymbol{\Theta}^k$, indicating that relations among repeated measures are fully captured by the latent growth factors $\boldsymbol{\eta}^k$. $\boldsymbol{\alpha}^k$ is the vector of latent factor means, $\mathbf{x}$ is the observed covariate vector and $\boldsymbol{\Gamma}^k$ is the matrix of regression coefficients of latent factors $\boldsymbol{\eta}^k$ on covariates $\mathbf{x}$. $\boldsymbol{\zeta}^k$ is the residual vector that also follows normal distribution with mean zero and covariance matrix $\boldsymbol{\Psi}^k$. The normality assumption of random effects implies that the individual variations are centered on the expected value of $\boldsymbol{\Lambda}^k\boldsymbol{\alpha}^k + \boldsymbol{\Lambda}^k\boldsymbol{\Gamma}^k\mathbf{x}$ within each latent class and they deviate from the center symmetrically.

2.1.2 Unconditional GMM

The inclusion of covariates was recommended in order to “correctly specify the model, find the proper number of latent classes, and correctly estimate class proportions and class membership” (Lubke & Muthén, 2007; Muthén, 2004).

However, a recent academic talk with Muthén suggested (Marsh, Ludtke, Trautwein & Morin, 2009) that the inclusion of covariates must satisfy a strong assumption; the covariates are strictly antecedent variables to the latent classes, indicating that the causal ordering must be from the covariates to the latent classes. Because it is difficult to test this assumption in practice, researchers should evaluate the inclusion of covariates carefully even with a strong justification to do so (Marsh et al., 2009). Considering that our primary research concern is how to determine the number of latent classes accurately rather than investigate the kind of relations among variables, and that covariates have been shown to present challenges for class enumeration (Tofighi & Enders, 2008), no covariate is considered in this study. Therefore, after covariates are removed from the equation (2), the function for unconditional GMM in matrix form becomes

$$\boldsymbol{\eta}^k = \boldsymbol{\alpha}^k + \boldsymbol{\zeta}^k$$

Now the individual variation in centered on the estimated intercept and slopes within each latent class.

2.1.3 Estimation of GMM

Maximum likelihood (ML) estimation is the dominant method for estimating mixture models (Yung, 1997). It is also used to estimate GMM through implementation of the EM algorithm (Muthén & Shedden, 1999). Following Tolvanen's (2008) derivation, the log-likelihood function of observed data for the GMM can be constructed as below:

$$\log L = \log\left[\prod_{i=1}^{n} L_i\right] = \sum_{i=1}^{n} \log L_i = \sum_{i=1}^{n} \log f(y_i)$$

where the density function is a mixture of K density functions for different latent classes as below

$$f(y_i) = \sum_{i=1}^{n} \pi^k f^k(y_i)$$

where π^k is the proportion of latent class k, whose density function follows a multivariate normal distribution:

$$f^k(y) \sim N(\boldsymbol{\mu}^k, \boldsymbol{\Sigma}^k)$$

where

$$\boldsymbol{\mu}^k = \boldsymbol{\Lambda}^k \boldsymbol{\alpha}^k$$

$$\boldsymbol{\Sigma}^k = \boldsymbol{\Lambda}^k \boldsymbol{\Psi}^k \boldsymbol{\Lambda}^k + \boldsymbol{\Theta}^k$$

and then the conditional density function is

$$f(y_i \mid c_i) = \sum_{k=1}^{K} p(c_{ik} = 1) f^k(y_i \mid c_{ik} = 1)$$

$c_{ik} = 1$ indicates i^{th} observation belongs to latent class k and $c_{ik} = 0$ otherwise.

$\sum_{k=1}^{K} p(c_{ik} = 1) = \sum_{k=1}^{k} p^k = 1$. This restriction is necessary for model identification.

Including the class information, the complete loglikelihood is

$$\log \prod_{i=1}^{n} f(y_i \mid c_i)$$

$$= \log \prod_{i=1}^{n} \prod_{k=1}^{K} [\pi(c_{ik} = 1) f(y_i \mid c_{ik})]^{c_{ik}}$$

$$= \sum_{i=1}^{n} \left[\log \prod_{k=1}^{K} \pi(c_{ik} = 1)^{c_{ik}} f(y_i \mid c_{ik})^{c_{ik}} \right]$$

$$= \sum_{i=1}^{n} \left[\log \prod_{k=1}^{K} \pi(c_{ik} = 1)^{c_{ik}} + \log \prod_{k=1}^{K} f(y_i \mid c_{ik})^{c_{ik}} \right]$$

$$= \sum_{i=1}^{n} \left[\sum_{k=1}^{K} c_{ik} \log \pi(c_{ik} = 1) + \sum_{k=1}^{K} c_{ik} \log f(y_i \mid c_{ik}) \right]$$

From the derivations of the above equation, we can infer that the estimation consists of two parts: estimating the sum of the weighted *K* class proportions and the sum of the weighted *K* density functions.

The EM algorithm includes an E-(expectation) step and an M-(maximization) step. In the E-step, the values of latent class information (i.e., posterior probabilities for each observation falling into each latent class after the first iteration) are considered missing and their expected values are estimated based on the starting values given in the first iteration and then the values from the M-step in following iterations. As expectations of the elements of the vector of class membership indicator variables c_{ik} they take the form of posterior probabilities of class membership. Then those posterior probabilities are inserted in the M-step to maximize the (expected) loglikelihood in this equation. Consequently, we get all the estimated parameters within each latent class at this iteration. After the M-step, the EM algorithm returns back to the E-step to obtain a new set of posterior probabilities. The iterations continue until some convergence criterion related to the complete-data log-likelihood is satisfied.

2.2. Methodological problems with GMM and suggested solutions

The increased popularity of GMM in the social sciences has invoked many methodological concerns, especially the enumeration of latent classes for this model, which is the first and a crucial step of applying GMM in practice. In fact, class enumeration is always a challenging issue for mixture modeling (e.g., latent class analysis, mixture confirmatory factor analysis). As experts emphasize, the application of GMM should be based on substantive theory (e.g., Muthén, 2003, 2004). A recent handbook for methodology in psychology explicitly states (Little, Card, Preacher, & McConnell, 2009) that to confirm a theory, researchers should clearly state "(1) why qualitatively distinct classes should exist, (2) how many classes should exist, and (3) what the functional form of the growth trajectories within each class should be," (pp.39) based on sufficient theoretical reasons.

However, usually this is not the case in practice. When a researcher believes in the existence of population heterogeneity in the developmental data, it is more likely that he/she will use an exploratory way to evaluate the number of latent classes for GMM. Unlike conventional structural equation models, testing the overall fit for GMM with different latent classes is not possible, as this model belongs to the mixture-modeling framework. Instead, researchers rely on statistical model indices to compare the relative fit of competing models with different latent classes to the data. This data-driven approach triggered much criticism on using GMM in the social sciences because spurious latent class might be generated from data and this problem becomes more serious when the key assumptions of GMM are violated.

To streamline following discussion of those methodological concerns, Table 2.2 provides a brief summary of all the methodological problems, authors' findings on the effects on class enumeration, and suggested solutions. Among them, the problems of local maxima and non-normality have received greater attention recently, but much less so for the other problems. Despite all these problems, GMM has become widely used for developmental study in the social sciences (e.g., psychopathology, Odgers, Moffitt, Broadbent, Dickson, Hancox, Harrington et al., 2008; organizational study, Wang & Bodner, 2007). Clearly, it is imperative and extremely significant to solve those methodological concerns regarding GMM to ensure this model as a promising approach for analyzing heterogeneous latent development process underlying data

Table 2.2 *Methodological problems, associated consequence on class enumeration and possible solutions*

Problems	Effects on class enumeration	Suggested solutions
Violation of within-class normality	overestimate	Second-order GMM (Grimm & Ram, 2009); Non-parametric version of a GMM (Muthén & Asparouhov, 2008; Kreuter & Muthén, 2008b); Skew-normal mixture model (Azzalini, 1985 & 2005; Chang, 2005)
Local Maxima	under-or overestimate	Multiple random starting values across a wide range of parameter space (Hipp & Bauer, 2006)
Violation of data missing at random (MAR)	might underestimate	Pattern mixture model or Probability weight (Bauer, 2007)
Violation of simple random sampling	might overestimate	Design-based or model-based approach (Hamilton, 2009)
Misspecification of within-class model (nonlinear relation is a special case)	overestimate	Unrestricted (or saturated) model (Yung, 1997; Bauer & Curran, 2004)

Bauer and Curran (2003, 2004) offered strong arguments against GMM. In their work in 2003, they showed that if the repeated measures are non-normal, a GMM with multiple latent classes always fits data better than a single-class latent growth model, whether or not the non-normality is caused by the mixture of multiple normal subpopulations or a unitary non-normal distribution. Even mild violation of normality may result in many artifact latent classes (Bauer & Curran, 2003; Tofighi & Enders, 2008). Actually, this phenomenon has been observed in mixture models assuming normal distributions for several decades (e.g., Maclean, Morton, Elston & Yee, 1976).

Several studies have been done to address the violation of the within-class normality assumption, as mentioned in the first chapter. Grimm and Ram (2009) posited that the latent construct of interest might be normally distributed, whereas its observed indicators might be non-normal due to ceiling, floor, or other possible measurement anomalies. Borrowing the idea from Hancock, Kuo, and Lawrence (2001), they proposed the second-order GMM, in which the factor scores indicated by observed variables were used as repeated measures across four occasions. As such, these latent constructs can provide more precise true-score distributions from the sample with non-normal data. We can see that this approach reduces the effect of measurement error, which directly deals with ε^k. In this way, the risk of generating spurious latent classes from non-normal data (not a mixture of multiple normal distributions) is reduced. However, there is one limitation of applying this model in practice: it requires many more observed variables (i.e., indicators) to build up this complex model.

Muthén and his colleagues proposed a non-parametric version of a GMM (NP-GMM) to accommodate non-normal random effects, which is denoted as ζ^k in the GMM model (Kreuter & Muthén, 2008b; Muthén & Asparouhov, 2008). Inspired by the idea of latent class growth analysis (LCGA), NP-GMM also does not rely on any distribution assumption for the random effect. Instead, it uses additional latent classes to capture the non-normal distribution within the *K* latent classes specified before. Unlike LCGA, only the *K* latent classes have substantive meaning in NP-GMM; those additional latent classes within them are just mathematical approximations to fit the non-normal data within the *K* GMM classes. In other words, practitioners do not have to interpret those additional latent classes as meaningful subpopulations. NP-GMM can be used to model non-normal data as long as the number of latent classes K and the non-normality of the random effects are known *a priori*. However, this approach does not completely solve the problem of overextraction of latent classes caused by non-normal data because the K latent classes are established prior to the estimation of NP-GMM.

Another potential method that might alleviate the overextraction of latent classes caused by nonnormal distributions is to change the underlying normal distribution to the skew normal distribution (Azzalini, 1985), in which a skewness parameter is introduced to loosen the normality assumption and thus the normal distribution becomes a special case. Chang (2005) applied this skew-normal mixture model to data with existence of skewness and successfully determined the number of components. By the same token, it is reasonable to assume this method could be used for the same purpose in GMM context.

The second problem associated with GMM is local maxima in the estimation process. Unlike a latent growth model for a homogeneous population, but similar to other finite mixture models, GMM could have a poorly behaved likelihood function often resulting in incorrect local solutions, as opposed to global maxima (e.g., Muthén & Shedden, 1999). Hipp and Bauer (2006) first presented an empirical study on the local optima problem in GMM for applied researchers and clearly recommended that it is necessary to vary the starting values extensively on the likelihood surface to obtain the global maxima. Almost at the same time, Mplus incorporated multiple random starting values across a wide range of the parameter space when estimating models. Moreover, Mplus version 6 can provide all the highest log-likelihood values and associated class proportion information from different solutions due to different starting values if users request “tech8” in the output. This function can give more diagnostic information for the appropriateness of the model.

In addition to the above problems, Bauer (2007) summarized other possible conditions that might prompt inappropriate estimation of latent classes. He found that if the missing data are modeled as random but in fact they are not, the number of latent classes might be underestimated because some smaller extreme classes could be under-represented in the observed data and hence become more difficult to recover the truth. Bauer (2007) also mentioned two possible corrections for this problem, using pattern mixture models or using probability weights to adjust for non-response and attrition.

In the same work, Bauer also pointed out that if the complex sampling is ignored and treated as simple random sampling, the number of latent classes might be incorrectly enumerated, such as the overextraction case in Wedel, ter Hofstede, and Steenkamp's (1998) work for finite mixture models in general. To alleviate the effect of violating this assumption, Hamilton (2009) conducted a simulation study to investigate using either design-based (i.e., weights) or a model-based approach (i.e., modeling stratification variables directly) or both to account for unequal probabilistic selection resulting from complex sampling design. However, neither approach can provide acceptable proportion of unbiased parameter estimates, though design-based performs better than the other. More importantly, she did not examine the effect of these adjustments on the accuracy of class enumeration.

Both Bauer (2007) and Bauer and Curran (2004) noticed that misspecification of the within-class model might also lead to spurious latent classes to capture the variance-covariance of the repeated measures. Moreover, Bauer (2007) pointed out that if nonlinear relation between exogenous predictors and the trajectory parameters within classes is treated as linear, more latent classes are required to approximate the data. Actually the nonlinear component is just one special case of model misspecification. To address this problem, a two-step modeling process was proposed to avoid that class overextraction solely induced by the model misspecification (Bauer & Curran, 2004; Yung, 1997). In the first step, the unrestricted (or saturated) models with different number of latent classes are estimated and compared according to the model fit indices, since no restriction is imposed on the within-class model structure and thus no within-class model misspecification would occur. Consequently, the possible spurious latent classes caused by an improperly specified model might, in theory, be avoided. Supposing the number of latent classes is correctly identified in the first step, in the second step the hypothesized models are fit to the data to see if the models can adequately capture the within-class mean and covariance structures underlying the data.

This idea is theoretically compelling. However, it has not been investigated for GMM and no empirical evidence is available to support this new decision rule. This study is designed to fill this gap. As GMM alone is prone to overextraction under certain misspecified model conditions as mentioned above, it is reasonable to suggest that an unrestricted although proper model could perform better as a preliminary tool for class determination of GMM. In the following section, a latent profile model, a completely unrestricted mixture model, is introduced in the first step to identify the number of latent classes.

2.3. Using Unrestricted or Less Restricted Mixture Model to Address Class Enumeration Problems Caused by Misspecified Within-Class Model

The latent profile model (LPM) was first developed by Gibson (1959). It is quite similar to latent class analysis (LCA) in the sense that they both use a model-based probabilistic approach to classify subjects into different groups (characterized by some distribution with unique set of parameters for each group) and can be tested with a number of model fit indices. Their difference lies in that LCA uses binary indicators while LPM uses continuous indicators. For this reason, LPM has been called "Latent class models with metrical manifest variables" (Bartholomew, 1987, pp.34). Comparing to traditional cluster analysis, LPM is advantageous because it does not require indicators on the same scales prior to their input into the analysis. The fundamental equations of LPM in matrix form can be written as

$$f(\mathbf{y}) = \sum_{k=1}^{K} \pi^k f(\mathbf{y} \mid \mu^k, \sum\nolimits^k \ )$$

The density function of LPM $f(y)$ is a sum of weighted group-based conditional distribution, each of which is defined by a mean vector μ^k and covariance matrix Σ^k. In social and behavioral science, the conditional distribution usually is assumed to be normal, but not limited to this form. π^k denotes each class proportion and so $\sum_{k=1}^{K} \pi^k = 1$. There are different ways to parameterize covariance matrix Σ^k as shown in Table 2.3.1.

Model E is chosen to fulfill the research goal in current study because there is no restriction imposed on the covariance, which makes LPM a completely unrestricted mixture model. As such it is a useful tool to study population heterogeneity (e.g., Hill, Degnan, Calkins, & Keane, 2006; Marsh et al., 2009).

However, as Bauer and Curran (2004) noted, a saturated (completely unrestricted) model has far more parameters to be estimated than the restricted model. Table 2.3.2 presents the number of parameters to be estimated in the three types of mixture models, linear GMM, UGMM (will be introduced later), and LPM. Clearly, LPM has many more parameters that need to be inferred from data than other two models. This is particularly clear in the models with 7 repeated measures. LPM doubles the number of parameters in UGMM, and almost triples as linear GMM.

Table 2.3.1 *Five parameterization ways of* Σ^k *for r indicators*

Model	Σ^k	Characteristics
A	$\begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \sigma_r^2 \end{bmatrix}$	Variance are allowed to differ across indicators within a class, but are constrained to be equal across classes; all covariances are zero.
B	$\begin{bmatrix} \sigma_1^2 & & & \\ \sigma_{21} & \sigma_2^2 & & \\ \vdots & \vdots & \ddots & \\ \sigma_{r1} & \sigma_{r2} & \cdots & \sigma_r^2 \end{bmatrix}$	Less restricted than Model A; covariance are freely estimated within a class, but are constrained to be equal across classes.
C	$\begin{bmatrix} \sigma_{1k}^2 & & & \\ 0 & \sigma_{2k}^2 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \sigma_{rk}^2 \end{bmatrix}$	Less restricted than Model A; variance are also freely estimated across classes
D	$\begin{bmatrix} \sigma_{1k}^2 & & & \\ \sigma_{21} & \sigma_{2k}^2 & & \\ \vdots & \vdots & \ddots & \\ \sigma_{r1} & \sigma_{r2} & \cdots & \sigma_{rk}^2 \end{bmatrix}$	Less restricted than Model C; covariance are freely estimated within a class, but are constrained to be equal across classes.
E	$\begin{bmatrix} \sigma_{1k}^2 & & & \\ \sigma_{21k} & \sigma_{2k}^2 & & \\ \vdots & \vdots & \ddots & \\ \sigma_{r1k} & \sigma_{r2k} & \cdots & \sigma_{rk}^2 \end{bmatrix}$	Least restricted model; variance and covariance are freely estimated within and across classes.

Note: this table is adapted from Pastor, Barron, Miller & Davis (2007)

Table 2.3.2 *The number of parameters to be estimated in the three types of mixture models with 4 and 7 repeated measures*

	LPM	UGMM	linear GMM
1-class	14/35	11/17	9/12
2-class	29/71	23/35	19/25
3-class	44/107	35/53	29/38

In statistical modeling, researchers always need to consider the bias-variance tradeoff (or “bias-variance dilemma”) as displayed in Figure 2.2 (e.g., A’Hearn & Komlos, 2003; Rice, Lumley, & Szpiro, 2008). In practice, whenever an incorrect restriction is imposed, fewer parameters are required and some degree of bias is induced. As long as researchers can find a balance point so that this restriction is close to the truth, the bias induced will be small while the reduction in variance will be substantial. In reality, the choice between restricted and unrestricted model estimation depends on the researcher’s degree of confidence in those restrictions. How to decide this trade-off is an empirical question, highly related to sample size (A’Hearn & Komlos, 2003). In the results section, it is observed that the model performance, especially LPM, is highly related to sample size.

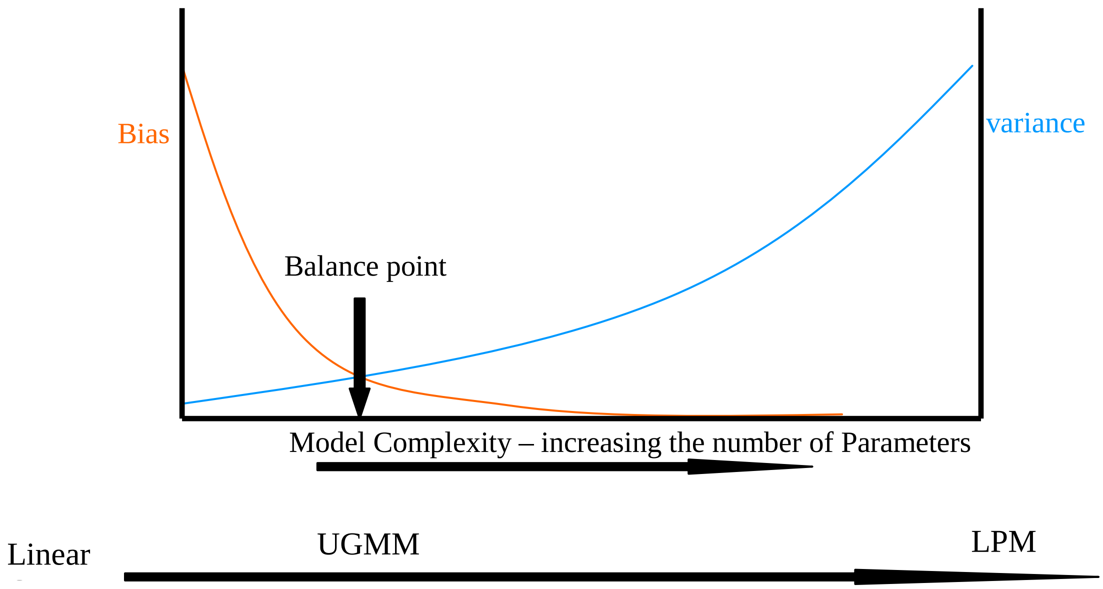

Figure 2.2 *The trade-off between bias and precision in statistical modeling*

Taking into account this rationale in our context, the linear GMM could be considered the most restricted model and put on the leftmost end of the horizontal line while the LPM is the least restricted model and could be put on the other end. Our preliminary results indicate that LPM does not always outperform linear GMM in class enumeration, possibly due to too many parameters to be estimated in LPM. For this reason, an Unstructured Growth Mixture Model (UGMM) is proposed as a balanced model to be compared with the other two in determining the number of latent classes. Compared to GMM, UGMM is partially unrestricted in the sense that the growth function is not restricted to be linear; compared to LPM, UGMM is more restricted since it still assumes the correlations among observed variables within each class are driven by latent growth factors.

As stated above, usually Λ is a matrix of fixed-factor loadings indicating fixed-growth function. As for UGMM, Λ does not follow a fixed pattern any more and needs to be estimated from data. Still, taking the GMM with four equally spaced time points as an example, the matrix of factor loadings becomes, $\mathbf{\Lambda} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & \lambda_{32} \\ 1 & \lambda_{42} \end{bmatrix}$ which indicates that the last two factor loadings need to be estimated from data and the growth function is not assumed to be linear, but rather piecewise linear. In this sense, UGMM is a less restricted model in comparison with the general linear GMM.

In sum, the primary purpose of this current study is to explore the performance of a LPM and an UGMM in selecting the number of latent classes compared to a general linear GMM across different experimental conditions as described in the Methods section. As such, this study can provide some practical guidance to practitioners in their empirical study using GMM.

2.4. Evaluating the number of latent classes for mixture models

For the purpose of comparing three types of mixture models, researchers need to refer to a number of statistical tests and fit indices, although none of them is considered a universally accepted criterion. Therefore, the suggested approach in practice is to look for converging evidence across multiple criteria. All the model fit indices used in this study can be categorized into three groups, information criteria, likelihood ratio tests and classification statistics.

2.4.1 Information Criteria

Information criteria are the biggest family of indices being used for model selection in this study. All of them follow the form as

$$\mathrm{IC} = -2LL + pernalty\ \ term$$

where the *LL* is the loglikelihood of the hypothesized model and the penalty term is determined by imposing different weights on parameterizations and/or sample size. Different choices of penalty term lead to different information criteria. All those information criteria used to compare mixture models in this study are summarized in Table 2.4.1. Models with lower values indicate a better fit to the data. We need to note that three new information criteria, DBIC, HQ, and HT-AIC, were first introduced in the context of GMM study because they have been investigated for determining the number of latent classes for latent class analysis under various experimental conditions (Yang & Yang, 2007). The information criteria that penalize for model complexity (i.e., the number of parameters) might be too conservative to scrutinize the potential latent classes. This is another reason that UGMM, as a potential solution for class enumeration, is studied in addition to the complex LPM.

Table 2.4.1 *Information Criteria used in this study*

Abbreviation	Information Criteria	Function Form	Key related paper	advantages or disadvantages
AIC	Akaike's information criterion	$-2LL+2p$	Akaike (1987)	Inconsistency for not considering sample size
BIC	Bayesian information criterion	$-2LL+p\ln(N)$	Schwarz (1978)	Consistent with increasing sample size
SABIC	Sample adjusted BIC	$-2LL+p\ln((N+2)/24)$	Sclove (1987) Yang (2006)	Good when model has large p or small N.
CAIC	Consistent version of AIC	$-2LL+p[\ln(N)+1]$	Bozdogan (1987)	Favor model with fewer parameters in comparison with BIC
SACAIC	Sample size adjusted CAIC	$-2LL+p[\ln((N+2)/24)+1]$	Tofighi and Enders (2008)	Favor model with fewer parameters in comparison with SABIC
DBIC	Draper's BIC	$-2LL+p[\ln(N)-\ln 2\pi]$	Draper (1995)	Good with small to moderate sample size
HQ	Hannan and Quinn's information criteria	$-2LL+2p[\ln(\ln(N))]$	Hannan and Quinn (1979)	Good with large sample size
HT-AIC	Hurvich and Tsai's AIC	$-2LL+2p+\frac{2(p+1)(p+2)}{N-p-2}$	Hurvich and Tsai (1989)	Good with small sample size

Note: LL is the model-based log-likelihood, p is the number of parameters, and N is the sample size.

2.4.2 Likelihood Ratio Tests

Compared to information criteria, likelihood ratio tests are more demanding because these statistics require bootstrapping or following certain asymptotic distributions in order to obtain the probabilistic statement (e.g., *p* value) regarding model selection. The commonly used ordinary likelihood ratio test (OLRT) is not applicable in GMM because this test can be used only for comparing nested models and not for mixture models with different numbers of latent classes. As summarized in Table 2.4.2, three other likelihood ratio tests are used in this study.

Several things need to be clarified for Table 2.3.2. First, $f(y \mid z;\theta)$ and $g(y \mid z;\gamma)$ are conditional probability density functions for two competing models. After substituting the observed values for the endogenous variable y and exogenous variables z and estimated model parameters $\hat{\theta}$ and $\hat{\gamma}$ for the two models, the $\frac{1}{n}VLMR$ can be calculated and is distributed as a sum of chi-square distributions if the two model-based density functions are equivalent, or a weighted sum of chi-square distributions if they are not (Henson, Reise, & Kim, 2007; Vuong, 1989). Second, p_k and p_{k-1} represent the numbers of parameters in the two competing *k*-1 and *k* class models. Both of VLMR and LMR are to be compared with critical values from their theoretical distributions under the null hypothesis that the two model-based probability density functions are equivalent. Lo, Mendell, and Rubin's (2001) work indicated that VLMR exhibited more Type I errors but more power than the LMR. The significance level alpha (α) is set to be 0.05 throughout this study. And the rate of accuracy over 90/95 percent will be considered as acceptable/good.

Table 2.4.2 *Likelihood ratio tests used in this study*

Abbreviation	Likelihood ratio tests	Function Form	Key related paper	Decision rule
VLMR	Vuong-Lo-Mendell-Rubin test	$\sum_{i=1}^{n}\log\left[\frac{f(y_i \mid z_i;\hat{\theta})}{g(y_i \mid z_i;\hat{\gamma})}\right]^2$	Lo, Mendell and Rubin (2001)	A significant result indicates *k* class model is superior to the *k*-1 class model
LMR	Lo-Mendell-Rubin test	$\frac{VLMR}{1+[(p_k - p_{k-1})\ln N]^{-1}}$	Lo, Mendell and Rubin (2001)	Same as above
BLRT	Bootstrapping likelihood ratio test	NA	McLachlan (1987)	Same as above

2.4.3 Classification-based Statistics

Unlike information criteria and likelihood ratio tests, classification-based statistics include the consideration of classification accuracy. After estimation of a mixture model, the chance of individuals arising from each latent class is measured by the estimated posterior probabilities. If each subject has a single high posterior probability for a certain class, this means the classification is unambiguous. Although this type of statistics can not be used as absolute fit indices because some mixture models *per se* have overlapping components, leading to ambiguous classification result, they could be used as comparative fit indices between models if the purpose is to select one out of several models that fit data equally well. Based on the previous summary (Henson et al., 2007; McLachlan & Peel, 2000), four classification-based statistics listed in Table 2.3.3 will be investigated in this study.

Table 2.4.3 *Classification-based statistics used in this study*

Abbreviation	Classification-based statistics	Function Form	Key related paper	Decision rule
NEC	Normalized entropy criterion	$\frac{E(k)}{LL(k)-LL(1)}$	Celeus and Soromenho (1996)	Close to 0 indicates better model fit
Entropy	Entropy	$1-\frac{E(k)}{N*\ln(k)}$	Ramaswamy, DeSarbo, Reibstein, and Robinson (1993); Lubke and Muthén (2007)	Close to 1 indicates better model fit; 0.6 indicates 80 percent or less accurate classification while 0.8 support 90 percent
CLC	Classification likelihood information criterion	$-2LL+2E(k)$	McLachlan and Peel (2000)	Lower value indicates better model fit
ICL-BIC	Integrated classification likelihood	$-2LL+2E(k)+p\mathrm{l}$	McLachlan and Peel (2000)	Lower value indicates better model fit

In Table 2.4.3, $E(k) = -\sum_{k=1}^{K}\sum_{i=1}^{N}\pi_i^k \ln(\pi_i^k) \geq 0$, π_i^k is the posterior probability that subject *i* belongs to latent class *k*, and $LL(k)$ and $LL(1)$ are the model maximum likelihoods for *k* class model and 1 class model (i.e., no mixture), respectively.

It is noteworthy that all the statistical indicators as introduced so far provide only a relative fit of competing models to data. Stated differently, we can infer that one model is better than another from these criteria or tests, but we are uncertain if this model is good enough to fit the observed data. Muthén (2003) tried to overcome this limitation by proposing the Multivariate Skewness Test (MST) and the Multivariate Kurtosis Test (MKT) for testing mixture models, analogous to the goodness of fit tests for structural equation models. A larger probability value (e.g., ≥ 0.05) means adequate model fit. However, Tofighi and Enders' (2008) simulation results implied MST and MKT perform poorly across all the experimental conditions they examined for GMM. They concluded that these two indices are model-dependent, at least in a GMM context. For these reasons, MST and MKT are not investigated in this study.

2.4.4 Previous studies of comparing relative model fit statistics

Only a few simulation studies examined the relative efficiency of the statistical indicators for class enumeration in a GMM context (Nylund, Asparouhov, & Muthén, 2007; Tofighi & Enders, 2008; Tolvanen, 2008). Tofighi and Enders' comprehensive simulation study recommended the SABIC and the LMR test in selecting the number of classes for GMM. Nylund et al. (2007), on the other hand, found that BLRT outperformed the other indices and that BIC was the most consistent information criterion among those considered. Henson et al. (2007) recommended using SABIC with latent variable mixture models but they found that no indices performed well when sample sizes were below 500. Tolvanen (2008) investigated the functionality of GMM with a limited sample size. His simulation results suggested BIC was more useful when the sample size was smaller than 500, whereas SABIC performed better when the sample size was larger than 500. These results are somewhat inconsistent or cover only some of the statistical indices aforementioned. While the current study is expected to shed some light on the relative efficiencies of a wider range of model fit indicators; the comparison of model fit indices is not this study's primary focus.

CHAPTER 3: METHOD

This simulation study investigates if and under what conditions LPM and UGMM can perform better than linear GMM in determining the number of latent classes. Data were generated from a GMM with model parameters specified a *priori* and then analyzed by GMM, LPM, and UGMM separately. By repeating this analysis within each model setting a large number of times, we can make an inference concerning the relative performance of these three types of models in accurately enumerating the latent classes for GMM.

3.1 Data generation

All sample data were simulated from a 2-class GMM population model in SAS IML. This population generating model is graphically depicted in Figure 3.1. Both the graph and the previous two-level equations indicate that no covariate is included in our study. The parameter values for this model are shown in Table 3.1.1, and include those of the Nylund et al. (2007) study for purposes of replication.

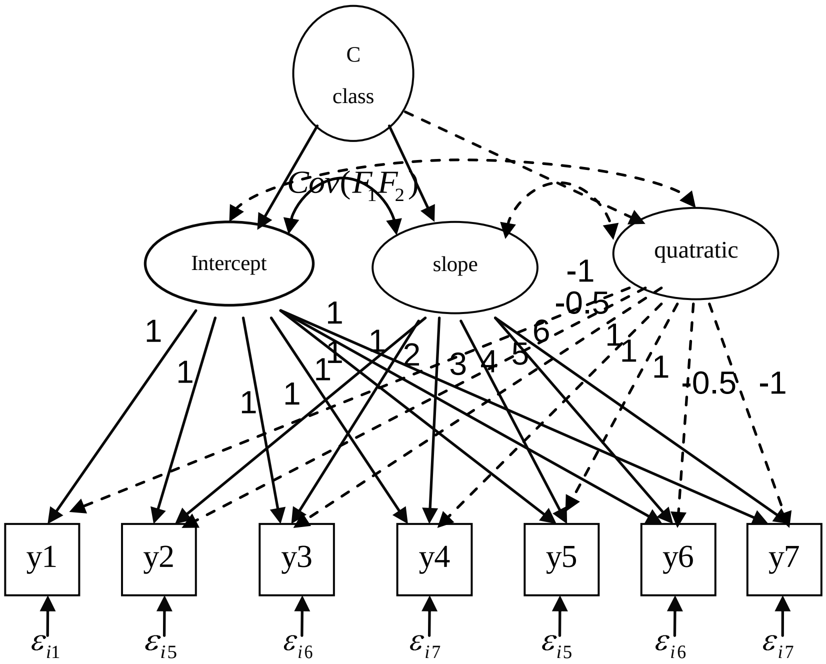

Figure 3.1. *Path diagram of the population growth mixture model used for data generation (Note: dashed lines indicate nonlinear components added into the misspecified model only)*

Table 3.1.1 Population growth mixture model specification

	Class 1		Class 2	
	mean	var	mean	var
Intercept(η_{0i}^{k})	2	0.25	1	0.25
Slope(η_{1i}^{k})	0.5	0.04	0	0.04
Quadratic(η_{2i}^{k})[a]	0.12	0.0016	-	-
Residual1:var(ε_{i1}^{k})	0	0.15	0	0.15
Residual2:var(ε_{i2}^{k})[b]	0	0.15	0	0.15
Residual3:var(ε_{i3}^{k})	0	0.2	0	0.2
Residual4:var(ε_{i4}^{k})[b]	0	0.2	0	0.2
Residual5:var(ε_{i5}^{k})	0	0.2	0	0.2
Residual6:var (ε_{i6}^{k})[b]	0	0.35	0	0.35
Residual7:var(ε_{i7}^{k})	0	0.35	0	0.35

[a] for misspecified model only [b] excluded for 4-measures model

During the process of data generation, five factors are manipulated in the 2×2×4×2×2 simulation design according to their potential impact and practical implications on class enumeration.

First, to examine if the LPM and UGMM outperform GMM in selecting the correct number of latent classes, both the properly and improperly specified population GMM were used to generate sample data. A quadratic term was added into the majority latent class in the population linear GMM; due to its small quantity (almost one-fifth of the slope and one- twentieth of the intercept), this subtle nonlinearity can not be detected by visual inspection of a spaghetti plot (i.e., trend line) of the sample data. As such, it is highly possible this growth pattern would be considered linear during estimation. Moreover, LPM and UGMM are still technically correct models since they do not assume a linear growth function, whereas the linear GMM is not the correct model. It is worth to emphasize that the inclusion of nonlinear component is just one type of misspecifying within-class model. Indeed, there are other possibilities for model misspecification, such as correlated error variance-covariance structure within a class.

Second, the number of repeated measures includes two levels, 4 and 7. Models with four measurement points are relatively simple and often seen in applications of LGM and GMM (Tolvanen, 2008). Including the condition of seven measurement occasions can accomplish two goals: 1) to clearly differentiate the effect of the number of repeated measures on the class enumeration and 2) to make the construction of the four measurement cases more convenient. The factor loadings t_j (i.e., the time variable) in the simpler model can take the values of 0, 2, 4, and 6, based on the more complex model with factor loadings ranging from the integers 0 to 6 (Tofighi & Enders, 2008).

Third, the total sample size was varied on values of 400, 700, 1000, and 2000. This factor takes these values according to a careful review of substantive GMM applications in Tofighi and Enders (2008). Hence, the results of our study can provide some guidelines for practitioners.

Fourth, class mixing proportions were 50/50 and 75/25. Two different mixing percentages of classes were chosen for their important influence on classification results in mixture models. Usually a model with a balanced mixing proportion performs better in enumerating the correct number of latent classes. To replicate the Nylund et al. (2007) study, we choose these two conditions.

Fifth, class separations along the intercept factor were chosen to be 2 and 3 standard deviations (SD) separately. Tofighi and Enders (2008) used approximately two and three SD between the latent intercept means representing “high separation” and “low separation” between classes. Nylund et al. (2007) only examined the condition of a two SD difference between intercept means. So class separation of two and three SD along the intercept factor is chosen to replicate their findings. This setting of class separation is equal to 3.5 and 5 squared Mahalanobis distance (a measure for the separation of two groups of objects) units, respectively, between the latent components of two latent classes according to the equation given by McLachlan (1999). This measure is defined by

$$\Delta^2 = (\alpha_1 - \alpha_2)^T \Psi^{-1} (\alpha_1 - \alpha_2)$$

Where superscript T denotes matrix transpose, Ψ denotes the common (nonsingular) covariance matrix for the two groups, and α_1 and α_2 are mean vectors of latent components for these two groups. If the Mahalanobis distance is measured at observable variable scale, it can be defined similarly as

$$d^2 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^T S^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$$

Where $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_2$ are mean vectors for the indicator variables of two groups and S is the pooled covariance matrix for the two groups of indicators. S equals to $\pi_1 S_1 + \pi_2 S_2$, in which π_1 and π_2 are mixing proportions for the two groups and S_1 and S_2 are group-based covariance matrices. This measure varies across the manipulated conditions: for two SD separation conditions, the squared Mahalanobis distance ranges from 2.9 to 3.4 with an average of 3.1; for three SD separation conditions, the measure ranges from 3.5 to 4.1 with an average of 3.7.

Only five factors are varied in the simulation design while others are held constant. As Table 3.2 shows, the full factorial design contains a total of 64 conditions, making it more complete than either of the two key preceding studies focusing on fit index performance (Nylund et al., 2007; Tofighi & Enders, 2008). For each condition, 100 replications were conducted to obtain a reliable result, just as Nylund et al. (2007) did. Hence, 6400 sample data sets were generated in total.

Table 3.1.2 *Simulation design*

Conditions	Manipulated factors				
	Class separation	Sample size	# of measures	mixing prop	model specification
1	2	2000	4	50/50	correct
2	2	2000	4	50/50	incorrect
3	2	2000	4	75/25	correct
4	2	2000	4	75/25	incorrect
5	2	2000	7	50/50	correct
6	2	2000	7	50/50	incorrect
7	2	2000	7	75/25	correct
8	2	2000	7	75/25	incorrect
9	2	1000	4	50/50	correct
10	2	1000	4	50/50	incorrect
11	2	1000	4	75/25	correct
12	2	1000	4	75/25	incorrect
13	2	1000	7	50/50	correct
14	2	1000	7	50/50	incorrect
15	2	1000	7	75/25	correct
16	2	1000	7	75/25	incorrect
17	2	700	4	50/50	correct
18	2	700	4	50/50	incorrect
19	2	700	4	75/25	correct
20	2	700	4	75/25	incorrect
21	2	700	7	50/50	correct
22	2	700	7	50/50	incorrect
23	2	700	7	75/25	correct
24	2	700	7	75/25	incorrect
25	2	400	4	50/50	correct
26	2	400	4	50/50	incorrect
27	2	400	4	75/25	correct
28	2	400	4	75/25	incorrect
29	2	400	7	50/50	correct
30	2	400	7	50/50	incorrect
31	2	400	7	75/25	correct
32	2	400	7	75/25	incorrect
33	3	2000	4	50/50	correct
34	3	2000	4	50/50	incorrect
35	3	2000	4	75/25	correct
36	3	2000	4	75/25	incorrect
37	3	2000	7	50/50	correct
38	3	2000	7	50/50	incorrect
39	3	2000	7	75/25	correct
40	3	2000	7	75/25	incorrect
41	3	1000	4	50/50	correct
42	3	1000	4	50/50	incorrect
43	3	1000	4	75/25	correct

44	3	1000	4	75/25	incorrect
45	3	1000	7	50/50	correct
46	3	1000	7	50/50	incorrect
47	3	1000	7	75/25	correct
48	3	1000	7	75/25	incorrect
49	3	700	4	50/50	correct
50	3	700	4	50/50	incorrect
51	3	700	4	75/25	correct
52	3	700	4	75/25	incorrect
53	3	700	7	50/50	correct
54	3	700	7	50/50	incorrect
55	3	700	7	75/25	correct
56	3	700	7	75/25	incorrect
57	3	400	4	50/50	correct
58	3	400	4	50/50	incorrect
59	3	400	4	75/25	correct
60	3	400	4	75/25	incorrect
61	3	400	7	50/50	correct
62	3	400	7	50/50	incorrect
63	3	400	7	75/25	correct
64	3	400	7	75/25	incorrect

3.2 Model Estimation

Three different mixture models with 1, 2, and 3 latent classes were used separately to analyze the 6400 data sets in Mplus Version 6 (Muthén & Muthén, 2008): LPM, UGMM, and a linear GMM. When the data set is generated from the population GMM without a quadratic term, all estimated mixture models have the correct within-class structure and their differences lie in their parameterizations; when the data are generated from a model with a quadratic term, LPM and UGMM still have technically correct within-class model specification while the linear GMM is not correct in the sense that it ignores the nonlinear relations underlying the data.

Estimation was carried out by using ML via an EM algorithm in Mplus. The default convergence criterion of complete-data log likelihood derivative for the EM algorithm is 0.001. For each of these mixture models, one-, two-, and three-class models were evaluated (i.e., under-extraction, proper extraction, and over-extraction). All parameters were allowed to be class-specific, so no cross-class model constraints were involved for any model. Note that properly specified linear GMMs had no quadratic component in the data for either class; misspecified models had a quadratic component in the data for the first class only. Finally, multiple sets of random start values were implemented in Mplus to avoid the irregularities on the likelihood surface and to differentiate local maxima from the global optimum for estimation of mixture models (e.g., McLachlan & Peel, 2000; Muthén & Muthén, 2001).

CHAPTER 4: RESULTS

Analyses for the total 64 conditions are summarized separately in Table A1 through Table A64 in the appendix. Note that all the 1-class and 2-class models converged properly; and it is not surprising to find nonconvergence did occur in some replications of estimating the 3-class mixture models since they are misspecified models (e.g., Nylund et al., 2007). One option is to simply discard these failed replications and summarize the results that providing a proper solution for the mixture models; the other is to treat nonconvergent replications in GMM as an indicator of model misfit and also evidence to support model with one fewer classes (Nylund et al., 2007; Tofighi & Enders, 2008). In following analysis, both ways are used to present the results.

Results are summarized in three parts. First, the general performance of the three types of mixture models and eleven model selection indices are presented. Second, the general effects of the manipulated factors on class enumeration are examined. Finally, the significant interaction effects among those factors in a given type of mixture model are also explored.

4.1 General Performance of Types of Mixture Models and Model Fit Indices

As stated before, nonconvergence is a problem for misspecified three-class mixture models. Among the three types of three-class mixture models, UGMM has the best convergence rate (95 out of 100 replications) while linear GMM has the worst (67 out of 100 replications) in this regard. As introduced above, two different ways of dealing with nonconvergent replications were used in Table

4.1.1 and Table 4.1.2 respectively, based on which the general performance of types of mixture models and model fit indices are summarized.

Table 4.1.1 provides the frequency summary of the number of latent class selected by each model fit index in three different types of mixture models, averaged over all the 64 manipulated conditions. Nonconvergent replications are treated as evidence for supporting two-class models because nonconvergence is assumed to be caused by misspecified three-class models. Thus Table 4.1.1 presents frequency information based on all the 100 replications. Moreover, the log likelihood derivative convergence criterion for the EM algorithm in Mplus is changed from the default value of .001 to .01 for some nonconvergent replications (not all of them due to time constraints) to see whether they could get converged. Unfortunately, the replications that had been re-examined still did not converge properly. However, if more efficient algorithm rather than those in Mplus were used in the future, it is possible that these nonconvergent replications might converge properly then and consequently some of them might not support 2-class model and the above assumption might not be valid.

Differently from Table 4.1.1, Table 4.1.2 excluded the nonconvergent replications and summarizes the percentage result based on convergent ones. Each cell frequency is divided by the total number of convergent replications for the same index within the same model. However, this method might be criticized that it rules out those data space for the nonconvergent cases, based on which the inference might be misleading.

Clearly, each method has its justification and flaw. Both are used to explore whether less restricted mixture models can more accurately identify the number of latent classes.

Table 4.1.1 Average Frequency of each class selected by each index for all the 64 conditions for all the replications (nonconvergent replications were included).

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (77 converged replications for 3-class model)													
1 class	0	26	7	21	2	9	6	0	-	5	-	4	-
2 class	24	72	91	79	84	90	89	28	28	95	81	96	55
3 class	76	2	1	0	14	1	5	72	72	-	19	-	45
UGMM (95 converged replications for 3-class model)													
1 class	0	15	1	9	0	2	1	0	-	2	-	2	-
2 class	**55**	85	**98**	91	**90**	**97**	**92**	**58**	23	98	**89**	**98**	**87**
3 class	45	0	2	0	10	1	7	42	77	-	11	-	13
Linear GMM (67 converged replications for 3-class model)													
1 class	0	5	0	3	0	0	0	0	-	0	-	2	-
2 class	36	**94**	87	**97**	62	93	71	37	**37**	**100**	74	**98**	**87**
3 class	64	0	13	0	38	7	29	63	63	-	26	-	13

Table 4.1.2 Average Percent of each class selected by each index for all the 64 conditions for all the replications (nonconvergent replications were excluded).

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (77 converged replications for 3-class model)													
1 class	0	36	10	29	2	13	8	0	-	5	-	4	-
2 class	1	62	88	70	77	85	85	7	7	95	74	96	41
3 class	99	2	2	1	22	1	7	93	93	-	26	-	59
UGMM (95 converged replications for 3-class model)													
1 class	0	16	1	10	0	2	1	0	-	2	-	2	-
2 class	**52**	84	**97**	90	**90**	**97**	**92**	**54**	**19**	98	**88**	**98**	**86**
3 class	48	0	2	0	10	1	7	46	81	-	12	-	14
Linear GMM (67 converged replications for 3-class model)													
1 class	0	7	0	4	0	0	0	0	-	0	-	2	-
2 class	3	**93**	81	**96**	43	90	54	4	5	**100**	59	**98**	81
3 class	97	0	19	0	57	10	46	96	95	-	41	-	19

Note: The highest frequency/percent selected by each index among the three types of mixture models are highlighted as bolded.

4.1.1. Comparison of three types of mixture models

In Table 4.1.1 and Table 4.1.2, three types of mixture model performances are compared in terms of the fit indices and those indices having highest frequency/probability of correct model selection among three models are highlighted in bold. Clearly, these two tables have almost identical pattern and very close values, making the results more valid because they do not reply on how to deal with nonconvergent replications. All the highest frequency/probabilities are clustered into UGMM and linear GMM. UGMM performs best in terms of most of the model fit indices we used.

More specifically, AIC, SACAIC, SABIC, DBIC, HQ, HT-AIC, LMR LRT (2 class versus 3 class) and BLRT within the UGMM perform best in selecting the correct 2-class model. Moreover, CAIC, BIC, Entropy, LMR LRT (1 class versus 2 class), and BLRT perform better in linear GMM than in UGMM. But they have the same or similar values in linear GMM and UGMM. All these findings support the hypothesis that less restricted models can more accurately identify the number of latent classes.

However, as an unrestricted mixture model assuming no specific within-class relations among variables, LPM does not outperform the linear GMM and UGMM on average (although LPM has close frequency values to the other two models in terms of some fit indices). This indicates that a completely unrestricted model might not win in this situation due to its over-parameterization (i.e., too many parameters to be estimated as shown in Table 2.3.2). As Table 2.4.1 presents, the number of parameters is a penalty component in the functions of all the information criteria, some of which put much weight on the number of parameters. Therefore, it is understandable that over-parameterization of LPM makes it less effective in class enumeration using these information criteria.

4.1.2. Comparison of model fit indices

All of the information criteria, likelihood ratio tests, and classification-based statistics previously introduced were included for the purpose of identifying the correct number of classes. Among the three different groups of model fit indices, we found all four classification-based statistics exhibited very limited utility with a low rate of accuracy in class determination. This is consistent with previous studies (e.g., Henson et al., 2007), thus entropy is retained as a representative classification measure, while the likelihood ratio tests and information criteria are used for the remainder of this work. Moreover, the performance of the LMR and VLMR are almost identical, with a difference of no more than 1 and therefore only LMR was presented in the tables.

An examination of Table 4.1.1 and Table 4.1.2 yields the similar general performance of those fit indices in class identification:

- Entropy and other classification-based statistics do not seem to be very useful indices as they tend to overestimate the number of classes for all the mixture models across all the cell conditions examined. So they are not recommended to determine the number of latent classes for mixture models.
- AIC and HT-AIC tend to overextract the number of latent classes with an

unacceptably low rate of accuracy across the three types of mixture models, which is consistent with previous published research (e.g., Nylund et al., 2007) and so they are not recommended for class enumeration in mixture modeling. Only in UGMM, both of them have more than 50% of chance to correctly select the two-class model.

- LMR and BLRT are sufficiently accurate when testing a 2-class versus a 1-class model across all the models and all the conditions. However, both are less accurate when testing the 2-class model against the 3-class model. BLRT (2 vs.3) has inflated Type I error rate up to .45 (.59 if excluding nonconvergent replications) in LPM. Both of the two likelihood ratio tests perform best in UGMM with Type I error rate of around .11 and .13 separately.
- CAIC and BIC have very similar patterns. Both tend to underestimates the number of latent classes in three types of mixture models. Both perform best in linear GMM and least in LPM. Generally speaking, BIC has higher rate of accuracy than CAIC. Given the fact that CAIC and BIC have the largest penalty terms for the number of parameters among all the indices, which make them tend to favor simple models over complex ones, it is understandable why they more often select the 1- or 2-class models over 3-class ones. This is consistent with previous studies (Hurvich & Tsai, 1989; Nylund et al., 2007).
- SACAIC and DBIC are almost perfect model selectors in UGMM because of

their highest probabilities of selecting 2-class models. Both of them work best in UGMM, slightly underestimate the number of latent classes in LPM and slightly overestimate in linear GMM across all the cell conditions.

- SABIC and HQ have very similar patterns. Both work best in UGMM and worse in linear GMM. In that sense, they favor less restricted models. Both tend to more often overestimate the number of latent classes, which is particularly true in linear GMM. HQ slightly outperform SABIC since it has higher rate of accuracy in all the three types of mixture models.

All these observations are briefly summarized in Table 4.1.2.1.

Table 4.1.2.1 *Usability of fit indices in determining the number of latent classes for GMM*

Model fit indices	recommendation	reason
classification-based statistics, HT-AIC and AIC	No	Likely to overestimate
BLRT and LMR LRT	Definitely yes	Sufficient power when testing 2- VS. 1-class model; Inflated type I error when testing 2- VS. 3-class model; both work best in less restricted UGMM
CAIC and BIC	Yes	BIC performs better than CAIC; tend to underestimate; both work best in most restricted model and have similar pattern
SACAIC and DBIC	Definitely yes	Almost perfect model selector in UGMM; both slightly underestimate in LPM and overestimate in linear GMM.
SABIC and HQ	Yes	HQ performs slightly better than SABIC; both work best in UGMM and worst in linear GMM; both tend to overestimate, especially in linear GMM

Table 4.2 One way ANOVA for the effect of design factors on model fit indices in selecting the true model across types of models and conditions.

Factors		AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT_AIC	Entropy	LMR_1V2	LMR_2V3	BLRT_1V2	BLRT_2V3
Class separation	F	0.05	20.27	5.50	20.71	0.71	10.76	0.95	0.04	1.20	10.57	0.75	13.50	0.64
	Sig.	0.81	0.00	0.02	0.00	0.40	0.00	0.33	0.85	0.27	0.00	0.39	0.00	0.43
	Eta squared	0.00	**0.10**	0.03	**0.10**	0.00	0.05	0.00	0.00	0.01	0.05	0.00	0.07	0.00
Sample size	F	2.00	18.45	21.64	14.81	24.01	15.40	1.44	2.64	2.42	14.05	1.49	16.57	0.92
	Sig.	0.12	0.00	0.00	0.00	0.00	0.00	0.23	0.05	0.07	0.00	0.22	0.00	0.43
	Eta squared	0.03	**0.23**	**0.26**	**0.19**	**0.28**	**0.20**	0.02	0.04	0.04	**0.18**	0.02	**0.21**	0.01
# measures	F	10.00	26.89	3.58	23.37	13.21	10.23	1.49	12.23	63.18	5.87	100.88	8.69	3.66
	Sig.	0.00	0.00	0.06	0.00	0.00	0.00	0.22	0.00	0.00	0.02	0.00	0.00	0.06
	Eta squared	0.05	**0.12**	0.02	**0.11**	0.07	0.05	0.01	0.06	**0.25**	0.03	**0.35**	0.04	0.02
Mixing proportion	F	0.21	1.10	0.90	0.43	0.15	0.63	0.47	0.05	0.24	1.00	0.30	0.90	0.18
	Sig.	0.65	0.29	0.34	0.51	0.70	0.43	0.49	0.82	0.63	0.32	0.59	0.34	0.67
	Eta squared	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Model specification	F	2.05	0.51	0.14	0.10	0.94	0.00	2.20	1.53	0.37	0.90	4.13	1.54	3.02
	Sig.	0.15	0.48	0.70	0.75	0.33	0.97	0.14	0.22	0.54	0.34	0.04	0.22	0.08
	Eta squared	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.02	0.01	0.02

4.2 The effect of design factors on class enumeration

Inspecting Table 4.2, in terms of Eta squared, which is a commonly used measure for effect size, value above 0.1 is considered practically significant throughout this study, three factors have practically significant effect on the accuracy of several model fit indices in selecting the correct two-class models across all three types of mixture models and sixty-four simulated conditions. They are class separations, sample size, and the number of repeated measures.

In this section, each manipulated factor is examined in terms of their impact on the accuracy of class determination, given the type of mixture models. Moreover, the practically significant interaction effect between the factors and the types of models are also displayed graphically and interpreted.

4.2.1 Class separation

Table 4.2.1.1(a & b) and Table 4.2.1.2(a & b) present the frequency/percent summary for the two different class separation conditions, two- and three-standard deviation differences between the two class-specific intercept means separately. Likewise, comparing three types of mixture models in terms of each model fit index, the two-class models with the highest chance of being selected are highlighted in bold. By means of visual inspection, it is clear that these two groups of tables have similar patterns with Table 4.1.1 and Table 4.1.2. Therefore, the previous observations regarding model fit indices can also be applied here.

Inspecting the two groups of tables, generally speaking, increasing the difference of latent intercept means directly lowers the chance of selecting the one-class model dramatically. This is particularly true in linear GMM, in which a one-class model is not chosen at all. This observation makes sense in that the larger class separation increases the power to detect the second class and thus reject the one-class model. Due to this reason, larger class separation increases the probability of selecting the correct two-class model for most of the fit indices.

However, there are a few exceptions in certain types of mixture models. First, AIC, SABIC, HT-AIC, and Entropy tend more often to overestimate the number of latent classes in models with larger class separation and so the probability for selecting two-class models decreases. Second, SACAIC and HQ select more three-class models in linear GMM. Third, the larger class separation does not help LMR and BLRT select two-class models over three-class ones. All of these exceptional indices share a common property that they have sufficient power to reject one-class models and tend to overestimate the number of latent class in the smaller class separation condition. That is to say, two SD class separation condition is enough to differentiate two different groups. As such, larger three SD class separation condition does not help separating the true two latent classes and would make overestimation even worse.

Furthermore, the statistically significant interaction effect between the types of models and class separation of four model fit indices, SACAIC, DBIC, HQ, and LMR_1V2, are examined and graphically displayed in Figure 4.2.1.1. But they are not practically significant in terms of the criterion of partial Eta squared value of 0.1. Their corresponding values are .06, .06, .04, and .03. The dashed black line and the solid red line represent the performances of fit index in two SD and three SD condition across types of models separately. As blue arrow shows, the larger class separation effect is most evident in LPM because the accuracy rate dramatically goes up as class separation increases. The class separation effect is least distinct in linear GMM. And on the contrary, SACAIC and HQ imply that larger class separation would slightly lower the accuracy rate in linear GMM. As the shaded circles show, the four indices perform best in UGMM, generally much better than in linear GMM.

Table 4.2.1.1a Average Frequency of each class selected by each index for 32 conditions with 2 SD class separations (nonconvergent replications are included)

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (75 converged replications for 3-class model)													
1 class	0	35	13	31	2	17	10	0	-	9	-	7	-
2 class	26	64	85	68	86	82	85	30	30	91	82	93	57
3 class	74	1	1	1	12	1	5	70	70	-	18	-	43
UGMM (94 converged replications for 3-class model)													
1 class	0	27	2	18	0	3	2	0	-	3	-	4	0
2 class	**58**	72	**97**	82	**91**	**96**	**92**	**60**	22	97	**89**	96	87
3 class	42	0	2	0	9	1	7	40	78	-	11	-	13
Linear GMM (71 converged replications for 3-class model)													
1 class	0	11	0	6	0	1	0	0	-	1	-	3	0
2 class	33	**89**	88	**93**	63	93	72	34	**33**	**99**	74	**97**	**88**
3 class	67	1	12	1	37	7	27	66	67	-	26	-	12

Table 4.2.1.1b Average Frequency of each class selected by each index for 32 conditions with 3 SD class separations (nonconvergent replications are included)

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (78 converged replications for 3-class model)													
1 class	0	17	1	11	0	2	1	0	-	1	-	1	0
2 class	22	81	**98**	89	83	97	93	27	27	99	80	100	53
3 class	78	3	1	0	17	0	6	73	73	-	20	-	47
UGMM (95 converged replications for 3-class model)													
1 class	0	3	0	1	0	0	0	0	-	0	-	0	0
2 class	**53**	97	**98**	99	**90**	**99**	**93**	**56**	24	**100**	**88**	**100**	**87**
3 class	47	0	2	0	10	1	7	44	76	-	12	-	13
Linear GMM (62 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	0
2 class	39	**100**	86	**100**	61	93	69	40	**41**	**100**	74	**100**	85
3 class	61	0	14	0	39	7	31	60	59	-	26	-	15

Table 4.2.1.2a Average percent of each class selected by each index for 32 conditions with 2 SD class separations (nonconvergent replications are excluded)

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (75 converged replications for 3-class model)													
1 class	0	50	19	44	3	23	15	0	-	9	-	8	-
2 class	2	49	79	55	77	75	79	7	7	91	75	92	42
3 class	98	1	3	1	20	2	7	93	93	-	25	-	58
UGMM (94 converged replications for 3-class model)													
1 class	0	29	2	19	0	3	2	0	-	3	-	4	-
2 class	**54**	71	**97**	81	**90**	**96**	**91**	**57**	**18**	97	**88**	96	**86**
3 class	46	0	2	0	10	1	7	43	82	-	12	-	14
Linear GMM (71 converged replications for 3-class model)													
1 class	0	14	0	7	0	1	0	0	-	1	-	3	-
2 class	5	**86**	84	**92**	49	90	60	6	5	**99**	63	**97**	84
3 class	95	1	16	1	51	9	40	94	95	-	37	-	16

Table 4.2.1.2b Average percent of each class selected by each index for 32 conditions with 3 SD class separations (nonconvergent replications are excluded)

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (78 converged replications for 3-class model)													
1 class	0	22	2	15	0	3	2	0	-	1	-	1	-
2 class	0	75	97	85	76	96	91	6	6	99	73	100	40
3 class	100	3	2	0	24	1	8	94	94	-	27	-	60
UGMM (95 converged replications for 3-class model)													
1 class	0	3	0	1	0	0	0	0	-	0	-	0	-
2 class	**49**	97	**98**	99	**89**	**99**	**92**	**52**	**20**	**100**	**87**	**100**	**86**
3 class	51	0	2	0	11	1	8	48	80	-	13	-	14
Linear GMM (62 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	2	**100**	79	**100**	37	89	48	2	5	**100**	55	**100**	78
3 class	98	0	21	0	63	11	52	98	95	-	45	-	22

Table 4.2.1.3. One way *ANOVA results for the frequency difference of model fit indices between two class separation conditions.*

		AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT_AIC	Entropy	LMR_1V2	LMR_2V3	BLRT_1V2	BLRT_2V3
LPM	F	1.49	4.33	9.11	8.15	0.48	9.87	4.64	0.71	0.44	5.81	1.52	5.11	0.00
	Sig.	0.23	0.04	0.00	0.01	0.49	0.00	0.04	0.40	0.51	0.02	0.22	0.03	0.96
	Eta Squared	0.00	0.08	**0.13**	0.00	**0.11**	0.01	0.00	0.09	0.09	**0.15**	**0.11**	0.00	0.07
UGMM	F	0.30	14.00	2.30	11.00	0.24	4.98	0.33	0.33	0.15	7.85	0.14	5.11	0.00
	Sig.	0.59	0.00	0.13	0.00	0.63	0.03	0.57	0.56	0.70	0.01	0.71	0.03	0.96
	Eta Squared	0.00	**0.18**	0.04	**0.15**	0.00	0.07	0.01	0.01	0.00	**0.11**	0.00	0.08	0.00
Linear GMM	F	5.24	9.05	0.16	7.65	0.47	0.01	5.83	5.88	10.84	7.52	0.08	4.46	1.67
	Sig.	0.03	0.00	0.69	0.01	0.49	0.92	0.02	0.02	0.00	0.01	0.78	0.04	0.20
	Eta Squared	0.02	0.07	**0.13**	**0.12**	0.01	**0.14**	0.07	0.01	0.01	0.09	0.02	0.08	0.03

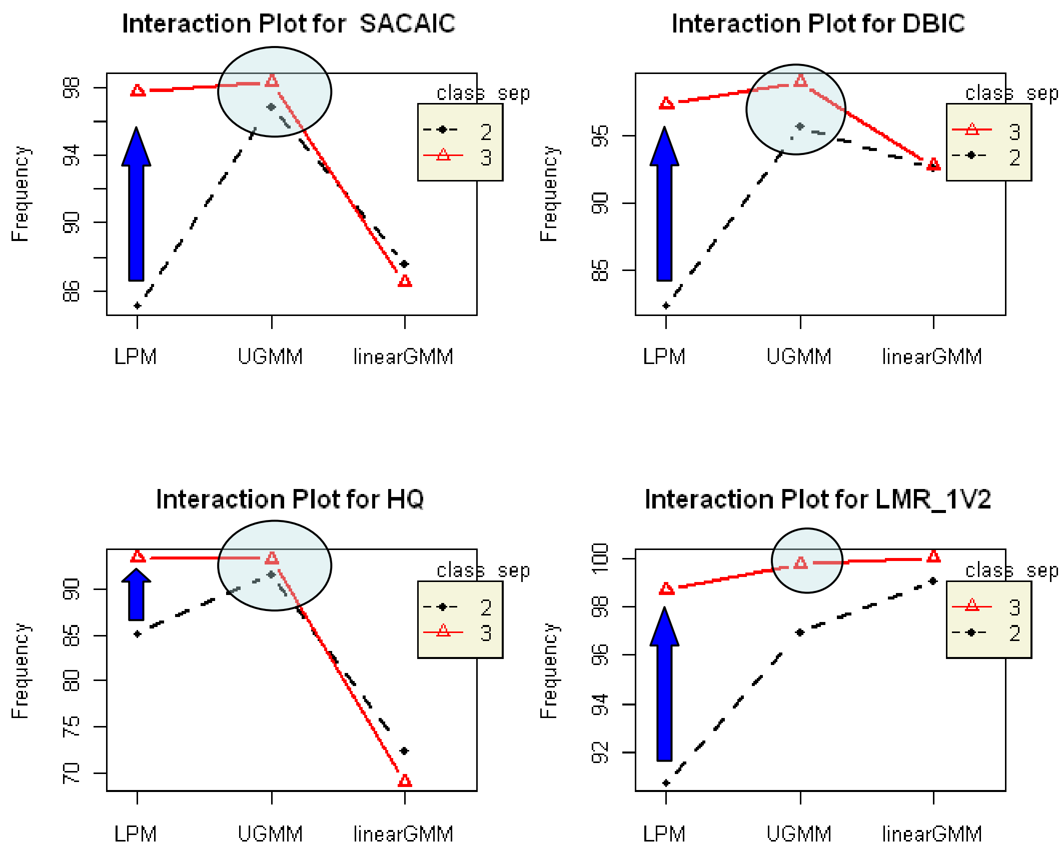

Figure 4.2.1.1 *Model fit indices with significant interaction effects between the types of models and class separations*

4.2.2 Sample size

In terms of two different ways of handling the nonconvergent replications, Table 4.2.2.1 (a through d) to Table 4.2.2.1 (a through d) present the frequency/percent summary under conditions of four different sample sizes.

Table 4.2.2.1 and Table 4.2.2.2 (a through d) indicate UGMM has a quite stable convergence rate, roughly around 95 out of 100. As expected, the convergence rate for LPM is lowest (62) at the smallest sample size of 400 and remains almost the same around 80 at or above sample size of 700. When sample size is sufficiently large (e.g., 700 in this case), the nonconvergence rate of 20% is highly possible to be caused by misspecified three-class models. As for linear GMM, a sample size of 400 is generally considered enough for model estimation. Increasing sample size provides more power to detect that the three-class model specification is not appropriate, which explains why the lowest convergence rate occurred in the case of 2000 sample size.

The ANOVA test result in Table 4.2.2.3 shows sample size has a significant impact on all the model selectors in certain model contexts. And based on two groups of tables with quite similar patterns, several conclusions regarding the impact of sample size on the performance of model fit indices can be drawn as below.

First, increasing sample size does not improve the accuracy of AIC and HT-AIC in identifying the number of latent classes. In fact, larger sample size shows a lower rate of accuracy, especially in LPM and UGMM. Moreover, the rates of selecting 2-class models for these two fit indices are unacceptably low (all less than 60 out of 100) so that AIC and HT-AIC are not suggested for the purpose of class enumeration.

Second, large sample size has a positive impact on the accuracy rate of CAIC, SACAIC, BIC, SABIC, and DBIC in all the three types of mixture models. That is to say, within each type of mixture model, increasing sample size could improve the performance of these fit indices in enumerating the correct 2-class models. CAIC reaches a satisfactory rate of accuracy in linear GMM when sample size is over 700; it needs 1000 to achieve a satisfactory rate in UGMM and 2000 in LPM. SACAIC and DBIC can achieve a satisfactory rate of accuracy with sample size of 400 and 700 separately in UGMM, but need 1,000 subjects in linear GMM and LPM to have the rate of accuracy over 95%. As for BIC, 700 is enough to reach the rate of accuracy over 95% in linear GMM and UGMM while it requires more, such as 2,000, to obtain a satisfactory rate in LPM. SABIC has acceptable rate of accuracy (over 90%) in UGMM when sample size is 700 and it needs 1000 to obtain the rate over 90% in LPM. Based on our data, SABIC only reaches the satisfactory rate of accuracy with the largest sample size 2,000.

Third, the relation between sample size and HQ's performance is not consistent. HQ has a satisfactory rate of accuracy in UGMM with a sample size 400 and 700, but it performs slightly worse when sample size increases to 1,000 and much worse at 2,000. As sample size increases from 400 to 1000 it performs better in LPM, but it tends to be worse with a sample size of 2,000. Therefore, this index does not have a clear asymptotic feature in this regard.

Fourth, the likelihood ratio tests LMR and BLRT exhibit clear asymptotic behavior when testing one-class versus two-class models (i.e., they tend to select two-class models as sample size increases). Both of them have sufficient power to reject a 1-class model with the smallest sample size of 400 in UGMM and linear GMM. When sample size reaches 700, both indices have over 95% of chance to make a correct decision regarding class determination in all the three types of mixture models. However, when testing three-class models against two-class models, both LMR and BLRT perform best and relatively stable in UGMM, but with a growing Type I error rate as sample size increases from 400 to 1000.

In a summary, it is not surprising to find that increasing sample size does help most fit indices more accurately identify the number of latent classes. But there are some exceptional cases; sample size does not improve the performance of AIC, HQ, and Entropy because their functions either remove or limit the effect of sample size: AIC does not include sample size in its penalty term while HQ and Entropy decrease this factor's effect using a logarithm or division function of sample size.

Examining the two groups of tables, we could summarize that the performance of these model fit measures based on sample size *N*.

- When *N* is equal to 400, SACAIC, DBIC, HQ, LMR, and BLRT have good rates of accuracy in identifying the number of latent class in a

UGMM setting. Only LMR and BLRT perform acceptably well when testing 1- versus 2-class linear GMM.

- When *N* increases to 700, SACAIC, BIC, DBIC, and HQ have satisfactory rates of more than 95% to select the two-class models in UGMM; CAIC and BIC also have satisfactory rate in linear GMM; SACAIC in LPM and SABIC in UGMM have acceptable rates of 90% to make right selections; LMR and BLRT has sufficient power to reject one-class model in all the three types of mixture models, but unfortunately they have inflated Type I error rates (mistakenly retain three-class models), which is particular worse in only in UGMM.
- When *N* equals to 1000, SACAIC, DBIC, LMR and BLRT (both testing 1- versus 2-class case) have satisfactory rates of accurate selection in all the three types of models; CAIC and BIC also have a rate of more than 95% in both UGMM and linear GMM; SABIC and HQ have good rates of more than 90% in both LPM and UGMM; LMR and BLRT have almost 90% chance to retain two-class models in UGMM.
- When considering the largest sample size 2000, CAIC, SACAIC, BIC, DBIC, LMR, and BLRT (testing 1- versus 2-class models) have sufficient rates of accuracy, more than 95%, in all the three types of mixture models; SABIC and HQ perform best in the unrestricted LPM and less accurate but acceptable in UGMM; LMR and BLRT perform best in UGMM, with 82% accuracy.

Comparing each of the four tables with Table 4.1.1, which has the general performance across all the sixty-four conditions, I find that Table 4.2.2.1a and Table 4.2.2.1b have very similar patterns with Table 4.1.1 while Table 4.2.2.1c and Table 4.2.2.1d conditioning on larger sample size exhibit different patterns. As stated before, LPM, a completely unrestricted model, does not outperform because there are many more parameters to be estimated than the other two types of mixture models based on the same set of data. However, when sample size is sufficiently large, the advantages of LPM become clear. In Table 4.2.2.1c, when sample size is 1,000, most of the model fit measures (except AIC, HQ, and HT-AIC, which are not useful for class enumeration) in LPM perform better than or equally well as the other two types of mixture models.

Figure 4.2.2.1 presents the model fit indices that exhibit a statistically significant effect between the types of mixture models and sample size. Among them, AIC, SACAIC, SABIC, HQ, HT-AIC, Entropy, LMR_1V2, BLRT_2V3 has Eta squared value more than 0.1, indicating a practically significant effect. Inspecting the characteristic of their patterns, they can essentially be classified into two groups: one group performs consistently better as sample size increases and the other does not.

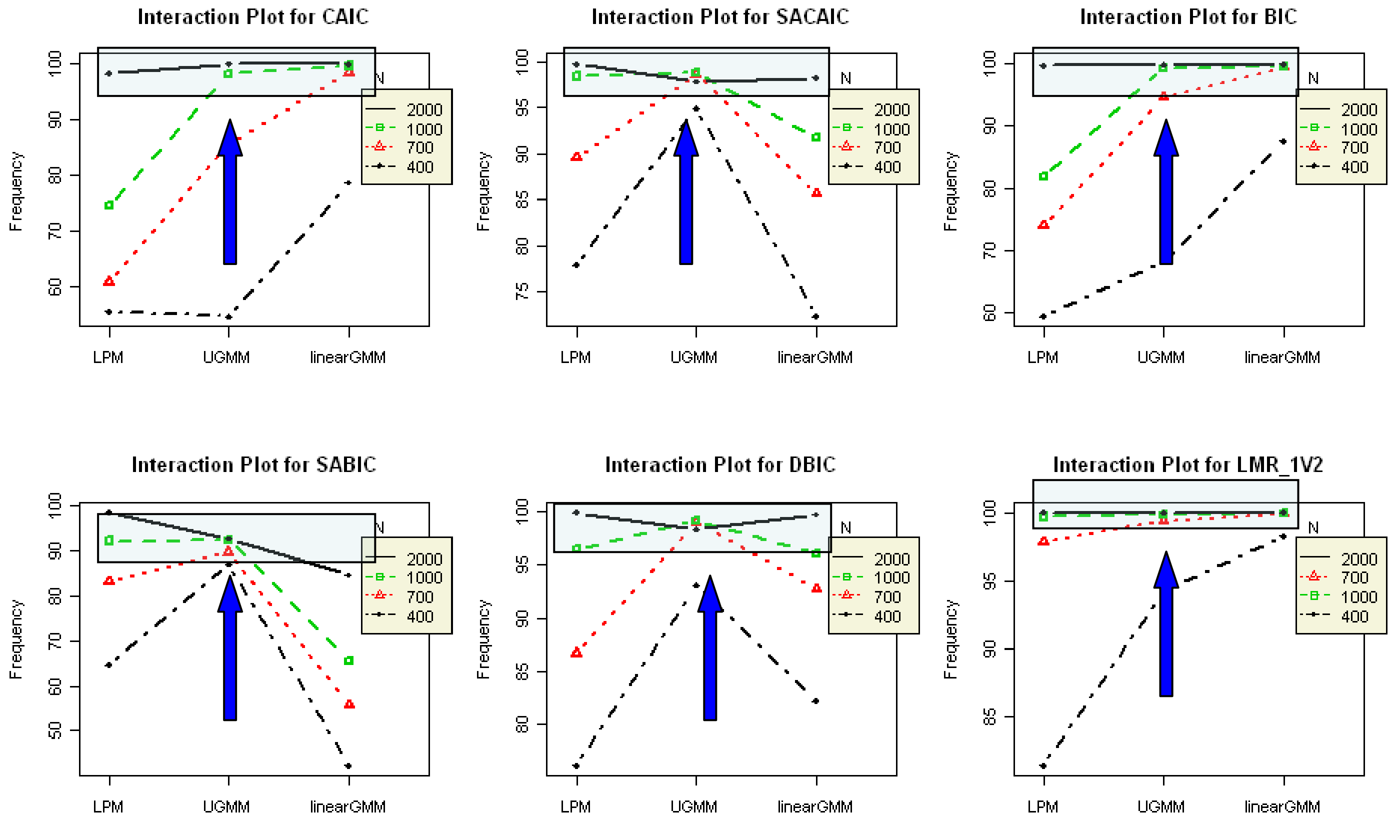

Figure 4.2.2.1a First group of model fit indices with significant interaction effects between the types of models and sample size

As shown in Figure 4.2.2.1a, Information criteria CAIC, SACAIC, BIC, SABIC, DBIC, LMR_1V2 belong to the first group because they have a similar pattern favoring large sample size. Blue arrows in the figure indicate that as sample size increase, they perform better in all the three types of models. When sample size approach 2,000, the performances of three types of mixture models are comparable, as evidenced by the shaded horizontal rectangular across the three mixture models. The advantage of UGMM is particularly clear in SACAIC, SABIC, DBIC and LMR_1V2 with higher or comparable probabilities when sample size ranging from 400 to 1,000.

In the second group, as Figure 4.2.2.1b shows, AIC, HQ, HT_AIC, Entropy, LMR_2V3 and BLRT_2V3 do not have the nice feature associated with sample size. Instead, AIC, HT_AIC, LMR_2V3 exhibit negative relationship with sample size in LPM and UGMM and positive in linear GMM. Among them, only LMR_2V3 shows an acceptable rate of accuracy in UGMM. As the shaded areas implied, UGMM performs best in the five fit indices except Entropy.

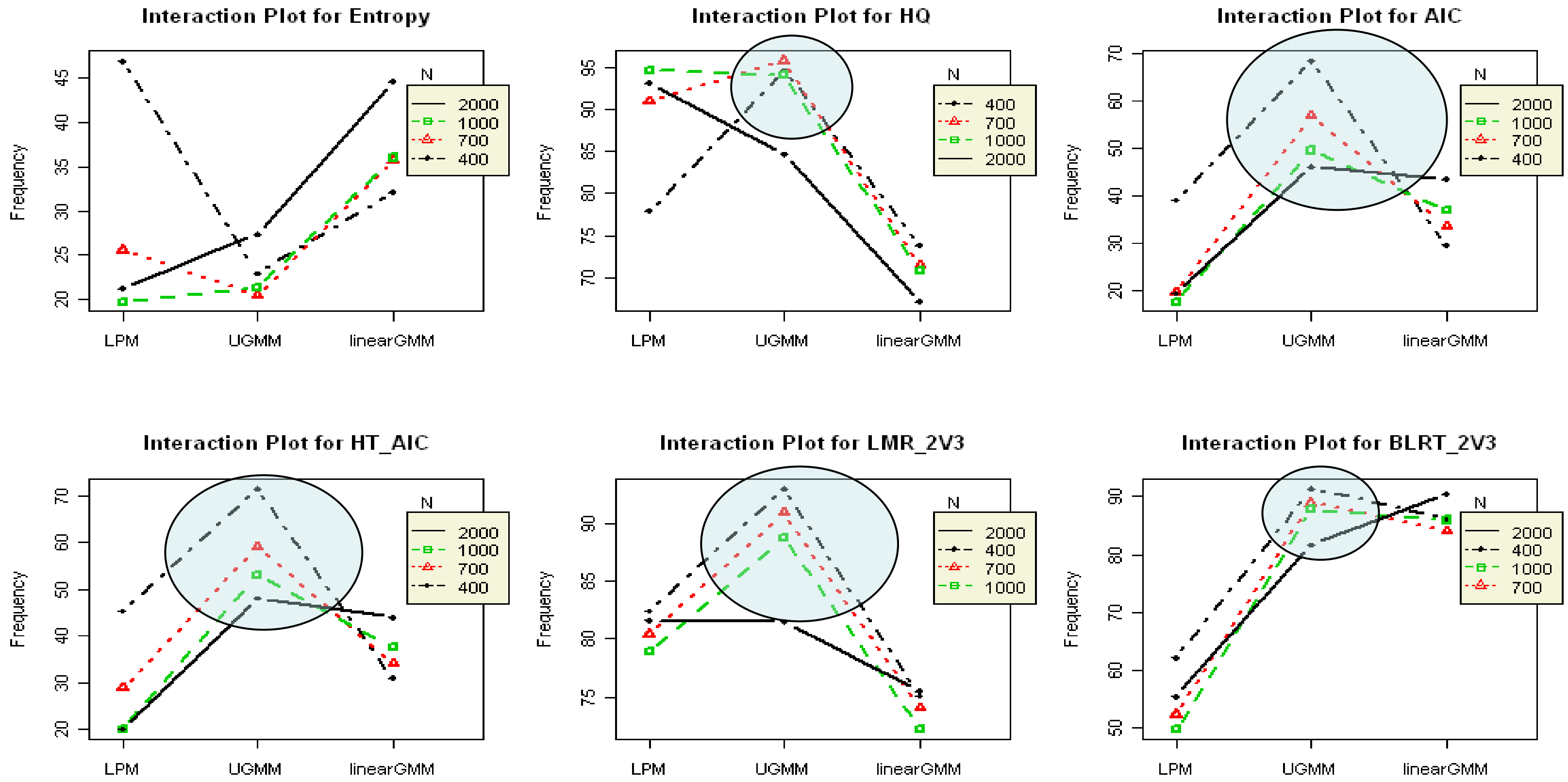

Figure 4.2.2.1b *Second group of model fit indices with significant interaction effects between the types of models and sample size*

Table 4.2.2.1a *Average frequency of each class selected by each index for 16 conditions with sample size of 400 (nonconvergent replications are included)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (62 converged replications for 3-class model)													
1 class	0	43	18	39	4	22	18	0	-	19	-	15	-
2 class	39	56	78	59	65	76	78	45	**47**	81	82	85	62
3 class	61	1	4	1	32	2	4	55	53	-	18	-	38
UGMM (95 converged replications for 3-class model)													
1 class	0	45	3	32	0	6	4	0	-	6	-	8	-
2 class	**69**	55	**95**	68	**87**	**93**	**95**	**72**	23	94	**93**	92	**91**
3 class	31	0	2	0	13	1	2	28	77	-	7	-	9
Linear GMM (75 converged replications for 3-class model)													
1 class	0	20	0	11	0	1	0	0	-	2	-	6	-
2 class	30	**79**	72	**88**	42	82	74	31	32	**98**	75	**94**	86
3 class	70	1	28	1	58	17	26	69	68	-	25	-	14

Table 4.2.2.1b *Average frequency of each class selected by each index for 16 conditions with sample size of 700 (nonconvergent replications are included)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (80 converged replications for 3-class model)													
1 class	0	39	10	26	1	13	4	0	-	2	-	1	-
2 class	20	61	90	74	83	87	91	29	26	98	80	99	52
3 class	80	0	0	0	16	0	5	71	74	-	20	-	48
UGMM (96 converged replications for 3-class model)													
1 class	0	14	0	5	0	0	0	0	-	1	-	0	-
2 class	**57**	86	**99**	95	**90**	**99**	**96**	**59**	20	99	**91**	**100**	**89**
3 class	43	0	1	0	10	1	4	41	80	-	9	-	11
Linear GMM (67 converged replications for 3-class model)													
1 class	0	2	0	1	0	0	0	0	-	0	-	0	-
2 class	34	**98**	86	**99**	56	93	71	34	**36**	**100**	74	**100**	84
3 class	66	0	14	0	44	7	29	66	64	-	26	-	16

Table 4.2.2.1c *Average frequency of each class selected by each index for 16 conditions with sample size of 1000 (nonconvergent replications are included)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (84 converged replications for 3-class model)													
1 class	0	20	1	18	0	3	0	0	-	0	-	0	-
2 class	18	75	98	82	**92**	96	**95**	20	20	**100**	79	**100**	50
3 class	82	5	0	0	8	0	5	80	80	-	21	-	50
UGMM (95 converged replications for 3-class model)													
1 class	0	2	0	0	0	0	0	0	-	0	-	0	-
2 class	**50**	**98**	**99**	99	**92**	**99**	94	**53**	21	**100**	**89**	**100**	**88**
3 class	50	0	1	0	8	1	6	47	79	-	11	-	12
Linear GMM (78 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	26	95	96	**100**	80	98	83	27	**26**	**100**	75	**100**	69
3 class	74	5	4	0	20	2	17	73	74	-	25	-	31

Table 4.2.2.1d *Average frequency of each class selected by each index for 16 conditions with sample size of 2000 (nonconvergent replications are included)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (81 converged replications for 3-class model)													
1 class	0	2	0	0	0	0	0	0	-	0	-	0	-
2 class	19	98	**100**	**100**	**98**	**100**	**93**	20	21	**100**	**82**	**100**	55
3 class	81	0	0	0	2	0	7	80	79	-	18	-	45
UGMM (93 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	**46**	**100**	98	**100**	93	98	85	**48**	27	**100**	**82**	**100**	82
3 class	54	0	2	0	8	2	15	52	73	-	19	-	18
Linear GMM (57 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	44	**100**	98	**100**	85	**100**	67	44	**45**	**100**	76	**100**	**90**
3 class	56	0	2	0	16	0	33	56	55	-	25	-	10

Table 4.2.2.2a *Average percent of each class selected by each index for 16 conditions with sample size of 400 (nonconvergent replications are excluded)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (62 converged replications for 3-class model)													
1 class	0	70	26	63	5	32	27	0	-	19	-	15	-
2 class	2	28	66	35	41	64	66	11	14	81	70	85	38
3 class	98	2	8	2	54	4	7	89	86	-	30	-	62
UGMM (95 converged replications for 3-class model)													
1 class	0	48	3	33	0	6	4	0	-	6	-	8	-
2 class	**66**	52	**95**	67	**86**	**93**	**94**	**69**	**19**	94	**92**	92	**91**
3 class	34	0	2	0	14	1	2	31	81	-	8	-	9
Linear GMM (75 converged replications for 3-class model)													
1 class	0	25	0	14	0	1	0	0	-	2	-	6	-
2 class	5	**73**	62	**85**	21	76	64	7	9	**98**	65	**94**	82
3 class	95	1	38	2	79	23	36	93	91	-	35	-	18

Table 4.2.2.2b *Average percent of each class selected by each index for 16 conditions with sample size of 700 (nonconvergent replications are excluded)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (80 converged replications for 3-class model)													
1 class	0	48	12	32	1	16	6	0	-	2	-	1	-
2 class	0	52	87	68	78	83	88	12	7	98	75	99	41
3 class	100	0	1	0	21	0	6	88	93	-	25	-	59
UGMM (96 converged replications for 3-class model)													
1 class	0	15	0	6	0	0	0	0	-	1	-	0	-
2 class	**54**	85	**99**	94	**89**	**99**	**96**	**57**	**17**	99	**90**	**100**	**88**
3 class	46	0	1	0	11	1	4	43	83	-	10	-	12
Linear GMM (67 converged replications for 3-class model)													
1 class	0	2	0	1	0	0	0	0	-	0	-	0	-
2 class	2	**98**	78	**99**	33	89	56	2	5	**100**	60	**100**	77
3 class	98	0	22	0	67	11	44	98	95	-	40	-	23

Table 4.2.2.2c *Average percent of each class selected by each index for 16 conditions with sample size of 1000 (nonconvergent replications are excluded)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (84 converged replications for 3-class model)													
1 class	0	25	2	23	0	5	0	0	-	0	-	0	-
2 class	2	69	98	77	90	94	94	4	4	100	74	100	40
3 class	98	6	0	0	10	0	6	96	96	-	26	-	60
UGMM (95 converged replications for 3-class model)													
1 class	0	2	0	0	0	0	0	0	-	0	-	0	-
2 class	**46**	98	**99**	99	**92**	**99**	**94**	**50**	**17**	**100**	**88**	**100**	**87**
3 class	54	0	1	0	8	1	6	50	83	-	12	-	13
Linear GMM (78 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	5	**100**	88	**100**	47	94	55	6	4	**100**	56	**100**	80
3 class	95	0	12	0	53	6	45	94	96	-	44	-	20

Table 4.2.2.2d *Average percent of each class selected by each index for 16 conditions with sample size of 2000 (nonconvergent replications are excluded)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (81 converged replications for 3-class model)													
1 class	0	2	0	0	0	0	0	0	-	0	-	0	-
2 class	0	98	**100**	**100**	**98**	**100**	**90**	1	2	**100**	76	**100**	45
3 class	100	0	0	0	2	0	10	99	98	-	24	-	55
UGMM (93 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	**40**	**100**	98	**100**	92	98	83	**42**	**22**	**100**	**80**	**100**	80
3 class	60	0	2	0	8	2	17	58	78	-	20	-	20
Linear GMM (57 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	1	**100**	97	**100**	72	**100**	41	1	2	**100**	55	**100**	**83**
3 class	99	0	3	0	28	0	59	99	98	-	45	-	17

Table 4.2.2.3 *ANOVA results for the frequency difference of model fit indices in selecting two-class models under conditions with four difference samples*

		AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT_AIC	Entropy	LMR_1V2	LMR_2V3	BLRT_1V3	BLRT_2V3
LPM	F	13.05	6.55	6.59	5.71	32.10	5.18	4.68	11.23	21.90	9.70	0.59	6.09	13.05
	Sig.	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.62	0.00	0.00
	Eta squared	**0.21**	**0.32**	**0.79**	**0.28**	**0.90**	**0.74**	**0.17**	**0.21**	**0.17**	**0.27**	0.03	**0.19**	0.07
UGMM	F	1.59	13.66	4.18	10.33	1.07	3.98	3.86	1.74	0.37	9.85	4.72	6.09	1.59
	Sig.	0.20	0.00	0.01	0.00	0.37	0.01	0.01	0.17	0.78	0.00	0.01	0.00	0.20
	Eta squared	0.07	**0.41**	**0.17**	**0.34**	0.05	**0.17**	**0.16**	0.08	0.02	**0.33**	**0.19**	**0.23**	**0.19**
Linear GMM	F	5.45	9.37	77.10	7.63	174.92	58.31	4.11	5.47	4.12	7.39	0.56	4.70	5.45
	Sig.	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.64	0.01	0.00
	Eta squared	**0.39**	**0.25**	**0.25**	**0.22**	**0.62**	**0.21**	**0.19**	**0.36**	**0.52**	**0.33**	0.03	**0.26**	**0.21**

4.2.3 Number of repeated measures

Table 4.2.3.1(a & b) and Table 4.2.3.2 (a & b) present the frequency and percent summary for four- and seven-measures conditions separately. Just like before, the model fit indices with highest probability of selection in the 2-class models are highlighted in bold. Table 4.2.3.2 has a very similar pattern with the general condition in Table 4.1 while Table 4.2.3.1 is slightly different in terms of a few exceptional indices, AIC, HQ, and BLRT for testing 1- versus 2-class models.

Generally speaking, increasing the number of repeated measures does not guarantee the improvement of the accuracy rate. Instead, many model selectors' values for two-class models in seven-measure models decrease. This is particularly clear in LPM, in which all the fit indices, except SABIC and LMR, perform better in selecting the two-class model in four-measure models than in seven-measure ones. Considering the seven-measure LPM has more parameters to be estimated than the four-measure LPM as shown in Table 4.1.1, we could understand why some information criteria achieve better class identifications in a four-measure LPM because they might penalize over-parameterization of a seven-measure LPM and thus disfavor complex models in this situation. Linear GMM has the least performance difference of fit indices between four and seven measure conditions. This finding is consistent with Tofighi and Enders' (2008) conclusion that the number of repeated measurements has only a relatively minor impact on the class enumeration.

ANOVA results in Table 4.2.3.3 also shows, due to LPM's complex parameterization, this model is most sensitive to the number of measures because most indices exhibit a significant (negative or positive) change in the accuracy rate. In contrast, linear GMM is the least sensitive one because its restricted parameterization makes seven-measure information redundant.

In both conditions with different repeated measures, SACAIC, DBIC, and BLRT (testing 1- vs. 2-class model only) in UGMM and BIC in linear GMM have satisfactory rates of accuracy (more than 95%). BLRT performs equally well in linear GMM for testing 1- versus 2-class models. CAIC can achieve acceptable rate of accuracy (more than 90%) in linear GMM. Moreover, both BIC and DBIC perform consistently well across the three types of mixture models with four repeated measures while LMR and BLRT are consistently good model selectors for testing 1- versus 2-class models across the three types of models with seven repeated measures.

Figure 4.2.3.1 presents model selectors exhibiting a significant interaction effect between the types of mixture models and the number of repeated measures. Only AIC, BIC, DBIC, HT-AIC, Entropy, BLRT_2V3 have partial Eta squared value more than 0.1, indicating a large effect size. Essentially they can be classified into two groups. In one group, seven-measure models generally perform better than four-measure models while it is the opposite case in the other group.

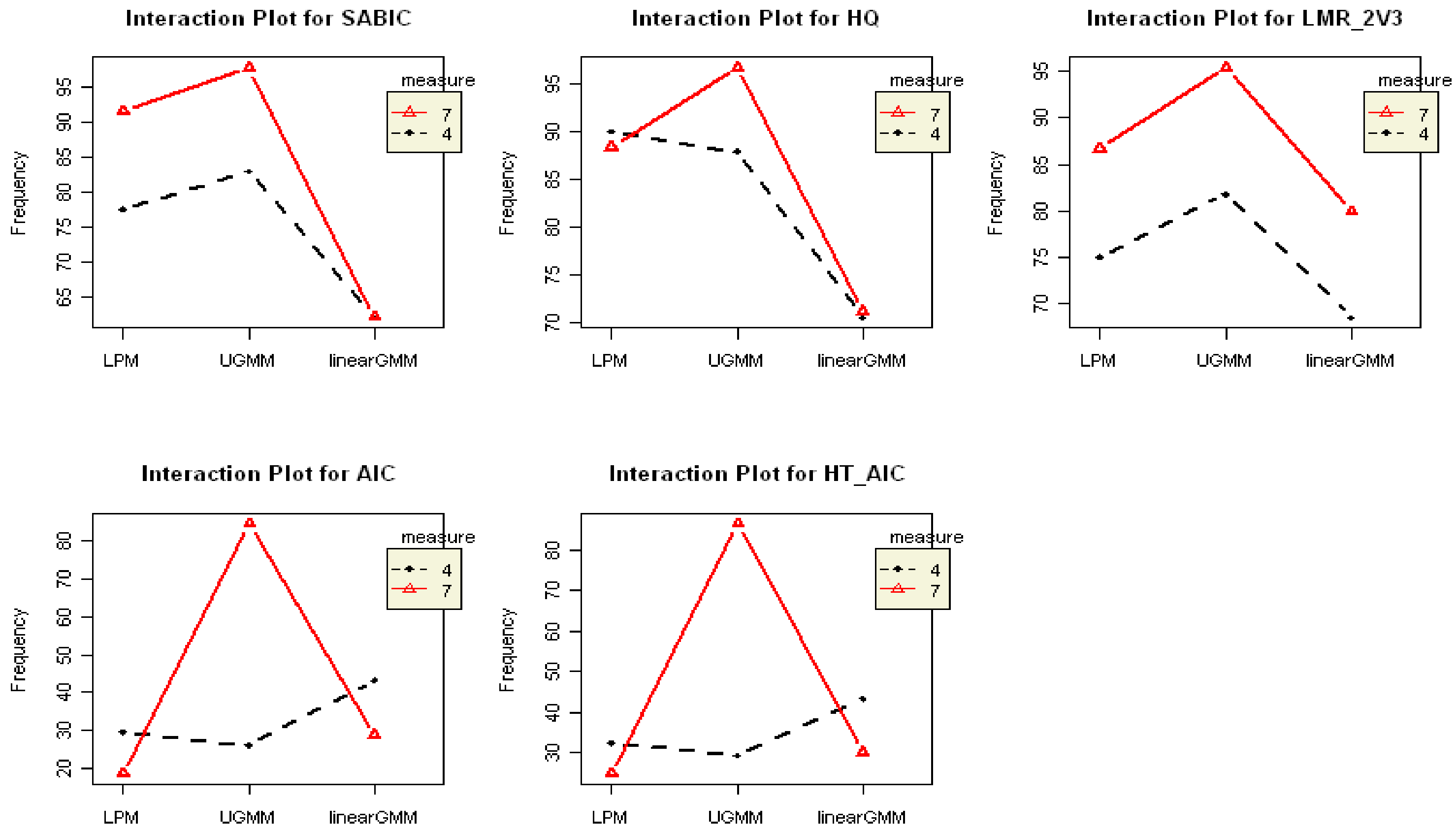

Figure 4.2.3.1(a) First group of model fit indices with significant interaction effects between the types of models and the number of measures

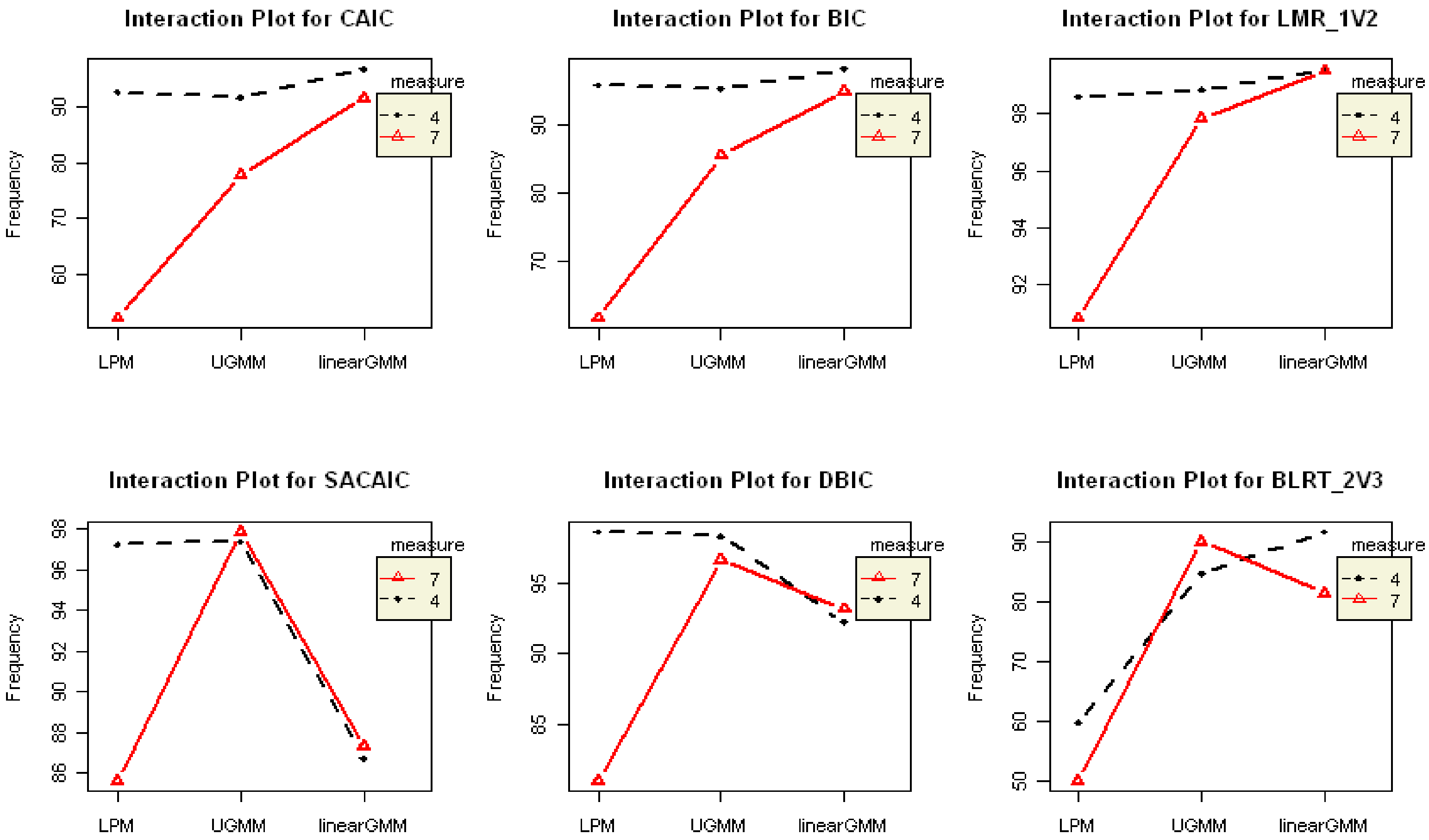

Figure 4.2.3.1(b) Second group of model fit indices with significant interaction effects between the types of models and the number of measures

Inspecting the first group of figures, it is clear that SABIC and LMR_2V3 have consistently higher rate of accuracy in models with seven measures across types of mixture models. The performance rate of over 95% is particularly satisfying in UGMM. HQ has very similar pattern with SABIC and LMR_2V3, except that its performance in LPM does not differ across the conditions with different measures.

AIC and HT-AIC have a similar pattern with a much higher rate of accuracy in UGMM with seven measures while consistently low across three types of mixture models with four measures and the other two mixture models with seven measures.

In the second group of figures, CAIC, BIC, LMR_1V2 have consistently high rates of accuracy across types of mixture models with four measures and dramatically increasing rates of accuracy from the least restricted LPM to the most restricted linear GMM. As stated before, LPM with seven measures needs to estimate many more parameters than the other two and so CAIC and BIC performs much worse in this model setting.

SACAIC and DBIC present much higher rates in four-measure LPM than seven-measure LPM. Both perform comparable across conditions with varying numbers of measurements in UGMM and linear GMM. BLRT_2V3 works satisfactorily in UGMM with seven measures and in linear GMM with four measures and much worse in LPM with different measures.

Table 4.2.3.1a *Average frequency of each class selected by each index for 32 conditions with 4 repeated measures (nonconvergent replications are included)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (71 converged replications for 3-class model)													
1 class	0	7	0	3	0	0	0	0	-	45	-	1	-
2 class	30	93	**97**	96	78	**99**	**90**	32	32	55	75	99	59
3 class	70	1	3	1	23	1	10	68	68	-	25	-	40
UGMM (92 converged replications for 3-class model)													
1 class	0	8	0	5	0	0	0	0	-	37	-	1	-
2 class	26	92	**97**	95	**83**	98	88	29	38	63	**82**	99	84
3 class	74	0	3	0	17	1	12	71	62	-	18	-	15
Linear GMM (59 converged replications for 3-class model)													
1 class	0	3	0	1	0	0	0	0	-	14	-	0	-
2 class	**43**	**97**	87	**98**	62	92	70	**43**	**44**	**86**	68	**100**	**91**
3 class	57	0	13	0	38	8	30	57	56	-	32	-	8

Table 4.2.3.1b *Average frequency of each class selected by each index for 32 conditions with 7 repeated measures (nonconvergent replications are included)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (82 converged replications for 3-class model)													
1 class	0	45	14	38	2	19	11	0	-	9	-	7	-
2 class	19	52	86	62	91	81	88	25	25	91	31	93	49
3 class	81	3	0	0	6	0	0	75	75	-	13	-	50
UGMM (97 converged replications for 3-class model)													
1 class	0	22	1	14	0	3	2	0	-	2	-	3	-
2 class	**85**	78	**98**	86	**98**	**97**	**97**	**87**	8	98	**95**	**97**	**89**
3 class	15	0	1	0	2	1	2	13	92	-	5	-	10
Linear GMM (74 converged replications for 3-class model)													
1 class	0	8	0	4	0	0	0	0	-	1	-	3	-
2 class	29	**92**	87	**95**	62	93	71	30	**31**	**99**	80	**97**	81
3 class	71	0	13	1	38	6	29	70	69	-	20	-	19

Table 4.2.3.2a *Average percent of each class selected by each index for 32 conditions with 4 repeated measures (nonconvergent replications are excluded)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (71 converged replications for 3-class model)													
1 class	0	13	0	8	0	0	0	0	-	1	-	1	-
2 class	1	86	95	91	65	97	86	5	5	99	64	99	43
3 class	99	1	5	1	35	2	14	95	95	-	36	-	57
UGMM (92 converged replications for 3-class model)													
1 class	0	9	0	5	0	0	0	0	-	1	-	1	-
2 class	**20**	91	**97**	95	**82**	**98**	**87**	**23**	**32**	99	**80**	99	83
3 class	80	0	3	0	18	2	13	77	68	-	20	-	17
Linear GMM (59 converged replications for 3-class model)													
1 class	0	4	0	2	0	0	0	0	-	0	-	0	-
2 class	3	**96**	79	**98**	37	88	48	3	4	**100**	46	**100**	86
3 class	97	0	21	0	63	12	52	97	96	-	54	-	14

Table 4.2.3.2b *Average percent of each class selected by each index for 32 conditions with 7 repeated measures (nonconvergent replications are excluded)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (82 converged replications for 3-class model)													
1 class	0	59	20	51	3	26	16	0	-	9	-	7	-
2 class	1	38	80	49	88	74	84	9	**8**	91	84	93	39
3 class	99	3	0	0	9	0	0	91	92	-	16	-	61
UGMM (97 converged replications for 3-class model)													
1 class	0	23	2	15	0	3	2	0	-	2	-	3	-
2 class	**84**	77	**98**	85	**98**	**96**	**97**	**86**	5	98	**95**	**97**	**89**
3 class	16	0	1	0	2	1	2	14	95	-	5	-	11
Linear GMM (74 converged replications for 3-class model)													
1 class	0	10	0	5	0	0	0	0	-	1	-	3	-
2 class	4	**90**	84	**94**	50	91	60	5	6	**99**	72	**97**	75
3 class	96	0	16	1	50	8	40	95	94	-	28	-	25

Table 4.2.3.3 *ANOVA results for the frequency difference of model fit indices in selecting two-class models between two conditions with four- and seven repeated measures*

		AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT_AIC	Entropy	LMR_1V2	LMR_2V3	BLRT_1V3	BLRT_2V3
LPM	F	11.47	37.48	7.56	27.97	14.22	13.97	0.18	3.17	3.67	5.51	78.92	4.13	18.96
	Sig.	**0.00**	**0.00**	**0.01**	**0.00**	**0.00**	**0.00**	0.67	0.08	0.06	**0.02**	**0.00**	**0.05**	**0.00**
	Eta squared	**0.16**	**0.38**	**0.11**	**0.31**	**0.19**	**0.18**	0.00	0.05	0.06	0.08	**0.56**	0.06	**0.23**
UGMM	F	390.77	3.91	0.27	3.13	71.15	1.15	11.80	345.33	86.88	0.89	53.95	2.25	7.68
	Sig.	**0.00**	**0.05**	0.61	0.08	**0.00**	0.29	**0.00**	**0.00**	**0.00**	0.35	**0.00**	0.14	**0.01**
	Eta squared	**0.86**	0.06	0.00	0.05	**0.53**	0.02	**0.16**	**0.85**	**0.58**	0.01	**0.47**	0.03	**0.11**
Linear GMM	F	47.37	1.72	0.05	1.73	0.00	0.24	0.20	41.26	31.46	0.03	75.33	2.75	30.42
	Sig.	**0.00**	0.19	0.83	0.19	0.98	0.63	0.66	**0.00**	**0.00**	0.86	**0.00**	0.10	**0.00**
	Eta squared	**0.43**	0.03	0.00	0.03	0.00	0.00	0.00	**0.40**	**0.34**	0.00	**0.55**	0.04	**0.33**

4.2.4 Mixing Proportions

Table 4.2.4.1 and Table 4.2.4.2 provide the frequency summary for the two groups conditioning on balanced and unbalanced sample sizes for the two latent classes separately. From the highlighted frequencies for all the model fit indices, it is clear that both tables have virtually identical patterns with the general performance summarized in Table 4.1. For this reason, the discussion for comparing three types of mixture models and model fit indices in section 4.1 can be applied here again.

Inspecting these two tables for the results of equal and unequal class proportions, neither one is overwhelmingly better than the other. ANOVA test results for the frequency difference of model selectors between the two class proportion conditions are summarized in Table 4.2.4.3. Clearly, varying this factor does not make any difference for all these model selectors. This is different from the Tofighi and Enders (2008) results, which indicated that a different mixing percentage can cause a dramatically different accuracy of class enumeration. More specifically, their model with extreme small proportion of 7% exhibited an unacceptable proportion of incorrect class identification. At least two reasons can explain this difference. First, the unbalanced mixing proportions in the current work are not extremely small; the smaller proportion reaches 25% of the total. Second, their results are based on two different sets of mixing proportions, conditioning on the other factors that held constant. The results in the current study come from a full-factorial design. The marginal effect of the mixing proportion is examined here, and so is its interaction effect later. Tueller and Lubke (2010) claimed that the BIC and SABIC perform worse in selecting the true model in conditions with lower sample sizes. But their competing models differed in within-class model structures, not the number of latent classes as in our case. We would expect that the difference between the balanced and unbalanced design might be clear if the minority class is extremely small. More research is required to know what the subtle cutting-point of mixing percentage is to make a difference in the accuracy of class enumeration. Considering this result in conjunction with Tofighi and Enders (2008) work, this cutting point is possibly between 7% and 25%, under the conditions that we have examined.

Some useful information about model fit indices can be summarized for practitioners. In both of the mixing proportion conditions, SACAIC, DBIC, LMR, and BLRT (testing 1-versus 2-class) in UGMM have satisfactory rates of accuracy. BIC and BLRT (testing 1-versus 2-class) in linear GMM and BLRT (testing 1- versus 2-class) in LPM also has almost perfect accuracy in this regard under both class proportion conditions. CAIC, DBIC, and LMR (testing 1- versus 2-class) in linear GMM, SACAIC in LPM, SABIC and HQ in UGMM, and LMR in both linear GMM and LPM have acceptable rates of accuracy across the mixing proportion conditions.

No model selector shows a significant interaction effect between the types of mixture models and mixing proportions.

Table 4.2.4.1a *Average frequency of each class selected by each index for 32 conditions with balanced sample size (nonconvergent replications are included)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (77 converged replications for 3-class model)													
1 class	0	27	8	21	2	11	7	0	-	4	-	5	-
2 class	24	69	90	78	83	88	88	28	28	96	44	95	54
3 class	76	3	2	1	15	1	6	72	72	-	21	-	45
UGMM (94 converged replications for 3-class model)													
1 class	0	17	1	11	0	2	1	0	-	1	-	2	-
2 class	**56**	83	**98**	89	**91**	**97**	**93**	**59**	21	**99**	**90**	**98**	**88**
3 class	44	0	1	0	9	1	6	41	79	-	10	-	11
Linear GMM (66 converged replications for 3-class model)													
1 class	0	7	0	4	0	0	0	0	-	10	-	2	-
2 class	37	**93**	86	**95**	61	92	69	38	**37**	90	73	**98**	86
3 class	63	1	14	1	39	7	31	62	63	-	27	-	13

Table 4.2.4.1b *Average frequency of each class selected by each index for 32 conditions with unbalanced sample size (nonconvergent replications are included)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (77 converged replications for 3-class model)													
1 class	0	25	6	21	1	8	5	0	-	6	-	3	-
2 class	24	75	93	79	86	91	90	29	29	94	82	97	54
3 class	76	0	1	0	14	0	5	71	71	-	18	-	45
UGMM (95 converged replications for 3-class model)													
1 class	0	13	1	8	0	1	1	0	-	2	-	1	-
2 class	**54**	87	**98**	92	**90**	**98**	**92**	**57**	25	98	**87**	**99**	**86**
3 class	46	0	2	0	10	1	8	43	75	-	13	-	14
Linear GMM (67 converged replications for 3-class model)													
1 class	0	4	0	2	0	0	0	0	-	1	-	1	-
2 class	35	**96**	88	**98**	63	93	72	36	**37**	**99**	75	**99**	**86**
3 class	65	0	12	0	37	7	28	64	63	-	25	-	14

Table 4.2.4.2a *Average percent of each class selected by each index for 32 conditions with balanced sample size (nonconvergent replications are excluded)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (77 converged replications for 3-class model)													
1 class	0	39	12	31	2	15	9	0	-	4	-	5	-
2 class	1	57	86	68	76	83	83	6	6	96	72	95	41
3 class	99	4	3	1	21	2	8	94	94	-	28	-	59
UGMM (94 converged replications for 3-class model)													
1 class	0	18	1	11	0	2	1	0	-	1	-	2	-
2 class	**52**	82	**97**	89	**90**	**97**	**92**	**55**	**17**	99	**89**	**98**	88
3 class	48	0	2	0	10	1	6	45	83	-	11	-	12
Linear GMM (66 converged replications for 3-class model)													
1 class	0	8	0	5	0	1	0	0	-	0	-	2	-
2 class	4	**91**	79	**94**	41	89	50	4	5	**100**	57	**98**	81
3 class	96	1	21	1	59	10	50	96	95	-	43	-	19

Table 4.2.4.2b *Average percent of each class selected by each index for 32 conditions with unbalanced sample size (nonconvergent replications are excluded)*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2vs.3)
LPM (77 converged replications for 3-class model)													
1 class	0	34	9	28	1	11	7	0	-	7	-	3	-
2 class	1	66	90	72	77	88	87	8	8	93	75	97	41
3 class	99	0	2	0	22	1	7	92	92	-	25	-	59
UGMM (95 converged replications for 3-class model)													
1 class	0	14	1	8	0	1	1	0	-	2	-	1	-
2 class	**51**	86	**97**	92	**89**	**98**	**91**	**54**	**21**	98	**86**	**99**	85
3 class	49	0	2	0	11	1	8	46	79	-	14	-	15
Linear GMM (67 converged replications for 3-class model)													
1 class	0	5	0	2	0	0	0	0	-	1	-	1	-
2 class	3	**94**	83	**97**	45	90	58	4	5	**99**	61	**99**	80
3 class	97	0	17	0	55	10	42	96	95	-	39	-	20

Table 4.2.4.3 *ANOVA results for the frequency difference of model fit indices in selecting two-class models between two different mixing proportions*

		AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT_AIC	Entropy	LMR_1V2	LMR_2V3	BLRT_1V3	BLRT_2V3
LPM	F	0.04	0.49	0.34	0.02	0.31	0.28	0.43	0.14	0.09	0.45	2.65	0.32	0.10
	Sig.	0.84	0.49	0.56	0.88	0.58	0.60	0.52	0.71	0.76	0.50	0.11	0.58	0.75
	Eta squared	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.04	0.01	0.00
UGMM	F	0.05	0.29	0.00	0.27	0.19	0.39	0.26	0.06	0.39	1.01	1.09	0.39	2.17
	Sig.	0.82	0.59	0.95	0.61	0.66	0.54	0.61	0.80	0.53	0.32	0.30	0.53	0.15
	Eta squared	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.02	0.02	0.01	0.03
Linear GMM	F	0.70	0.62	1.05	1.18	0.24	0.24	5.60	0.63	0.07	0.75	0.99	0.29	0.21
	Sig.	0.41	0.43	0.31	0.28	0.63	0.63	0.02	0.43	0.79	0.39	0.32	0.59	0.65
	Eta squared	0.01	0.01	0.02	0.02	0.00	0.00	0.08	0.01	0.00	0.01	0.02	0.00	0.00

4.2.5 Within-class model specification

The frequency summary in Table 4.2.5.1 and Table 4.2.5.2 present information about two groups conditioning on within-class models, properly and improperly specified. Again, after visual inspection, we found both tables have identical patterns with the general performance pattern summarized in Table 4.1. All the discussion about the types of mixture models and various model selectors in section 4.1 can also be applied. They are not repeated here for the sake of brevity.

As described in Chapter 3, the nonlinear component introduced to the majority class is subtle so that the growth pattern could often be considered linear mistakenly. In comparing these two tables, it is worthwhile to know which model or model selector(s) can function well in class enumeration on the two conditions that models are specified properly or improperly (taking nonlinear growth as linear). Most fit indices in Table 4.2.5.1 have higher rates of accuracy than that in Table 4.2.5.2, in which the model estimation is conducted with misspecified within-class models. As seen in Table 4.2.5.1a versus Table 4.2.5.1b and Table 4.2.5.2a versus Table 4.2.5.2b, the likelihood ratio tests, BLRT and LMR, both tend to overestimate the number of latent class, which is the effect of nonlinear component.

In addition, the ANOVA test is conducted to check whether the frequency rate of model selectors in selecting two-class models between the properly and improperly specified models is significantly different or not. Although most model selectors perform better in the properly specified model, the very few significant cases in Table 4.2.5.3 indicate this performance gap is not huge, probably due to the very subtle nonlinear component introduced in the population model.

Moreover, there are several exceptional indices (e.g., CAIC) that have better performance in the improperly specified within-class model than in the properly specified one. One common property shared by these exceptions is that they underestimated the number of latent classes conditioning on the properly specified within-class models. As Bauer and Curran (2004) summarized, nonlinear relations among observed or latent variables might lead to a spurious latent class. Some model fit indices in Table 4.2.5.1, such as CAIC and SACAIC in UGMM or BIC in linear GMM, underestimate the number of latent class, but they might extract spurious latent class due to the existence of nonlinearity and therefore their performance improve to some extent as shown in Table 4.2.5.2.

Due to the nonlinear component added to the population model, the indices overestimated the number of latent classes in Table 4.2.5.1, which will decrease the accuracy rate in Table 4.2.5.2 because more replications were incorrectly classified into three-class group. This finding also confirms the Bauer and Curran study result that a spurious latent class can be extracted because of nonlinear relations.

Some information about model fit indices for practitioners’ use is summarized as follows. In both model specification conditions, SACAIC and DBIC in

UGMM, BIC in linear GMM, and two likelihood ratio tests for 1- against 2-class models perform well with satisfactory accuracy rate. These model selectors seem to be robust to mild nonlinearity in this case. CAIC and DBIC in linear GMM and SACAIC in LPM have acceptable rates of accuracy.

Only Entropy and BLRT_2V3 exhibit a significant interaction effect between types of mixture models and the within-class model specification as shown in Figure 4.2.5.1. However, neither of them has partial Eta squared value more than 0.1, which indicates their interaction effect is not practically significant. Entropy performs poorly across the models and model specification conditions, particularly worse in less restricted UGMM. Examining its efficiency under different conditions, Entropy always favored the most restricted linear GMM. By the same token as introduced before, the most restricted model linear GMM, as long as the bias is acceptably small, might have great precision in estimates, such as posterior probability associated with each subject, resulting in larger Entropy values. However, entropy *per se* is not useful because of its low rate of accuracy in identifying the number of latent classes in GMM. Generally speaking, BLRT_2V3 performs better in estimating data in which no nonlinear component is embedded, as evidenced by the fact that the broken line is always above the solid line. It works best in UGMM when the nonlinear factor does not exist in data. The results in linear GMM are identical across two different model specifications embedded in data. This also implies the nonlinear effect introduced is quite small in magnitude and so the advantages of LPM and UGMM are not distinct.

Table 4.2.5.1a Average frequency of each class selected by each index for 32 conditions with properly specified within-class model (nonconvergent replications are included)

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (73 converged replications for 3-class model)													
1 class	0	26	9	21	2	11	6	0	-	7	-	5	-
2 class	27	71	91	79	85	89	89	31	32	93	83	95	58
3 class	73	3	1	0	13	0	4	69	68	-	17	-	42
UGMM (94 converged replications for 3-class model)													
1 class	0	18	1	11	0	2	1	0	-	2	-	3	-
2 class	**59**	82	**98**	89	**93**	**98**	**95**	**62**	16	98	**90**	97	**91**
3 class	41	0	1	0	7	0	4	38	84	-	10	-	9
Linear GMM (64 converged replications for 3-class model)													
1 class	0	7	0	4	0	0	0	0	-	1	-	2	-
2 class	37	**93**	89	**96**	63	94	73	37	**38**	**99**	75	**98**	87
3 class	63	0	11	0	37	6	27	63	62	-	25	-	13

Table 4.2.5.1b Average frequency of each class selected by each index for 32 conditions with improperly specified within-class model (nonconvergent replications are included)

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (80 converged replications for 3-class model)													
1 class	0	26	6	21	0	8	5	0	-	4	-	3	-
2 class	21	73	92	78	84	91	89	26	25	96	79	97	52
3 class	79	1	2	1	16	1	6	74	75	-	21	-	48
UGMM (95 converged replications for 3-class model)													
1 class	0	13	0	8	0	1	1	0	-	1	-	1	-
2 class	**51**	87	**97**	92	**87**	**97**	**89**	**54**	30	99	**87**	**99**	84
3 class	49	0	2	0	13	2	10	46	70	-	13	-	16
Linear GMM (69 converged replications for 3-class model)													
1 class	0	4	0	2	0	0	0	0	-	0	-	1	-
2 class	35	**95**	86	**97**	62	92	69	36	**36**	**100**	73	**99**	**87**
3 class	65	1	14	1	38	8	31	64	64	-	27	-	13

Table 4.2.5.2a Average percent of each class selected by each index for 32 conditions with properly specified within-class model (nonconvergent replications are excluded)

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (73 converged replications for 3-class model)													
1 class	0	40	13	33	3	17	10	0	-	7	-	5	-
2 class	0	58	85	67	76	83	84	6	7	93	75	95	42
3 class	100	3	2	0	21	0	6	94	93	-	25	-	58
UGMM (94 converged replications for 3-class model)													
1 class	0	19	1	12	0	2	1	0	-	2	-	3	-
2 class	**56**	81	**98**	88	**93**	**97**	**95**	**58**	**11**	98	**89**	97	**90**
3 class	44	0	1	0	7	0	4	42	89	-	11	-	10
Linear GMM (64 converged replications for 3-class model)													
1 class	0	9	0	5	0	1	0	0	-	1	-	2	-
2 class	2	**91**	82	**95**	43	90	56	3	4	**99**	60	**98**	80
3 class	98	0	17	0	57	9	44	97	96	-	40	-	20

Table 4.2.5.2b Average percent of each class selected by each index for 32 conditions with improperly specified within-class model (nonconvergent replications are excluded)

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (80 converged replications for 3-class model)													
1 class	0	33	7	26	1	10	6	0	-	4	-	3	-
2 class	1	66	90	73	77	88	85	8	6	96	73	97	40
3 class	99	1	3	1	22	2	8	92	94	-	27	-	60
UGMM (95 converged replications for 3-class model)													
1 class	0	13	0	8	0	1	1	0	-	1	-	1	-
2 class	**48**	87	**97**	92	**87**	**97**	**89**	**51**	**27**	99	**86**	**99**	**83**
3 class	52	0	3	0	13	2	11	49	73	-	14	-	17
Linear GMM (69 converged replications for 3-class model)													
1 class	0	4	0	2	0	0	0	0	-	0	-	1	-
2 class	5	**95**	80	**97**	44	89	52	6	6	**100**	58	99	81
3 class	95	1	20	1	56	11	48	94	94	-	42	-	19

Table 4.2.5.3 *ANOVA results for the frequency difference of model fit indices in selecting two-class models between two model specification conditions*

		AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT_AIC	Entropy	LMR_1V2	LMR_2V3	BLRT_1V3	BLRT_2V3
LPM	F	2.97	0.06	0.17	0.01	0.09	0.17	0.00	1.00	3.74	0.50	3.62	0.32	6.31
	Sig.	0.09	0.81	0.68	0.93	0.77	0.68	0.96	0.32	0.06	0.48	0.06	0.58	0.01
	Eta squared	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.06	0.01	0.06	0.01	0.09
UGMM	F	1.00	0.40	1.09	0.19	5.30	0.00	4.46	0.96	8.57	0.35	1.25	0.86	14.42
	Sig.	0.32	0.53	0.30	0.66	0.02	0.98	0.04	0.33	0.00	0.56	0.27	0.36	0.00
	Eta squared	0.02	0.01	0.02	0.00	0.08	0.00	0.07	0.02	**0.12**	0.01	0.02	0.01	**0.19**
Linear GMM	F	0.22	0.29	1.20	0.34	0.07	1.10	6.56	0.21	0.61	1.50	1.84	0.93	0.00
	Sig.	0.64	0.59	0.28	0.56	0.80	0.30	0.01	0.65	0.44	0.23	0.18	0.34	0.99
	Eta squared	0.00	0.00	0.02	0.01	0.00	0.02	0.10	0.00	0.01	0.02	0.03	0.01	0.00

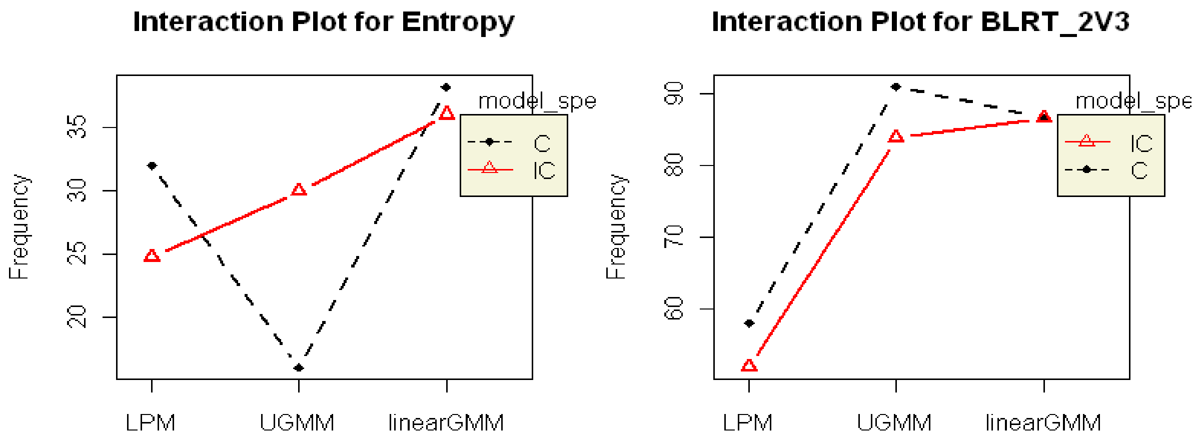

Figure 4.2.5.1 The significant interaction effects between the types of models and model specifications

4.3. Significant Interaction Effect between Factors in a Given Mixture Model

Two-way ANOVA tests were conducted for the purpose of examining whether there are interaction effects between the manipulated factors on the performance of model selectors conditioning on types of mixture models. For the sake of brevity, only significant results are listed and interpreted. Interaction effects involving mixing proportion and within-class model specifications are not presented here because none of their interaction terms is significant.

4.3.1 Sample size X *Class separation*

As Figure 4.3.1.1 shows, five model fit indices are statistically and practically significant in LPM, in terms of their *p* values (below 0.5) and partial Eta squared values (above 0.1) respectively. Except entropy, SACAIC, HQ, LMR_1V2 and BLRT_1V2 follow a similar interaction pattern. While these four indices work consistently well across different sample sizes under the condition of high class separation with a Mahalanobis distance of 5, only when the sample size reaches around or above 700 do they perform acceptably well (over 90%) under lower class separation condition.

Figure 4.3.1.2 indicates that all five indices, CAIC, BIC, DBIC, LMR_1V2, and BLRT_1V2, exhibiting a statistically and practically significant interaction effect between sample size and class separation in UGMM, have a similar pattern as those indices in LPM. CAIC requires larger sample size (e.g., 1000) to achieve an acceptable rate of accuracy (90%) than the other four indices do (700 or less).

As Figure 4.3.1.3 shows, again, four indices with a statistically and practically significant interaction effect have a similar pattern with those in other types of mixture models. If the class separation is very large, such as 5 standardized Mahalanobis distance units in this case, sample size of 400 is large enough to accurately identify the number of latent classes. If the class separation is 3.5 Mahalanobis distance units, sample size of 700 is enough for the purpose of class identification.

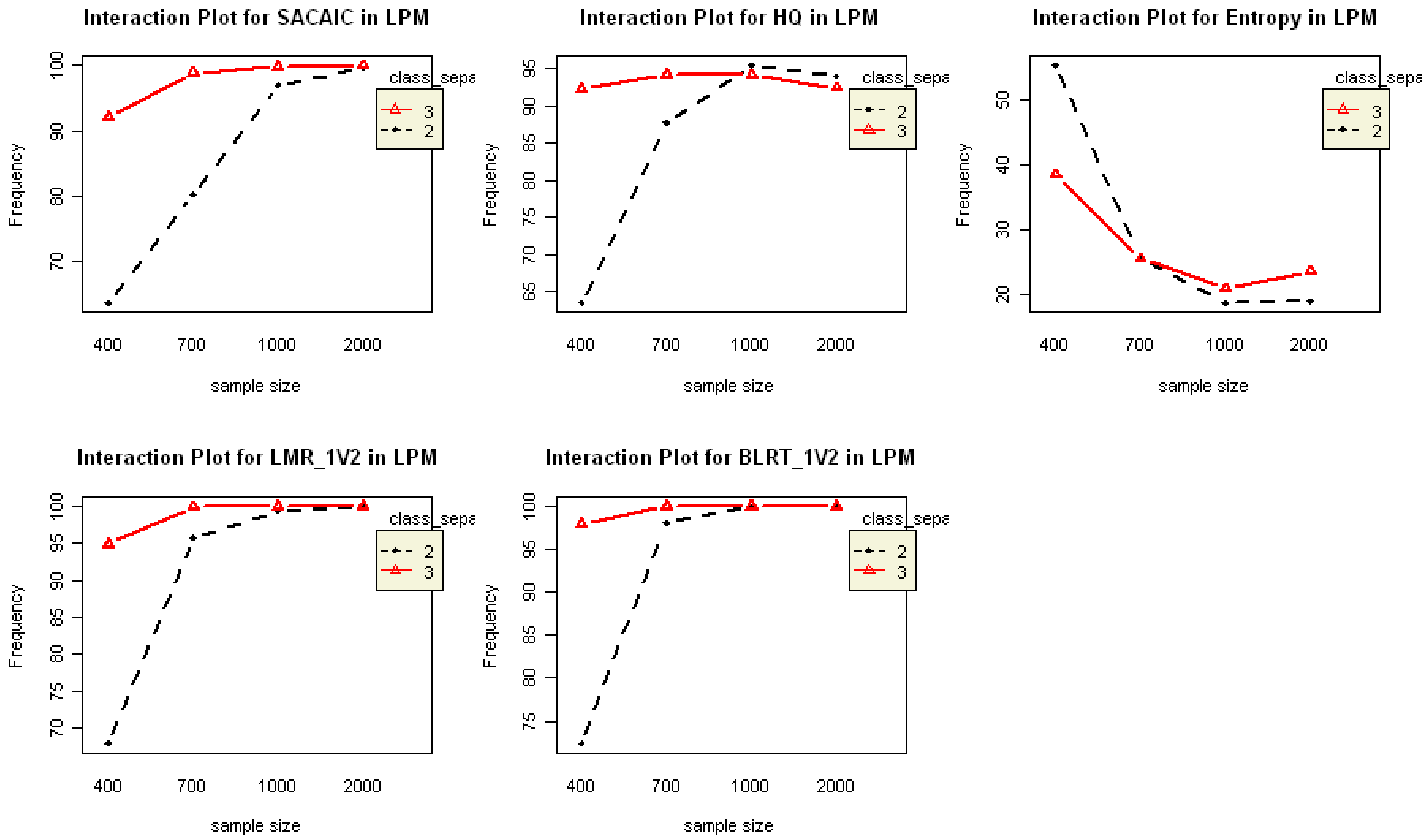

Figure 4.3.1.1 *Significant Interaction (sample size X class separation) Plot for model selectors in LPM*

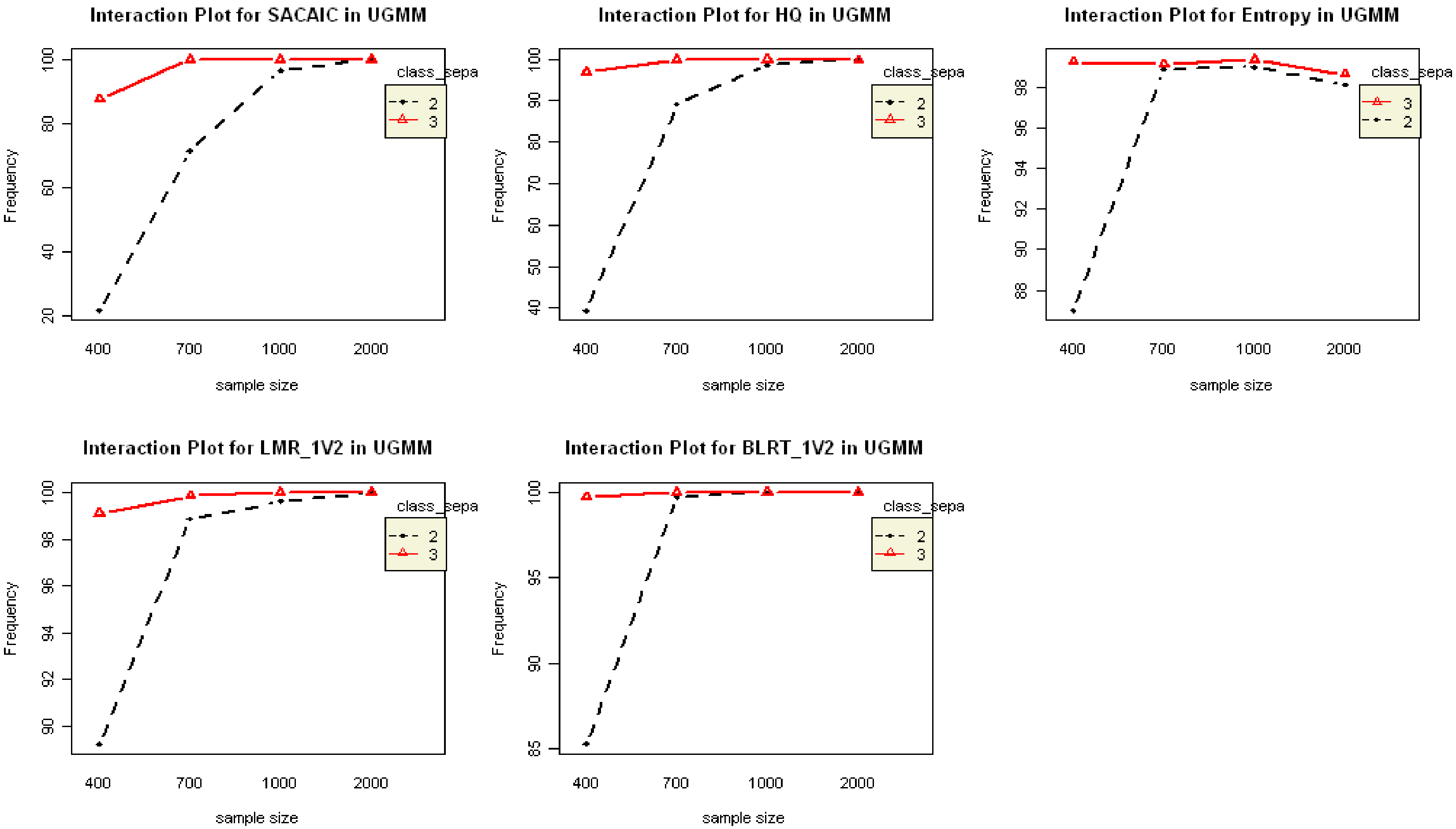

Figure 4.3.1.2 *Significant Interaction (sample size X class separation) Plot for model selectors in UGMM*

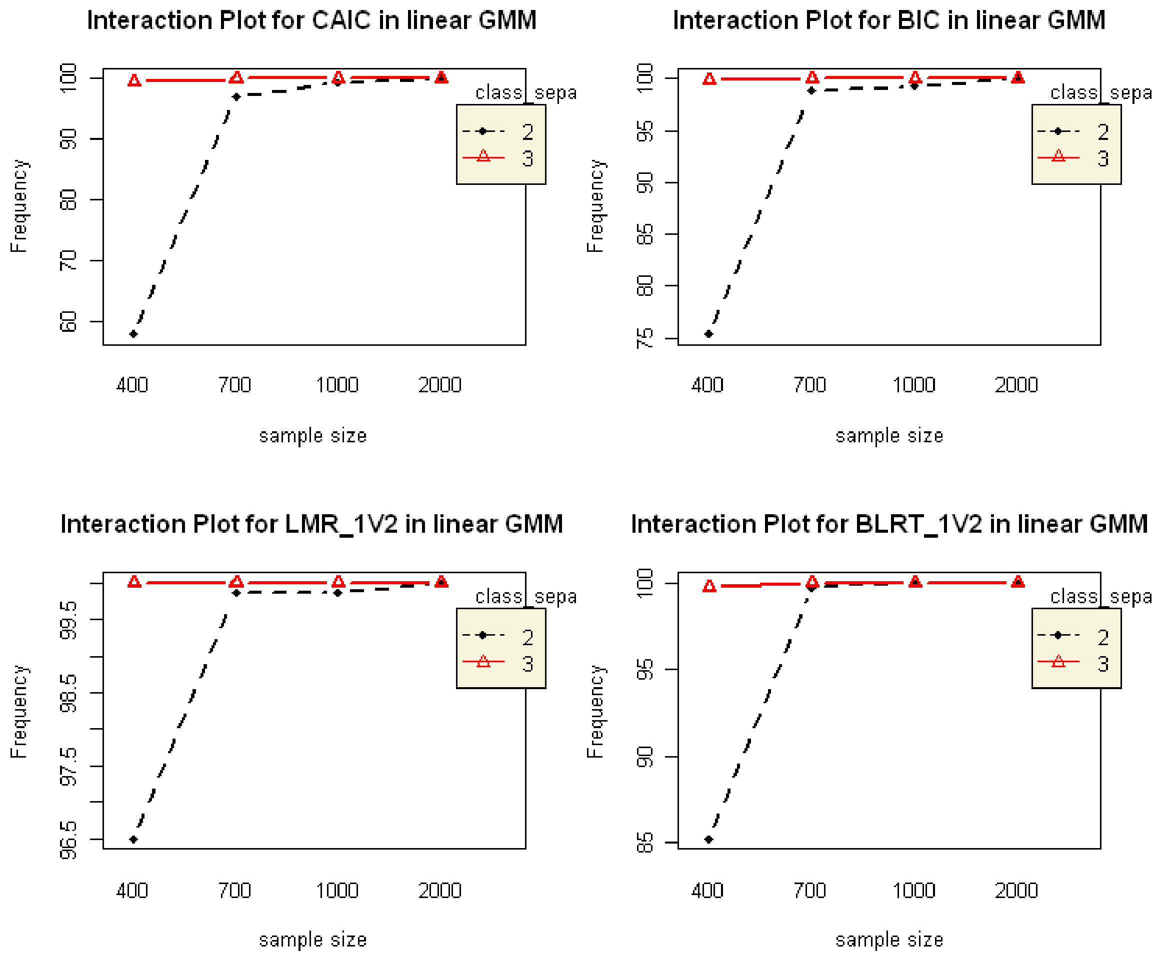

Figure 4.3.1.3 *Significant Interaction (sample size X class separation) Plot for model selectors in Linear GMM*

4.3.2 Sample size X *Number of measures*

Figure 4.3.2.1 presents the statistically and practically significant interaction plots for eight model selectors in LPM. Six of them, CAIC, SACAIC, BIC, DBIC, LMR_1V2, BLRT_1V2 follow a similar pattern that they tend to have higher rate of correctly identifying the number of latent classes in the four-measure LPM rather than in the seven-measure LPM. This is partly due to the great demand of sample size for the highly parameterized LPM with seven measures.

Very different from the other six indices, SABIC and HQ are two exceptional cases. SABIC works better in seven-measure model over four-measure ones. As summarized in section 2.4.1, this measure favors model with large number of parameters and as such its special pattern does make sense. As for HQ, sample size of 700 is a cutting point, below which HQ performs better in four-measure model and above which HQ works better with an acceptable rate of accuracy in seven-measure model.

Figure 4.3.2.2 shows that the interaction pattern for model indices in UGMM distinctly different from those in LPM. First, generally seven-measure models win in this type of mixture model. This is probably due to the relatively lower requirement for sample size of this model. Second, the trend line of accuracy rate is not positively associated with sample size, which is also different from LPM. As summarized in Section 4.2.2, SABIC generally performs better as sample size increases across all the conditions. Since SABIC favors complex models with more parameters, sample size of 400 is enough for it achieving the ceiling effect in more complex seven-measure UGMM, and as discussed in Section 4.2.2, HQ, HT_AIC and LMR_2V3 perform worse as sample size increases in UGMM.

In linear GMM, only DBIC and BLRT_2V3 exhibit a statistically and practically significant interaction effect between sample size and the number of measures, as displayed in Figure 4.3.2.3. As sample size approaches 700, DBIC achieves good accuracy rate in both conditions with different numbers of measures. BLRT_2V3 performs much better in linear GMM with four repeated measures than in those with seven measures.

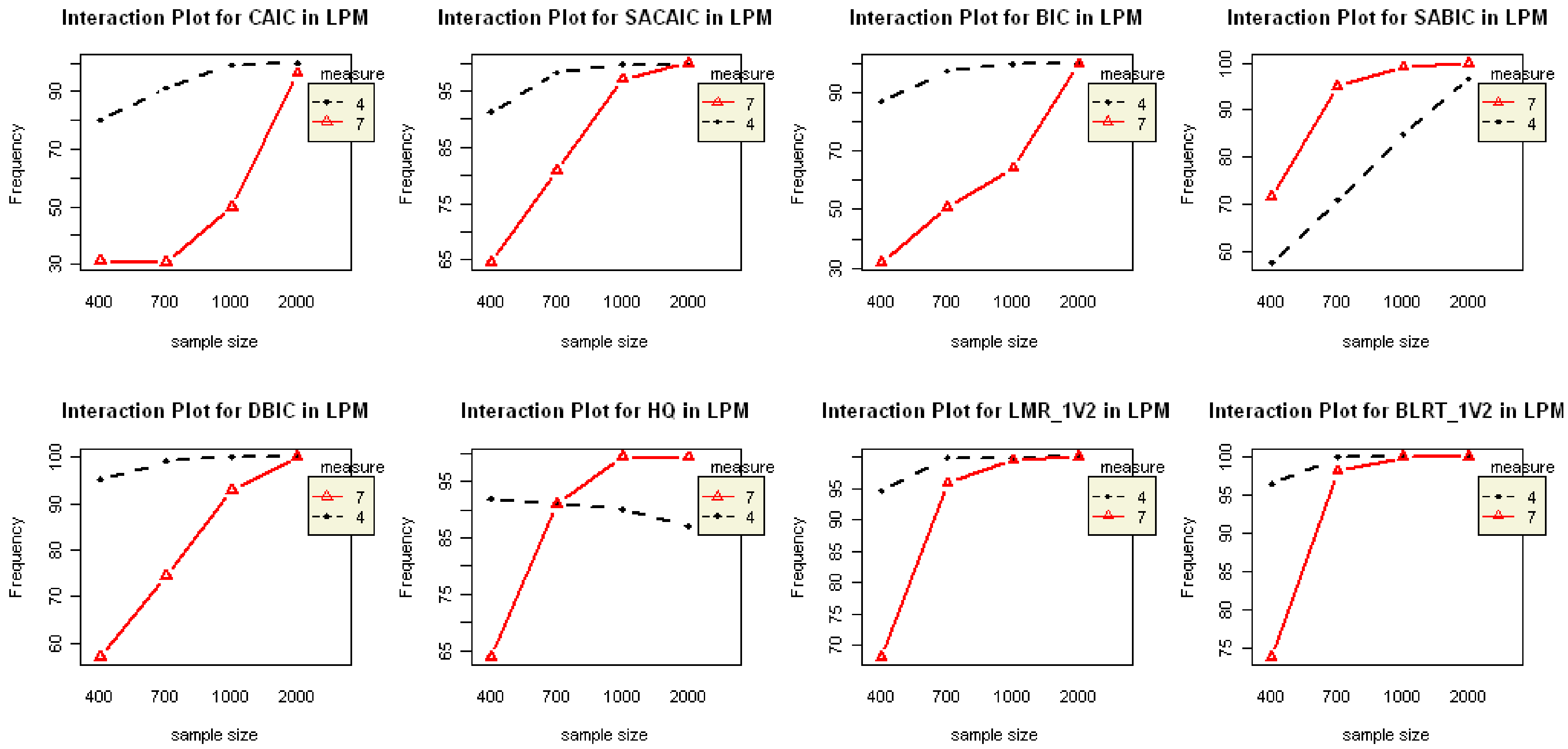

Figure 4.3.2.1 *Significant Interaction (sample size X the number of measures) Plot for model selectors in LPM*

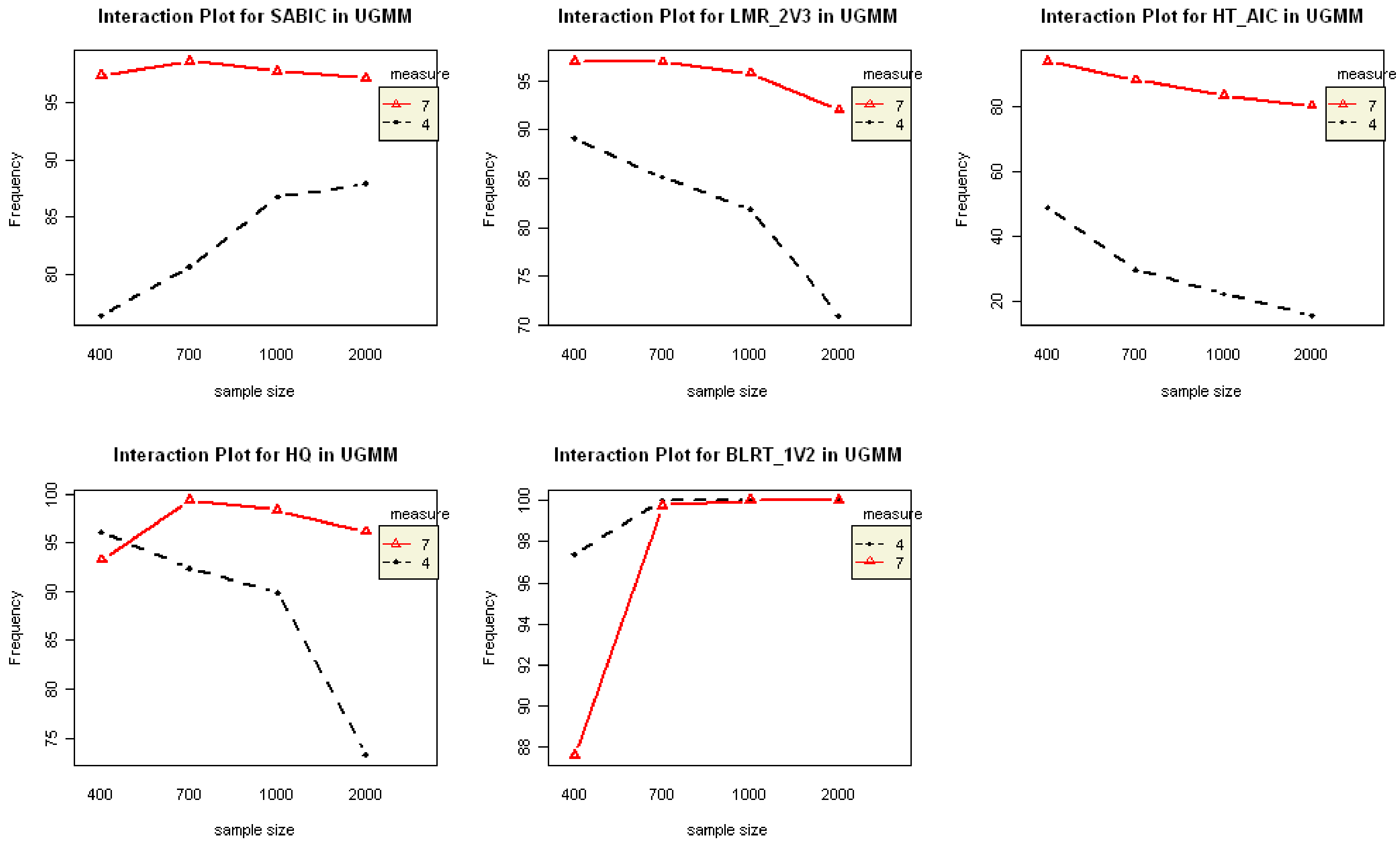

Figure 4.3.2.2 *Significant Interaction (sample size X the number of measures) Plot for model selectors in UGMM*

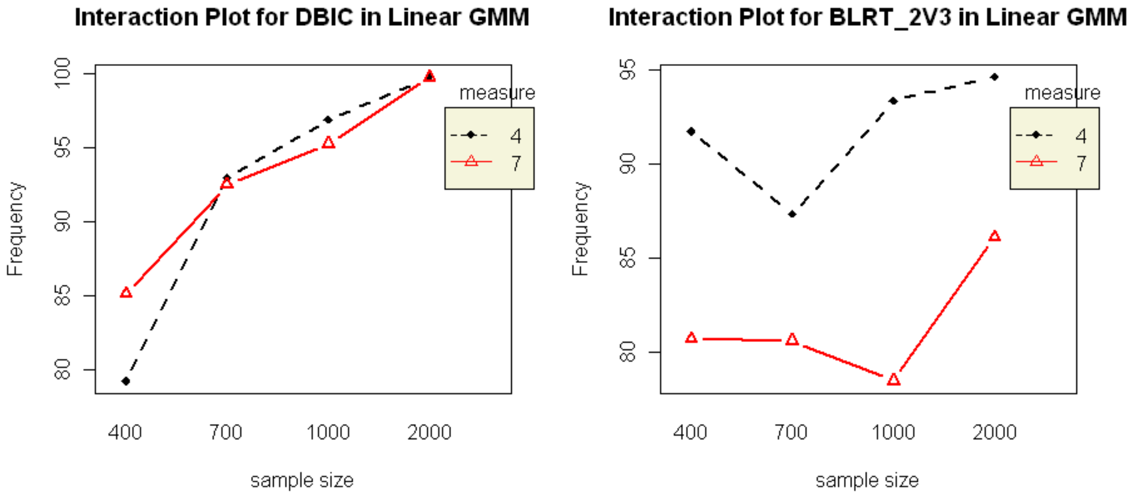

Figure 4.3.2.3 *Significant Interaction (sample size X the number of measures) Plot for model selectors in Linear GMM*

4.3.3 Class separation X *Number of measures*

Four model-fit indices in Figure 4.3.3.1 exhibit statistically and practically significant interaction effect of class separation and the number of repeated measures in LPM. Their accuracy rates go up dramatically as class separation increases from 2 SD to 3SD in seven-measure LPM, but not sensitive to this change in the models with four-measure.

As Figure 4.3.3.2 shows, only SACAIC has a statistically and practically significant interaction effect in UGMM. And SACAIC has a very satisfactory rate of accuracy across conditions with different combinations of class separation and number of measures. Increasing class separation does not help this index correctly enumerate the number of latent class in four-measure UGMM. On the contrary, larger class separation does have a significant effect on improving rate of accuracy in seven-measure UGMM.

There is no significant interaction effect between class separation and the number of measures for model selectors in Linear GMM.

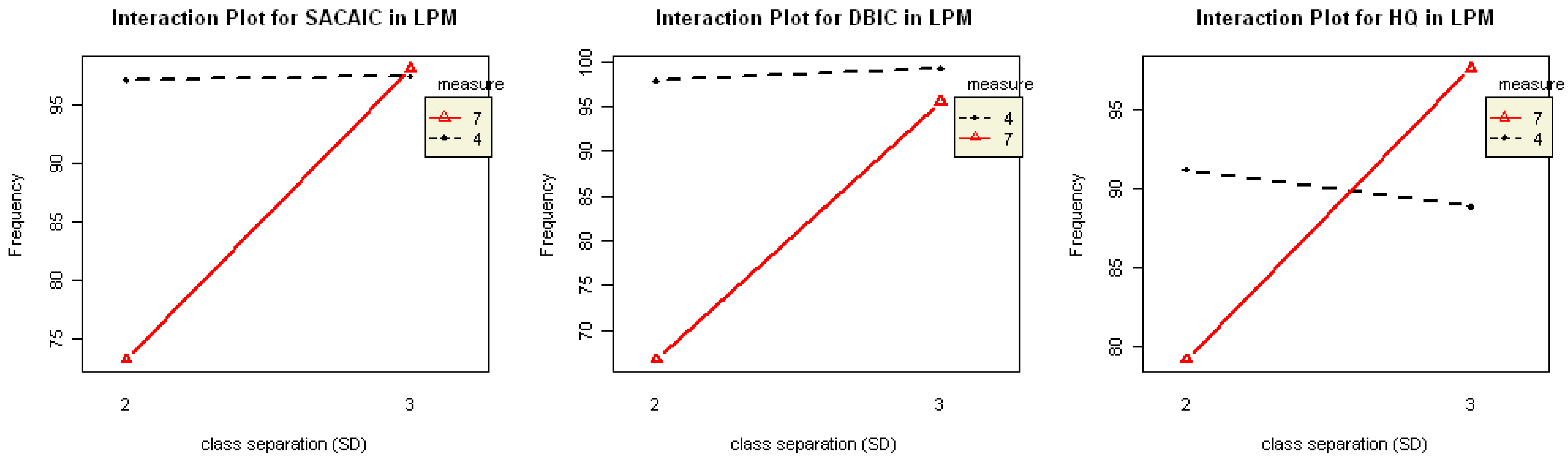

Figure 4.3.3.1 *Significant Interaction (Class separation X the number of measures) Plot for model selectors in LPM*

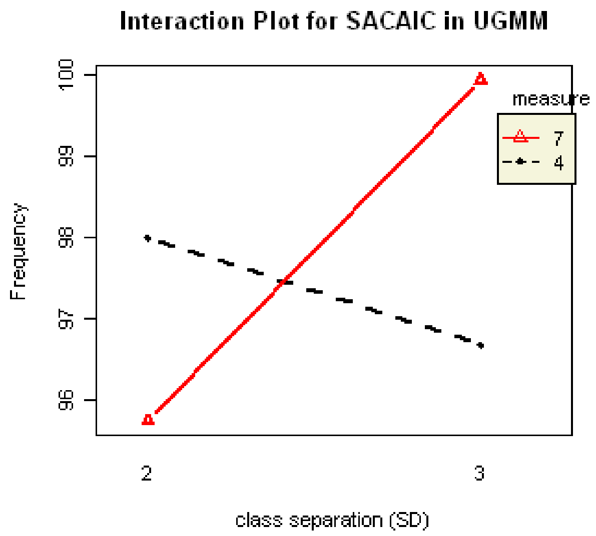

Figure 4.3.3.2 *Significant Interaction (Class separation X the number of measures) Plot for model selectors in UGMM*

CHAPTER 5: DISCUSSION

"It is a capital mistake to theorize before one has data"

. —Arthur Conan Doyle, "Sherlock Homes"

Although class enumeration in application of growth mixture model is recommended by some researchers to be confirmatory in nature, practitioners often use this model in an exploratory way in reality because theory could be too ambiguous to tell exactly how many classes exist underlying the data, or researchers do not know how robust this theory is to be applied to different dataset. That is why practitioners using GMM need to explore the data and rely on model fit indices to make a decision with respect to the number of latent classes. However, there is no universally accepted index that can accomplish this task so far.

In addition to studying the relative efficiency of a wide range of model fit indices in class enumeration, more importantly, the current study has provided an alternative modeling strategy of assessing the number of latent classes for GMM. Both theoretical and empirical reasons for using less restricted models in this regard were presented.

As stated before, how to balance bias and precision is always an important issue in statistical modeling. More flexible models, like UGMM and LPM, can lower the chance of bias occurring caused by model misspecification. But, estimating them requires larger sample sizes to detect the heterogeneity underlying the data and obtain a reliable result regarding class determination. Between the least restrictive LPM and the most restrictive linear GMM, UGMM is only one kind of compromise choice and there must be numerous ways to construct less restricted mixture models, depending on various ways to impose model restrictions to data. A practical suggestion arising from this study is that practitioners, based on existing theory, their experience or belief, ought to think about which part of the within-class model structure that is uncertain and thuse should be loosened. By doing so, the chance of bias caused by model misspecification is reduced.

After pooling all the mixture models into Mplus to be estimated, just as other type of mixture models, nonconvergence is a problem that needs to be addressed in current study, which is particularly true for three-class LPM and three-class linear GMM with low convergence rate on average. To make the arguments herein convincing, as presented in the results section, two different ways were used to summarize the results, one exclude those nonconvergent replications and the other include them as evidence supporting two-class models. Both methods have its limitations. And both types of results are very similar making the conclusions more credible.

As the results section shows, different model fit indices might perform well in different mixture models with varying restrictions. After considering associated factors, such as class separation and sample size, practitioners must make a decision regarding using which models in conjunction with which model selector(s) to maximize the chance of correctly identifying the number of latent classes for mixture models. Some observations are given below based on the conditions examined in this work.

- The results summarized in Chapter 4 show that AIC, HT-AIC, and Entropy are not useful for class enumeration in GMM studies because of their general 30%-90% incorrect identification. Others might be superior in different mixture models under conditions with different combinations of manipulated factors. In general, most indices would perform best in UGMM as Table 4.1 implies. More specifically, BIC, LMR_1V2 and BLRT_1V2 in linear GMM could work well; SACAIC, DBIC, LMR_1V2 and BLRT_1V2 in UGMM can provide sufficiently accurate identification on the number of latent class.

- Larger class separation can improve the performance of the useful indices. Table 4.2.1.1 and Table 4.2.1.2 indicate SACAIC and DBIC in UGMM, and LMR_1V2, and BLRT_1V2 in both UGMM and linear GMM can obtain sufficient rate of accuracy (over 95%) across class separation conditions.
- Sample size plays an important role in this process because it directly influences the performance of model indices and does so through other factors. If the sample size at hand is sufficiently large, for example 2,000, Table 4.2.2.4 indicates that most indices perform satisfyingly best in LPM. But, if the sample size varies from 400 to 1000, based on the conditions investigated here, UGMM together with SACAIC and DBIC, or linear GMM with LMR_1V2 could achieve satisfactory rates of accuracy for our purpose. As discussed in Section 4.3, three types of models with 2 SD class separations and seven-measure LPM demand larger sample size to achieve good rate of accuracy.
- The effect of the number of measures is highly associated with sample size. Increasing this factor does not necessarily improve the rate of accuracy. Instead, it might lower the performance of model selectors if the sample size is not sufficiently large. SACAIC, DBIC, LMR, and BLRT in UGMM, and BIC and BLRT_1V2 in linear GMM perform equally well (over 95%) under both conditions with 4 and 7 measures, respectively.
- The mixing proportion and within-class model specification set up in my

simulation design might be too mild to show a significant difference on the performance of model selectors in types of mixture models. More investigations are necessary for these two factors.

- Most fit indices used for class enumeration, more or less, perform better in the less restricted UGMM. This finding provides evidence supporting our conjecture that less restricted models might perform better in selecting the correct number of latent class for GMM prior to direct application of linear GMM, even when within-class model is appropriately specified. We could expect that the advantage of UGMM might be more distinct if the within-class model misspecification is more serious.

The practical suggestions this study could offer to the practitioners who use GMM is that they can try less restricted mixture models, UGMM first. If sample size is sufficiently large (e.g., 2000), LPM is also recommended for the same purpose. If different combinations of mixture models and model fit indices lead to the same number of latent class, researchers have more confidence about the result of class enumeration and then further consider what kind of growth function can fit the data; if these combinations indicate different number of latent classes, holding other conditions constant, the results from less restricted UGMM or LPM is more reliable. Moreover, researchers can make this decision by incorporating other information, such as substantial theory, or graphical inspection of data.

Based on several research works on procedures for applying GMM (Connell & Frye, 2006; Muthen, 2004; Wang & Bodner, 2007), Ram and Grimm (2009) viewed GMM an exploratory technique and formulated four steps for conducting a GMM analysis, in which a single-group growth curve model is obtained prior to class enumeration. However, as stated in Section 2.2, within-class model misspecification might lead to spurious latent classes. Due to the exploratory nature of applying GMM in practice, it is more reasonable to determine the number of latent classes before specifying the within-class model structure. Based on the current study, less restricted models are suggested to be used first to lower the chance of incorrect class enumeration. Figure 5 summarizes a "roadmap" for determining the number of latent class in GMM based on the conditions examined in this study.

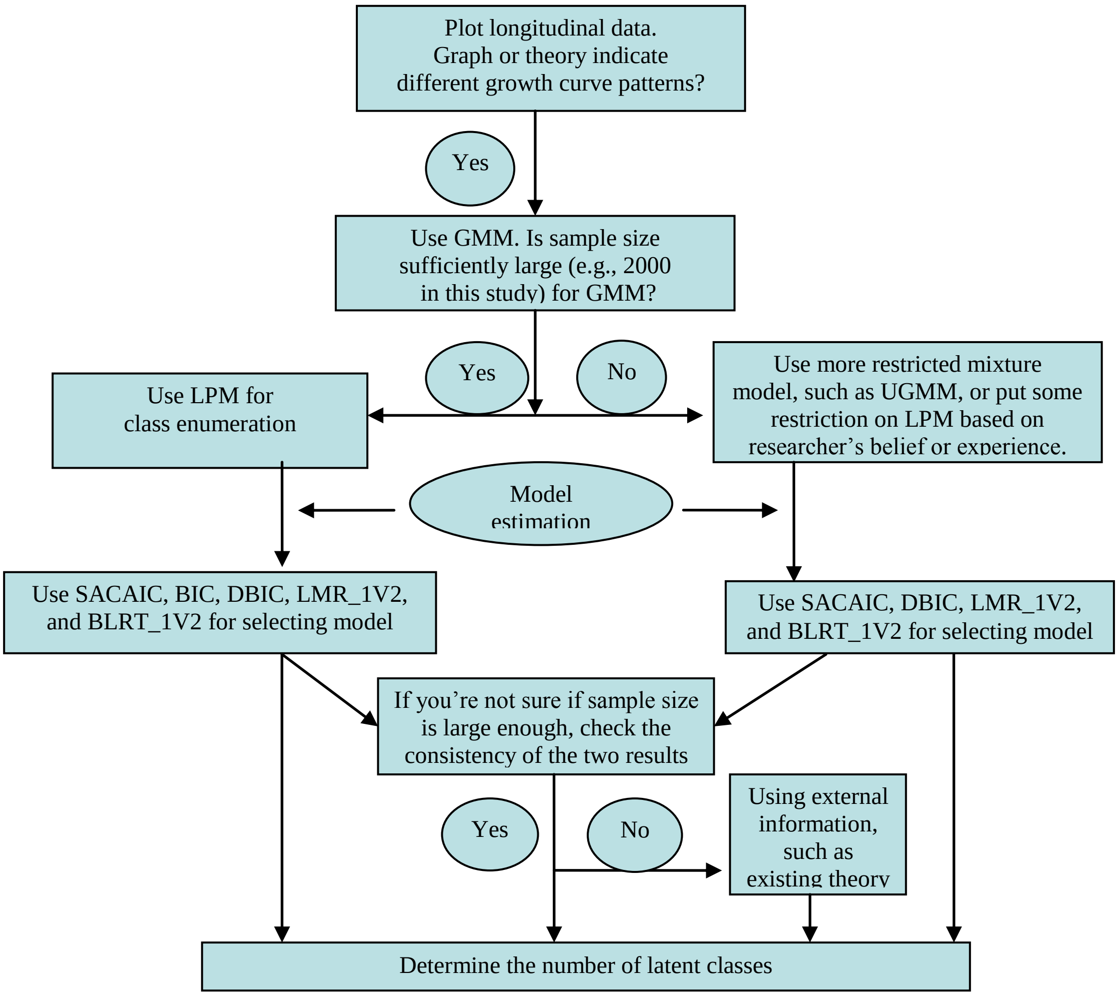

Figure 5. A roadmap for class enumeration in application of GMM

In sum, based on the conditions examined in this study, the less restricted mixture model, UGMM, can be considered as a promising way to partly solve class enumeration problems caused by within-class model misspecification because it can provide more a reliable result in selecting the correct number of classes than linear GMM. Surely this finding has important implications for class enumeration for other types of mixture models. But it needs further investigation to know how effective the less restricted model could play for the same purpose in different contexts.

Just like any other methodological studies, there are some limitations and associated possible future research directions in this study.

- Only two-class true model was used to generate data. Therefore, this study provides some information about how indices work to distinguish two-class from other class models when two-class model is true, but it does not tell how often they would still choose two-class model when a three- or four-class model is true. In other words, this study tells researchers about true positive and false negatives, but nothing about true negatives and false positives with respect to two-class model. To clarify this inquiry, more research needs to be done.
- As stated before, the manipulated settings for two design factors, mixing proportion and within-class model specification are too mild and so they do not have substantial effect on the performance of model fit indices in selecting the number of latent classes. More variations of the two current factors could be further investigated, such as more extreme proportion for minority group or larger nonlinear component.
- Due to time constraints, some other possible influential factors are not included in this simulation, such as correlation between latent intercept and slope factors and covariates for latent factors, etc. Usually the correlation between intercept and slope are correlated to some extent and so the degree of

correlation is worthy of further investigation. Although Tofighi and Enders (2008) results indicate covariates have detrimental effect on the class enumeration in linear GMM, their effect in less restricted mixture models, UGMM and LPM, are unknown. They might play a more important role in less restricted model because these models loosen the restrictions imposed to the variable relations and covariates can bring some useful information to facilitate researcher's understanding to the associations among variables and thus to more accurately identify the number of latent classes.

- UGMM is just one type of balancing model between the most unrestricted and restricted mixture models. Many other variations could be considered. Different mixture model could be used for different latent classes. For example, one class could follow linear growth function, while the other could use unstructured growth function; or one class could let all the parameter be freely estimated while the other put some equality constraints to some parameters.

Appendices A: Results for each condition listed in simulation design, as shown in Table 3.2

Table A 3. *Number of classes selected by each index in condition 1*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
					LPM (71 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	71	69	71	68	70	63	0	0	100	50	100	40
3 class	71	0	2	0	3	1	8	71	71	-	21	-	31
					UGMM (85 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	6	85	85	85	83	85	73	6	8	100	63	100	71
3 class	79	0	0	0	2	0	12	79	77	-	22	-	14
					Linear GMM (46 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	46	42	46	32	45	18	0	0	100	24	100	40
3 class	46	0	4	0	14	1	28	46	46	-	22	-	6

Table A 4. *Number of classes selected by each index in condition 2*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
					LPM (68 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	68	68	68	66	68	56	0	0	100	43	100	25
3 class	68	0	0	0	2	0	12	68	68	-	25	-	43
					UGMM (87 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	2	87	85	87	81	85	63	4	43	100	65	100	66
3 class	85	0	2	0	6	2	24	83	44	-	22	-	21
					Linear GMM (51 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	51	49	51	38	51	15	0	5	100	25	100	49
3 class	51	0	2	0	13	0	36	51	46	-	26	-	2

Table A 3. *Number of classes selected by each index in condition 3*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (69 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	69	68	69	67	69	62	0	2	100	51	100	33
3 class	69	0	1	0	2	0	7	69	67	-	18	-	36
UGMM (86 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	5	86	86	86	83	86	72	8	4	100	68	100	70
3 class	81	0	0	0	3	0	14	78	82	-	18	-	16
Linear GMM (54 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	54	54	54	42	54	28	0	0	100	27	100	47
3 class	54	0	0	0	12	0	26	54	54	-	27	-	7

Table A 4. *Number of classes selected by each index in condition 4*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (79 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	79	79	79	74	79	62	0	0	100	51	100	34
3 class	79	0	0	0	5	0	17	79	79	-	28	-	45
UGMM (92 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	1	92	88	92	65	89	38	1	69	100	48	100	52
3 class	91	0	4	0	27	3	54	91	23	-	44	-	40
Linear GMM (58 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	58	58	58	46	58	27	0	2	100	31	100	53
3 class	58	0	0	0	12	0	31	58	56	-	27	-	5

Table A 5. *Number of classes selected by each index in condition 5*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (90 converged replications for 3-class model)													
1 class	0	22	0	4	0	0	0	0	-	0	-	0	-
2 class	0	68	90	86	90	90	88	1	0	100	78	100	39
3 class	90	0	0	0	0	0	2	89	90	-	12	-	51
UGMM (95 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	86	95	95	95	95	95	94	86	5	100	90	100	88
3 class	9	0	0	0	0	0	1	9	90	-	3	-	5
Linear GMM (70 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	1	70	67	70	51	70	32	1	1	100	41	100	56
3 class	69	0	3	0	19	0	38	69	69	-	29	-	14

Table A 6. *Number of classes selected by each index in condition 6*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (91 converged replications for 3-class model)													
1 class	0	1	0	0	0	0	0	0	-	0	-	0	-
2 class	0	90	91	91	91	91	91	1	1	100	69	100	26
3 class	91	0	0	0	0	0	0	90	90	-	22	-	65
UGMM (97 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	71	97	97	97	96	97	95	74	4	100	95	100	84
3 class	26	0	0	0	1	0	2	23	93	-	2	-	13
Linear GMM (63 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	1	63	61	63	49	63	29	3	1	100	42	100	50
3 class	62	0	2	0	14	0	34	60	62	-	21	-	13

Table A 7. *Number of classes selected by each index in condition 7*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
						LPM (94 converged replications for 3-class model)							
1 class	0	5	0	0	0	0	0	0	-	0	-	0	-
2 class	0	89	94	94	94	94	94	3	1	100	84	100	46
3 class	94	0	0	0	0	0	0	91	93	-	10	-	48
						UGMM (95 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	67	95	95	95	94	95	93	72	1	100	84	100	80
3 class	28	0	0	0	1	0	2	23	94	-	7	-	11
						Linear GMM (73 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	1	73	73	73	60	73	44	2	3	100	53	100	60
3 class	72	0	0	0	13	0	29	71	70	-	20	-	13

Table A 8. *Number of classes selected by each index in condition 8*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
						LPM (92 converged replications for 3-class model)							
1 class	0	1	0	0	0	0	0	0	-	0	-	0	-
2 class	0	91	92	92	92	92	90	1	1	100	79	100	34
3 class	92	0	0	0	0	0	2	91	91	-	13	-	58
						UGMM (96 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	60	96	84	96	78	86	77	61	20	100	73	100	64
3 class	36	0	12	0	18	10	19	35	76	-	21	-	30
						Linear GMM (71 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	1	71	70	71	59	71	42	3	1	100	53	100	67
3 class	70	0	1	0	12	0	29	68	70	-	18	-	4

Table A 9. *Number of classes selected by each index in condition 9*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
						LPM (84 converged replications for 3-class model)							
1 class	0	6	0	2	0	0	0	0	-	0	-	0	-
2 class	0	78	84	82	72	84	77	3	1	100	54	100	42
3 class	84	0	0	0	12	0	7	81	82	-	29	-	42
						UGMM (92 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	1	-	0	-
2 class	21	92	92	92	86	92	87	26	14	99	81	100	81
3 class	71	0	0	0	6	0	5	66	77	-	10	-	10
						Linear GMM (55 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	55	50	55	24	54	30	0	0	100	24	100	49
3 class	55	0	5	0	31	1	25	55	55	-	31	-	6

Table A 10. *Number of classes selected by each index in condition 10*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
						LPM (92 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	10	92	91	92	74	92	82	10	0	100	59	100	44
3 class	82	0	1	0	18	0	10	82	92	-	32	-	48
						UGMM (98 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	23	98	96	98	87	97	88	32	43	100	85	100	85
3 class	75	0	2	0	11	1	10	66	55	-	13	-	13
						Linear GMM (78 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	27	78	72	78	44	75	44	27	5	100	36	100	72
3 class	51	0	6	0	34	3	34	51	73	-	42	-	6

Table A 11. *Number of classes selected by each index in condition 11*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
					LPM (81 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	1	-	0	-
2 class	0	81	81	81	70	81	75	0	3	99	56	100	41
3 class	81	0	0	0	11	0	6	81	78	-	25	-	40
					UGMM (92 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	21	92	92	92	85	92	89	25	12	100	81	100	86
3 class	71	0	0	0	7	0	3	67	80	-	11	-	6
					Linear GMM (53 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	53	48	53	28	52	32	0	1	100	25	100	49
3 class	53	0	5	0	25	1	21	53	52	-	28	-	4

Table A 12. *Number of classes selected by each index in condition 12*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
					LPM (92 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	13	92	92	92	77	92	82	13	1	100	64	100	36
3 class	79	0	0	0	15	0	10	79	91	-	28	-	56
					UGMM (100 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	1	-	0	-
2 class	12	100	99	100	84	99	88	14	51	99	76	100	79
3 class	88	0	1	0	16	1	12	86	49	-	24	-	21
					Linear GMM (77 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	19	77	68	77	47	72	51	19	5	100	37	100	64
3 class	58	0	9	0	30	5	26	58	72	-	40	-	13

Table A 13. *Number of classes selected by each index in condition 13*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
							LPM (49 converged replications for 3-class model)						
1 class	0	49	13	49	1	29	1	0	-	1	-	0	-
2 class	0	0	36	0	48	20	48	4	2	99	40	100	19
3 class	49	0	0	0	0	0	0	45	47	-	9	-	30
							UGMM (99 converged replications for 3-class model)						
1 class	0	20	0	4	0	0	0	0	-	0	-	0	-
2 class	95	79	99	95	99	99	99	96	1	100	95	100	95
3 class	4	0	0	0	0	0	0	3	98	-	2	-	2
							Linear GMM (70 converged replications for 3-class model)						
1 class	0	1	0	0	0	0	0	0	-	0	-	0	-
2 class	4	69	63	70	40	67	44	5	8	100	54	100	55
3 class	66	0	7	0	30	3	26	65	62	-	16	-	15

Table A 14. *Number of classes selected by each index in condition 14*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
							LPM (96 converged replications for 3-class model)						
1 class		95	1	86		8			-		-		-
2 class	3	1	93	9	93	86	94	7	2	100	71	100	34
3 class	93		2	1	3	2	2	89	94	-	25	-	62
							UGMM (99 converged replications for 3-class model)						
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	70	95	93	94	90	93	91	74	10	100	94	100	88
3 class	29	4	6	5	9	6	8	25	89	-	5	-	11
							Linear GMM (91 converged replications for 3-class model)						
1 class	0	0	0	0	0	0	0	0	-		-		-
2 class	9	86	76	85	51	80	61	9	5	100	57	100	72
3 class	82	5	15	6	40	11	30	82	86	-	34	-	19

Table A 15. *Number of classes selected by each index in condition 15*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (92 converged replications for 3-class model)													
1 class	0	92	6	88	1	16	1	0	-	2	-	0	-
2 class	0	0	86	4	91	76	91	9	5	98	82	100	38
3 class	92	0	0	0	0	0	0	83	87	-	10	-	54
UGMM (95 converged replications for 3-class model)													
1 class	0	3	0	1	0	0	0	0	-	0	-	0	-
2 class	74	92	95	94	93	95	94	76	2	100	79	100	74
3 class	21	0	0	0	2	0	1	19	93	-	9	-	14
Linear GMM (81 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	2	81	73	81	48	78	52	2	3	100	62	100	67
3 class	79	0	8	0	33	3	29	79	78	-	19	-	14

Table A 16 *Number of classes selected in condition 16*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (82 converged replications for 3-class model)													
1 class	0	78	1	62	0	2	0	0	-	1	-	0	-
2 class	0	4	81	20	81	80	82	2	1	99	70	100	20
3 class	82	0	0	0	1	0	0	80	80	-	12	-	62
UGMM (97 converged replications for 3-class model)													
1 class	0	1	0	1	0	0	0	0	-	1	-	0	-
2 class	71	96	97	96	92	97	94	74	12	99	86	100	70
3 class	26	0	0	0	5	0	3	23	85	-	4	-	20
Linear GMM (71 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	1	-	0	-
2 class	0	71	59	71	37	65	41	0	1	99	55	100	54
3 class	71	0	12	0	34	6	30	71	70	-	16	-	17

Table A 17. *Number of classes selected by each index in condition 17*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (80 converged replications for 3-class model)													
1 class	0	37	0	18	0	1	0	0	-	1	-	0	-
2 class	0	43	80	62	64	79	76	3	6	99	62	100	49
3 class	80	0	0	0	16	0	4	77	74	-	18	-	31
UGMM (95 converged replications for 3-class model)													
1 class	0	14	0	4	0	0	0	0	-	0	-	0	-
2 class	36	81	93	91	74	93	85	41	11	100	80	100	82
3 class	59	0	2	0	21	2	10	54	84	-	15	-	13
Linear GMM (61 converged replications for 3-class model)													
1 class	0	3	0	0	0	0	0	0	-	0	-	0	-
2 class	2	58	50	61	19	58	36	2	1	100	35	100	37
3 class	59	0	11	0	42	3	25	59	60	-	26	-	24

Table A 18. *Number of classes selected by each index in condition 18*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (82 converged replications for 3-class model)													
1 class	0	18							-		-		-
2 class	0	64	80	82	57	82	76		7	100	60	100	41
3 class	82	0	2	0	25		6	82	75	-	22	-	41
UGMM (97 converged replications for 3-class model)													
1 class	0	2	0	0	0	0	0	0	-	1	-	0	-
2 class	20	95	94	97	75	97	87	21	32	99	81	100	83
3 class	77	0	3	0	22	0	10	76	65	-	16	-	14
Linear GMM (70 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	70	48	70	20	60	38	0	3	100	29	100	61
3 class	70	0	22	0	50	10	32	70	67	-	41	-	9

Table A 19. *Number of classes selected by each index in condition 19*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (70 converged replications for 3-class model)													
1 class	0	11	0	4	0	0	0	0	-	0	-	0	-
2 class	0	59	69	66	46	69	62	1	2	99	48	100	29
3 class	70	0	1	0	24	1	8	69	68	-	22	-	41
UGMM (96 converged replications for 3-class model)													
1 class	0	4	0	1	0	0	0	0	-	1	-	0	-
2 class	40	92	94	95	83	95	93	44	13	99	83	100	83
3 class	56	0	2	0	13	1	3	52	83	-	12	-	12
Linear GMM (66 converged replications for 3-class model)													
1 class	0	1	0	0	0	0	0	0	-	0	-	0	-
2 class	0	65	55	66	19	58	38	0	2	100	35	100	60
3 class	66	0	11	0	47	8	28	66	64	-	31	-	6

Table A 20. *Number of classes selected by each index in condition 20*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (78 converged replications for 3-class model)													
1 class	0	4	0	0	0	0	0	0	-	0	-	0	-
2 class	0	74	76	78	51	77	69	0	6	100	53	100	32
3 class	78	0	2	0	27	1	9	78	72	-	25	-	46
UGMM (97 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	3	-	0	-
2 class	18	97	96	97	79	96	92	22	47	97	86	100	91
3 class	79	0	1	0	18	1	5	75	50	-	11	-	6
Linear GMM (62 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	1	-	0	-
2 class	0	62	53	62	15	58	37	0	2	99	35	100	57
3 class	62	0	9	0	47	4	25	62	60	-	27	-	5

Table A 21. *Number of classes selected by each index in condition 21*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (75 converged replications for 3-class model)													
1 class	0	75	58	75	10	69	37	0	-	5	-	12	-
2 class	0	0	17		61	6	38	8	2	95	62	88	32
3 class	75	0	0	0	4	0	0	67	73	-	13	-	43
UGMM (99 converged replications for 3-class model)													
1 class	0	79	1	40	0	2	0	0	-	0	-		-
2 class	93	20	97	59	97	96	98	94	2	100	92	100	92
3 class	6	0	1	0	2	1	1	5	97	-	5	-	5
Linear GMM (80 converged replications for 3-class model)													
1 class	0	11	0	5	0	0	0	0	-	0	-	2	-
2 class	3	69	73	75	37	76	58	3	6	100	63	98	71
3 class	77	0	7	0	43	4	22	77	74	-	17	-	9

Table A 22. *Number of classes selected by each index in condition 22*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (87 converged replications for 3-class model)													
1 class	0	87	30	87	1	45	10	0	-	2	-	0	-
2 class	0	0	57	0	82	42	76	15	3	98	73	100	30
3 class	87	0	0	0	4	0	1	72	84	-	14	-	57
UGMM (96 converged replications for 3-class model)													
1 class	0	42	0	11	0	1	0	0	-	0	-	0	-
2 class	81	54	96	85	94	95	96	82	5	100	92	100	82
3 class	15	0	0	0	2	0	0	14	91	-	1	-	11
Linear GMM (84 converged replications for 3-class model)													
1 class	0	3	0	1	0	0	0	0	-	0	-	0	-
2 class	4	81	69	83	42	78	60	5	6	100	65	100	69
3 class	80	0	15	0	42	6	24	79	78	-	19	-	15

Table A 23. *Number of classes selected by each index in condition 23*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (77 converged replications for 3-class model)													
1 class	0	77	49	76	6	62	20	0	-	17	-	3	-
2 class	0	0	28	1	70	15	57	11	5	82	62	96	35
3 class	77	0	0	0	2	0	0	66	71	-	15	-	43
UGMM (94 converged replications for 3-class model)													
1 class	0	56	0	25	0	1	0	0	-	2	-	2	-
2 class	83	38	94	69	93	93	93	83	0	97	79	98	75
3 class	11	0	0	0	1	0	1	11	93	-	7	-	11
Linear GMM (77 converged replications for 3-class model)													
1 class	0	6	0	2	0	0	0	0	-	0	-	0	-
2 class	1	71	59	74	37	66	48	1	3	100	56	100	57
3 class	76	0	18	1	40	11	29	76	74	-	21	-	20

Table A 24. *Number of classes selected by each index in condition 24*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (84 converged replications for 3-class model)													
1 class	0	84	16	83	0	29	3	0	-	9	-	0	-
2 class	0	0	68	1	83	55	81	8	5	91	73	100	27
3 class	84	0	0	0	1	0	0	76	78	-	10	-	57
UGMM (98 converged replications for 3-class model)													
1 class	0	31	0	6	0	0	0	0	-	2	-	0	-
2 class	83	67	97	92	95	98	97	85	7	98	87	100	76
3 class	15	0	1	0	3	0	1	13	91	-	4	-	15
Linear GMM (85 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	3	85	71	85	46	78	63	5	5	100	67	100	62
3 class	82	0	14	0	39	7	22	80	80	-	18	-	23

Table A 25. *Number of classes selected by each index in condition 25*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
					LPM (35 converged replications for 3-class model)								
1 class	0	35	3	29	0	4	3	0	-	15	-	16	-
2 class	2	0	28	6	13	29	28	3	6	84	22	83	17
3 class	33	0	4	0	22	2	4	32	29	-	13	-	18
					UGMM (86 converged replications for 3-class model)								
1 class	0	75	1	50	0	7	3	0	-	4	-	15	-
2 class	45	11	81	36	68	77	79	49	12	93	75	82	75
3 class	41	0	4	0	20	2	4	37	74	-	13	-	13
					Linear GMM (62 converged replications for 3-class model)								
1 class	0	42	0	27	0	4	1	0	-	4	-	8	-
2 class	1	20	43	35	12	48	43	1	4	96	37	92	56
3 class	61	0	19	0	50	10	18	61	58	-	25	-	6

Table A 26. *Number of classes selected by each index in condition 26*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
					LPM (59 converged replications for 3-class model)								
1 class	0	29	0	16	0	0	0	0	-	5	-	3	-
2 class	2	8	36	21	14	37	36	3	3	95	39	97	23
3 class	57	22	23	22	45	22	23	56	56	-	20	-	36
					UGMM (100 converged replications for 3-class model)								
1 class	0	66	0	43	0	2	0	0	-	4	-	2	-
2 class	42	34	94	57	75	96	95	47	32	96	89	98	93
3 class	58	0	6	0	25	2	5	53	68	-	11	-	7
					Linear GMM (98 converged replications for 3-class model)								
1 class	0	18	0	4	0	0	0	0	-	0	-	1	-
2 class	12	68	48	81	20	58	52	12	8	100	49	99	81
3 class	86	12	50	13	78	40	46	86	90	-	49	-	17

Table A 27. *Number of classes selected by each index in condition 27*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (29 converged replications for 3-class model)													
1 class	0	23	0	21	0	0	0	0	-	16	-	8	-
2 class	0	6	26	8	5	29	27	1	4	84	14	92	7
3 class	29	0	3	0	24	0	2	28	25	-	15	-	22
UGMM (85 converged replications for 3-class model)													
1 class	0	63	0	31	0	1	0	0	-	13	-	1	-
2 class	48	22	84	54	68	83	84	52	11	86	73	98	68
3 class	37	0	1	0	17	1	1	33	74	-	11	-	16
Linear GMM (67 converged replications for 3-class model)													
1 class	0	23	0	9	0	0	0	0	-	7	-	1	-
2 class	2	44	49	58	10	55	49	2	7	90	35	96	59
3 class	65	0	18	0	58	12	18	65	60	-	33	-	9

Table A 28. *Number of classes selected by each index in condition 28*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (43 converged replications for 3-class model)													
1 class	0	26	0	13	0	0	0	0	-	6	-	1	-
2 class	0	17	38	30	7	41	38	0	4	92	20	97	13
3 class	43	0	5	0	36	2	5	43	39	-	23	-	30
UGMM (92 converged replications for 3-class model)													
1 class	0	42	0	17	0	0	0	0	-	4	-	3	-
2 class	36	50	89	75	73	91	89	42	37	96	80	97	80
3 class	56	0	3	0	19	1	3	50	55	-	11	-	11
Linear GMM (64 converged replications for 3-class model)													
1 class	0	2		0	0	0	0	0	-	2	-	0	-
2 class	2	62	37	64	10	46	38	2	6	98	39	100	58
3 class	62	0	27	0	54	18	26	62	58	-	25	-	6

Table A 29. *Number of classes selected by each index in condition 29*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
					LPM (61 converged replications for 3-class model)								
1 class	0	61	61	61	25	61	61	1	-	55	-	67	-
2 class	4	0	0	0	31	0	0	17	22	39	54	27	32
3 class	57	0	0	0	11	0	0	43	39	-	13	-	35
					UGMM (95 converged replications for 3-class model)								
1 class	0	95	20	90	1	39	22	0	-	17	-	42	-
2 class	91	0	75	5	93	56	73	91	9	82	83	57	86
3 class	4	0	0	0	2	0	0	4	86	-	5	-	2
					Linear GMM (80 converged replications for 3-class model)								
1 class	0	72	1	53	0	7	3	0	-	4	-	39	-
2 class	4	8	57	27	28	61	59	7	3	95	60	60	64
3 class	76	0	22	0	53	12	18	73	77	-	21	-	17

Table A 30. *Number of classes selected by each index in condition 30*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
					LPM (78 converged replications for 3-class model)								
1 class	0	77	72	76	11	75	72	0	-	36	-	50	-
2 class	5	1	5	2	41	2	5	18	15	64	60	50	34
3 class	73	0	1	0	26	1	1	60	62	-	17	-	44
					UGMM (100 converged replications for 3-class model)								
1 class	0	97	9	88	1	21	11	0	-	7	-	19	-
2 class	90	2	90	11	96	78	88	93	7	93	93	81	89
3 class	10	1	1	1	3	1	1	7	93	-	3	-	7
					Linear GMM (97 converged replications for 3-class model)								
1 class	0	59	0	32	0	3	0	0	-	2	-	12	-
2 class	19	31	72	56	48	78	72	22	10	98	81	88	78
3 class	78	7	25	9	49	16	25	75	87	-	16	-	19

Table A 31. *Number of classes selected by each index in condition 31*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (67 converged replications for 3-class model)													
1 class	0	67	63	67	15	66	64	0	-	69	-	46	-
2 class	1	0	4	0	35	1	3	14	18	29	62	52	36
3 class	66	0	0	0	17	0	0	53	48	-	4	-	31
UGMM (95 converged replications for 3-class model)													
1 class	0	94	12	86	1	19	14	0	-	22	-	24	-
2 class	91	1	83	9	93	76	81	91	1	75	83	75	83
3 class	4	0	0	0	2	0	0	4	92	-	2	-	2
Linear GMM (80 converged replications for 3-class model)													
1 class	0	60	1	33	0	2	1	0	-	7	-	22	-
2 class	4	20	59	47	20	67	60	6	4	93	62	78	64
3 class	76	0	20	0	60	11	19	74	76	-	18	-	16

Table A 32. *Number of classes selected by each index in condition 32*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (75 converged replications for 3-class model)													
1 class	0	75	56	75	5	68	57	0	-	55	-	30	-
2 class	0	0	19	0	51	7	18	16	16	44	65	69	34
3 class	75	0	0	0	20	0	0	59	59	-	11	-	42
UGMM (100 converged replications for 3-class model)													
1 class	0	92	5	79	0	8	6	0	-	15	-	12	-
2 class	90	8	95	21	97	92	94	90	6	85	86	88	80
3 class	10	0	0	0	3	0	0	10	94	-	6	-	12
Linear GMM (79 converged replications for 3-class model)													
1 class	0	42	0	17	0	0	0	0	-	2	-	13	-
2 class	6	37	50	62	28	64	51	8	13	97	66	86	58
3 class	73	0	29	0	51	15	28	71	66	-	14	-	22

Table A 33. *Number of classes selected by each index in condition 33*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
						LPM (73 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	73	73	73	70	73	56	0	0	100	47	100	29
3 class	73	0	0	0	3	0	17	73	73	-	26	-	44
						UGMM (91 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	8	91	91	91	88	91	85	9	9	100	67	100	82
3 class	83	0	0	0	3	0	6	82	82	-	24	-	9
						Linear GMM (48 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	48	46	48	31	48	8	0	1	100	16	100	40
3 class	48	0	2	0	17	0	40	48	47	-	32	-	8

Table A 34. *Number of classes selected by each index in condition 34*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
						LPM (60 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	60	60	60	55	60	42	0	1	100	33	100	28
3 class	60	0	0	0	5	0	18	60	59	-	27	-	32
						UGMM (86 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	86	85	86	73	85	53	0	44	100	58	100	61
3 class	86	0	1	0	13	1	33	86	42	-	27	-	24
						Linear GMM (41 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	41	41	41	26	41	9	0	2	100	16	100	38
3 class	41	0	0	0	15	0	32	41	39	-	25	-	3

Table A 35. *Number of classes selected by each index in condition 35*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (75 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	75	75	75	73	75	64	0	4	100	54	100	37
3 class	75	0	0	0	2	0	11	75	71	-	21	-	38
UGMM (90 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	7	90	89	90	86	90	78	7	10	100	65	100	76
3 class	83	0	1	0	4	0	12	83	80	-	25	-	14
Linear GMM (48 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	48	45	48	37	48	23	0	0	100	19	100	42
3 class	48	0	3	0	11	0	25	48	48	-	29	-	6

Table A 36. *Number of classes selected by each index in condition 36*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (75 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	75	75	75	70	75	62	0	2	100	51	100	38
3 class	75	0	0	0	5	0	13	75	73	-	24	-	37
UGMM (92 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	92	79	90	53	82	33	0	77	100	42	100	40
3 class	92	0	13	2	39	10	59	92	16	-	50	-	52
Linear GMM (52 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	52	51	52	36	51	16	0	1	100	20	100	46
3 class	52	0	1	0	16	1	36	52	51	-	32	-	6

Table A 37. *Number of classes selected by each index in condition 37*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (83 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	83	83	83	83	83	82	0	2	100	73	100	38
3 class	83	0	0	0	0	0	1	83	81	-	10	-	45
UGMM (98 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	84	98	98	98	97	98	96	86	11	100	94	100	90
3 class	14	0	0	0	1	0	2	12	87	-	4	-	8
Linear GMM (55 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	55	51	55	38	53	15	0	0	100	35	100	36
3 class	55	0	4	0	17	2	40	55	55	-	20	-	19

Table A 38. *Number of classes selected by each index in condition 38*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (88 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	88	88	88	88	88	88	2	6	100	72	100	43
3 class	88	0	0	0	0	0	0	86	82	-	16	-	45
UGMM (96 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	76	96	96	96	95	96	92	80	7	100	84	100	79
3 class	20	0	0	0	1	0	4	16	89	-	12	-	17
Linear GMM (55 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	55	51	55	31	55	19	0	1	100	28	100	40
3 class	55	0	4	0	24	0	36	55	54	-	27	-	15

Table A 39. *Number of classes selected by each index in condition 39*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (89 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	89	89	89	89	89	88	0	4	100	81	100	38
3 class	89	0	0	0	0	0	1	89	85	-	8	-	51
UGMM (98 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	71	98	98	98	98	98	98	75	6	100	84	100	84
3 class	27	0	0	0	0	0	0	23	92	-	11	-	11
Linear GMM (64 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	64	64	64	42	64	29	0	1	100	44	100	46
3 class	64	0	0	0	22	0	35	64	63	-	20	-	18

Table A 40. *Number of classes selected by each index in condition 40*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (91 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0		-	0	-	0	-
2 class	0	91	91	91	91	91	91	0	3	100	78	100	48
3 class	91	0	0	0	0	0	0	91	88	-	13	-	43
UGMM (97 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	78	97	96	97	96	97	96	80	2	100	93	100	88
3 class	19	0	1	0	1	0	1	17	95	-	4	-	9
Linear GMM (60 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	2	60	59	60	43	60	28	3	3	100	43	100	45
3 class	58	0	1	0	17	0	32	57	57	-	17	-	15

Table A 41. *Number of classes selected by each index in condition 41*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (70 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	70	70	70	55	70	61	0	3	100	40	100	20
3 class	70	0	0	0	15	0	9	70	67	-	30	-	50
UGMM (90 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	3	90	88	90	79	89	83	6	13	100	75	100	76
3 class	87	0	2	0	11	1	7	84	77	-	15	-	14
Linear GMM (45 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	45	40	45	7	43	18	0	0	100	12	100	38
3 class	45	0	5	0	38	2	27	45	45	-	33	-	7

Table A 42. *Number of classes selected by each index in condition 42*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (81 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	81	80	81	60	81	66	0	3	100	46	100	27
3 class	81	0	1	0	21	0	15	81	78	-	35	-	54
UGMM (86 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	86	83	86	65	84	68	1	21	100	54	100	62
3 class	86	0	3	0	21	2	18	85	65	-	32	-	24
Linear GMM (49 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	49	42	49	17	44	18	0	1	100	17	100	44
3 class	49	0	7	0	32	5	31	49	48	-	32	-	5

Table A 43. *Number of classes selected by each index in condition 43*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (77 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	77	77	77	60	77	65	0	2	100	48	100	28
3 class	77	0	0	0	17	0	12	77	75	-	29	-	49
UGMM (92 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	9	92	92	92	78	92	83	12	16	100	74	100	81
3 class	83	0	0	0	14	0	9	80	76	-	18	-	11
Linear GMM (60 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	60	53	60	20	57	31	0	0	100	26	100	55
3 class	60	0	7	0	40	3	29	60	60	-	34	-	5

Table A 44. *Number of classes selected by each index in condition 44*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	0-	0	-	0	-
2 class	0	83	83	83	71	83	73	0	4	100	57	100	38
3 class	83	0	0	0	12	0	10	83	79	-	26	-	45
UGMM (converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	2	91	87	91	71	89	74	4	47	100	69	100	73
3 class	89	0	4	0	20	2	17	87	44	-	22	-	18
Linear GMM (converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	56	50	56	20	51	27	0	1	100	22	100	49
3 class	56	0	6	0	36	5	29	56	55	-	34	-	7

Table A 45. *Number of classes selected by each index in condition 45*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (90 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	7	90	90	90	90	90	4	10	100	77	100	39
3 class	90	83	0	0	0	0	0	86	80	-	13	-	51
UGMM (96 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	73	96	96	96	95	96	95	80	9	100	92	100	90
3 class	23	0	0	0	1	0	1	16	87	-	3	-	5
Linear GMM (75 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	75	62	75	33	72	42	1	5	100	50	100	46
3 class	75	0	13	0	42	3	33	74	70	-	25	-	29

Table A 46. *Number of classes selected by each index in condition 46*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (89 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	89	89	89	89	89	89	3	4	100	73	100	32
3 class	89	0	0	0	0	0	0	86	85	-	16	-	57
UGMM (95 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	76	95	95	95	95	95	95	80	4	100	92	100	82
3 class	19	0	0	0	0	0	0	15	91	-	3	-	13
Linear GMM (63 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	1	63	50	63	23	59	26	1	0	100	41	100	52
3 class	62	0	13	0	40	4	37	62	63	-	22	-	11

Table A 47. *Number of classes selected by each index in condition 47*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
					LPM (90 converged replications for 3-class model)								
1 class	0	4	0	1	0	0	0	0	-	0	-	0	-
2 class	0	86	90	89	90	90	90	4	9	100	79	100	37
3 class	90	0	0	0	0	0	0	86	81	-	11	-	53
					UGMM (96 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	76	96	96	96	96	96	96	78	0	100	87	100	89
3 class	20	0	0	0	0	0	0	18	96	-	6	-	4
					Linear GMM (72 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	3	72	65	72	42	69	43	6	6	100	57	100	58
3 class	69	0	7	0	30	3	29	66	66	-	15	-	14

Table A 48. *Number of classes selected by each index in condition 48*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
					LPM (93 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	93	93	93	92	93	92	1	5	100	86	100	43
3 class	93	0	0	0	1	0	1	92	88	-	7	-	50
					UGMM (99 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	86	99	99	99	98	99	99	87	2	100	95	100	88
3 class	13	0	0	0	1	0	0	12	97	-	2	-	9
					Linear GMM (76 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	76	69	76	40	71	44	3	7	100	52	100	23
3 class	76	0	7	0	36	5	32	73	69	-	24	-	53

Table A 49. *Number of classes selected by each index in condition 49*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	82	80	82	39	81	70	0	1	100	42	100	29
3 class	82	0	2	0	43	1	12	82	80	-	39	-	53
UGMM (92 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	20	92	90	92	82	91	88	22	22	100	82	100	81
3 class	72	0	2	0	10	1	4	70	70	-	10	-	11
Linear GMM (converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	55	33	55	8	45	15	0	0	100	25	100	51
3 class	55	0	22	0	47	10	40	55	55	-	30	-	4

Table A 50. *Number of classes selected by each index in condition 50*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (81 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	81	80	81	49	81	73	0	2	100	60	100	38
3 class	81	0	1	0	32	0	8	81	79	-	21	-	43
UGMM (89 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	6	89	86	89	60	87	76	8	27	100	69	100	70
3 class	83	0	3	0	29	2	13	81	62	-	19	-	18
Linear GMM (56 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	56	38	56	6	47	20	0	4	100	21	100	44
3 class	56	0	18	0	50	9	36	56	52	-	35	-	12

Table A 51. *Number of classes selected by each index in condition 51*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (69 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	69	66	69	36	69	59	0	5	100	44	100	27
3 class	69	0	3	0	33	0	10	69	64	-	25	-	42
UGMM (95 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	1	-	0	-
2 class	20	95	92	95	76	93	87	23	28	99	76	100	85
3 class	75	0	3	0	19	2	8	72	67	-	18	-	9
Linear GMM (67 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	67	51	67	12	59	36	0	2	100	28	100	29
3 class	67	0	16	0	55	8	31	67	65	-	39	-	38

Table A 52. *Number of classes selected by each index in condition 52*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (74 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	74	71	74	41	71	59	74	6	100	48	100	36
3 class	74	0	3	0	33	3	15	0	67	-	25	-	38
UGMM (93 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	10	93	91	93	70	91	85	12	44	100	74	100	77
3 class	83	0	2	0	23	2	8	81	49	-	18	-	15
Linear GMM (49 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	49	38	49	9	45	25	0	3	100	18	100	46
3 class	49	0	11	0	40	4	24	49	46	-	31	-	3

Table A 53. *Number of classes selected by each index in condition 53*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (85 converged replications for 3-class model)													
1 class	0	78	0	30	0	0	0	0	-	0	-	0	-
2 class	0	7	85	55	83	85	85	6	5	100	67	100	23
3 class	85	0	0	0	2	0	0	79	80	-	18	-	62
UGMM (95 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	80	95	95	95	95	95	95	82	12	100	94	100	91
3 class	15	0	0	0	0	0	0	13	83	-	1	-	4
Linear GMM (57 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	1	57	40	57	21	49	30	1	1	100	41	100	38
3 class	56	0	17	0	36	8	27	56	56	-	16	-	19

Table A 54. *Number of classes selected by each index in condition 54*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (87 converged replications for 3-class model)													
1 class	0	60	0	14	0	0	0	0	-	0	-	0	-
2 class	0	27	87	73	83	87	86	3	7	100	62	100	25
3 class	87	0	0	0	4	0	1	84	79	-	24	-	62
UGMM (94 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0		-	0	-	0	-
2 class	83	94	94	94	93	94	93	83	3	100	93	100	80
3 class	11	0	0	0	1	0	1	11	91	-	0	-	13
Linear GMM (85 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	3	75	55	75	28	63	39	3	4	100	53	100	48
3 class	72	0	20	0	47	12	36	72	71	-	22	-	27

Table A 55. *Number of classes selected by each index in condition 55*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
							LPM (81 converged replications for 3-class model)						
1 class	0	52	0	17	0	0	0	0	-	1	-	0	-
2 class	0	29	81	64	80	81	81	6	12	99	73	100	37
3 class	81	0	0	0	1	0	0	75	68	-	8	-	44
							UGMM (99 converged replications for 3-class model)						
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	84	99	99	99	99	99	99	85	0	100	86	100	85
3 class	15	0	0	0	0	0	0	14	99	-	5	-	6
							Linear GMM (71 converged replications for 3-class model)						
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	1	71	60	71	29	64	44	4	8	100	49	100	49
3 class	70	0	11	0	42	7	27	67	63	-	22	-	22

Table A 56. *Number of classes selected by each index in condition 56*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
							LPM (93 converged replications for 3-class model)						
1 class	0	43	0	11	0	0	0	0	-	0	-	0	-
2 class	0	50	93	82	89	93	93	10	14	100	79	100	34
3 class	93	0	0	0	4	0	0	83	79	-	14	-	59
							UGMM (99 converged replications for 3-class model)						
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	83	99	99	99	97	99	98	86	1	100	96	100	83
3 class	16	0	0	0	2	0	1	13	98	-	2	-	15
							Linear GMM (68 converged replications for 3-class model)						
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	3	68	61	68	27	63	38	4	5	100	47	100	48
3 class	65	0	7	0	41	5	30	64	63	-	21	-	20

Table A 57. *Number of classes selected by each index in condition 57*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (71 converged replications for 3-class model)													
1 class	0	15	0	2	0	0	0	0	-	0	-	0	-
2 class	0	56	64	69	17	71	64	1	5	100	45	100	26
3 class	71	0	7	0	54	0	7	70	65	-	25	-	45
UGMM (93 converged replications for 3-class model)													
1 class	0	2	0	0	0	0	0	0	-	0	-	0	-
2 class	36	91	89	93	61	92	91	39	32	100	81	100	81
3 class	57	0	4	0	32	1	2	54	61	-	12	-	12
Linear GMM (53 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	53	26	53	2	37	29	0	1	100	28	100	46
3 class	53	0	27	0	51	16	24	53	52	-	25	-	7

Table A 58. *Number of classes selected by each index in condition 58*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (80 converged replications for 3-class model)													
1 class	0	3	0	0	0	0	0	0	-	0	-	0	-
2 class	0	77	70	80	15	77	71	0	5	100	46	100	46
3 class	80	0	10	0	65	3	9	80	75	-	34	-	34
UGMM (91 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	-	1	-	0	-
2 class	15	91	88	91	61	90	89	20	39	99	82	100	82
3 class	76	0	3	0	30	1	2	71	51	-	8	-	8
Linear GMM (59 converged replications for 3-class model)													
1 class	0	0	0	0	0	0	0	0	0-	0	-	0	-
2 class	0	59	21	59	2	37	22	0	6	100	21	100	52
3 class	59	0	38	0	57	22	37	59	53	-	38	-	7

Table A 59. *Number of classes selected by each index in condition 59*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
						LPM (61 converged replications for 3-class model)							
1 class	0	4	0	1	0	0	0	0	-	1	-	0	-
2 class	1	57	52	60	11	59	54	1	7	99	39	100	23
3 class	60	0	9	0	50	2	7	60	53	-	21	-	38
						UGMM (96 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	38	96	91	96	77	94	91	43	37	100	82	100	86
3 class	58	0	5	0	19	2	5	53	59	-	13	-	9
						Linear GMM (58 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	58	29	58	1	36	31	0	0	100	33	100	51
3 class	58	0	29	0	57	22	27	58	58	-	25	-	7

Table A 60. *Number of classes selected by each index in condition 60*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
						LPM (58 converged replications for 3-class model)							
1 class	0	2	0	0	0	0	0	0	-	0	-	0	-
2 class	0	56	52	58	15	55	53	2	4	100	31	100	21
3 class	58	0	6	0	43	3	5	56	54	-	27	-	37
						UGMM (95 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	3	-	0	-
2 class	31	95	88	95	68	93	89	38	57	97	87	100	86
3 class	64	0	7	0	27	2	6	57	37	-	8	-	9
						Linear GMM (64 converged replications for 3-class model)							
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	0	63	29	63	7	42	31	0	2	100	29	100	57
3 class	64	1	35	1	57	22	33	64	62	-	35	-	7

Table A 61. *Number of classes selected by each index in condition 61*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (61 converged replications for 3-class model)													
1 class	0	61	14	61	1	28	15	0	-	6	-	2	-
2 class	0	0	47	0	34	33	46	11	0	94	49	98	15
3 class	61	0	0	0	26	0	0	50	61	-	12	-	46
UGMM (99 converged replications for 3-class model)													
1 class	0	32	0	7	0	0	0	0	-	0	-	0	-
2 class	94	67	99	92	97	99	99	95	2	100	93	100	91
3 class	5	0	0	0	2	0	0	4	97	-	3	-	5
Linear GMM (86 converged replications for 3-class model)													
1 class	0	1	0	0	0	0	0	0	-	0	-	0	-
2 class	7	85	62	86	20	75	65	9	9	100	67	100	65
3 class	79	0	24	0	66	11	21	77	77	-	19	-	21

Table A 62. *Number of classes selected by each index in condition 62*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
LPM (66 converged replications for 3-class model)													
1 class	0	66	8	64	0	17	10	0	-	6	-	4	-
2 class	0	0	58	2	43	49	56	7	5	94	54	96	12
3 class	66	0	0	0	23	0	0	59	61	-	12	-	54
UGMM (100 converged replications for 3-class model)													
1 class	0	23	0	9	0	0	0	0	-	1	-	0	-
2 class	87	77	100	91	97	100	100	89	2	99	98	100	92
3 class	13	0	0	0	3	0	0	11	98	-	2	-	8
Linear GMM (85 converged replications for 3-class model)													
1 class	0	2	0	0	0	0	0	0	-	0	-	0	-
2 class	3	83	55	85	20	74	57	6	11	100	64	100	65
3 class	82	0	30	0	65	11	28	79	74	-	21	-	20

Table A 63. *Number of classes selected by each index in condition 63*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
					LPM (62 converged replications for 3-class model)								
1 class	0	62	5	62	0	17	5	0	-	18	-	0	-
2 class	0	0	57	0	38	45	57	13	11	82	51	100	26
3 class	62	0	0	0	24	0	0	49	51	-	11	-	36
					UGMM (98 converged replications for 3-class model)								
1 class	0	26	0	6	0	0	0	0	-	2	-	2	-
2 class	96	72	98	92	96	98	98	96	0	98	100	86	84
3 class	2	0	0	0	2	0	0	2	98	-	0	-	4
					Linear GMM (81 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	6	81	60	81	25	67	61	8	12	100	65	100	63
3 class	75	0	21	0	56	14	20	73	69	-	16	-	18

Table A 64. *Number of classes selected by each index in condition 64*

	AIC	CAIC	SACAIC	BIC	SABIC	DBIC	HQ	HT-AIC	Entropy	LMR LRT (1 vs.2)	LMR LRT (2 vs.3)	BLRT (1 vs.2)	BLRT (2 vs.3)
					LPM (82 converged replications for 3-class model)								
1 class	0	82	4	79	0	10	4	0	-	10	-	10	-
2 class	0	0	78	3	59	72	78	5	10	90	58	90	25
3 class	82	0	0	0	23	0	0	77	72	-	23	-	57
					UGMM (99 converged replications for 3-class model)								
1 class	0	16	0	3	0	0	0	0	0	0	-	0	-
2 class	91	83	99	96	98	99	99	94	3	100	90	100	80
3 class	8	0	0	0	1	0	0	5	96	-	3	-	13
					Linear GMM (85 converged replications for 3-class model)								
1 class	0	0	0	0	0	0	0	0	-	0	-	0	-
2 class	8	85	58	85	26	68	58	9	15	100	66	100	64
3 class	77	0	27	0	59	17	27	76	70	-	19	-	21

Appendices B: Two-way ANOVA Results

Table B1: Types of mixture model X Class separation

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	33223.229[a]	5	6644.646	15.057	.000
	CAIC	31625.609[b]	5	6325.122	9.633	.000
	SACAIC	6203.792[c]	5	1240.758	9.033	.000
	BIC	23146.417[d]	5	4629.283	9.930	.000
	SABIC	28756.047[e]	5	5751.209	26.883	.000
	DBIC	5707.375[f]	5	1141.475	7.367	.000
	HQ	18725.062[g]	5	3745.012	29.259	.000
	HT_AIC	30704.688[h]	5	6140.938	13.159	.000
	Entropy	7802.417[i]	5	1560.483	6.251	.000
	LMR_1V2	1953.187[j]	5	390.637	6.118	.000
	LMR_2V3	6765.089[k]	5	1353.018	18.015	.000
	BLRT_1V2	1338.417[l]	5	267.683	3.664	.003
	BLRT_2V3	44191.875[m]	5	8838.375	104.940	.000
Intercept	AIC	284284.083	1	284284.083	644.190	.000
	CAIC	1346197.547	1	1346197.547	2.050E3	.000
	SACAIC	1625456.021	1	1625456.021	1.183E4	.000
	BIC	1508752.083	1	1508752.083	3.236E3	.000
	SABIC	1196850.422	1	1196850.422	5.594E3	.000
	DBIC	1671786.750	1	1671786.750	1.079E4	.000
	HQ	1357777.687	1	1357777.687	1.061E4	.000
	HT_AIC	322752.000	1	322752.000	691.585	.000
	Entropy	166970.021	1	166970.021	668.807	.000

	LMR_1V2	1826370.187	1	1826370.187	2.860E4	.000
	LMR_2V3	1266362.755	1	1266362.755	1.686E4	.000
	BLRT_1V2	1826760.333	1	1826760.333	2.500E4	.000
	BLRT_2V3	1119046.687	1	1119046.687	1.329E4	.000
type_mixture	AIC	32027.823	2	16013.911	36.288	.000
	CAIC	15421.594	2	7710.797	11.743	.000
	SACAIC	3613.948	2	1806.974	13.155	.000
	BIC	10543.510	2	5271.755	11.308	.000
	SABIC	28475.094	2	14237.547	66.550	.000
	DBIC	1882.781	2	941.391	6.076	.003
	HQ	17449.031	2	8724.516	68.162	.000
	HT_AIC	29551.344	2	14775.672	31.661	.000
	Entropy	6464.823	2	3232.411	12.948	.000
	LMR_1V2	804.500	2	402.250	6.300	.002
	LMR_2V3	6653.323	2	3326.661	44.295	.000
	BLRT_1V2	214.542	2	107.271	1.468	.233
	BLRT_2V3	43881.031	2	21940.516	260.506	.000
class_sepa	AIC	33.333	1	33.333	.076	.784
	CAIC	14822.755	1	14822.755	22.575	.000
	SACAIC	892.687	1	892.687	6.499	.012
	BIC	10800.000	1	10800.000	23.165	.000
	SABIC	254.380	1	254.380	1.189	.277
	DBIC	1850.083	1	1850.083	11.941	.001
	HQ	212.521	1	212.521	1.660	.199
	HT_AIC	22.687	1	22.687	.049	.826
	Entropy	341.333	1	341.333	1.367	.244
	LMR_1V2	728.521	1	728.521	11.409	.001
	LMR_2V3	81.380	1	81.380	1.084	.299
	BLRT_1V2	990.083	1	990.083	13.552	.000
	BLRT_2V3	200.083	1	200.083	2.376	.125

type_mixture * class_sepa	AIC	1162.073	2	581.036	1.317	.271
	CAIC	1381.260	2	690.630	1.052	.351
	SACAIC	1697.156	2	848.578	6.178	.003
	BIC	1802.906	2	901.453	1.934	.148
	SABIC	26.573	2	13.286	.062	.940
	DBIC	1974.510	2	987.255	6.372	.002
	HQ	1063.510	2	531.755	4.154	.017
	HT_AIC	1130.656	2	565.328	1.211	.300
	Entropy	996.260	2	498.130	1.995	.139
	LMR_1V2	420.167	2	210.083	3.290	.039
	LMR_2V3	30.385	2	15.193	.202	.817
	BLRT_1V2	133.792	2	66.896	.916	.402
	BLRT_2V3	110.760	2	55.380	.658	.519
Error	AIC	82082.688	186	441.305		
	CAIC	122127.844	186	656.601		
	SACAIC	25548.188	186	137.356		
	BIC	86715.500	186	466.212		
	SABIC	39792.531	186	213.938		
	DBIC	28817.875	186	154.935		
	HQ	23807.250	186	127.996		
	HT_AIC	86803.312	186	466.684		
	Entropy	46435.562	186	249.654		
	LMR_1V2	11876.625	186	63.853		
	LMR_2V3	13969.156	186	75.103		
	BLRT_1V2	13589.250	186	73.060		
	BLRT_2V3	15665.437	186	84.223		
Total	AIC	399590.000	192			
	CAIC	1499951.000	192			
	SACAIC	1657208.000	192			
	BIC	1618614.000	192			
	SABIC	1265399.000	192			
	DBIC	1706312.000	192			
	HQ	1400310.000	192			
	HT_AIC	440260.000	192			

	Entropy	221208.000	192
	LMR_1V2	1840200.000	192
	LMR_2V3	1287097.000	192
	BLRT_1V2	1841688.000	192
	BLRT_2V3	1178904.000	192
Corrected Total	AIC	115305.917	191
	CAIC	153753.453	191
	SACAIC	31751.979	191
	BIC	109861.917	191
	SABIC	68548.578	191
	DBIC	34525.250	191
	HQ	42532.312	191
	HT_AIC	117508.000	191
	Entropy	54237.979	191
	LMR_1V2	13829.812	191
	LMR_2V3	20734.245	191
	BLRT_1V2	14927.667	191
	BLRT_2V3	59857.313	191

a. R Squared = .288 (Adjusted R Squared = .269)

b. R Squared = .206 (Adjusted R Squared = .184)

c. R Squared = .195 (Adjusted R Squared = .174)

d. R Squared = .211 (Adjusted R Squared = .189)

e. R Squared = .419 (Adjusted R Squared = .404)

f. R Squared = .165 (Adjusted R Squared = .143)

g. R Squared = .440 (Adjusted R Squared = .425)

h. R Squared = .261 (Adjusted R Squared = .241)

i. R Squared = .144 (Adjusted R Squared = .121)

j. R Squared = .141 (Adjusted R Squared = .118)

k. R Squared = .326 (Adjusted R Squared = .308)

l. R Squared = .090 (Adjusted R Squared = .065)

m. R Squared = .738 (Adjusted R Squared = .731)

Table B2: Types of mixture model X Sample size

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	43257.167[a]	11	3932.470	9.825	.000
	CAIC	58971.391[b]	11	5361.036	10.181	.000
	SACAIC	14566.729[c]	11	1324.248	13.870	.000
	BIC	36752.792[d]	11	3341.163	8.226	.000
	SABIC	54168.016[e]	11	4924.365	61.638	.000
	DBIC	10489.625[f]	11	953.602	7.141	.000
	HQ	21944.062[g]	11	1994.915	17.441	.000
	HT_AIC	42834.625[h]	11	3894.057	9.387	.000
	Entropy	15923.729[i]	11	1447.612	6.801	.000
	LMR_1V2	5054.687[j]	11	459.517	9.426	.000
	LMR_2V3	8086.932[k]	11	735.176	10.463	.000
	BLRT_1V2	3840.792[l]	11	349.163	5.669	.000
	BLRT_2V3	46421.313[m]	11	4220.119	56.536	.000
Intercept	AIC	284284.083	1	284284.083	710.229	.000
	CAIC	1346197.547	1	1346197.547	2.557E3	.000
	SACAIC	1625456.021	1	1625456.021	1.703E4	.000
	BIC	1508752.083	1	1508752.083	3.715E3	.000
	SABIC	1196850.422	1	1196850.422	1.498E4	.000
	DBIC	1671786.750	1	1671786.750	1.252E4	.000
	HQ	1357777.688	1	1357777.688	1.187E4	.000
	HT_AIC	322752.000	1	322752.000	777.993	.000
	Entropy	166970.021	1	166970.021	784.424	.000
	LMR_1V2	1826370.188	1	1826370.188	3.746E4	.000
	LMR_2V3	1266362.755	1	1266362.755	1.802E4	.000
	BLRT_1V2	1826760.333	1	1826760.333	2.966E4	.000
	BLRT_2V3	1119046.688	1	1119046.688	1.499E4	.000
type_mixture	AIC	32027.823	2	16013.911	40.008	.000
	CAIC	15421.594	2	7710.797	14.644	.000
	SACAIC	3613.948	2	1806.974	18.926	.000
	BIC	10543.510	2	5271.755	12.979	.000

	SABIC	28475.094	2	14237.547	178.210	.000
	DBIC	1882.781	2	941.391	7.050	.001
	HQ	17449.031	2	8724.516	76.277	.000
	HT_AIC	29551.344	2	14775.672	35.617	.000
	Entropy	6464.823	2	3232.411	15.186	.000
	LMR_1V2	804.500	2	402.250	8.251	.000
	LMR_2V3	6653.323	2	3326.661	47.346	.000
	BLRT_1V2	214.542	2	107.271	1.742	.178
	BLRT_2V3	43881.031	2	21940.516	293.934	.000
N	AIC	3562.875	3	1187.625	2.967	.033
	CAIC	34969.766	3	11656.589	22.137	.000
	SACAIC	8150.104	3	2716.701	28.455	.000
	BIC	21002.958	3	7000.986	17.237	.000
	SABIC	18990.391	3	6330.130	79.234	.000
	DBIC	6812.458	3	2270.819	17.006	.000
	HQ	955.271	3	318.424	2.784	.042
	HT_AIC	4743.292	3	1581.097	3.811	.011
	Entropy	2017.771	3	672.590	3.160	.026
	LMR_1V2	2533.104	3	844.368	17.320	.000
	LMR_2V3	480.057	3	160.019	2.277	.081
	BLRT_1V2	3122.375	3	1040.792	16.898	.000
	BLRT_2V3	868.188	3	289.396	3.877	.010
type_mixture * N	AIC	7666.469	6	1277.745	3.192	.005
	CAIC	8580.031	6	1430.005	2.716	.015
	SACAIC	2802.677	6	467.113	4.893	.000
	BIC	5206.323	6	867.720	2.136	.051
	SABIC	6702.531	6	1117.089	13.982	.000
	DBIC	1794.385	6	299.064	2.240	.041
	HQ	3539.760	6	589.960	5.158	.000
	HT_AIC	8539.990	6	1423.332	3.431	.003
	Entropy	7441.135	6	1240.189	5.826	.000
	LMR_1V2	1717.083	6	286.181	5.870	.000
	LMR_2V3	953.552	6	158.925	2.262	.040
	BLRT_1V2	503.875	6	83.979	1.363	.232

	BLRT_2V3	1672.094	6	278.682	3.733	.002
Error	AIC	72048.750	180	400.271		
	CAIC	94782.062	180	526.567		
	SACAIC	17185.250	180	95.474		
	BIC	73109.125	180	406.162		
	SABIC	14380.562	180	79.892		
	DBIC	24035.625	180	133.531		
	HQ	20588.250	180	114.379		
	HT_AIC	74673.375	180	414.852		
	Entropy	38314.250	180	212.857		
	LMR_1V2	8775.125	180	48.751		
	LMR_2V3	12647.312	180	70.263		
	BLRT_1V2	11086.875	180	61.594		
	BLRT_2V3	13436.000	180	74.644		
Total	AIC	399590.000	192			
	CAIC	1499951.000	192			
	SACAIC	1657208.000	192			
	BIC	1618614.000	192			
	SABIC	1265399.000	192			
	DBIC	1706312.000	192			
	HQ	1400310.000	192			
	HT_AIC	440260.000	192			
	Entropy	221208.000	192			
	LMR_1V2	1840200.000	192			
	LMR_2V3	1287097.000	192			
	BLRT_1V2	1841688.000	192			
	BLRT_2V3	1178904.000	192			
Corrected Total	AIC	115305.917	191			
	CAIC	153753.453	191			
	SACAIC	31751.979	191			
	BIC	109861.917	191			
	SABIC	68548.578	191			
	DBIC	34525.250	191			
	HQ	42532.312	191			

HT_AIC	117508.000	191
Entropy	54237.979	191
LMR_1V2	13829.812	191
LMR_2V3	20734.245	191
BLRT_1V2	14927.667	191
BLRT_2V3	59857.313	191

a. R Squared = .375 (Adjusted R Squared = .337)
b. R Squared = .384 (Adjusted R Squared = .346)
c. R Squared = .459 (Adjusted R Squared = .426)
d. R Squared = .335 (Adjusted R Squared = .294)
e. R Squared = .790 (Adjusted R Squared = .777)
f. R Squared = .304 (Adjusted R Squared = .261)
g. R Squared = .516 (Adjusted R Squared = .486)
h. R Squared = .365 (Adjusted R Squared = .326)
i. R Squared = .294 (Adjusted R Squared = .250)
j. R Squared = .365 (Adjusted R Squared = .327)
k. R Squared = .390 (Adjusted R Squared = .353)
l. R Squared = .257 (Adjusted R Squared = .212)
m. R Squared = .776 (Adjusted R Squared = .762)

Table B3: Types of mixture model X Number of measures

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	91773.792[a]	5	18354.758	145.078	.000
	CAIC	45425.172[b]	5	9085.034	15.599	.000
	SACAIC	5774.229[c]	5	1154.846	8.269	.000
	BIC	31196.979[d]	5	6239.396	14.753	.000
	SABIC	35107.922[e]	5	7021.584	39.055	.000
	DBIC	6823.250[f]	5	1364.650	9.163	.000
	HQ	18760.062[g]	5	3752.012	29.357	.000
	HT_AIC	85825.313[h]	5	17165.063	100.771	.000
	Entropy	24668.354[i]	5	4933.671	31.034	.000
	LMR_1V2	1781.562[j]	5	356.312	5.501	.000
	LMR_2V3	13890.526[k]	5	2778.105	75.504	.000
	BLRT_1V2	1009.854[l]	5	201.971	2.699	.022

	BLRT_2V3	47571.813[m]	5	9514.363	144.046	.000
Intercept	AIC	284284.083	1	284284.083	2.247E3	.000
	CAIC	1346197.547	1	1346197.547	2.311E3	.000
	SACAIC	1625456.021	1	1625456.021	1.164E4	.000
	BIC	1508752.083	1	1508752.083	3.567E3	.000
	SABIC	1196850.422	1	1196850.422	6.657E3	.000
	DBIC	1671786.750	1	1671786.750	1.122E4	.000
	HQ	1357777.687	1	1357777.687	1.062E4	.000
	HT_AIC	322752.000	1	322752.000	1.895E3	.000
	Entropy	166970.021	1	166970.021	1.050E3	.000
	LMR_1V2	1826370.187	1	1826370.187	2.820E4	.000
	LMR_2V3	1266362.755	1	1266362.755	3.442E4	.000
	BLRT_1V2	1826760.333	1	1826760.333	2.441E4	.000
	BLRT_2V3	1119046.687	1	1119046.687	1.694E4	.000
type_mixture	AIC	32027.823	2	16013.911	126.575	.000
	CAIC	15421.594	2	7710.797	13.239	.000
	SACAIC	3613.948	2	1806.974	12.938	.000
	BIC	10543.510	2	5271.755	12.465	.000
	SABIC	28475.094	2	14237.547	79.191	.000
	DBIC	1882.781	2	941.391	6.321	.002
	HQ	17449.031	2	8724.516	68.263	.000
	HT_AIC	29551.344	2	14775.672	86.744	.000
	Entropy	6464.823	2	3232.411	20.333	.000
	LMR_1V2	804.500	2	402.250	6.210	.002
	LMR_2V3	6653.323	2	3326.661	90.413	.000
	BLRT_1V2	214.542	2	107.271	1.434	.241
	BLRT_2V3	43881.031	2	21940.516	332.175	.000
measure	AIC	5764.083	1	5764.083	45.560	.000
	CAIC	19060.255	1	19060.255	32.727	.000
	SACAIC	588.000	1	588.000	4.210	.042
	BIC	12033.333	1	12033.333	28.452	.000
	SABIC	4456.380	1	4456.380	24.787	.000
	DBIC	1764.187	1	1764.187	11.845	.001
	HQ	330.750	1	330.750	2.588	.109

	HT_AIC	7105.333	1	7105.333	41.713	.000
	Entropy	13534.083	1	13534.083	85.133	.000
	LMR_1V2	414.187	1	414.187	6.394	.012
	LMR_2V3	7190.755	1	7190.755	195.432	.000
	BLRT_1V2	652.687	1	652.687	8.723	.004
	BLRT_2V3	1131.021	1	1131.021	17.123	.000
type_mixture * measure	AIC	53981.885	2	26990.943	213.339	.000
	CAIC	10943.323	2	5471.661	9.395	.000
	SACAIC	1572.281	2	786.141	5.629	.004
	BIC	8620.135	2	4310.068	10.191	.000
	SABIC	2176.448	2	1088.224	6.053	.003
	DBIC	3176.281	2	1588.141	10.663	.000
	HQ	980.281	2	490.141	3.835	.023
	HT_AIC	49168.635	2	24584.318	144.328	.000
	Entropy	4669.448	2	2334.724	14.686	.000
	LMR_1V2	562.875	2	281.438	4.345	.014
	LMR_2V3	46.448	2	23.224	.631	.533
	BLRT_1V2	142.625	2	71.313	.953	.387
	BLRT_2V3	2559.760	2	1279.880	19.377	.000
Error	AIC	23532.125	186	126.517		
	CAIC	108328.281	186	582.410		
	SACAIC	25977.750	186	139.665		
	BIC	78664.938	186	422.930		
	SABIC	33440.656	186	179.788		
	DBIC	27702.000	186	148.935		
	HQ	23772.250	186	127.808		
	HT_AIC	31682.687	186	170.337		
	Entropy	29569.625	186	158.976		
	LMR_1V2	12048.250	186	64.776		
	LMR_2V3	6843.719	186	36.794		
	BLRT_1V2	13917.812	186	74.827		
	BLRT_2V3	12285.500	186	66.051		
Total	AIC	399590.000	192			
	CAIC	1499951.000	192			

	SACAIC	1657208.000	192
	BIC	1618614.000	192
	SABIC	1265399.000	192
	DBIC	1706312.000	192
	HQ	1400310.000	192
	HT_AIC	440260.000	192
	Entropy	221208.000	192
	LMR_1V2	1840200.000	192
	LMR_2V3	1287097.000	192
	BLRT_1V2	1841688.000	192
	BLRT_2V3	1178904.000	192
Corrected Total	AIC	115305.917	191
	CAIC	153753.453	191
	SACAIC	31751.979	191
	BIC	109861.917	191
	SABIC	68548.578	191
	DBIC	34525.250	191
	HQ	42532.312	191
	HT_AIC	117508.000	191
	Entropy	54237.979	191
	LMR_1V2	13829.812	191
	LMR_2V3	20734.245	191
	BLRT_1V2	14927.667	191
	BLRT_2V3	59857.313	191

a. R Squared = .796 (Adjusted R Squared = .790)

b. R Squared = .295 (Adjusted R Squared = .277)

c. R Squared = .182 (Adjusted R Squared = .160)

d. R Squared = .284 (Adjusted R Squared = .265)

e. R Squared = .512 (Adjusted R Squared = .499)

f. R Squared = .198 (Adjusted R Squared = .176)

g. R Squared = .441 (Adjusted R Squared = .426)

h. R Squared = .730 (Adjusted R Squared = .723)

i. R Squared = .455 (Adjusted R Squared = .440)

j. R Squared = .129 (Adjusted R Squared = .105)

k. R Squared = .670 (Adjusted R Squared = .661)

l. R Squared = .068 (Adjusted R Squared = .043)
m. R Squared = .795 (Adjusted R Squared = .789)

Table B4: Types of mixture model X Mixing proportions

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	32176.042[a]	5	6435.208	14.399	.000
	CAIC	16373.484[b]	5	3274.697	4.434	.001
	SACAIC	3845.417[c]	5	769.083	5.126	.000
	BIC	10821.229[d]	5	2164.246	4.064	.002
	SABIC	28642.609[e]	5	5728.522	26.700	.000
	DBIC	2030.125[f]	5	406.025	2.324	.045
	HQ	17763.500[g]	5	3552.700	26.679	.000
	HT_AIC	29729.375[h]	5	5945.875	12.599	.000
	Entropy	6654.604[i]	5	1330.921	5.202	.000
	LMR_1V2	909.687[j]	5	181.937	2.619	.026
	LMR_2V3	6983.026[k]	5	1396.605	18.891	.000
	BLRT_1V2	292.167[l]	5	58.433	.743	.592
	BLRT_2V3	44055.625[m]	5	8811.125	103.715	.000
Intercept	AIC	284284.083	1	284284.083	636.075	.000
	CAIC	1346197.547	1	1346197.547	1.823E3	.000
	SACAIC	1625456.021	1	1625456.021	1.083E4	.000
	BIC	1508752.083	1	1508752.083	2.833E3	.000
	SABIC	1196850.422	1	1196850.422	5.578E3	.000
	DBIC	1671786.750	1	1671786.750	9.569E3	.000
	HQ	1357777.687	1	1357777.687	1.020E4	.000
	HT_AIC	322752.000	1	322752.000	683.901	.000
	Entropy	166970.021	1	166970.021	652.674	.000
	LMR_1V2	1826370.187	1	1826370.187	2.629E4	.000
	LMR_2V3	1266362.755	1	1266362.755	1.713E4	.000
	BLRT_1V2	1826760.333	1	1826760.333	2.322E4	.000
	BLRT_2V3	1119046.687	1	1119046.687	1.317E4	.000
type_mixture	AIC	32027.823	2	16013.911	35.831	.000
	CAIC	15421.594	2	7710.797	10.440	.000
	SACAIC	3613.948	2	1806.974	12.044	.000
	BIC	10543.510	2	5271.755	9.900	.000

	SABIC	28475.094	2	14237.547	66.361	.000
	DBIC	1882.781	2	941.391	5.388	.005
	HQ	17449.031	2	8724.516	65.516	.000
	HT_AIC	29551.344	2	14775.672	31.309	.000
	Entropy	6464.823	2	3232.411	12.635	.000
	LMR_1V2	804.500	2	402.250	5.791	.004
	LMR_2V3	6653.323	2	3326.661	44.997	.000
	BLRT_1V2	214.542	2	107.271	1.363	.258
	BLRT_2V3	43881.031	2	21940.516	258.259	.000
mix_prop	AIC	126.750	1	126.750	.284	.595
	CAIC	888.380	1	888.380	1.203	.274
	SACAIC	150.521	1	150.521	1.003	.318
	BIC	247.521	1	247.521	.465	.496
	SABIC	53.130	1	53.130	.248	.619
	DBIC	114.083	1	114.083	.653	.420
	HQ	105.021	1	105.021	.789	.376
	HT_AIC	31.688	1	31.688	.067	.796
	Entropy	67.687	1	67.687	.265	.608
	LMR_1V2	72.521	1	72.521	1.044	.308
	LMR_2V3	32.505	1	32.505	.440	.508
	BLRT_1V2	70.083	1	70.083	.891	.347
	BLRT_2V3	56.333	1	56.333	.663	.417
type_mixture * mix_prop	AIC	21.469	2	10.734	.024	.976
	CAIC	63.510	2	31.755	.043	.958
	SACAIC	80.948	2	40.474	.270	.764
	BIC	30.198	2	15.099	.028	.972
	SABIC	114.385	2	57.193	.267	.766
	DBIC	33.260	2	16.630	.095	.909
	HQ	209.448	2	104.724	.786	.457
	HT_AIC	146.344	2	73.172	.155	.856
	Entropy	122.094	2	61.047	.239	.788
	LMR_1V2	32.667	2	16.333	.235	.791
	LMR_2V3	297.198	2	148.599	2.010	.137
	BLRT_1V2	7.542	2	3.771	.048	.953

	BLRT_2V3	118.260	2	59.130	.696	.500
Error	AIC	83129.875	186	446.935		
	CAIC	137379.969	186	738.602		
	SACAIC	27906.563	186	150.035		
	BIC	99040.687	186	532.477		
	SABIC	39905.969	186	214.548		
	DBIC	32495.125	186	174.705		
	HQ	24768.812	186	133.166		
	HT_AIC	87778.625	186	471.928		
	Entropy	47583.375	186	255.825		
	LMR_1V2	12920.125	186	69.463		
	LMR_2V3	13751.219	186	73.931		
	BLRT_1V2	14635.500	186	78.685		
	BLRT_2V3	15801.687	186	84.955		
Total	AIC	399590.000	192			
	CAIC	1499951.000	192			
	SACAIC	1657208.000	192			
	BIC	1618614.000	192			
	SABIC	1265399.000	192			
	DBIC	1706312.000	192			
	HQ	1400310.000	192			
	HT_AIC	440260.000	192			
	Entropy	221208.000	192			
	LMR_1V2	1840200.000	192			
	LMR_2V3	1287097.000	192			
	BLRT_1V2	1841688.000	192			
	BLRT_2V3	1178904.000	192			
Corrected Total	AIC	115305.917	191			
	CAIC	153753.453	191			
	SACAIC	31751.979	191			
	BIC	109861.917	191			
	SABIC	68548.578	191			
	DBIC	34525.250	191			
	HQ	42532.312	191			

HT_AIC	117508.000	191
Entropy	54237.979	191
LMR_1V2	13829.812	191
LMR_2V3	20734.245	191
BLRT_1V2	14927.667	191
BLRT_2V3	59857.313	191

a. R Squared = .279 (Adjusted R Squared = .260)

b. R Squared = .106 (Adjusted R Squared = .082)

c. R Squared = .121 (Adjusted R Squared = .097)

d. R Squared = .098 (Adjusted R Squared = .074)

e. R Squared = .418 (Adjusted R Squared = .402)

f. R Squared = .059 (Adjusted R Squared = .034)

g. R Squared = .418 (Adjusted R Squared = .402)

h. R Squared = .253 (Adjusted R Squared = .233)

i. R Squared = .123 (Adjusted R Squared = .099)

j. R Squared = .066 (Adjusted R Squared = .041)

k. R Squared = .337 (Adjusted R Squared = .319)

l. R Squared = .020 (Adjusted R Squared = -.007)
m. R Squared = .736 (Adjusted R Squared = .729)

Table B5: Types of mixture model X Model specifications

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	33625.542[a]	5	6725.108	15.314	.000
	CAIC	15897.109[b]	5	3179.422	4.290	.001
	SACAIC	3825.354[c]	5	765.071	5.096	.000
	BIC	10684.917[d]	5	2136.983	4.008	.002
	SABIC	29034.484[e]	5	5806.897	27.334	.000
	DBIC	2021.188[f]	5	404.238	2.313	.046
	HQ	18185.437[g]	5	3637.087	27.786	.000
	HT_AIC	30822.563[h]	5	6164.513	13.227	.000
	Entropy	10441.542[i]	5	2088.308	8.869	.000
	LMR_1V2	908.875[j]	5	181.775	2.617	.026
	LMR_2V3	7104.026[k]	5	1420.805	19.389	.000
	BLRT_1V2	335.667[l]	5	67.133	.856	.512
	BLRT_2V3	45300.688[m]	5	9060.138	115.768	.000
Intercept	AIC	284284.083	1	284284.083	647.363	.000
	CAIC	1346197.547	1	1346197.547	1.816E3	.000
	SACAIC	1625456.021	1	1625456.021	1.083E4	.000
	BIC	1508752.083	1	1508752.083	2.830E3	.000
	SABIC	1196850.422	1	1196850.422	5.634E3	.000
	DBIC	1671786.750	1	1671786.750	9.567E3	.000
	HQ	1357777.687	1	1357777.687	1.037E4	.000
	HT_AIC	322752.000	1	322752.000	692.525	.000
	Entropy	166970.021	1	166970.021	709.108	.000
	LMR_1V2	1826370.187	1	1826370.187	2.629E4	.000
	LMR_2V3	1266362.755	1	1266362.755	1.728E4	.000
	BLRT_1V2	1826760.333	1	1826760.333	2.329E4	.000
	BLRT_2V3	1119046.687	1	1119046.687	1.430E4	.000
type_mixture	AIC	32027.823	2	16013.911	36.466	.000
	CAIC	15421.594	2	7710.797	10.404	.000
	SACAIC	3613.948	2	1806.974	12.035	.000
	BIC	10543.510	2	5271.755	9.887	.000

	SABIC	28475.094	2	14237.547	67.019	.000
	DBIC	1882.781	2	941.391	5.387	.005
	HQ	17449.031	2	8724.516	66.652	.000
	HT_AIC	29551.344	2	14775.672	31.704	.000
	Entropy	6464.823	2	3232.411	13.728	.000
	LMR_1V2	804.500	2	402.250	5.790	.004
	LMR_2V3	6653.323	2	3326.661	45.396	.000
	BLRT_1V2	214.542	2	107.271	1.367	.257
	BLRT_2V3	43881.031	2	21940.516	280.349	.000
model_spec	AIC	1230.187	1	1230.187	2.801	.096
	CAIC	411.255	1	411.255	.555	.457
	SACAIC	24.083	1	24.083	.160	.689
	BIC	58.521	1	58.521	.110	.741
	SABIC	338.672	1	338.672	1.594	.208
	DBIC	.187	1	.187	.001	.974
	HQ	487.688	1	487.688	3.726	.055
	HT_AIC	936.333	1	936.333	2.009	.158
	Entropy	105.021	1	105.021	.446	.505
	LMR_1V2	65.333	1	65.333	.940	.333
	LMR_2V3	441.047	1	441.047	6.019	.015
	BLRT_1V2	120.333	1	120.333	1.534	.217
	BLRT_2V3	936.333	1	936.333	11.964	.001
type_mixture * model_spec	AIC	367.531	2	183.766	.418	.659
	CAIC	64.260	2	32.130	.043	.958
	SACAIC	187.323	2	93.661	.624	.537
	BIC	82.885	2	41.443	.078	.925
	SABIC	220.719	2	110.359	.519	.596
	DBIC	138.219	2	69.109	.395	.674
	HQ	248.719	2	124.359	.950	.389
	HT_AIC	334.885	2	167.443	.359	.699
	Entropy	3871.698	2	1935.849	8.221	.000
	LMR_1V2	39.042	2	19.521	.281	.755
	LMR_2V3	9.656	2	4.828	.066	.936
	BLRT_1V2	.792	2	.396	.005	.995

	BLRT_2V3	483.323	2	241.661	3.088	.048
Error	AIC	81680.375	186	439.142		
	CAIC	137856.344	186	741.163		
	SACAIC	27926.625	186	150.143		
	BIC	99177.000	186	533.210		
	SABIC	39514.094	186	212.441		
	DBIC	32504.062	186	174.753		
	HQ	24346.875	186	130.897		
	HT_AIC	86685.438	186	466.051		
	Entropy	43796.438	186	235.465		
	LMR_1V2	12920.938	186	69.467		
	LMR_2V3	13630.219	186	73.281		
	BLRT_1V2	14592.000	186	78.452		
	BLRT_2V3	14556.625	186	78.261		
Total	AIC	399590.000	192			
	CAIC	1499951.000	192			
	SACAIC	1657208.000	192			
	BIC	1618614.000	192			
	SABIC	1265399.000	192			
	DBIC	1706312.000	192			
	HQ	1400310.000	192			
	HT_AIC	440260.000	192			
	Entropy	221208.000	192			
	LMR_1V2	1840200.000	192			
	LMR_2V3	1287097.000	192			
	BLRT_1V2	1841688.000	192			
	BLRT_2V3	1178904.000	192			
Corrected Total	AIC	115305.917	191			
	CAIC	153753.453	191			
	SACAIC	31751.979	191			
	BIC	109861.917	191			
	SABIC	68548.578	191			
	DBIC	34525.250	191			
	HQ	42532.312	191			

HT_AIC	117508.000	191
Entropy	54237.979	191
LMR_1V2	13829.812	191
LMR_2V3	20734.245	191
BLRT_1V2	14927.667	191
BLRT_2V3	59857.313	191

a. R Squared = .292 (Adjusted R Squared = .273)

b. R Squared = .103 (Adjusted R Squared = .079)

c. R Squared = .120 (Adjusted R Squared = .097)

d. R Squared = .097 (Adjusted R Squared = .073)

e. R Squared = .424 (Adjusted R Squared = .408)

f. R Squared = .059 (Adjusted R Squared = .033)

g. R Squared = .428 (Adjusted R Squared = .412)

h. R Squared = .262 (Adjusted R Squared = .242)

i. R Squared = .193 (Adjusted R Squared = .171)

j. R Squared = .066 (Adjusted R Squared = .041)

k. R Squared = .343 (Adjusted R Squared = .325)

l. R Squared = .022 (Adjusted R Squared = -.004)

m. R Squared = .757 (Adjusted R Squared = .750)

Table B6: Sample size X Class separation in LPM

Tests of Between-Subjects Effects[n]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	5747.734[a]	7	821.105	6.837	.000
	CAIC	23287.500[b]	7	3326.786	3.962	.001
	SACAIC	9575.734[c]	7	1367.962	7.493	.000
	BIC	24046.734[d]	7	3435.248	5.217	.000
	SABIC	10543.359[e]	7	1506.194	13.629	.000
	DBIC	11319.484[f]	7	1617.069	5.946	.000
	HQ	6343.234[g]	7	906.176	5.833	.000
	HT_AIC	8004.609[h]	7	1143.516	5.841	.000
	Entropy	8901.359[i]	7	1271.623	12.390	.000
	LMR_1V2	6828.937[j]	7	975.562	11.059	.000
	LMR_2V3	339.359[k]	7	48.480	.766	.618
	BLRT_1V2	5173.000[l]	7	739.000	9.315	.000

	BLRT_2V3	2111.687[m]	7	301.670	3.855	.002
Intercept	AIC	37008.141	1	37008.141	308.166	.000
	CAIC	334662.250	1	334662.250	398.609	.000
	SACAIC	534909.391	1	534909.391	2.930E3	.000
	BIC	397372.641	1	397372.641	603.425	.000
	SABIC	456807.016	1	456807.016	4.134E3	.000
	DBIC	516421.891	1	516421.891	1.899E3	.000
	HQ	509260.641	1	509260.641	3.278E3	.000
	HT_AIC	51927.016	1	51927.016	265.239	.000
	Entropy	51472.266	1	51472.266	501.524	.000
	LMR_1V2	574185.062	1	574185.062	6.509E3	.000
	LMR_2V3	418447.266	1	418447.266	6.615E3	.000
	BLRT_1V2	590592.250	1	590592.250	7.444E3	.000
	BLRT_2V3	193380.062	1	193380.062	2.471E3	.000
class_sepa	AIC	293.266	1	293.266	2.442	.124
	CAIC	4590.063	1	4590.063	5.467	.023
	SACAIC	2537.641	1	2537.641	13.900	.000
	BIC	7077.016	1	7077.016	10.747	.002
	SABIC	129.391	1	129.391	1.171	.284
	DBIC	3645.141	1	3645.141	13.403	.001
	HQ	1048.141	1	1048.141	6.747	.012
	HT_AIC	213.891	1	213.891	1.093	.300
	Entropy	102.516	1	102.516	.999	.322
	LMR_1V2	1008.063	1	1008.063	11.427	.001
	LMR_2V3	92.641	1	92.641	1.465	.231
	BLRT_1V2	756.250	1	756.250	9.532	.003
	BLRT_2V3	175.563	1	175.563	2.243	.140
N	AIC	4924.547	3	1641.516	13.669	.000
	CAIC	17341.875	3	5780.625	6.885	.000
	SACAIC	4905.547	3	1635.182	8.957	.000
	BIC	13537.547	3	4512.516	6.852	.001
	SABIC	10308.797	3	3436.266	31.094	.000
	DBIC	5462.797	3	1820.932	6.696	.001
	HQ	2851.172	3	950.391	6.118	.001

	HT_AIC	6819.047	3	2273.016	11.610	.000
	Entropy	7656.422	3	2552.141	24.867	.000
	LMR_1V2	3843.312	3	1281.104	14.523	.000
	LMR_2V3	111.172	3	37.057	.586	.627
	BLRT_1V2	2532.375	3	844.125	10.640	.000
	BLRT_2V3	1371.313	3	457.104	5.841	.002
class_sepa * N	AIC	529.922	3	176.641	1.471	.232
	CAIC	1355.563	3	451.854	.538	.658
	SACAIC	2132.547	3	710.849	3.894	.013
	BIC	3432.172	3	1144.057	1.737	.170
	SABIC	105.172	3	35.057	.317	.813
	DBIC	2211.547	3	737.182	2.711	.054
	HQ	2443.922	3	814.641	5.244	.003
	HT_AIC	971.672	3	323.891	1.654	.187
	Entropy	1142.422	3	380.807	3.710	.017
	LMR_1V2	1977.563	3	659.188	7.473	.000
	LMR_2V3	135.547	3	45.182	.714	.548
	BLRT_1V2	1884.375	3	628.125	7.917	.000
	BLRT_2V3	564.813	3	188.271	2.406	.077
Error	AIC	6725.125	56	120.092		
	CAIC	47016.250	56	839.576		
	SACAIC	10223.875	56	182.569		
	BIC	36877.625	56	658.529		
	SABIC	6188.625	56	110.511		
	DBIC	15229.625	56	271.958		
	HQ	8699.125	56	155.342		
	HT_AIC	10963.375	56	195.775		
	Entropy	5747.375	56	102.632		
	LMR_1V2	4940.000	56	88.214		
	LMR_2V3	3542.375	56	63.257		
	BLRT_1V2	4442.750	56	79.335		
	BLRT_2V3	4382.250	56	78.254		
Total	AIC	49481.000	64			
	CAIC	404966.000	64			

	SACAIC	554709.000	64
	BIC	458297.000	64
	SABIC	473539.000	64
	DBIC	542971.000	64
	HQ	524303.000	64
	HT_AIC	70895.000	64
	Entropy	66121.000	64
	LMR_1V2	585954.000	64
	LMR_2V3	422329.000	64
	BLRT_1V2	600208.000	64
	BLRT_2V3	199874.000	64
Corrected Total	AIC	12472.859	63
	CAIC	70303.750	63
	SACAIC	19799.609	63
	BIC	60924.359	63
	SABIC	16731.984	63
	DBIC	26549.109	63
	HQ	15042.359	63
	HT_AIC	18967.984	63
	Entropy	14648.734	63
	LMR_1V2	11768.937	63
	LMR_2V3	3881.734	63
	BLRT_1V2	9615.750	63
	BLRT_2V3	6493.937	63

a. R Squared = .461 (Adjusted R Squared = .393)

b. R Squared = .331 (Adjusted R Squared = .248)

c. R Squared = .484 (Adjusted R Squared = .419)

d. R Squared = .395 (Adjusted R Squared = .319)

e. R Squared = .630 (Adjusted R Squared = .584)

f. R Squared = .426 (Adjusted R Squared = .355)

g. R Squared = .422 (Adjusted R Squared = .349)

h. R Squared = .422 (Adjusted R Squared = .350)

i. R Squared = .608 (Adjusted R Squared = .559)

j. R Squared = .580 (Adjusted R Squared = .528)

k. R Squared = .087 (Adjusted R Squared = -.027)

l. R Squared = .538 (Adjusted R Squared = .480)

m. R Squared = .325 (Adjusted R Squared = .241)
n. type_mixture = LPM

Table B7: Sample size X Class separation in UGMM

Tests of Between-Subjects Effects[n]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	5157.609[a]	7	736.801	.712	.662
	CAIC	41643.000[b]	7	5949.000	32.576	.000
	SACAIC	275.438[c]	7	39.348	3.400	.004
	BIC	24628.938[d]	7	3518.420	26.501	.000
	SABIC	373.750[e]	7	53.393	.483	.843
	DBIC	1001.734[f]	7	143.105	5.700	.000
	HQ	1479.938[g]	7	211.420	1.849	.096
	HT_AIC	5483.688[h]	7	783.384	.776	.610
	Entropy	632.609[i]	7	90.373	.206	.983
	LMR_1V2	766.438[j]	7	109.491	17.032	.000
	LMR_2V3	1282.188[k]	7	183.170	2.011	.070
	BLRT_1V2	1508.937[l]	7	215.562	8.939	.000
	BLRT_2V3	827.484[m]	7	118.212	1.883	.090
Intercept	AIC	196359.766	1	196359.766	189.760	.000
	CAIC	459006.250	1	459006.250	2.513E3	.000
	SACAIC	609570.562	1	609570.562	5.268E4	.000
	BIC	523814.062	1	523814.062	3.945E3	.000

	SABIC	522006.250	1	522006.250	4.718E3	.000
	DBIC	607425.391	1	607425.391	2.420E4	.000
	HQ	545751.562	1	545751.562	4.773E3	.000
	HT_AIC	214600.562	1	214600.562	212.672	.000
	Entropy	33902.016	1	33902.016	77.137	.000
	LMR_1V2	618975.562	1	618975.562	9.629E4	.000
	LMR_2V3	502326.562	1	502326.562	5.514E3	.000
	BLRT_1V2	615832.562	1	615832.562	2.554E4	.000
	BLRT_2V3	489125.391	1	489125.391	7.790E3	.000
class_sepa	AIC	301.891	1	301.891	.292	.591
	CAIC	9555.063	1	9555.063	52.322	.000
	SACAIC	33.063	1	33.063	2.857	.097
	BIC	4830.250	1	4830.250	36.381	.000
	SABIC	25.000	1	25.000	.226	.636
	DBIC	178.891	1	178.891	7.126	.010
	HQ	42.250	1	42.250	.370	.546
	HT_AIC	333.063	1	333.063	.330	.568
	Entropy	62.016	1	62.016	.141	.709
	LMR_1V2	126.563	1	126.563	19.688	.000
	LMR_2V3	14.062	1	14.062	.154	.696
	BLRT_1V2	217.563	1	217.563	9.021	.004
	BLRT_2V3	.141	1	.141	.002	.962
N	AIC	4656.422	3	1552.141	1.500	.225
	CAIC	21052.750	3	7017.583	38.427	.000
	SACAIC	159.562	3	53.188	4.596	.006
	BIC	10923.062	3	3641.021	27.424	.000
	SABIC	333.250	3	111.083	1.004	.398
	DBIC	399.672	3	133.224	5.307	.003
	HQ	1276.062	3	425.354	3.720	.016
	HT_AIC	4960.562	3	1653.521	1.639	.191
	Entropy	456.922	3	152.307	.347	.792
	LMR_1V2	371.812	3	123.938	19.279	.000
	LMR_2V3	1218.062	3	406.021	4.457	.007
	BLRT_1V2	667.688	3	222.562	9.229	.000

	BLRT_2V3	823.922	3	274.641	4.374	.008
class_sepa * N	AIC	199.297	3	66.432	.064	.979
	CAIC	11035.188	3	3678.396	20.142	.000
	SACAIC	82.812	3	27.604	2.386	.079
	BIC	8875.625	3	2958.542	22.284	.000
	SABIC	15.500	3	5.167	.047	.986
	DBIC	423.172	3	141.057	5.619	.002
	HQ	161.625	3	53.875	.471	.704
	HT_AIC	190.063	3	63.354	.063	.979
	Entropy	113.672	3	37.891	.086	.967
	LMR_1V2	268.063	3	89.354	13.900	.000
	LMR_2V3	50.062	3	16.687	.183	.907
	BLRT_1V2	623.688	3	207.896	8.621	.000
	BLRT_2V3	3.422	3	1.141	.018	.997
Error	AIC	57947.625	56	1034.779		
	CAIC	10226.750	56	182.621		
	SACAIC	648.000	56	11.571		
	BIC	7435.000	56	132.768		
	SABIC	6196.000	56	110.643		
	DBIC	1405.875	56	25.105		
	HQ	6402.500	56	114.330		
	HT_AIC	56507.750	56	1009.067		
	Entropy	24612.375	56	439.507		
	LMR_1V2	360.000	56	6.429		
	LMR_2V3	5101.250	56	91.094		
	BLRT_1V2	1350.500	56	24.116		
	BLRT_2V3	3516.125	56	62.788		
Total	AIC	259465.000	64			
	CAIC	510876.000	64			
	SACAIC	610494.000	64			
	BIC	555878.000	64			
	SABIC	528576.000	64			
	DBIC	609833.000	64			
	HQ	553634.000	64			

	HT_AIC	276592.000	64
	Entropy	59147.000	64
	LMR_1V2	620102.000	64
	LMR_2V3	508710.000	64
	BLRT_1V2	618692.000	64
	BLRT_2V3	493469.000	64
Corrected Total	AIC	63105.234	63
	CAIC	51869.750	63
	SACAIC	923.438	63
	BIC	32063.938	63
	SABIC	6569.750	63
	DBIC	2407.609	63
	HQ	7882.438	63
	HT_AIC	61991.438	63
	Entropy	25244.984	63
	LMR_1V2	1126.438	63
	LMR_2V3	6383.438	63
	BLRT_1V2	2859.437	63
	BLRT_2V3	4343.609	63

a. R Squared = .082 (Adjusted R Squared = -.033)

b. R Squared = .803 (Adjusted R Squared = .778)

c. R Squared = .298 (Adjusted R Squared = .211)

d. R Squared = .768 (Adjusted R Squared = .739)

e. R Squared = .057 (Adjusted R Squared = -.061)

f. R Squared = .416 (Adjusted R Squared = .343)

g. R Squared = .188 (Adjusted R Squared = .086)

h. R Squared = .088 (Adjusted R Squared = -.025)

i. R Squared = .025 (Adjusted R Squared = -.097)

j. R Squared = .680 (Adjusted R Squared = .640)

k. R Squared = .201 (Adjusted R Squared = .101)

l. R Squared = .528 (Adjusted R Squared = .469)

m. R Squared = .191 (Adjusted R Squared = .089)
n. type_mixture = UGMM

Table B8: Sample size X Class separation in Linear GMM

Tests of Between-Subjects Effects[n]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	2390.500[a]	7	341.500	3.602	.003
	CAIC	12123.984[b]	7	1731.998	24.041	.000
	SACAIC	5925.109[c]	7	846.444	31.815	.000
	BIC	4156.984[d]	7	593.855	15.303	.000
	SABIC	15210.750[e]	7	2172.964	77.954	.000
	DBIC	2765.500[f]	7	395.071	24.041	.000
	HQ	582.359[g]	7	83.194	2.956	.010
	HT_AIC	2238.859[h]	7	319.837	3.764	.002
	Entropy	2552.438[i]	7	364.634	3.833	.002
	LMR_1V2	84.187[j]	7	12.027	14.721	.000
	LMR_2V3	133.000[k]	7	19.000	.289	.956
	BLRT_1V2	1002.437[l]	7	143.205	6.491	.000
	BLRT_2V3	546.109[m]	7	78.016	.951	.475
Intercept	AIC	82944.000	1	82944.000	874.821	.000
	CAIC	567950.641	1	567950.641	7.884E3	.000
	SACAIC	484590.016	1	484590.016	1.821E4	.000
	BIC	598108.891	1	598108.891	1.541E4	.000
	SABIC	246512.250	1	246512.250	8.843E3	.000
	DBIC	549822.250	1	549822.250	3.346E4	.000
	HQ	320214.516	1	320214.516	1.138E4	.000
	HT_AIC	85775.766	1	85775.766	1.009E3	.000
	Entropy	88060.562	1	88060.562	925.735	.000
	LMR_1V2	634014.062	1	634014.062	7.761E5	.000
	LMR_2V3	352242.250	1	352242.250	5.356E3	.000
	BLRT_1V2	620550.062	1	620550.062	2.813E4	.000
	BLRT_2V3	480422.266	1	480422.266	5.858E3	.000
class_sepa	AIC	600.250	1	600.250	6.331	.015
	CAIC	2058.891	1	2058.891	28.579	.000
	SACAIC	19.141	1	19.141	.719	.400
	BIC	695.641	1	695.641	17.926	.000

	SABIC	126.562	1	126.562	4.540	.038
	DBIC	.563	1	.563	.034	.854
	HQ	185.641	1	185.641	6.596	.013
	HT_AIC	606.391	1	606.391	7.136	.010
	Entropy	1173.062	1	1173.062	12.332	.001
	LMR_1V2	14.063	1	14.063	17.213	.000
	LMR_2V3	5.062	1	5.062	.077	.782
	BLRT_1V2	150.063	1	150.063	6.802	.012
	BLRT_2V3	135.141	1	135.141	1.648	.205
N	AIC	1648.375	3	549.458	5.795	.002
	CAIC	5155.172	3	1718.391	23.852	.000
	SACAIC	5887.672	3	1962.557	73.767	.000
	BIC	1748.672	3	582.891	15.021	.000
	SABIC	15050.875	3	5016.958	179.981	.000
	DBIC	2744.375	3	914.792	55.668	.000
	HQ	367.797	3	122.599	4.356	.008
	HT_AIC	1503.672	3	501.224	5.899	.001
	Entropy	1345.562	3	448.521	4.715	.005
	LMR_1V2	35.062	3	11.688	14.306	.000
	LMR_2V3	104.375	3	34.792	.529	.664
	BLRT_1V2	426.187	3	142.062	6.439	.001
	BLRT_2V3	345.047	3	115.016	1.402	.252
class_sepa * N	AIC	141.875	3	47.292	.499	.685
	CAIC	4909.922	3	1636.641	22.718	.000
	SACAIC	18.297	3	6.099	.229	.876
	BIC	1712.672	3	570.891	14.711	.000
	SABIC	33.313	3	11.104	.398	.755
	DBIC	20.563	3	6.854	.417	.741
	HQ	28.922	3	9.641	.343	.795
	HT_AIC	128.797	3	42.932	.505	.680
	Entropy	33.813	3	11.271	.118	.949
	LMR_1V2	35.063	3	11.688	14.306	.000
	LMR_2V3	23.563	3	7.854	.119	.948
	BLRT_1V2	426.188	3	142.063	6.439	.001

	BLRT_2V3	65.922	3	21.974	.268	.848
Error	AIC	5309.500	56	94.812		
	CAIC	4034.375	56	72.042		
	SACAIC	1489.875	56	26.605		
	BIC	2173.125	56	38.806		
	SABIC	1561.000	56	27.875		
	DBIC	920.250	56	16.433		
	HQ	1576.125	56	28.145		
	HT_AIC	4758.375	56	84.971		
	Entropy	5327.000	56	95.125		
	LMR_1V2	45.750	56	.817		
	LMR_2V3	3682.750	56	65.763		
	BLRT_1V2	1235.500	56	22.063		
	BLRT_2V3	4592.625	56	82.011		
Total	AIC	90644.000	64			
	CAIC	584109.000	64			
	SACAIC	492005.000	64			
	BIC	604439.000	64			
	SABIC	263284.000	64			
	DBIC	553508.000	64			
	HQ	322373.000	64			
	HT_AIC	92773.000	64			
	Entropy	95940.000	64			
	LMR_1V2	634144.000	64			
	LMR_2V3	356058.000	64			
	BLRT_1V2	622788.000	64			
	BLRT_2V3	485561.000	64			
Corrected Total	AIC	7700.000	63			
	CAIC	16158.359	63			
	SACAIC	7414.984	63			
	BIC	6330.109	63			
	SABIC	16771.750	63			
	DBIC	3685.750	63			
	HQ	2158.484	63			

	HT_AIC	6997.234	63
	Entropy	7879.438	63
	LMR_1V2	129.937	63
	LMR_2V3	3815.750	63
	BLRT_1V2	2237.937	63
	BLRT_2V3	5138.734	63

a. R Squared = .310 (Adjusted R Squared = .224)
b. R Squared = .750 (Adjusted R Squared = .719)
c. R Squared = .799 (Adjusted R Squared = .774)
d. R Squared = .657 (Adjusted R Squared = .614)
e. R Squared = .907 (Adjusted R Squared = .895)
f. R Squared = .750 (Adjusted R Squared = .719)
g. R Squared = .270 (Adjusted R Squared = .179)
h. R Squared = .320 (Adjusted R Squared = .235)
i. R Squared = .324 (Adjusted R Squared = .239)
j. R Squared = .648 (Adjusted R Squared = .604)
k. R Squared = .035 (Adjusted R Squared = -.086)
l. R Squared = .448 (Adjusted R Squared = .379)
m. R Squared = .106 (Adjusted R Squared = -.005)
n. type_mixture = Linear GMM

Table B9: Sample size X Number of measures in LPM

Tests of Between-Subjects Effects[n]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	7344.234[a]	7	1049.176	11.456	.000
	CAIC	51512.250[b]	7	7358.893	21.930	.000
	SACAIC	9003.484[c]	7	1286.212	6.672	.000
	BIC	39372.734[d]	7	5624.676	14.615	.000
	SABIC	14302.859[e]	7	2043.266	47.105	.000
	DBIC	13930.234[f]	7	1990.033	8.831	.000
	HQ	6917.484[g]	7	988.212	6.811	.000
	HT_AIC	8352.859[h]	7	1193.266	6.295	.000
	Entropy	9135.109[i]	7	1305.016	13.255	.000
	LMR_1V2	6720.937[j]	7	960.134	10.651	.000
	LMR_2V3	2407.609[k]	7	343.944	13.066	.000

	BLRT_1V2	4594.000[l]	7	656.286	7.319	.000
	BLRT_2V3	3053.437[m]	7	436.205	7.100	.000
Intercept	AIC	37008.141	1	37008.141	404.096	.000
	CAIC	334662.250	1	334662.250	997.317	.000
	SACAIC	534909.391	1	534909.391	2.775E3	.000
	BIC	397372.641	1	397372.641	1.033E3	.000
	SABIC	456807.016	1	456807.016	1.053E4	.000
	DBIC	516421.891	1	516421.891	2.292E3	.000
	HQ	509260.641	1	509260.641	3.510E3	.000
	HT_AIC	51927.016	1	51927.016	273.941	.000
	Entropy	51472.266	1	51472.266	522.786	.000
	LMR_1V2	574185.063	1	574185.063	6.370E3	.000
	LMR_2V3	418447.266	1	418447.266	1.590E4	.000
	BLRT_1V2	590592.250	1	590592.250	6.586E3	.000
	BLRT_2V3	193380.063	1	193380.063	3.148E3	.000
N	AIC	4924.547	3	1641.516	17.924	.000
	CAIC	17341.875	3	5780.625	17.227	.000
	SACAIC	4905.547	3	1635.182	8.482	.000
	BIC	13537.547	3	4512.516	11.725	.000
	SABIC	10308.797	3	3436.266	79.218	.000
	DBIC	5462.797	3	1820.932	8.081	.000
	HQ	2851.172	3	950.391	6.550	.001
	HT_AIC	6819.047	3	2273.016	11.991	.000
	Entropy	7656.422	3	2552.141	25.921	.000
	LMR_1V2	3843.313	3	1281.104	14.212	.000
	LMR_2V3	111.172	3	37.057	1.408	.250
	BLRT_1V2	2532.375	3	844.125	9.413	.000
	BLRT_2V3	1371.313	3	457.104	7.440	.000
measure	AIC	1947.016	1	1947.016	21.260	.000
	CAIC	26487.563	1	26487.563	78.935	.000
	SACAIC	2150.641	1	2150.641	11.155	.001
	BIC	18940.641	1	18940.641	49.216	.000
	SABIC	3122.016	1	3122.016	71.974	.000
	DBIC	4882.516	1	4882.516	21.668	.000

	HQ	43.891	1	43.891	.303	.584
	HT_AIC	922.641	1	922.641	4.867	.031
	Entropy	819.391	1	819.391	8.322	.006
	LMR_1V2	961.000	1	961.000	10.661	.002
	LMR_2V3	2173.891	1	2173.891	82.583	.000
	BLRT_1V2	600.250	1	600.250	6.694	.012
	BLRT_2V3	1521.000	1	1521.000	24.757	.000
N * measure	AIC	472.672	3	157.557	1.720	.173
	CAIC	7682.813	3	2560.938	7.632	.000
	SACAIC	1947.297	3	649.099	3.367	.025
	BIC	6894.547	3	2298.182	5.972	.001
	SABIC	872.047	3	290.682	6.701	.001
	DBIC	3584.922	3	1194.974	5.303	.003
	HQ	4022.422	3	1340.807	9.241	.000
	HT_AIC	611.172	3	203.724	1.075	.367
	Entropy	659.297	3	219.766	2.232	.094
	LMR_1V2	1916.625	3	638.875	7.087	.000
	LMR_2V3	122.547	3	40.849	1.552	.211
	BLRT_1V2	1461.375	3	487.125	5.432	.002
	BLRT_2V3	161.125	3	53.708	.874	.460
Error	AIC	5128.625	56	91.583		
	CAIC	18791.500	56	335.562		
	SACAIC	10796.125	56	192.788		
	BIC	21551.625	56	384.850		
	SABIC	2429.125	56	43.377		
	DBIC	12618.875	56	225.337		
	HQ	8124.875	56	145.087		
	HT_AIC	10615.125	56	189.556		
	Entropy	5513.625	56	98.458		
	LMR_1V2	5048.000	56	90.143		
	LMR_2V3	1474.125	56	26.324		
	BLRT_1V2	5021.750	56	89.674		
	BLRT_2V3	3440.500	56	61.438		
Total	AIC	49481.000	64			

	CAIC	404966.000	64
	SACAIC	554709.000	64
	BIC	458297.000	64
	SABIC	473539.000	64
	DBIC	542971.000	64
	HQ	524303.000	64
	HT_AIC	70895.000	64
	Entropy	66121.000	64
	LMR_1V2	585954.000	64
	LMR_2V3	422329.000	64
	BLRT_1V2	600208.000	64
	BLRT_2V3	199874.000	64
Corrected Total	AIC	12472.859	63
	CAIC	70303.750	63
	SACAIC	19799.609	63
	BIC	60924.359	63
	SABIC	16731.984	63
	DBIC	26549.109	63
	HQ	15042.359	63
	HT_AIC	18967.984	63
	Entropy	14648.734	63
	LMR_1V2	11768.937	63
	LMR_2V3	3881.734	63
	BLRT_1V2	9615.750	63
	BLRT_2V3	6493.937	63

a. R Squared = .589 (Adjusted R Squared = .537)

b. R Squared = .733 (Adjusted R Squared = .699)

c. R Squared = .455 (Adjusted R Squared = .387)

d. R Squared = .646 (Adjusted R Squared = .602)

e. R Squared = .855 (Adjusted R Squared = .837)

f. R Squared = .525 (Adjusted R Squared = .465)

g. R Squared = .460 (Adjusted R Squared = .392)

h. R Squared = .440 (Adjusted R Squared = .370)

i. R Squared = .624 (Adjusted R Squared = .577)

j. R Squared = .571 (Adjusted R Squared = .517)

k. R Squared = .620 (Adjusted R Squared = .573)

l. R Squared = .478 (Adjusted R Squared = .412)

m. R Squared = .470 (Adjusted R Squared = .404)
n. type_mixture =LPM

Table B10: Sample size X Number of measures in UGMM

Tests of Between-Subjects Effects[n]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	55201.547[a]	3	18400.516	139.686	.000
	CAIC	13643.375[b]	3	4547.792	7.138	.000
	SACAIC	158.063[c]	3	52.688	4.130	.010
	BIC	7493.063[d]	3	2497.688	6.099	.001
	SABIC	3768.125[e]	3	1256.042	26.900	.000
	DBIC	415.297[f]	3	138.432	4.169	.010
	HQ	1505.563[g]	3	501.854	4.722	.005
	HT_AIC	53501.188[h]	3	17833.729	126.030	.000
	Entropy	15104.547[i]	3	5034.849	29.791	.000
	LMR_1V2	162.813[j]	3	54.271	3.379	.024
	LMR_2V3	3084.313[k]	3	1028.104	18.698	.000
	BLRT_1V2	407.812[l]	3	135.937	3.327	.025
	BLRT_2V3	519.297[m]	3	173.099	2.716	.053
Intercept	AIC	196359.766	1	196359.766	1.491E3	.000
	CAIC	459006.250	1	459006.250	720.455	.000
	SACAIC	609570.562	1	609570.562	4.779E4	.000
	BIC	523814.062	1	523814.062	1.279E3	.000

	SABIC	522006.250	1	522006.250	1.118E4	.000
	DBIC	607425.391	1	607425.391	1.829E4	.000
	HQ	545751.562	1	545751.562	5.135E3	.000
	HT_AIC	214600.562	1	214600.562	1.517E3	.000
	Entropy	33902.016	1	33902.016	200.595	.000
	LMR_1V2	618975.562	1	618975.562	3.854E4	.000
	LMR_2V3	502326.562	1	502326.562	9.136E3	.000
	BLRT_1V2	615832.562	1	615832.562	1.507E4	.000
	BLRT_2V3	489125.391	1	489125.391	7.674E3	.000
class_sepa	AIC	301.891	1	301.891	2.292	.135
	CAIC	9555.062	1	9555.062	14.998	.000
	SACAIC	33.062	1	33.062	2.592	.113
	BIC	4830.250	1	4830.250	11.795	.001
	SABIC	25.000	1	25.000	.535	.467
	DBIC	178.891	1	178.891	5.387	.024
	HQ	42.250	1	42.250	.398	.531
	HT_AIC	333.062	1	333.062	2.354	.130
	Entropy	62.016	1	62.016	.367	.547
	LMR_1V2	126.562	1	126.562	7.880	.007
	LMR_2V3	14.062	1	14.062	.256	.615
	BLRT_1V2	217.562	1	217.562	5.325	.024
	BLRT_2V3	.141	1	.141	.002	.963
measure	AIC	54463.891	1	54463.891	413.457	.000
	CAIC	3080.250	1	3080.250	4.835	.032
	SACAIC	4.000	1	4.000	.314	.578
	BIC	1540.562	1	1540.562	3.762	.057
	SABIC	3510.562	1	3510.562	75.183	.000
	DBIC	43.891	1	43.891	1.322	.255
	HQ	1260.250	1	1260.250	11.858	.001
	HT_AIC	52555.562	1	52555.562	371.406	.000
	Entropy	14731.891	1	14731.891	87.167	.000
	LMR_1V2	16.000	1	16.000	.996	.322
	LMR_2V3	2970.250	1	2970.250	54.019	.000
	BLRT_1V2	100.000	1	100.000	2.447	.123

	BLRT_2V3	478.516	1	478.516	7.507	.008
class_sepa * measure	AIC	435.766	1	435.766	3.308	.074
	CAIC	1008.062	1	1008.062	1.582	.213
	SACAIC	121.000	1	121.000	9.486	.003
	BIC	1122.250	1	1122.250	2.740	.103
	SABIC	232.562	1	232.562	4.981	.029
	DBIC	192.516	1	192.516	5.798	.019
	HQ	203.062	1	203.062	1.911	.172
	HT_AIC	612.562	1	612.562	4.329	.042
	Entropy	310.641	1	310.641	1.838	.180
	LMR_1V2	20.250	1	20.250	1.261	.266
	LMR_2V3	100.000	1	100.000	1.819	.183
	BLRT_1V2	90.250	1	90.250	2.209	.142
	BLRT_2V3	40.641	1	40.641	.638	.428
Error	AIC	7903.688	60	131.728		
	CAIC	38226.375	60	637.106		
	SACAIC	765.375	60	12.756		
	BIC	24570.875	60	409.515		
	SABIC	2801.625	60	46.694		
	DBIC	1992.312	60	33.205		
	HQ	6376.875	60	106.281		
	HT_AIC	8490.250	60	141.504		
	Entropy	10140.438	60	169.007		
	LMR_1V2	963.625	60	16.060		
	LMR_2V3	3299.125	60	54.985		
	BLRT_1V2	2451.625	60	40.860		
	BLRT_2V3	3824.312	60	63.739		
Total	AIC	259465.000	64			
	CAIC	510876.000	64			
	SACAIC	610494.000	64			
	BIC	555878.000	64			
	SABIC	528576.000	64			
	DBIC	609833.000	64			
	HQ	553634.000	64			

	HT_AIC	276592.000	64
	Entropy	59147.000	64
	LMR_1V2	620102.000	64
	LMR_2V3	508710.000	64
	BLRT_1V2	618692.000	64
	BLRT_2V3	493469.000	64
Corrected Total	AIC	63105.234	63
	CAIC	51869.750	63
	SACAIC	923.438	63
	BIC	32063.938	63
	SABIC	6569.750	63
	DBIC	2407.609	63
	HQ	7882.438	63
	HT_AIC	61991.438	63
	Entropy	25244.984	63
	LMR_1V2	1126.438	63
	LMR_2V3	6383.438	63
	BLRT_1V2	2859.437	63
	BLRT_2V3	4343.609	63

a. R Squared = .875 (Adjusted R Squared = .868)

b. R Squared = .263 (Adjusted R Squared = .226)

c. R Squared = .171 (Adjusted R Squared = .130)

d. R Squared = .234 (Adjusted R Squared = .195)

e. R Squared = .574 (Adjusted R Squared = .552)

f. R Squared = .172 (Adjusted R Squared = .131)

g. R Squared = .191 (Adjusted R Squared = .151)

h. R Squared = .863 (Adjusted R Squared = .856)

i. R Squared = .598 (Adjusted R Squared = .578)

j. R Squared = .145 (Adjusted R Squared = .102)

k. R Squared = .483 (Adjusted R Squared = .457)

l. R Squared = .143 (Adjusted R Squared = .100)

m. R Squared = .120 (Adjusted R Squared = .076)

n. type_mixture = UGMM

Table B11: Sample size X Number of measures in Linear GMM

Tests of Between-Subjects Effects[n]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	5108.250[a]	7	729.750	15.768	.000
	CAIC	6487.484[b]	7	926.783	5.367	.000
	SACAIC	6075.359[c]	7	867.908	36.281	.000
	BIC	2262.234[d]	7	323.176	4.449	.001
	SABIC	15267.000[e]	7	2181.000	81.167	.000
	DBIC	2894.000[f]	7	413.429	29.242	.000
	HQ	562.359[g]	7	80.337	2.819	.014
	HT_AIC	4434.609[h]	7	633.516	13.844	.000
	Entropy	4014.938[i]	7	573.563	8.311	.000
	LMR_1V2	35.437[j]	7	5.062	3.000	.010
	LMR_2V3	2357.500[k]	7	336.786	12.933	.000
	BLRT_1V2	787.437[l]	7	112.491	4.343	.001
	BLRT_2V3	2185.359[m]	7	312.194	5.920	.000
Intercept	AIC	82944.000	1	82944.000	1.792E3	.000
	CAIC	567950.641	1	567950.641	3.289E3	.000
	SACAIC	484590.016	1	484590.016	2.026E4	.000
	BIC	598108.891	1	598108.891	8.234E3	.000
	SABIC	246512.250	1	246512.250	9.174E3	.000
	DBIC	549822.250	1	549822.250	3.889E4	.000
	HQ	320214.516	1	320214.516	1.123E4	.000
	HT_AIC	85775.766	1	85775.766	1.874E3	.000
	Entropy	88060.563	1	88060.563	1.276E3	.000
	LMR_1V2	634014.063	1	634014.063	3.757E5	.000
	LMR_2V3	352242.250	1	352242.250	1.353E4	.000
	BLRT_1V2	620550.063	1	620550.063	2.396E4	.000
	BLRT_2V3	480422.266	1	480422.266	9.109E3	.000
N	AIC	1648.375	3	549.458	11.872	.000
	CAIC	5155.172	3	1718.391	9.950	.000
	SACAIC	5887.672	3	1962.557	82.040	.000
	BIC	1748.672	3	582.891	8.024	.000

	SABIC	15050.875	3	5016.958	186.709	.000
	DBIC	2744.375	3	914.792	64.703	.000
	HQ	367.797	3	122.599	4.301	.008
	HT_AIC	1503.672	3	501.224	10.953	.000
	Entropy	1345.562	3	448.521	6.499	.001
	LMR_1V2	35.063	3	11.688	6.926	.000
	LMR_2V3	104.375	3	34.792	1.336	.272
	BLRT_1V2	426.188	3	142.063	5.485	.002
	BLRT_2V3	345.047	3	115.016	2.181	.100
measure	AIC	3335.062	1	3335.062	72.061	.000
	CAIC	435.766	1	435.766	2.523	.118
	SACAIC	5.641	1	5.641	.236	.629
	BIC	172.266	1	172.266	2.371	.129
	SABIC	.250	1	.250	.009	.924
	DBIC	14.063	1	14.063	.995	.323
	HQ	6.891	1	6.891	.242	.625
	HT_AIC	2795.766	1	2795.766	61.095	.000
	Entropy	2652.250	1	2652.250	38.433	.000
	LMR_1V2	.063	1	.063	.037	.848
	LMR_2V3	2093.063	1	2093.063	80.378	.000
	BLRT_1V2	95.063	1	95.063	3.670	.061
	BLRT_2V3	1691.266	1	1691.266	32.069	.000
N * measure	AIC	124.813	3	41.604	.899	.448
	CAIC	896.547	3	298.849	1.731	.171
	SACAIC	182.047	3	60.682	2.537	.066
	BIC	341.297	3	113.766	1.566	.208
	SABIC	215.875	3	71.958	2.678	.056
	DBIC	135.562	3	45.188	3.196	.030
	HQ	187.672	3	62.557	2.195	.099
	HT_AIC	135.172	3	45.057	.985	.407
	Entropy	17.125	3	5.708	.083	.969
	LMR_1V2	.312	3	.104	.062	.980
	LMR_2V3	160.062	3	53.354	2.049	.117
	BLRT_1V2	266.187	3	88.729	3.426	.023

	BLRT_2V3	149.047	3	49.682	.942	.427
Error	AIC	2591.750	56	46.281		
	CAIC	9670.875	56	172.694		
	SACAIC	1339.625	56	23.922		
	BIC	4067.875	56	72.641		
	SABIC	1504.750	56	26.871		
	DBIC	791.750	56	14.138		
	HQ	1596.125	56	28.502		
	HT_AIC	2562.625	56	45.761		
	Entropy	3864.500	56	69.009		
	LMR_1V2	94.500	56	1.688		
	LMR_2V3	1458.250	56	26.040		
	BLRT_1V2	1450.500	56	25.902		
	BLRT_2V3	2953.375	56	52.739		
Total	AIC	90644.000	64			
	CAIC	584109.000	64			
	SACAIC	492005.000	64			
	BIC	604439.000	64			
	SABIC	263284.000	64			
	DBIC	553508.000	64			
	HQ	322373.000	64			
	HT_AIC	92773.000	64			
	Entropy	95940.000	64			
	LMR_1V2	634144.000	64			
	LMR_2V3	356058.000	64			
	BLRT_1V2	622788.000	64			
	BLRT_2V3	485561.000	64			
Corrected Total	AIC	7700.000	63			
	CAIC	16158.359	63			
	SACAIC	7414.984	63			
	BIC	6330.109	63			
	SABIC	16771.750	63			
	DBIC	3685.750	63			
	HQ	2158.484	63			

HT_AIC	6997.234	63
Entropy	7879.438	63
LMR_1V2	129.937	63
LMR_2V3	3815.750	63
BLRT_1V2	2237.937	63
BLRT_2V3	5138.734	63

a. R Squared = .663 (Adjusted R Squared = .621)
b. R Squared = .401 (Adjusted R Squared = .327)
c. R Squared = .819 (Adjusted R Squared = .797)
d. R Squared = .357 (Adjusted R Squared = .277)
e. R Squared = .910 (Adjusted R Squared = .899)
f. R Squared = .785 (Adjusted R Squared = .758)
g. R Squared = .261 (Adjusted R Squared = .168)
h. R Squared = .634 (Adjusted R Squared = .588)
i. R Squared = .510 (Adjusted R Squared = .448)
j. R Squared = .273 (Adjusted R Squared = .182)
k. R Squared = .618 (Adjusted R Squared = .570)
l. R Squared = .352 (Adjusted R Squared = .271)
m. R Squared = .425 (Adjusted R Squared = .353)
n. type_mixture = Linear GMM

Table B12: Class separation X Number of measures in LPM

Tests of Between-Subjects Effects[n]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	2255.297[a]	3	751.766	4.415	.007
	CAIC	31518.625[b]	3	10506.208	16.253	.000
	SACAIC	7101.547[c]	3	2367.182	11.185	.000
	BIC	28893.297[d]	3	9631.099	18.041	.000
	SABIC	3882.672[e]	3	1294.224	6.043	.001
	DBIC	11538.922[f]	3	3846.307	15.375	.000
	HQ	2824.672[g]	3	941.557	4.624	.006
	HT_AIC	1244.172[h]	3	414.724	1.404	.250
	Entropy	923.797[i]	3	307.932	1.346	.268
	LMR_1V2	2410.062[j]	3	803.354	5.150	.003
	LMR_2V3	2332.547[k]	3	777.516	30.113	.000

	BLRT_1V2	1776.750[l]	3	592.250	4.533	.006
	BLRT_2V3	1721.562[m]	3	573.854	7.215	.000
Intercept	AIC	37008.141	1	37008.141	217.321	.000
	CAIC	334662.250	1	334662.250	517.717	.000
	SACAIC	534909.391	1	534909.391	2.528E3	.000
	BIC	397372.641	1	397372.641	744.351	.000
	SABIC	456807.016	1	456807.016	2.133E3	.000
	DBIC	516421.891	1	516421.891	2.064E3	.000
	HQ	509260.641	1	509260.641	2.501E3	.000
	HT_AIC	51927.016	1	51927.016	175.787	.000
	Entropy	51472.266	1	51472.266	225.016	.000
	LMR_1V2	574185.062	1	574185.062	3.681E3	.000
	LMR_2V3	418447.266	1	418447.266	1.621E4	.000
	BLRT_1V2	590592.250	1	590592.250	4.520E3	.000
	BLRT_2V3	193380.062	1	193380.062	2.431E3	.000
class_sepa	AIC	293.266	1	293.266	1.722	.194
	CAIC	4590.062	1	4590.062	7.101	.010
	SACAIC	2537.641	1	2537.641	11.991	.001
	BIC	7077.016	1	7077.016	13.257	.001
	SABIC	129.391	1	129.391	.604	.440
	DBIC	3645.141	1	3645.141	14.571	.000
	HQ	1048.141	1	1048.141	5.147	.027
	HT_AIC	213.891	1	213.891	.724	.398
	Entropy	102.516	1	102.516	.448	.506
	LMR_1V2	1008.062	1	1008.062	6.463	.014
	LMR_2V3	92.641	1	92.641	3.588	.063
	BLRT_1V2	756.250	1	756.250	5.788	.019
	BLRT_2V3	175.562	1	175.562	2.207	.143
measure	AIC	1947.016	1	1947.016	11.433	.001
	CAIC	26487.562	1	26487.562	40.976	.000
	SACAIC	2150.641	1	2150.641	10.162	.002
	BIC	18940.641	1	18940.641	35.479	.000
	SABIC	3122.016	1	3122.016	14.578	.000
	DBIC	4882.516	1	4882.516	19.517	.000

	HQ	43.891	1	43.891	.216	.644
	HT_AIC	922.641	1	922.641	3.123	.082
	Entropy	819.391	1	819.391	3.582	.063
	LMR_1V2	961.000	1	961.000	6.161	.016
	LMR_2V3	2173.891	1	2173.891	84.195	.000
	BLRT_1V2	600.250	1	600.250	4.594	.036
	BLRT_2V3	1521.000	1	1521.000	19.123	.000
class_sepa * measure	AIC	15.016	1	15.016	.088	.768
	CAIC	441.000	1	441.000	.682	.412
	SACAIC	2413.266	1	2413.266	11.403	.001
	BIC	2875.641	1	2875.641	5.387	.024
	SABIC	631.266	1	631.266	2.948	.091
	DBIC	3011.266	1	3011.266	12.037	.001
	HQ	1732.641	1	1732.641	8.509	.005
	HT_AIC	107.641	1	107.641	.364	.548
	Entropy	1.891	1	1.891	.008	.928
	LMR_1V2	441.000	1	441.000	2.827	.098
	LMR_2V3	66.016	1	66.016	2.557	.115
	BLRT_1V2	420.250	1	420.250	3.217	.078
	BLRT_2V3	25.000	1	25.000	.314	.577
Error	AIC	10217.562	60	170.293		
	CAIC	38785.125	60	646.419		
	SACAIC	12698.062	60	211.634		
	BIC	32031.062	60	533.851		
	SABIC	12849.312	60	214.155		
	DBIC	15010.188	60	250.170		
	HQ	12217.688	60	203.628		
	HT_AIC	17723.812	60	295.397		
	Entropy	13724.938	60	228.749		
	LMR_1V2	9358.875	60	155.981		
	LMR_2V3	1549.188	60	25.820		
	BLRT_1V2	7839.000	60	130.650		
	BLRT_2V3	4772.375	60	79.540		
Total	AIC	49481.000	64			

	CAIC	404966.000	64
	SACAIC	554709.000	64
	BIC	458297.000	64
	SABIC	473539.000	64
	DBIC	542971.000	64
	HQ	524303.000	64
	HT_AIC	70895.000	64
	Entropy	66121.000	64
	LMR_1V2	585954.000	64
	LMR_2V3	422329.000	64
	BLRT_1V2	600208.000	64
	BLRT_2V3	199874.000	64
Corrected Total	AIC	12472.859	63
	CAIC	70303.750	63
	SACAIC	19799.609	63
	BIC	60924.359	63
	SABIC	16731.984	63
	DBIC	26549.109	63
	HQ	15042.359	63
	HT_AIC	18967.984	63
	Entropy	14648.734	63
	LMR_1V2	11768.937	63
	LMR_2V3	3881.734	63
	BLRT_1V2	9615.750	63
	BLRT_2V3	6493.937	63

a. R Squared = .181 (Adjusted R Squared = .140)
b. R Squared = .448 (Adjusted R Squared = .421)
c. R Squared = .359 (Adjusted R Squared = .327)
d. R Squared = .474 (Adjusted R Squared = .448)
e. R Squared = .232 (Adjusted R Squared = .194)
f. R Squared = .435 (Adjusted R Squared = .406)
g. R Squared = .188 (Adjusted R Squared = .147)
h. R Squared = .066 (Adjusted R Squared = .019)
i. R Squared = .063 (Adjusted R Squared = .016)
j. R Squared = .205 (Adjusted R Squared = .165)
k. R Squared = .601 (Adjusted R Squared = .581)
l. R Squared = .185 (Adjusted R Squared = .144)
m. R Squared = .265 (Adjusted R Squared = .228)
n. type_mixture =LPM

Table B13: Class separation X Number of measures in UGMM

Tests of Between-Subjects Effects[n]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	55201.547[a]	3	18400.516	139.686	.000
	CAIC	13643.375[b]	3	4547.792	7.138	.000
	SACAIC	158.063[c]	3	52.688	4.130	.010
	BIC	7493.063[d]	3	2497.688	6.099	.001
	SABIC	3768.125[e]	3	1256.042	26.900	.000
	DBIC	415.297[f]	3	138.432	4.169	.010
	HQ	1505.563[g]	3	501.854	4.722	.005
	HT_AIC	53501.188[h]	3	17833.729	126.030	.000
	Entropy	15104.547[i]	3	5034.849	29.791	.000
	LMR_1V2	162.813[j]	3	54.271	3.379	.024
	LMR_2V3	3084.313[k]	3	1028.104	18.698	.000
	BLRT_1V2	407.812[l]	3	135.937	3.327	.025
	BLRT_2V3	519.297[m]	3	173.099	2.716	.053
Intercept	AIC	196359.766	1	196359.766	1.491E3	.000
	CAIC	459006.250	1	459006.250	720.455	.000
	SACAIC	609570.562	1	609570.562	4.779E4	.000
	BIC	523814.062	1	523814.062	1.279E3	.000

	SABIC	522006.250	1	522006.250	1.118E4	.000
	DBIC	607425.391	1	607425.391	1.829E4	.000
	HQ	545751.562	1	545751.562	5.135E3	.000
	HT_AIC	214600.562	1	214600.562	1.517E3	.000
	Entropy	33902.016	1	33902.016	200.595	.000
	LMR_1V2	618975.562	1	618975.562	3.854E4	.000
	LMR_2V3	502326.562	1	502326.562	9.136E3	.000
	BLRT_1V2	615832.562	1	615832.562	1.507E4	.000
	BLRT_2V3	489125.391	1	489125.391	7.674E3	.000
class_sepa	AIC	301.891	1	301.891	2.292	.135
	CAIC	9555.062	1	9555.062	14.998	.000
	SACAIC	33.062	1	33.062	2.592	.113
	BIC	4830.250	1	4830.250	11.795	.001
	SABIC	25.000	1	25.000	.535	.467
	DBIC	178.891	1	178.891	5.387	.024
	HQ	42.250	1	42.250	.398	.531
	HT_AIC	333.062	1	333.062	2.354	.130
	Entropy	62.016	1	62.016	.367	.547
	LMR_1V2	126.562	1	126.562	7.880	.007
	LMR_2V3	14.062	1	14.062	.256	.615
	BLRT_1V2	217.562	1	217.562	5.325	.024
	BLRT_2V3	.141	1	.141	.002	.963
measure	AIC	54463.891	1	54463.891	413.457	.000
	CAIC	3080.250	1	3080.250	4.835	.032
	SACAIC	4.000	1	4.000	.314	.578
	BIC	1540.562	1	1540.562	3.762	.057
	SABIC	3510.562	1	3510.562	75.183	.000
	DBIC	43.891	1	43.891	1.322	.255
	HQ	1260.250	1	1260.250	11.858	.001
	HT_AIC	52555.562	1	52555.562	371.406	.000
	Entropy	14731.891	1	14731.891	87.167	.000
	LMR_1V2	16.000	1	16.000	.996	.322
	LMR_2V3	2970.250	1	2970.250	54.019	.000
	BLRT_1V2	100.000	1	100.000	2.447	.123

	BLRT_2V3	478.516	1	478.516	7.507	.008
class_sepa * measure	AIC	435.766	1	435.766	3.308	.074
	CAIC	1008.062	1	1008.062	1.582	.213
	SACAIC	121.000	1	121.000	9.486	.003
	BIC	1122.250	1	1122.250	2.740	.103
	SABIC	232.562	1	232.562	4.981	.029
	DBIC	192.516	1	192.516	5.798	.019
	HQ	203.062	1	203.062	1.911	.172
	HT_AIC	612.562	1	612.562	4.329	.042
	Entropy	310.641	1	310.641	1.838	.180
	LMR_1V2	20.250	1	20.250	1.261	.266
	LMR_2V3	100.000	1	100.000	1.819	.183
	BLRT_1V2	90.250	1	90.250	2.209	.142
	BLRT_2V3	40.641	1	40.641	.638	.428
Error	AIC	7903.688	60	131.728		
	CAIC	38226.375	60	637.106		
	SACAIC	765.375	60	12.756		
	BIC	24570.875	60	409.515		
	SABIC	2801.625	60	46.694		
	DBIC	1992.312	60	33.205		
	HQ	6376.875	60	106.281		
	HT_AIC	8490.250	60	141.504		
	Entropy	10140.438	60	169.007		
	LMR_1V2	963.625	60	16.060		
	LMR_2V3	3299.125	60	54.985		
	BLRT_1V2	2451.625	60	40.860		
	BLRT_2V3	3824.312	60	63.739		
Total	AIC	259465.000	64			
	CAIC	510876.000	64			
	SACAIC	610494.000	64			
	BIC	555878.000	64			
	SABIC	528576.000	64			
	DBIC	609833.000	64			
	HQ	553634.000	64			

	HT_AIC	276592.000	64
	Entropy	59147.000	64
	LMR_1V2	620102.000	64
	LMR_2V3	508710.000	64
	BLRT_1V2	618692.000	64
	BLRT_2V3	493469.000	64
Corrected Total	AIC	63105.234	63
	CAIC	51869.750	63
	SACAIC	923.438	63
	BIC	32063.938	63
	SABIC	6569.750	63
	DBIC	2407.609	63
	HQ	7882.438	63
	HT_AIC	61991.438	63
	Entropy	25244.984	63
	LMR_1V2	1126.438	63
	LMR_2V3	6383.438	63
	BLRT_1V2	2859.437	63
	BLRT_2V3	4343.609	63

a. R Squared = .875 (Adjusted R Squared = .868)

b. R Squared = .263 (Adjusted R Squared = .226)

c. R Squared = .171 (Adjusted R Squared = .130)

d. R Squared = .234 (Adjusted R Squared = .195)

e. R Squared = .574 (Adjusted R Squared = .552)

f. R Squared = .172 (Adjusted R Squared = .131)

g. R Squared = .191 (Adjusted R Squared = .151)

h. R Squared = .863 (Adjusted R Squared = .856)

i. R Squared = .598 (Adjusted R Squared = .578)

j. R Squared = .145 (Adjusted R Squared = .102)

k. R Squared = .483 (Adjusted R Squared = .457)

l. R Squared = .143 (Adjusted R Squared = .100)

m. R Squared = .120 (Adjusted R Squared = .076)

n. type_mixture = UGMM

Table B14: Class separation X Number of measures in Linear GMM

Tests of Between-Subjects Effects[n]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AIC	3935.375[a]	3	1311.792	20.907	.000
	CAIC	2909.797[b]	3	969.932	4.393	.007
	SACAIC	33.047[c]	3	11.016	.090	.966
	BIC	1046.797[d]	3	348.932	3.963	.012
	SABIC	159.875[e]	3	53.292	.192	.901
	DBIC	23.625[f]	3	7.875	.129	.943
	HQ	192.547[g]	3	64.182	1.959	.130
	HT_AIC	3402.297[h]	3	1134.099	18.928	.000
	Entropy	3825.563[i]	3	1275.188	18.874	.000
	LMR_1V2	14.187[j]	3	4.729	2.451	.072
	LMR_2V3	2098.125[k]	3	699.375	24.431	.000
	BLRT_1V2	340.187[l]	3	113.396	3.585	.019
	BLRT_2V3	1955.797[m]	3	651.932	12.289	.000
Intercept	AIC	82944.000	1	82944.000	1.322E3	.000
	CAIC	567950.641	1	567950.641	2.572E3	.000
	SACAIC	484590.016	1	484590.016	3.939E3	.000
	BIC	598108.891	1	598108.891	6.792E3	.000
	SABIC	246512.250	1	246512.250	890.371	.000
	DBIC	549822.250	1	549822.250	9.008E3	.000
	HQ	320214.516	1	320214.516	9.773E3	.000
	HT_AIC	85775.766	1	85775.766	1.432E3	.000
	Entropy	88060.562	1	88060.562	1.303E3	.000
	LMR_1V2	634014.062	1	634014.062	3.286E5	.000
	LMR_2V3	352242.250	1	352242.250	1.230E4	.000
	BLRT_1V2	620550.062	1	620550.062	1.962E4	.000
	BLRT_2V3	480422.266	1	480422.266	9.056E3	.000
class_sepa	AIC	600.250	1	600.250	9.567	.003
	CAIC	2058.891	1	2058.891	9.324	.003
	SACAIC	19.141	1	19.141	.156	.695

	BIC	695.641	1	695.641	7.900	.007
	SABIC	126.562	1	126.562	.457	.502
	DBIC	.562	1	.562	.009	.924
	HQ	185.641	1	185.641	5.666	.020
	HT_AIC	606.391	1	606.391	10.121	.002
	Entropy	1173.062	1	1173.062	17.362	.000
	LMR_1V2	14.062	1	14.062	7.289	.009
	LMR_2V3	5.062	1	5.062	.177	.676
	BLRT_1V2	150.062	1	150.062	4.744	.033
	BLRT_2V3	135.141	1	135.141	2.547	.116
measure	AIC	3335.062	1	3335.062	53.154	.000
	CAIC	435.766	1	435.766	1.973	.165
	SACAIC	5.641	1	5.641	.046	.831
	BIC	172.266	1	172.266	1.956	.167
	SABIC	.250	1	.250	.001	.976
	DBIC	14.062	1	14.062	.230	.633
	HQ	6.891	1	6.891	.210	.648
	HT_AIC	2795.766	1	2795.766	46.662	.000
	Entropy	2652.250	1	2652.250	39.255	.000
	LMR_1V2	.062	1	.062	.032	.858
	LMR_2V3	2093.062	1	2093.062	73.115	.000
	BLRT_1V2	95.062	1	95.062	3.006	.088
	BLRT_2V3	1691.266	1	1691.266	31.881	.000
class_sepa * measure	AIC	.062	1	.062	.001	.975
	CAIC	415.141	1	415.141	1.880	.175
	SACAIC	8.266	1	8.266	.067	.796
	BIC	178.891	1	178.891	2.032	.159
	SABIC	33.062	1	33.062	.119	.731
	DBIC	9.000	1	9.000	.147	.702
	HQ	.016	1	.016	.000	.983
	HT_AIC	.141	1	.141	.002	.962
	Entropy	.250	1	.250	.004	.952
	LMR_1V2	.062	1	.062	.032	.858
	LMR_2V3	.000	1	.000	.000	1.000

	BLRT_1V2	95.062	1	95.062	3.006	.088
	BLRT_2V3	129.391	1	129.391	2.439	.124
Error	AIC	3764.625	60	62.744		
	CAIC	13248.562	60	220.809		
	SACAIC	7381.938	60	123.032		
	BIC	5283.312	60	88.055		
	SABIC	16611.875	60	276.865		
	DBIC	3662.125	60	61.035		
	HQ	1965.938	60	32.766		
	HT_AIC	3594.938	60	59.916		
	Entropy	4053.875	60	67.565		
	LMR_1V2	115.750	60	1.929		
	LMR_2V3	1717.625	60	28.627		
	BLRT_1V2	1897.750	60	31.629		
	BLRT_2V3	3182.938	60	53.049		
Total	AIC	90644.000	64			
	CAIC	584109.000	64			
	SACAIC	492005.000	64			
	BIC	604439.000	64			
	SABIC	263284.000	64			
	DBIC	553508.000	64			
	HQ	322373.000	64			
	HT_AIC	92773.000	64			
	Entropy	95940.000	64			
	LMR_1V2	634144.000	64			
	LMR_2V3	356058.000	64			
	BLRT_1V2	622788.000	64			
	BLRT_2V3	485561.000	64			
Corrected Total	AIC	7700.000	63			
	CAIC	16158.359	63			
	SACAIC	7414.984	63			
	BIC	6330.109	63			
	SABIC	16771.750	63			
	DBIC	3685.750	63			

HQ	2158.484	63
HT_AIC	6997.234	63
Entropy	7879.438	63
LMR_1V2	129.937	63
LMR_2V3	3815.750	63
BLRT_1V2	2237.937	63
BLRT_2V3	5138.734	63

a. R Squared = .511 (Adjusted R Squared = .487)

b. R Squared = .180 (Adjusted R Squared = .139)

c. R Squared = .004 (Adjusted R Squared = -.045)

d. R Squared = .165 (Adjusted R Squared = .124)

e. R Squared = .010 (Adjusted R Squared = -.040)

f. R Squared = .006 (Adjusted R Squared = -.043)

g. R Squared = .089 (Adjusted R Squared = .044)

h. R Squared = .486 (Adjusted R Squared = .461)

i. R Squared = .486 (Adjusted R Squared = .460)

j. R Squared = .109 (Adjusted R Squared = .065)

k. R Squared = .550 (Adjusted R Squared = .527)

l. R Squared = .152 (Adjusted R Squared = .110)

m. R Squared = .381 (Adjusted R Squared = .350)
n. type_mixture = Linear GMM

Bibliography

A'Hearn, B. A. & Komlos, J. (2003). *Improvements in maximum likelihood estimators of truncated normal samples with prior knowledge of* α *. A simulation based study with application to historical height samples.* Unpublished Working Paper, University of Munich, Germany.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*, 317-332.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, *12*, 171-8.

Azzalini,A. (2005). The Skew-Normal Distribution and Related Multivariate Families. S*candinavian Journal of Statistics,*32, 2.

Bartholomew, D. J. (1987). *Latent variable models and factor analysis*. London: Griffin.

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 the overextraction of latent trajectory classes. *Psychological Methods*, *8*, 338-363.

Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods*, *9*, 3-29.

Bozdogan, H. (1987). Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, *52,* 345-370.

Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification, 13,* 195-212.

Chang, I. (2005). *Bayesian inference on mixture models and their applications.* (Doctoral dissertation, Texas A&M University). Retrieved from http://repository.tamu.edu/bitstream/handle/1969.1/3990/etd-tamu-2005A-STAT-Chan.pdf?sequence=1.

Connell, A.M., & Frye, A.A. (2006). Growth mixture modeling in developmental psychology: Overview and demonstration of heterogeneity in developmental trajectories of adolescent antisocial behaviour. *Infant and Child Development, 15*, 609–621.

Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society: Series B., 57*, 45-97.

Douglas, K., & Liu, M. (2009, June). Different trajectories in learning to read in U.S. elementary schools. Paper presented at the annual meeting of the Society for the Scientific Study of Reading, Boston, MA.

Gibson, W. A. (1959). Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. *Psychometrika, 24,* 229-252.

Greenbaum, P. E., Del Boca, F. K., Darkes, J., Wang, C-P., & Goldman, M. S. (2005). Variation in the drinking trajectories of freshman college students. *Journal of Consulting and Clinical Psychology*, *73*, 229-238.

Grimm, K. J., & Ram, N. (2009). A second-order growth mixture model for developmental research. *Research in Human Development*, 6, 121-143.

Hamilton, J. (2009). *An investigation of growth mixture models when data are collected with unequal selection probabilities: A Monte Carlo study.* (Doctoral dissertation, University of Maryland). Retrieved from http://hdl.handle.net/1903/9613.

Hancock, G. R., & Lawrence, F. R. (2006). Using latent growth models to evaluate longitudinal change. In G. R. Hancock & R. O. Mueller (Eds.), *Structural Equation Modeling: A Second Course* (pp. 171-196). Greenwood, CT: Information Age Publishing, Inc.

Hancock, G. R., Kuo, W. L., & Lawrence, F.R. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling: A Multidisciplinary Journal*, *8,* 470-489.

Hannan, E. J. & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society, 41,*190-195.

Haughton, D. (1997). Packages for estimating finite mixtures: a review. *The American Statistician 51,* 194-205.

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: A Multidisciplinary Journal*, *14*, 202-226.

Hill, A. L., Degnan, K. A., Calkins, S. D., & Keane, S. P. (2006). Profile of externalizing behavior problems for boys and girls across preschools: The role of emotion regulation and inattention. *Developmental Psychology, 42,* 913-928.

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

Hurvich, C. M., & Tsai, C. L. (1989). Regression and time series model selection in small samples, *Biometrika 76***,** 297-307.

Kreuter, F., & Muthén, B. (2008a). Analyzing criminal trajectory profiles: Bridging multilevel and group-based approaches using growth mixture modeling. *Journal of Quantitative Criminology, 24*, 1-31.

Kreuter, F., & Muthén, B. (2008b). Longitudinal modeling of population heterogeneity: Methodological challenges to the analysis of empirically derived criminal trajectory profiles. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (pp. 53-75). Greenwich, CT: Information Age Publishing.

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.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data, *Biometrics 38*, 963-974.

Little, T. D., Card, N. A., Preacher, K. J., & McConnell, E. (2009). Modeling longitudinal data from research on adolescence. In R.M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (3rd ed.). Hoboken: Wiley.

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika, 88,* 767-778.

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: A Multidisciplinary Journal*, *14,* 26-47.

Maclean, C. J., Morton, N. E., Elston, R. C., & Yee, S. (1976). Skewness in commingled distributions. *Biometrics, 32*, 695-699.

Marsh, H. W., Ludtke, O., Trautwein, U., & Morin, A. J. S. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person- and variable-centered appropaches to theoretical models of self-concept. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*, 191-225.

McLachlan, G. J. (1987). On bootstrapping the likelihood ratio test statistic for the number of components in a normal mixture. *Applied Statistics, 36*, 318-324.

McLachlan, G. J. (1999). Mahalanobis distance. *Resonance 4,* 20-26.

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

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

Muthén, B. (2001). *Two-part growth mixture modeling.* University of California, Los Angeles.

Muthén, L. K., & Muthén, B. O. (2001). *Mplus user's guide* Los Angeles: Muthén and Muthén.

Muthén, B. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, *29,* 81-117.

Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling. *Psychological Methods, 8,* 369-377.

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

Muthén, B. & Asparouhov, T. (2008). Growth mixture modeling: Analysis with non-Gaussian random effects. In Fitzmaurice, G., Davidian, M., Verbeke, G. & Molenberghs, G. (eds.), *Longitudinal Data Analysis* (pp. 143-165). Boca Raton: Chapman & Hall/CRC Press.

Muthén, L. K., & Muthén, B. O. (2008). *Mplus user's guide* (5th Ed.). Los Angeles: Muthén and Muthén.

Muthén, B., Brown, C. H., Hunter, A., Cook, I. A. & Leuchter, A. F. (2011). General approaches to analysis of course: Applying growth mixture modeling to randomized trials of depression medication. In P.E. Shrout (ed.), *Causality and psychopathology: Finding the determinants of disorders and their cures* (pp. 159-178). New York: Oxford University Press.

Nylund, K. L., 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: A Multidisciplinary Journal*, *14*, 535-569.

Odgers, C. L., Moffitt, T. E., Broadbent, J. M., Dickson, N., Hancox, R. J., Harrington, H., et al. (2008). Female and male antisocial trajectories: From childhood origins to adult outcomes. *Development and Psychopathology, 20*, 673-716.

Pastor, D. A., Barron, K.E., Miller, B.J., & Davis, S.L. (2007). A latent profile analysis of college students' achievement goal orientation. *Contemporary Educational Psychology, 32,* 8-47.

Rice, K., Lumley, T. and Szpiro, A. (2008). Trading bias for precision: Decision theory for intervals and sets. UW Biostatistics Working Paper Series, Working Paper 336.

Ramaswamy, V., DeSarbo, W. S., Reibstein, D. J., & Robinson, W. T. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science*, *12*, 103–124.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461-464.

Sclove, L. S. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrics, 52,* 333–343.

Soromenho, G. (1994). Comparing approaches for testing the number of components in a finite mixture models. *Computational Statistics,* 9, 65-78

Stoolmiller, M., Kim, H. K., & Capaldi, D. M. (2005). The course of depressive symptoms in men from early adolescence to young adulthood: Identifying latent trajectories and early predictors. *Journal of Abnormal Psychology*, *114,* 331-345

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

Tolvanen, A. (2008). *Latent growth mixture modeling: A simulation study.* (Doctoral dissertation, University of Jyvaskyla). Retrieved from http://www.statmodel.com/download/rep111.pdf.

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

Verbeke, G. & Lesaffre, E. (1996). A linear mixed-effects model with heterogeneity in the random-effects population. *Journal of the American Statistical Association 91*, 217-221.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrics*, *57*, 307-333.

Wang, M., & Bodner, T. E. (2007) Growth mixture modeling: identifying and predicting unobserved subpopulations with longitudinal data. *Organizational Research Methods*, *10*, 635-656.

Wedel, M., ter Hofstede, F., & Steenkamp, J-B. E. M. (1998). Mixture model analysis of complex samples. *Journal of Classification*, *15,* 225-244.

Yang, C. C., (2006). Evaluating latent class analysis models in qualitative phenotype Identification. *Computational Statistics & Data Analysis*, *50,* 1090-1104.

Yang, C. C & Yang, C.C. (2007). Separating latent classes by information criteria. *Journal of Classification*, *24*, 183-203.

Yung, Y. F. (1997). Finite mixtures in confirmatory factor-analysis models.

Psychometrika, 62.